

Senate Select Committee on Adopting Artificial Intelligence

Responses to Additional Questions on Notice

Amazon Commercial Services Pty Ltd and Amazon Web Services Australia Pty Ltd (AWS) (collectively, Amazon) provide the following responses to additional written questions on notice provided by the Committee's secretariat on 9 September 2024.

Question 1

At the hearing Amazon was asked “do you believe there should be a regulatory transparency requirement on Amazon and your competitors”, to which Ms Foster said “Yes, I agree that it should happen.”

Please advise (answering separately for each):

- a. Would Amazon support a regulated requirement to disclose the sources of data used to train its AI models?*
- b. Would Amazon support a regulated requirement to disclose what copyrighted data has been used in the training of AI models?*
- c. Would Amazon support a regulated requirement to disclose what licenced data has been used in the training of AI models?*
- d. Would Amazon support a regulated requirement to disclose what personal information has been used in the training of AI models?*
- e. Would Amazon support a regulated requirement to disclose what data has been used to train AI models that is data stored or otherwise accessible by Amazon from customers of Amazon and AWS?*
- f. Would Amazon support a regulated requirement to disclose how data sources have been selected and curated?*

We recognise that transparency can play an important role in fostering trust in AI and helping people make more informed decisions about their use of our AI products and services. We have therefore been proactive in implementing mechanisms that enhance public understanding and awareness of our AI services. At Amazon, we were among the first to provide service cards, disclosing key information about our AI services to our customers. The AI service cards for our generative AI services provide robust disclosures about our models' capabilities and limitations, performance metrics, intended use cases, and responsible AI considerations. For example, please refer to the [service card](#) for Titan Text Premier.

We are invested in working with governments and international organisations to support the development of thoughtful, well-calibrated AI transparency regulations and standards that balance the values promoted by any required disclosures against the risks of such disclosure (including cybersecurity and national security risks) and the proprietary rights of businesses in maintaining their confidential business information. Transparency requirements should also consider the proper audiences for disclosure, as certain information might be safely made available to regulators, but would create undue risk to disclose publicly.

Question 2

At the hearing, both Mr Levey and Ms Foster at various points referred to Amazon using “publicly available” content or data to train its AI models, in addition to data that is licenced and data that is open sourced.

Please advise (answering separately for each):

- a. How does Amazon define data that is publicly available for the purposes of hoovering into data training sets?*
- b. Does publicly available data include content on social media pages that are not hidden by privacy settings? If not, how does Amazon exclude this data?*
- c. Does publicly available data include pirated content? If not, how does Amazon exclude this data?*
- d. Does publicly available data include data that is copyrighted, trademarked or patented? If not, how does Amazon exclude this data?*

Publicly available material could be defined as text, video, and image data that is generally available to the public on the internet without the need to log-in to access the material.

We don't disclose specific sources of our training data, but as a rightsholder ourselves, we take IP-related concerns seriously and respect the rights of artists and creators. Rightsholders have tools they can use to signal their desire to not have their content collected; robots.txt exclusion protocol is one such widely-used industry standard.

Question 3

At the hearing, Amazon was asked about the recent finding of the CNIL that Amazon had set up an "excessively intrusive system for monitoring employee activity and performance." The ruling found that "three indicators processed by the company are illegal", namely the "stow machine gun" indicator, "idle time" indicator, and "latency under ten minutes" indicator.

Please advise (answering separately for each):

- a. Does Amazon utilise the "stow machine gun" indicator it uses in France, in Australia?*
- b. Does Amazon utilise the "idle time" indicator it uses in France, in Australia?*
- c. Does Amazon utilise the "latency under ten minutes" indicator it uses in France, in Australia?*

Generally speaking, warehouse management systems are industry standard and are necessary for ensuring the safety, quality, and efficiency of operations and to track the storage of inventory and processing of packages on time and in line with customer expectations.

(a) No.

(b) No.

(c) No.

