Reasonable steps in digital platform regulation – what is reasonable and to whom?

Februrary 2025 | **Policy briefing**

Reset·Tech
AUSTRALIA

Contents

Reset.Tech Australia is an independent, non-partisan policy research lab committed to driving public policy advocacy, research and civic engagement to strengthen our democracy within the context of technology. We are the Australian affiliate of Reset, a global initiative working to counter digital threats to democracy.

Introduction

The concept of ‘reasonable steps’ is embedded throughout Australian digital platform regulation.

For example;

- Under the *Online Safety Act 2021* (OSA), digital platforms are already required to:
    - Take reasonable steps under the Basic Online Safety Expectations to ensure end users can safely use the service, to minimise users’ exposure to illegal and harmful material as defined in the Act and, to prevent children from accessing Class 2 materials (largely pornography)
    - Take all reasonable steps to remove content classified as child-bullying, adult cyber-abuse, image-based abuse, or content that falls under the online contents scheme (e.g. Class 1A & 1B materials)

- Under the *Online Safety Amendment (Social Media Minimum Age) Act 2024* social media platforms will have to ‘take reasonable steps to prevent children who have not reached a minimum age from having accounts’ by the end of 2025

- Under the exposure draft of the Scams Prevention Framework — more formally the Treasury Laws Amendment Bill 2024 (Scams Prevention Framework) — digital platforms will have an obligation to:
    - Take reasonable steps to prevent scams,
    - Take reasonable steps to detect scams, and,
    - Take reasonable steps to disrupt scams.

- If the Digital Duty of Care passes, we expect platforms will have to take reasonable steps to identify and mitigate against foreseeable risks

- It also features prominently across the *Privacy Act* and the Australian Privacy Principles, where guidance and legal precedent help to determine what could be expected when it comes to Australians' data protection.

Despite being a widespread concept in the digital regulatory framework, determining what is or is not a reasonable step is not a clear-cut process. As the 2024 *eSafety Commissioner v X Corp* case highlighted, determining reasonableness can be highly contested, with regulators, digital platforms, politicians and the public often not seeing eye-to-eye (see Figure 1). However, as discussions about reviewing and strengthening Australia’s OSA and other regulatory frameworks for digital platforms continue, the concept of ‘reasonable steps’ in the online world is likely to be reinforced across the regulatory landscape.

This policy briefing reflects discussions from a roundtable of 17 experts from academia and civil society in January 2025, where we examined the concept of reasonable steps, debating its usefulness and whether and how it could be improved. The event was held under the Chatham House Rule, meaning this briefing presents an overview of the discussion, without attributing specific comments.

The complexities of determining what is a reasonable step: lessons from *eSafety Commissioner v X Corp*

One of the issues in the 2024 *eSafety Commissioner v X Corp*[1] case was an argument about what constituted 'reasonable steps'. The eSafety Commissioner had issued a s 109 notice to the X platform to remove 65 posts (or URLs) containing video footage of a knife attack on Bishop Mar Mari Emmanuel in Sydney, which had been deemed a terrorist attack. Under the *Online Safety Act*, the eSafety Commissioner has the power to issue s 109 notices (removal notices) to require platforms to remove Class 1 material (illegal and restricted content).

While X complied with the order by removing access to the posts for users within Australia, the posts remained accessible to international users (or users operating with a non-Australian IP address). At the core of the dispute was what activity constituted the taking of 'all reasonable steps to ensure the removal of the material from the service'. X Corp argued that geo-blocking, such that end-users physically located in Australia could not view the material, was sufficient. The eSafety Commissioner argued it was not, and sought further remedies.

In the end, Justice Kennett ruled in favour of X Corp, noting that while the s 109 removal notice was valid under the Act and the 65 posts needed to be blocked for Australian users, requiring global removal was not within the expectation of all reasonable steps. Justice Kennett noted:

> 'Identification of the steps that are 'reasonable' in this sense may involve consideration of expense, technical difficulty, the time permitted for compliance (which may be short: see s 109(2)) and the other interests that are affected.'

