

Aroturuki Tamariki's proposed use of Microsoft 365 Copilot

Privacy Impact Assessment

Contents

1.	Executive summary.....	3
2.	Background	5
3.	Scope of assessment.....	5
4.	How Microsoft 365 Copilot works	6
5.	How Microsoft processes content associated with the use of Copilot	7
6.	Personal information to which Copilot will have access.....	16
7.	Technical and non-technical controls	16
8.	Privacy assessment	17
9.	Privacy and related risks and recommendations	26
10.	Addendum post-pilot – 26 May 2026	29

Document Control

Title	Aroturuki Tamariki's proposed use of Microsoft 365 Copilot – Privacy Impact Assessment
Author	Richard Best
Date and Version	17 March 2026. Final.
	26 May 2026. Updated with an Addendum to address implications arising from a pilot that took place from 13 April to 1 May 2026.

1. Executive summary

1.1 What is being assessed

- (a) Aroturuki Tamariki is proposing to allow staff to use Microsoft 365 Copilot for certain everyday, back-office activities.
- (b) This Privacy Impact Assessment (**PIA**) assesses privacy-related risks relating to that proposed use of Copilot.

1.2 What Copilot is and the AI models it uses

- (a) Microsoft 365 Copilot is an AI-powered tool that can help with work-related tasks. Users enter a prompt in Copilot and Copilot responds with AI-generated information.
- (b) Originally, Copilot utilised an Azure OpenAI instance maintained by Microsoft as its AI model. It still does but, in late 2025, Microsoft added Anthropic's Claude as an alternative selectable model. Aroturuki Tamariki has indicated it may turn on the ability to select Claude.
- (c) When Azure OpenAI is used as the processing model, all processing occurs within the Microsoft 365 service boundary. Two exceptions to this are where the web search feature is enabled and where Copilot agent extensions are used to provide Copilot with access to external or third-party systems. The web search feature does not appear to create material heightened risk, and Aroturuki Tamariki does not currently propose to use Copilot agent extensions.
- (d) When Claude is selected, currently processing will occur outside the Microsoft 365 service boundary. This is because the Claude models are hosted by Anthropic (typically running on AWS or Google Cloud) rather than natively within the core Microsoft 365 infrastructure. However, Microsoft now treats Anthropic as a subprocessor, and Microsoft's contractual commitments apply to the use of either processing model except, in the case of Claude, in-country processing commitments (which in any event are not absolute).
- (e) The practical consequence is that Microsoft's commitments to not using customer data to train the AI models it uses, and to not use or disclose customer data for its own purposes unrelated to service provision, apply to both models. Offshore processing beyond Aroturuki Tamariki's Australian instance of Microsoft 365 may occur when Claude is selected, but that remains a possibility when Azure OpenAI is selected because, even when the in-country processing option has been taken up, processing outside Australia may occur when there are capacity constraints. Therefore, in either case, there are residual jurisdictional risks that Aroturuki Tamariki will need to accept if it is to deploy Copilot. Other New Zealand government agencies deploying Copilot must have accepted these risks (whether knowingly or otherwise).

1.3 Personal information that could be processed

Importantly, whilst Copilot may have access to some personal information, that information will be limited. Copilot's access will be limited to information that staff are already authorised to view within Microsoft 365, excluding monitoring data and sensitive repositories designated as off-limits. It will not have access to personal information about tamariki, rangatahi, and their whānau, nor to personal information held in libraries relating to monitoring, human resources, and other staffing matters. It is understood that any personal

information Copilot may draw on via Microsoft Graph will be confined to routine organisational data such as staff and Board contact details, user preferences and settings, user activity and usage data, and personal information contained in users' own emails, calendars, and documents (excluding personal information of tamariki, rangatahi and their whānau, which is stored in Te Pataka or is in libraries marked as off-limits for Copilot).

1.4 Privacy Act compliance and risk

- (a) If Copilot is used with the above access controls in place and staff comply with Aroturuki Tamariki's AI Policy and AI Usage Guide, then the use of Copilot will not result in contravention of the Privacy Act's information privacy principles. There are, however, some privacy-related risks that Aroturuki Tamariki can take active steps to mitigate. Those risks arise from staff not complying with the AI Policy and AI Usage Guide, and either:
 - (i) using Copilot to collect new personal information about people or entering kinds of personal information into Copilot that they should not be entering into Copilot;
 - (ii) not checking the accuracy of a Copilot output that does legitimately contain personal information before using or permissibly disclosing the output; or
 - (iii) using the basic personal information that might be entered or included in a prompt (e.g., some staff names for the purpose of getting Copilot to help with a meeting agenda) for some unusual and unexpected purpose.
- (b) These risks can be mitigated through a combination of staff training (which will be required before staff can use Copilot) and spot checks. These recommendations are set out in section 8 of this PIA, and are collected together with other recommendations in section 9.
- (c) Whilst not legally required, it is also recommended that Aroturuki Tamarik consider preparation and publication of a GenAI Transparency Statement.

1.5 Oversight Act and Information Rules compliance

- (a) The proposed use of Copilot does not raise concerns relating to the information provisions in Part 3 of the Oversight of Oranga Tamariki System Act 2022 or the Information Rules under that Act.
- (b) Most proposed uses can be described as purely back-office, administrative, or incidental in nature and will not require consultation with stakeholders under rule 7.2 of the Information Rules. It is recommended, however, that Aroturuki Tamariki review its list of proposed use cases and not proceed with use cases that could shape monitoring analysis, monitoring evaluative judgments, monitoring prioritisation, or monitoring outputs (if any) before undertaking the consultation referred to in rule 7.2. Aroturuki Tamariki is best placed to make this judgement.

1.6 Other matters

- (a) It is also recommended that Aroturuki Tamariki take advantage of the Copilot in-country processing option (meaning Copilot processing will usually, but not invariably, occur in Australia), confirm that the Copilot feedback feature will be disabled before the roll-out of Copilot to staff, and adopt a formal position that staff are not to use

Copilot agent extensions without approval from the AI Officer and until appropriate assessments have been done.

- (b) Any technical control issues that may emerge from Aroturuki Tamariki's certification of Copilot will need to be considered separately to this PIA. Technical control issues are not expected but they are beyond the scope of this PIA.

2. Background

- 2.1 Aroturuki Tamariki is proposing to allow staff to use Microsoft 365 Copilot for certain everyday, back-office activities. Such use would be subject to the agency's AI Policy and AI Usage Guide.
- 2.2 The deployment of Copilot would be through Aroturuki Tamariki's enterprise subscription to Microsoft 365.
- 2.3 It is proposed that an AI Working Group will pilot the use of Copilot before wider roll-out to staff. The purpose of the pilot is to identify any unforeseen issues or risks that may need to be addressed before the wider roll-out occurs.
- 2.4 Aroturuki Tamariki also proposes to hold training sessions for staff, to help them understand the uses of Copilot that are permitted, those that are not, and constraints on the use of other AI tools for work-related activity.

3. Scope of assessment

- 3.1 This PIA assesses privacy-related risks relating to Aroturuki Tamariki's proposed use of Copilot, by reference to:
 - (a) the Privacy Act 2020; and
 - (b) the Oversight of Oranga Tamariki System Act 2022 and the Information Rules under that Act.
- 3.2 The PIA also takes into account Aroturuki Tamariki's AI Policy and AI Usage Guide.
- 3.3 The PIA is not a "privacy, human rights, and ethics assessment" utilising the likes of MSD's PHRAE framework, nor is it an Algorithm Impact Assessment using the Algorithm Toolkit developed for StatsNZ.
- 3.4 It is understood that, in addition to this PIA, Aroturuki Tamariki is or will be undertaking separate certification and accreditation of its proposed use Microsoft 365 Copilot. That exercise, rather than this PIA, can be expected to consider the presence of relevant NZISM controls.
- 3.5 The PIA is undertaken based on information available and provided as at February 2026. The PIA may need to be updated if:
 - (a) further use cases are proposed beyond those communicated;
 - (b) there are technical or legal changes adversely affecting the use of Copilot, Azure OpenAI, or Anthropic's Claude;
 - (c) material incidents occur relating to Aroturuki Tamariki's or other New Zealand government agencies' use of Copilot;