Question 4

In addition to the three indicators described in the previous question, what other data does Amazon collect about the performance of workers in Amazon warehouses?

Like many businesses, we collect information about the performance of workers. This includes information such as rewards and recognition for outstanding performance, as well as identified growth opportunities.

Question 5

At the hearing, Amazon was asked whether Amazon uses any form of artificial intelligence to surveil workers, to which Mr Levey responded, "no". Mr Levey also said: "we don't in any way surveil or monitor union activities in our Australian facilities."

Please advise (answering separately for each):

a. Does Amazon collect data, through CCTV or other video recording technology, of people, including union officials and workers, entering, exiting or being present within Amazon warehouses?

b. Is it correct that anywhere within an Amazon warehouse is subject to video recording?

c. Is facial recognition software automatically or manually applied to all or some of the video footage captured in Amazon warehouses? If yes, please provide details of when this is done and for what purpose.

d. Noting that Amazon warehouses have electronic gates that anyone entering or exiting the premises needs to scan through, using either a staff or visitor pass, does Amazon collect data through this system of people, including union officials and workers, entering or exiting Amazon warehouses?

e. Does Amazon monitor or engage in any other activity associated with obtaining knowledge of the social media activity of its warehouse employees?

f. Does Amazon keep a record, either in Australia or overseas, of visits of union officials to Amazon premises in Australia?

g. Does Amazon keep a record, either in Australia or overseas, of engagements that Amazon warehouse workers in Australia have with unions?

h. Would Amazon support requirements for workers to be informed about the data their employer or contracting entity collects about them and their work performance, including the nature of the data, who has access to the data, and the purpose for which it is collected?

i. Would Amazon support requirements for this data to be made available to the worker at their request?

j. Would Amazon support restrictions on the unreasonable surveillance of workers, particularly where AI is involved?

While it is not apparent how some of these matters, including details of specific interactions with union officials, are relevant to the terms of reference, we have responded to the questions generally.

Like any other responsible business, we maintain a level of security within our operations to help keep our employees, buildings, visitors and inventory safe. This includes the use of CCTV cameras, which are operated in accordance with applicable legal requirements. CCTV footage is not used for the purpose of surveilling union officials, who may be on site holding discussions with employees. Amazon has a visitors register to ensure the safety and well-being of all visitors, and so we can ensure they are informed and cared for in the event there was an emergency situation, such an evacuation. Amazon keeps records of entry notices submitted by union officials (including, for example, notices for entries authorised by section 484 of the Fair Work Act 2009 (Cth)).

Amazon Australia does not use facial recognition software in our fulfilment centres. Amazon does not monitor or engage in activity associated with obtaining knowledge of the social media activity of warehouse employees, unless reviewing social media is reasonably required - for example, to investigate allegations raised by another employee in relation to inappropriate social media activity in connection with their employment.

Amazon agrees that employers should have open communications with their employees about their performance, with reference to relevant performance data. Ordinarily, line management and HR have access to data relevant to performance discussions, and the employee is aware of this. Amazon supports workers being reasonably informed about the basis for any performance concerns. Our focus is on supporting our teams and delivering for our customers, not monitoring workers. We therefore are supportive of reasonable limits regarding surveillance.

Question 6

Mr Levey said at the hearing “We facilitate dozens of lawful union rights of entry in our sites around Australia all the time.” Noting Mr Levey’s familiarity with the number of union entries to Amazon sites, please advise what data in addition to the number of entries, that Amazon collects about union activity at Amazon sites.

Please refer to our response to Question 5.

Question 7

At the hearing Ms Foster said that Amazon provides an indemnification guarantee to users of its AI services, including around alleged copyright infringement. Please provide details of the scope of this guarantee and the coverage it provides, and whether it differs in Australia to other markets.

Building customer trust is core to AWS. We have been on a journey with our customers since our inception, and with the growth of generative AI, we remain committed to building innovative technology together. To enable customers to harness the power of our generative AI, they need to know they are protected.

