Pearcey Oration

September 4th, 2024

Dr. Ian Oppermann

MBA (Lon), PhD (Syd). FIEEE, FIEAust, FTSE, FACS, FRSN, GAICD

We have a tiger by the tail...

It is an honour to be invited here this evening to offer the 2024 Pearcey Oration.

Let me begin by acknowledging the traditional custodians of the lands on which we stand, the Wurundjeri Woi-wurrung and Bunurong/Boon Wurrung peoples of the Kulin Nation and pay respect to their Elders past and present. Let me extend those respects to any aboriginal or Torres Strait Islanders with us this evening.

This year marks 75 years since the famous CSIRAC digital computer ran its first line of code in 1949. I will ask you to hold on to that thought.

What does it mean to have a tiger by the tail?

Having a tiger by the tail means finding yourself in a situation that's unexpectedly difficult to manage but impossible to abandon without facing significant risks. Imagine holding onto the tail of a wild tiger; you can't let go without danger, and holding on is equally perilous. This idiom is often used to describe scenarios where someone has taken on a challenge or responsibility that turns out to be more formidable than anticipated.

So, what is the tiger in our case?

Well, we can take our pick really. AI is the one people are talking about from a new technology perspective. AI seems to have fallen from the sky in late 2022 and in that small window of time, has caused people to stand amazed at the seemingly humanlike ability to create text, images and even movies from simple prompts. Consider the following prompt:

> *A stylish woman walks down a Tokyo street filled with warm glowing neon and animated city signage. She wears a black leather jacket, a long red dress, and black boots, and carries a black purse. She wears sunglasses and red lipstick. She walks confidently and casually. The street is damp and reflective, creating a mirror effect of the colourful lights. Many pedestrians walk about.*

That slightly long, yet very natural description leads a tool from OpenAI to produce a remarkable video clip of exactly that. Add your favourite brand and you have an advertisement. Add some tunes and you have a music video, add a plot twist and you have the next potential Hollywood, Bollywood, or TikTok blockbuster. No wonder artists, actors and creative industry members are concerned.

But AI can also do a whole lot of good.

My most recent experiences have largely been with government, so let me turn to the potential of AI in government for a moment.

One of the most significant challenges facing governments today is the widening gap between what citizens expect and what the government can deliver— addressing latent demand. This issue is particularly evident in areas like human services, where the need is high, but the capacity to meet that need is always limited.

Health services provide the starkest example.

The total demand for health services is essentially infinite. As our population grows and ages, and as we all live longer with chronic diseases that would have once dispatched us, the strain on healthcare systems only intensifies. Similar scenarios play out in other areas like justice and education, where the gap between what people need and what the system can provide continues to widen.

Health, however, is the big one. Way back in 2010, the Commonwealth Treasury predicted that if left unchecked, health and hospital spending would consume all tax revenues collected directly by all state governments by 2045. That is just over 20 years from now.

While much has been done to address this doomsday scenario, health continues to take up a larger proportion of state budgets each year. In the last NSW budget, Health spending was easily the largest sector with just over $31bn or 25% of the total budget. Education came in second at just under $24bn. This trend scales when you look across the country.

Automation and AI can play a crucial role in addressing this latent demand. By speeding up and joining together process stages, and then making them more efficient, AI can help governments and service providers respond more quickly and effectively to the needs of citizens.

For example, AI-powered tools can analyse vast amounts of data to predict where demand will be highest, allowing governments to allocate resources more efficiently. Automation can also handle routine tasks, freeing up human workers to focus on more complex and value-added activities. This not only helps to bridge the gap between demand and supply but also improves the overall quality of service delivery.

AI powered digital automation can also help governments become more agile and responsive. By leveraging AI and data analytics, governments can quickly adapt to changing circumstances and respond to emerging needs. This is particularly important in today's fast-paced world, where the ability to act quickly can make a significant difference in outcomes in particular in cases of natural or man-made disasters.

Is there a way we can trust our digital tiger?

AI comes with increasing inability to trust our digital world polluted as it is by "fake news".

I am going to use the term "data product" quite a few times. By "data product", I mean anything produced from data whether it be a summary, a chart, an image, a video or a report.

Increasingly, the value of data and data products are impacted by concerns that they have been inappropriately sourced (for example scraped), inappropriately governed (for example access is provided to minors), or inappropriately manipulated (for example, faked or hallucinated content).

When engaging with data products in the digital world, we need to answer important questions about data and products derived from data:

1. How can I determine if data is fit for the purposes I plan to use it for?
2. How can I provide guidance / restrictions / prohibitions for future uses of the products I create from this data?
3. How can I enforce restrictions / prohibitions for future uses of the products I create from data?
4. How can I determine if a data product has been manipulated in ways that I did not expect?

