

AI impacts on the profession

Let's chat about ChatGPT and other AI and how these tools will likely impact on the profession of Librarianship.

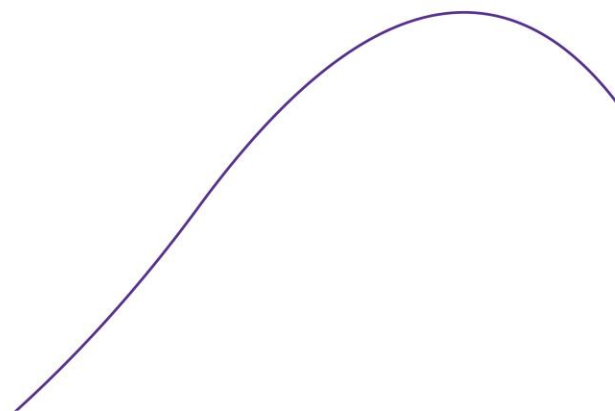
Presenter: Michelle Pitman – Grampians Health Library

Asia-Pacific Health, Law and
Special libraries conference,

13-15 June 2023



What is Large Language Model AI?





Hi ChatGPT



Hello! How can I assist you today?



Can you paraphrase the following quote from Vaswani et al (2017) into a Grade 9 reading level, please? "An attention function can be described as mapping a query and a set of key-value pairs to an output, where the query, keys, values, and output are all vectors. The output is computed as a weighted sum of the values, where the weight assigned to each value is computed by a compatibility function of the query with the corresponding key."



Sure! Here's a paraphrased version of the quote that should be more suitable for a Grade 9 reading level:

"An attention function is like a map that connects a question and a group of information to find an answer. Imagine the question, information, and answer as different paths on a map. The answer is determined by adding up the information, but each piece of information has a different importance, or weight. This weight is calculated based on how well the question matches with each piece of information."

Please note that while I have aimed to simplify the language, some concepts may still be challenging for a Grade 9 reader. If you need further clarification, feel free to ask!