Discussing what might be considered 'reasonable steps' in the context of the 'comity of nations', Justice Kennett noted that the apparent extraterritoriality provisions in the *Online Safety Act* did not 'control the meaning of 'all reasonable steps'. In short, the expectations of what can be considered 'reasonable steps' are complex, and various provisions within the *Online Safety Act* may give rise to different reasonable expectations.[2]

Figure 1: Debates about determining 'reasonable steps' arising from the 2024 *eSafety Commissioner v X Corp Federal Court case*

[1] *eSafety Commissioner v X Corp* [2024] FCA 499

[2] See Marcus Smith, Mark Nolan, & John Gaffey (2024). 'Online safety and social media regulation in Australia: eSafety Commissioner v X Corp.' *Griffith Law Review,* 33(1), 2–18. https://doi.org/10.1080/10383441.2024.2405760

1. Reasonable steps in Australian digital regulation

'Reasonable steps' is a very common test. We see it scattered across our laws, in multiple places in the *Corporations Act*, for instance, and closer to home, in the current *Online Safety Act*, and likely an enhanced role in the next iteration of it as well, and the *Privacy Act*.

When things go wrong or a regulator decides to take action against a company under a 'reasonable steps' expectation, the litigation will involve the judge turning their mind to what is reasonable under the circumstances for the company to have done to comply with their obligations. It is worth noting that 'reasonable steps' appears to operate more as a *defence* rather than an *obligation*. Further, it is not an absolute or unconditional obligation to achieve a particular objective.

It is a neutral expectation designed to achieve a balance, giving rise to obligations that are specific to their commercial or technical contexts and are standard expressions of mitigation that (we hope) society would expect a company in the circumstances to take. It is assessed by a standard of reasonableness, rather than an analysis of efficacy or demonstrable avoidance of harm. This point may be quite important in the context of matters that are critical to user safety, where it is arguably insufficient and at times irrelevant to prove that the company acted reasonably in the circumstances. Would that defence hold up in circumstances of disaster or crisis? It is hard to say.

A. Reasonable steps in the *Online Safety Act*

There are three key considerations in thinking about what might be considered 'reasonable' within the *Online Safety Act*.

i. Elements in the test

Before we consider what reasonableness is in the OSA, it is worth examining how Federal Court judges have looked at this elsewhere. Figure 2 presents a framework from a case about financial services and the obligations of licensees. While it is functionally quite different, it is likely that we'd see a similar analytical model deployed in judicial reasoning applied to a future OSA.

Totality of factors	**No expectation to take all possible steps:** It is not an obligation to take all reasonable steps and it does not require identification or performance of either the totality of all possible steps or alternatively the following of one 'correct' approach. **Must examine all the steps taken:** The steps taken by a licensee must be reasonable in their totality. It is not a defence to merely point to one or two reasonable steps taken. The steps, in context and when examined together, must be reasonable. **Reasonable steps do not have to be not optimal steps.**

Contextually dependent	**Business and activities:** It is a question of fact that depends on the circumstances of the business and activities carried out by the licensee and its authorised representative(s). **Experience:** The nature of the representative informs what steps are reasonable. Experienced ARs may reasonably attract less oversight and supervision.
Erroneous understanding is not a defence	The question is an objective one and is made on the understanding that the licensee knows the law. Erroneous understandings of the law do not provide a defence as the taking of reasonable steps inherently implies that the licensee informs itself as to the nature of the obligation.
Knowledge of all the circumstances	Despite the objectivity of the test, a licensee's knowledge of all the circumstances informs what it reasonably could have done, including with regard to the difficulty, practicality and costs of any steps the licensee could have taken.
Importance and purpose of statutory provision	The interpretation of 'reasonable steps' must have regard to the importance and purpose of the provision. Given the civil penalties that may be imposed for contravention of section 963F and the fact it only applies to retail clients, it should be read as safeguarding a vulnerability that should be taken into account when determining what actions should be taken to ensure compliance.
Regulatory context as a whole	The broader context of the licensing regime should be considered. As licensees have a broad remit to appoint the ARs they choose, this is accompanied by a strict obligation to supervise the actions of their ARs.

