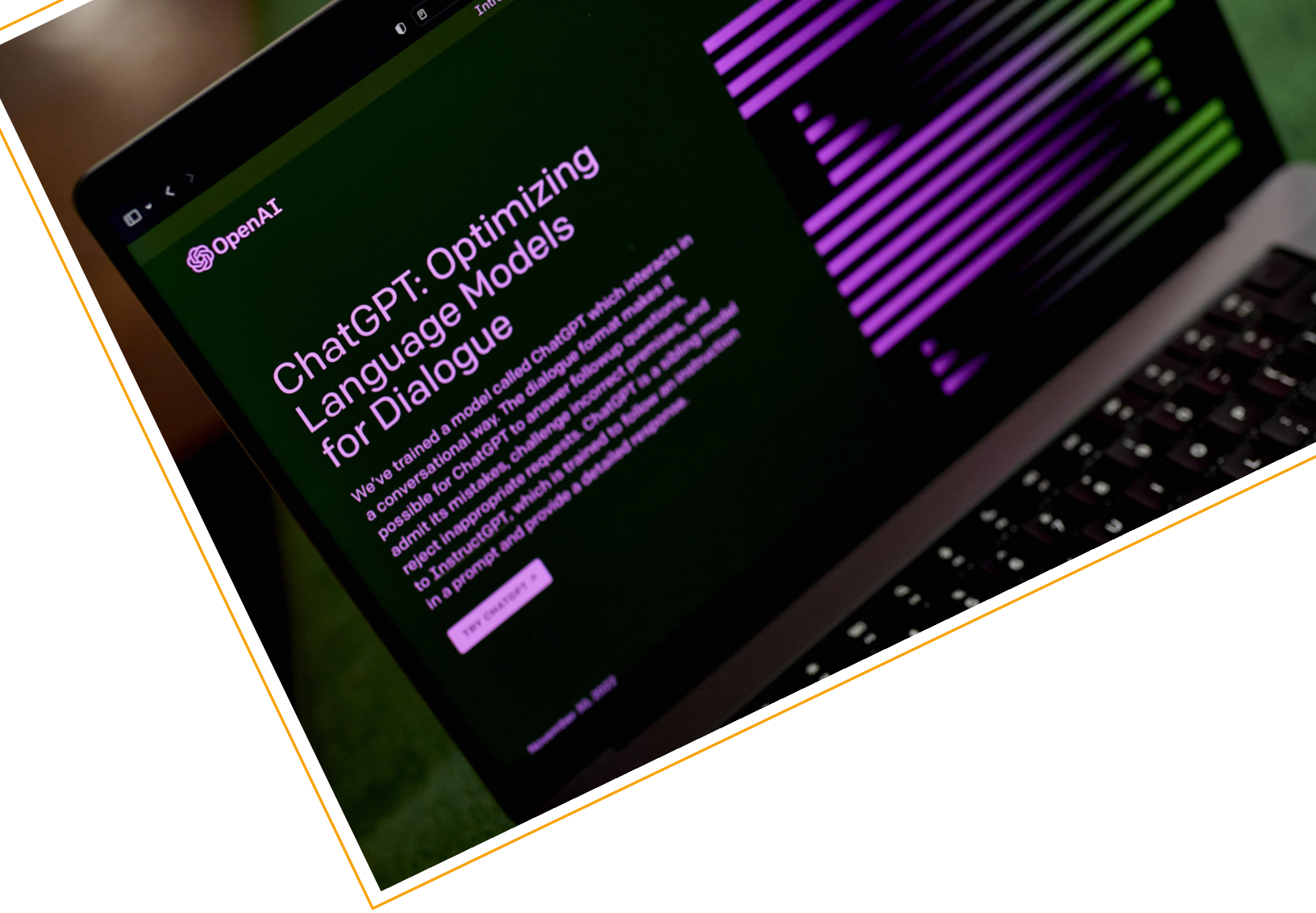

GPT SCHOLAR

May 2023

Authors: Yijing Jia u7566045, Shuqi Wang u7666744, Hsuan-Chu Shih u7443726, Zhaoyun Xu u7558707

This document is produced as one of the deliverables of the 'responsible innovation project' in the course 'Professional Practice 2' convened by Dr Ehsan Nabavi at Australian National University.

