phn
NORTH WESTERN
MELBOURNE
An Australian Government Initiative

Emerging business trends in primary care: AI, confidentiality and privacy

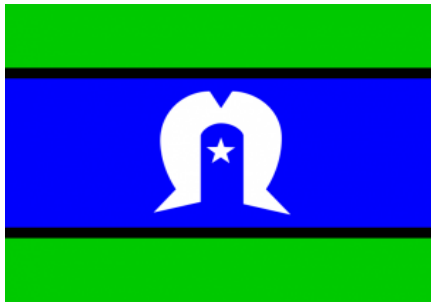
Wednesday 19th February 2025

The content in this session is valid at date of presentation

Acknowledgement of Country

North Western Melbourne Primary Health Network would like to acknowledge the Traditional Custodians of the land on which our work takes place, The Wurundjeri Woi Wurrung People, The Boon Wurrung People and The Wathaurong People.

We pay respects to Elders past, present and emerging as well as pay respects to any Aboriginal and Torres Strait Islander people in the session with us today.



Housekeeping – Zoom Webinar

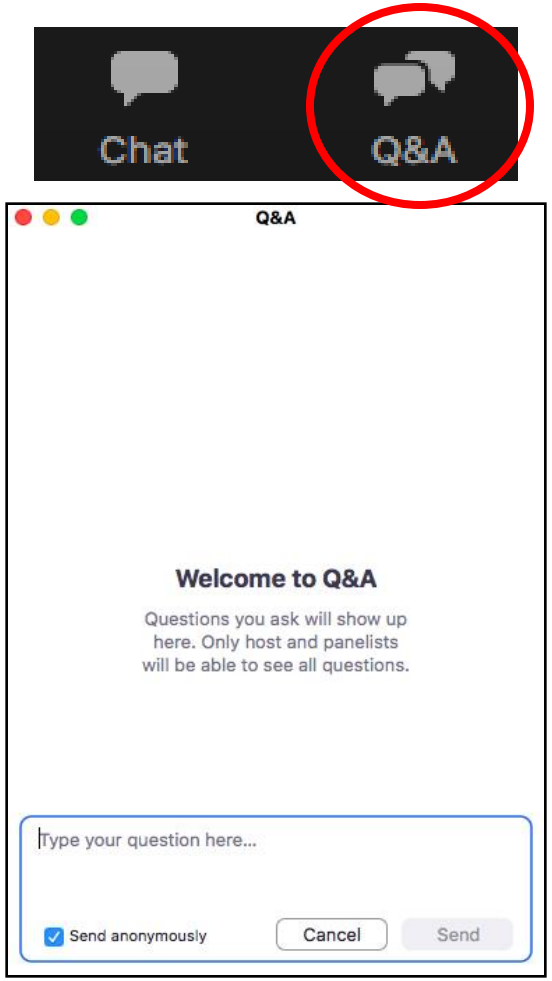
All attendees are muted

Please ask questions via the Q&A box only

Q&A will be at the end of the presentation

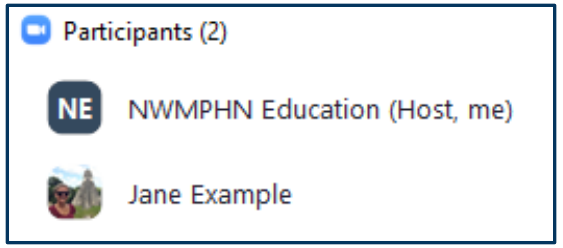
This session is being recorded, you will receive a link to this recording and copy of slides in post session correspondence.

Questions will be asked anonymously to protect your privacy



Housekeeping – Zoom Webinar

Please ensure you have joined the session using the same name as your event registration (or phone number, if you have dialled in)



NWMPHN uses Zoom's participant list to mark attendance and certificates and CPD will not be issued if we cannot confirm your attendance.

If you are not sure if your name matches, please send a Chat message to 'NWMPHN Education' to identify yourself.

Speaker

Miroslav Doncevic

Miroslav Doncevic is the Managing Director of Digital Medical Systems, which has provided Infotech services and support to Australian health care clinics since 1990.

He has partnered with peak health care organisations such as PHNs, RACGP, AAPM, and ASCD for many years, providing IT and cybersecurity education to doctors and practice managers through webinars and articles on key topics.

He completed a Master of Cyber Security in 2024, holds a Graduate Certificate in Cyber Security, and is a Certified NIST CSF Practitioner.

Implementing Generative AI in Primary Healthcare: Fundamentals

Miroslav Doncevic MCyberSec, Grad Cert Cyber Sec, Cert NIST CSF Practitioner
Managing Director

DMS
DIGITAL MEDICAL SYSTEMS
CONFIDENT IT

www.dms-it.com.au

19 February 2025

Key learning objectives:

Generative AI promises to transform, and turbo charge clinical and administrative workflows with significant productivity benefits. Gen AI medical "scribe, dictation, and transcription" tools are increasingly used by Australian clinicians. AI tools for administrative workflows are also becoming more available, for example, Microsoft Copilot M365.

There are very real risks in using Generative AI that must be mitigated, including; exposing sensitive data to privacy violations, data leakage, intellectual property loss, employee misuse, AI hallucinations, inaccuracies, and other unintended consequences, all with significant legal, ethical, privacy, and cyber security risks.

Gen AI in private healthcare practice must be implemented and used in alignment with AI data governance risk and compliance guidelines, which includes preparation and classification of data, securing sensitive data, optimizing data management, and ensuring stringent user role based access controls and permissions to avoid the pitfalls. These essential AI governance implementation steps and guardrails are required to mitigate these risks for safe and secure Gen AI use.

Part One Introduction to Generative AI

1. What is Generative AI and how does it work?
2. Case studies of Gen AI in healthcare
3. Privacy and cyber security obligations in healthcare, and the risks of Gen AI
4. Essential guidelines for safe and secure use of Gen AI

Part Two Practical Productivity with Microsoft Copilots for Admin

5. Introduction to Microsoft M365 Copilot and demonstrations

The only constant is *change*:

” The illiterate of the 21st century will not be those who cannot read and write, but those who cannot learn, unlearn, and relearn.”

Alvin Toffler,

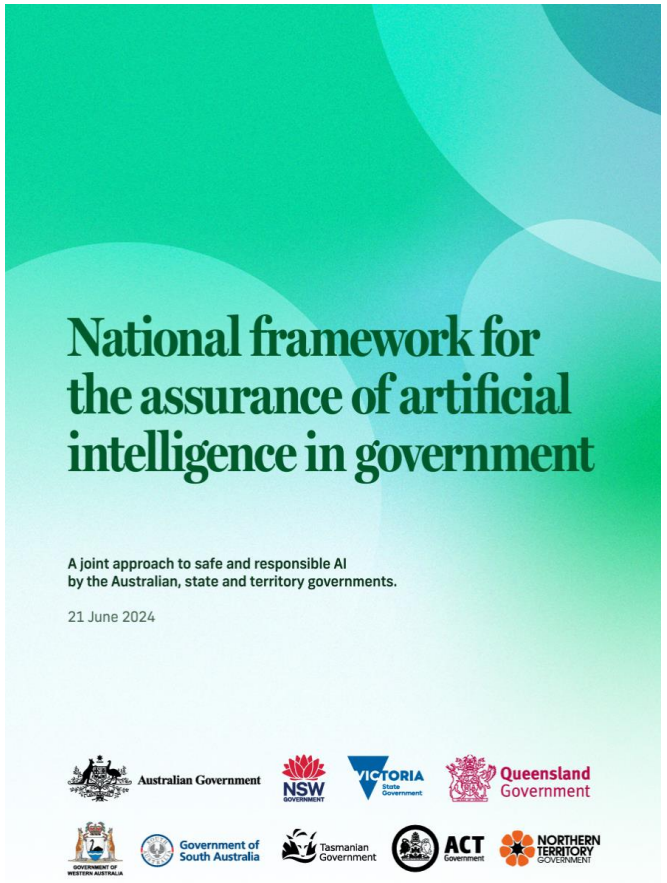
Future Shock

A definition of Generative AI:

”Generative AI is a type of AI that can recognize, summarize, translate, predict, and generate text and other content based on knowledge gained from large datasets.”

Source: https://www.ama-assn.org/system/files/ama-ai-principles.pdf

What is an AI system?



What is an AI system?

In November 2023, OECD member countries approved this revised definition of an AI system:

'A machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.'

Source: https://www.finance.gov.au/government/public-data/data-and-digital-ministers-meeting/national-framework-assurance-artificial-intelligence-government

What is an AI system?

"Artificial intelligence in healthcare

AI can be defined as 'computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision making, and translation between languages'[1].

Some AI tools available for health practitioners are designed specifically for healthcare and have been developed for a therapeutic purpose, for example, to diagnose and treat patients or clients. Many more are general purpose and are being applied in a healthcare setting. Some professions are increasingly using new AI such as medical scribing tools to support workload management and efficiency in practice to develop or edit documents.

There are different types of AI including machine learning which encompasses generative AI, natural language processing and computer vision"…

Source: https://www.ahpra.gov.au/Resources/Artificial-Intelligence-in-healthcare.aspx

The promise of Gen AI:

”The development of AI is as fundamental as the creation of the microprocessor, the personal computer, the Internet, and the mobile phone. It will change the way people work, learn, travel, get health care, and communicate with each other. Entire industries will reorient around it. Businesses will distinguish themselves by how well they use it.”

Gates, B., (2023). The Age of AI has begun

Source: https://www.gatesnotes.com/The-Age-of-AI-Has-Begun

The promise of Gen AI:

Bill Gates

”You’ll simply tell your device, in everyday language, what you want to do. And depending on how much information you choose to share with it, the software will be able to respond personally because it will have a rich understanding of your life.

In the near future, anyone who’s online will be able to have a personal assistant powered by artificial intelligence that’s far beyond today’s technology.

This type of software – something that responds to natural language and can accomplish many different tasks based on its knowledge of the user is called an agent.

…Agents are not only going to change how everyone interacts with computers. They’re also going to upend the software industry, bringing about the biggest revolution in computing since we went from typing commands to tapping on icons.”

Gates, B., (2023). Gates, B., (2023), AI is about to completely change how you use computers

Source: https://www.gatesnotes.com/AI-agents

The promise of Gen AI:

Credit: Shutterstock

“AI is likely to have just as profound an impact as electricity. As AI becomes embedded in devices, tools and systems, it becomes invisible to us.”

Dr Ian Oppermann
NSW Government’s Chief Data Scientist and an Industry Professor at the University of Technology Sydney (UTS)

Source: https://thepolicymaker.jmi.org.au/the-ai-genie-is-out-of-the-bottle-heres-how-to-regulate-it/

The promise of Gen AI:

Credit: Shutterstock

“What we are playing with is like using electricity – don’t put a fork in the toaster, socket…

Understanding data driven tools is critical: a chatbot is not a web page!”

Dr Ian Oppermann
NSW Government’s Chief Data Scientist and an Industry Professor at the University of Technology Sydney (UTS)

Source: AI in Healthcare presentation at Australian Health Care Week Conference Sydney, 20 March 2024

The promise of Gen AI:

Credit: Shutterstock

“For any decision which matters, there must always be an empowered, capable and responsible human in the loop ultimately making that decision.”

Dr Ian Oppermann
NSW Government’s Chief Data Scientist and an Industry Professor at the University of Technology Sydney (UTS)

Source: https://thepolicymaker.jmi.org.au/the-ai-genie-is-out-of-the-bottle-heres-how-to-regulate-it/

The promise of Gen AI:

Credit: Shutterstock

“AI represents a meaningful new tool that can help unlock a piece of the unrealised $1 trillion of improvement potential present in the [healthcare] industry.

It can do so by automating tedious and error-prone operational work, bringing years of clinical data to a clinician’s fingertips in seconds, and by modernising health system infrastructure.”

Source: McKinsey & CO., (2023)., Tackling healthcare’s biggest burdens with generative AI

Source: https://www.mckinsey.com/industries/ healthcare/our-insights/tackling-healthcares- biggest-burdens-with-generative-ai

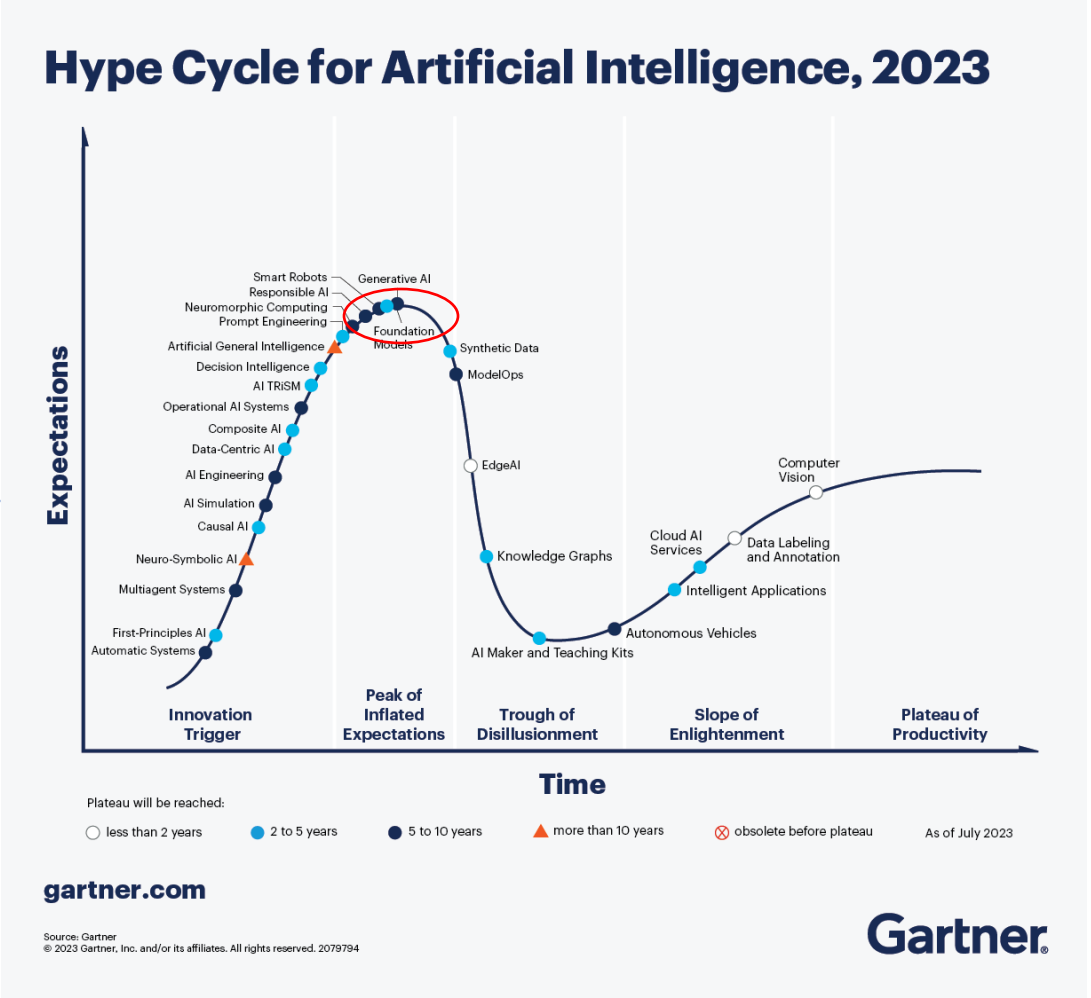
AI Concepts

Generative AI:

The AI referred to here is more specifically Artificial General Intelligence, or more commonly called Generative AI (GAI), and it's component parts;

- Machine Learning (ML)
- Large Language Models (LLMs)
- Human Language AI
- Deep Learning, powered by Neural Networks.

Generative AI is no longer some futuristic tech development, Gen AI is here.