Data and data products are all around us. Once thought of as rows and columns in spreadsheets, or words on paper, data comes in a myriad of forms including biometric sensors, voice recordings and video images. The recent releases of powerful AI tools have changed the scale and scope of the products that can be generated from data. Analytic insights, predictors, classifiers and anomaly detectors, have been augmented with report generators, automatic summary compilers, image generators and many more.

The problem is compounded by the abundance of AI driven “deep fake” technologies along the lines discussed earlier. AI which can generate synthetic images, documents, videos, or audio based on original source data can be used to fool us rather than entertain us.

These technologies have the potential to entertain but have also been used to defraud unsuspecting users of the faked data products. They also have the potential to create great harm. The faked images of an aircraft crashed into the US Pentagon in 2023 led to a real-world impact on the US stock market, or the generation of synthetic pornographic images earlier in 2024 by a school child led to physiological harm of the subjects and others who saw the images. Whilst these examples give an insight into the types of harms not previously imagined by use of such technologies, the scope of possible harms is beyond imagining. The data products can be very convincing, and potentially very damaging, even if eventually identified as synthetically generated fakes.

The challenge becomes:

- the ability to assess how an AI derived data “product” can be appropriately used,
- to create guidance, restrictions or even prohibitions on how data products can be used,
- to create environments to enforce those restrictions and prohibitions, and
- to create systematic ways to identify if data products have been manipulated in ways which are unexpected leading to ways of rapidly limiting the use of those data products.

In all uses (or sharing) of data and data products, the context of use (or sharing) matters. Context includes the nature of the use, who is using, the environment of the use and what happens once the data or data products are used. These multiple dimensions or “degrees of freedom” create a very wide range of possible considerations.

So, let’s talk about the data tiger …

I have talked quite a bit about AI, but data is the lifeblood of the modern economy. It impacts, enables and personalises how we work, play and engage socially and is also crucial for the operation of government and the economy. Banks and financial services companies can be described as data and digital services organisations with some bricks and mortar operations. Value comes from creating, using, protecting and sharing data. Use of data is a very wide and vague topic, incorporating analysis, storage, aggregation, dissemination and deletion.

The dilemma often faced by people who want access to data is how to build a trusted data sharing framework in the absence of one. The question of ‘Can I have access to your data?’ will very often be met with a firm, polite but negative response of ‘No’, often backed by the statement ‘because of the Privacy Act’ – the BOTPA reason. This is particularly true if the data is about people.

At the other end of that sharing or use relationship is “can I trust the data or data product you have tried to share with me?”. When that data product is a chart or an analytical insight, it is easier to apply tests on source of origin, governance and methods of analysis. When the data product is the output of complex AI systems, there is little in the way of frameworks to test the image, video, voice message or other generated result.

Ultimately data sharing and use is an act of trust, and trust is either developed within a trusted relationship or through demonstration of trustworthy capability that encompasses technical and governance capability, as well as authorisation frameworks and clarity of purpose. Data sharing and use is not a single transaction, but parties who share data are a step in what may be a very complex data life cycle. As the number of stages of the life cycle increase, trust between parties becomes increasingly difficult to maintain. Trust between parties can be replaced with controls and scrutiny to ensure appropriate use of data across the stages of the data life cycle.

And then there is the cyber security tiger...

You cannot talk about data sharing and use without talking about cyber security.

Cybersecurity has been an arms race since "cyber" was invented. Open systems can be eaves-dropped on, disrupted or degraded, so we have continuously developed new and smarter ways of protecting our online systems. Every clever new protection puts the defender a step ahead, but also creates the irresistible challenge for attackers to innovate and counter the protection. This was once a contest between human minds with nation states applying their brightest to develop (or to crack) codes and security systems, or of groups trying to defend themselves from determined ideologically minded hackers or sometimes from kids having mischievous fun.

Once you had to be interesting to be a cyber-attack target. You held an important position in society, held valuable data or were doing something that someone else strongly objected to. That is because it took time and real mental effort to find a way past your cyber protections, to achieve the result you wanted, and to not get caught.

It is no longer the case that you need to be interesting to be a target. It is also no longer the case that it is a contest of minds, at least not at the operational level.

AI has long been used to probe and test the defences of systems, as well as to counter such probing and testing. There are famous stories of malware being loaded onto USB sticks which then wreak havoc when inserted into sensitive systems by unsuspecting humans. We have all heard of bot-based denial of service attacks where automated agents flood an online service with requests and overwhelm the server.