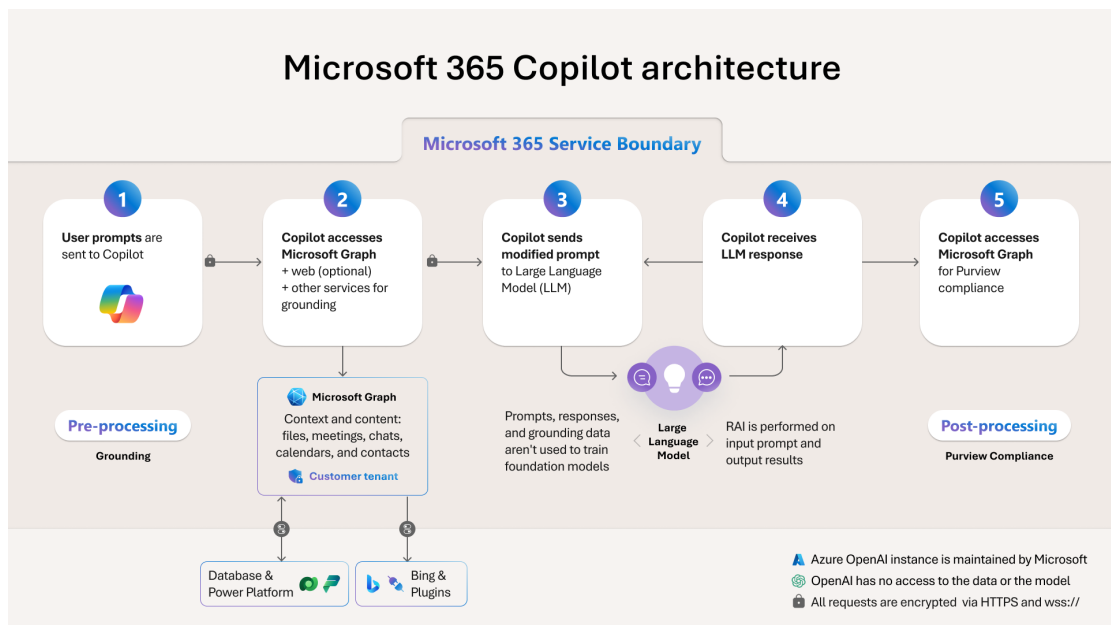
- (d) Aroturuki Tamariki's certification and accreditation of Copilot reveals material technical control-related risks not assessed in this PIA; or
- (e) DIA, the forthcoming Government Digital Delivery Agency, or the Office of the Privacy Commissioner releases guidance specifically on the use of Microsoft Copilot.

4. How Microsoft 365 Copilot works

- 4.1 Microsoft 365 Copilot is an AI-powered tool that can help with work-related tasks. Users enter a prompt in Copilot and Copilot responds with AI-generated information. The responses are in real-time and can include work content that users have permission to access and, if web search is enabled, internet-based content.
- 4.2 Originally, Copilot utilised an Azure OpenAI instance maintained by Microsoft as its large language model of choice. There was no ability to select an alternative model. However, in late 2025, Microsoft added Anthropic's Claude as an alternative selectable model. Admins have the ability to turn this 'alternative model' feature on or off. Aroturuki Tamariki's AI Officer has advised that, unless data control and privacy risks are significant, Aroturuki Tamariki would likely wish to turn this feature on.

Processing within Microsoft 365 service boundary when Azure OpenAI model is used

- 4.3 With two notable exceptions discussed below, user prompts, contextual information that Copilot sources from internal documentation (known as 'grounding'), processing by the large language model, and Copilot's provision of a response from the large language model (an Azure OpenAI instance maintained by Microsoft), all occur within the Microsoft 365 service boundary. It is understood that no customer data is sent outside of that service boundary to OpenAI. This is explained diagrammatically in the image below:



4.4 As Microsoft puts it:¹

"When you enter prompts using Microsoft 365 Copilot, the information contained within your prompts, the data they retrieve, and the generated responses remain within the

¹ See <https://learn.microsoft.com/en-us/copilot/microsoft-365/microsoft-365-copilot-privacy>

Microsoft 365 service boundary, in keeping with our current privacy, security, and compliance commitments. Microsoft 365 Copilot uses Azure OpenAI services for processing, not OpenAI's publicly available services. Azure OpenAI doesn't cache customer content and Copilot modified prompts for Microsoft 365 Copilot."

- 4.5 This service architecture is reassuring in the sense that, when the Azure instance of the OpenAI model(s) is called upon, but subject to the situations discussed below, one can be confident that Microsoft is not relying on a range of third party services outside the Microsoft 365 service boundary or sending prompts and associated customer content to third party large language model (**LLM**) providers like OpenAI.

Processing when Anthropic Claude models are used

- 4.6 Claude models are hosted by Anthropic rather than natively within the core Microsoft 365 infrastructure. This means users' grounded prompts are transmitted to Anthropic-managed infrastructure (typically running on AWS or Google Cloud) for at least the duration of the request, outside of the Microsoft service boundary. This topic is discussed further at paragraphs 5.22-5.24 below.

5. How Microsoft processes content associated with the use of Copilot

Sources

- 5.1 This issue has been considered by reference to two sources:
- (a) first, by reference to what is understood to be the applicable online legal terms; and
 - (b) second, by reference to representations made by Microsoft on its website(s).
- 5.2 Each is addressed in turn. The key point that flows from the discussion is that Microsoft does not use customer data for purposes unrelated to provision of the services and does not use it to train its own or third party LLMs.

Applicable legal terms

- 5.3 It is understood that the use of Microsoft 365 Copilot in enterprise settings is governed by the Microsoft Product Terms > Universal Licence Terms.² Those terms state that the Data Protection Addendum (**DPA**) "sets forth [the parties'] obligations with respect to the processing and security of Customer Data and Personal Data by the Online Services".
- 5.4 The DPA:
- (a) defines Customer Data as "all data, including all text, sound, video, or image files, and software, that are provided to Microsoft by, or on behalf of, Customer through use of the Online Service. Customer Data does not include Professional Services Data"; and
 - (b) states that Microsoft "will use and otherwise process Customer Data, Professional Services Data, and Personal Data only as described and subject to the limitations provided below (a) **to provide** Customer the Products and Services in accordance with Customer's documented instructions and (b) for **business operations** incident to providing the Products and Services to Customer."

² Available at <https://www.microsoft.com/licensing/terms/product/ForOnlineServices/all>

5.5 To understand what this means, one needs to understand what Microsoft means in (a) by "to provide" a Product, and what it means in (b) by "business operations". It defines both of these terms.

5.6 The DPA states that "to provide" a Product consists of:

- “• Delivering functional capabilities as licensed, configured, and used by Customer and its users, including providing personalized user experiences;
- Troubleshooting (preventing, detecting, and repairing problems); and
- Keeping Products up to date and performant, and enhancing user productivity, reliability, efficacy, quality, and security.”

5.7 Microsoft also makes it clear that "[w]hen providing Products and Services, Microsoft will not use or otherwise process Customer Data ... or Personal Data for: (a) user profiling, (b) advertising or similar commercial purposes, or (c) market research aimed at creating new functionalities, services, or products or any other purpose, unless such use or processing is in accordance with Customer's documented instructions".

5.8 Turning to the meaning of "business operations", the DPA states (the emphasis is mine):

“business operations’ means the processing operations authorized by customer in this section.

Customer authorizes Microsoft:

- to create aggregated statistical, non-personal data from data containing pseudonymized identifiers (such as usage logs containing unique, pseudonymized identifiers); and
- to calculate statistics related to Customer Data or Professional Services Data

in each case without accessing or analyzing the content of Customer Data or Professional Services Data and limited to achieving the purposes below, each as incident to providing the Products and Services to Customer.

Those purposes are:

- billing and account management;
- compensation such as calculating employee commissions and partner incentives;
- internal reporting and business modeling, such as forecasting, revenue, capacity planning, and product strategy; and
- financial reporting.

When processing for these business operations, Microsoft will apply principles of data minimization and will not use or otherwise process Customer Data, Professional Services Data, or Personal Data for: (a) user profiling, (b) advertising or similar commercial purposes, or (c) any other purpose, other than for the purposes set out in this section. In addition, as with all processing under this DPA, processing for business operations remains subject to Microsoft's confidentiality obligations and commitments under Disclosure of Processed Data.”

