ASIC
Australian Securities &
Investments Commission

Beware the gap: Governance arrangements in the face of AI innovation

REPORT 798 | OCTOBER 2024

CONTENTS

About this report

ASIC reviewed how 23 AFS licensees and credit licensees are using and planning to use artificial intelligence, how they are identifying and mitigating associated consumer risks, and their governance arrangements. This report outlines the key findings from that review.

About ASIC regulatory documents

In administering legislation ASIC issues the following types of regulatory documents: consultation papers, regulatory guides, information sheets and reports.

Disclaimer

This report does not constitute legal advice. We encourage you to seek your own professional advice to find out how the *Corporations Act 2001, National Consumer Credit Protection Act 2009* and other applicable laws apply to you, as it is your responsibility to determine your obligations. Examples in this report are purely for illustration; they are not exhaustive and are not intended to impose or imply particular rules or requirements. For privacy reasons, the names of case-study subjects have been changed.

Foreword

Artificial intelligence (AI) is transforming many aspects of our lives, including how we engage with financial products and services. The potential benefits to business and individuals are enormous – digital innovations including AI are estimated to contribute $315 billion to Australia's GDP by 2030[1].

To fully realise those benefits, we must balance innovation and protection. The integrity of our financial system – and the safety of the consumers who interact with it – relies on us finding the right balance.

For some time, ASIC has been reminding licensees that existing obligations apply to their use of AI. ASIC has also been building an understanding of how AI is actually being used in the sectors we regulate.

This report is ASIC's first examination of the ways Australian financial services (AFS) and credit licensees are implementing AI where it impacts consumers. Concerningly, it finds that there is the potential for a governance gap.

Simply put, some licensees are adopting AI more rapidly than their risk and governance arrangements are being updated to reflect the risks and challenges of AI. There is a real risk that such gaps widen as AI use accelerates and this magnifies the potential for consumer harm.

While the approach to using AI where it impacts consumers has mostly been cautious for licensees, it is worrying that competitive pressures and business needs may incentivise industry to adopt more complex and consumer-facing AI faster than they update their frameworks to identify, mitigate and monitor the new risks and challenges this brings.

As the race to maximise the benefits of AI intensifies, it is critical that safeguards match the sophistication of the technology and how it is deployed. All entities who use AI have a responsibility to do so safely and ethically.

Our review comes at a pivotal time in the development of AI regulation in Australia. We support the Australian Government's Voluntary AI Safety Standard and intention to introduce mandatory guardrails ensuring testing, transparency and accountability for AI in high-risk settings.

However, licensees and those who govern them should not take a wait-and-see approach to legislative and regulatory reform. Current licensee obligations, consumer protection laws and director duties are technology neutral and licensees need to ensure that their use of AI does not breach any of these provisions.

ASIC's work to engage with and monitor licensees' AI use will continue, particularly as we consider how they embed the requirements of any future AI-specific regulatory obligations.

I call on industry to consider the findings of this review and reflect on the questions posed to ensure that innovation is balanced with the responsible, safe and ethical use of this technology.

Joseph Longo
ASIC Chair

[1] Department of Industry, Science and Resources, List of Critical Technologies in the National Interest: AI Technologies

Executive summary

Artificial intelligence has the potential to transform the provision of financial services and credit in Australia. It provides opportunities for more efficient, accessible and tailored products and services. However, AI can also amplify existing risks to consumers and introduce new ones. Potential harms include bias and discrimination, provision of false information, exploitation of consumer vulnerabilities and behavioural biases, and the erosion of consumer trust. To help shape our understanding of risk to consumers and to inform our regulatory response, we reviewed the use of AI by 23 AFS and credit licensees.

Our review

We analysed 624 AI use cases that 23 licensees in the banking, credit, insurance and financial advice sectors were using, or developing, as at December 2023. These were use cases that directly or indirectly impacted consumers and included generative AI and advanced data analytics (ADA) models.

As part of our review, we also asked licensees about their risk management and governance arrangements for AI, and how they planned to use AI in the future. We met with 12 of the licensees in June 2024 to discuss their use cases and governance arrangements.

What we found

We observed a rapid acceleration in the volume of AI use cases. We also observed a shift towards more complex and opaque types of AI such as generative AI.

But on the whole, the way licensees used AI was quite cautious in terms of decision making and interactions with consumers: AI generally augmented rather than replaced human decision making and there was only limited direct interaction between AI and consumers.

The majority of licensees told us they are planning to increase their use of AI. Given the fast-moving nature of AI and competitive pressures in industry, there is potential for the way AI is used and the associated risk to consumers to shift quickly.

We are concerned that not all licensees are well positioned to manage the challenges of their expanding AI use. Some licensees were updating their governance arrangements at the same time as increasing their use of AI. And in the case of two licensees, AI governance arrangements lagged AI use.

Governance and risk management arrangements are, by their nature, slow to change. It is therefore likely that any gap between the use of AI and governance arrangements will widen as AI adoption increases. This could leave licensees unprepared if they want to respond quickly but safely to innovations from competitors.

KEY STATISTICS

- 57% of all use cases were less than two years old or in development.
- 61% of licensees in the review planned to increase AI use in the next 12 months.
- 92% of generative AI use cases reported were less than a year old, or still to be deployed. Generative AI made up 22% of all use cases in development.
- Only 12 licensees had policies in place for AI that referenced fairness or related concepts such as inclusivity and accessibility.
- Only 10 licensees had policies that referenced disclosure of AI use to affected consumers.

Executive summary

OUR FINDINGS

Use of AI

FINDING 1: The extent to which licensees used AI varied significantly. Some licensees had been using forms of AI for several years and others were early in their journey. Overall, adoption of AI is accelerating rapidly (see page 11).

FINDING 2: While most current use cases used long-established, well-understood techniques, there is a shift towards more complex and opaque techniques. The adoption of generative AI, in particular, is increasing exponentially. This can present new challenges for risk management (see page 13).

FINDING 3: Existing AI deployment strategies were mostly cautious, including for generative AI. AI augmented human decisions or increased efficiency; generally, AI did not make autonomous decisions. Most use cases did not directly interact with consumers (see page 15).

Risk management and governance

FINDING 4: Not all licensees had adequate arrangements in place for managing AI risks (see page 19).

FINDING 5: Some licensees assessed risks through the lens of the business rather than the consumer. We found some gaps in how licensees assessed risks, particularly risks to consumers that are specific to the use of AI, such as algorithmic bias (see page 20).

FINDING 6: AI governance arrangements varied widely. We saw weaknesses that create the potential for gaps as AI use accelerates (see page 24).

FINDING 7: The maturity of governance and risk management did not always align with the nature and scale of licensees' AI use – in some cases, governance and risk management lagged the adoption of AI, creating the greatest risk of consumer harm (see page 29).

FINDING 8: Many licensees relied heavily on third parties for their AI models, but not all had appropriate governance arrangements in place to manage the associated risks (see page 31).

We observed a rapid acceleration in the volume of AI use cases, and a shift towards more complex and opaque types of AI such as generative AI. But on the whole, the way licensees used AI was quite cautious. We found some gaps in how licensees assessed risks to consumers from AI, and for some licensees, governance arrangements lagged their AI use. This creates risk of consumer harm.

Executive summary

Where to from here for licensees?

ASIC supports innovation in the financial system that is balanced with appropriate consumer protections and market integrity safeguards.

While licensees' deployment strategies were somewhat cautious, there is fertile ground for consumer harm where use of AI leaps ahead of governance arrangements and controls.

We expect licensees to carefully consider their readiness to deploy AI safely and responsibly. Decisions that licensees make now about how they will govern their AI use will determine whether they establish solid foundations on which to deliver the expected benefits and manage risks to themselves and their customers.

Many licensees told us that they were updating their governance arrangements in relation to AI. This is welcome, but there is more to do. AI presents novel challenges, and licensees' governance arrangements should lead their AI use as it increases and evolves.

Licensees should consider the findings of this report, and the questions on pages 35–36, to help them consider their readiness to deploy AI safely, responsibly and in compliance with existing obligations.

Licensees' obligations and resources for licensees

The regulatory framework for financial services and credit is technology neutral. Licensees need to consider their existing regulatory obligations before deploying AI. In particular, licensees need to consider the general licensee obligations, directors' duties, and consumer protection provisions, including prohibitions against unconscionable conduct and false or misleading representations (see page 34).

There are a number of resources that licensees can draw on as they deploy AI, such as the recently issued Voluntary AI Safety Standard. This standard gives practical guidance to all Australian organisations on how to safely use and innovate with AI.

Licensees who invest the time now will also be in a better position to comply with any future AI-specific regulatory obligations.

The future regulatory landscape

The landscape of AI regulation in Australia is evolving. The Australian Government recently consulted on how it proposes to define 'high-risk AI', and the introduction of mandatory guardrails to promote the safe design, development and deployment of high-risk AI use. The proposed guardrails include requirements related to testing, transparency and accountability of AI.

ASIC supports the introduction of regulatory measures to mandate guardrails for the use of AI in high-risk settings. The findings of this review have informed our contribution to the Government's proposals.

ASIC's focus

We remain focused on advancing digital and data resilience and safety, targeting technology-enabled misconduct and the poor use of AI. Understanding and responding to the use of AI across the entities we regulate is a key priority for ASIC.

We will:

- continue to monitor how our regulated population uses AI, and the adequacy of their risk management and governance processes
- contribute to the Australian Government's development of AI-specific regulation
- engage and collaborate with domestic and international regulator counterparts, and
- where necessary and appropriate, take enforcement action if licensees' use of AI results in breaches of their obligations.

AI presents novel challenges, and licensees' governance arrangements should lead their AI use as it increases and evolves. Licensees should review their arrangements in line with our findings.