Composition of the Pile by Category

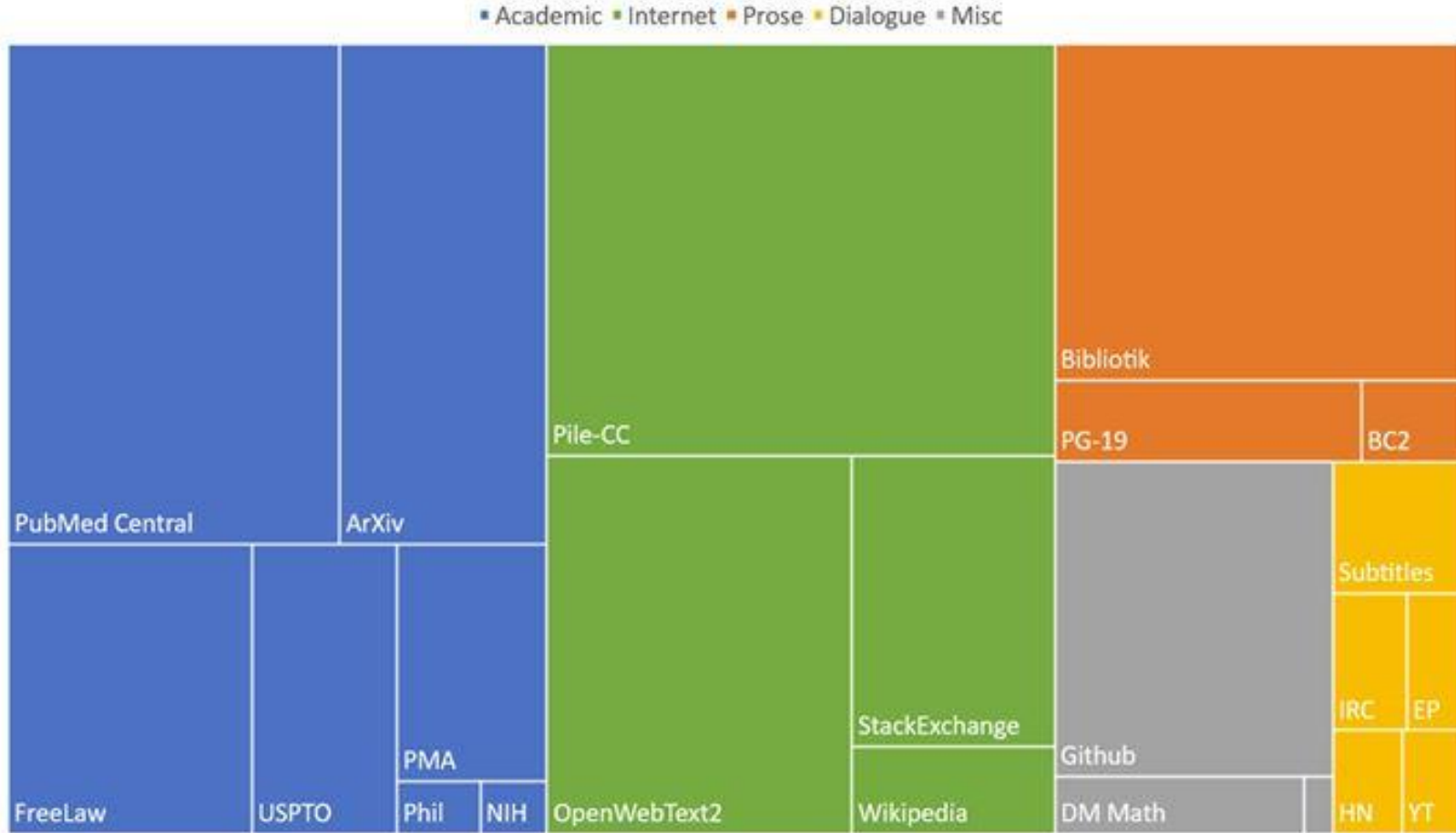


Figure 1: Treemap of Pile components by effective size.

Gao, L. et al (2020). The Pile: An 800GB dataset of diverse text for language modelling [Preprint]. ArXive.org. <https://doi.org/10.48550/arXiv.2101.00027>



Unknown. (2016, March 15). "Her Theodore GIF." <https://tenor.com/view/her-theodore-joaquin-phoenix-scarlett-johansson-samantha-gif-5203383>

The world's now awash with LLM AI tools.

OpenAI - GPT 3 and GPT 3.5, ChatGPT, ChatGPT Plus; GPT4 and InstructGPT

Google – LaMDA, Bard & BERT

Meta – RoBERTa & ESMFold

Then there's also:

EleutherAI, Cohere, Alpa, InstructGPT, Azure, NLP Cloud, AI21 Studio, XLNet, NVIDIA NeMo, NVIDIA MegatronLM, GooseAI, FreedomGPT, LUIS (Microsoft), Claude, Sparrow, Manubot, SciNote, Grover, Meena, Turing-NLG,

Text to speech generators:

Lovo.ai,
speechify,
Murf,
Speechmaker,
Listnr,
Synthesys,
Play.ht,
Sonantic,
Woord,
Voicera,

But wait... there's more!

Text to image generators:

DALL·E & DALL·E 2:

Bing Image Creator - based on DALL·E

Stable Diffusion

GetIMG (based on StableDiffusion),

NightCafe,

Deep Dream Generator,

Art Breeder,

Stablecog,

DeepAI,

StarryAI,

CF Spark,

Fotor,

Runway ML,

WOMBO Dream,

Specialty area AI tools

Law

Contract Review,
Contract Analytics,
Litigation Prediction,
Everlaw,
DISCO,
Catalyst,
Exterro,
Brainspace
Discovery,
DoNotPay,
Automio,
BillyBot

Education

Perplexity,
Curipod,
Education Co-Pilot,
Yippity:
Quillbot (*amazing!*):
Grammarly,
Conker,
Hemmingway Editor (*extremely
useful!*).

Healthcare

Merative (formerly IBM Watson Health)
Viz.ai
Enlitic & Enlitic Curie™
Regard
Evidence Hunt
Twill
Elicit
Note – Health Informatics/AI in medicine
journals are growing in number and
relevance.