Online representations

5.9 Turning now to online representations specifically in relation to Copilot, Microsoft states that:

- (a) Microsoft 365 Copilot only surfaces (for grounding) organisational data to which individual users have at least view permissions;
- (b) Copilot respects the organisational customer's identity model and permissions, inherits its sensitivity labels, applies its retention policies, supports audit of interactions, and follows its administrative settings;
- (c) prompts, responses, and data accessed through Microsoft Graph are not used to train foundation LLMs, including those used by Microsoft 365 Copilot, and Microsoft says it "won't use your data except as you instruct";
- (d) when a user interacts with Microsoft 365 Copilot (using apps such as Word, PowerPoint, Excel, OneNote, Loop, or Whiteboard), Microsoft stores data about these interactions, which includes the user's prompt and Copilot's response, including citations to any information used to ground Copilot's response. This data is said to be processed and stored in alignment with contractual commitments with the customer's other content in Microsoft 365. The data is encrypted while it's stored and it is not (as noted above) used to train foundation LLMs;
- (e) admins can use Content search or Microsoft Purview to view and manage this stored data, and they can use Microsoft Purview to set retention policies for the data related to chat interactions with Copilot;
- (f) an organisational customer's users can delete their Copilot activity history, which includes their prompts and the responses Copilot returns;
- (g) Microsoft does not claim ownership of the output of the service; at the same time, it does not make a determination on whether a customer's output is copyright protected or enforceable against other users. This is because generative AI systems may produce similar responses to similar prompts or queries from multiple customers; and
- (h) whilst abuse monitoring, which includes human review of content, is available in Azure OpenAI, Microsoft 365 Copilot services have opted out of it.

No use by Microsoft for own purposes

5.10 For the reasons above (but subject to comments below on the provision of feedback), in the author's view there is no real risk that Microsoft will use personal information that might be included in prompts for its own purposes unrelated to provision of the services.

No disclosure by Microsoft for own purposes

5.11 Microsoft's DPA is also reasonably clear that, except as required by law or to act in accordance with the customer's instructions, Microsoft will not disclose Customer Data to third parties. The DPA states:

"Microsoft will not disclose or provide access to any Processed Data except: (1) as Customer directs; (2) as described in this DPA; or (3) as required by law. For purposes of this section, "Processed Data" means: (a) Customer Data; (b) Professional Services Data; (c) Personal Data; and (d) any other data processed by Microsoft in connection

with the Products and Services that is Customer's confidential information under Customer's agreement. All processing of Processed Data is subject to Microsoft's obligation of confidentiality under Customer's agreement.

Microsoft will not disclose or provide access to any Processed Data to law enforcement unless required by law. ...

Upon receipt of any other third-party request for Processed Data, Microsoft will promptly notify Customer unless prohibited by law. Microsoft will reject the request unless required by law to comply. If the request is valid, Microsoft will attempt to redirect the third party to request the data directly from Customer.

...

Microsoft will not provide any third party: (a) direct, indirect, blanket, or unfettered access to Processed Data; (b) platform encryption keys used to secure Processed Data or the ability to break such encryption; or (c) access to Processed Data if Microsoft is aware that the data is to be used for purposes other than those stated in the third party's request."

- 5.12 One can reasonably conclude, therefore, that Microsoft will not disclose Customer Data for its own purposes.

Anthropic is a subprocessor

- 5.13 When Claude models were first introduced as an option within Copilot, if an organisation wished to utilise those models, it needed to accept Anthropic's commercial terms. However, in December 2025, that approach was deprecated.³ Now, Anthropic is treated as a subprocessor, and use of the Anthropic models is covered by and subject to the Microsoft Product Terms and the Microsoft Data Protection Addendum (DPA) discussed above.

- 5.14 However, there is still at least one significant difference: when using Claude, Microsoft's in-country processing commitments do not apply. Microsoft has said:⁴

"Anthropic models deployed in Microsoft offerings (including Microsoft 365 Copilot, Researcher, Copilot Studio, Agent Mode in Excel, and Word, Excel, and PowerPoint agents) are currently excluded from the EU Data Boundary, *and when applicable, in-country processing commitments.*"

Offshore processing

- 5.15 It is understood that Aroturuki Tamariki's Microsoft 365 tenant is hosted in Australia, i.e., that Australia is the selected data residency/processing location. Whilst this means that most processing can be expected to occur in Australia, Microsoft does make it clear that, when Copilot is used, AI inference could on occasion occur offshore:⁵

"Microsoft 365 Copilot calls to the LLM are routed to the closest data centers in the region, but also can call into other regions where capacity is available during high

³ See <https://learn.microsoft.com/en-us/copilot/microsoft-365/connect-to-ai-subprocessor>

⁴ Above n 3.

⁵ See above n 1.

utilization periods. ... Customers outside the EU may have their queries processed in the US, EU, or other regions."

- 5.16 In late 2025 Microsoft announced that that it was "making in-country data processing for customers' Microsoft 365 Copilot interactions available in 15 countries around the world" and that, by the end of 2025, Microsoft would be offering "customers in four countries—Australia, the United Kingdom, India, and Japan—the option to have Microsoft 365 Copilot interactions processed in-country".
- 5.17 **Recommendation – REC1:** It is recommended that Aroturuki Tamariki ensure that this option is selected.
- 5.18 However, even with this option selected, it seems Microsoft may still route prompts internationally when capacity in Australia is limited, because its announcement post said this:
- "This offer is designed to enable customers, particularly those in government and highly regulated industries, to access Microsoft 365 Copilot with an additional option for governance, security, and regulatory compliance. With in-country processing, Copilot interactions are processed, *under normal operations*, in data centers located within a nation's borders, giving customers greater control over their data."
- 5.19 One cannot be certain, therefore, that prompts will *only* be processed in Australia. They could still be processed offshore in abnormal circumstances.⁶ Microsoft does not provide a fixed or exhaustive list of fallback countries.
- 5.20 Whilst such offshore processing could (it seems) occur when capacity in Australia is constrained:
- (a) Microsoft still stores interaction history (prompts and outputs) in accordance with its data location commitments relating to other customer content, i.e., in Australia; and
 - (b) Microsoft's documentation indicates that any offshore processing would involve transient processing only, rather than persistent storage of Copilot interaction content outside Australia.
- 5.21 It follows that, whilst there is residual jurisdictional risk in relation processing occurring in unknown countries when capacity in Australia is constrained, in practical terms the risk is probably low. The risk is probably low because, not only is the processing transient, but Microsoft remains subject to contractual commitments described above, Copilot interactions are encrypted in transit and at rest, and processing outside Australia (when using Azure OpenAI) is the exception rather than the norm.
- 5.22 It is also important to note that Microsoft's announcement in late 2025 related specifically to "[p]rompts and responses from Microsoft 365 Copilot large language models hosted on Azure OpenAI". There is no indication that the so-called 'sovereignty option' referred to in the announcement covers prompts and outputs processed by Anthropic's Claude when that model is selected for processing. To the contrary, as noted above, Claude models are hosted

⁶ External commentary notes that Microsoft's use of the words "under normal operations" in its in-country processing language reflects an exception model where, under conditions such as capacity constraints, security incidents, or other operational disruptions, processing may be routed outside the country. See <https://windowsforum.com/threads/in-country-copilot-processing-microsofts-step-toward-regulated-ai.388009/>

by Anthropic rather than natively within the core Microsoft 365 infrastructure, with users' grounded prompts being transmitted to Anthropic-managed infrastructure (typically running on AWS or Google Cloud) for processing. It is understood that much of this third party infrastructure is in the United States, but could be in other countries as well. An Anthropic article/FAQ for commercial customers states:⁷

“Anthropic utilizes multiple cloud service providers to process customer data, as documented in our Subprocessor list. By default, we may process customer data in select countries in the US, Europe, Asia and Australia.”