Executive summary

CASE STUDY

Beware the gap between AI use and governance

One licensee cited 10 AI use cases in scope, but adoption appeared to front-run their governance and risk management arrangements. The licensee had no overarching AI strategy setting out how and why the licensee had decided to use AI in its operations. The licensee produced no policies setting out standards to guide the design, deployment and oversight of AI, and had not articulated the key risks associated with AI and ADA in their risk management framework (e.g. a lack of explainability for complex models). None of the licensee's use cases were risk rated.

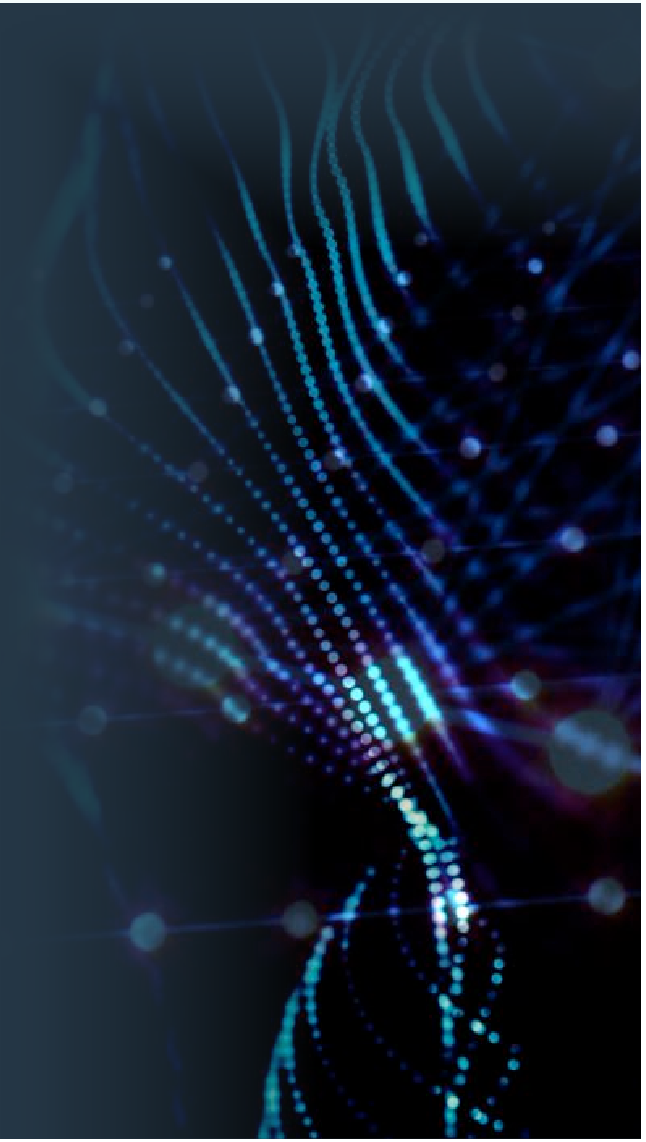
The licensee used an AI model to predict consumer credit default risk by producing a risk score. The score produced by the model was one input into credit decisions. It had the potential to result in consumers being refused credit or offered less credit than they otherwise would have been.

An internal report to a senior committee dated approximately 10 months after deployment of the model stated that it was developed with 'limited understanding' of the third-party platform used, there was 'incomplete model documentation with missing critical elements', and 'poor governance and a lack of a monitoring process'.

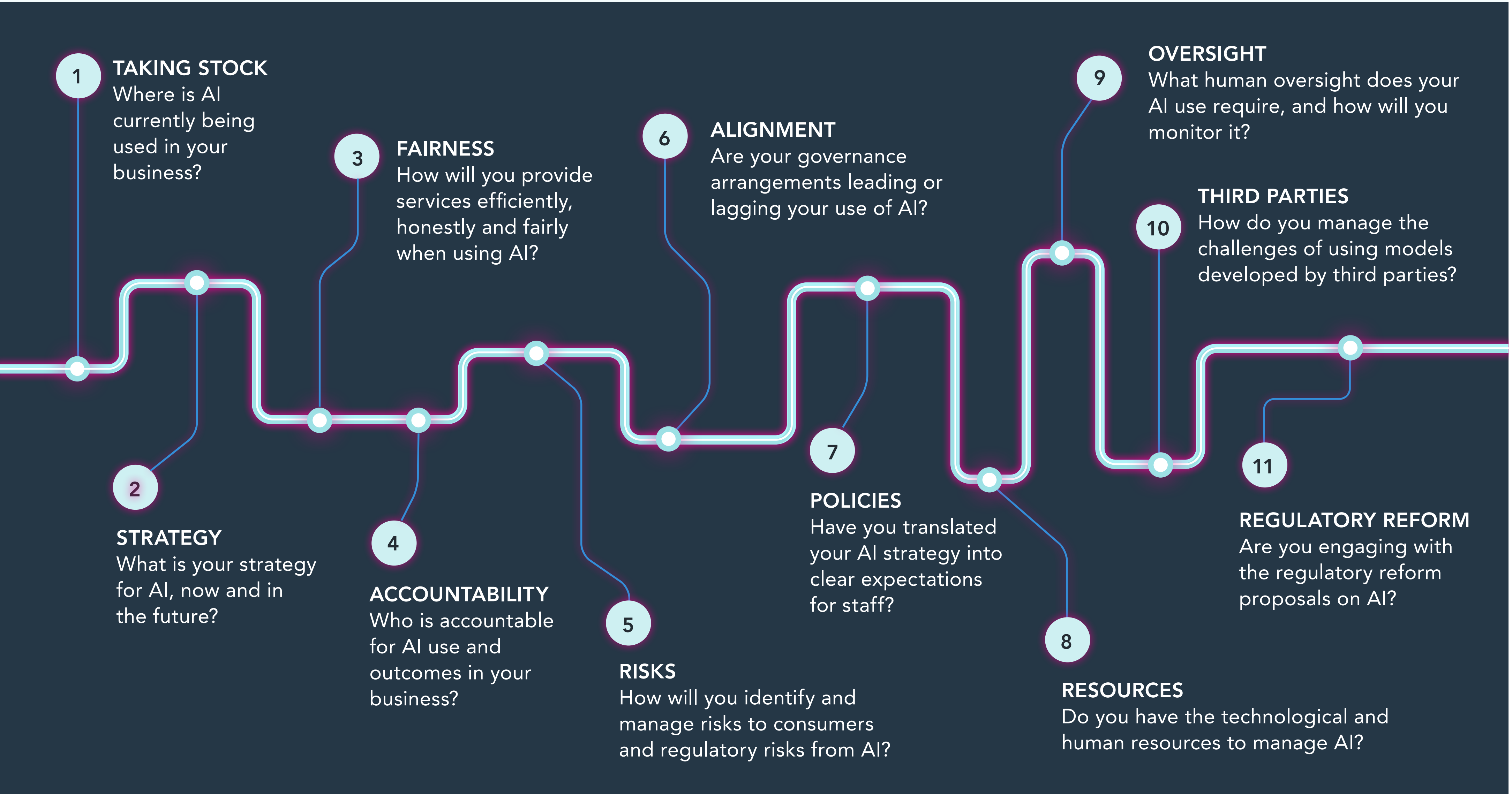
The report further described the model as a 'black box with no ability to explain the variables in the scorecard or the impact they are having on an applicant's score'.

Although the licensee's report stated that 'the model has been stable', it noted that its ability to monitor the model was limited. The report proposed 'to revise the [model], to ensure it is explainable, documented, and has a robust governance process in place'. The licensee continued to use the model for several months before replacing it with a simpler model, to ensure scoring outcomes and the model were explainable.

Despite the issues identified with the above AI model, the licensee reported having plans to expand their use of AI. They also noted that if they did not engage with these capabilities, they would be 'left behind' by competitors. The licensee referred to ongoing work to update their governance and risk management frameworks. However, this example exemplifies the risk in proceeding to adopt AI without adequate foundations in place, and the risk that gaps between use cases and governance will remain or widen in the face of competitive pressures.



AI governance: Questions for licensees



1 **TAKING STOCK**
Where is AI currently being used in your business?

2 **STRATEGY**
What is your strategy for AI, now and in the future?

3 **FAIRNESS**
How will you provide services efficiently, honestly and fairly when using AI?

4 **ACCOUNTABILITY**
Who is accountable for AI use and outcomes in your business?

5 **RISKS**
How will you identify and manage risks to consumers and regulatory risks from AI?

6 **ALIGNMENT**
Are your governance arrangements leading or lagging your use of AI?

7 **POLICIES**
Have you translated your AI strategy into clear expectations for staff?

8 **RESOURCES**
Do you have the technological and human resources to manage AI?

9 **OVERSIGHT**
What human oversight does your AI use require, and how will you monitor it?

10 **THIRD PARTIES**
How do you manage the challenges of using models developed by third parties?

11 **REGULATORY REFORM**
Are you engaging with the regulatory reform proposals on AI?

For more details, see pages 35–36

Why look at AI?

The use of AI in financial services and credit creates the potential for significant benefits to consumers, such as more efficient, accessible and tailored products and services.

But AI can amplify existing risks and create new risks to consumers.

The potential risks to consumers

Unfair or unintended discrimination due to biased training data or algorithm design: Biased AI outputs could have disproportionate, negative impacts on vulnerable individuals or groups, including financial exclusion (for example, being denied access to credit or insurance, or paying a higher price).

Incorrect information provided to consumers about products or services: AI models can provide information or advice that appears correct, but contains factual errors or fallacies. This exposes consumers to the risk of harm from relying upon such misleading or false information.