First: the threats

Ethics – AI risks identified by Center for AI Safety (<https://www.safe.ai/ai-risk>)

1. Weaponization
2. Misinformation
3. Proxy Gaming
4. Enfeeblement
5. Value lock-in
6. Emergent goals
7. Deception
8. Power-seeking behaviour



Image generated by DALL-E, 30/03/2023 from the prompt "A Health Librarian who hates ICT"

These risks include:

- Ethics – governance & policy.
 - Data errors & hallucinations
 - Deliberate & unintended misinformation,
 - Inherent biases including gendered, information, racial profiling and so on.
 - Obfuscation about pre-training data including manipulating AI outputs to control humans
 - Privacy & data security
 - Copyright and Intellectual Property Rights
 - Risk of eliminating human jobs & dramatically changing how we work.
 - Also a risk of humans becoming unable to start or complete some types of tasks we've always previously been able to do
-

Other threats include:

- Climate impact of energy footprints
 - Anthropomorphizing the Machine – AKA what we end up feeling for and believing about these machines.
 - Cost of investment, implementation, governance
 - Commercialisation of AI will make it difficult to control
 - AI-written papers.
 - Targeting & steering users/consumers through curated recommendation algorithms
 - Malware, phishing, cyberattack
 - Evading CAPTCHA
 - Infringement of IP and copyright when training AI
-

Opportunities for Health, Law & Special Libraries

Can AI tools help maximise the value of our profession as librarians?

- **Creative ideas generation across all library domains of interest**
 - **Spot retracted and spurious articles in literature search results.**
 - **Eliminate duplicates more efficiently**
 - **Professional Development opportunities**
 - **Analytics uses – great for tracking our stats.**
 - **Library service use impact analysis - faster and clearer?**
-

- **Streamlining, simplifying & supporting aspects of the scholarly publishing lifecycle**
 - **Efficiency improvements to scholarly review and ethics**
 - **Scholarly writing/research support**
 - **Faster identification of predatory, retracted or fake papers**
 - **Managing rote and repetitive reference queries and tasks**
 - **Empowering the profession to develop an expertise presence at governance and government policy levels.**
-

- **More services for our users – yet to be imagined!**
 - **Marketing and communications:**
 - **Collection management:**
 - **Prompt Engineering:**
 - **Easier to find stuff (we just need to make sure it is the “right stuff”)**
-

Possibilities -

AI will bridge some gaps between research and practice – but not all

- Increasingly transformative personalised medicine
 - Robotics – especially in medicine
 - Case analysis – both in Law and Medicine
 - Educational tools (by the gazillions!).
 - Identifying new drugs & vaccines
 - Medical trends analysis
 - Legal trends analysis
 - Outpatient monitoring apps & services
 - Genomic data analysis
 - Billing and accounting
 - Discovery and synthesis of data and information
 - Chatbots for accessibility and access
 - Apps including medical, legal and educational
 - Chronic illness management
 - Contracts & procurement management
 - Client and student records management
 - Student knowledge assessment
 - New Taxonomies & knowledge mapping
 - Billing and accounting
 - Discovery and synthesis of data and information
 - Pandemic preparedness at scale
 - Mental health support at public health level
-

Learn: Experiment: Guide:

- Research about AI and AI information Governance
 - Defining best-evidence corpora for training AI systems
 - Risk Management, ethics and AI standards development in Library services
 - Discovery & data synthesis
 - Managing rapid change in the research lifecycle and systematic review process.
 - Experimentation with AI in real world use cases (within excellence guardrails)
 - User design [UX]
 - New Classification schemas?
 - Delivering information and digital literacy to users - now more important than ever
 - Accreditation and professional development
 - Corpora authentication
-

Recommended reading & exploration

AI4LAM. (2022). Artificial Intelligence for Libraries, Archives & Museums [YouTube]. <https://www.youtube.com/@ai4lam120/videos>

Cox, A. (2021). Research Report: The impact of AI, machine learning, automation and robotics on the information professions. CILIP. <https://www.cilip.org.uk/page/researchreport>

Centre for AI Safety. (2023). 8 examples of AI risk [Weblog]. <https://www.safe.ai/ai-risk>

European Union Intellectual Property Office. (2022). Study on the impact of artificial intelligence on the infringement and enforcement of copyright and designs: Impact of technology deep dive report 1. EUIPO. https://euipo.europa.eu/tunnel-web/secure/webdav/quest/document_library/observatory/documents/reports/2022_Impact_AI_on_the_Infringement_and_Enforcement_CR_Designs/2022_Impact_AI_on_the_Infringement_and_Enforcement_CR_Designs_Full_Report.pdf

IFLA Artificial Intelligence SIG. (2022, December 29). 23 resources to get up to speed on AI in 2023 [Weblog]. International Federation of Library Associations and Institutions. <https://www.ifla.org/23-resources-to-get-up-to-speed-on-ai-in-2023/>

Narayanan, A., & Kapoor. S. (2023). AI Snake Oil [Weblog]. Substack.com. <https://aisnakeoil.substack.com/>

Thank you

References:

AI4LAM. (2022). Artificial Intelligence for Libraries, Archives & Museums [YouTube].