Sophisticated "intelligent" anomaly detection has been the basis of cybersecurity systems for a long time. Something unusual in the operation of a part of the system, or the behaviour of an authorised user can trigger a closer inspection or lead to rapidly escalating levels of defence. Increasingly what is "unusual" is identified by AI and alerted to a human, but even then, some quarantine actions may have been carried out by the time the alert is made.

Imagine however that you no longer needed to be interesting, or important, or seriously disliked to be a cyber target. Imagine you just had to be connected to the outside world.

This is of course the world we live in today. With the ability to use AI to easily generate and deploy malware, then every connected device is a target: from your laptop to your smart watch; from your employer to your bank. We will all have noticed the rapid increase in the number of "suspected spam" phone calls and increasingly sophisticated phishing emails. I am embarrassed to say how many I have responded to in the last few months.

The flood of AI generated attacks is arguably making the current way we do things harder and harder to justify. I will not answer a mobile call if I do not have the number stored in my phone (as so presumably I know the caller) so my behaviour has changed. This is not helpful for the genuine person on the other end trying to reach me.

Of course, it does not stop there. The more our world becomes digital and connected, the more avenues we create for AI based malware to approach us, learn about us, and then possibly do bad things to us. We reveal a great deal about ourselves from our digital interactions, and much of it unintentionally. Information such as where we are, when we are there, where we came from, if we were with someone else and much more can all be captured (and so potentially revealed) by our engagement with our connected intelligent devices.

What about a digital identity tiger?

Arguably, the most important thing that can be captured is our identify, either in the form of a real unique identifier (such as a passport number, driver's license or Medicare number), or in the form of a combination of parameters which identify that we are the same person who appeared in some system earlier. Identities enable us to verify who we are to many different systems. They can allow access or can be used to authorise use of resources including funds. Loss or theft of identity is a serious problem.

So surely, identity is one type of data we should really take care to protect. If so, why would we ever build systems or databases which store a whole lot of identities in the same place? Also, why would we be so careless as to allow datasets to come together without a good understanding of whether someone can be reidentified from the connection of these different data elements?

The answer to the first question is that this is how we have always done things. There is a mindset which says, 'put it all in one place and protect it'. This centralisation approach does allow a focus for the protections which are applied but unfortunately, does create very attractive datasets for people (and bots) to try to crack in to. We have seen data breach after data breach arise from this way of doing things.

The answer to the second question is that we do not understand what "identify" really means in the online world. We are unable to answer the question of when someone is "reasonably identifiable" when we have a fragment of their pattern of life data. The risk is always that an attacker has some other data fragments which makes an individual identifiable and given the number of data breaches we have already experienced which involve personal information, this is a real possibility. Unfortunately, the consequence of this challenge is that many datasets are over protected (access and sharing not allowed) or under protected (risk not fully understood).

So, what is to be done?

One thing for sure it that we need standards for data sharing and use, and standards for what "reasonably identifiable" means.

For nearly a decade, I worked within the NSW government seeking to develop frameworks for general data sharing within and across government based on experience and expertise from the research community, state and commonwealth government colleagues, industry, and the world of international standards.

Year after year, the good folks at the Australian Computer Society (ACS) supported workshops digging into the "why" of data sharing challenges, some years producing whitepapers which tackled the problem at the next level of complexity and the next level of specificity. This led to a total of 5 technical whitepapers, the development of a simple tool for measuring the amount of personal information in linked, "deidentified", people centred datasets, the "PIF" (Personal Information Factor), and a set of frameworks for generalised data sharing which allowed domain specific sensitivities to be considered.

The PIF tool was and is used by NSW in numerous open data situations. Importantly, it was used every single day for almost 3 years to assess the level of personal information in deidentified, but unit-record level, COVID case data before that data was released to the public. If the data products created from the raw data were determined to be appropriately protected, then they were released. If not, greater levels of protections were applied.

All this work, and a lot of late-night meetings, ultimately led to the creation of a pair of international data standards, both of which were published in April this year:

- ISO/IEC 5207:2024 (Terminology and use cases)
- ISO/IEC 5212:204 (Guidance for data use).

Those names simply roll off the tongue, don't they?

A next target for standardisation is the measure of personal information. The PIF shows that something can be done, but it is certainly not sufficient in its current form.

The second major challenge is to stop linking datasets and storing data in one "honey pot" in the first place, especially those datasets with identity information. Data fabrics supported by data virtualisation have become increasingly more sophisticated, ironically perhaps, aided by use of AI and process automation. It is increasingly possible to only ever build data overlays or connect through metadata without the underlaying data needed to be shipped across and joined in a database. The metadata can also be assessed to do some sort of PIF assessment without the data ever coming together.