Gartner Hype Cycle for AI (Source: https://www.gartner.com/en/articles/what-s-new-in-artificial-intelligence-from-the-2023-gartner-hype-cycle)

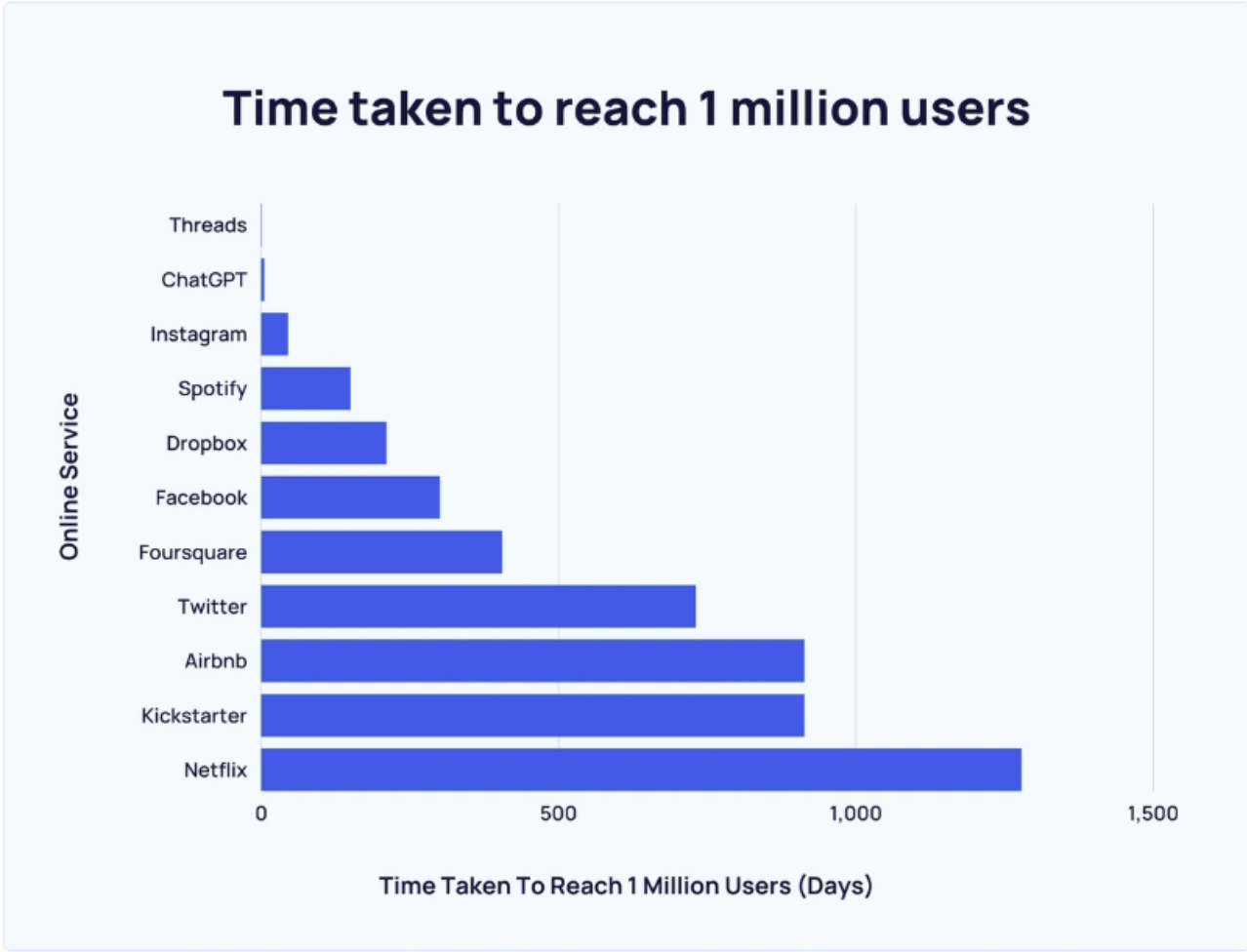
Generative AI: ChatGPT by OpenAI

The release of ChatGPT 3.5 in November 2022 with natural language processing using Deep Learning through transformer neural networks trained on Large Language Models (LLMs), has spearheaded the rapid release of chatbots such as Microsoft Bing/Copilot, Google Bard/Gemini, and many other Gen AI tools with more to come…

Generative AI: ChatGPT by OpenAI

Top 5 ChatGPT User Statistics (Sept 2024)

- ChatGPT currently has over 180 million users
- In just 5 days, ChatGPT surpassed 1 million users
- ChatGPT gets approximately between 207-600 million visits per month
- Around 65% of ChatGPT's social media traffic come via YouTube
- Around 12% of ChatGPT's users are American



Source: https://explodingtopics.com/blog/chatgpt-users

Generative AI: ChatGPT by OpenAI

The GPT stands for *"**G**enerative **P**re-trained **T**ransformer"*

Generative: It means that the model has the ability to ***generate*** text or other forms of output. In the case of ChatGPT, it can generate human-like responses to prompts or questions.

Pre-trained: Before being used for specific tasks, such as answering questions in a chat-based format, the model is trained on a large dataset. This training process helps the model learn patterns, grammar, and context from the text it is exposed to.

Transformer: The transformer is a type of neural network architecture that plays a key role in GPT models. It enables the model to understand the relationships and dependencies between words in a piece of text. Transformers have been successful in various natural language processing tasks, including language translation, summarization, and question answering.

Source: https://community.clari.com/ai-chatgpt-and-revai-83/what-does-the-gpt-mean-in-chatgpt-692

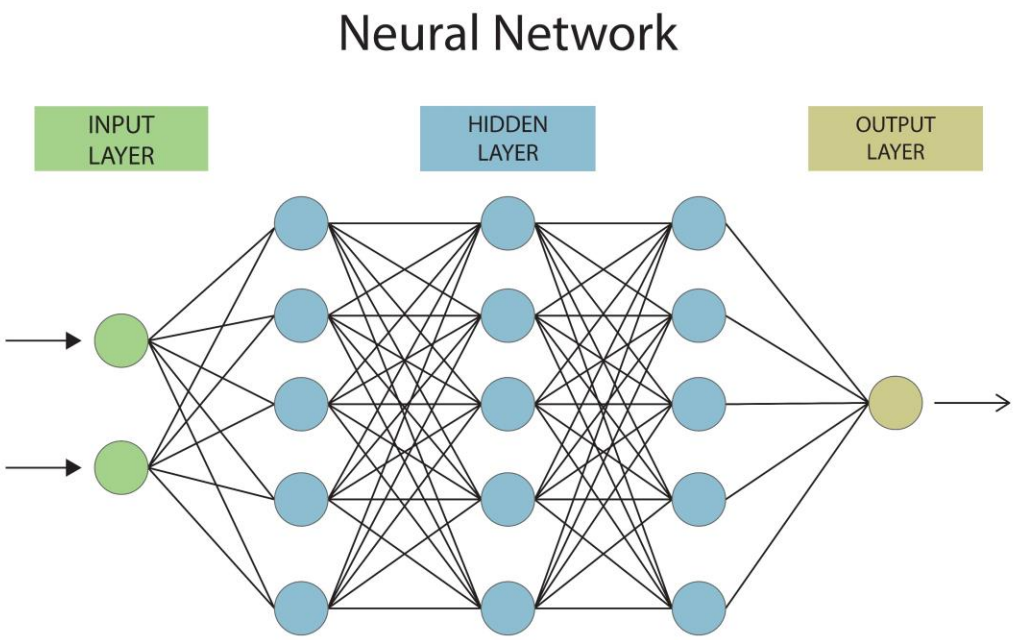
Generative AI:

The AI referred to here is more specifically Artificial General Intelligence, or more commonly called Generative AI (GAI), and it's component parts;

- **Machine Learning (ML)**
- **Large Language Models (LLMs)**
- **Human Language AI**
- **Deep Learning, powered by Neural Networks.**

"the essence of an artificial neuron is nothing but this simple equation from elementary school, Z(X)=W*X + B, where **x** is the input, **w** is a weight, **b** is a bias term and the result or output is Z(x). This allows the AI system to map the input value **x** to some preferred output value **Z**(x)."

Source: https://arvinash.com/how-the-brain-of-an-ai-works-shockingly-simple-but-genius/



Generative AI:

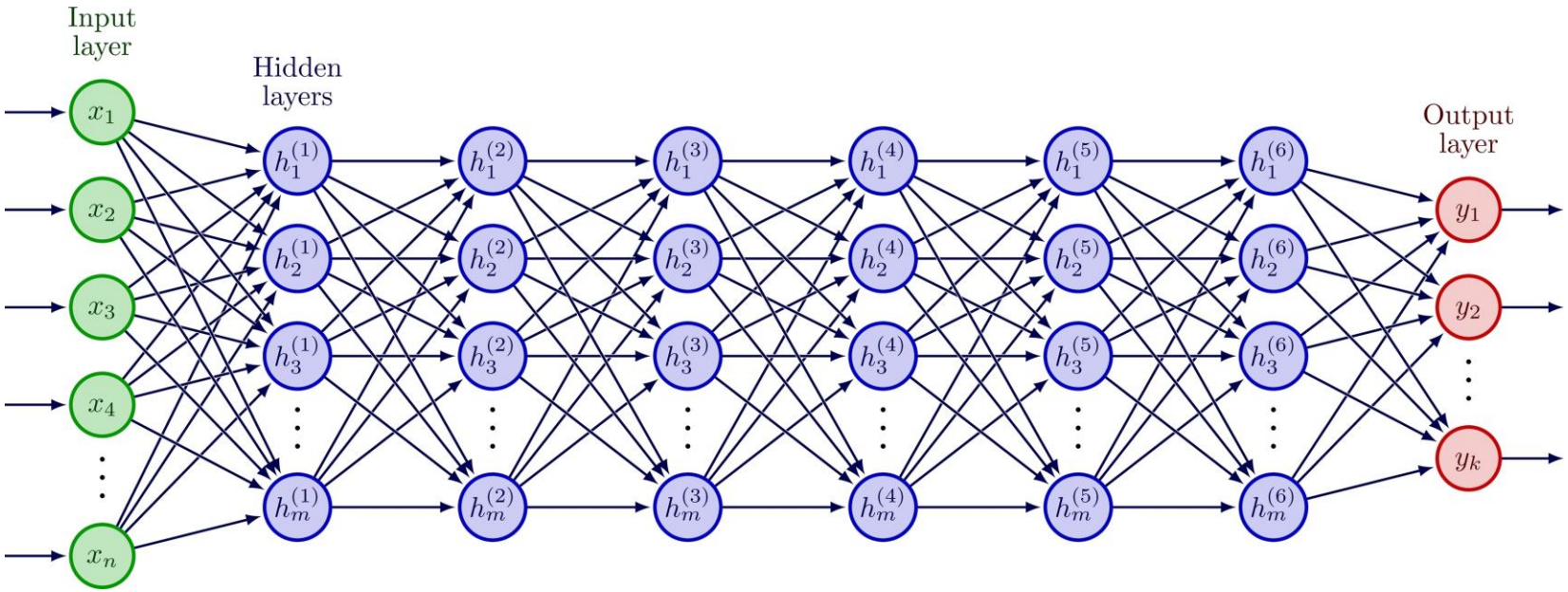
$$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))]$$

V(D,G) = Value function (just trying to maximize this)
D(x) = Probability x is real data
G(z) = Generated data
x = real data
z = noise
Note - Tildas just represent the variable's distribution.

Source: https://medium.com/@randyip9/the-math-behind-generative-ai-for-software-engineers-1-must-know-formula-d2a5db9269a8

Generative AI (GAI), and it's component parts:

- **Machine Learning (ML)**
- **Large Language Models (LLMs)**
- **Human Language AI**
- **Deep Learning, powered by Neural Networks.**



Credit: Shutterstock

What's a LLM

LLM, Large language model

Trained on gigantic corpus of data, billions of parameters

Concepts

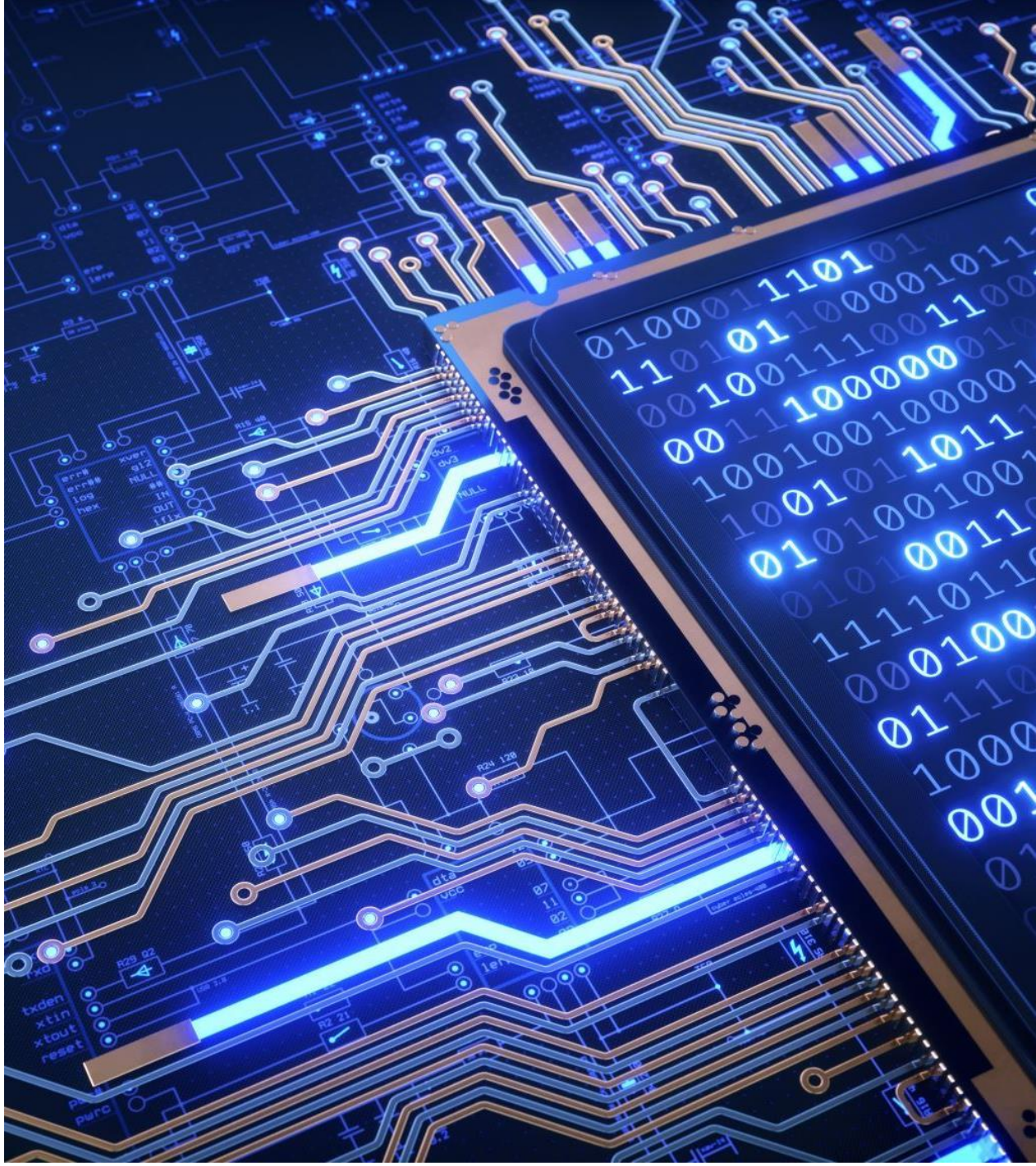
- Transformers
- Parameters
- Nodes

"Computes result" over "looking up facts"