AWS offers indemnification protections on intellectual property (IP) infringement from AI outputs, equivalent to other generative AI providers. AWS offers an uncapped IP indemnity for copyright claims arising from generative output of Amazon Titan models and other generally available Amazon generative AI services listed in Section 50.10 of the [Service Terms](#) (the “Indemnified Generative AI Services”). This means that customers are protected from third-party claims alleging copyright infringement by the output generated by the Indemnified Generative AI Services in response to inputs or other data provided by the customer. Customers must also use the services responsibly, such as not inputting infringing data or disabling a service’s filtering features. In addition, our standard IP indemnity for use of the services protects customers from third-party claims alleging IP infringement (including copyright claims) by the services and the data used to train them. These indemnity protections are global, and do not differ in Australia to other jurisdictions.

Question 8

At the hearing, Amazon was asked whether Amazon captures audio recorded from Alexa devices in Australian homes, to which Mr Levey said: “Amazon does use some limited number of voice recordings at a global level to improve the service. That’s something which all customers have the opportunity or the ability to opt out of. On your specific question of whether that applies to Australian customers, I don’t have that data.”

Please advise (answering separately for each):

- a. What is meant by the use of data to “improve the service” – which service and in what way is it improved by this data?*
- b. Are Alexa users in Australia subject to the use of their voice recordings to train Amazon’s AI products as outlined by Mr Levey?*
- c. Does this apply to every Alexa device in Australia, or just some? And if just some, how many and how were they selected?*
- d. Where this takes place, is every word the Alexa device listens to used for this purpose, or just some? And if just some, how is this curated?*
- e. How are Alexa users informed about this process?*
- f. How are Alexa users informed about their ability to opt out of this process?*

When customers first register an Alexa-enabled product like an Amazon Echo device or the Alexa app, we inform them that Amazon processes and retains their audio, interactions, and other data in the cloud to provide and improve our services. During setup, we provide customers with links to information about [Alexa, Echo Devices, and Your Privacy](#), the [Alexa and Alexa Device FAQs](#), and the [Alexa Terms of Use](#), all of which provide additional information about how customers can use Alexa, what happens when customers speak to Alexa, and how customers can manage their privacy settings. Among other things, the Alexa and Alexa Device FAQs explain that Alexa and Echo devices are designed to record as little audio as possible, customers can review voice recordings associated with their Amazon account, and they can delete and manage Amazon’s use of those recordings at any time. Our Alexa and Alexa Device FAQs and other linked materials also explain how Amazon uses voice recordings. See, for example, our FAQs on “How are my voice recordings used?” and “How do my voice recordings and text transcripts improve Alexa?”

Question 9

At the hearing, Mr Levey was asked whether the audio data captured by Alexa (including data used for the purpose referred to in the previous question, and data more generally) is used to improve services in other Amazon divisions, to which Mr Levey said “I can’t speak specifically to a voice recording on Alexa.”

Other than the primary purpose of understanding and responding to an user’s query to Alexa, and the secondary purpose referred to in the previous question of “improving the service” provided by Alexa, is any data collected by Alexa (including transcriptions of audio) used for any other purpose by Amazon? For example, for the purpose of training LLMs, or to inform other Amazon services or products? Please provide details of all other uses.

Please see responses to Questions 8 and 10.

Question 10

Does Amazon’s privacy policy and terms of service/use permit Amazon to transfer, distribute, share or sell audio recorded by Alexa, or transcriptions of audio recorded by Alexa, to a third party?

a. If yes – please also advise whether this has taken place in Australia, the third parties involved, the scope of data involved, and the purpose.

At Amazon, our customers are at the center of everything we do. We value our customers’ trust in us. We never sell their personal data, and we design our products and services to protect their privacy and to put them in control of their experience. The Privacy Notice available on [Amazon.com.au](#) describes how Amazon collects and processes personal information through Amazon Services, including our websites, devices, products, services, and online and physical stores, and the circumstances in which we share customers’ personal information with third parties, including when customers enable third-party skills through our Alexa services and when necessary to fulfill a customer request. To learn more about Alexa skills and how Alexa uses voice recordings and transcripts, see the [Alexa and Alexa Device FAQs](#).