<https://www.youtube.com/@ai4lam120/videos>

IFLA Artificial Intelligence SIG. (2022, December 29). 23 resources to get up to speed on AI in 2023 [Weblog].

International Federation of Library Associations and Institutions. <https://www.ifla.org/23-resources-to-get-up-to-speed-on-ai-in-2023/>

Budler, L. C., Gosak, L., & Stiglic, G. (2022). Review of artificial intelligence-based question-answering systems in healthcare. *WIREs Data Mining and Knowledge Discovery* 13(2). <https://doi.org/10.1002/widm.1487>

Centre for AI Safety. (2023). 8 examples of AI risk [Weblog]. New York Times. <https://www.safe.ai/ai-risk>

Cox, A. (2021, May). The impact of AI, machine learning, automation and robotics on the information professions: a report for CILIP. CILIP, UK. <https://www.cilip.org.uk/page/researchreport>

Croft, D. (2023, May 18). OpenAI CEO concerned by the dangers of AI, calls for government regulation [Article].

CyberSecurity Connect. <https://www.cybersecurityconnect.com.au/policy/9069-openai-ceo-concerned-by-the-dangers-of-ai-calls-for-government-regulation>

References:

Dempsey, L. (2023, May 24). Generative AI and large language models: background and contexts [Weblog]. LorcanDempsey.net. <https://www.lorcandempsey.net/intro-gen-ai/>

Ditch That Textbook [Weblog]. (2023, May 2). 30 AI tools for the classroom [Web page]. <https://ditchthattextbook.com/ai-tools/>

European Union Intellectual Property Office. (2022). Study on the impact of artificial intelligence on the infringement and enforcement of copyright and designs: Impact of technology deep dive report 1. EUIPO. https://euipo.europa.eu/tunnel-web/secure/webdav/guest/document_library/observatory/documents/reports/2022_Impact_AI_on_the_Infringement_and_Enforcement_CR_Designs/2022_Impact_AI_on_the_Infringement_and_Enforcement_CR_Designs_FullR_en.pdf

Future of Life Institute. (2023, May 22). Pause giant AI experiments: An open letter. <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>

Future of Life Institute. (2023, April 19 Revised). Policymaking in the pause: what can policymakers do now to combat risks from advanced AI systems. https://futureoflife.org/wpcontent/uploads/2023/04/FLI_Policymaking_In_The_Pause.pdf

Gao, C. A., Howard, F. M., Markov, N. S., [...], & Pearson, A. T. (2023). Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers [PrePrint]. *bioRxiv.org*: 1-18. <https://www.biorxiv.org/content/10.1101/2022.12.23.521610v1>

References:

Gao, L., Biderman, S., Black, S. et al. (2020). The Pile: An 800GB Dataset of Diverse text for language modelling. *EleutherAI*. arXiv.org. <https://arxiv.org/pdf/2101.00027.pdf>

Gozalo-Brizuela, R., & Carrido-Merchán, E.C. (2023, January 11). ChatGPT is not all you need. A State of the Art Review of Large Generative AI Models [PrePrint]. *arXiv.org: Quantitative Methods Department, Universidad Pontificia Comillas, Madrid, Spain*. <https://doi.org/10.48550/arXiv.2301.04655>

Hugging Face. (2023). Natural Language Processing [Web page]. <https://huggingface.co/learn/nlp-course/chapter1/2?fw=pt%2031/3/2023>.

Kendrick, C. (2023, January 26). Guest Post – The efficacy of ChatGPT: Is it time for the Librarians to go home? The Scholarly Kitchen [Weblog]. <https://scholarlykitchen.sspnet.org/2023/01/26/guest-post-the-efficacy-ofchatgpt-is-it-time-for-the-librarians-to-go-home/>

Kung, T. H., Cheatham, M., ChatGPT., [...], & Tseng, V. (2022, December, 21). Performance of ChatGPT on USMLE: Potential for AI-assisted Medical Education Using Large Language Models [PrePrint]. *medRxiv.org*. <https://doi.org/10.1101/2022.12.19.22283643>

References:

Loyal, L. (2023, May 31). AI intensifies fight against 'paper mills' that churn out fake research. *Nature*: 1476-4687.
<https://doi.org/10.1038/d41586-023-01780-w>

Lund, B., & Ting, W. (2023, January 22). Chatting about ChatGPT: How May AI and GPT Impact Academia and Libraries? *Library Hi Tech News*. <http://dx.doi.org/10.2139/ssrn.4333415>

