Introducing mandatory guardrails for AI in high-risk settings: proposals paper

Make a submission

Consult hub

Response received at:
October 4, 2024 at 05:02 PM GMT+10

Response ID:
sbm30e75903cb583b66cef4c

1 Do you agree to the Privacy Collection Statement?
Yes I agree

2 Do you consent to the Department's use of Microsoft Copilot and Azure OpenAI to analyse your response to this consultation, including to process personal information provided?
Yes, I consent

3 Please indicate how and if you want your submission published.
Public

4 Published name
Australian Library and Information Association

5 First name
Trish

6 Last name
Hepworth

7 Email

Trish.Hepworth@alia.org.au

8 Phone

+61401838244

9 Who are you answering on behalf of?

Organisation

10 Organisation

Australian Library and Information Association (ALIA)

11 What is the size of your organisation?

Small business (less than 20 employees)

12 What sector best describes you or your organisation?

Peak or professional body (including unions)

13 What state or territory do you live in?

Australian Capital Territory

14 Postcode

2604

15 What area best describes where you live?

City

16 1. Do the proposed principles adequately capture high-risk AI?

No

17 Please provide any additional comments.

Please note that for all answers in this submission, due to the tight timeframe for consultation we were not able to properly consult with Aboriginal and Torres Strait Islander experts in our sector and our responses to this submission, especially regarding ICIP and data sovereignty have not been endorsed. We strongly recommend that the

Department speak with experts in data sovereignty and ICIP from the library and archive sector, which have both deep experience in related areas and hold some of the most significant collections containing ICIP, Indigenous data and traditional knowledge in Australia.

With that said, ALIA strongly recommends that explicit reference to ICIP, traditional knowledge and Indigenous data are included in the principles. This principle should make it clear that the use of ICIP, traditional knowledge and Indigenous data are in and of themselves high risk uses, not dependent on the intended use of these materials. The addition of this principle needs to be in conjunction with a strengthening of guardrails to specifically address the safeguards needed when dealing with these materials.

18 ## Are there any principles we should add or remove?

Yes

19 ## Please provide any additional comments.

A specific principle regarding any use of Indigenous Cultural and Intellectual Property (ICIP), traditional knowledge or Indigenous data.

20 ## Please identify any:

Not answered

21 ## 2. Do you have any suggestions for how the principles could better capture harms to First Nations people, communities and Country?

Yes

22 ## Please provide any additional comments.

When First Nations people are not able to exercise ownership and control over their ICIP, traditional knowledge and Indigenous data then this in and of itself can cause harm. Libraries, as custodians of significant collections containing these materials are very well aware that often legal rights and cultural rights are held by different people, and also that much of the content may not even be mislabelled (often with racist terminology) and therefore not easily identifiable as containing ICIP, traditional knowledge or Indigenous data. The library sector has a number of key protocols and guidelines to try to deal with materials in a respectful and appropriate way, in consultation with communities, and acknowledges there is still much work to be done to ensure Indigenous sovereignty in collections.

The principles should make it clear that the use of these materials in and of itself can cause harm, and then the guidelines should be strengthened to ensure that any use of ICIP, traditional knowledge and Indigenous data are done appropriately as befits the communities, materials and uses. This doesn't mean that these materials can never be used, and there are several good examples of projects utilising AI in a respectful ways, for example in language revival, however there needs to be an onus on developers and deployers to treat ICIP, traditional knowledge and Indigenous data appropriately.

See for example the library sector guidelines the ATSILIRN protocols https://atsilirn.aiatsis.gov.au/protocols.php

Example ICIP protocols https://www.nla.gov.au/using-library/indigenous-cultural-and-intellectual-proper ty

Indigenous Data Sovereignty Communique Communique-Indigenous-Data-Sovereignty-Summit.pdf (squarespace.com)

23 3. Do the proposed principles, supported by examples, give enough clarity and certainty on high-risk AI settings and high-risk AI models? Is a more defined approach, with a list of illustrative uses, needed?

YES the principles give enough clarity and certainty

24 If you prefer a principles-based approach, what should we address in guidance to give the greatest clarity?

While we general support a principles based approach, we are also open to a defined list, and for that list to include specific reference to uses involving ICIP and specific reference to democracy (not just rule of law)

25 If you prefer a list-based approach (similar to the EU and Canada), what use cases should we include? How can this list capture emerging uses of AI?

Not answered

26 How can this list capture emerging uses of AI?

Not answered

27 4. Are there high-risk use cases that government should consider banning in its regulatory response (for example, where there is an unacceptable level of risk)?

Not answered

28 If so, how should we define these?

Not answered

29 Please provide any additional comments.

Not answered

30 5. Are the proposed principles flexible enough to capture new and emerging forms of high-risk AI, such as general-purpose AI (GPAI)?

Not answered

31 Please provide any additional comments.

Not answered

32 6. Should mandatory guardrails apply to all GPAI models?