Manipulation of consumer sentiment or exploitation of behavioural biases: AI can allow for faster iteration of marketing and advertising material, and bespoke micro-targeting. AI can play on customers' feelings and restrict or manipulate their choices.

Breaches of data privacy and security: AI models may contain or reproduce confidential or sensitive information without the prior and informed consent of impacted individuals. AI models can also be vulnerable to cyber attacks and data leaks.

An erosion of consumer trust and confidence due to a lack of:

- **explainability** – AI models may use techniques that are too complex to be understood and explained by humans and be trained on data that is too vast and complex for humans to process, resulting in a 'black box', where decisions may not be traceable.
- **transparency** – Consumers may not be informed when AI has been used to make decisions that impact them, or when they are interacting with AI and AI generated information, and
- **contestability** – Consumers may not be provided with a process and the necessary information to contest the outcome of a decision facilitated by AI. Contestability is further undermined if consumers are unaware of the use of AI.

MANAGING RISKS FROM AI

Risks are very specific to each AI use case. For example, they can arise from the data input, from the technique or model used, as well as from the purpose, context, and level of automation of the models. Risks can also arise throughout the AI lifecycle and can change over time.

Because AI operates at scale, using vast amounts of data, risks can be amplified, and have the potential to cause harm at scale.

This means that AI creates new challenges for licensees in managing risks to consumers from AI. While this review did not test the outcomes from individual AI use cases, we have made observations on whether licensees are prepared to manage the risk of harms from the use of AI.

FINDINGS:

Use of AI

FINDING 1

The extent of AI use varied significantly but overall adoption is accelerating

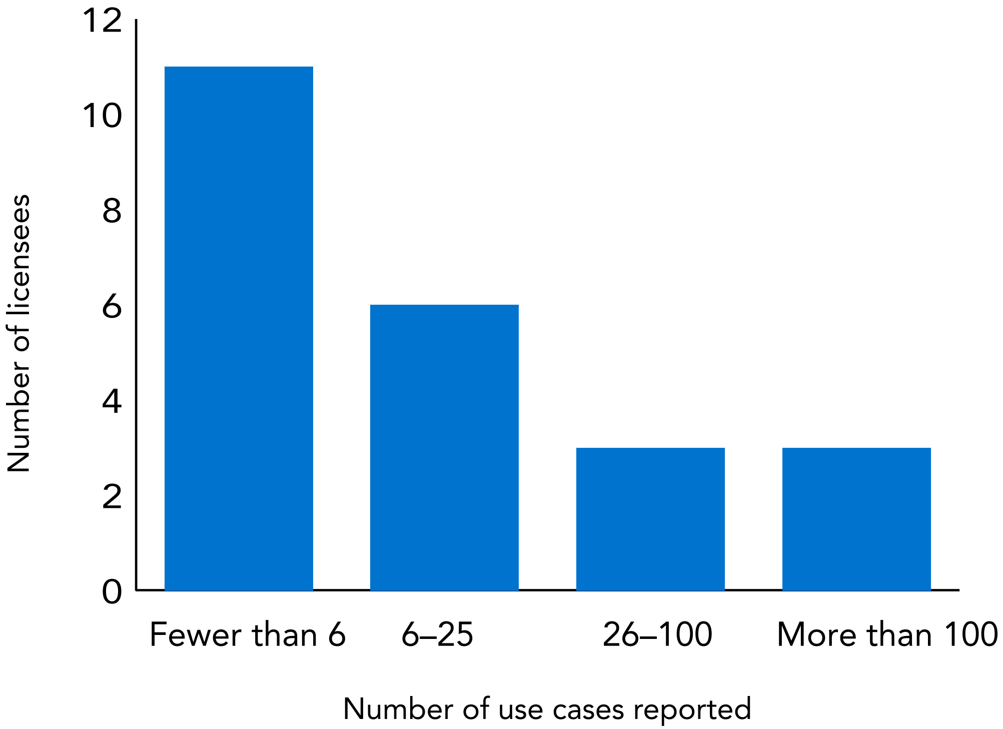
What we did

We collected data from 23 licensees on the number of AI use cases in use or in development (as at December 2023) where AI interacted with or impacted consumers.

What we found

- All but two licensees reported at least one AI use case that directly or indirectly impacted consumers.
- The number of use cases each licensee reported varied significantly – see Figure 1.

Figure 1: Use cases reported by licensees



Note: See Table 2 for the data shown in this figure (accessible version).

FINDING 1 (continued)

The extent of AI use varied significantly but overall adoption is accelerating

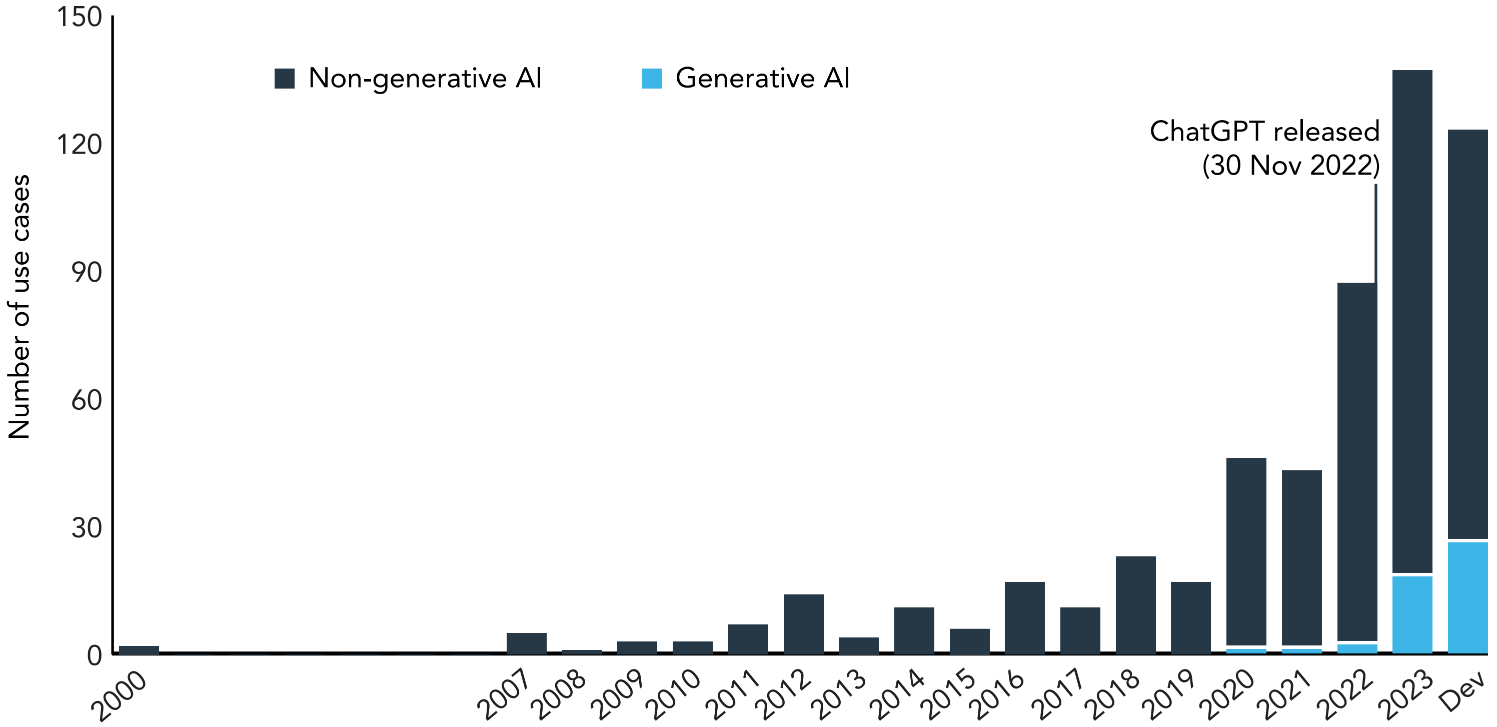
What we did

We reviewed a total of 624 use cases (see Appendix 1) and mapped them to the year they were deployed.

What we found

- AI adoption is increasing rapidly: 57% of all use cases reported were less than two years old or in development. Of the 624 use cases reported to us, 20% were still in development and had not yet been deployed.
- The adoption of generative AI is, unsurprisingly, a very recent development: 92% of generative AI use cases were deployed in 2022 or 2023, or in development as at December 2023.
- We can expect the pace of change to continue: 61% of licensees in the review told us they planned to increase their use of AI in the next 12 months. The remainder were planning to maintain their current AI use.

Figure 2: Number of AI use cases by deployment year



Note 1: See Table 3 for the data shown in this figure (accessible version)

Note 2: Dev = Advised to be in development by the licensee as at Dec 2023 – see Appendix 1 for more information. The development dates of 12 use cases were not provided or did not have a clear date and are not reflected in this graph. This graph includes use cases reported as 'in production' or 'in development' as at Dec 2023. It does not include use cases built and decommissioned before the data collection, or use cases where the model technique was not specified.

