



**Submission by the
Financial Rights Legal Centre**

Department of Industry, Innovation and Science

Artificial Intelligence, Australia's Ethical
Framework – A Discussion Paper

May 2019

About the Financial Rights Legal Centre

The Financial Rights Legal Centre is a community legal centre that specialises in helping consumers understand and enforce their financial rights, especially low income and otherwise marginalised or vulnerable consumers. We provide free and independent financial counselling, legal advice and representation to individuals about a broad range of financial issues. Financial Rights operates the National Debt Helpline, which helps NSW consumers experiencing financial difficulties. We also operate the Insurance Law Service which provides advice nationally to consumers about insurance claims and debts to insurance companies, and the Mob Strong Debt Help services which assist Aboriginal and Torres Strait Islander Peoples with credit, debt and insurance matters. Financial Rights took close to 25,000 calls for advice or assistance during the 2017/2018 financial year.

Financial Rights also conducts research and collects data from our extensive contact with consumers and the legal consumer protection framework to lobby for changes to law and industry practice for the benefit of consumers. We also provide extensive web-based resources, other education resources, workshops, presentations and media comment.

Sign up to our E-flyer at www.financialrights.org.au

National Debt Helpline 1800 007 007

Insurance Law Service 1300 663 464

Mob Strong Debt Help 1800 808 488

Monday – Friday 9.30am-4.30pm

Introduction

Thank you for the opportunity to comment on Department of Industry, Innovation and Science's (the Department's) Artificial Intelligence, Australia's Ethical Framework discussion paper.

The Financial Rights Legal Centre's (Financial Rights') interest in the issues being raised by the discussion paper and the development of Australia's ethical framework on Artificial Intelligence (AI) is borne of the already increasing use of AI in the financial services sector and its expected impact on financially vulnerable consumers and their rights.

This submission is based upon our casework experience with people experiencing financial hardship and what we have started to witness with respect to the increased use of new AI technologies.

In this submission, Financial Rights provides further details of the increasing use of AI technologies in the financial services sector and their current and expected impact on consumers.

We argue that AI's use in financial services reveal further ethical issues that need to be considered by the Department in developing its AI Ethics Framework – namely that the Department needs to take into account the ability for AI to automate the profiling of consumers for profit and the embedding of unfair and exclusionary price discrimination models into financial services both of which have the strong potential to lead to increased economic inequality and financial exclusion.

In examining the draft core principles put forward, Financial Rights generally supports the eight identified but argue for the principals to be expanded to incorporate notions of well-being, inclusion, and trust.

Finally we recommend that the AI Ethics Toolkit explicitly refer to the utility of product/consumer testing with diverse demographics in order to better meet the core principles.

Financial services and Artificial Intelligence

Financial Rights notes that the Discussion Paper details a number of examples of AI in practice. We wish to add examples of current and expected impacts of AI technologies on the financial services sector and consumers for consideration in developing its AI Ethical Framework.

We note that the discussion paper defines AI as:

A collection of interrelated technologies used to solve problems autonomously and perform tasks to achieve defined objectives without explicit guidance from a human being.

In the financial services sector new computing power and technology has led to:

- an expansion of the data collection of their own customers as well as from external sources both conventional (e.g. government databases and transactional data), and unconventional (e.g. social media, emails etc.);
- advanced data processing techniques; and
- advanced analytical, artificial intelligence and algorithmic techniques including predictive analytics.

AI is consequently well suited to exploitation in the financial services sector given AI's ability to recognise patterns, predictively anticipate future events based on large sets of data and make decisions based on this information

The burgeoning financial technology (or FinTech) sector is creating products, services and tools that are transforming ways the sector undertakes risk assessment, detects and manages fraud and assists consumer manage their finances.

New and emerging services involving some element of AI technologies include:

- mobile and online banking;
- Open Banking applications using consumer transaction data to assist in a series of services including but not limited to account switching, mortgage search services;
- new personal financial management services (such as Money Dashboard);
- investment and wealth management services with automated or robo-advisers services such as Wealthfront;
- new lending and unsecured credit services based on data led credit-scoring and risk profiling (e.g. Afterpay, Defer It);
- encrypted digital wallets that stored bank, debit or credit card detailing for online payments (e.g. PayPal and AliPay);
- neo banks and FinTech savings banks such as AliPay's Yu'eBao;
- offline mobile payments such as Apple Pay, Android Pat or Ali Pay used at retail locations; and
- credit scoring and social scoring – utilising financial and social datasets from non-traditional sources such as Facebook and other social media to create measures of credit worthiness, outside of the “traditional” credit reporting and scoring.