- 5.23 At the date of this assessment, Microsoft does not appear to have provided a definitive list of the countries in which Anthropic processing could occur.⁸ It follows that, when Claude is selected as the applicable model, there is a measure of jurisdictional risk as the processing could occur in unknown countries.
- 5.24 In addition, although Anthropic's published materials indicate that customer API inputs and outputs are not used to train foundation models by default and may be retained for a limited operational period (historically described by Anthropic as up to 30 days), it is not possible from publicly available documentation to determine whether this retention practice applies to Copilot-mediated requests, or whether Microsoft has in place a zero data retention arrangement with Anthropic. Where Anthropic agrees to a zero data retention agreement with an API customer, Anthropic says it does not retain inputs or outputs except where required to comply with law or prevent misuse;⁹ however, as stated, it is not publicly known whether Microsoft has entered into such an arrangement for Copilot.
- 5.25 **Recommendation – REC2:** It is recommended that Aroturuki Tamariki consider whether to accept the residual jurisdictional risks referred to above for both:
- (a) Copilot using Azure OpenAI as the LLM; and
 - (b) Copilot using Anthropic Claude as the LLM,
- noting in each case that:
- (c) Microsoft is still subject to a range of contractual obligations designed to protect Copilot prompts and outputs;
 - (d) those prompts and outputs are encrypted in transit and at rest;
 - (e) where processing or storage occurs outside Australia, such handling is represented by Microsoft as transient for Azure OpenAI inference fallback, and by Anthropic as subject to limited operational retention for API processing (historically described as up to 30 days), noting that the precise retention characteristics for Copilot-mediated requests are not publicly confirmed; and

⁷ See <https://privacy.claude.com/en/articles/7996890-where-are-your-servers-located-do-you-host-your-models-on-eu-servers>

⁸ Note that 'Europe' and 'Asia' are entire regions. Some countries in those regions are considered safe, but some are not.

⁹ See <https://privacy.claude.com/en/articles/8956058-i-have-a-zero-data-retention-agreement-with-anthropic-what-products-does-it-apply-to>

- (f) both Microsoft and Anthropic are established international vendors with mature security and compliance programmes and relevant industry-standard certifications.

Consideration of other areas of potential residual risk

5.26 In addition to the offshore processing referred to above, there are three respects in which residual risk may remain, relating to:

- (a) using Copilot with the web search feature enabled;
- (b) providing feedback on Copilot outputs; and
- (c) connecting Copilot to agents.

Web search

5.27 As Microsoft explains:

- (a) Microsoft 365 Copilot has an optional feature that allows Copilot to reference web content when responding to user prompts. Allowing Copilot to reference web content can improve the quality of Copilot responses by grounding them in the latest information from the web.
- (b) When web search is enabled, Copilot parses the user's prompt and identifies terms where information from the web would improve the quality of the response. Based on these terms, Copilot generates a search query that it sends to the Bing search service asking for more information.
- (c) This generated search query is different from the user's original prompt: it consists of a few words informed by the user's prompt. The following information is not included in the generated search query sent to the Bing search service:
 - (i) the user's entire prompt, unless the prompt is very short (for example, "local weather");
 - (ii) entire Microsoft 365 files (for example, emails or documents) or files uploaded into Copilot;
 - (iii) any identifying information based on the user's Microsoft Entra ID (for example, username, domain, or tenant ID).
- (d) Whilst the generated search query does not include the entirety of Microsoft 365 documents associated with the prompt, it may be informed by data within a Microsoft 365 document under the following conditions:
 - (i) when a user enters a prompt into Copilot inside a Microsoft 365 application (for example, writing a prompt into Copilot in Microsoft Word while a relevant document is open); or
 - (ii) when the user explicitly references a specific document in their prompt.
- (e) The user's prompts and Copilot's responses (distinct from the web search query) are stored within Microsoft 365 and never leave the Copilot service boundary without customer direction.

- (f) After Copilot receives additional information from the Bing search service, this information is used to compose the response returned to the user.
 - (g) Organisational admins can determine whether the web search feature is enabled or disabled. If enabled, users can also toggle web search on and off. If an admin has disabled web search, users cannot see the toggle.
- 5.28 The web search feature has been mentioned because Microsoft gives some examples where small amounts of data (such a staff member's name or words from a confidential document) could be sent with the search query to Bing, and therefore outside the Microsoft 365 service boundary.
- 5.29 Microsoft explains that the DPA referred to above does not apply to the use of generated web search queries in Copilot or Bing. However, there is a "Microsoft 365 Copilot and Microsoft 365 Copilot Chat" section of the Product Terms which provides the following additional commitments about the generated search queries sent to the Bing search service:
- “• Microsoft has no rights in Query Data other than as needed to provide the services,
 - Query Data is not used to improve Bing,
 - Query Data is not used to create advertising profiles or track user behavior,
 - Query Data is not shared with advertisers or otherwise beyond Microsoft and its contracted suppliers who are subject to terms no less protective than these provisions,
 - Query Data is not used to train generative AI foundation models, and
 - Query Data is treated as Customer confidential information and protected by appropriate technical and organizational measures.”
- 5.30 On the basis of these statements, it would seem the use of web search poses minimal risk from a use and disclosure perspective.

Providing feedback on Copilot outputs

- 5.31 The ability to provide feedback on Copilot outputs is another feature that can be enabled or disabled by an organisation's admins.
- 5.32 When enabled, a user can select thumbs-up or thumbs-down on a response from Copilot. After a user selects one of the thumbs, a feedback pane appears and asks for more information, on what the user liked (if they selected thumbs-up) or on what went wrong (if they selected thumbs-down).
- 5.33 Microsoft explains that, in the feedback pane, Microsoft might ask for permission to collect additional data that is specific to the user's Copilot experience. If the user agrees to this collection, some or all of the following data will be collected when they select Submit:
- (a) *Prompt*: What the user asked Copilot to do. Microsoft might also collect modifications to the prompt made by the Copilot system, additional metadata fields (such as file type) that are based only on the original prompt, and/or a link to a prompt template, if the user used one to create their Copilot prompt.
 - (b) *Generated response*: What Copilot returned to the user in response to their prompt.

- (c) *Relevant content samples*: Content that helps Microsoft understand why Copilot gave the user that generated response. These samples could include the following:
- (i) The history of the user's chat with Copilot.
 - (ii) Files or other content that Copilot used to formulate the generated response, which might include content not shown or referenced in the final response. This is content that the user already has permission to access.
 - (iii) The actual prompt, which includes Microsoft proprietary content, that was sent to the LLM.
- (d) *Additional log files*: Log files about Microsoft programs or services the user is using. These files might contain the content of the user's files, their name or email address, their IP address, or other personal data. Microsoft says these files might be the only way to determine why Copilot gave the response it did.

5.34 Microsoft says it does not "use ... feedback to train the foundation models used by Copilot with Microsoft 365 apps", but may use it to improve Copilot. It says:

"With your feedback, we can do the following:

- Improve Copilot product quality, including resolving issues more quickly with specific information about what you experienced.
- Prioritize feature requests that will enhance your Copilot experience."

5.35 It is understood that Aroturuki Tamariki will disable the feedback feature to reduce any risk arising from Microsoft's review of feedback that could contain confidential organisational or personal information.

5.36 **Recommendation – REC3**: It is recommended that the disabling of this feature be confirmed internally in writing.

Agents

5.37 An agent in Microsoft 365 Copilot is a specialised component or extension that helps Copilot retrieve information or perform actions by connecting to other data sources, applications, or APIs. When these agents are used to access external or third-party systems, they may handle data beyond Microsoft's standard enterprise boundaries.

5.38 For this reason, Microsoft advises organisations to review each agent's privacy statement and terms of use to understand how it collects, stores, and shares information.

5.39 It is understood that, at least in the early stages of using Copilot, Aroturuki Tamariki is not using agents. If that were to change and third party agents or systems were to be used, one would need to check all applicable privacy statements and terms of use, and potentially extend this PIA and any cloud risk assessment.

5.40 **Recommendation – REC4**: It is recommended:

- (a) that Aroturuki Tamariki adopt a formal position that staff are not to use agents with Copilot without approval from the AI Officer and only after required privacy impact and cloud risk assessments have been completed and considered; and
- (b) if possible, that staff be technically prevented from using agents without IT approval.

6. Personal information to which Copilot will have access

6.1 It is understood that Copilot:

- (a) will have access to general, back-office, administrative data that staff members using it are already authorised to view when using Microsoft 365 applications (in other words, access is governed by the permissions attached to a user); but
- (b) will *not* have access to personal information of tamariki, rangatahi and their whānau that Aroturuki Tamariki has collected through its monitoring function, either because it is stored in Te Pataka¹⁰ or because certain file libraries will be marked as off-limits for Copilot.

6.2 The libraries marked as off-limits for Copilot relate to monitoring, human resources, and other staff-related matters.

6.3 Copilot will not, therefore, have access to personal information relating to, among other things, staff recruitment, performance, or disciplinary processes.