FINDING 2

Most current use cases applied long-established, well-understood techniques. But there was a shift towards more complex and opaque techniques, including generative AI

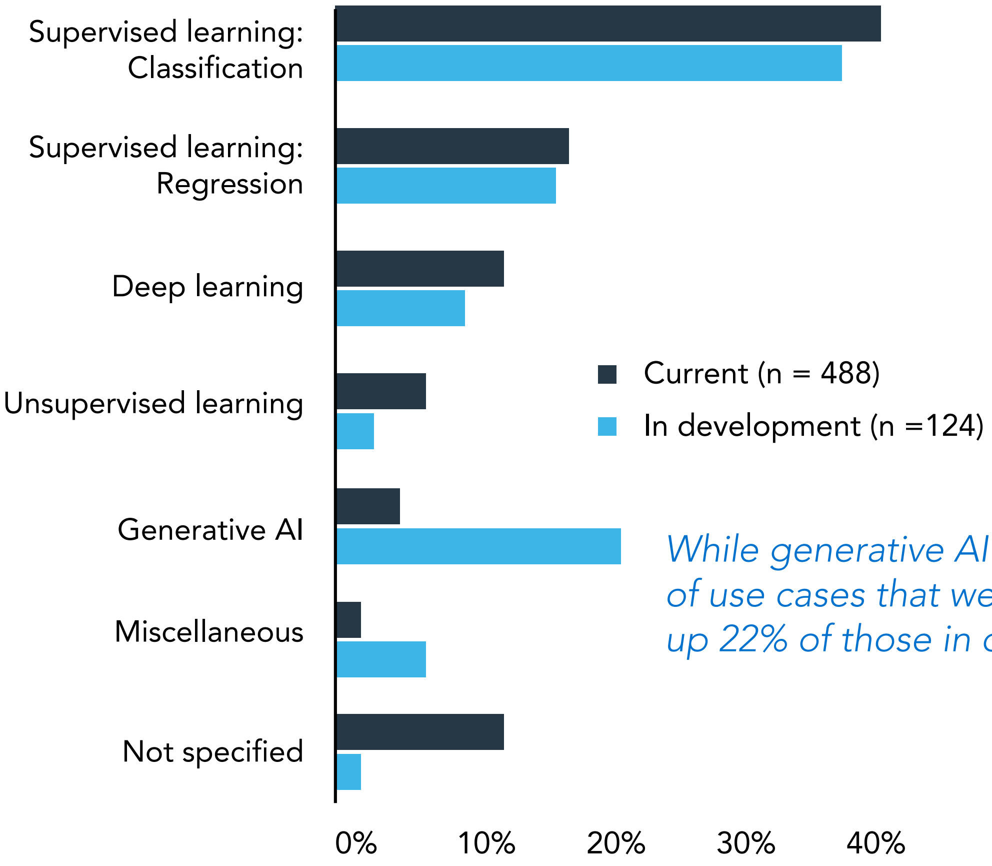
What we did

We assessed the complexity of model techniques used in each of the 624 use cases. Complex and opaque techniques can pose additional challenges for oversight. Challenges include understanding and explaining how AI obtains its results, determining whether results are reliable and accurate, and knowing whether outputs are unfairly biased or discriminatory.

What we found

- The majority of current use cases relied on well-known and established machine-learning techniques that produced explainable and interpretable results.
- We observed an increase in the use of more complex and opaque techniques (such as neural networks used in deep learning and generative AI), which are used for the processing and analysis of large volumes of images, audio and text data – see Figure 3. Together these represent 32% of the use cases we saw under development.
- The use of generative AI is set to increase exponentially. While generative AI made up only 5% of use cases that were in use, it made up 22% of those in development.

Figure 3: Model techniques by status



While generative AI made up only 5% of use cases that were in use, it made up 22% of those in development

Note: See Table 4 for the data shown in this figure (accessible version)

Most current use cases applied long-established, well-understood techniques. But there was a shift towards more complex and opaque techniques, including generative AI

How different model techniques were used

Supervised learning: Classification models were mostly used to predict if a consumer was likely to take out a financial product using explainable models such as logistic regression.

Supervised learning: Regression models were primarily used to derive prices, rates or forecast future time series.

Deep learning models were mostly used for natural language processing and optical character recognition, primarily when scanning analogue form data to speed up loan, insurance, or other form-heavy business processes.

Unsupervised learning models were mostly used for detecting strange or anomalous patterns in areas such as internal audit and fraud detection.

Generative AI models were used to generate first drafts of materials, or responses to customers in carefully constrained circumstances – see page 15 for more information.

Miscellaneous models were mostly non-predictive models, such as search engine optimisation or pattern matching.

'Not specified' models were models where licensees did not disclose the model technique. In some cases, these were models built by third parties, and licensees did not have this information.

WHAT IS GENERATIVE AI?

Generative AI is a type of AI that focuses on creating or generating novel content such as images, text, music, video, designs or recommendations.

Unlike traditional AI techniques that produce output that is programmed or copied from existing data, generative AI techniques are designed to generate output based on patterns, structures and examples learned from large data sets during the training process.

Generative AI models have certain characteristics that make them particularly prone to risks of harm. For example, they:

- tend to use large amounts of data for the training of the model. The presence of incomplete data in training sets mean that models have the potential to provide biased or inappropriate results
- can generate outputs that are false or inaccurate
- can use complex techniques that are not easily interpretable or explainable, and
- can be subject to novel cyber attacks.

FINDING 3

The way AI was used was mostly cautious

What we did

We looked at what licensees were using AI for, the role AI played in decision making and the types of data used by the 624 use cases. This allowed us to make some observations around risks posed by the use of AI. We also compared what we saw against use cases observed overseas or in literature.

What we found

- AI use was mostly cautious. Generally, AI was used to assist or augment human decision making or increase efficiency, rather than make autonomous decisions.
- Most AI use was internal facing. Where AI did directly interact with consumers, it generally operated within set parameters or alongside specific rules.

Decision making

Generally, models were not providing ungoverned outputs or replacing human judgement.

Decision making generally involved either:

Non-automated decisions – decisions where the model produced the output with a final check or verification performed by a human. For example:

- income/expense verifications for credit applications, and
- suggested responses for customer service staff.

Automated decisions – decisions made without human intervention, but operating within specific criteria, thresholds and rules set by humans. For example:

- credit score calculations that had to meet thresholds set by humans, and that operated alongside other set rules or checks (e.g. serviceability), and
- models that predict the likelihood of a transaction being fraudulent, which were referred to a human for review if they exceed a defined threshold.

Sources of data

Most data used by these models tended to be from internal sources. For example:

- customer financial information, such as transaction history or asset holdings, or
- details provided by customers when they applied for loans, lodged claims or requested quotes for financial products.

HOW GENERATIVE AI WAS BEING USED

Most current uses of generative AI or those in development were internal facing; they involved supporting staff and creating operational efficiencies.

In the limited instances where generative AI was used to interact with consumers, it was used within prescribed parameters (i.e. pre-vetted chatbot responses; chatbots deployed in limited circumstances).

Generating first drafts of documents, such as correspondence or marketing material

Call analysis; summarisation of call transcripts and consumer correspondence (e.g. for hardship identification)

Chatbots for internal use, and for customer engagement

Internal assistance: retrieving internal policies

Key uses of AI among licensees

Table 1 sets out key uses of AI among the licensees in our sample. At the time of our review and within the sample in scope, we did not see examples of the more concerning uses of AI observed in literature or overseas, such as the use of unconventional third-party data sources (e.g. social media activity) to inform credit or insurance decisions, or the use of generative AI models to produce targeted marketing messages to consumers to maximise sales, based on consumers' perceived emotional responses. This is point-in-time information (December 2023) and could change quickly, given the pace of innovation.

Table 1: Key uses among licensees

Area of use	Most common uses	Emerging uses (less commonly observed and/or in development)
Credit decisioning and management	Predicting credit default risk to support a decision, either by producing a score or rating where a minimum threshold must be met to proceed, or with other rules in automated decisioning. Monitoring existing credit holders to inform contact and collection strategies.	Accuracy improvements for decisioning, including to predict probability of recovery for defaults or arrears, and to prioritise customer contact.
Marketing	Analysing a consumer's spending patterns to segment them into specific groups so that they receive relevant marketing messages or offers. Optimising marketing communications and engagement by predicting best forms and times for contact.	Generative AI generating draft marketing copy for review.
Customer engagement and customer value proposition	Chatbots to answer simple customer questions based on pre-scripted responses. Cash flow forecasting and budgeting tools to assist customers with personal finances and to engage with their finances and with AI tools. Predicting credit card or product rewards offers likely to be of interest for customers.	Use of generative AI by customer-contact staff to summarise key information from customer complaints so they can respond to complaints in a more efficient and timely manner. The optimisation of consumer-facing apps and website layouts for ease of customer use based on browsing history and most-used features of the app.

Area of use	Most common uses	Emerging uses (less commonly observed and/or in development)
Fraud detection	Fraud detection activities, including transaction monitoring, and identification of fraudulent documents, and applications or claims. Use of biometric information for identity verification.	Identifying possible mule accounts and instances of account takeover. Identifying customers who may be susceptible to scams, to proactively prevent them.
Business efficiencies and compliance	Internal process efficiency, such as business analytics, quality assurance, and assistance for staff. Document indexing or data enrichment to improve information extraction from documents and support efficiencies in decision making. Triaging incoming complaints to enable more efficient complaints handling. Call transcription analytics that assist in quality assurance reviews of customer contact staff to ensure that treatment of customer issues and queries is within established quality and compliance standards.	Anomaly detection to identify internal errors or non-compliance and to efficiently target internal audit activities. Automated data cleaning, verification and integrity checks to correct for any potential errors such as spelling mistakes or incorrect labels in consumer form application data. Identification of financial hardship or vulnerability indicators in conversations missed by staff.
Pricing optimisation	Predicting the likelihood that a customer will switch to a competitor to drive targeted retention offers. To assist in determining discretionary discount offers on products upon a customer's request for a review.	
Insurance	Actuarial models for risk, cost and demand modelling. Supporting the claims process: Claims triaging, decision engines to support claims staff, document indexation, identifying claims for cost recovery. Identifying lapse propensity and prompts to contact consumers. Automating a component of the claims decisioning process, but humans remain responsible for overall claims decision.	Use of machine learning to increase efficiencies in the underwriting process, focused on automating the extraction of information and summarising key information about a customer's application. The use of generative AI and natural language processing techniques to extract and summarise key information from claims, emails and other key documents.