Question 11

At the hearing, Mr Levey said “the licensed content we use in Australia comes from a variety of trusted and reputable sources across news, sport, health and other areas. For example, we have licensing

arrangements with the ABC, with news.com.au and with other news providers as well to ensure that when someone asks that we're surfacing the most relevant information from a trusted source that we can." Mr Levey was then asked about whether the licencing arrangements expressly permit Amazon to use the licenced data for AI, to which Mr Levey responded: "I'm not familiar with the details of those contracts."

Please advise (answering separately for each):

- a. How many such agreements Amazon have in Australia?
- b. Are they limited to news organisations or do they include other content creators? Please expand if the latter.
- c. Do the terms of those licencing arrangements expressly permit Amazon to train the AI which powers Alexa with that content?
- d. Do the terms of those licencing arrangements expressly permit Amazon to train all of Amazon's AI models with that content?

The Alexa business sources licensed content from a variety of trusted sources in news, health, sports and other areas to help ensure Alexa can provide factual, helpful information in response to customer queries. We do not comment publicly on specific commercial arrangements.

Questions 12 and 13

At the hearing Mr Levey said Amazon's use of data has "got a focus on transparency and control." Amazon was then asked about the ability for users of Alexa or other Amazon products and services to seek for their data to be withdrawn from a dataset used to train AI, which Ms Foster then referred to as "disgorgement" and said "Disgorgement is kind of an evolving science, and we're working on ways to think about how we can do disgorgement in a more kind of practical way."

Please advise (answering separately for each):

- a. Can an Australian Alexa user whose data has been used to train Alexa or another Amazon AI product or service, request to have their data removed from that data set?
 - i. If yes – please advise the process for how this can occur, how many such requests Amazon has received in Australia, and how many of those requests have been wholly fulfilled.
- b. Can an Australian content creator, whose content has been used to train Amazon's LLM or other AI products or services without their knowledge and consent, request to have this data removed from that data set?
 - i. If yes – please advise the process for how this can occur, how many such requests Amazon has received in Australia, and how many of those requests have been wholly fulfilled.

Please see our response to Question 8. As noted there, customers can manage the use of their Alexa voice recordings to improve our services and develop new features by visiting Settings > Alexa Privacy > Manage Your Alexa Data in the Alexa app and Echo Show devices. In addition, customers can delete their voice recordings and the transcripts of those recordings at any time, including by enabling a setting to delete their voice recordings automatically.

Question 14

At the hearing Mr Levey mentioned that Alexa users have the ability to “opt-out of” their data being used to train its AI services. If an Australian Amazon customer has an Echo device, a Kindle device, is subscribed to Amazon Prime, shops on Amazon Marketplace and uses Prime Video, Amazon Music, Audible, and Prime Gaming (noting these are just a fraction of the products and services Amazon provides), then please advise (answering separately for each):

- a. How many different terms of use and privacy policies would they need to read to be aware of how Amazon collects and uses their data?*
- b. How many words would they need to read across these policies?*
- c. Does Amazon believe its customers genuinely read all of these policies and are making a genuine informed decision to not opt-out of their data being used in this way? And if so, why can it not be offered on an opt-in basis instead?*

Amazon’s [Privacy Notice](#) describes how Amazon collects and processes personal information through Amazon Services, including our websites, devices, products, services, and online and physical stores. The same Privacy Notice applies across Amazon Services and is written in plain English to explain to customers how we collect and process their personal information. Some of our businesses operated by different entities might have specific Privacy Notices, but they follow practices at least as protective as those reflected on Amazon’s Privacy Notice. For specific products and services, we may also provide additional information to help customers understand how we use their personal information in connection with those products and services. For example, in the [Alexa and Alexa Device FAQs](#), we explain to customers when we send voice recordings to the cloud, how they can review and delete those voice recordings, and how they can manage the use of their voice recordings to improve our services and develop new features.