There is also a sub-class of FinTech known as insurance technology or InsurTech. InsurTech is using AI technologies in three key ways.

Firstly it is using AI to build behaviour into premium pricing. Connected devices and telematics technology (e.g. Fitbit), connected home technologies (e.g. Amazon Alexa) and what is known as the “Internet of Things” (e.g. connected smoke alarms, locks, fridges and light switches) are also being put to specific use by the insurance sector.

Telematics technologies involve the use of GPS technology and increased information processing power to collected and transmit information and data to insurers directly. Telematics devices being used by insurers include:

- Motor vehicle telematics – devices in vehicles that can record GPS location data as well as information from a vehicle’s engine management system to monitor all aspects of driving style QBE, for example, offers “Insurance Box for young drivers”. Here, drivers install an electronic device or “black box” in their car that transmits back to the insurer a detailed breakdown of their driving habits in areas such as their braking, acceleration, steering, cornering, speed and night driving.¹ QBE then calculate a “DriveScore” rating to evaluate the driver. The higher the DriveScore the less the policyholder will pay for insurance. The lower the score, the more the driver pays.
- Home telematics – devices can monitor the use and supply of a range of utilities as well as security of a home. Smart smoke alarms, water leak and freeze detectors are already being used overseas by insurers.
- Health monitors – fitness monitors such as Apple Watch and FitBit can record the location, movement, activities and other health information. AIA vitality² is an example of a product that enables a life insured to gain benefits such as discounts and rewards via the earning of “vitality points” for activities undertaken.³ Others include Asteron Life Plus Health Rewards and Bupa Living Well.

Life insurers are using genetic testing technology in their underwriting provided to them under disclosure laws, an ability borne of increased computing processing power, new hardware and data analytics.

AI is also being used in insurance to personalise the customer experience through the use of chatbots and other tools to improve the sales experience.

And finally AI is being used to ‘enhance’ the claims handling process including fraud detection through data analysis and machine learning, and speeding up the settling of claims. Many of the FinTech and InsurTech services are using algorithms and AI for automated decision making.

Ethical implications of the use of AI in financial services

FinTech products and services’ utility arises from a near total reliance on data – largely a consumer’s personal financial data - their transactions history, credit history, biometrics etc. FinTechs and InsurTechs are also integrating financial data with other data about individuals drawn from social media and other sources – information that people would consider have nothing to do with their financial status. InsurTech is tracking people’s every movement and drawing conclusions about a person’s identity and their life derived from the use of their car.

This increased collection of data is feeding the creation of a “financial identity” – a concept increasingly used by financial institutions to get to know their customer more.

¹ <https://www.qbe.com.au/news/car/how-insurance-box-works>

² <https://www.aiavitality.com.au>

³ <https://www.aiavitality.com.au/vmp-au/rewards>

Financial institutions have for years stored and verified customer identities and attributes through “Know Your Customer” systems i.e. the process by which banks or other financial institutions identify their customers in order to evaluate the possible legal and other risks. They therefore have a commercial incentive to collect more and more accurate information about their individual customers. The World Economic Forum in 2016 has in fact argued that financial institutions “should champion efforts to build digital identity systems, driving the building and implementation of identity platforms.”⁴

However the development of an increasingly accurate financial identity built by data has serious consequences for consumers.

Some positive impacts include enabling increased access to financial services and potentially empowering consumers in increasing their own financial literacy, behavior or wellbeing.

There are however a series of impacts upon consumers – particularly consumers experiencing financial vulnerability or hardship – that are of significant concern to Financial Rights. We detail the following identified harms. While some of these cleave to ethical issues already raised in the Discussion Paper, there are new and further dimensions that we believe need to be considered in developing an Ethical Framework.

Profiling for profit: Increased economic inequality and financial exclusion

Financial Rights is concerned that with the rise of AI in FinTech, we will see increased occurrences of consumers being ‘profiled for profit’, which will see more people experiencing financial difficulties or hardship being offered unsuitable (but highly profitable) products.

Target marketing of products to particular groups of consumers is not new. In consumer lending, technology can be used to identify consumers who are likely to be profitable, tailor and price products that the most profitable customers are likely to accept, and develop strategies to reduce the likelihood that the most profitable customers will close their accounts.