Figure 2: A framework for understanding 'reasonable steps' derived from financial service licences and obligations[3]

ii. Interaction between s 109 and s 111

Removal notices issued under the OSA include language where the provider must take all reasonable steps to ensure the removal of offending material from their service. However it has been noted in the courts that there is a tension between the power for issuing removal notices in s 109 and the qualifying language in s 111 (see Figure 3).

s 109(e)	s 111
'Take all reasonable steps to ensure the removal of the material from the service'	'To the extent that the person is capable of doing so'

Figure 3: Differences in the Online Safety Act 2021 between s 109 obligations and s 111 clarifying powers

Arguably, s 111 dilutes the expectation for reasonable steps by injecting an idea of capability – presumably this was introduced to handle the burden on smaller services, but it could easily be

[3]From *Australian Securities and Investments Commission v R M Capital Pty Ltd* (2024) FCA 151

weaponised by recalcitrant operators, so we should be watching that space in reform conversations. It was foreshadowed in the eSafety and X Corp litigation as something that may need clarification in the future.

iii. Reasonableness will involve an assessment of both the company and the decision-maker

Justice Kennett turned their mind to reasonableness in the X Corp matter, and it was an assessment that looked to the efforts of the company and the substance of the regulatory decision-making. These elements of administrative law principles are a quirk of our legal system that potentially puts us in a different position from our counterparts in the European Union (EU), who are not subject to the same standards of review.

Beyond this, the X litigation highlights the importance of a granular and practical understanding of what platforms can and cannot do among civil society and advocacy groups. As we enter the year of setting enforceable expectations for how platforms design and operate their services – admittedly in only the narrow terrain of age access for now rather than substantive safety – there is a role for advocates to be able to effectively identify what 'reasonable steps' platforms may actually be able to take.

B. Reasonable steps in the *Privacy Act*

Reasonableness runs through the *Privacy Ac*t and the Australian Privacy Principles (APP).[4] Like the other examples, they are not defined in the instruments– they take what is referred to as their 'ordinary meaning'. What is reasonable is determined based on what a reasonable person would expect in the circumstances. This test involves some examination of current standards and practices.[5] The APP Guidelines cite two authorities, as below, noting that there may well be a 'conflicting range of objective circumstances to be considered, and the factors in support of a conclusion should outweigh those against'.[6] The authorities are:

- George v Rockett,[7] where the High Court observed that whether there are 'reasonable grounds' to support a course of action 'requires the existence of facts which are sufficient to [persuade] a reasonable person', and
- McKinnon v Secretary, Department of Treasury,[8] where the test 'involves an evaluation of the known facts, circumstances and considerations that may bear rationally upon the issue in question'.

Helpfully, we do have recent guidance from the Privacy Commissioner on what reasonable steps are involved for the purposes of privacy protection in Australia. In a recent decision,[9] a government agency was held to have failed to take reasonable steps to protect personal information from unauthorised disclosure, in so doing breaching APP 11.1. The Commissioner determined:

---

[4]See APP Guidelines, B.108.
[5]*Jones v Bartlett* [2000] HCA 56 [57] – [58] (Gleeson CJ); *Bankstown Foundry Pty Ltd v Braistina* [1986] HCA 20 [12] (Mason, Wilson and Dawson JJ).
[6]APP Guidelines, B.108.
[7](1990) 170 CLR 104 at 112 (Mason CJ, Brennan, Deane, Dawson, Toohey, Gaudron & McHugh JJ)
[8](2006) 228 CLR 423 at 430 (Gleeson CJ & Kirby J).
[9]*'ATQ' and CEO of Services Australia* (Privacy) [2025] AICmr 19 (23 January 2025)

> I am of the view that the steps taken by the respondent during the relevant period were not reasonable in the circumstances because they failed to protect the complainant's personal information from unauthorised disclosure on multiple separate occasions over an extended period. The repetition of incidents of a similar nature despite the respondent taking such steps suggests those measures were inadequate and ineffective in appropriately protecting the complainant's personal information during the relevant period.[10]

The Commissioner provided examples of what steps should have been taken, indicating the technical knowledge required by decision-makers for these kinds of determinations, and the expectations likely to fall to APP entities for similar issues of protecting personal information in a Privacy Act environment:

- Implementing robust procedures for the intertwinement of customer records
- Conducting routine reviews or audits of existing processes, in order to proactively monitor for privacy risks of the nature identified, and
- Providing additional, regular training and guidance to staff.