6.4 The information that *could* be drawn upon by Copilot through Microsoft Graph that is or might be considered personal information may include:

- (a) names, email addresses and other contact information for staff, Board members, and contact representatives of other organisations with whom Aroturuki Tamariki interacts, including contacts at certain IT support suppliers (such as Datacom and Kordia);
- (b) user preferences and settings;
- (c) user activity and usage data; and
- (d) personal information that could be included in a user's email, calendar, or documents created in the likes of Word and Excel (but excluding, as noted above, personal information of tamariki, rangatahi and their whānau that is stored in Te Pataka or is in libraries marked as off-limits for Copilot).

7. Technical and non-technical controls

Technical controls

7.1 It is understood that a range of technical controls are or will be in place to protect data from unauthorised use or disclosure, including the following:

- (a) Copilot is configured to prevent prompts and outputs being used to train the large language models that Copilot utilises;
- (b) Aroturuki Tamariki will not enable third party integrations;
- (c) through user authentication and role-based access permissions, when a user uses Copilot, Copilot will only be able to access data to which the user otherwise has access (and excluding the information referred to in paragraphs 6.1(b) and 6.2 above); and

¹⁰ Te Pataka is Aroturuki Tamariki's Salesforce instance. It is where monitoring information is held, including the details surrounding visits and times, and the interview notes collected during the monitoring process.

- (d) as part of the Microsoft 365 suite of applications, data processed via Copilot will be protected by controls in place relating to Microsoft 365 (which has been the subject of previous certification and accreditation), including encryption of data in transit and at rest.

Non-technical controls

- 7.2 Copilot will only have access via Microsoft Graph to the limited amounts of personal information referred to in paragraph 6 above. However, the fact remains that staff members could copy and paste personal information to which they have access into prompts (beyond what is authorised) which would then be processed by Copilot.
- 7.3 To mitigate this risk, prior to being granted access to Copilot, staff will be required to:
 - (a) confirm they have read and agree to comply with the AI Policy and AI Usage Guide; and
 - (b) attend AI training.
- 7.4 The AI Policy:
 - (a) is clear that "[n]o personal information about tamariki, rangatahi, or whānau may be entered into AI systems";
 - (b) states that "except to the extent permitted by authorised use cases, no personal information about kaimahi may be entered into AI systems";
 - (c) instructs staff to "not enter or upload sensitive, personal, confidential, or other classified information into any AI tool or platform that is not approved for that data type by ELT"; and
 - (d) states that "[u]sing an AI system in violation of this policy – particularly in relation to confidential or personal information – is a serious breach of protocol and may result in disciplinary action".
- 7.5 It is also understood that Aroturuki Tamariki will monitor compliance with its AI Policy and AI Usage Guide.

8. Privacy assessment

- 8.1 As noted earlier, this PIA assesses privacy-related risks relating to Aroturuki Tamariki's proposed use of Copilot by reference to the Privacy Act and the Oversight of Oranga Tamariki System Act 2022 and the Information Rules under that Act. Each is considered in turn.

Privacy Act 2020

8.2 Threshold issue – Does section 11 of the Act apply

- (a) If an agency is proposing to include people's personal information in prompts to an AI tool, it is important to consider whether section 11 of the Privacy Act applies. This is important because, if section 11 *does* apply, the agency will not be 'disclosing' the personal information to the AI tool provider for the purposes of the Act, and that will mean some of the Act's information privacy principles (**IPPs**) will not apply to the transfer of information to the AI tool.
- (b) To understand this, we need to look at the content of section 11.

- (c) Under section 11, if an agency (let's call it A) uses a service provider to store or process personal information, and the service provider does *not* use or disclose the personal information for its own purposes, the information is deemed to be held by A, and there is no 'disclosure' to the service provider for the purposes of the Act (that is what this PIA means when referring to section 11 'applying'). Various cloud services were intended to be covered by this provision. However, if the service provider *does* use or disclose the information for its own purposes, then the personal information is to be treated as being held by both A and the service provider. In that situation, the preferable interpretation of section 11 is that, in transferring the information to the service provider, A would be 'disclosing' it to the service provider for the purposes of the Act.
- (d) If section 11 *does* apply and there is no 'disclosure' to the AI tool provider, then a couple of things follow in relation to the inclusion of personal information in prompts:
 - (i) first, when the agency collects the personal information directly from individuals, the agency most probably¹¹ does not need – under IPP3 – to mention that it is using the AI tool to process the personal information; and
 - (ii) second, issues will not arise under IPP11 (which imposes limits on the disclosure of personal information) or IPP12 (which relates to disclosure of personal information outside New Zealand).
- (e) An important factor in determining whether section 11 applies is whether the AI tool provider will use the agency's inputs and outputs for further training of the underlying models. If it will, section 11 will not apply, and the agency *will* be 'disclosing' the personal information to the AI tool/provider.
- (f) In relation to Aroturuki Tamariki's proposed use of Copilot, it is clear that any personal information included in prompts will *not* be used to train the models on which Copilot relies.
- (g) That is not necessarily the end of the analysis, though, because there could still be a 'disclosure' for the purposes of the Act if the AI tool provider uses or discloses the inputs and outputs in other ways that amount to 'for its own purposes'. To determine if this is the case, it can be important to review the provider's terms of use, privacy statement, and any other relevant documentation.
- (h) Based on the analysis in section 5 above, one can be confident that neither Microsoft nor Anthropic (where Claude is used as the preferred model) will 'use' or 'disclose' any personal information included in prompts for purposes unrelated to provision of the Copilot service to Aroturuki Tamariki users. Therefore, for all material intents and purposes, they will not use or disclose the information for their own purposes.
- (i) It follows that:
 - (i) individuals whose personal information might be included in prompts to Copilot

¹¹ This PIA uses the term "most probably" because this issue has not, to the author's knowledge, been considered by the Office of the Privacy Commissioner, the Human Rights Review Tribunal, or the courts. The position stated here is, however, considered to be correct. If there is no 'disclosure' to another entity for the purposes of the Act, then logically it would seem to follow that there is no intended 'recipient' for the purposes of IPP3.

do not need to be informed of that possibility under IPP3; and

- (ii) in relation to the transfer of information to Microsoft and Anthropic, no ‘disclosure’ issues arise under IPP11 or IPP12.

8.3 Information privacy principles

The table below assesses whether the proposed use of Copilot will be consistent with applicable information privacy principles in section 22 of the Act and associated risks in that regard.

Collection of personal information		
IPP	Analysis	Status ¹²
IPP1 Purpose of collection Only collect personal information if it's necessary to do so for a lawful purpose connected with the agency's functions. If the lawful purpose for which personal information about an individual is collected does not require the collection of their identifying information, the agency "may not require the individual's identifying information".	Under the proposed use cases, Aroturuki Tamariki staff will not, through their use of Copilot, be collecting new information about identifiable individuals from or via Copilot, and the agency's AI Policy makes it clear that "[u]sing AI to collect personal information about individuals from third party sources" is "strictly not allowed". There is always a risk that staff <i>could</i> seek to collect new personal information about individuals, as AI tools like Copilot <i>can</i> be used to collect new personal information about individuals (even when the generated information is derivative, opinionative, or evaluative) based on data to which the tools have access. However, as long as agency staff are advised of the prohibition on doing so during the training and comply with the AI Policy, no issues will arise under IPP1. Recommendations: • REC5: Ensure the staff training covers the prohibitions in the AI Policy, and remind staff of the statement in the Policy that "[u]sing an AI system in violation of [the] policy – particularly in relation to confidential or personal information – is a serious breach of protocol and may result in disciplinary action". • REC6: Undertake periodic, randomised spot checks as to whether staff are seeking to collect new personal information about individuals. It is noted that Copilot may, through Microsoft Graph, draw on personal information in the data sources to which it has access through Microsoft Graph (to ground or provide context for prompts), but that does not entail the collection by Aroturuki Tamariki of new personal information. Copilot would merely be drawing on information already held by the agency.	ORANGE
IPP2 Source Personal information must be collected from the individuals concerned, unless an exception applies.	As per the discussion of IPP1, the use of Copilot does not of itself involve the collection by agency staff of new personal information. Therefore, as long as agency staff comply with the AI Policy, no issues will arise under IPP2.	GREEN

¹² Key: Green means no issues under the IPP. Orange means there is an issue or risk but it may be able to be addressed or mitigated. Red means there is an issue or risk that cannot be sufficiently addressed or mitigated at this time. Colour names are included for those who cannot see the colours.

IPP3 Collection from subject

If an agency collects personal information directly from individuals it must, unless an exception applies, take reasonable steps to inform the individuals of the matters listed in IPP3, which include, among other things, that the information is being collected, the purpose(s) of collection, and the intended recipients.