Question 15

At the hearing Amazon was asked whether content stored on its Prime Video, Kindle, Audible and other publishing platforms, including content created by Australians, has been used to train Amazon’s AI, to which Mr Levey responded: “I’m not familiar with the answer to that question”, and Ms Foster said “I would need to verify, but I believe the answer is no.” Have works, including Australian works, published on Kindle been used to train Amazon’s AI products or services? If yes – please advise (answering separately for each):

- a. Has every work on Kindle been used for this purpose, or just some? And if the latter, how were the works split up?*
- b. Which AI products or services have been trained on this data?*
- c. How was consent obtained from creators for this? Please provide the specific terms in the specific applicable policy.*
- d. Can the AI product or service be used to mimic or imitate works published on Kindle? If no – please advise what steps Amazon has taken to prevent this.*

We don’t disclose specific sources of our training data, but as a rightsholder ourselves, we take IP-related concerns seriously and respect the rights of artists and creators. Rightsholders have tools they can use to signal their desire to not have their content collected; robots.txt exclusion protocol is one such widely-used industry standard.

Question 16

Have works, including Australian works, published through Kindle Direct Publishing been used to train Amazon's AI products or services? If yes – please advise (answering separately for each):

- a. Has every work published through Kindle Direct Publishing been used for this purpose, or just some? And if the latter, how were the works split up?*
- b. Which AI products or services have been trained on this data?*
- c. How was consent obtained from creators for this? Please provide the specific terms in the specific applicable policy.*
- d. Can the AI product or service be used to mimic or imitate works published through Kindle Direct Publishing? If no – please advise what steps Amazon has taken to prevent this.*

Please see response to question 15.

Question 17

Have works, including Australian works, published on Audible been used to train Amazon's AI products or services? If yes – please advise (answering separately for each):

- a. Has every work on Audible been used for this purpose, or just some? And if the latter, how were the works split up?*
- b. Which AI products or services have been trained on this data?*
- c. How was consent obtained from creators for this? Please provide the specific terms in the specific applicable policy.*
- d. Can the AI product or service be used to mimic or imitate works published on Audible? If no – please advise what steps Amazon has taken to prevent this.*

Please see response to question 15.

Question 18

Have works, including Australian works, published on Prime Video been used to train Amazon's AI products or services? If yes – please advise (answering separately for each):

- a. Has every work on Prime Video been used for this purpose, or just some? And if the latter, how were the works split up?*
- b. Which AI products or services have been trained on this data?*
- c. How was consent obtained from creators for this? Please provide the specific terms in the specific applicable policy.*
- d. Can the AI product or service be used to mimic or imitate works published on Prime Video? If no – please advise what steps Amazon has taken to prevent this.*

Please see response to question 15.

Question 19

Have works, including Australian works, published on Twitch been used to train Amazon's AI products or services? If yes – please advise (answering separately for each):

- a. Has every work on Twitch been used for this purpose, or just some? And if the latter, how were the works split up?*
- b. Which AI products or services have been trained on this data?*

- c. How was consent obtained from creators for this? Please provide the specific terms in the specific applicable policy.*
- d. Can the AI product or service be used to mimic or imitate works published on Twitch? If no – please advise what steps Amazon has taken to prevent this.*

Please see response to question 15.