[10] *'ATQ' and CEO of Services Australia* (Privacy) [2025] AICmr 19 (23 January 2025), [43].

2. Reasonable steps or outcomes?

There are two key ways of thinking about regulatory models in the online world to date: outcomes-based regulation, and prescriptive regulation. The trend in most jurisdictions is towards outcomes-based legislation.

Outcomes-based legislation applies to foreseeable risk, rather than the precise bad actor or precise 'risk'. An outcome-based approach creates an obligation to do what is reasonable to mitigate risk, which is referred to as a fault-based negligence standard. The main thrust of the approach is requirements to 'do what you can', rather than imposing an absolute obligation to get it *right.* In evaluating an outcome based approach in court, where a harm has occurred and led to a case, the court would examine the circumstances behind the particular harm and evaluate whether in that instance the response to the risk was proportionate. This allows the courts to have regard for the magnitude of the risk, its foreseeability and the resources available to the defendant to mitigate against it. This includes evaluating the possible mitigation techniques available to defendants, and their cost and viability. In this sense, an outcomes-based approach is not about absolute standards: instead, it focuses on proportionate responses. Focusing on proportionality generally allows for legitimate business interests.

Effective outcomes-based regulation necessitates clear boundaries laid out in regulatory guidance. In general, outcomes-based systems tend to match broad, overarching obligations, with an expectation to do what is reasonable in the circumstances (the legislation may also define priority areas, such as illegal content or harms to minors, for example). But the 'breadth' of the obligations, and what should be considered reasonable, is elucidated through soft law regulatory guidance. Such guidance ensures a degree of standards-setting to nudge effective compliance, but without losing the flexibility of a reasonable steps approach.

Additionally, for an outcomes-based approach to work, reporting and oversight are critical. To achieve this, what we have seen in the past is legislative instruments that require companies to openly report on risk (largely via risk assessments), matched in turn by regulatory or other scrutiny of these risk assessments. That is, external scrutiny of risk assessments, risk mitigations, the viability of the responses and assessment of the robustness of the measures need to occur to enable meaningful evaluation of 'reasonableness'.

The alternative to an outcomes-based approach is a prescriptive approach. This is implemented by legislating for an absolute duty to ensure safety (as opposed to a fault-based duty), with some defence. This approach is a strong measure, creating obligations that mean that 'getting it wrong' becomes illegal. This places a reverse onus on companies. As yet, no country has imposed a strict safety duty when it comes to digital platforms; however, there is an adjacent precedent. In the UK, banks and their responsibilities with respect to scams involve absolute duties, as do product safety regimes in the EU. For example, the *EU General Product Safety Regulation* places a strict duty on 'economic operators' to 'place or make available on the market only safe products.[11] While not without precedent, the application of absolute duties has so far been narrow in the world of digital regulation.

[11]Article 5, Regulation (EU) 2023/988 of the European Parliament and of the Council of 10 May 2023 on general product safety

3. Reasonable steps and a duty of care

The UK's *Online Safety Act* (UK OSA) was operationalised last November or December when Ofcom, the regulator published the first set of codes of practice. This fired the starter's gun on enforcement regimes and colours what we know about the nascent risk assessment and mitigation obligations placed on digital platforms. The Act itself rests largely on two key sets of duties of care rather than an overarching duty of care; duties around illegal content online and duties around the prevention of harm to minors.