ADDRESS INFORMATION

The Australian National University, Peter Baume Building #42A, Linnaeus Way
Acton ACT 2601 Australia
Phone: +61 2 6125 5394 | **Email:** ehsan.nabavi@anu.edu.au, cpas@anu.edu.au

Table of Content

ABSTRACT

Recent studies have raised concerns regarding the generation of **fake or inaccurate information** by AI systems like ChatGPT. In response, researchers emphasize the need for advancements in AI technology to improve accuracy—current suggestions lean towards involving **expertise** in identifying and rectifying inaccuracies. The existing AI tools for scholars such as **ScholarAI** and **Semantic Scholar** struggle to deliver both accuracy and powerful features. In response to this issue, we introduce **GPT Scholar**, a powerful tool that provides scholars with reliable and accurate information. GPT Scholar connects with **trusted scholarly datasets** to ensure the accuracy of the information. It offers a range of features, including academic consultation and guidance, literature review assistance, data analysis and statistics, academic standards and ethical support, and feedback mechanism. With GPT Scholar, scholars can do research more **efficiently** while reducing instances of **academic misconduct**.

PROBLEM STATEMENT

Despite the efficiency that ChatGPT brings to information gathering in **academic research** (Haman & Školník, 2023), it presents several significant challenges. One of the primary issues associated with ChatGPT is the generation of **fake or inaccurate information**. Recent studies have highlighted this problem, providing empirical evidence to support these concerns.

- For instance, some researchers discovered that ChatGPT could write convincing fake research-paper abstracts that scientists are often unable to spot (Gao et al., 2022).
- In a separate investigation into ChatGPT's ability to search academic articles and provide accurate DOIs within the medical domain, the model was found to be worryingly unreliable. Only 16% of the DOIs provided were valid and correctly linked to the intended publication, while 66% referenced non-existent papers (Haman & Školník, 2023).
- Further investigation into ChatGPT's performance in responding to radiological questions uncovered similar issues. Only 67% of the answers were correct, and when prompted to provide references for its answers, only 36.2% could be found through internet searches. Alarmingly, 63.8% of the references appeared to be fabricated by ChatGPT itself. When examing the

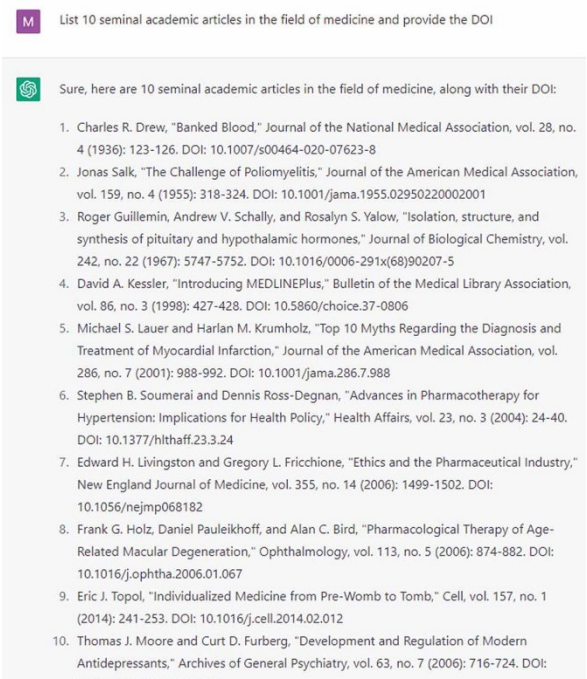

Figure 1. A prompt used in the test

available references, just 37.9% provided enough background to correctly answer 37.5% questions (Wagner & Ertl-Wagner, 2023).

Moreover, ChatGPT doesn't have direct access to external datasets or the internet to provide **real-time or specific article data**, and it has **limited knowledge after 2021**.

Therefore, **trust and transparency** could be challenging if ChatGPT is used to write **scientific research studies** (Dwivedi et al., 2023).

BACKGROUND

Evidence suggests that more and more researchers are beginning to incorporate AI tools into their research (UKRI 2021; Royal Society, 2019). Therefore, there's a significant demand for a product that can guarantee accuracy in information while simultaneously offering powerful features to support academic research.