When an agency is proposing to use an AI tool like Copilot in a manner that may entail the entry into prompts (either directly or via something like Microsoft Graph) of personal information that is yet to be collected from individuals, one may need to consider whether those individuals need to be informed that their information may be shared with the AI tool provider. The answer to this question turns on whether there will be a “disclosure” to the AI tool provider for the purposes of the IPPs, and that requires consideration of section 11 of the Act.

Section 11 has been discussed above. As noted in that discussion, section 11 applies to Aroturuki Tamariki’s use of Copilot, because we can be confident that neither Microsoft, nor Anthropic (where Claude is used as the preferred model), will ‘use’ or ‘disclose’ any personal information included in prompts for their own purposes. The main legal consequence of that is that Aroturuki Tamariki’s entry of personal information into Copilot (to the extent this may occur, either directly or via Microsoft Graph, which is expected to be minimal) will not constitute a “disclosure” of personal information for the purposes of the IPPs.

On this basis, there is a good argument that:

- Microsoft and Anthropic are not “intended recipients” for the purposes of IPP3; and, therefore
- even in relation to personal information Aroturuki Tamariki is yet to collect from individuals (e.g., as part of new staff onboarding) which may at some point be included in prompts to Copilot, IPP3 does not require the agency to inform those individuals that their information may be used in Copilot.

For these reasons, no issues arise under IPP3.

There is a separate question, however, as to whether it would be desirable for Aroturuki Tamariki to be transparent with staff and others about potential use of certain categories of personal information in AI tools like Copilot. In the present context, this is not a Privacy Act question. Rather, it is a question that stems from the desirability of ensuring people have trust and confidence in how public sector agencies like Aroturuki Tamariki handle the personal information they hold.

It is also consistent with the Government's *Responsible AI Guidance for the Public Service: GenAI*, which states:¹³

"Be open with your people and the public about why and how you're using GenAI. Public service agencies need to be able to explain what information goes into AI systems and what the results are used for.

...

A lack of transparency can lead to harmful outcomes, public distrust, and no-one being responsible for the final decision. Transparency involves communicating clearly that you're using AI and why you're using it."

GREEN

¹³

See <https://www.digital.govt.nz/standards-and-guidance/technology-and-architecture/artificial-intelligence/responsible-ai-guidance-for-the-public-service-genai/customer-experience/transparency>

	Recommendation – REC7: Consider preparation and publication of a GenAI Transparency Statement.	
IPP3A Collection of personal information other than from individual concerned (Comes into force on 1 May 2026.) If an agency collects personal information other than from the individuals concerned (that is, from a third party source), it must, unless an exception applies, take reasonable steps to inform the individuals of the same kinds of things listed in IPP3 (there are a couple of differences but for the most part the information is the same).	IPP3A is not yet in force but will be soon (1 May 2026) and so the comments below proceed on the basis that it is in force. No issues arise under IPP3A. As per the discussion of IPP1, the use of Copilot does not of itself involve the collection by agency staff of new personal information. Therefore, as long as agency staff comply with the AI Policy, no issues will arise under IPP3A. Note that if agency staff were to collect new personal information from or through Copilot, IPP3A might apply to require disclosure, even when section 11 applies and there is no 'disclosure' from Aroturuki Tamariki to Microsoft or Anthropic or vice versa. This is because IPP3A does not refer to collecting information "from another person, agency or organisation". The precise wording is: "If an agency collects personal information about an individual other than from the individual concerned...". That can be interpreted as meaning any other source.	GREEN
IPP4 Manner of collection When collecting personal information, an agency needs to do so by a means that is lawful and that, in the circumstances (particularly where personal information is being collected from children or young persons) is fair and does not intrude to an unreasonable extent upon the personal affairs of those individuals.	As noted above, Aroturuki Tamariki staff will not, through their use of Copilot, be collecting new information about identifiable individuals. For this reason, no issues arise under IPP4 in relation to the use of Copilot. (For completeness, it is noted that there can be situations in which issues might arise under IPP4 where an agency collects personal information either from vulnerable individuals to process with an AI tool or directly from an AI tool about individuals. That is not the case here, but the point is noted for future reference.)	GREEN
Storage and security		
IPP	Analysis	Status
IPP5 Storage and security An agency needs to ensure, in relation to the personal information it holds, that there are reasonable safeguards against: loss; access, use, modification or disclosure not authorised by the agency; and other misuse. When the information needs to be given to another party for provision of a service to the agency, the agency needs to do everything reasonably within its power to prevent unauthorised use or disclosure.	The contractual controls and the technical and non-technical controls in place to protect any personal information that could be included in a prompt (either directly or via Microsoft Graph) are referred to in paragraphs 4, 5, and 7 above. In addition, it is understood that Aroturuki Tamariki has previously certified and accredited its Microsoft 365 environment and will be undertaking certification and accreditation processes for Copilot. On the basis of the controls referred to above and other information currently available, one can be confident that Aroturuki Tamariki is and will be complying with IPP5. (This PIA does not assess steps that may be considered necessary or desirable to address residual risks that might be identified in the certification of Copilot.)	GREEN
Access and correction		
IPP	Analysis	Status
IPP6 Access	Aroturuki Tamariki's proposed use of Copilot does not raise issues under IPP6. Not only is minimal (if any) personal information likely	GREEN

Individuals are entitled to access personal information that an agency holds about them and, if access is granted, they must be advised that they may request correction of that information. The right of access is subject to a range of exceptions in Part 4 of the Act.	to be used in prompts (either directly or through Microsoft Graph), but any source personal information that is used can be expected to remain available at its original source, and outputs that users wish to save will be stored within the Microsoft 365 environment.	
IPP7 Correction Individuals are entitled to request correction of the personal information an agency holds about them.	Aroturuki Tamariki's proposed use of Copilot does not raise issues under IPP7. Again, any source personal information that is used (if any at all) can be expected to remain available at its original source, and outputs that users wish to save will be stored within the Microsoft 365 environment. Correction requests can be made and processed in the normal way.	GREEN
Use and retention of personal information		
IPP	Analysis	Status
IPP8 Accuracy An agency must not use or disclose personal information without taking any steps that are, in the circumstances, reasonable to ensure it is accurate, up to date, complete, relevant, and not misleading.	To comply with IPP8 in relation to use, Aroturuki Tamariki staff should take reasonable steps in the circumstances to ensure that any personal information permissibly entered into prompts and any personal information in outputs is accurate before the information is used. Given: - the low volume and basic nature of any personal information that might be processed by Copilot; and - the AI Policy's requirement that "AI-generated content must be checked to ensure it is accurate and appropriate for the task it is to be used for", this accuracy risk is considered to be low. That said, there is a known cognitive bias that can affect some (certainly not all) people using AI tools. The cognitive bias is to assume the AI tool is right. Confident language often used by large language models can generate or reinforce this bias. Recommendation – REC8-A: Remind staff during their required AI training that Copilot-generated outputs are not necessarily accurate, and need to be checked for accuracy before use.	ORANGE
IPP9 Retention An agency that holds personal information must not keep that information for longer than is required for the purposes for which the information may lawfully be used.	There is unlikely to be any issue under IPP9. Any personal information included in Copilot inputs and outputs – which is expected to be minimal – will be stored in Aroturuki Tamariki's Microsoft 365 environment as part of those inputs and outputs. Those inputs and outputs can be expected to be stored or deleted in accordance with Aroturuki Tamariki's Public Records Act obligations and any disposal authority under that Act.	GREEN
IPP10 Limits on use Personal information that an agency obtains in connection with one purpose must not be used for another, unless an exception applies.	There is unlikely to be any issue under IPP10, as staff are most likely to be using any personal information in prompts (either directly or via Microsoft Graph) – which is expected to be minimal – consistently with the purposes for which it was obtained by Aroturuki Tamariki in the first place. That said, there is always a risk of staff using personal information in prompts in unexpected ways that are not consistent with IPP10.	ORANGE