Ultimately however, we need to give identity back to individual people and change our thinking about identity from "prove to me that you are this unique person, and then I will access information about you" to "do you have a valid credential or licence to access this asset, product or service". The more we can move to asking questions of an individual's carefully protected data, rather than asking to see the underlying data itself, the more we can ease friction in the digital world.

That leaves us with the big issue of individuals now needing to defend themselves from cyber-attacks designed to access or misuse their personal data. Again, AI can help here. All of those efforts to detect unauthorised access or detect anomalous behaviours can now be focussed on the behaviours of one individual who provides access to others for a range of known (or knowable) purposes. How and to whom they grant access to can be templated, the type of questions which can be asked of the data can be strictly controlled, the entire history of access and questions asked can be assessed for possible reidentification risks, and the behaviour of the user used to personalise protection services. There is good work happening in this space but still much to be done.

Let's Look a little further out.

I would like to reframe a little now, and talk about the world of 2030. That is the timeframe for the realisation of the UN Sustainable Development goals. These goals describe how we want the world to look in 2030.

We once talked about 2030 as the distant future. For any one of the 800 or so children born in Australia on this day today, they will turn just 6 during 2030.

However, given the rapid pace of technological change, a lot will still happen in that time. In the meantime, many parts of government in Australia are struggling to engage with AI effectively and to reap its benefits.

As stated earlier, for many, AI seems to have suddenly fallen from the sky with the advent of ChatGPT, generative AI and large language models. However, AI has been around for a long time. Initially, AI was something that could do a narrow task very well: adapting over time, filtering information and improving over time with a specific task. Recently, it has been applied to a much wider range of areas. The most successful applications we have seen so far focus on addressing the complexity of systems, joining up disconnected processes, navigating policies, and summarizing documents.

Over the years to 2030, it is inevitable that the public will increasingly demand that government do things differently. As services in every other sector become increasingly personalized, joined-up and responsive to individual needs, it is inevitable that the public will expect, even demand this of government. This is an expectation that AI can help with. The uses of data, digital channels and AI that help companies engage with customers can also help governments treat the public in a much more customer-centric way and reflect the "life journeys" of individuals as opposed to forcing people to face the siloed structures reflecting individual government agencies. Creating seamless interactions with government around "life journeys" or significant events such as the birth of a child, the death of a loved one, transitioning to retirement, or even just getting a job are all areas where government services can be delivered in a customer-centric way.

There are many such life journeys and significant events, but not all of them are ones we choose or have the option of engaging with government on. Some areas of government, like justice, are impossible to avoid — you can't opt out if you're arrested, fined, or summoned to court. However, these areas could also benefit remarkably from use of data, digital channels and AI to create more customer-centric experiences.

If government services don't meet public expectations, people seek alternatives. The regulated taxi system was completely disrupted by Uber offering a different type of ridesharing service. Governments around the world then played catch-up trying to adapt the taxi model to this new approach. Similar disruptions happened in other sections of the “Gig economy” with governments everywhere caught off guard. Disruption could and will happen in other areas and AI could both help cause the disruption and offer solutions.

As we approach 2030, it’s important for us to have a view on what we want the world to look like – those outcomes - and then work backward to determine the steps needed to achieve that vision.

Imagine a government that allows for seamless interactions - where engaging with those life journeys are as simple as booking a hotel online. AI should eliminate the complexity and friction in engaging with government and increasingly integrating different government functions. AI can allow government to look like it really is built around the unique needs of every single one of us. This is a join-up view of government which is also cybersecure, privacy preserving, and does not over-reach in terms of just how personal governments get with each of us. After all, who wants a highly personalized infringement notice?

However, there are many other applications where governments can explore using AI to deliver “old” services in new ways. Recently, NSW announced the use of AI-driven vision software to detect if passengers are wearing seat belts. This application enforces an existing law in a novel way, by examining every car passing under the camera to check if passengers are complying. The AI recommends images it “believes” is showing infringing behaviour to a human for final decision on issuing a fine. This builds on the successful roll out of AI used to detect drivers using mobile phones and similarly, making recommendations to human operators to issue fines.

The challenge is to strike a balance along a number of dimensions. We need to ensure we do not inadvertently create a surveillance state and that, ultimately, people are still making the decision about the course of action even if recommended by an algorithm. It also requires the impersonal nature of due process to balance with the “personalised” nature of joined-up government. We may well choose to have barriers between what government knows about us when it comes to health and social services, and what government knows about us from a tax and justice perspective.