Mann, D. L. (2023). Artificial Intelligence discusses the role of Artificial Intelligence in Translational Medicine: A JACC: Basic to Translational Science interview with ChatGPT. *JACC: Basic to Translational Science* 8(2): 221-223.
<https://doi.org/10.1016/j.jacbts.2023.01.001>

National Institute of Standards and Technology [NIST]. (2023, January). Artificial Intelligence Risk Management Framework (AI RMF 1.0) [NIST AI 100-1]. U.S. Department of Commerce.
<https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf>

Narayanan, A., & Kapoor. S. (2023). AI Snake Oil [Weblog]. Substack.com. <https://aisnakeoil.substack.com/>

Newman, J. (2023, January). A taxonomy of trustworthiness for artificial intelligence: connecting properties of trustworthiness with risk management and the AI lifecycle [CLTC White Paper Series]. *Center for Long-term Cybersecurity*. UC Berkley. https://cltc.berkeley.edu/wp-content/uploads/2023/01/Taxonomy_of_AI_Trustworthiness.pdf?mc_cid=7c9372e38f&mc_eid=59e9a960b2

References:

Rajpurkar, P., Chen, E., Banerjee, O. et al. (2022). AI in health and medicine. *Nat Med* 28, 31–38 (2022).
<https://doi.org/10.1038/s41591-021-01614-0>

Ravi, N., Chaturvedi, P., Huerta, E.A. et al. (2022). FAIR principles for AI models with a practical application for accelerated high energy diffraction microscopy. *Sci Data*, 9(657): <https://doi.org/10.1038/s41597-022-01712-9>

Shanahan, M. (2022, December, Revised February 2023). Talking about large language models [PrePrint]. arXiv.org.
<https://arxiv.org/pdf/2212.03551.pdf>

Shen, Y., Heacock, L., Elias, J., Hentel, K.D., et al. (2023). [ChatGPT and Other Large Language Models Are Double-edged Swords](#). *Radiology* 2023, 307(2). <https://doi.org/10.1148/radiol.230163>

Simon, J. (2021). Large language models: a new Moore's Law? (Weblog). Hugging Face [Website].
<https://huggingface.co/blog/large-language-models>

Stokel-Walker, C., & Van Noorden, R. (2023). The promise and peril of generative AI. *Nature* 614: 214-216.
<https://doi.org/10.1038/d41586-023-00340-6>

Pividori, M., Greene, C. S. (2023, January 23). A publishing infrastructure for AI-assisted academic authoring [PrePrint]. bioRxiv.org. Creative Commons Attribution, CC BY-SA 4.0. <https://doi.org/10.1101/2023.01.21.525030>

References:

Vaswani, A., Shazeer, N., Parmar, N. et al. (2017). Attention is all you need. Proceedings from the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach California, USA. Advances in neural information processing systems. *ArXiv.org*. <https://arxiv.org/pdf/1706.03762.pdf>

Wikipedia. (n.d.). Existential risk from artificial general intelligence. https://en.wikipedia.org/wiki/Existential_risk_from_artificial_general_intelligence

Wolfram, S. (2023, February 14). What is ChatGPT doing... and why does it work? [Weblog]. Stephen Wolfram Writings. <https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/>

Yang, M. (2023, May 26). Scientists use AI to discover new antibiotic to treat deadly superbug. The Guardian. <https://www.theguardian.com/technology/2023/may/25/artificial-intelligence-antibiotic-deadly-superbug-hospital>

Zhou, H., & Hunter, S. I. (2023, June 5). Guest Post - Accessibility powered by AI: How Artificial Intelligence can help universalize access to digital content. The Scholarly Kitchen [Weblog]. https://scholarlykitchen.sspnet.org/2023/06/05/guest-post-accessibility-powered-by-ai-how-artificial-intelligence-can-help-universalize-access-to-digital-content/?informz=1&nbd=&nbd_source=informz

Zielinski, C., Winker, M., Aggarwal, R., Ferris, L., Heinemann, M., Lapeña, J., Pai, S., Ing, E., Citrome, L. (2023). Chatbots, ChatGPT, and Scholarly Manuscripts - WAME Recommendations on ChatGPT and Chatbots in Relation to Scholarly Publications. *Afro-Egyptian Journal of Infectious and Endemic Diseases*, 13(1), 75-79. <https://doi.org/10.21608/aeji.2023.282936>