Non-deterministic, even if you can control it somewhat

Input, uses natural language to interact

Understands many spoken languages



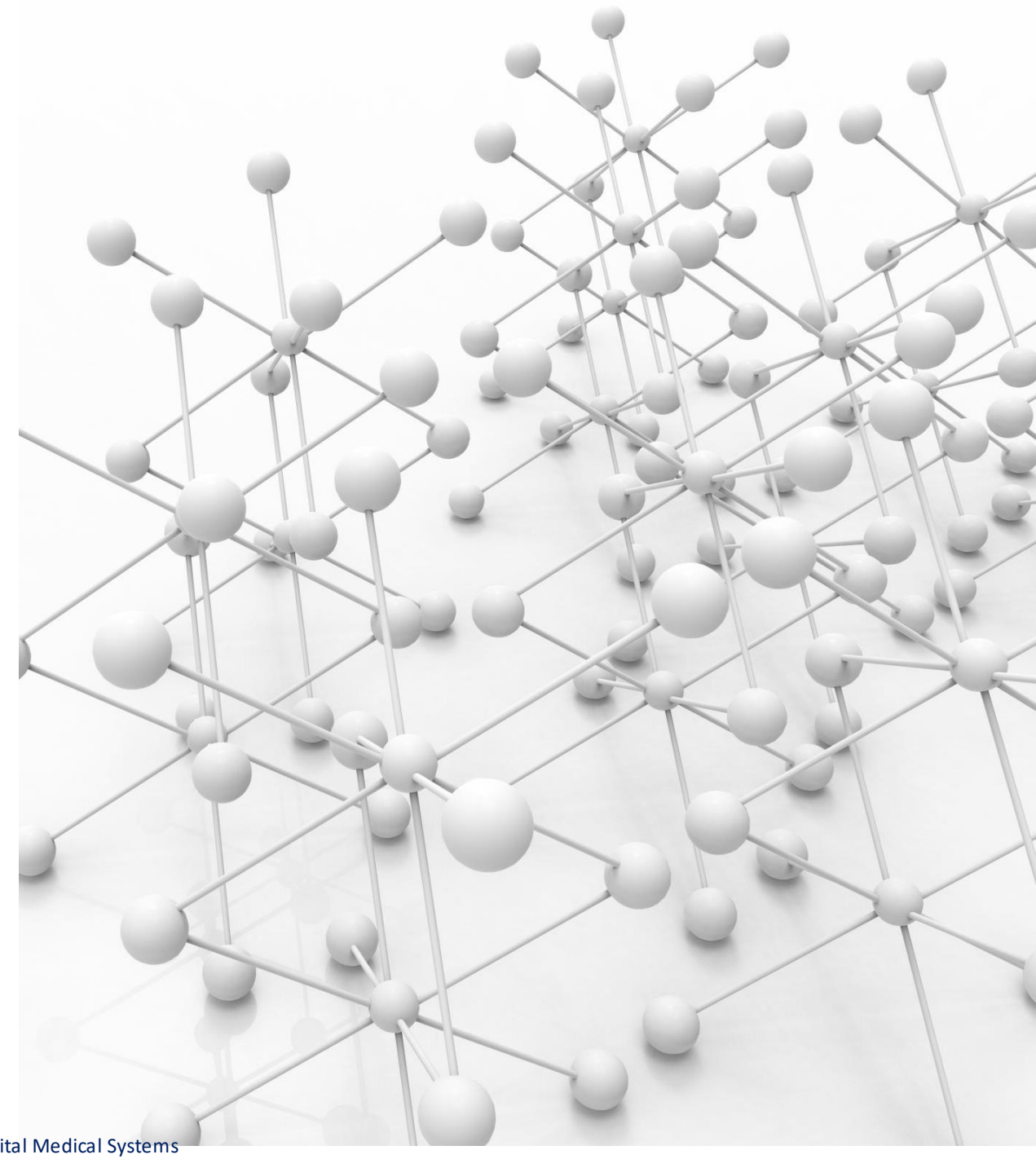
LLM, parameters

Based on neural networks, layer of Nodes

A model that predicts the next word in a sequence of words

Parameters

- Each Node has a weight (parameter), adjusts during training process
- Bias (parameter) also used
- Parameters = rough indicator of size and complexity
- More parameters = can learn more complex patterns



LLM model types

BERT, RoBERTa, RoBERT, ALBERTA

T5

XLNet

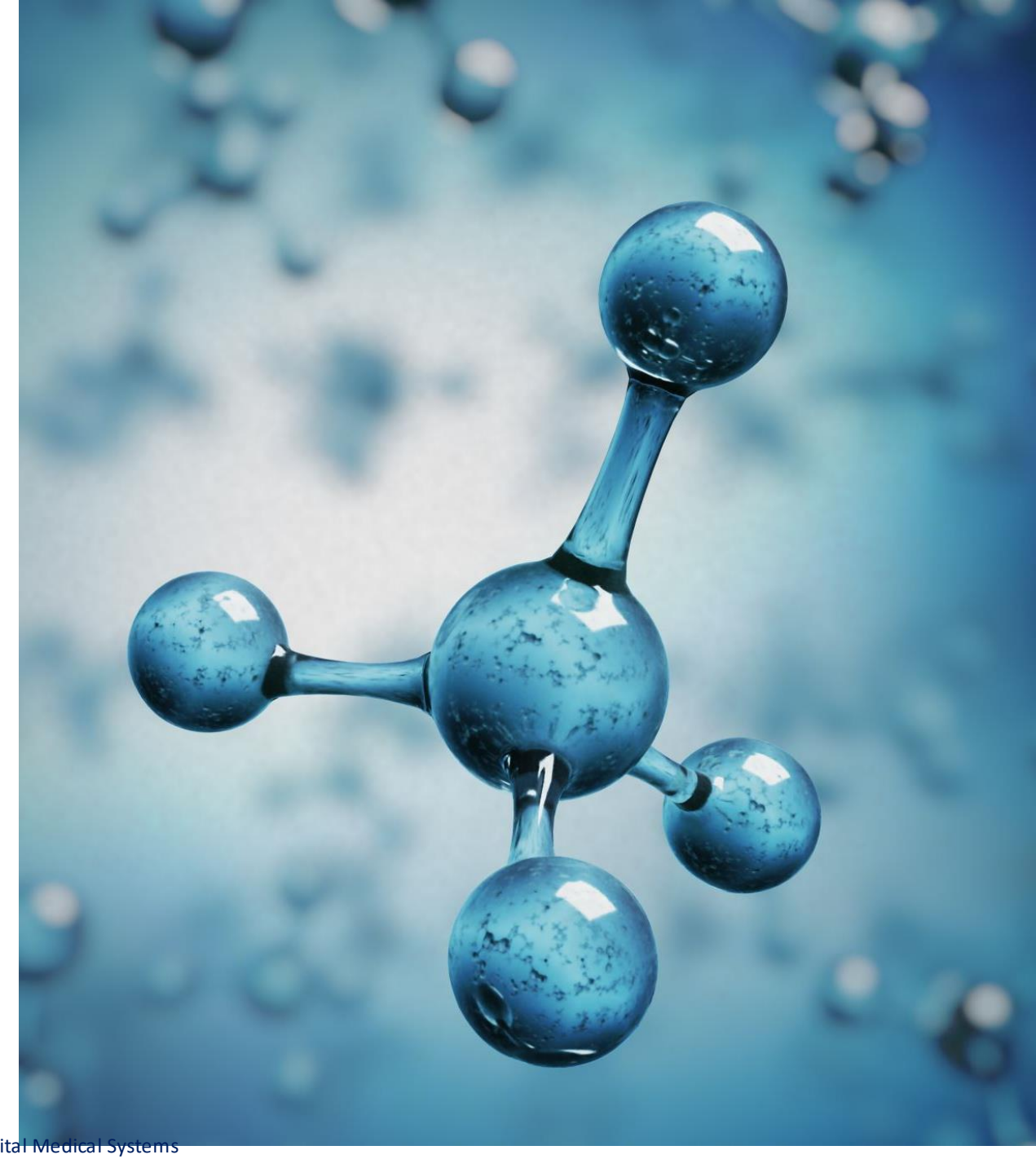
ELECTRA

LLAMA by Meta

GPT, Generative Pretrained Transformer

GPT series

- GPT-1, 117 million parameters
- GPT-2, 1.5 billion parameters
- GPT-3, 175 billion parameters
- GPT-4 176 trillion parameters



Prompts

The subtleties of prompting (prompt design)

Prompt Copy

```
Please give me the history of humans.
```

Prompt Copy

```
Please give me the history of humans in 3 sentences.
```

Why is prompt engineering important?

Prompt engineering is a critical skill for anyone working with LLM AI models

A good prompt produces the desired outputs

Improvements you can perform to make the results more predictable:

- **Specificity**
- **Structure**
- **Examples**
- **Tell the AI what to do**
- **Context**
- **Message Roles**
- **Words of Encouragement**

Gen AI – Healthcare Case Studies

Gen AI - hospital case study

Microsoft News

Taiwan hospital deploys AI copilots to lighten workloads for doctors, nurses and pharmacists

By Chen May Yee

“TAINAN, Taiwan – Since April, the pharmacists have been getting some help from a generative AI assistant, or copilot, built with Microsoft’s Azure OpenAI Service.

…”One click on the A+ Pharmacy copilot button on a screen brings up a patient’s clinical information, summarized from multiple databases on a single interface – medication lists, surgical records, allergy history, lab tests as well as nursing, medical and surgical records, along with a patient’s ID number, bed number and diagnosis.

One tab flags dangerous drug interactions. A pharmacist can also click on a photograph of a particular medication to see if it’s covered by insurance before prescribing the drug.

“**The design logic fits how pharmacists work**,” said department head Hui-Chen Su.

Su said **time saved with the copilot means one pharmacist can now see 30 patients a day, up from 15. It also “allows pharmacists more time to care for patients with complex needs**,” she added.”

Source: https://https://news.microsoft.com/source/asia/features/taiwan-hospital-deploys-ai-copilots-to-lighten-workloads-for-doctors-nurses-and-pharmacists/?OCID=lock1/

Gen AI: - hospital case study

Microsoft News

Taiwan hospital deploys AI copilots to lighten workloads for doctors, nurses and pharmacists

By Chen May Yee

…”It’s just one example of how hospitals are starting to use generative AI to help Taiwan’s chronically overstretched health care workers do their jobs particularly in imaging.

Source: https://https://news.microsoft.com/source/asia/features/taiwan-hospital-deploys-ai-copilots-to-lighten-workloads-for-doctors-nurses-and-pharmacists/?OCID=lock1/

AI in medical settings has been around for a few years,

This new wave aims to cut down mountains of paperwork – whether it’s retrieving information from multiple internal databases, summarizing and generating medical reports or providing patient education materials.”…

Gen AI: - AI Scribes

PULSE+IT

Share

Scribes pave the way for AI acceptance in healthcare

22 January 2025

By Reesh Lyon

“The use of AI scribes by GPs has the potential to pave the way for wider public acceptance of AI in healthcare, according to a report into the use of AI scribes in Australian primary care settings…

…The AI Medical Scribes in Primary Care report was published by Tobias strategic design consultancy and authored by social robotics expert Dr Meg Tonkin and registered nurse Nicole Jess, a strategic designer with an interest in consumer care and clinical workflows.

The report says AI scribes are among the first AI tools to gain traction in Australian healthcare, and predicts they will play a “crucial role” in establishing trust and “paving the way for broader AI acceptance in the healthcare space.”

While AI scribes were becoming widespread in Australian primary care, experiences with their implementation, use and potential benefits were still largely undocumented.

The report looked at a number of points of contact between users (GPs and patients) and the AI scribe, including pre-appointment, arrival at appointment, beginning appointment, during appointment and concluding appointment. There was also a detailed section on the importance of informed consent. ”…

Source: https://www.pulseit.news/Australian-digital-health/Scribes-pave-the-way-for-AI-acceptance-in-healthcare/

Gen AI: - AI Scribes

Home › News › Technology › Software

Digital Nation

By Kate Weber
Feb 18 2025 6:39AM

Ramsay Health Care pilots AI-powered clinical documentation tool

Reduces time spent on manual notetaking in clinical settings.

Ramsay Health Care is piloting an AI-powered clinical documentation tool titled ‘Ramsay Scribe’ Ramsay Health Care pilots AI-powered clinical documentation tool

The tool is designed to enable reduce the time clinicians spend on manual notetaking.

“This improves efficiency, supports compliance with documentation standards, and most importantly, enhances the patient experience by allowing clinicians to dedicate more attention to direct care," Ramsay’s group chief data and digital officer Dr Rachna Gandhi told iTnews.

“Additionally, it provides more structured and complete patient records, which can contribute to better decision-making and continuity of care."

Ramsay Scribe is currently being trialed at St Andrew’s Ipswich Private Hospital in Queensland, with the operator describing the tool as “another step in our commitment to leveraging AI for better patient outcomes, supporting our dedicated people, and driving sustainable transformation”.

Ramsay Health Care created the tool after identifying a need to enhance clinical documentation efficiency while reducing administrative burden on clinicians.

“By leveraging AI-powered tools like Ramsay Scribe, we aim to improve workflow productivity, reduce burnout, and allow clinicians to focus more on patient care rather than paperwork," Gandhi said.

“The growing advancements in AI and speech-to-text technology provided an opportunity to implement a solution that enhances both accuracy and speed in documentation."

Initial research and planning began in mid-2024, progressing to prototyping and testing in "mid-late 2024" and the pilot in January 2025.

"We plan to continue our rollout across key sites in coming months," Gandhi said.”...

Source: https://www.itnews.com.au/news/ramsay-health-care-pilots-ai-powered-clinical-documentation-tool-615061

Gen AI: - Australian AI Scribes

Avant
By doctors for doctors

VoiceBox Intelligent Transcription

Streamline your medical correspondence from dictation to delivery.

https://avant.org.au/practice/voicebox-intelligent-transcription

Lyrebird Health

Automate all of your clinical documentation

Lyrebird listens to your conversations with patients, writes down what's said, and creates all the notes and paperwork you need.

https://www.lyrebirdhealth.com/au

i-scribe

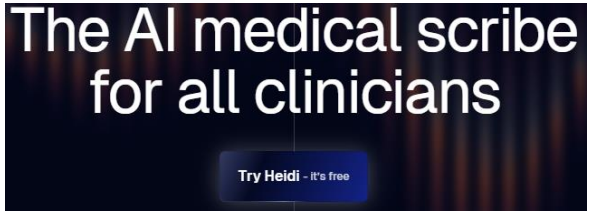
Transcribe. Generate. Share.

Your AI medical scribe, designed and built by Australian doctors.

Whether you're a surgeon, physician, or an allied health professional, i-scribe simplifies your documentation process, allowing you more time to focus on what matters most – your patients.

https://i-scribe.com.au/

Heidi



https://www.heidihealth.com/au

mAIscribe

MEDICAL AI SCRIBE – BY DOCTORS FOR DOCTORS

Save hours each day with mAIscribe.

Accurate, detailed and **customised** clinical documentation plus GP co-billing recommendations. It's like having your own personal medical secretary.

https://maiscribe.com.au/

Gen AI: - AI in diagnostics


PULSE+IT

News / Allied HealthAustralian Digital Health

Share

Artificial intelligence diagnosing lung diseases

29 January 2025

By Heather Fletcher

Source: https://www.pulseit.news/australian-digital-health/artificial-intelligence-diagnosing-lung-diseases

"Researchers from Charles Darwin University (CDU), United International University, and Australian Catholic University (ACU) have developed and trained an AI model to analyse lung ultrasound videos and diagnose respiratory diseases.

The model examined each video frame to find important features of the lungs and assessed the order of the video frames to understand the patterns of the lungs over time. It then identified specific patterns indicating different lung diseases and based on this information, classified the ultrasound into a diagnosis category such as normal, pneumonia, COVID-19 and other lung diseases.

Co-author and CDU adjunct Associate Professor Niusha Shafiabady said the model had an accuracy of 96.57 per cent, with the AI analyses verified by medical professionals…"


PULSE+IT

News / Australian Digital Health

Share

Wound care boosted by artificial intelligence

15 January 2025

By Reesh Lyon

Source: https://www.pulseit.news/australian-digital-health/wound-care-boosted-by-artificial-intelligence/

"Amplar Home Health patients will soon benefit from AI-powered wound care using iPads and smartphones.

Amplar has announced it will incorporate the Net Health Tissue Analytics platform, an application it says can transform iPads and smartphones into "sophisticated imaging platforms" that allow clinicians to use artificial intelligence and computer vision for advanced wound assessment.

Clinicians using the app can take photos and videos of a wound, with the computer vision technology then calculating the surface area and volume in millimeters.

The app can accurately calculate the dimensions and colour composition of various wound types and allow clinicians to track healing progress more accurately…"

Gen AI: - AI diagnoses

COMPUTERWORLD

AI chatbots outperform doctors in diagnosing patients, study finds

news analysis
13 Feb 2025 6 mins

"A new study comparing whether chatbots or physicians can diagnose patients more accurately and quickly found that AI is more often better at the task.

Chatbots quickly surpassed human physicians in diagnostic reasoning — the crucial first step in clinical care — according to a new study published in the journal Nature Medicine.

The study suggests physicians who have access to large language models (LLMs), which underpin generative AI (genAI) chatbots, demonstrate improved performance on several patient care tasks compared to colleagues without access to the technology.