Recently, I heard about AI being used in another country to automatically issue fines when someone drives into a no-stopping zone. While this is arguably an efficient closed-loop use of data and AI, but it seems ruthlessly unforgiving and heavy handed. We need to carefully consider the implications of such surveillance and enforcement of legislation. Currently, enforcement involves a lot of human factors such as judgement, awareness, and presence. How do we balance efficiency with these human elements? Keeping our eyes on that vision for 2030 and the details underpinning the headline outcomes we are striving for.

Looking ahead to 2030, I hope AI will handle the mundane tasks, making systems more cohesive and addressing any inconsistencies or gaps. Ideally, AI will highlight issues for humans to address. I also hope that while governments use AI to fulfil their roles, they do so with real exercise of judgement and using frameworks of assurance, ensuring laws are applied and services delivered without creating a dystopian surveillance state.

We are doing some things right: Consistent Standards and Domestic Collaboration

We’re going to arrive at 2030 either way. On the positive side, we’re making progress in ensuring consistency at a national level regarding AI assurance. The Australian Government has recently introduced the Australian AI Assurance Framework, which builds on the NSW AI Assurance Framework. The NSW framework provides risk-based principles that guide how AI should be managed. Consistency across states is crucial — having different systems in South Australia compared to NSW and Victoria would be problematic.

Additionally, it’s encouraging that standards for data sharing and AI management are being developed. These standards are essential for setting clear definitions, and consistent guidelines and expectations as we move forward with AI technologies.

Where are we going with our tigers?

Let’s look at road mapping the next five-and-a-half years.

To reach the level of seamlessness outlined in the 2030 vision, we need to start by simplifying the complex web of government interactions. Currently, navigating government is challenging due to the sheer volume of information and the three layers of government in Australia, which can create an "embarrassment of choices." For instance, the housing crisis in Australia highlights this complexity: different local governments, state regulations, and Commonwealth considerations all contribute to a convoluted process for housing development approvals.

To improve, we need to focus on several key areas:

> **Information Access:** First, we need to make finding and understanding government rules and regulations easier. This means simplifying access to information and creating a more user-friendly way to engage with it. This is not “plain English” websites, this is personalised navigation tools.
>
> **Data Integration:** Next, we should aim to integrate and digitize data across various government levels. For example, in the housing sector, having a digital system that allows users to visualize their projects, engage in an intuitive way with planning requirements, automatically submit for approval, and track progress would streamline the entire process.
>
> **Reducing Complexity for the end user:** Finally, addressing the friction points in current systems — where rules and processes create confusion — will be crucial. This involves digitizing rules and processes, connecting them seamlessly, and identifying and eliminating points of friction or "sludge."

By focusing on these areas, we can transform government interactions to be as intuitive and efficient as possible, making it easier for people to navigate and engage with government services. The goal is to create a future where the process is clear and straightforward, eliminating unnecessary complexity and ensuring that everyone knows exactly what steps to take next.

And the future?

I very briefly referenced 2030 using the metaphor of a child born today, and used this to explore a little of what that world will look like and some of the aspects that we would like to influence in terms of outcomes.

This child’s future is digital, ubiquitously connected and critically dependent on technology. The child is also likely to live well past 2030 and even into the 22nd century. What a thought on the 75th anniversary year of CSIRAC.

As technology and digital solutions continue to play a key role in driving the economy and society forward, they become increasingly embedded into business operations, across key service offerings and into our personal lives.

By 2030, it will have become a self-reinforcing process which is being accelerated by increased use of AI to make sense of the rising tide of data, to continue to locally optimise services delivery and to increasingly personalise.

By 2030, our complete dependence on technology makes cyber security crucial to navigating the associated risks and opportunities ahead. Combined with the evolving complexity and sophistication of cyber security threats, together increase our level of vulnerability – at a national, organisational and individual level.

Future services have the potential to deliver enormous benefits, however their very nature highlights challenges when contemplated within existing regulatory frameworks. As systems develop, the management of privacy and consent when collecting, sharing and using these datasets will need to be considered alongside the technological capability. New methods for providing and handling consent, new frameworks for sharing and using data, and new considerations for security in highly complex networks will need to be considered.

As we reach 2040, our child born today will have grown into their teens. This world will hopefully still be shaped by the work we are doing today. As we reach 2050, they will be shaping their present and future. Data and AI will certainly still be playing a role. Hopefully also, the outcomes lens will still be firmly fixed. As they reach the 75th anniversary of tonight's address and the 150th year since CSIRAC first ran a line of code in 2099, they may well look back on our time and wonder…

What were we doing with our tigers?