Yes

33 Please provide any additional comments.

We understand from the definition in the paper that GPAI would included generative AI models. If GPAI models are not included then a significant strengthening of the principles would be needed. In particular this would undermine the important and welcome moves to improve transparency around both inputs and outputs. As such we support the guardrails applying to all GPAI.

34 7. What are suitable indicators for defining GPAI models as high-risk?

Define high-risk against the principles

35 What technical capability should it be based on?

Not answered

36 8. Do the proposed mandatory guardrails appropriately mitigate the risks of AI used in high-risk settings?

No

37 Please provide any additional comments.

ALIA strongly support the moves towards greater transparency for both inputs (guardrail 3) and outputs (guardrail 6). Guardrail 1 should specifically require accountability processes and policies to include ICIP, traditional knowledge and Indigenous data with appropriate policies and processes, and these must be implemented.

38 Are there any guardrails that we should add or remove?

Yes

39 What guardrails should we add or remove?

Other: "Add ICIP in as a specific guardrail or in its own guardrail "

40 Please provide any additional comments

Not answered

41 9. How can the guardrails incorporate First Nations knowledge and cultural protocols to ensure AI systems are culturally appropriate and preserve Indigenous Cultural and Intellectual Property?

Guardrail 1 should specifically require accountability processes and policies to include ICIP, traditional knowledge and Indigenous data with appropriate policies and processes, and these must be implemented.

Guardrail 3 could be strengthened by specifically requiring ICIP, traditional knowledge and Indigenous data to be proactively identified (as opposed to consideration of the principles of Indigenous Data Sovereignty and ICIP) and dealt with appropriately, which may include the possibility that data is not able to be ingested.

These should be in line with key protocols and not contradict the Australian Government's Framework for the Governance of indigenous Data https://www.niaa.gov.au/news-and-media/framework-governance-indigenous-data

We note the in Guardrail 3 the requirement for data to be legally obtained. We note that this is a necessary but not sufficient requirement. In particular we note that permissions for data use may often have been given before the advent of recent advances in AI and should not necessarily be assumed to cover new uses in all cases. We also stress that due

to historical and ongoing circumstances, the legal “ownership” of much ICIP, traditional, knowledges and Indigenous data is not held by the appropriate First Nations people or communities. An undue stress on legal rights without recognition of other rights, such as ICIP, has the potential to override First Nations interests in material, for example if a publisher makes bulk deals with AI companies. Identification and active seeking of consent should be part of this guardrail.
We note that there are a number of projects working to use ICIP in AI projects in an appropriate way, for example the international collaboration between UTS, Kings College London and the university of Glasgow https://www.kcl.ac.uk/research/ireal
As this is a developing area, strategic government funding for research into how to appropriately develop and deploy AI models with ICIP, traditional knowledge and Indigenous data would be highlight beneficial and should inform continued guidance.

42 10. Do the proposed mandatory guardrails distribute responsibility across the AI supply chain and throughout the AI lifecycle appropriately?

Not answered

43 Please provide any additional comments

Not answered

44 Which of the guardrails should be amended?

Not answered

45 Please provide any additional comments

Not answered

46 11. Are the proposed mandatory guardrails sufficient to address the risks of GPAI?

Not answered

47 Please provide any additional comments

Not answered

48 How could we adapt the guardrails for different GPAI models, for

example low-risk and high-risk GPAI models?

Not answered

49 12. Do you have suggestions for reducing the regulatory burden on small-to-medium sized businesses applying guardrails?

Yes

50 Please provide any additional comments

It may be important for small organisations (of which we are one) to be able to rely on the representation of developers and deployers of tools to aid in compliance.
General community AI literacy will also benefit all, as will specific targeted education initiatives and resources.

51 13. Which legislative option do you feel will best address the use of AI in high-risk settings?

A whole of economy approach – introducing a new cross-economy AI Act

52 What opportunities should the Government take into account in considering each approach?

Not answered

53 14. Are there any additional limitations of options outlined in this section which the Australian Government should consider?

Not answered

54 Please provide any additional comments.

Not answered

55 15. Which regulatory option(s) will best ensure that guardrails for high-risk AI can adapt and respond to step-changes in technology?

A whole of economy approach – introducing a new cross-economy AI Act

56 Please provide any additional comments.

Not answered

57 16. Where do you see the greatest risks of gaps and inconsistencies with Australia's existing laws for the development and deployment of AI?

Not answered

58 Which regulatory option best addresses this?

Not answered

59 Please explain why.

Not answered

60 Have you removed any identifying information from your submission?

Not answered

61 Upload 1

Not answered

62 Upload 2

Not answered

63 Make a general comment

As noted at the beginning of the submission, while we consulted with First Nations experts in the preparation fot this submission, due to time constraints we were not able to fully consult and this submission has not been endorsed officially by the ALIA Aboriginal and Torres Strait Islander Expert Advisory Group. We would be happy to support further First Nations consultation with the library, information and archival sectors who have significant experience in dealing with ICIP, traditional knowledge and Indigenous data in collection, including its use with machine learning and related technologies.

We stress that alongside the guardrails it is imperative that whole of community AI literacy, including information an media literacy, is supported.