The study also found that physicians using chatbots spent more time on patient cases and made safer decisions than those without access to the genAI tools.

The research, undertaken by more than a dozen physicians at Beth Israel Deaconess Medical Center (BIDMC), showed genAI has promise as an "open-ended decision-making" physician partner.

"However, this will require rigorous validation to realize LLMs' potential for enhancing patient care," said Dr. Adam Rodman, director of AI Programs at BIDMC. "Unlike diagnostic reasoning, a task often with a single right answer, which LLMs excel at, management reasoning may have no right answer and involves weighing trade-offs between inherently risky courses of action."

The conclusions were based on evaluations about the decision-making capabilities of 92 physicians as they worked through five hypothetical patient cases. They focused on the physicians' management reasoning, which includes decisions on testing, treatment, patient preferences, social factors, costs, and risks.

When responses to their hypothetical patient cases were scored, the physicians using a chatbot scored significantly higher than those using conventional resources only. Chatbot users also spent more time per case — by nearly two minutes — and they had a lower risk of mild-to-moderate harm compared to those using conventional resources (3.7% vs. 5.3%). Severe harm ratings, however, were similar between groups.

"My theory," Rodman said, "[is] the AI improved management reasoning in patient communication and patient factors domains; it did not affect things like recognizing complications or medication decisions. We used a high standard for harm — immediate harm — and poor communication is unlikely to cause immediate harm.""...

Source: https://www.computerworld.com/article/3823233/ai-chatbots-outperform-doctors-in-diagnosing-patients-study-finds.htm

Gen AI: - Bendigo Primary Care Centre case study

Bendigo Primary Care Centre is a large GP and multi-disciplinary community owned Super Clinic with that has grown from 38 staff in 2023 to 115 staff.

Callum Wright, BPCC Executive Officer , says, “Generative AI tools Copilot and ChatGPT have saved me 16 hours per month preparing the Monthly Management report, which fluctuates from 30 to 70 pages, with generally 20 pages of data metrics and dashboards within that.”

“Using Copilot has allowed me to focus on growing the business, with more time for analysis and planning, not focusing on writing facts. More time for finding GPs, registrars, and alliances bringing in specialists, allied health, metal health practitioners. AI has provided me the time capacity for those relationship linkages…”

Gen AI: - Bendigo Primary Care Centre case study

“Copilot AI is just another resource that reports to me.

AI writes the reports so I can manage the meetings.

I like the way Copilot AI writes more clearly, more concisely, and quicker that I do, and so not requiring as much cognitive effort.

Must do’s are re-reading all the output, checking the facts, no patient info, no sensitive financials, no IP, etc.

Great for policy development, responding to customer complaints – interestingly softer correspondence with AI, newsletters, etc., and has reduced the cost of compliance reporting, e.g., RACGP, PHN, Health department, etc.

I don’t like the inserting of American spelling – z instead of s…

We use AI for repeatable and auditable elements, not for answers to complex problems”

Privacy & cybersecurity obligations in healthcare for AI

Policy for responsible use of AI:

Contents

Source: https://www.digital.gov.au/policy/ai/policy

AI risks:

Source: https://www.digital.gov.au/policy/ai/policy p17

"**What is the risk that the use of AI**:

- negatively affects public accessibility or inclusivity of government services
- unfairly discriminates against individuals or communities
- perpetuates stereotypes or demeaning representations of individuals or communities
- causes harm to individuals, communities, businesses or the environment
- results in privacy concerns due to the sensitivity of the data being manipulated, parsed or transformed by the system
- results in security concerns due to the sensitivity or classification of the data being manipulated, parsed or transformed by the system
- results in security concerns due to the implementation, sourcing or characteristics of the system
- influences decision-making that affects individuals, communities, businesses or the environment
- poses a reputational risk or undermines public confidence in government
- results in intellectual property concerns due to the system manipulating, transforming or reproducing material for which a third party owns copyright.

Australian Government Guidance on use of GEN AI for APS:

Australian Government
Digital Transformation Agency

Australian Government Architecture

Search

Home | How to use | Explore the AGA | AGA and New Investments | Contact

Home > Interim guidance on government use of public generative AI tools - November 2023

Interim guidance on government use of public generative AI tools - November 2023

Updated on 22 November 2023

Notice

This guidance will be iterative. It is provided for government agencies to implement within their organisation. APS staff should follow their agency's policies and guidance on using generative AI tools in the first instance.

Feedback from public consultation on the responsible use of AI in Australia will be used to inform consideration across government on appropriate regulatory and policy responses that may include future iterations of this guidance.

Guidance for Australian Public Service (APS) staff

Generative AI tools present new and innovative opportunities for government. However, due to their rapid evolution and uptake, the risks involved in their use need to be considered and assessed.

The breadth of government activities includes developing policy advice for ministers, delivering programs to industry, providing services to the community and providing regulatory oversight. As such, the risk of using generative AI tools for official activities is context-specific and requirements will differ depending on how they are deployed.

Users should first and foremost align with their departmental or agency ICT obligations and policies. The DTA encourages departments and agencies to review their policies related to AI in line with this advice.

This guidance will be supplemented in due course with a risk framework to assist with the risk assessment process.

Golden rules

As you consider using generative AI tools in your work, you should assess the potential benefits and risks for each use case and take appropriate steps to mitigate them.

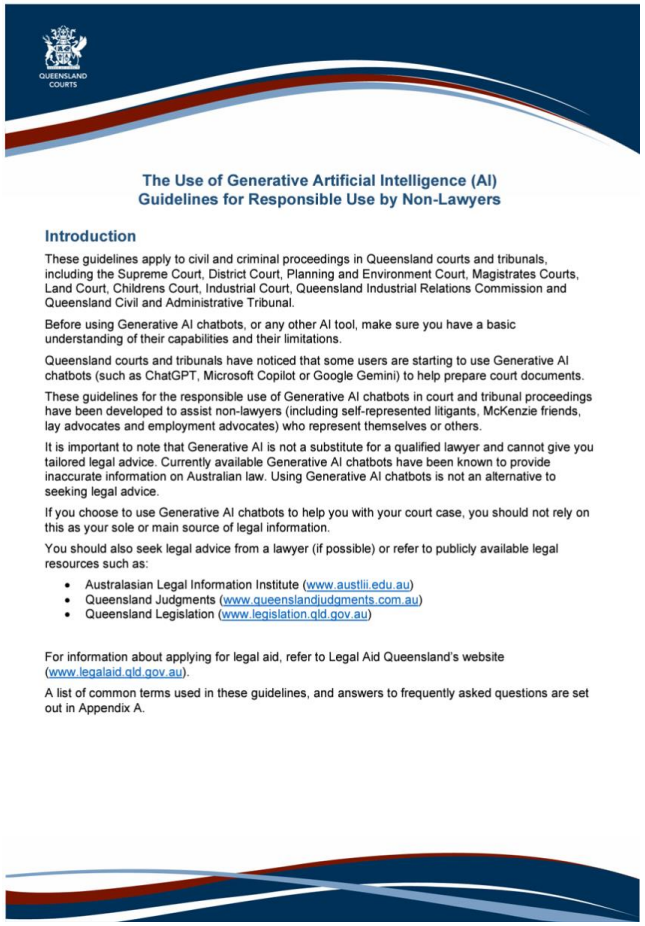
The principles, tactical guidance and use cases that follow will guide responsible application of these tools. Above all, apply these two **golden rules.**

- **You should be able to explain, justify and take ownership of your advice and decisions.**

- **Assume any information you input into public generative AI tools[1] could become public. Don't input anything that could reveal classified, personal or otherwise sensitive information.**

.

Queensland Courts Guidelines for Use of Gen AI:



The Use of Generative Artificial Intelligence (AI)
Guidelines for Responsible Use by Non-Lawyers

Introduction

These guidelines apply to civil and criminal proceedings in Queensland courts and tribunals, including the Supreme Court, District Court, Planning and Environment Court, Magistrates Courts, Land Court, Childrens Court, Industrial Court, Queensland Industrial Relations Commission and Queensland Civil and Administrative Tribunal.

Before using Generative AI chatbots, or any other AI tool, make sure you have a basic understanding of their capabilities and their limitations.

Queensland courts and tribunals have noticed that some users are starting to use Generative AI chatbots (such as ChatGPT, Microsoft Copilot or Google Gemini) to help prepare court documents.

These guidelines for the responsible use of Generative AI chatbots in court and tribunal proceedings have been developed to assist non-lawyers (including self-represented litigants, McKenzie friends, lay advocates and employment advocates) who represent themselves or others.

It is important to note that Generative AI is not a substitute for a qualified lawyer and cannot give you tailored legal advice. Currently available Generative AI chatbots have been known to provide inaccurate information on Australian law. Using Generative AI chatbots is not an alternative to seeking legal advice.

If you choose to use Generative AI chatbots to help you with your court case, you should not rely on this as your sole or main source of legal information.

You should also seek legal advice from a lawyer (if possible) or refer to publicly available legal resources such as:

- Australasian Legal Information Institute (www.austlii.edu.au)
- Queensland Judgments (www.queenslandjudgments.com.au)
- Queensland Legislation (www.legislation.qld.gov.au)

For information about applying for legal aid, refer to Legal Aid Queensland's website (www.legalaid.qld.gov.au).

A list of common terms used in these guidelines, and answers to frequently asked questions are set out in Appendix A.

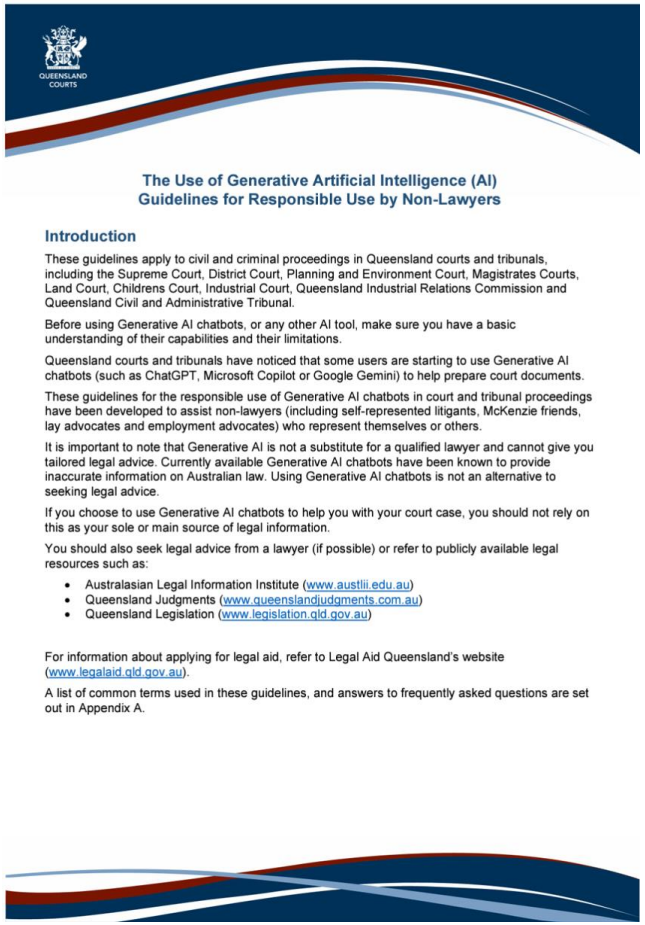
"**Before using Generative AI chatbots (or any other AI tool) make sure you have a basic understanding of their capabilities and their limitations. Despite the name, Generative AI chatbots are not actually intelligent in the ordinary human sense. Nor is the way in which they provide answers analogous to the human reasoning process.** **It is important to note:**

- **Generative AI chatbots are built on Large Language Models (LLMs). LLMs analyse a large amount of training text to predict the probability of the next best word in a sentence given the context. Just as Google offers to autocomplete your search, LLMs autocomplete repeatedly to form words, sentences, and paragraphs of text.**
- **LLMs have been further trained on ideal human written responses to prompts, and on survey results, about which responses sound most natural or best mimic human dialogue.**
- **This means the answers which Generative AI chatbots generate is what the chatbot predicts to be the most likely combination of words (based on the documents and data that it holds as source information), not necessarily the most accurate answer.**

And because their responses are on probability-derived calculations about the next best word in context, these tools are unable to reliably answer questions that require a nuanced understanding of language content. They have no intrinsic understanding of what any word they output means, nor a conception of truth."

Source: https://www.courts.qld.gov.au/going-to-court/using-generative-ai

Queensland Courts Guidelines for Use of Gen AI:



The Use of Generative Artificial Intelligence (AI)
Guidelines for Responsible Use by Non-Lawyers

Introduction

These guidelines apply to civil and criminal proceedings in Queensland courts and tribunals, including the Supreme Court, District Court, Planning and Environment Court, Magistrates Courts, Land Court, Childrens Court, Industrial Court, Queensland Industrial Relations Commission and Queensland Civil and Administrative Tribunal.

Before using Generative AI chatbots, or any other AI tool, make sure you have a basic understanding of their capabilities and their limitations.

Queensland courts and tribunals have noticed that some users are starting to use Generative AI chatbots (such as ChatGPT, Microsoft Copilot or Google Gemini) to help prepare court documents.

These guidelines for the responsible use of Generative AI chatbots in court and tribunal proceedings have been developed to assist non-lawyers (including self-represented litigants, McKenzie friends, lay advocates and employment advocates) who represent themselves or others.

It is important to note that Generative AI is not a substitute for a qualified lawyer and cannot give you tailored legal advice. Currently available Generative AI chatbots have been known to provide inaccurate information on Australian law. Using Generative AI chatbots is not an alternative to seeking legal advice.

If you choose to use Generative AI chatbots to help you with your court case, you should not rely on this as your sole or main source of legal information.

You should also seek legal advice from a lawyer (if possible) or refer to publicly available legal resources such as:

- Australasian Legal Information Institute (www.austlii.edu.au)
- Queensland Judgments (www.queenslandjudgments.com.au)
- Queensland Legislation (www.legislation.qld.gov.au)

For information about applying for legal aid, refer to Legal Aid Queensland's website (www.legalaid.qld.gov.au).

A list of common terms used in these guidelines, and answers to frequently asked questions are set out in Appendix A.

Source: https://www.courts.qld.gov.au/going-to-court/using-generative-ai

"Some capabilities of Generative AI chatbots

Generative AI chatbots cannot give you reliable legal advice that is tailored to your specific case. However, they may be able to help you by identifying and explaining laws and legal principles that might be relevant to your situation.

Generative AI chatbots may be able to help you prepare some basic legal documents. For example, they may be able to help you organise the facts into a clearer structure or suggest suitable headings. They can also help with formatting and provide suggestions on grammar, tone, vocabulary and writing style.

Some limitations of Generative AI chatbots

Generative AI chatbots are not search engines. They do not provide answers from authoritative databases, but rather generate new text using a complex algorithm, based on the prompts they receive and the data with which they have been 'trained'. Generally, the text used to train public Generative AI chatbots comes from various internet sources, such as webpages, online books, and social media posts.

This means the output which Generative AI chatbots generate is what the chatbot predicts to be the most likely combination of words (based on the documents and data that it holds as source information), not necessarily the most correct or accurate answer.

The currently available Generative AI chatbots have limited 'training' on Australian law and court procedure. Even when the training for Generative AI chatbots improves, there will be a limitation based on the currency of the data on which they have been trained."

Australia's AI Ethics Principles:

1. Human, societal and environmental wellbeing

Throughout their lifecycle, AI systems should benefit individuals, society and the environment

2. Human-centred values

AI systems should respect human rights, diversity and the autonomy of individuals.

3. Fairness

AI systems should be inclusive and accessible and should not involve or result in unfair discrimination against individuals, communities or groups.

4. Privacy protection and security

AI systems should respect and uphold privacy rights of individuals and ensure the protection of data.

5. Reliability and safety

Throughout their lifecycle, AI systems should reliably operate in accordance with their intended purpose.

6. Transparency and explainability

There should be transparency and responsible disclosure so people can understand when they are being significantly impacted by AI, and can find out when an AI system is engaging with them.