Consumers struggling with debt are however often the most profitable customers for banks and lenders. It is often argued that it is not in the interests of lenders to extend credit to people who are unable to repay. However, our experience suggests that many consumers struggle for years at a time to make repayments to their credit accounts without ever reaching the point of default, but paying significant amounts of interest. These customers are very profitable for lenders, despite the fact that repayments are causing further financial hardship.

Another example of profiling for profit is the current non-AI based practice (used by many pay day lenders) whereby customers are asked to provide access to their bank statements via third party account aggregation software for responsible lending assessment purposes. Providing access to this ‘screen scraping’ technology can amount to a breach of the terms and conditions of a customer’s bank account, and can put customers at risk of losing their protections under the ePayments Code. We are concerned that some lenders may be keeping these links open

⁴ World Economic Forum & Deloitte (2016) “A Blueprint for Digital Identity: The Role of Financial Institutions in Building Digital Identity”: page 28
http://www3.weforum.org/docs/WEF_A_Blueprint_for_Digital_Identity.pdf

after the initial lending assessment has been completed, so that they can direct their marketing of further loans to consumers specifically when their account is empty and their need for cash is likely to be at a high point.

What is of more significant concern is that with the automating of these processes through an Open Banking regime and the application of AI to this, there will be significant room for increased exploitation. Consumer advocates in the United Kingdom, have already raised concerns that 'Open Banking enables lenders to continually monitor accounts and take repayment as soon as income is detected'. These are real risks that are poorly understood by consumers and unlikely to be dealt with by disclosure and consent because of the take it or leave nature of the service.

Price discrimination on low-income households

Much of the promise of FinTech is that more tailored products and services will be made available with lower fees or lower loan interest rates for many banking customers. However, the flip side to lower fees and interest rates for some is that costs will increase for others. These 'others' will undoubtedly be Australia's most vulnerable, disadvantaged and financially stressed households.

Those in more precarious financial situations – again identified as such by their data driven financial identities - will likely be unfairly charged higher amounts for credit, or be pushed to second-tier and high cost fringe lenders. In other words, the consumers who can afford it the least will pay the most be it via higher interest rates or higher fee products. There are serious fairness considerations at play here. As banks and credit providers are increasingly able to use consumer data and technology to better automate the targeting of particular financial services offers to 'profitable' consumers, we will likely see an increased use of 'risk-based pricing'. This may result in some lenders targeting 'riskier' borrowers with higher interest rates. While risk based pricing has effectively existed in Australia in the non-bank sector for some years, we have recently seen the first example of risk based pricing by a major bank.

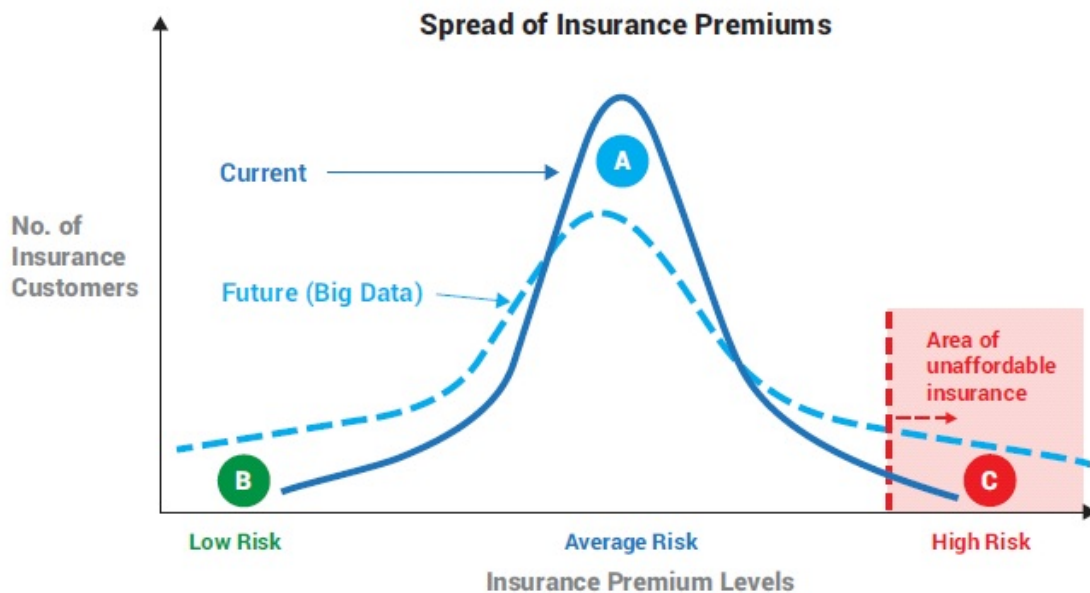
A 2015 report by United States organisation Data Justice raises concerns that enabling advertisers to offer goods at different prices to different people to extract the maximum price from each individual consumer. The report found that such price discrimination not only raised prices overall for consumers, but particularly hurts low-income and less technologically savvy households. In fact, the ability to segment the market further will likely mean that firms can 'cherry pick' the most commercially viable consumers and exclude others (or charge them more).