FINDINGS:

Risk management and governance

FINDING 4

There were gaps in arrangements for managing some AI risks

What we did

We asked licensees about how they identify and manage AI risks, including risks to consumers. We also reviewed any frameworks, policies and procedures that supported this.

What we found

- Approximately half of licensees had specifically updated their risk management policies or procedures to address AI risks. Other licensees relied on their existing policies or procedures without making changes.
- Licensees generally had documented policies or procedures for managing risks that are relevant to, but not specific to AI – such as those associated with privacy and security, and data quality.
- There were gaps in arrangements for managing some of the more unique AI risks, and for managing challenges such as transparency and contestability.

How licensees approached risk management

Licensees took different approaches to managing risk from AI. Approximately half of the 23 licensees had made specific changes to their risk management arrangements to reflect the characteristics of AI. They had updated existing policies with AI-specific content, or had created bespoke AI-related policies, standards or guidance.

However, it was not clear in all cases that these documents considered all AI risks, or that they were operationalised consistently; some were limited to generative AI and some provided only guiding principles, without establishing clear standards.

Most of the remaining licensees indicated that they relied on existing risk management frameworks and documents such as codes of conduct, or IT policies. Some of these licensees told us they had considered the adequacy of their existing documentation in light of their AI use – but in other cases, it was not clear that the reliance on existing materials was the result of a deliberate decision.

What policies included (or didn't include)

Nearly all licensees produced policies that broadly referenced risks that are relevant, but not specific to AI, such as privacy and security and data quality.

Only 12 of the licensees in the review had AI policy documents, guidance or checklists that referenced fairness, or related concepts such as risks of discrimination or bias against individuals, communities or groups. In some cases, references were principles-based, and it was not clear how consideration of these principles was embedded into operations.

Only 10 of the licensees had documented requirements or principles in place about disclosure to consumers when they were interacting with or affected by AI. Of these, some only prompted consideration of whether disclosure is appropriate and did not prescribe an approach to disclosing.

No licensees appeared to have implemented specific contestability arrangements for AI, though some noted this concept in principle. Some licensees referred to the availability of internal dispute resolution; though take-up of this in relation to AI will be impacted by the fact that consumers would not necessarily be aware AI was being used.

FINDING 5

There were gaps in licensees' assessments of AI risks

What we did

We asked licensees to set out the risks they had identified for each AI use case, how they mitigated these, and the frequency and type of monitoring they did.

What we found

We found some gaps in licensees' assessment of risks:

- Some licensees considered the risks of AI through a business lens rather than focusing on potential harm to consumers, and they did not consistently identify AI-specific risks such as algorithmic bias, or fully consider the impact of AI use on their regulatory obligations.
- We observed some weaknesses in how licensees provided meaningful human oversight, and in how they monitored for and responded to unexpected model outputs.
- We observed that licensees' consideration of transparency and contestability was relatively immature.

Business vs consumer risk

Many AI use cases were driving business efficiencies, and/or providing outputs to accountable human decision makers. These characteristics reduced the potential risk of consumer harm, which likely accounted for some licensees' more limited identification of consumer harm.

However, this was not the case for all use cases, and we identified gaps in how some licensees considered risks to consumers. For example, some licensees identified the risk of an incorrect model output, but noted the consumer could contest it, or staff could override it. However, they did not consider the potential harm if the model output caused a consumer to abandon their transaction altogether, potentially without knowing they could contest it (or indeed that AI was used).

In some instances, licensees were focused on business risk, and did not fully consider and manage the effects of their models on consumers. In those instances, mitigation and monitoring activity was also skewed towards business risk rather than consumer risk. For example, we observed instances where licensees used overseas-developed models for identity verification. They identified the business risk of failing to identify fraud and escalated cases that failed verification for manual review. However, their responses did not identify the potential for some groups to be disproportionately impacted, if overseas-developed models had not been adequately trained on a data set that was representative of the licensee's Australian customers.

We observed instances where licensees focused on business over consumer risk, or where the use of AI could have implications for licensees' compliance with existing conduct and consumer protection obligations, but this was not identified as a possible risk.

FINDING 5 (continued)

There were gaps in licensees' assessments of AI risks

Impact on regulatory obligations

We observed instances where AI use cases could have potential implications for licensees' compliance with existing conduct and consumer protection obligations, but this was not identified as a possible risk.

For example, customer segmentation by AI models in marketing could potentially identify customers who are not in a product's target market and lead to breaches of the design and distribution obligations.

This risk was generally not identified, though when prompted, licensees referred to existing controls to ensure compliance with design and distribution obligations, or to human oversight (i.e. a 'human in the loop').

Failure to consider the impact on regulatory obligations is particularly a risk where decisions about AI models or use cases are made without input or oversight by risk and compliance functions.

Few licensees considered algorithmic bias

Very few licensees proactively identified risks of algorithmic bias in their responses about particular use cases, or indicated they actively tested for bias.

Algorithmic bias describes systematic and repeatable errors in a computer system that create unfair outcomes, such as privileging one category over another in ways different from the intended function of the algorithm. In some cases, this was likely due to algorithmic bias being less relevant, given the nature of their use cases.

More licensees demonstrated that they had considered and mitigated this risk when we specifically queried this, but one licensee indicated that they did not test for bias.

Some licensees indicated they were aware that possible algorithmic bias in their data sets could influence outcomes, but they did not appear to test for a disparity in outcomes on an ongoing basis.

CASE STUDY

No evidence of consideration of impact on regulatory obligations

A licensee recognised that a model was under-predicting risk for a particular customer cohort. In response, the licensee adjusted the settings for the model, which, among other things, had the effect of increasing pricing offered to that cohort. The licensee used the outputs from the updated model even where it acknowledged that some consumers within the cohort could potentially be eligible for a lower price, based on outputs from other assessments that were subsequently introduced. We did not see evidence that the licensee had considered the flow-on impacts of this approach for consumers, in the context of the general obligation to provide financial or credit services efficiently, honestly and fairly.

There were gaps in licensees' assessments of AI risks

Transparency and contestability for consumers is a complex area and consideration of this was relatively immature

The review highlighted that the question of whether the use of AI should be disclosed to consumers is a challenging one, as well as the question of what information should be provided and when. Transparency is important as it allows for generally greater engagement and informed decision-making, but there are limits to the effectiveness of disclosure in protecting consumers.

A small number of licensees had considered whether their use of AI should be disclosed to consumers, and many of these appeared to consider the appropriateness of disclosure on a case-by-case basis.

Few licensees identified that a lack of transparency and contestability about the use of AI could erode consumer trust. This stance on disclosure likely reflects the nature of their use cases, with few licensees using AI to make automated decisions or interact directly with consumers. However, this potentially also reflects a lack of maturity in considering these issues.

Discussions with licensees highlighted transparency and contestability as a challenging area, with licensees questioning:

- how much AI had to be involved in an interaction or decision before it should be disclosed
- whether consumers would find disclosure useful, and
- whether it was necessary to introduce transparency now, given some models had been in use for a long time.

CASE STUDY

Not disclosing AI use to a consumer making a claim

One licensee used a third-party AI model to assist with indexing documents submitted for insurance claims, which included sensitive personal information, to improve efficiencies for claims staff.

The licensee identified that consumers may be concerned about their documents being sent to a third party to be read by AI, but decided not to specifically disclose this to consumers. The licensee's documents explained that its privacy policy stated that consumers' data would be shared with third parties, and the data was at all times kept in Australia.

But consumers were not specifically informed that some sharing of information involved AI, or about whether they could opt out.

While the AI use in this case only involved the provision of administrative support functions to human claims assessors, rather than any AI-driven decisions, it illustrates the complexity of the issue and the potential for loss of consumer trust.

There were gaps in licensees' assessments of AI risks

The importance of meaningful human oversight

We asked licensees to provide information about whether their models operated with a 'human in the loop'. Most licensees told us that this was the case for most of their models. In practice, however, this ranged from using the model's output to inform a human decision-maker, to referring exceptions to a human for review, to having human involvement in training and to periodically testing the model's operation.

What constitutes meaningful human oversight depends on the nature of the use case. Some licensees had purposefully decided that a human would be involved in and accountable for each decision where AI was involved and had documents affirming the accountability of humans for decisions. Other licensees conducted periodic checks of models in line with established controls, to identify issues such as model drift.

But in some cases, licensees' arrangements did not appear to provide sufficient human oversight, particularly where licensees did not fully understand models, as seen in the case study on page 7.

Monitoring and responding to issues was not consistent

Most licensees were monitoring their models, but practices varied widely, and we identified some gaps:

- A small number of licensees who were in the early stages of their AI journeys reported only testing at the pre-deployment and deployment stages and relying on trigger-based reviews for post-deployment monitoring. Better practice was licensees conducting periodic reviews of the model data and model output to ensure continual oversight.
- In many cases, monitoring practices focused on testing outputs against business metrics, rather than a more comprehensive analysis that considered possible consumer harm.
- Where licensees identified shifts or unexpected outputs, they differed in their response. Better practice was to conduct root cause analysis, including testing for consumer impact. Poorer practice was a licensee simply amending model thresholds to bring outputs within their business risk appetite, without root cause analysis or assessment of customer impact.

CASE STUDY

Different approaches when issues arose with AI models

Licensee A and Licensee B each deployed models to predict consumer credit default risk by producing a risk score.

Poorer practice: When scores were recalibrated by the external vendor on whose platform the model was built, Licensee A noted:

- 'time did not allow for thorough testing', and
- 'no documentation was created to ascertain the impact of this change'.