7. Contestability

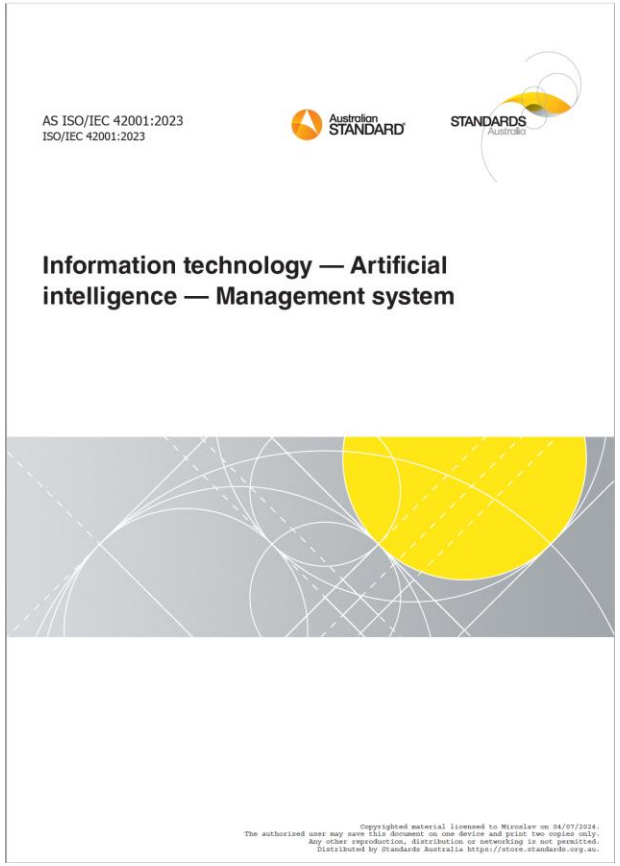
When an AI system significantly impacts a person, community, group or environment, there should be a timely process to allow people to challenge the use or outcomes of the AI system.

8. Accountability

Those responsible for the different phases of the AI system lifecycle should be identifiable and accountable for the outcomes of the AI systems, and human oversight of AI systems should be enabled.

Source: https://www.finance.gov.au/government/public-data/data-and-digital-ministers-meeting/national-framework-assurance-artificial-intelligence-government

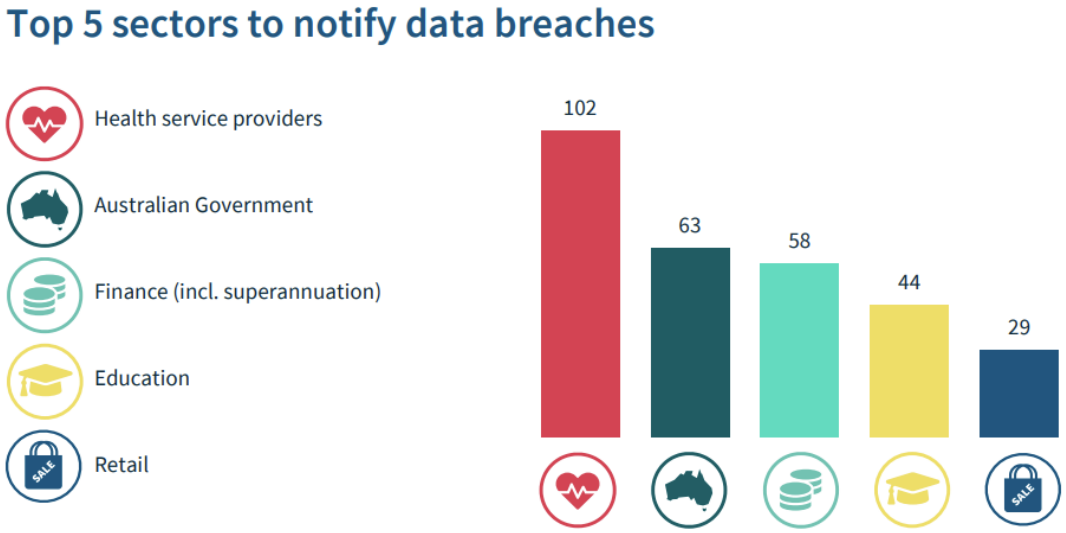
International Standards for AI: AS ISO/IEC 42001:2023:



Contents

Healthcare sector tops the Notifiable Data Breaches Report, *again*: January – June 2024:

"The latest Office of Australian Information Commissioner (OAIC) Notifiable Data Breaches Report for January – June 2024 identifies healthcare providers at the top of the list, *again*" …



Source: https://www.oaic.gov.au/__data/assets/pdf_file/0013/242050/Notifiable-data-breaches-report-January-to-June-2024.pdf

Healthcare sector tops the Notifiable Data Breaches Report, *again*: January – June 2024:

Reasonable steps

Several of the Australian Privacy Principles (APPs), such as APPs 1, 8 and 11, and NDB scheme provisions require an entity to take 'reasonable steps' to comply with an obligation. In particular:

- APP 11.1 requires entities to take reasonable steps to protect personal information from misuse, interference and loss, as well as unauthorised access, modification or disclosure.
- APP 11.2 states an organisation must take reasonable steps to destroy or de-identify information it no longer needs for any purpose for which the information may be used or disclosed under the APPs.

As outlined in the OAIC's *Australian Privacy Principles guidelines* and *Guide to securing personal information*, what constitutes reasonable steps depends on circumstances such as:

- the nature of the entity, its size and resources
- the volume and sensitivity of personal information concerned
- possible adverse consequences for an individual in case of a breach.

Source: https://www.oaic.gov.au/__data/assets/pdf_file/0013/242050/Notifiable-data-breaches-report-January-to-June-2024.pdf

Generative AI Risks:

Beyond the hype cycle and the predicted productivity gains, there are many expert voices raising concerns and urging government regulation to prevent GAI causing harm to humans.

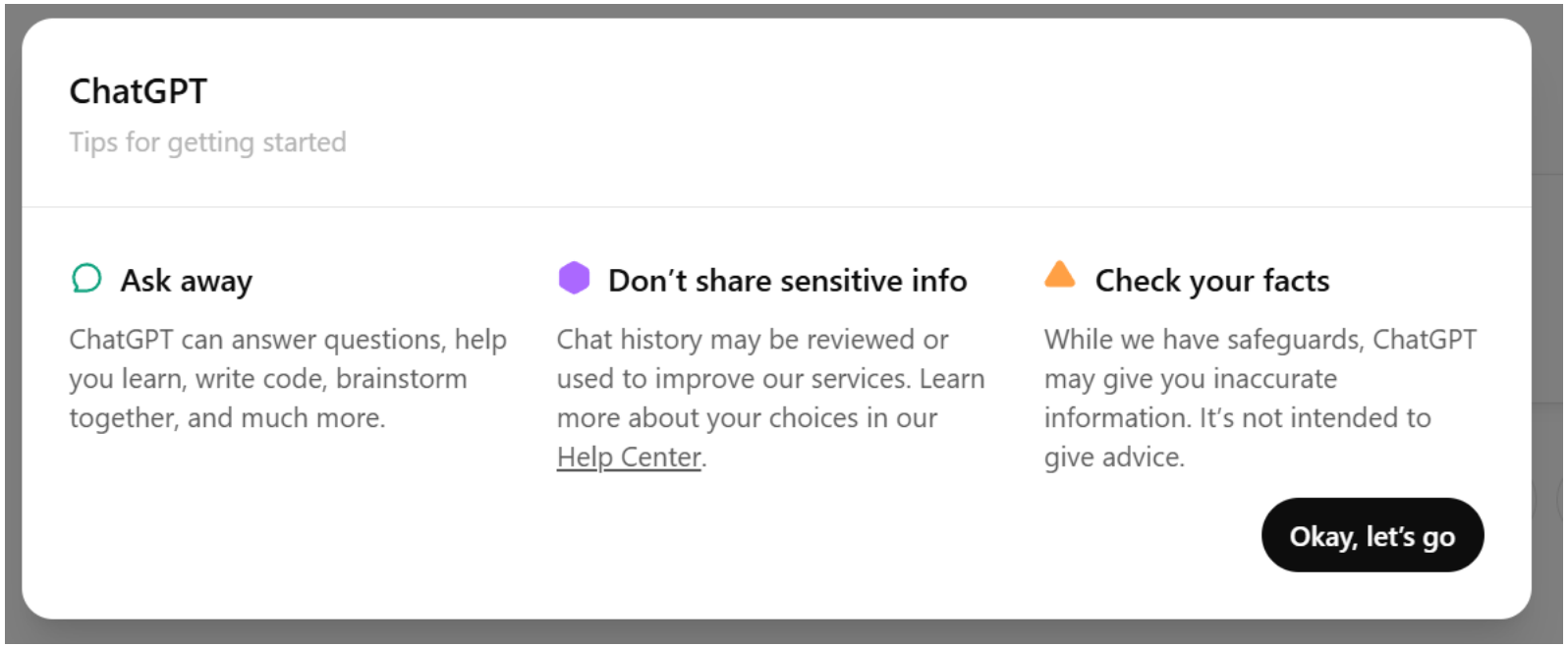
AI limitations are well documented with negative aspects of GAI, for example, AI fabrications, the so called 'AI hallucinations', AI biases, privacy concerns, misinformation, and disinformation, deep fakes, AI driven cyber attacks, etc.

There are very real ethical, legal, indemnity, privacy, confidentiality, intellectual property, *and* cyber security concerns with using Gen AI tools in a healthcare setting.

Cyber threat actors are also using Gen AI to find vulnerabilities, for social engineering attacks, and to hack into networks for data theft and extortion…

PROCEED WITH CAUTION!

Generative AI Risks: ChatGPT warnings



Gen AI: - DeepSeek

Australia bans DeepSeek on government devices

By Staff Writer
Feb 5 2025 6:28AM

Citing security concerns

Australia has banned DeepSeek from all government devices over concerns that the Chinese artificial intelligence startup poses security risks.

"The secretary of the Department of Home Affairs issued a mandatory direction for all government entities to "prevent the use or installation of DeepSeek products, applications and web services and where found remove all existing instances of DeepSeek products, applications and web services from all Australian Government systems and devices," the statement said.

Home Affairs Minister Tony Burke said DeepSeek posed an "unacceptable risk" to government technology and the immediate ban was "to protect Australia's national security and national interest."

The ban does not extend to devices of private citizens.

Tech stocks worldwide plunged after the launch of DeepSeek last month - apparently costing a fraction of rival AI models and requiring less sophisticated chips - raised questions over the West's huge investments in chipmakers and data centres.

Australia's decision to ban Deepseek follows similar action in Italy, while other countries in Europe and elsewhere are also looking into the AI firm.

Taiwan banned government departments from using DeepSeek earlier this week.

Prime Minister Anthony Albanese's government imposed a government-wide ban on Chinese social media app TikTok two years ago over security concerns.."...

Source: https://https://www.itnews.com.au/news/australia-bans-deepseek-on-government-devices-614763

Cyber Risks 2024 : Generative AI

THE WEEKEND AUSTRALIAN

Sunday, March 3, 2024 | Today's Paper | Mind Games

≡ All sections HOME THE NATION WORLD BUSINESS COMMENTARY SPORT ARTS All

HOME / NATION / POLITICS

Save

ChatGPT for criminals to turbo charge scams, say Australian Federal Police

EXCLUSIVE

By DAVID MURRAY
NATIONAL CRIME CORRESPONDENT

7:11PM JANUARY 30, 2024
3 COMMENTS

Australians are at risk from malicious AI that could lead to a new generation of scams, federal police warn. Picture: AAP

An Australian Federal Police submission to a federal cybercrime inquiry flags AI models such as **FraudGPT** and **WormGPT** as a growing threat.

The models are similar to ChatGPT, but are devoid of restrictions on answering questions about illegal activity.

They provide a suite of tools that can craft spear-phishing emails "with perfect grammar and spelling", aimed at stealing sensitive information such as login details, the AFP says.

The tools can also assist in voice phishing for "hi Mum, hi Dad scams", business email compromise attacks, generating malware and testing for security vulnerabilities.

"The development of malicious AI models by threat actors is in its early stages, but already proving effective and lowers the entry threshold for burgeoning cyber criminals who may lack the technical proficiencies or resources to establish their own cyber criminal tradecraft," the submission states.

The Australian Institute of Criminology also warns: "AI is already being leveraged by criminal actors to upscale and enhance criminal activities, exploit human-centric vulnerabilities and lower the barriers and costs to engaging in criminal activities."

With Australians reeling from major hacks on corporations including Medibank and Optus, the federal parliamentary joint committee on law enforcement is conducting an inquiry into the capability of agencies to respond to cybercrime.

When WormGPT's existence was revealed in July, it was described by cyber security firm SlashNext as being "similar to ChatGPT but with no ethical boundaries or limitations".

FraudGPT has reportedly been advertised on the dark web as an "unrestricted alternative for ChatGPT".

The AFP says current trends such as Ransomware-as-a-Service (RaaS) and Malware-as-a-Service (MaaS) have allowed more people to launch cyber attacks and scams. "The frequency and severity of cybercrime incidents are expected to increase as a result, placing new demands on the AFP as a law enforcement agency," the submission states.

Source: https://www.theaustralian.com.au/nation/politics/chatgpt-for-criminals-to-turbo-charge-scams-say-australian-federal-police/news-story/6dfcbdc500381f3569762a68adf12dfe

Cyber Risks 2024 : Generative AI

THE WEEKEND AUSTRALIAN

Sunday, March 3, 2024 | Today's Paper | Mind Games

≡ All sections HOME THE NATION WORLD BUSINESS COMMENTARY SPORT ARTS All

HOME / NATION / POLITICS

Save

ChatGPT for criminals to turbo charge scams, say Australian Federal Police

EXCLUSIVE

By DAVID MURRAY
NATIONAL CRIME CORRESPONDENT

7:11PM JANUARY 30, 2024

3 COMMENTS

Australians are at risk from malicious AI that could lead to a new generation of scams, federal police warn. Picture: AAP

Other recent hacks were carried out on Victoria's court system and the St Vincent's Health network.

Malicious AI was also used to create increasingly realistic deepfakes, lifelike child abuse material and believable disinformation content, the AFP says.

"AI is an emerging technology in child exploitation matters and as it improves, the material is becoming more lifelike. This type of material is known as deepfakes, which involve manipulating images, audio and video using AI."

The AIC's submission to the inquiry states almost one in every two Australians surveyed had been a victim of at least one type of cybercrime in the previous 12 months.

"Cybercrime targeting individual computer users is most frequently a high volume, low yield crime," the AIC states.

"The high rate of victimisation means that, even with the relatively small median losses per victim, the overall cost to Australian individuals is likely to be enormous."

The impact of losses was "potentially catastrophic and can have long-term effects on - victims".

Cyber criminals were quick to adopt emerging technologies for criminal purposes, the AIC states.

"Artificial intelligence has the potential to facilitate better-targeted, more frequent and widespread criminal attacks, and is already being used for password guessing, CAPTCHA-breaking and voice cloning," the submission states.

Source: https://www.theaustralian.com.au/nation/politics/chatgpt-for-criminals-to-turbo-charge-scams-say-australian-federal-police/news-story/6dfcbdc500381f3569762a68adf12dfe

Cyber Risks 2024 : Generative AI

Cyber attacks are more sophisticated than ever, and IT leaders feel ill-equipped to handle emerging threats

Source: https://www.keepersecurity.com/en_GB/top-data-threats-insight-report/

Cyber Risks 2025 : Generative AI

SC Media
A CRA Resource

Network Security, Vulnerability Management, Ransomware

2025 Forecast: AI to supercharge attacks, quantum threats grow, SaaS security woes

January 1, 2025
By Stephen Weigand

Artificial intelligence will super-charge familiar 2024 threats in 2025, putting new wrinkles on old security challenges such as phishing, insider threats and ransomware.

Meanwhile artificial intelligence (AI) itself will increasingly be thrust in hacker crosshairs with AI models themselves coming under innovative new and constant threats such as malicious prompt injections from bad actors and large language model (LLM) data tampering by adversaries.

Red flag warnings also include a growing number of cyberattacks that have real world implications as hacks continue to impact the virtual and now the physical world. This year, experts shared with SC Media, that 2025 will be shaped by the rise and use of quantum computing attack techniques by adversaries aimed at exploiting existing and emerging encryption tech.