It is clear that the result of the price discrimination enabled by these technologies in the financial services sector is a downward spiral of debt. A self-fulfilling prophecy ensues. A consumer's low credit rating decreases from a default, which in turn feeds an algorithm of credit-worthiness leading to higher interest rates and further financial difficulty and further defaults.

In the insurance sector, the increased use of big data analysis and automated processing allowed by increased computing power will enable insurers to increasingly distinguish

between risks on an increasingly granular level. This will lead to the higher risks only being able to be insured for higher prices or on worse terms. According to the Actuaries Institute

At the extreme, some policyholders will have their risks assessed as so high that the price will be prohibitive or insurers will decline to provide cover. The following diagram illustrates the effect that increasing data will have on insurance premiums.



Overall, there will be fewer insureds treated as “average” risk (area A) and paying average premiums. They will increasingly be classed as either lower or higher than average. Greater numbers of insureds will thus be recognised as being lower risk and given lower insurance premiums (area B). Conversely there will be more consumers falling into the higher risk category, ultimately reaching the “unaffordable” levels of insurance premiums (area C).

...

In response, some people may mitigate or avoid the risk. Others who find the insurance premiums for their risk to be unaffordable may have to take the risk themselves. If the risk event does happen, they will suffer financially. The more people change from insured to uninsured status because of price increases arising from more targeted use of data, the greater the burden will be on the public purse or on others outside the insurance system.⁵

Unfair and exclusionary price discrimination practices in insurance and the broader financial services sector should be a cause for serious concern where it contributes to lower-income people paying higher prices than others, or where pricing discrimination negatively affects particularly marginalised groups. In the insurance sector, people who need insurance the most may increasingly find they have been excluded completely as a result of issues which may be completely beyond their control. These are key issues of fairness and equity.

⁵ Page 19-20, Actuaries Institute, The Impact of Big Data on the Future of Insurance <https://actuaries.asn.au/Library/Opinion/2016/BIGDATAGPWEB.pdf>

Indirect Discrimination

Closed proprietary algorithms used by FinTechs and InsurTechs to automatically calculate say an individual's credit worthiness or the interest rate they are offered could also potentially lead to situations where consumers are denied access to crucial products and services based on accurate or inaccurate data without the ability to determine why or to correct underlying assumptions.

Algorithmic bias or discrimination is already well documented⁶ and arises when an algorithm used in a piece of technology – say a FinTech product or service – that reflects the implicit or explicit values of those who are involved in coding, collecting, selecting, or using data to establish and develop an algorithm.

Credit scoring, social scoring or e-scoring algorithms for example can produce feedback loops where somebody from a particular suburb where a lot of people default can be given lower credit ratings due to that association. Statistical correlations used by actuaries between a person's postcode (here geographical information standing in for a particular race, ethnicity or culture); their language patterns on social media; their potential to pay back a loan; or, keep a job; can lead to significant discrimination being built into opaque black box algorithm technology.

Cybercrime, identity theft and material theft

As our financial services sector becomes more and more reliant on FinTech and InsurTech with more and more accurate information about financial identities – the fuel on which AI depends - individuals become increasingly vulnerable to cybercrime.

Firstly consumers are vulnerable to identity theft. With increasingly sensitive and accurate data being held by FinTechs, breaches of these datasets make it easier for criminals to use this identifying information to undertake subsequent crimes, financial or otherwise.