Better practice: When Licensee B's model produced unexpected scores, Licensee B:

- detected this as part of routine weekly monitoring
- conducted a root cause analysis to address the underlying issue, and
- investigated to identify any consumer impact (and found none).

FINDING 6

AI governance arrangements varied widely. We identified weaknesses that create the potential for gaps as AI use accelerates

What we did

The effectiveness of governance and risk management frameworks in relation to AI is a key factor in determining what risks a licensee's AI use poses. We therefore reviewed each licensee's approach to governance and the maturity of their governance arrangements.

What we found

- The maturity of AI governance and oversight varied significantly. Licensees sat somewhere on a spectrum of maturity of AI governance.
- We also identified some weak points in governance arrangements that will impact how licensees are able to manage risks from AI use, particularly if they accelerate adoption.

Maturity of governance arrangements

We identified three broad categories of approaches to governance that formed a spectrum from least to most mature:

- The least mature took a latent approach that had not considered AI-specific governance and risk.
- The most mature took a strategic, centralised approach.
- Licensees falling in between generally adopted decentralised approaches that leveraged existing frameworks.

Licensees weren't always entirely within one of the three categories, but sat somewhere along the spectrum.

The most mature licensees developed strategic, centralised AI governance approaches. The least mature licensees had not considered AI risks and governance, with no or few formal arrangements.

FINDING 6 (continued)

AI governance arrangements varied widely. We identified weaknesses that create the potential for gaps as AI use accelerates

Least mature

Most mature

Latent

At the time of our review, the least mature licensees had not considered AI risks and governance, with no or few formal arrangements.

Where these licensees used AI, they relied entirely on their existing frameworks. Any weaknesses in those translated into weaknesses in AI governance.

Leveraged and decentralised

Some licensees leveraged their existing governance and risk management arrangements to govern AI. While these licensees had considered the risks and opportunities of AI, their approaches tended to be decentralised, and determined by various parts of the business, based on their requirements.

These licensees generally did some or all of the following:

- considered that AI risks were covered by existing risk classes and did not include AI explicitly in their risk appetite statement
- relied on individual business lines to propose one-off AI use cases to address business needs
- demonstrated ownership and accountability for AI at a model or business unit level, but did not always have a senior executive accountable overall
- had pre-existing governance arrangements, policies and procedures for well-established forms of AI
- had documented AI and/or data ethics principles. Licensees varied in how well they incorporated these into relevant existing policies and operationalised them in practice.

The efficacy of the leveraged, decentralised approaches depended on:

- whether the licensee had considered its AI strategy and risk appetite
- the robustness of existing governance and risk management arrangements, and
- the nature and extent of the licensee's AI use.

Strategic and centralised

The more mature licensees developed strategic, centralised AI governance approaches. These licensees generally:

- had a clearly articulated AI strategy
- included AI explicitly in their risk appetite statement
- demonstrated clear ownership and accountability for AI at an organisational level, including an AI-specific committee or council
- reported to the board about AI strategy, risk and use
- had AI-specific policies and procedures that reflected a risk-based approach, and these spanned the whole AI lifecycle
- incorporated consideration of AI ethics principles in the above, and
- told us they were investing in resources, skills and capability.

FINDING 6 (continued)

AI governance arrangements varied widely. We identified weaknesses that create the potential for gaps as AI use accelerates

AI GOVERNANCE ARRANGEMENTS: BETTER AND POORER PRACTICES OBSERVED

AI strategy:

Better AI strategies set out clear objectives and principles for AI use, and considered the skills, capabilities and technological infrastructure required to deliver on the strategy.

Poorer AI strategies did not align AI use with desired outcomes and objectives or inform organisational risk appetite.

Board reporting:

Better practices included periodic reporting to the board/relevant board committee on holistic AI risk.

Poorer practices included ad-hoc reporting on a subset of AI-related risks, or none at all.

Oversight:

Seven licensees had, or were in the process of, setting up a committee or council to oversee AI.

Better practices were cross-functional, executive-level committees with clear responsibility and decision-making authority over AI use and governance.

Poorer practices included committees that met infrequently and had a poorly defined mandate.

AI ethics principles:

Twelve licensees had incorporated some of the eight Australian AI Ethics Principles in their AI policies and procedures. However, in some cases the references were high level and it was unclear how principles were to be applied in practice across the AI lifecycle. Licensees did not necessarily refer to all eight ethics principles – they were weaker in considering the disclosure of AI outputs and contestability.

Poorer practices included licensees relying on their organisational codes of conduct or other general policies instead of any explicit AI ethics principles.

AI governance arrangements varied widely. We identified weaknesses that create the potential for gaps as AI use accelerates

Weaknesses in governance and risk management

We identified the following weak points in governance, which can indicate that licensees' arrangements have not been fully operationalised or are starting to lag their AI use. These are particularly relevant for licensees taking a leveraged and decentralised approach.

Licensees and their boards may not have clear visibility of their AI use

Some licensees required extra time to collate use cases to respond to our notices. We suspect that in some cases a lack of an AI inventory, or the recording of models in several dispersed model registers, contributed to this.

CASE STUDY

Incomplete model register

One licensee required all models as defined in its Model Risk Policy to be entered into a model register and had developed a Model Risk Management System to maintain its register and manage model lifecycle workflow activities. However, in responding to ASIC's request, the licensee identified models missing from the register and failures to comply with the Model Risk Management System, suggesting that the licensee's centralised oversight remained incomplete.

FINDING 6 (continued)

AI governance arrangements varied widely. We identified weaknesses that create the potential for gaps as AI use accelerates

Evolving arrangements lead to complexity and fragmentation

Some licensees' AI governance frameworks and policies were spread across several documents, which had developed iteratively in response to particular issues and AI implementations, creating a risk of gaps and inconsistencies. These licensees may have difficulty overseeing their AI use and compliance with complex and fragmented frameworks, especially as AI use increases.

Evolving expectations are not applied to existing models

In some cases, licensees' expectations evolved, for example around the application of ethical considerations to consumer-facing models. However, updated policies and procedures were not necessarily applied to the existing suite of models, nor was there an expectation that they do so. Applying evolving policies to all existing models is important to ensure that they are implemented consistently.

CASE STUDY

Failure to apply evolving policies

One licensee had introduced a requirement that disclosure to consumers be considered in the context of the ethical principle around transparency. When we queried how they had considered disclosure for a particular consumer-facing model with a direct impact on consumers making an insurance claim, they said that while they had considered the costs and benefits of disclosure to the consumers at the time of the model's inception several years ago, there had been no formal process for this.

They indicated that they had not applied their current policy to models already in use. They told us that '[they] would certainly consider the question [about explainability and transparency] for new deployments … It wasn't in our process then; it is now'.

FINDING 7

The maturity of governance and risk management did not always align with the nature and scale of licensees’ AI use

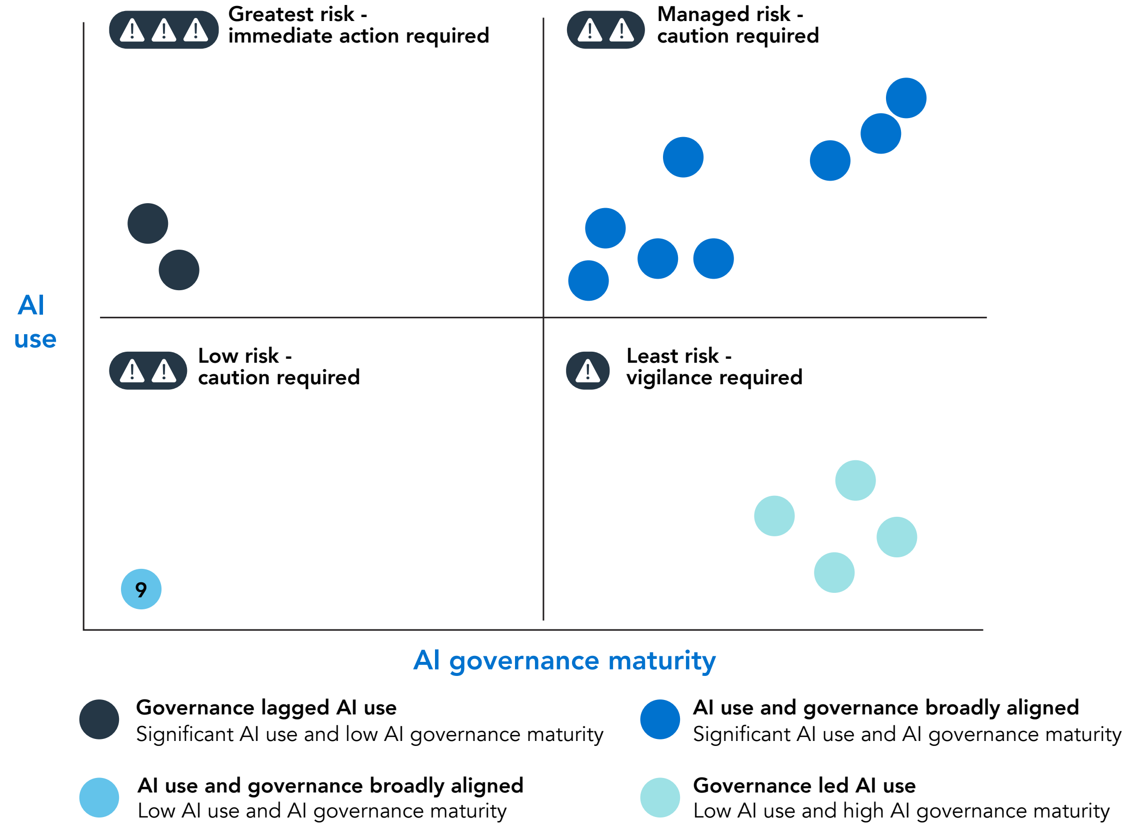
What we did

We compared the maturity of licensees’ governance arrangements to the nature and scale of their AI use to identify potential risks and gaps.