Question 20

The Australian Association of Voice Actors (AAVA) has raised concerns about narrators being required by studio intermediaries, on behalf of Audible, to sign terms that give Amazon the power to steal their voices and use generative AI to replicate them without further recognition or payment. At the hearing Mr Levey repeatedly evaded direct questions about these concerns, so please advise (answering separately for each):

- a. Do the terms described at the hearing, and described at the earlier inquiry appearance of the AAVA and in their submission, give Audible/Amazon (or users or partners of Audible or Amazon) the right to train generative AI models on the voices of signatories to those terms?*
- b. Do those terms give Audible/Amazon, or users or commercial partners of Audible or Amazon, the right to then use that generative AI to produce voices or narration based on the data input of the signatories' voices?*
- c. Do those terms extinguish or seek to extinguish any right the signatory has to seek refuse or control how their voice is used to train and generate AI voices?*
- d. Do those terms extinguish or seek to extinguish any right the signatory has to seek recognition or payment for these uses of their voice?*
- e. Do signatories have any recourse to preventing Amazon or Audible from using their voices for these purposes?*

Audible has not digitally replicated any individual voice actor's voice to narrate audiobook content without that voice actor's consent, recognition or payment. Audible's voiceover contracts include a "name and likeness" provision that already limits Audible's use of the actors' name and likeness (including their voice).

Question 21

Does Amazon, Audible, or other related entities, currently publish audiobooks narrated by AI-generated voices? If yes, please advise (answering separately for each):

- a. When did this practice begin?*
- b. How many audiobooks have been narrated with AI-generated voices?*
- c. Are the models trained on voices obtained from signatories to the aforementioned terms and/or other narrators on Audible?*
- d. Are voice actors informed when their voice is used to generate an artificially-voiced audiobook on the Audible platform?*
- e. Have voice actors been credited when their voice is used to generate an artificially-voiced audiobook on the Audible platform?*
- f. Have voice actors been remunerated when their voice is used to generate an artificially-voiced audiobook on the Audible platform?*

Neither Audible nor Amazon publish AI-narrated audiobooks for sale on Audible.com.au or Amazon.com.au.

These additional questions were sent to Amazon by the Committee Chair, on behalf of the Australian Writers' Guild, on 9 September 2024

Question 1

How can creative workers opt out of having your company and platforms 'train' its AI models on their works or materials?

Question 2

What data will you make available to creative representatives and creators, so they can be confident their work has not been input into 'training' data sets?

Question 3

If your company does not permit an opt out of material being used to 'train' AI without consent, credit and compensation, how do you currently avoid issues of copyright infringement in Australia of Australian and New Zealand-Aotearoa works and materials?

Question 4

What are your plans to remunerate creators whose work is being input to the benefit of your company (with or without their consent)?

Question 5

How do you safeguard the moral rights of creators, either in Amazon's own AI use or the access of AI models via your platforms?

Question 6

Is content on 'Prime Video' ever transcribed, whether using AI or not, and that transcription subsequently used to 'train' AI? If so, is creator consent sought and what terms are offered to creators?

Question 7

Is content on Audible or Kindle used to 'train' AI? If so, is creator consent sought and what terms are offered to creators?

Question 8

What if any warranties are users of your platforms offered by you against secondary liability from infringement, if it is conducted by, with, or using your platforms?

Question 9

In your view, are public broadcasters distinguished in any way from commercial broadcasters in terms of how 'publicly available' their content is, and are they treated differently in terms of your 'training' scope for AI?

Question 10

How do you safeguard Indigenous Cultural and Intellectual Property (ICIP) from inclusion in 'training' data, infringement and inappropriate use, and what protocols are in place for managing appropriate use of First Nations works?

We don't disclose specific sources of our training data, but as a rightsholder ourselves, we take IP-related concerns seriously and respect the rights of artists and creators. Rightsholders have tools they can use to signal their desire to not have their content collected; robots.txt exclusion protocol is one such widely-used industry standard.

We are also committed to providing customers with tools and resources to develop and use AI responsibly. For example, AWS policies prohibit customers from using our AI/ML services to, among other things: violate the rights of others; engage in intentional disinformation or deception; depict a person's voice or likeness without their consent or other appropriate rights, including unauthorised impersonation and non-consensual sexual imagery; or harm or abuse of a minor, including child sexual exploitation.