The vulnerability of the data protection systems in place also facilitates actual theft of property – that is the hacking of FinTech systems to access payment systems and steal money. According to Juniper Research, fraudulent online transactions will reach a value of \$25.6 billion by 2020⁷ In Australia online credit card fraud, with transactions made using stolen card details hitting \$417.6 million in 2016, more than doubling since 2011.⁸

⁶ See Cathy O'Neil, Weapons of Math Destruction, 2017

⁷ "Online Transaction Fraud to More than Double to \$25BN by 2020' Juniper Research UK, May 2016.

⁸ Lucy Cormack, Carol Saffer, Online credit card fraud on the rise, accounting for 78 per cent of total card fraud in Australia, SMH, 3 August 2017 <https://www.smh.com.au/business/consumer-affairs/online-credit-card-fraud-on-the-rise-accounting-for-78-per-cent-of-total-card-fraud-in-australia-20170802-gxnwd7.html>

The news⁹ that UK company Cambridge Analytica legitimately gathered some personal data from Facebook accounts and concurrently illegitimately gathered other people's data, and then, when found out and were requested to delete the data, did not, has raised public consciousness over the potential for data to be misused in various ways. Combined with the never ending list of significant and high profile data breaches at Equifax, Ashley Madison, Yahoo and more, consumer awareness of how vulnerable consumers are is increasing every day.

Core principles for AI

The application of AI in the financial services sector leads to a series of ethical issues additional to those found in the discussion paper that require consideration in developing a generally applicable AI Ethical Framework. While the paper does touch upon the potential for discrimination (direct, indirect or otherwise) and the impact upon security and privacy, the Department needs to take into account the ability for AI to automate the profiling of consumers for profit and embed of unfair and exclusionary price discrimination models into financial services both of which have the strong potential to lead to increased economic inequality and financial exclusion.

In examining the draft core principles put forward, Financial Rights generally supports the eight identified. We however provide the following observations on draft core principles 1, 5 and 8 re: Generating net benefits, fairness and transparency & explainability and propose incorporating principles of well-being, inclusion, and trust.

Core principle 1 – Generates net benefits

Core principle 1 states:

Generates net-benefits: *The AI system must generate benefits for people that are greater than the costs.*

This principle recognises that there will be some costs in the use of AI but that the benefits must outweigh these. As guidance to AI developers and corporations this leaves far too much wiggle room to make a judgement on the impact of the costs. One financial service provider's cost benefit analysis may not be the community's analysis. Does the cost of exploitation and harm to financially vulnerable sections of the Australian community outweigh benefits of frictionless convenience to the middle class? Many in the financial services sector would argue no. We would disagree. We have already seen the consequences of this lack of meeting community expectations in the financial service sector in the Royal Commission into Financial Services. A pure cost-benefit analysis embeds a zero sum game into AI.

⁹ 'I made Steve Bannon's psychological warfare tool': meet the data war whistleblower, *The Guardian*, 18 March 2018 <https://www.theguardian.com/news/2018/mar/17/data-war-whistleblower-christopher-wylie-faceook-nix-bannon-trump>

A better approach may be to co-opt the concept of well-being already referred to in the discussion paper. The first line of the Discussion Paper states:

Artificial Intelligence (AI) has the potential to increase our well-being; lift our economy; improve society by, for instance, making it more inclusive; and help the environment by using the planet's resources more sustainably.

This potential for increased well-being as a principal can and should be embedded as a core principal. AI should serve the best interests of humanity and increase the well-being of *all* Australians – individually and collectively. The goals and incentives of AI algorithms should serve and support positive human outcomes.

In designing AI technology to, for example, enhance credit provision, the focus should be on meeting the best interests of consumers, meeting and prioritising responsible lending obligations over profit. AI technology in insurance should lead to consumers choosing an insurance product that best suits their needs, covers their risks and adheres to a genuine consumer/insurer risk mitigation partnership. Consumers should have agency in a transparent process and not simply be used to meet and exceed sales targets.

A cost-benefit analysis empowers corporations to place profit over people. Increasing well-being requires that we all benefit from AI.

Core principle 5 - Fairness

Core principle 5 states:

***Fairness:** The development or use of the AI system must not result in unfair discrimination against individuals, communities or groups. This requires particular attention to ensure the “training data” is free from bias or characteristics which may cause the algorithm to behave unfairly.*

Fairness is a key principle that needs to be built into AI. However, as currently drafted this principle establishes the concept of fairness in the negative – that is, one free from bias. Turning this around to a more proactive principle would ensure that those designing and developing AI will know how to build fairness into AI systems.