	Recommendation – REC9: During staff’s required AI training: - remind them of the limited use cases for which personal information may be entered into Copilot prompts (either directly or via Microsoft Graph); and - inform them that, when personal information can legitimately be entered into prompts in accordance with the AI Policy and permitted use cases, that personal information should not be used or processed in ways that are unusual or unexpected, that is, that might be considered inconsistent with the purposes for which Aroturuki Tamariki obtained it.	
Disclosure of personal information		
IPP	Analysis	Status
IPP8 Accuracy An agency must not use or disclose personal information without taking any steps that are, in the circumstances, reasonable to ensure it is accurate, up to date, complete, relevant, and not misleading.	To comply with IPP8 in relation to disclosure, Aroturuki Tamariki staff should take steps to ensure that any outputs that contain personal information are accurate before the information is disclosed outside of Aroturuki Tamariki (if this ever happens at all). As with the consideration of IPP8 above in the context of the use of personal information, given: - the low volume and basic nature of any personal information that might be processed by Copilot; and - the AI Policy’s requirement that “AI-generated content must be checked to ensure it is accurate and appropriate for the task it is to be used for”, this accuracy risk is considered to be low. Recommendation – REC8-B: Remind staff during their required AI training that Copilot-generated outputs are not necessarily accurate, and need to be checked for accuracy before any permissible disclosure outside of Aroturuki Tamariki.	ORANGE
IPP11 Limits on disclosure An agency must not disclose personal information it holds unless an exception in IPP11 applies.	As explained above in the discussion of section 11, the use of Copilot does not involve the ‘disclosure’ of information to Microsoft or Anthropic for Privacy Act purposes. Therefore, no issues arise under IPP11 in relation to the inclusion of personal information (which in any event is expected to be minimal) in prompts, whether directly or via Microsoft Graph.	GREEN
IPP12 Disclosure overseas IPP12 regulates the disclosure of personal information outside New Zealand. It seeks to ensure that, when information is disclosed offshore, there are comparable safeguards to those in the Privacy Act. In essence, an agency can only disclose personal information to a foreign person or entity (i.e., overseas), in reliance on certain listed IPP11 exceptions, if one of a number of conditions is satisfied.	As explained above in the discussion of section 11, the use of Copilot does not involve the ‘disclosure’ of information to Microsoft or Anthropic for Privacy Act purposes. Therefore, no issues arise under IPP12 in relation to the inclusion of personal information in prompts and the associated transfer of personal information offshore (primarily, but not necessarily only, to Australia (in the case of Microsoft) and it seems the United States (in the case of Anthropic)).	GREEN

Unique identifiers		
IPP	Analysis	Status
IPP13 Unique identifiers IPP13 regulates the ‘assignment’ of unique identifiers (i.e., an identifier other than the individual’s name that uniquely identifies the individual).	No issues arise under IPP13 in relation to the proposed use of Copilot, as Aroturuki Tamariki is not proposing to assign unique identifiers to individuals for such purposes.	GREEN

Oversight of Oranga Tamariki System Act 2022 and the Information Rules under that Act

- 8.4 Information-related provisions are found in Part 3 of the Oversight of Oranga Tamariki System Act (**Oversight Act**) and in the Information Rules made by the Monitor in accordance with section 49 of that Act.
- 8.5 On the author’s understanding of the proposed use cases, Aroturuki Tamariki’s proposed use of Copilot does not appear to be inconsistent with any provision of Part 3 of the Act or the Information Rules.
- 8.6 However, the Monitor does need to be mindful of the consultation obligation in rule 7.2 of the Information Rules, which states:
- 7.2 The Monitor will not implement any algorithm (other than for the purpose of de-identifying information) or predictive modelling in connection with the performance of its functions without:
- (a) first consulting representative agencies and organisations with an interest in, or who could be affected by, the use of the algorithm or modelling; and
 - (b) complying with the Algorithm Charter for Aotearoa New Zealand. *[sic]*
 - (c) considering the Algorithm Impact Assessment process and applying the Algorithm Impact Assessment toolkit.
- 8.7 Under rule 7.3, the consultees referred to in rule 7.2(a):
- (a) must include the Office of the Privacy Commissioner, Te Puni Kōkiri, StatsNZ, Oranga Tamariki, Mana Mokopuna – Children’s Commissioner, and the Centre for Data Ethics and Innovation; and
 - (b) may include such other individuals, agencies and organisations as the Monitor considers appropriate (including representatives of, for example, communities, cultural groups or children or young persons).
- 8.8 Rule 7.2 was written when the Monitor was a departmental agency, that is, before it became an independent Crown entity. On a purposive and contextual reading, rule 7.2 is directed at algorithms or predictive modelling implemented as part of the Monitor’s statutory monitoring functions under section 14 of the Oversight Act.

8.9 Section 14 states:

14 Monitoring function

- (1) The function of the Monitor is to monitor the performance of the Oranga Tamariki system in the context of its interface with other systems.
- (2) For the purposes of subsection (1), the function includes (without limitation)—
 - (a) assessing compliance with the Oranga Tamariki Act 1989, national care standards regulations, and other regulations and standards made under that Act by the chief executive of Oranga Tamariki and approved providers;
 - (b) assessing the quality and impacts of service delivery, service mix, service resourcing, and practices on the experiences of children, young people, families, and whānau;
 - (c) assessing outcomes for children, young people, families, and whānau who receive services or support through the Oranga Tamariki system, and changes in outcomes over time, with particular regard to Māori children and young people and their whānau.

8.10 Whilst Copilot is, technically, an algorithmic system, rule 7.2 is concerned with the Monitor implementing algorithmic or predictive systems as part of its monitoring methodology. The use of general-purpose AI functionality embedded in office productivity software for routine internal administrative tasks (such as summarising notes, preparing agendas, or summarising internal email threads) does not amount to the “implementation” of an algorithm or predictive modelling in connection with the performance of the Monitor’s statutory monitoring functions under section 14 of the Oversight Act as contemplated by rule 7.2. Although the phrase “in connection with” in rule 7.2 is broad, in the author’s view it requires a substantive functional nexus with the monitoring function, which is not satisfied by routine internal back-office use of Copilot.

8.11 As an independent Crown entity, the Monitor has the functions in section 14 of the Oversight Act (the focus of rule 7.2) and, under section 14(1)(c) of the Crown Entities Act, “any functions that are incidental and related to, or consequential on, [these] functions”. The use of general-purpose AI functionality embedded in office productivity software for routine internal administrative tasks is activity falling within the latter, incidental functions, which were not and are not the focus of rule 7.2.

8.12 It follows that the consultation, Algorithm Charter and related obligations in rule 7.2 are not engaged by such uses of Copilot.

8.13 There is, however, an important caveat. The position above depends on Copilot not being used in any way that materially shapes monitoring analysis, monitoring evaluative judgments, monitoring prioritisation, or monitoring outputs. For example, if Copilot were used for:

- (a) drafting evaluative judgments about the performance of providers in the Oranga Tamariki system;
- (b) synthesising evidence in a way that materially shapes findings;
- (c) prioritising monitoring targets or themes;

- (d) generating risk characterisations of cohorts or providers;
- (e) producing analytic content that feeds into published monitoring outputs; or
- (f) any other use that materially influences monitoring judgments, prioritisation, or published outputs,

Aroturuki Tamariki would be much closer to implementing algorithmic assistance in connection with the performance of its section 14 function, and the consultation obligation in rule 7.2 could be triggered.

8.14 Recommendation – REC10: It is recommended that Aroturuki Tamariki:

- (a) review its list of proposed use cases; and
- (b) not proceed with any such use cases that fall within the kinds of activities listed or similar to those in paragraph 8.13 without first undertaking the consultation required by rule 7.2.

9. Privacy and related risks and recommendations

- 9.1 The table below sets out risks that are considered to be material and, where possible, how those risks can be mitigated. It captures risks and mitigations set out in the analysis above of how Copilot works, the Privacy Act's IPPs, and the Oversight Act and Information Rules. Where steps can be taken to mitigate a risk, an opinion is given on whether the mitigation is sufficient.