The promises and dangers of artificial intelligence in cybersecurity

Generative AI will upend traditional security methods — and vastly increase the amount of zero days to the detriment of many, says Sanjeev Verma, co-founder of PreVeil…

Threat actors will exploit AI by manipulating private data, Daniel Rapp, Proofpoint's chief AI and data officer…

Protect against AI-assisted threats; plan for AI-powered threats, Troy Bettencourt, head of IBM X-Force and global partner…

Cybersecurity will see a 'trust and verify' approach to coding with AI, says Andrea Malagodi, Sonar CIO…

Malicious use of multimodal AI will create entire attack chains, says Corey Nachreiner, CISO at WatchGuard…

Bad actors will develop synthetic online personalities for financial gain, says Tyler Swinehart, Ironscales director of global IT and security…

System prompt vulnerabilities is Achilles' heel for LLM security, says Elad Schulman, Skyhigh Security CEO…

AI-powered ransomware attacks will rise, says Art Ukshini, Permiso associate threat research…

Threat actors turn to AI-driven cloud threats, says Marina Segal, Tamnoon CEO

Source: https://www.scworld.com/feature/cybersecurity-threats-continue-to-evolve-in-2025-driven-by-ai

All existing cyber attacks are now powered by Gen AI:

- *Social Engineering:*
  - *Phishing, Spearphishing, Whaling, Business Email Compromise, SMSishing, Vishing, Impersonation*
- *Ransomware*
- *Triple Extortion – Multi Faceted Extortion*
- *Supply Chain attacks*
- *Data Leakage/Insider Threat*
- *Credential Theft*
- *Network Perimeter and Endpoint security:*

  Working From Home?
- *Denial of Service (DoS) and*
- *Distributed Denial of Service (DDoS)*
- *Insider Threat*

Gen AI specific risks:

- *Prompt Injection Attacks*
- *Prompt Flooding*
- *Denial of Service (DoS) and*
- *Distributed Denial of Service (DDoS)*
- *Data Leakage/Insider Threat*
- *Excessive Agency*
- *Training Data Poisoning*
- *Breach of Intellectual Property/Copyright laws*
- *Misuse*
- *Socially Questionable Outcomes*
- *Supply Chain Commercial issues*
- *Unauthorized Access to systems*
- *Non-compliance with AI / ML regulations*

AI Hallucinations Case Study 1:

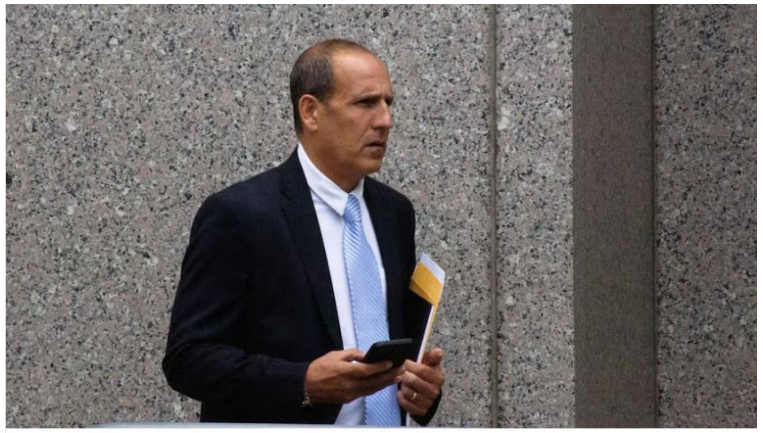
ABC NEWS

Just In | Watch Live | Politics | World | Business | Analysis | Sport | Science | Health | Entertainment | Lifestyle | Fact Check | Other

This US lawyer used ChatGPT to research a legal brief with embarrassing results. We could all learn from his error

ABC RN / By Damien Carrick and Sophie Kesteven for the Law Report, with additional reporting from Reuters.

Posted Sat 24 Jun 2023 at 11:15am



Steven Schwartz, pictured outside a Manhattan court in June, used ChatGPT to research a legal case. *(Getty: New York Daily News)*

A New York-based lawyer has been fined after he misused the artificial intelligence chatbot, ChatGPT, relying on it for research for a personal injury case.

Last week Steven A. Schwartz, fellow lawyer Peter LoDuca and law firm Levidow, Levidow & Oberman, were fined US$5,000 (AU$7,485) for submitting fake citations in a court filing.

The judge found the lawyers acted in bad faith and made "acts of conscious avoidance and false and misleading statements to the court."

In a written opinion, Judge P. Kevin Castel said lawyers had to ensure their filings were accurate, even though there was nothing "inherently improper" about using artificial intelligence in assisting with legal work.

“Technological advances are commonplace and there is nothing inherently improper about using a reliable artificial intelligence tool for assistance,” Castel wrote.

“But existing rules impose a gatekeeping role on attorneys to ensure the accuracy of their filings.” Schwartz, who has more than 30 years' experience practicing law in the US, was part of a legal team acting for a man suing the airline Avianca. The client, Roberto Mata, had claimed that he was injured after a metal serving cart hit his knee during a flight.

Unfortunately for the client, Schwartz did his legal research for the case using ChatGPT without fact checking if the cases he cited in his brief, involving other airlines and personal injuries, were real or not.

Turns out they weren't.

"He did ask ChatGPT whether one of the cases was real but was happy enough when ChatGPT said yes," Professor Lyria Bennett Moses tells ABC RN's Law Report.

"In fact, ChatGPT told him that they could all be found on reputable databases, and he didn't do any checking outside of the ChatGPT conversation to confirm the cases were real — he didn't look any of them up on a legal database."

Professor Moses is the director of the UNSW Allens innovation hub. She says the lesson here is to use this platform with caution.

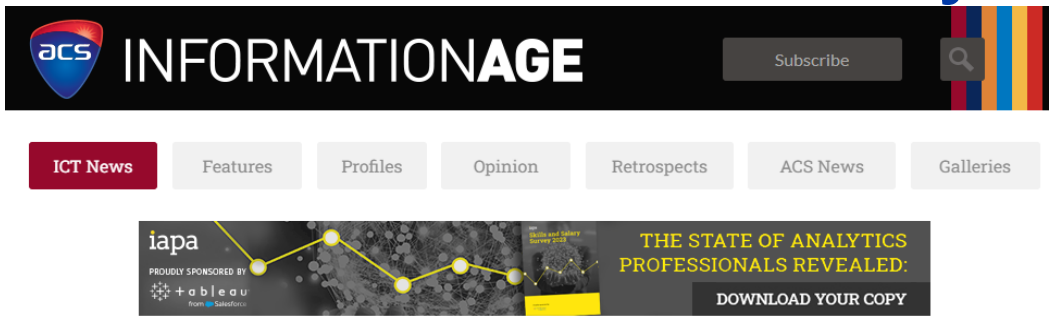
"[Schwartz] stated in the court [hearing], 'I just never could imagine that ChatGPT would fabricate cases.' So, what it showed is a real misunderstanding of the technology," she explains.

"[ChatGPT] has no truth filter at all. It's not a search engine. It's a text generator working on a probabilistic model. So, as another lawyer at the firm pointed out, it was a case of ignorance and carelessness, rather than bad faith."

Schwartz's lack of due diligence when researching his brief in this personal injury case has caused him great embarrassment, particularly as his hearing has drawn worldwide attention.

Source: https://www.abc.net.au/news/2023-06-24/us-lawyer-uses-chatgpt-to-research-case-with-embarrassing-result/102490068

AI Hallucinations Case Study 2:



Australian academics caught in generative AI scandal

Submitted fake information to Senate submission.

By Denham Sadler on Nov 06 2023 03:13 PM

"A group of Australian academics has "unreservedly apologised" for including factually incorrect allegations about big consulting firms – produced by a generative AI tool – to a Senate inquiry submission.

In a letter to the Senate, Emeritus Professor James Guthrie AM – a professor in the Department of Accounting and Corporate Governance at Macquarie University – admitted to having used Google Bard AI to research information for a submission to a Parliamentary inquiry into the conduct of the Big 4 consulting firms, with numerous false claims being generated, as reported by The Guardian.

The other academics in the group are Professor John Dumay (Macquarie University), Professor Jane Andrew (University of Sydney Business School), and Dr Erin Twyford (University of Wollongong), however they were quick to distance themselves from the scandal, making it very clear in their revised submission that the blame was squarely on Professor Guthrie.

"Our original submission contained factual errors due to Professor Guthrie's use of the Google Bard Large Language model generator referenced in the original submission," they wrote.

"These errors related to claims of Deloitte, KPMG, EY and PwC audit activities and involvement in financial scandals."

Bard is Google's conversational generative AI tool that is in direct competition with OpenAI's ChatGPT tool."

"What the academics said

The academics' submission included case studies about alleged wrongdoing by large consulting firms that had been produced by Bard but were entirely fictional.

It is believed to be the first time that a Parliamentary Committee has been forced to grapple with the use of generative AI in research and writing submissions to inquiries, which are covered by Parliamentary Privilege and free from any defamation action.

Included in the submission were a number of case studies about consulting and accounting giant Deloitte, including that it had been involved in a "NAB financial planning scandal", was sued by the liquidators of collapsed construction firm Probuild, had audited cafe chain Patisserie Valerie, and was auditing Westpac at the time of a scandal.

All of these claims are false.

In a letter to the Senate, Deloitte general counsel Tala Bennett said that there had never been a "Deloitte NAB financial planning scandal", that it was not the auditor of Probuild but had been the administrator and is not being sued in relation to this, and that it has never audited Patisserie Valerie or Westpac.

"Deloitte supports academic freedom and constructive discourse in relation to those matters currently before the Committee, however, it considers that it is important to have factually incorrect information corrected," Bennett said in a letter to the Committee.

"It is disappointing that this has occurred, and we look forward to understanding the Committee's approach to correcting this information."

The submission also made false claims about Big 4 consultancy KPMG, also generated by Google Bard.

The submission falsely claimed that KPMG had been involved in a "KPMG 7-Eleven wage theft scandal" that led to the resignation of several of its partners.

It also incorrectly said that KPMG had audited the Commonwealth Bank during a financial planning scandal.

But KPMG was not involved with either of these scandals…."

Source: https://ia.acs.org.au/article/2023/australian-academics-caught-in-generative-ai-scandal.html

AI Hallucinations Case Study 3:

ABC NEWS

Just In | Watch Live | Politics | World | Business | Analysis | Sport | Science | Health | Entertainment | Lifestyle | Fact Check | Other

An AI-generated image of a Victorian MP raises wider questions on digital ethics

By Joseph Dunstan and Mikaela Ortolan

Posted Thu 1 Feb 2024 at 9:25am, updated Thu 1 Feb 2024 at 3:18pm

Nine News apologised to Ms Purcell after broadcasting the digitally altered image of her on its nightly news (left). *(Nine News /*

It was an image broadcast for just a few moments on a Melbourne TV news bulletin, but it's since attracted international attention.

The digitally altered image of Victorian Animal Justice Party MP Georgie Purcell used to introduce a story on Victorian duck hunting saw the white dress she was wearing in the original photo swapped for a top exposing her midriff.

"Note the enlarged boobs and outfit to be made more revealing. Can't imagine this happening to a male MP," she tweeted.

After Ms Purcell called it out, media outlets including CNN and the BBC picked up the story, fuelling debate on the reach of generative artificial intelligence in our lives.

In its swift apology to Ms Purcell, broadcaster Nine News said the alteration had occurred due to an "automation by Photoshop" while resizing the original photo.

Ms Purcell says she's not sure she buys Nine's explanation but is happy to move on, provided a lesson is learnt by everyone to ensure it never happens again.

So how might the image have been made and what lessons should we take from it?

Expert says Nine's explanation is plausible

Nine News has told the ABC the alteration to Ms Purcell's midriff occurred when using Adobe Photoshop's "generative expand" tool.

The tool allows users to make an image bigger — the program uses AI to make assumptions or guesses about how that image might be best filled out with new material.

Nine said the image it used to produce the graphic was a more tightly cropped version of the Bendigo Advertiser's photo of Ms Purcell.

This image appears in some online image searches, and is cropped above Ms Purcell's waist.

Source: https://ia.acs.org.au/article/2023/australian-academics-caught-in-generative-ai-scandal.html

AI Hallucinations Case Study 3:

ABC NEWS

Just In | Watch Live | Politics | World | Business | Analysis | Sport | Science | Health | Entertainment | Lifestyle | Fact Check | Other

An AI-generated image of a Victorian MP raises wider questions on digital ethics

By Joseph Dunstan and Mikaela Ortolan

Posted Thu 1 Feb 2024 at 9:25am, updated Thu 1 Feb 2024 at 3:18pm

An image of a man created with AI generation below the neck, creating three alternative torsos. *(ABC News)*

TJ Thomson, a senior lecturer in digital media at RMIT, said it was plausible that a cropped image could produce a range of different torsos when expanded using AI.

"If you are giving Photoshop less data about the body ... it has to imagine a lot more about what's below the torso," he said.

"So it can have a lot more creative input into what is below the torso."

To demonstrate the tool in question, we uploaded a photo of a consenting ABC employee to Photoshop, cut off just above the chin.

When the program was presented with the image and asked to perform the "generative expand" action, it produced several variants of the torso.

The clothing in each torso differed, including a buttoned-up shirt and one with several of the top buttons undone.

In one version, the program appeared to struggle to generate the hands, presenting a jumble of fingers at the end of the man's arms.

Source: https://www.abc.net.au/news/2024-02-01/georgie-purcell-ai-image-nine-news-apology-digital-ethics/103408440

AI Hallucinations Case Study 4:

Businesses liable for AI chatbot mistakes

Air Canada's "remarkable" excuses don't fly with tribunal.
By David Braue on Mar 04 2024 10:07 PM

"Companies using AI chatbots to handle customer enquiries must honour the advice they give customers, a civil tribunal has ruled, after an Air Canada customer was given incorrect instructions by a chatbot that the airline tried to argue is "responsible for its own actions".

After his grandmother died in November 2022, Vancouver resident Jake Moffatt booked an airfare to Toronto and was advised by the company's website chatbot that he could apply within 90 days for retroactive reimbursement of the fare difference between normal and bereavement fares.

Airline policy, however, actually requires those receiving bereavement fares to apply for the lower rates before the booking was made – a point that Air Canada used in denying Moffatt's subsequent application for a refund.

A series of emails ensued, and in early February an Air Canada representative conceded that the chatbot had provided "misleading words" and that, despite Moffatt's correctly following its guidance, the airline would not pay the refund.

Yet Air Canada, the British Canada Civil Resolution Tribunal (CRT) ultimately held in awarding Moffatt $926 ($C812) in damages and costs, argued that the chatbot "is a separate legal entity that is responsible for its own actions" – a "remarkable" suggestion that, CRT found, had not been explained by Air Canada and made no logical sense."

"While a chatbot has an interactive component," the tribunal found, "it is still just a part of Air Canada's website. It should be obvious to Air Canada that it is responsible for all the information on its website."

"It makes no difference whether the information comes from a static page or chatbot," the ruling notes, adding that the airline "did not take reasonable care to ensure its chatbot was accurate".

Don't blame the AI when things go wrong

Amidst surging corporate adoption of conversational chatbots – often powered by generative AI (genAI) engines that are well known to make mistakes, make things up and be subject to manipulation – the Air Canada decision is a warning for businesses that simply adding conversational capabilities doesn't reduce their obligation to provide accurate information to customers.

...Bad information from a genAI chatbot "will directly lead to the death of a customer" by 2027, Gartner has warned as the technology becomes an increasingly significant business risk that must be managed to ensure accuracy and consistency with business policies, and to prevent inadvertent breaches of privacy, consumer, and other laws.