We propose that the concept of fairness should be tied to the principle of inclusion. Inclusion is the proactive state of ensuring people feel they belong, can engage, participate and be connected irrespective of race, gender ability or any other attribute that can be perceived as difference.

The design and development of AI should embrace the concept of inclusion, embed the diversity of the Australian community into their work and enable expansive participation of all Australians. AI should not lead to exclusion of different groups of Australians in economic and social life. This means:

- accounting for the diverse and capabilities needs of target user bases in designing and implementing AI;
- ensuring that datasets used are representative of different groups;

- paying specific attention to inequalities and the needs of vulnerable communities;
- including diversity within the people and organisations responsible for the creation and implementation of AI; and
- seeking out a diversity of views from a variety of cohorts (including vulnerable communities) in consumer testing and design.

In the general insurance context, for example, AI should promote accessibility to insurance for all Australians including vulnerable Australians such as those from a cultural or linguistically diverse background, those with physical disabilities, those experiencing financial hardship, those in remote and regional communities, older Australians and those suffering from a mental illness. AI's ability to super-charge risk assessment, segment populations and price discriminate needs to be balanced and tempered by the principle of inclusion. In this way AI can produce better outcomes for all individuals and communities, not just those who won the lottery of life.

Inclusion means treating Australians in all their diversity and varying needs in fair ways. It means removing unfair and exclusionary price discrimination and other forms of discrimination in insurance products including the use of algorithms in selling, pricing and distributing.

Core principle 6 - Transparency & Explainability

Core principle 6 states

Transparency & Explainability. People must be informed when an algorithm is being used that impacts them and they should be provided with information about what information the algorithm uses to make decisions.

Related to the principle of transparency is the principle of trust, which we propose should be included as a core principle.

Trust in an AI-driven digital environment is critical for consumers and general insurers alike. An algorithm or piece of AI technology may be transparent and may be able to be explained to somebody but consumers still need to know that they can trust the process.

Consumers want to know that the sum insured calculator is independent and fair. They want to know that the insurer is highlighting and telling them the information that they need to know. They do not want to be tricked or shocked later down the track. They want to know that the fine print won't get them in the end and that nothing is being hidden from them.

Full transparency can simply lead to some corporations simply providing information overload – confusing and potentially misleading consumers. Integrating the concept of trust ensures that a reliable environment is created – one that builds confidence and promotes authentic engagement.

A toolkit for Ethical AI

Financial Rights supports the development of tools for businesses to use to assist them in implementing the principles of ethical AI into their own practices.

We wish to add one element into the mix that is not mentioned explicitly in the Discussion Paper and that is the use of product/consumer testing. Done well consumer testing of new AI products and services on diverse demographics has the potential to provide businesses with real insights into the impact of AI technologies before they are introduced or implemented. Consumer testing can identify negative impacts upon certain populations before they are committed.

It is vital that financial services design products, processes and service that include algorithmic and AI element for the broad diversity of customers and potential customers. In order to do this, consumer testing for all new products must involve consumers of a variety of backgrounds. Businesses should be obliged to identify a series of vulnerable groups of people and consumer-test all new products with a variety of types of customers. For instance, products should undertake consumer testing specifically to gauge whether or not they will be appropriate for, and whether or not they will be harmful for, low income customers consumers or people experiencing other forms of vulnerability.

An example of poor implementation of a new technology is the Commonwealth Banks introduction of *Albert* machines. Several years ago, Commonwealth bank introduced a new type of EFTPOS terminal with a touch screen – known as *Albert*. This machine was clearly not designed with the needs of some types of people experiencing vulnerability in mind: the machine is close to entirely unusable by blind customers. These customers at the point of purchase are backed into a corner where they can either forego purchasing the items, or disclose their pin code to another person. This is clearly deeply inappropriate. Consumer testing of this product with various demographics of customers could have avoided this.

Similar unintentional issues will arise out of the introduction of untested AI products and services. It is vital that businesses are guided to avoid problems and meet core principles before they arise. Product or consumer testing is the key way to forestall any negative impacts.

Concluding Remarks

Thank you again for the opportunity to comment. If you have any questions or concerns regarding this submission please do not hesitate to contact Financial Rights on (02) 9212 4216.

Kind Regards,



Karen Cox
Chief Executive Officer

Financial Rights Legal Centre
Direct: (02) 8204 1340
E-mail: Karen.Cox@financialrights.org.au