What we found

- We expected to see a clear correlation between those licensees with the most mature frameworks and the greatest AI use. Instead, the picture was more nuanced.
- For some licensees, their governance arrangements led their AI use. For most, AI governance and use was broadly aligned, but where they were updating their governance arrangements in parallel with increased AI use, this created a risk. For a small number of licensees, governance arrangements lagged their use of AI.
- As AI use accelerates, there is a risk that the gap between AI deployment and appropriate governance arrangements will widen.

Figure 4: Licensees’ AI governance maturity relative to AI use



Note: For an accessible version of this figure, see page 30.

The maturity of governance and risk management did not always align with the nature and scale of licensees' AI use

Significant AI use with low AI governance maturity: AI governance arrangements lagged AI use

Two licensees had started deploying consumer-impacting AI use cases without considering AI challenges in a systematic way or making changes to their existing governance and risk management arrangements.

Weaknesses in existing frameworks meant their arrangements were not adequate to manage AI risks. This cohort represents the greatest source of risk.

Both licensees in this cohort were relatively small. Neither was using or developing generative AI.

Significant AI use with high AI governance maturity: AI use and governance broadly aligned

For many licensees, governance arrangements and models were broadly aligned, but potentially coming under pressure, especially as AI use accelerates.

Eight licensees within this cohort had significant consumer-impacting use cases and mature governance arrangements relative to others. This cohort included licensees of various sizes. Most of the generative AI use cases that were in use (18 of the 22) and in development (23 of 26) belonged to licensees in this cohort.

The challenge for these licensees will be to maintain the adequacy of their arrangements and ensure they are fully operationalised as their AI use grows in scale and complexity, particularly if their approach to AI governance is already fragmented.

Low AI use with low governance maturity: AI use and governance broadly aligned

Nine licensees had limited AI use and had not put specific AI governance arrangements in place. Most licensees in this cohort had few use cases and had limited plans to expand, but some were considering uplifts to their governance frameworks to prepare for future AI use. There was only one very limited use of generative AI among this cohort.

Licensees in this cohort do not currently present significant risk, but risks could emerge if their posture towards AI changes without first establishing appropriate governance arrangements.

Low AI use with high AI governance maturity: Governance arrangements led AI use

Four licensees had relatively mature frameworks and yet did not have significant consumer-impacting models. This suggests that their decision to progress carefully is a deliberate one and is informed by a well-considered AI strategy and a thorough assessment of risk.

This cohort had particularly well-advanced governance frameworks relative to their use cases.

This cohort was starting to explore generative AI but was cautious in deployment and had appropriate frameworks in place.

VIGILANCE REQUIRED

Of the 23 licensees reviewed, 14 were planning to increase their use of AI. Of these, 13 were also planning, or had commenced, an uplift in AI governance. Only one appeared to have uplifted their governance before their anticipated uptick in AI use.

These figures underline the need for licensees to regularly review whether their governance arrangements are aligned to the scale and complexity of their use, and to consider the potential for gaps if AI uptake outpaces governance uplifts.

Licensees should be regularly reviewing and updating their governance and risk management arrangements to ensure the arrangements do not fall behind their evolving AI use.

FINDING 8

Many licensees relied heavily on third parties for their AI models, but not all had appropriate governance arrangements in place to manage risk

What we did

We asked licensees to identify which models were developed by third parties, and how they managed these relationships. Using third parties to develop or deploy models can bring significant benefits, such as overcoming limitations of resourcing and technical skills, especially for smaller licensees.

However, improperly managed third-party models can introduce risks, such as a lack of transparency and control, and security and privacy concerns.

There are additional challenges in risk management and oversight where licensees do not have insight into the operation and training of models.

What we found

30% of all use cases in our review had models that were developed by third parties.

Some licensees relied heavily on third parties for their models:

- For four licensees, 100% of the models in their use cases were developed by a third party.
- For 13 licensees, 50% or more models were developed by a third party.

Some licensees did not have robust third-party management procedures.

Better practices saw licensees setting the same expectations for models developed by third parties as for internally developed models.

FINDING 8 (continued)

Many licensees relied heavily on third parties for their AI models, but not all had appropriate governance arrangements in place to manage risk

CASE STUDY

Poorer practice oversight of third-party models

Most models reported by one particular licensee were developed by third parties. This licensee was not able to identify the AI technique used for all of its models and acknowledged the challenges: 'In [our] experience vendors are hesitant to provide details beyond standard marketing literature … due to intellectual property concerns.'

The licensee described processes for understanding accuracy and fitness for purpose of third-party models, but did not produce a third-party supplier policy or documented process for validating, monitoring and reviewing third-party models.

CASE STUDY

Better practice oversight of third-party models

One licensee had supplier risk frameworks in place that complemented its model risk requirements for third-party developed models, and set clear expectations, including to:

- obtain proof of independent validation from the supplier and validate the model internally before use
- establish service-level agreements to ensure models are implemented appropriately, including back-ups and disaster recovery plans, and
- establish a process to be notified of model changes, to obtain performance monitoring results and to consider fourth-party risks.

The licensee reported: 'All third-party models are subject to the same governance principles [as internally developed models].'

Where to from here for licensees?

Licensees must consider their existing regulatory obligations

What licensees need to do to comply with their existing regulatory obligations when using AI depends on the nature, scale and complexity of their business. It also depends on the strength of their existing risk management and governance practices. This means there is no one-size-fits-all approach for the responsible use of AI.

The regulatory framework for financial services and credit is technology neutral. There are several existing regulatory obligations that are relevant to licensees' safe and responsible use of AI – in particular, the general licensee obligations, consumer protection provisions and directors' duties. For example:

- **Licensees must do all things necessary to ensure that financial services or credit services are provided in a way that meets all of the elements of 'efficiently, honestly and fairly'**. Licensees should consider how their AI use may impact their ability to do so; for example, if AI models bring risks of unfairly biased or discriminatory treatment of consumers, or if the licensees are not able to explain AI outcomes or decisions.
- **Licensees must not engage in unconscionable conduct.** Licensees must ensure that their AI use does not result in acting unconscionably towards consumers. Licensees must ensure that AI is not used to unfairly exploit consumer vulnerabilities or behavioural biases. It is also critical that licensees mitigate and manage the risks of unfair bias and discrimination of vulnerable consumers from AI use.
- **Licensees must not make false or misleading representations.** Licensees must ensure that the representations they make about their AI use, model performance and outputs are consistent with how they operate. If licensees choose to rely on AI-generated representations when supplying or promoting financial services, they must ensure that those representations are not false or misleading.
- **Licensees should have measures for complying with their obligations, including their general obligations,** and these should be documented, implemented, monitored and regularly reviewed. If the use of AI poses new risks or challenges to complying with obligations, licensees should identify and update relevant compliance measures.
- **Licensees must have adequate technological and human resources**. Licensees should consider whether there are staff with the skills and experience to understand the AI used, and who can review AI-generated outputs. Licensees should have sufficient technological resources to maintain data integrity, protect confidential information, meet current and anticipated future operational needs (including in relation to system capacity), and comply with all legal obligations.
- **Licensees must have adequate risk management systems.** Licensees should consider how the use of AI changes their risk profile, whether this requires changes to their risk management frameworks, and whether they are still meeting their risk management obligations in light of their use of AI.
- **Licensees remain responsible for outsourced functions,** and they should have measures in place to choose suitable service providers, to monitor their performance, and deal appropriately with any actions by such providers. Licensees should consider how these expectations apply if they use third-party providers at any stage in the AI lifecycle.
- **Company directors and officers must discharge their duties with a reasonable degree of care and diligence.** These duties extend to the adoption, deployment and use of AI. Directors and officers should be aware of the use of AI within their companies, the extent to which they rely on AI-generated information to discharge their duties and the reasonably foreseeable associated risks.

AI governance: Questions for licensees

1. Taking stock:

- Where on your AI journey are you? Do you know where AI is used in your organisation?
- Do you have an AI inventory, and are you confident that it is being adequately maintained?

2. AI strategy:

- Are you clear where you are going, and why?
- Do you have a clear and documented strategy for what you want to achieve with AI, now and in the future? How does this align with your business objectives and risk appetite?

3. Ethics and fairness:

- What ethical challenges does your use of AI raise?
- How do you meet your obligations to provide financial services and credit efficiently, honestly and fairly when using AI?

4. Accountability:

- Who is accountable for AI use and outcomes, at model level and overall? Do they get the reporting they need to do their job?
- How are you measuring consumer outcomes from AI? Are you delivering benefits and avoiding harms?
- For accountable entities under the Financial Accountability Regime (FAR), have you considered the use of AI in key functions when assigning accountable persons and establishing clear lines of accountability?

5. Risk:

- Are you clear on conduct and regulatory compliance risk from AI, particularly risk to consumers? What is your risk tolerance?
- How are you identifying, mitigating and monitoring risk throughout the AI lifecycle?
- How will you document this, and monitor adherence to it?
- Do you have staff from multiple disciplines involved in assessing risk, and not just technical experts?
- Are your assessments of risk, your controls and your monitoring still adequate if your risk profile changes with your use of AI?

6. Alignment:

- Do your governance arrangements lead your AI use, now and for your future AI plans?
- How do the risks and ethical challenges change with a move towards more complex and opaque AI, such as generative AI? What changes do you need to make as a result to ensure your governance leads your use?

7. Policies and procedures:

- Have you translated your AI strategy and assessment of risk into policies for your staff, setting clear expectations through the AI lifecycle? Is your approach risk based? Do policies lead your AI use?
- Are your AI policies and procedures fit for purpose, now and for anticipated future use?
- How do you ensure your staff adhere to your AI policies and procedures?