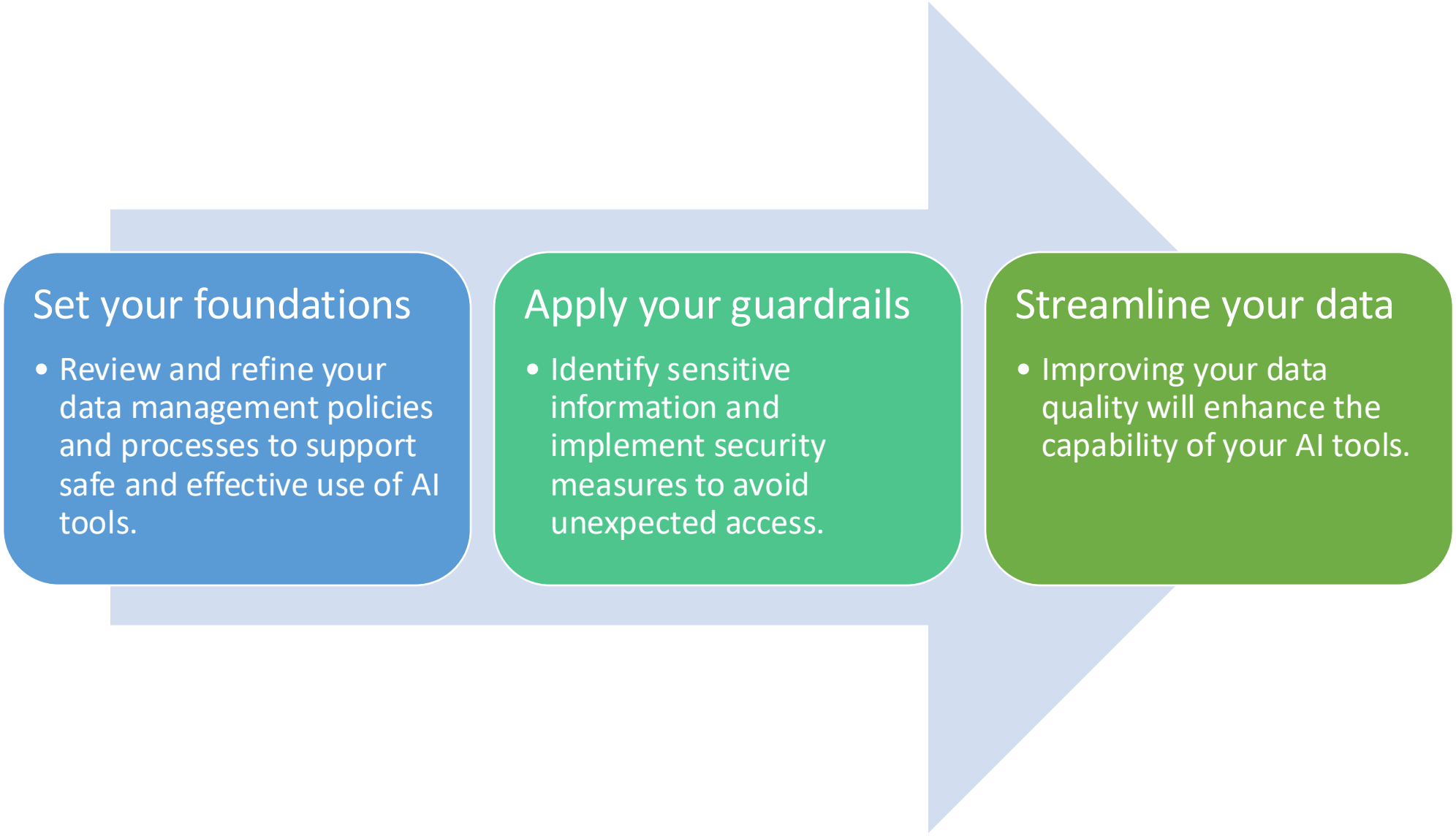
...Respondents cite widespread concerns about visibility into how AI is used, transparency of underlying algorithms, industry standards, and human validation of AI's outputs – not to mention the concerns identified in a recent Reveal survey of 585 software developers, in which 40.7 per cent said actually integrating AI into their software development process would be the biggest challenge in software development during 2024."

Source: https://ia.acs.org.au/article/2024/businesses-liable-for-ai-chatbot-mistakes.html

Gen AI – Considerations for use in Primary Care

Guardrails for Gen AI are *essential*:

Critical Steps



Check the output of Gen AI:

Credit: Shutterstock

“Only use AI to do work you can easily verify, and be sure to check it’s work”

Cordilia James
Wall Street Journal Tech Writer

Source: https://www.wsj.com/tech/ai/the-smartest-way-to-use-ai-at-work-ce921ff4

Gen AI – Practical Productivity in Primary Care

Once Gen AI tools, have been implemented with appropriate data safety and security guardrails, practice managers can use these tools to generate and transform data, emails, documents, spreadsheets, presentations, and meetings.

make

Revolutionize your work with AI and Make

Turn smarts into power, using Make to add AI into your business workflows.

Make's drag-and-drop automation tools allow you to gather info to prompt AI and seamlessly direct outputs to the right places.

https://www.make.com/en/ai-automation

rayven

Workflow Automation Software.

Automate workflows + streamline processes **in minutes.**

https://www.rayven.io/workflow-automation-software

n8n

Secure, AI-native workflow automation

The world's most popular workflow automation platform for technical teams

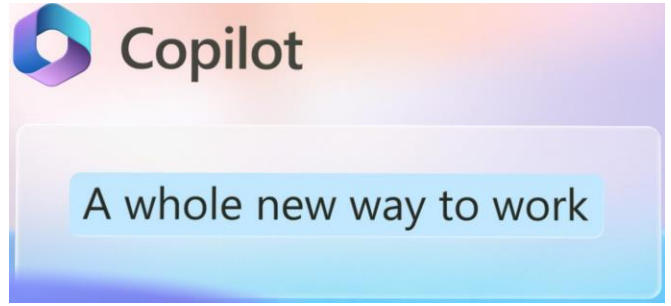
https://n8n.io/features/

Poe

The best AI, all in one place

https://poe.com/about

Microsoft



https://adoption.microsoft.com/en-us/copilot/smb/

Google AI

Making AI helpful for everyone

https://ai.google/

Gemini

The Gemini ecosystem represents Google's most capable AI

https://ai.google/get-started/gemini-ecosystem/

Copilot for Microsoft 365

The Extended Copilot Family

Azure & AI

Copilot

OpenAI


Power Platform


D365


M365


Security


Windows


Bing


GitHub


Azure OpenAI Service


Cognitive Services

How does Copilot for Microsoft 365 work?

Copilot for Microsoft 365

Microsoft 365 Apps

Chat

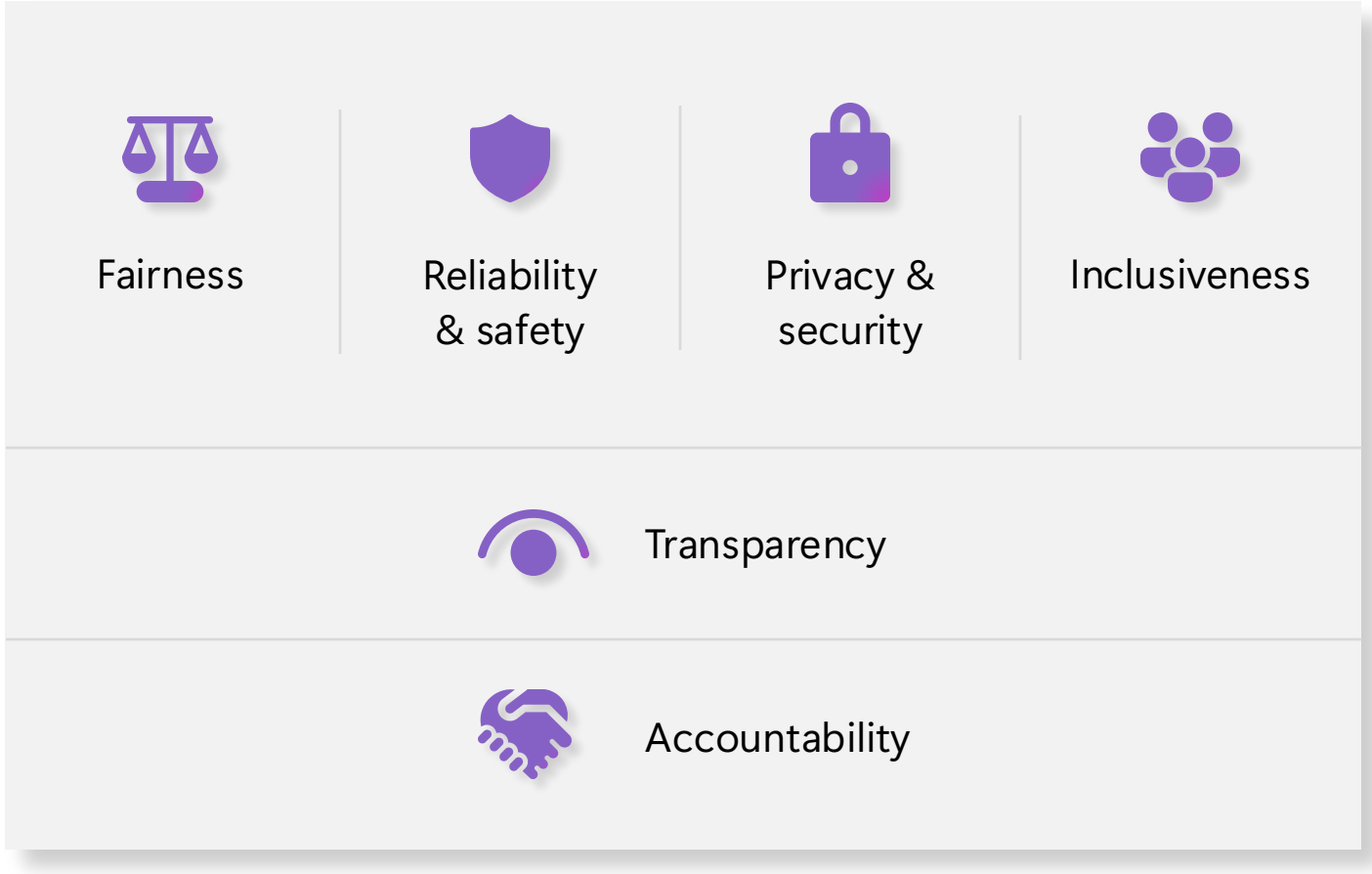
Microsoft Graph - Your Data -

Semantic Index

Copilot feature comparison

	Copilot (For individuals)	Copilot Pro	Copilot (For business)	Copilot for Microsoft 365
Foundational Capabilities	●	●	●	●
Web Grounding	●	●	●	●
Commercial Data Protection	●	●	●	●
Priority Model Access		●		●
Copilot in Outlook, Word, Excel, PowerPoint, and OneNote		●		●
Copilot in Teams				●
Microsoft Graph Grounding				●
Enterprise-Grade Data Protection				●
Customization		Copilot GPT Builder		Copilot Studio

Microsoft's AI principles



Microsoft Cloud — AI you can trust

Your data is **your** data.

Your data is **not** used to train the OpenAI foundation models without permission.

Your data is **protected** by the most comprehensive enterprise compliance and security controls.

Microsoft's approach to privacy

You control your data

You know where your data is located

We secure your data at rest and in transit

We defend your data

Shared responsibilities of security for AI usage for Copilot for Microsoft 365

Access to production environment

Set up access controls that strictly restrict standing access to customer's data or production environment

Protect data

Encrypt data at rest and in transit based on industrial standards (BitLocker, TLS, etc.)

Access to AI apps

Set up access control policy, leveraging identity and access management solutions

Protect data entered and created by AI apps

Apply sensitivity label that associates with encryption and Data Loss Prevention controls to files with sensitive data

User accountability of AI use

Implement and operate organisation-wide AI security training

Action

Security and Compliance controls for Copilot for Microsoft 365

Essential security controls

Copilot + M365 Business Standard

- Multi-factor Authentication
- Audit Logging
- Basic content and keyword search

Comprehensive security controls

Copilot + Microsoft 365 Business Premium

Everything in M365 Business Standard, plus:

- Conditional Access
- Sensitivity labels
- Data loss prevention policies
- Unified endpoint management
- eDiscovery, litigation hold and retention policies

Note: Not all features/products shown.

Action

Securing and governing Copilot for Microsoft 365

Baseline

Copilot for Microsoft 365
+ Office 365 E3

- **Multi-factor authentication**
- **Audit logging**

Core

Copilot for Microsoft 365
+ Microsoft 365 E3
+ SharePoint Advanced Management

- **Conditional Access**
- **Manual sensitivity labels**
- **Data Loss Prevention policies**
- **Advanced SharePoint sitewide access controls and reporting**
- **Unified endpoint management**

Best-in-class

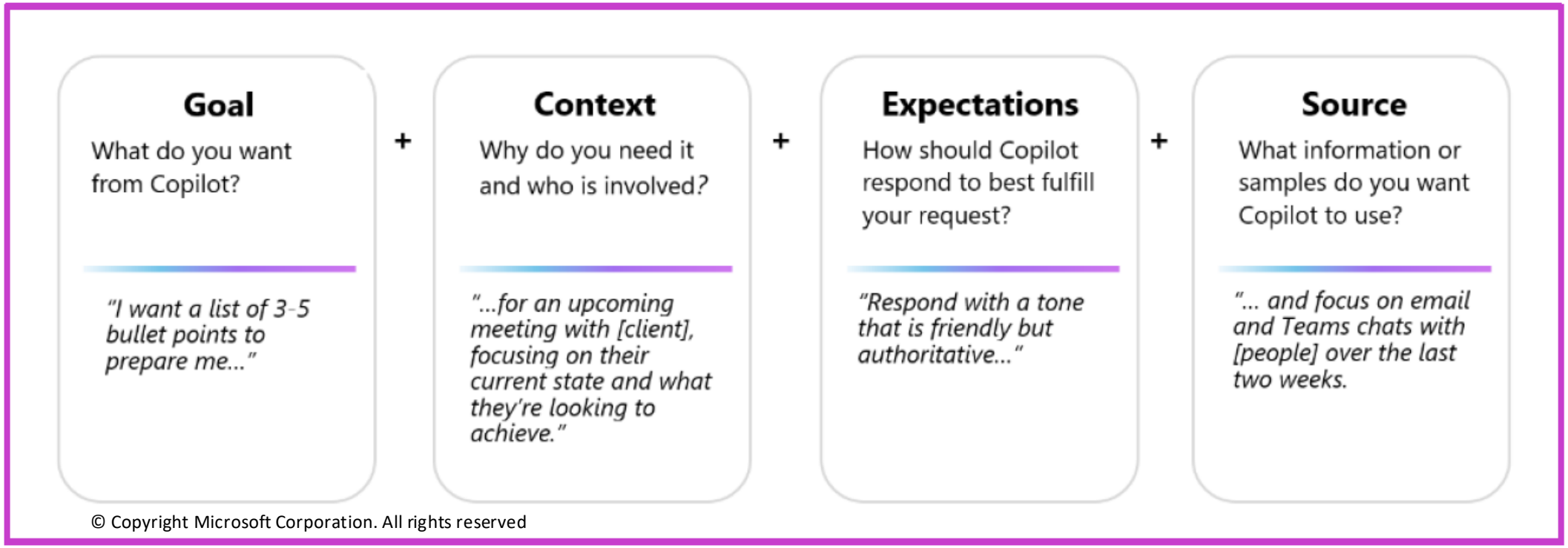
Copilot for Microsoft 365
+ Microsoft 365 E5
+ SharePoint Advanced Management

- **Conditional Access based on identity risk**
- **Automatically apply sensitivity labels**
- **Automatically remove inactive content**
- **Prevent data leak on endpoint devices**
- **Detect non-compliant usage**

Microsoft Copilot M365:

Examine how to build an effective prompt

Prompts can include four parts—the goal, context, expectations, and source, as described in the following image:



Microsoft 365 Copilot: Summary

Prompting do's and don'ts

Get the most out of Copilot and avoid common pitfalls by learning **what to do** and **what not to do** when writing prompts.

Do's

Be clear and specific.
Provide specific instructions to Copilot, such as topic, purpose, tone, and required length.

Keep it conversational.
Give feedback to Copilot based on the quality of its responses to help the AI learn and match your preferences.

Give examples.
Use clear and specific keywords or phrases when asking Copilot to write a piece of text for you. This helps it generate more relevant and creative copy.

Ask for feedback.
Requesting feedback from Copilot helps it to understand your needs and preferences, and to provide you with more relevant, helpful responses.

Write legibly.
Use correct punctuation, capitalization, and grammar when writing prompts, as this will help the AI produce better quality text and responses.

Check for accuracy.
Occasionally, Copilot may make mistakes. Always check Copilot's responses for accuracy, grammar, and style, and watch out for irrelevant or inappropriate content.

Provide details.
Provide Copilot with contextual details to help it generate more accurate, consistent responses. For example, the genre, characters, and plot to a story.

Be polite.
Using kind and respectful language when chatting with Copilot helps foster collaboration and improves the AI's responsiveness and performance.

Don'ts

Be vague.
When prompting Copilot, avoid using vague language, and be as clear as possible to receive better-quality responses.

Request inappropriate or unethical content.
Copilot is not responsible for the content or the consequences of your writing. You should respect local laws, rules, and the rights of others.