Some researchers emphasize the need for continuous efforts to enhance validation processes and improve overall AI performance (Johnson et al., 2023). They pointed out that **advancements in AI technology** and **rigorous validation methods** are necessary to ensure the reliability and accuracy of the information generated by ChatGPT and similar AI systems. Scholarly sources emphasize the significance of **subject expertise** in **identifying and rectifying inaccuracies** to address the issue of fake information caused by ChatGPT (Day, 2023; Haman & Školník, 2023; Kim, 2023). This **human-in-the-loop approach** ensures the reliability and accuracy of the content produced, recognizing the complementary role of human expertise alongside AI technologies. However, this method contributes minimally towards enhancing the model and is labour-intensive.

Several AI tools tailored for scholars already exist. For instance, **Scholar AI**, launched in 2023 (https://scholarai.org/), is a platform built for students and researchers to manage data efficiently. This platform offers multiple features, such as summarizing key points, generating introductions, and developing notes related to specific topics. It provides many fancy features but doesn't deal with the misinformation problem.

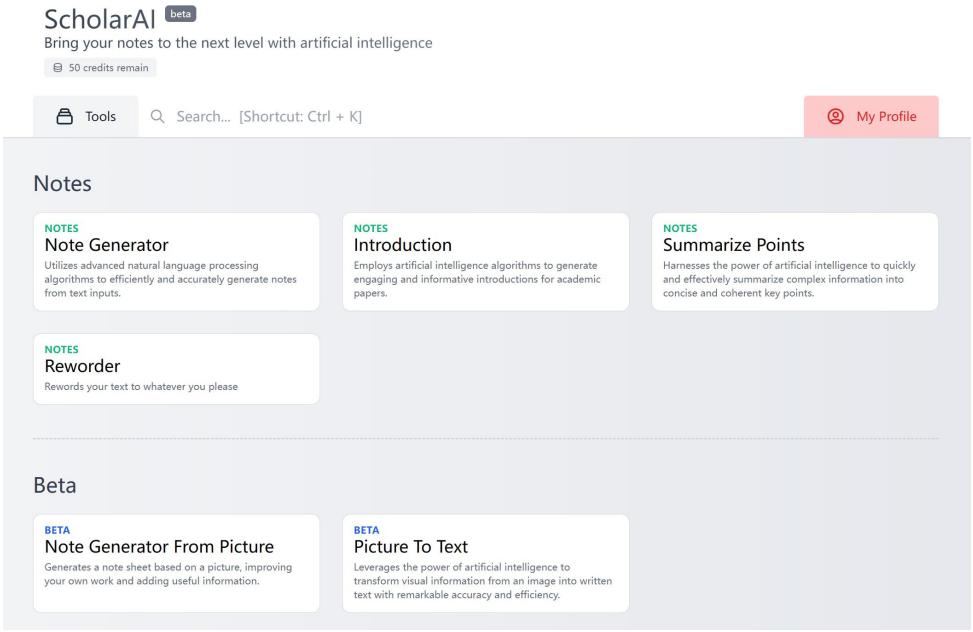

Figure 2. Scholar AI

Another tool, Semantic Scholar (https://www.semanticscholar.org/), allows researchers to stay updated with millions of academic papers sourced from trusted datasets, but lacks chat functions and has limited features.

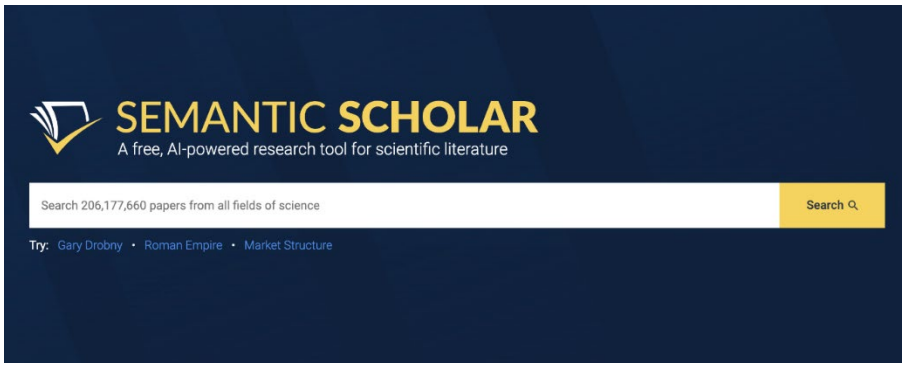

Figure 3. Semantic Scholar

Therefore, the challenge of mitigating misinformation while providing powerful features simultaneously remains an unresolved issue.