Underneath these two duties, platforms have two secondary duties that function as the way platforms demonstrate compliance:

- First, regulated services have an obligation to perform a risk assessment. Platforms regulated under the Act have to look at how content and design of services are built on and relate to risk associated with both primary duties, and are required to identify mitigation strategies. This is the 'risk assessment' approach of the UK OSA,which might be relevant to Australia, and requires platforms regulated under the Act to undertake a suitable and sufficient assessment when it comes to their illegal content duty and risk to minors duty. The UK OSA does focus on content, but it also takes a systemic approach to risks regarding this content. For example, in the Act, a key harm is how quickly illegal content spreads, which involves understanding platforms' algorithms and functionalities, but also requirements on takedowns and so on.

- Second, platforms regulated under the Act are required to comply with the Codes of practice as developed by Ofcom. Ofcom have done extensive work developing highly detailed Codes of practice that set out the measures that platforms must take. This includes, for example, looking at the risks of group chats, messaging, liking functions and so on. Platforms need to look at the prescriptions within these Codes and ensure they apply to their service. In effect these Codes prescriptively outline the reasonable steps platforms must take. Platforms regulated under the Act need to make sure that they are compliant with the Codes of practice, and *if* they are compliant, then they have 'safe harbour'. In effect, it is assumed that platforms are in compliance with the UK OSA if they meet the prescriptive steps outlined in the Codes.

One of the big concerns arising in the UK is the gap between these two duties. Platforms regulated under the Act must undertake a risk assessment activity but are given 'safe harbour' if they comply with Ofcom's Codes. However, there is room for a disjuncture here; what is identified by a platform in their risk assessment may not appear in a Code. This disjuncture means that a platform could have identified multiple risks in their risk assessment activity without being required to adequately mitigate against them because they are offered safe harbour for complying with the Codes. Put another way, Codes can be incomplete when it comes to the specific risks of a specific platform. This has to an extent undermined the fundamentals of a duty of care approach.

An overarching duty of care, linked to requirements to comprehensively identify risks and implement effective risk mitigations according to the risk assessment, could avoid this dilemma.

Discussion

The discussion focussed on three themes.

i. How to make reasonable steps work in the court system

Reasonable steps may offer leeway to digital platforms to evade responsibility, but it also has utility because it offers flexibility in the interpretation of obligations.

This helps allow courts to consider a range of factors when determining reasonableness. If a decision comes before a court, it allows the judiciary to make relevant inquiries into what tools platforms have available to them, what their knowledge is, what external research has identified as risks and best practices and what justifications may exist that reduce or increase burdens on platforms. It also allows other considerations to be brought into the fold. For example, the discussion around digital encryption and automated scanning for illegal content is nuanced, balancing risks around the spread of illegal content (which may be reduced by scanning) with the expectations of privacy and cybersecurity (which may be reduced by breaking encryption to enable scanning). Where there are requirements for platforms to take reasonable steps to prevent the spread of illegal content, courts may be able to consider justifiable departures based on privacy and cybersecurity in deciding the reasonableness of encryption.

This can also help ensure in ensuring digital platforms are adequately covered by requirements. A reasonable steps approach is not as prescriptive as legislation which can also help to avoid regulatory arbitrage, where a requirement is so strictly defined that platforms can define themselves as outside of the barriers, in effect claiming to be 'outside of' the rules. Further, a flexible approach — such as those provided by 'reasonable steps' — prevents platforms arguing that they did not know about a rule or were technical outside of the rules coverage, and instead focuses on what is reasonable, whether it was foreseeable, where people affected might be more vulnerable (e.g. children or people affected by disability).

However, capitalising on this flexibility requires an understanding of, and documented evidence around what platforms might know about potential risks and what measures platforms could deploy to mitigate them. This creates a role for civil society in both advocacy around reasonable steps and evidence generation. Civil society groups need to be informed and articulate enough to express what risks are and what mitigation measures could be taken, and to evaluate the effectiveness of these. The X Corp argument, as outlined in Figure 1 above, is a good case study of this. The debate about the reasonableness of geoblocking versus content takedown was not well understood, which potentially hampered the quality of the policy debate in Australia.

ii. The role of research and civil society in supporting courts and regulators

Despite the need for civil society research and expertise, there were questions about the capacity of Australian independent researchers to meet this need. High-quality research into this is complicated because digital platforms largely operate as 'black boxes', and there are limited protections and technical experimental skills in Australia to undertake this research.