Use slang, jargon, or informal language.
This may cause Copilot to give low-quality, inappropriate or unprofessional responses.

Give conflicting instructions.
Prompting Copilot to perform a task that includes multiple or conflicting pieces of information in the same request can confuse the AI and result in lower quality responses.

Interrupt or change topics abruptly.
This could disrupt Copilot's writing process. Always close or finish a task before starting a new one. When starting a new task, write "New task."

Microsoft Copilot M365:

Review prompting best practices

Craft clear and specific prompts for optimal Copilot performance. Here are some best practices:

Provide clear and concise prompts

- Make your request unambiguous
- Keep it brief without sacrificing clarity
- Use positive instructions and "if-then" statements
- Give examples for desired output, leveraging AI's imitation capability

Experiment with different styles

- Get creative with tone, language, and styles
- Explore analogies, poems, or historical allegories for diverse results

Give Copilot a point of view from which to answer

- Give Copilot a context by providing a point of view
- Roleplay scenarios or ask for responses in a specific style or persona

Know what to avoid when creating a prompt

- Avoid vague or overly general prompts
- Don’t cram too many questions into one prompt
- Don't assume Copilot has the context

Microsoft Copilot M365:

Review prompting best practices (continued)

Craft clear and specific prompts for optimal Copilot performance. Here are some best practices:

Understand Copilot's limitations

- Break down complex tasks for better results
- Optimize for repetition in simple, repetitive tasks
- Be aware of potential misinterpretation of ambiguous prompts
- Aim for clarity
- Review and validate Copilot's responses

Be polite

- Follow basic etiquette to generate respectful, collaborative outputs
- Practice politeness to enhance AI responsiveness and performance
- Start prompts with "please" and express gratitude

Be curious

- Ask preliminary questions
- Use curiosity to better understand problems and provide specific instructions

Iterate, iterate, iterate

- Don't give up after one prompt; it's a dynamic process
- Provide feedback for improvement
- Experiment, iterate, and refine prompts until satisfied

Copilot for Microsoft 365 - demo

Microsoft Copilot M365 resource links: https://learn.microsoft.com/en-us/copilot/

Microsoft Copilot https://learn.microsoft.com/en-us/copilot/overview
Find documentation about Microsoft Copilot with enterprise data protection.

Microsoft 365 Copilot https://learn.microsoft.com/en-us/copilot/microsoft-365/
Learn about Microsoft 365 Copilot and how your organization can use this copilot for work.

Copilot for Security https://learn.microsoft.com/en-us/copilot/security/
Use Copilot for Security to bring the full power of OpenAI architecture to defend your organization at machine speed and scale.

Copilot in Azure (preview) https://learn.microsoft.com/en-us/azure/copilot/
Learn how to manage operations from cloud to edge with an AI assistant.

Copilot experiences in Dynamics 365 https://learn.microsoft.com/en-us/dynamics365/copilot/
Find information about Copilot and generative AI in Dynamics 365 apps, including training and documentation.

Copilot experiences for your industry https://learn.microsoft.com/en-us/industry/copilot/
Find information about Copilot and generative AI in Industry Solutions, including training and documentation.

Copilot experiences in Power Platform https://learn.microsoft.com/en-us/power-platform/copilot/
Find information about Copilot and generative AI in Power Platform, including training and documentation.

Microsoft Copilot Studio https://learn.microsoft.com/en-us/microsoft-copilot-studio/
Discover how to build AI-driven copilots easily with Microsoft Copilot Studio.

Microsoft Azure AI Studio https://learn.microsoft.com/en-us/azure/ai-studio/
Build cutting-edge, market-ready, responsible applications for your organization with AI.

GitHub Copilot https://docs.github.com/copilot/
Use GitHub Copilot to get autocomplete-style suggestions from an AI pair programmer as you code.

GitHub Copilot in Visual Studio https://learn.microsoft.com/en-us/visualstudio/ide/visual-studio-github-copilot-install-and-states
Learn about GitHub Copilot Completions and GitHub Copilot Chat in Visual Studio.

Miroslav Doncevic

MCyberSec, Grad Cert Cyber Security, Cert NIST CSF Practitioner

Managing Director

miroslav@dms-it.com.au

DMS
DIGITAL MEDICAL SYSTEMS
CONFIDENT IT

www.dms-it.com.au

Resources for Primary Care

Who can support you

- Your IT provider
- Your accreditation agency
- Your indemnity insurer

AHPRA - Meeting your professional obligations when using Artificial Intelligence in healthcare

RACGP - Artificial Intelligence (AI) Scribes

ACCRM – Artificial Intelligence in healthcare

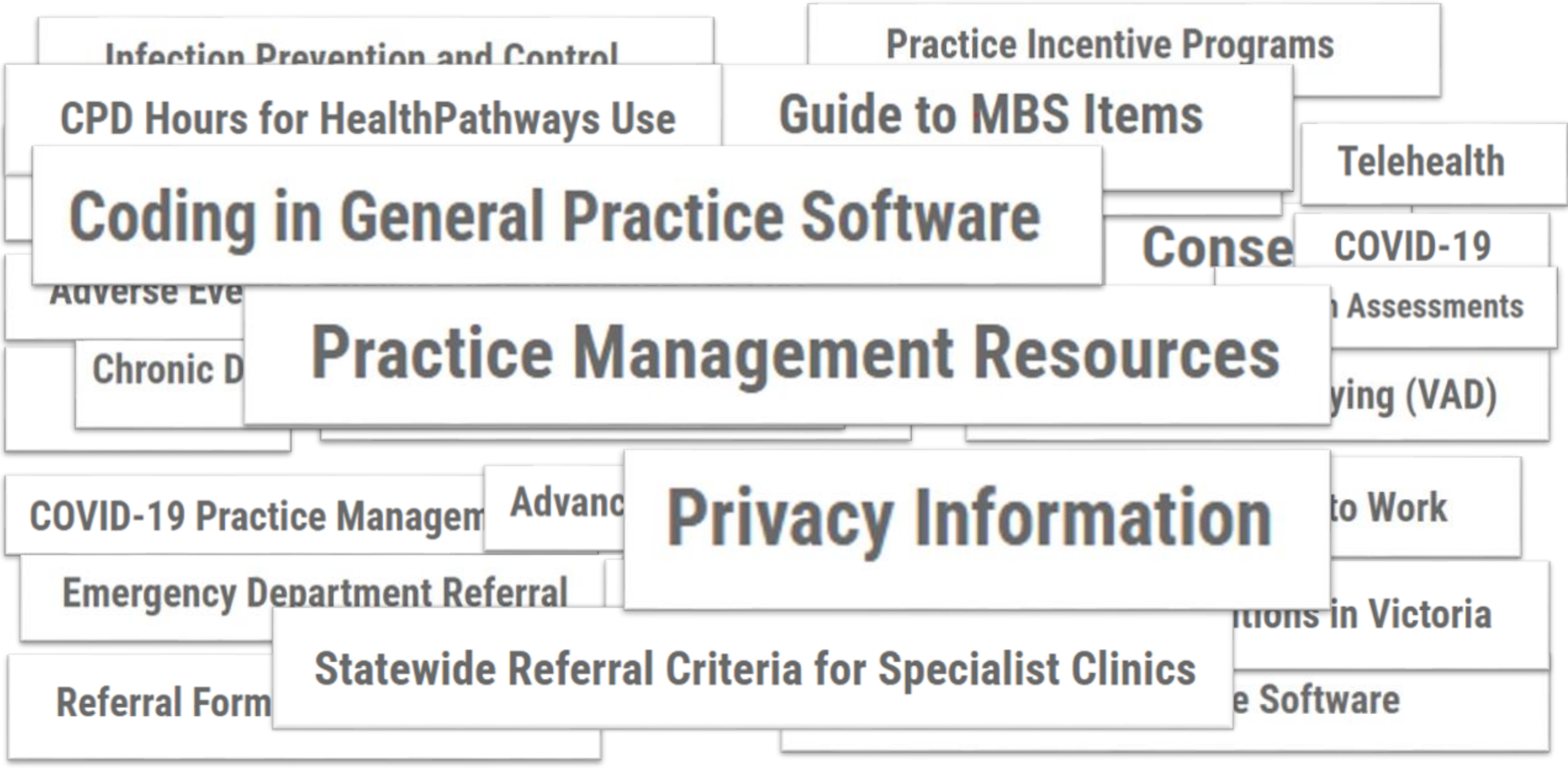
Australian College of Nursing – Artificial Intelligence position statement

HealthPathways Melbourne

Pathways are written by GP clinical editors with support from local GPs, hospital-based specialists and other subject matter experts

- clear and concise, evidence-based medical advice
- Reduce variation in care
- how to refer to the most appropriate hospital, community health service or allied health provider.
- what services are available to my patients

HealthPathways – Everything you need!



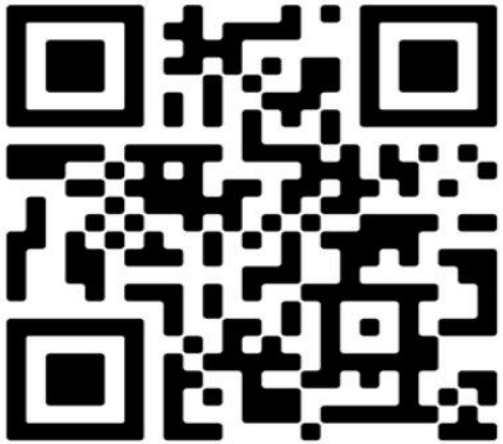
Practice Management Resources

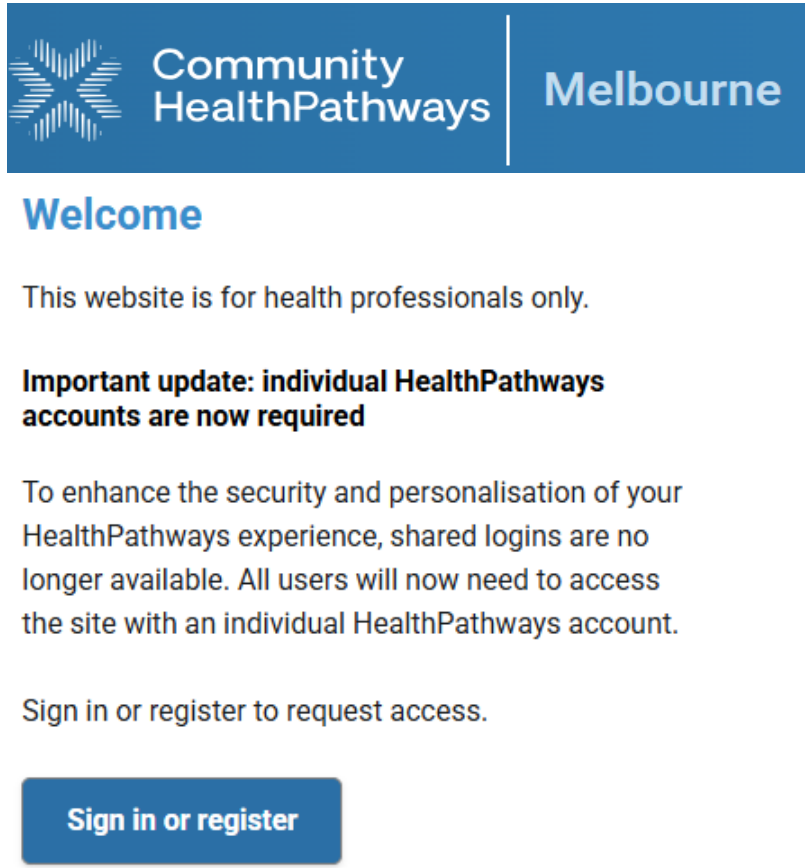
Aboriginal and Torres Strait Islander health
After hours services
Disaster planning and management
Forms templates and useful links
Health assessments and care plans
Immunisation and vaccines
Infection prevention and control
Legal and ethical
Managing drugs in general practice
MBS items
Practice Incentive Programs

Accessing HealthPathways

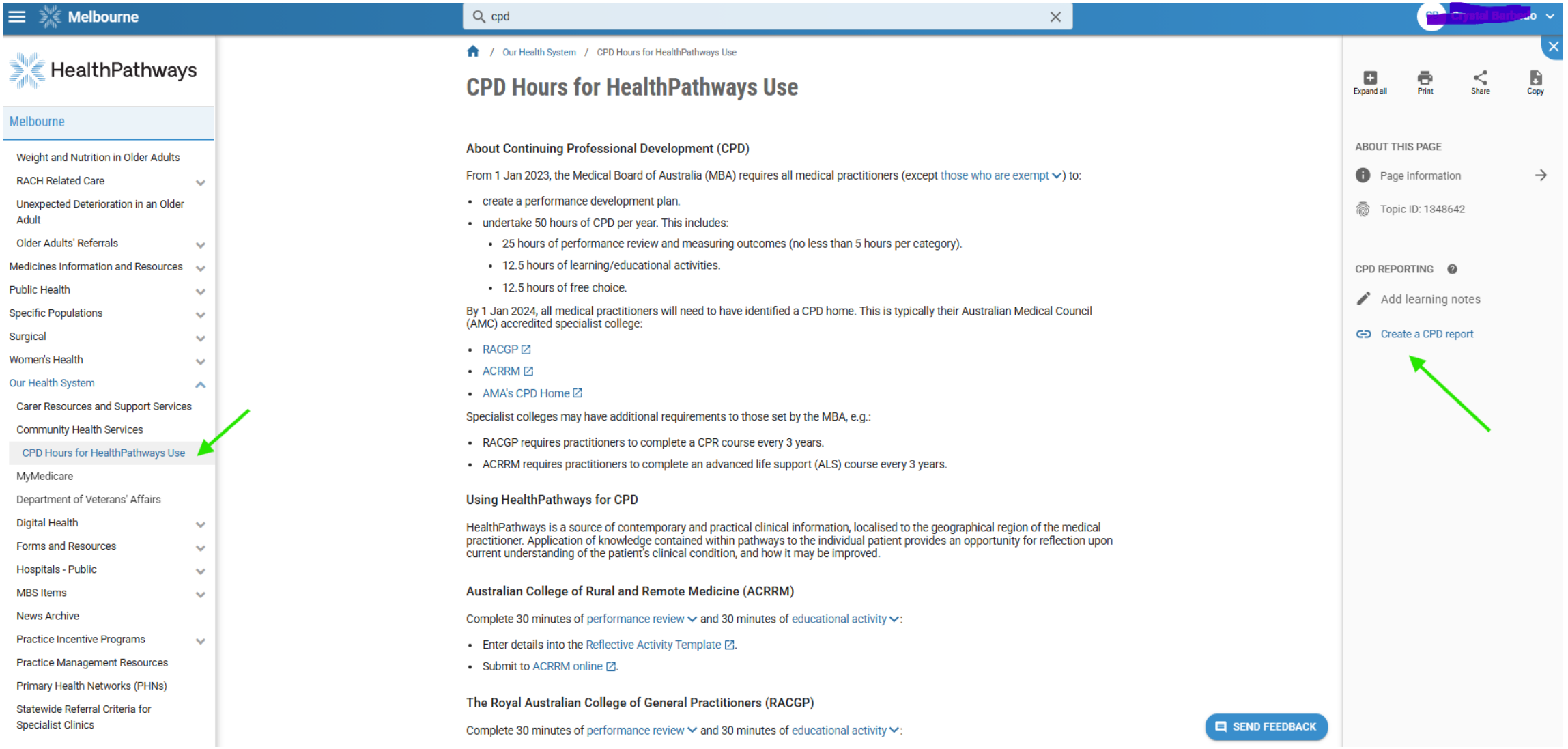
Please click on the **Sign in or register** button to create your individual account or scan the QR code below.

If you have any questions, please email the team info@healthpathwaysmelbourne.org.au.





HealthPathways – CPD Hours for HealthPathways Use



Questions?

Session Conclusion

We value your feedback, let us know your thoughts.

Scan this QR code

You will receive a post session email within a week which will include slides and resources discussed during this session.

Attendance certificate will be received within 4-6 weeks.

RACGP CPD hours will be uploaded within 30 days.

To attend further education sessions, visit,

https://nwmphn.org.au/resources-events/events/

This session was recorded, and you will be able to view the recording at this link within the next week.

https://nwmphn.org.au/resources-events/resources/