AI governance: Questions for licensees

8. Resourcing:

- Do you have the right technological and human resources at all levels? How do you ensure your resources remain fit for purpose as AI use accelerates and evolves?
- How do you ensure your staff at all levels, including compliance and internal audit staff, have the skills and voice to engage with AI decisions and monitoring in their roles?

9. Oversight and monitoring:

- Are you clear on what human oversight you expect? Do you have procedures for when things go wrong?
- Do you have an action plan if a model is found to be producing unexpected outputs?
- Have you considered the adequacy of your business continuity, backup and disaster recovery plans for AI systems?

10. Third parties:

- How do you manage the challenges of relying on third parties?
- How will you validate, monitor and review third-party AI models?

11. Regulatory reform:

- Are you engaging with the regulatory reform proposals on AI?

Appendices

APPENDIX 1

Review methodology and definitions

Definition of AI

We defined 'AI' broadly, to include both:

- **advanced data analytics** – the autonomous or semi-autonomous examination of data or content using sophisticated techniques and tools, beyond those of traditional business intelligence (BI), to discover deeper insights, make predictions and generate recommendations, and
- **generative AI** – a category of AI that focuses on creating or generating novel content in forms such as image, text, music, video, designs and recommendations. Generative AI systems are designed to produce output that is not explicitly programmed or copied from existing data, but rather is generated based on patterns, structures and examples learned from large datasets during the training process.

We adopted this broad definition because risks to consumers are not limited to newer, more complex techniques that are the subject of widespread debate, such as generative AI. If governance is inadequate, and risks are not well identified, mitigated, and monitored, consumer harm can arise even from techniques or models that have been used for many years.

Review scope

We reviewed the current and planned uses of AI, as at December 2023, by a sample of 23 licensees. The licensees were drawn from the banking, credit, general and life insurance, and financial advice sectors. We looked at use cases where AI interacted with or impacted consumers.

The sample was *not* representative of AI use generally, or of the sectors in the review. We selected licensees that we identified as most likely to be using AI, based on their business model and ASIC's intelligence. Some were found to be early on the AI journey in our review.

We limited the scope of our review to AI use cases that directly or indirectly impacted consumers. The scope did *not* include:

- back-office functions
- investing, markets and trading activities, or
- models used for compliance with laws administered by other regulators.

The review was intended to provide ASIC with an understanding of how licensees are using and planning to use AI, and how they are considering and mitigating associated risks. We did not test for consumer outcomes from individual AI models.

Review methodology

We reviewed information for 624 use cases provided by the 23 licensees:

- For licensees with a relatively small number of use cases, we reviewed detailed information for all of their use cases.
- For licensees with a larger number of AI use cases, we reviewed detailed information for a subset of their use cases, selected by ASIC.

We also asked the 23 licensees to respond to questions and provide relevant documents to enable ASIC to understand their AI strategies, policies, processes and practices. We reviewed licensees' responses and supporting material – these covered governance and oversight, risk management, consumer benefits, harms and outcomes, monitoring, reporting, and future plans.

We held meetings with 12 of the licensees in the review during June 2024, to ask for further context and clarification about their use of AI and their governance arrangements.

APPENDIX 1

Review methodology and definitions

The nature of the use cases and their models varied significantly. In preparing this report, we have considered the context and operation of the models and provided generalised views. In some circumstances – such as where a use case's model contributed to decisions affecting consumers – we have characterised 'explainable' and 'interpretable' attributes positively. However, we acknowledge that there is an inherent trade-off between the complexity and explainability of a model, and that a more complex model is not inherently riskier than a simpler model. The risks of AI are heavily context dependent.

Data provided by licensees

Model techniques

As highlighted by the 'not specified' category in Figure 3, some licensees did not provide detail about the model technique used in a use case due to commercial sensitivities or a lack of transparency from third-party providers. We have used the category 'not specified' for these use cases or, where possible, assigned the use case models into categories based on model characteristics inferred from information provided to us. As such, there may be some small variances present in the actual model types and categories.

Number of use cases

Licensees had varying approaches to responding to our request for use case information. Certain licensees responded to our request by providing one use case per line item, while some larger licensees provided multiple use cases in a single line item. Unless the number was specifically confirmed by the licensee, we have based the number of licensees' use cases on the number of line items provided. As such, there may be some licensees with a greater number of use cases in scope than set out in Figure 1.

Model development year

Licensees had varying approaches to responding to our request for information about the date of deployment for use cases. Some licensees preferred to provide us with the date they updated a model rather than the original date of deployment. For consistency, we have chosen to take the earlier date when reporting this data. There may also be a selection of use cases that were decommissioned before we requested the information. Since they were not currently in use or in deployment, these models would have been omitted from the sample and are not reflected in Figure 2.

Use cases and their corresponding models were classified as 'current' if they were operational and/or working on live data at the time of the review.

Use cases and their corresponding models were classified as 'in development' if the model had not yet been deployed by the licensee for use on live, real-time data streams, was part of a pilot study, was still being built, or was scheduled for deployment.

APPENDIX 2

Accessible data points

Table 2: Number of use cases reported by licensees

Number of use cases reported	Number of licensees
Fewer than 6	11
6–25	6
26–100	3
More than 100	3

Note: This is the data shown in Figure 1.

Table 4: Model techniques by status

Model techniques	Current (n = 488)	In development (n = 124)
Supervised learning: Classification	42%	39%
Supervised learning: Regression	18%	17%
Deep learning	13%	10%
Unsupervised learning	7%	3%
Generative AI	5%	22%
Miscellaneous	2%	7%
Not specified	13%	2%

Note: This is the data shown in Figure 3.

Table 3: Number of AI use cases by deployment year

Deployment year	Use cases: non-generative AI	Use cases: generative AI
2000	2	0
2001	0	0
2002	0	0
2003	0	0
2004	0	0
2005	0	0
2006	0	0
2007	5	0
2008	1	0
2009	3	0
2010	3	0
2011	7	0
2012	14	0
2013	4	0
2014	11	0
2015	6	0
2016	17	0
2017	11	0
2018	23	0
2019	17	0
2020	44	1
2021	41	1
2022	84	2
2023	118	18
Dev	96	26

Note: This is the data shown in Figure 2.

APPENDIX 3

Key terms

ADA	Advanced data analytics – the autonomous or semi-autonomous examination of data or content using sophisticated techniques and tools, beyond those of traditional business intelligence, to discover deeper insights, make predictions and generate recommendations
AI	Artificial intelligence – a collection of interrelated technologies that can be used to solve problems autonomously and perform tasks to achieve defined objectives. In some cases, this is done without explicit guidance from a human being. For the purposes of this report, AI includes advanced data analytics and generative AI
AI lifecycle	Includes the following stages: design, data and modelling phase; verification and validation phase; deployment phase; and operating and monitoring phase. These phases may take place in an iterative manner and are not necessarily sequential
algorithm	A set of instructions that guide a computer in performing specific tasks or solving problems. Algorithms can range from simple tasks, like sending reminders, to complex problem solving, which is crucial in AI and machine learning
algorithmic bias	Systematic and repeatable errors in a computer system that create unfair outcomes, such as privileging one category over another in ways different from the intended function of the algorithm
contestability	The ability for the outputs or use of an AI system to be challenged by people impacted by that AI system
deep learning	A machine-learning technique that uses interconnected layers of 'neurons' to learn and understand patterns in data, especially in tasks like image recognition and speech synthesis. 'Deep' refers to the fact that the circuits are typically organised into many layers, which means that computation paths from inputs to outputs have many steps. Deep learning is currently the most widely used approach for applications such as visual object recognition, machine translation, speech recognition, speech synthesis and image synthesis
explainability	The ability of an AI system to be comprehended and trusted by humans. Explainable AI allows an understanding of how an AI system has produced a specific output
generative AI	A category of AI that focuses on creating or generating novel content in forms such as images, text, music, video, designs and recommendations. Generative AI systems are designed to produce output that is not explicitly programmed or copied from existing data, but rather is generated based on patterns, structures and examples learned from large datasets during the training process
licensee	A person who holds an Australian financial services licence under section 913B of the *Corporations Act 2001* and/or an Australian credit licence under section 35 of the *National Consumer Credit Protection Act 2009*
machine learning	A branch of AI and computer science that focuses on the development of systems that are able to learn and adapt without following explicit instructions, imitating the way that humans learn, gradually improving their accuracy, by using algorithms and statistical models to analyse and draw inferences from patterns in data

APPENDIX 3

Key terms

model	A machine-learning or AI algorithm that has been trained to do a particular task
model technique	A simplified way of referring to a model's particular algorithm to perform a certain task alongside the underlying structure or design of a machine-learning model. Also referred to as a model's architecture in technical terms
natural language processing	A branch of AI with techniques to help computers understand, interpret and manipulate human language
neural networks	Computer models inspired by the human brain's structure. These interconnected artificial neurons, organised in layers, learn from data to make predictions in machine learning, underpinning deep learning
optical character recognition	A process that converts an image of text into a machine-readable text format
supervised learning	A sub-category of machine learning where algorithms learn from labelled data to make predictions or classifications, often with high accuracy
training data	The data used in the first instance to develop a machine-learning model, from which the model creates and refines its rules
transparency	The disclosure provided to people about when they are engaging with an AI system or when AI has been used to make decisions that impact them
use case	A model or several models that are applied to a specific context – for example, a logistic regression model applied to predict a customer's likelihood of default
unsupervised learning	A sub-category of machine learning where algorithms group data objects based on similarities, without prior category specifications

ASIC
Australian Securities &
Investments Commission

asic.gov.au