A core issue discussed is that this is still seen as a 'lawyers' game, which may limit the role of civil society in effectively advocating for solutions. The extent to which the focus of the legislation addresses the activity and knowledge of digital platforms, this creates an onus on regulators to be up to speed with what platforms should know, do know, and we know they know. Civil society could again play a role in generating knowledge and supporting regulators, but they

too would be required to be across debates about what risks are and what mitigations are possible and deployed elsewhere.

Without transparency around what risks are reasonably foreseeable, what mitigation measures are possible and evaluations of their effectiveness, it is extremely difficult for courts to assess whether steps are reasonable. In effect, we need to have a more informed debate about the measurement and ascertainment of harm, what is reasonable to that harm, what platforms could know and do about that harm, and how this plays out for users to more effectively implement a reasonable steps approach. Civil society could be a great resource in this.

How courts could use this evidence is also worth consideration. Evidence would need to come in clearly documented forms that judges would find compelling. For example, this may include levels of user-complaints (noting that reporting rates may be significantly under the control of digital platforms themselves), how user interfaces are created and interpreted, influences on user-behaviour and, in short, evidence about the risks that exist as well as, and distinct from, the harms that these cause. Getting the research right is an essential part of the process of getting judicial determination. The Attorneys-General led litigation in the US perhaps provides a model for this; they were extremely open to expert testimony and experimental research to test and evaluate risks and safety features.

There are distinct challenges in this. For example, it can be difficult to understand who the cohorts of users are who may be affected; it can be even harder to understand which interventions or mitigation measures are targeted at which users, as well as basic research principles such as differing terminologies to describe harms and mitigations.

Core to the issues that Australian courts may experience is a lack of skills and experience in generating this sort of research and evidence in Australia. The field needs a massive uplift to support judicial determination on these issues. This is made more complicated by the lack of legal protections for researchers, given the litigious nature of some of the larger platforms and the lack of legal protections for researchers in Australia. Researchers with the technical skills may not wish to operate within Australia.

iii. Balancing considerations for reasonableness

Understanding what is reasonable generally requires balancing the nature and magnitude of the harm the reasonable steps aim to mitigate; big risks require big steps, whereas smaller risks might not. The digital world presents some unusual challenges in this regard because the scale and size of the risks can be both unknown and or framed differently. For example, under the current *Online Safety Amendment (Social Media Minimum Age) Act 2024*, the risks appear as a ‘black and white’ outcome, either young people can make an account on a social media platform or they cannot. Within this, there are no clear ‘greater or lesser’ harms. It requires a novel rethink.

In this context, it is helpful if harm could be defined within legislation, so that key areas can be prioritised. Legislators would, in effect, be identifying what is important, to help guide what is considered a proportionate response or a reasonable step. The OSA should prioritise key areas, so that regulators and courts can evaluate what might be reasonable steps outside of the legislation.

We see this in the EU, where the *Digital Services Act* outlines the key risks that platforms must mitigate against, and to an extent in the UK OSA, where they have identified key priority areas tied to duties of care. However, if we look at the consumer duty in the UK, we see a more productive, tiered approach, with an overarching duty that provides the flexibility not found in rules and adds in soft law guidance, which can more rapidly be adjusted.

Conclusions

The concept of reasonable steps looks set to be increasingly important as Australia's *Online Safety Act* is reviewed, and other digital regulations are implemented. The approach has benefits because it is flexible, but to be meaningfully implemented, it requires:

- Broad overarching obligations laid out in regulation. Legislation that sets out priority areas or identifies key harms would be beneficial, and it must include requirements to assess risk in a publicly evaluable manner and to mitigate against risks identified.

- Clear regulatory guidance about what should be considered reasonable steps. This guidance is necessary to add details about how broad obligations ought to be achieved, but learning from the UK's experience, compliance with regulatory guidance alone should not offer safe harbour.

- Scrutiny and oversight about what reasonable steps are available to platforms. This requires a strong public debate, informed by a resourced and equipped academia and civil society and up-to-date technical research about the functionalities of digital platforms.