Risk	Recommended mitigations	Sufficient?
In-country (Australian) versus offshore processing: Failure to take advantage of Microsoft's recent introduction in some countries, including Australia, of an in-country processing option for Copilot.	REC1: Take advantage of the Copilot in-country processing option (meaning Copilot processing will usually, but not invariably, occur in Australia when Azure OpenAI is the selected model).	Minimises offshore processing (beyond Australia) when Azure OpenAI is used but residual risk remains (see next item)
Residual offshore jurisdiction processing risk: Where Azure OpenAI is the selected model, processing could still occur offshore during times of capacity constraint in Australia, and where Claude is the preferred model processing either will or is more likely to occur offshore.	REC2: Consider whether to accept the residual jurisdictional risks referred to, for both: (a) Copilot using Azure OpenAI as the LLM; and (b) Copilot using Anthropic Claude as the LLM, noting in each case that: (c) Microsoft is still subject to a range of contractual obligations designed to protect Copilot prompts and outputs; (d) those prompts and outputs are encrypted in transit and at rest; (e) where processing or storage occurs outside Australia, such handling is represented by Microsoft as transient	-- (Not a mitigation; rather, an acceptance of residual risk)

	for Azure OpenAI inference fallback, and by Anthropic as subject to limited operational retention for API processing (historically described as up to 30 days), noting that the precise retention characteristics for Copilot-mediated requests are not publicly confirmed; and (f) both Microsoft and Anthropic are established international vendors with mature security and compliance programmes and relevant industry-standard certifications.	
Feedback feature not disabled: Feature not disabled, resulting in potential review by Microsoft of confidential organisational or personal information included in feedback.	REC3: Confirm internally in writing the disabling of the feedback feature.	Yes
Use of Copilot agent extensions: Use of agents without approval and therefore potentially sharing personal information with third party services under unassessed conditions.	REC4: Adopt a formal position that staff are not to use agents with Copilot without approval from the AI Officer and only after required privacy impact and cloud risk assessments have been completed and considered. Also, if possible, lock down the technical ability for staff to use agents without IT approval.	Yes
IPP1: Risk that staff could seek to collect new personal information about individuals from or via Copilot, as AI tools like Copilot can be used to collect new personal information about individuals (even when the generated information is derivative, opinionative, or evaluative) based on data to which they have access.	REC5: Ensure the staff training covers the prohibitions in the AI Policy, and remind staff of the statement in the Policy that “[u]sing an AI system in violation of [the] policy – particularly in relation to confidential or personal information – is a serious breach of protocol and may result in disciplinary action”. REC6: Undertake periodic, randomised spot checks as to whether staff are seeking to collect new personal information about individuals.	Yes
Transparency: Whilst in the present context not a legal Privacy Act risk, it may be desirable for Aroturuki Tamariki to be transparent with staff and others about potential use of certain categories of personal information in AI tools like Copilot. This stems from the desirability of ensuring people have trust and confidence in how public sector agencies like Aroturuki Tamariki handle the personal information they hold.	REC7: Consider preparation and publication of a GenAI Transparency Statement.	Yes

IPP8: Copilot outputs containing personal information (which, on the proposed use cases, should be a rare occurrence) are not checked for accuracy before use or disclosure.	REC8-A+B: Remind staff during their required AI training that Copilot-generated outputs are not necessarily accurate and need to be checked for accuracy before use and before any permissible disclosure outside of Aroturuki Tamariki.	Yes
IPP10: Risk of staff using personal information (to the extent permitted by the permissible use cases) in prompts in unexpected ways that are not consistent with IPP10.	REC9: During staff's required AI training: • remind them of the limited use cases for which personal information may be included in Copilot prompts; and • inform them that, when personal information can legitimately be entered into prompts in accordance with the AI Policy and permitted use cases, that personal information should not be used or processed in ways that are unusual or unexpected, that is, that might be considered inconsistent with the purposes for which Aroturuki Tamariki obtained it.	Yes
Consultation failure: Not consulting when required under rule 7.2 of the Information Rules.	REC10: It is recommended that Aroturuki Tamariki: • review its list of proposed use cases; and • not proceed with any such use cases that fall within the kinds of activities listed or similar to those in paragraph 8.13 without first undertaking the consultation required by rule 7.2.	Yes

10. Addendum post-pilot – 26 May 2026

10.1 Context

- (a) A group of 14 Aroturuki Tamariki staff participated in a pilot of Copilot from 13 April to 1 May 2026. The purpose of the pilot was to trial and test Copilot within the parameters of the agency's AI Policy and Usage Guidance.
- (b) The pilot was designed to identify practical benefits and potential challenges with the agency's organisational settings, and to test the security settings to ensure that Copilot was only able to access information as intended.
- (c) Two issues were identified that have a bearing on this Privacy Impact Assessment and that warrant this Addendum. The first issue concerns information access. The second relates to working within the boundaries of the agency's AI Policy.

10.2 Issue 1: Copilot's ability to access off-limits documents

- (a) Pilot participants confirmed they could not get Copilot to access documents outside of their personal access settings within Microsoft 365. For example, the participants could not get Copilot to access personnel files that those participants cannot usually access due to technical access restrictions.
- (b) However, controls testing carried out by one member of staff identified the need for an additional layer of control within the agency's document libraries. This issue concerns libraries that staff can ordinarily access but which Aroturuki Tamariki had decided to prevent Copilot from accessing (for example, a document library containing engagement notes).
- (c) During testing, three staff members were asked to deliberately try to access information within library documents that it was thought Copilot should not be able to access.
- (d) One staff member was able to access documents they should not have been able to access. This revealed an unforeseen gap in controls: the then-current control during the pilot prevented Copilot from conducting a general search of a library designated as off-limits, but if a staff member requested a specific document that they were (in non-Copilot settings) able to access, Copilot could access it even if the document resided in a library that was marked as off-limits for Copilot search.
- (e) On the face of it, settings within Copilot suggested these libraries would have been locked down. However, further research revealed that sensitivity labels need to be applied not only to the document library, but also to each individual document within it. It is understood that Aroturuki Tamariki is implementing these labels across all relevant libraries and documents and will develop training guidance for staff on the issue. In addition, because this issue could be experienced by other agencies, Aroturuki Tamariki has shared this learning with the Chief Government Digital Officer.
- (f) The identification of this issue demonstrates the importance of agencies conducting controlled pilots of AI tools before rolling them out to staff. The agency's decisions to conduct a pilot and to deliberately test the limits of what Copilot could access are what enabled this control gap to be identified and remedied.
- (g) If not identified and remedied before Copilot roll-out, this control gap might have enabled staff to act in a manner that exceeded the boundaries of the agency's AI

Policy or decisions made by management as to libraries that would be marked off-limits to Copilot. However, if left unchecked and even if such documents were to contain personal information, infringement of the IPPs arising from Copilot use would be unlikely. There are two reasons for this conclusion:

- (i) first, the staff members concerned already have access to the relevant documents in non-Copilot settings; and
 - (ii) second, the other controls on Microsoft's use of input and output data prevent the data from being misused.
- (h) **Recommendation:** Once Aroturuki Tamariki has implemented the applicable sensitivity labels across all relevant libraries and documents, it is recommended that the agency repeat the testing that identified the issue to ensure it cannot recur.

10.3 Issue 2: Limits on using Copilot with de-identified quotes from interviews

- (a) Pilot findings revealed that one staff member had used Copilot for thematic analysis that involved taking keywords from de-identified interview quotes and asking Copilot for ideas on how to group them thematically. Full quotes were *not* entered into Copilot. Only non-personal keywords were entered (for example, keywords relating to sport). The purpose of the staff member using Copilot in this way was to support their work in writing an agency 'share back' document.
- (b) The staff member noted in a post-pilot meeting that it would be helpful to be able to put full de-identified quotes into Copilot and ask Copilot to group them, while confirming that he had not used it in that way.
- (c) What was done during the pilot (as described in paragraph 10.3(a)) does not raise issues under the IPPs. No personal information was being used. For the same reason, it does not appear to raise issues under the AI Policy. It was 'productivity support' (an approved use) and did not infringe the prohibition on using AI systems to process personal information about tamariki, rangatahi or whānau without approval from the AI Officer and ELT.
- (d) By contrast, putting full de-identified quotes into Copilot and asking Copilot to group them might fall on the other side of the line. There are two risk-related reasons for this:
 - (i) First, there is a probably low risk of such quotes still containing personal information about tamariki, rangatahi or whānau (a risk which, it is understood, the agency goes to great lengths to mitigate). This is more a question of re-identification risk than a question of the de-identified quotes containing any obvious identifiers. However, given the risk, it is **recommended** that staff not be permitted to enter full interview quotes into Copilot without approval from the AI Officer and ELT.
 - (ii) Second, there is also a probably small risk that entering full de-identified quotes into Copilot for the purpose of assisting with thematic analysis of quotes for a share back report could be perceived as engaging rule 7.2 of the Information Rules and therefore requiring prior consultation as described in that rule. It is not clear that rule 7.2 would be engaged, and this Addendum is not the place to consider that issue in detail. It is **recommended** that Aroturuki Tamariki consider this question further before deciding to allow staff to enter de-identified quotes from interviews into Copilot.