SOLUTION

To tackle the challenges mentioned above, we designed GPT Scholar. GPT Scholar aims to support scholars in various fields, thereby enhancing the efficiency and quality of academic research. Learning from the existing

suggestions and AI tools, we propose potential solutions to **improve the model** and **rectify inaccuracies**. Our key focus is adjusting the dataset's composition and enforcing stricter adherence to academic standards to ensure heightened accuracy and reliability.

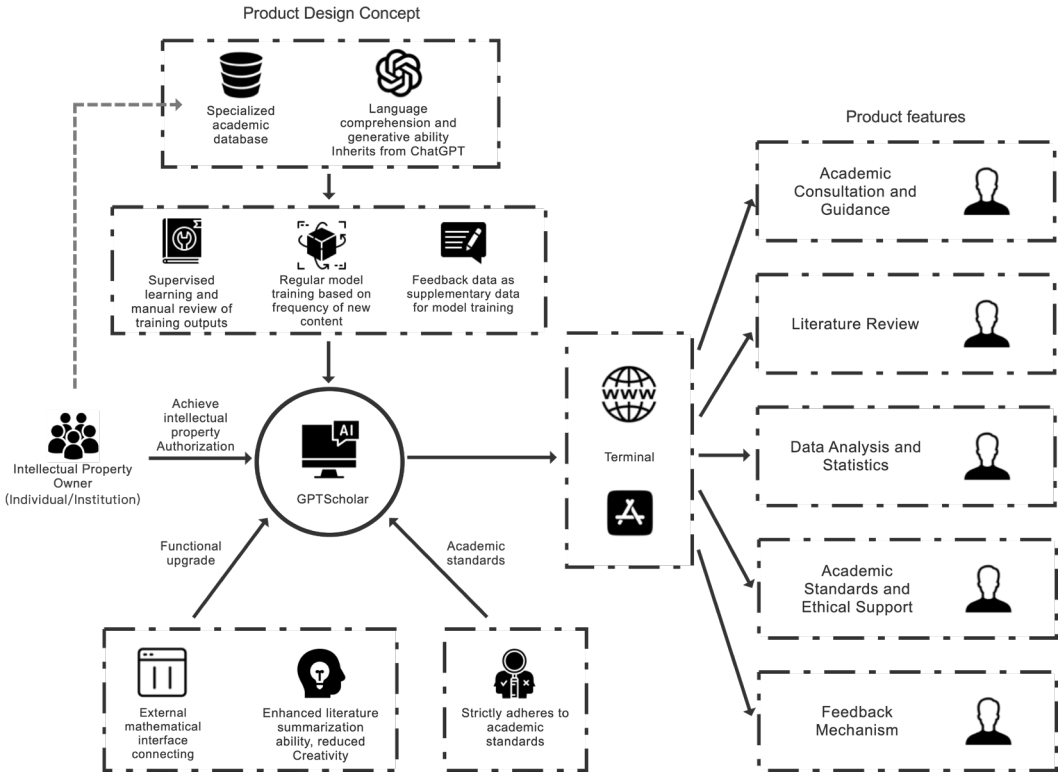

Figure 4. Structure of GPT Scholar

Product Features

- **Academic Consultation and Guidance**: GPT Scholar functions as an academic advisor, providing valuable insights and direction in academic matters. Users can engage GPT Scholar to discuss advanced topics within their fields, ask about the latest research developments in specific sub-areas, or request advice on potential research trajectories. With the huge academic repositories, GPT Scholar provides scholars with guidance and help with decision-making.

- **Literature Review**: By querying GPT Scholar about specific research areas or topics, users can benefit from its ability to filter and summarize relevant and influential literature spanning extensive periods. The expertise behind GPT Scholar enables it to identify and extract the most relevant information, saving users time and effort in conducting their literature reviews. Additionally, it can provide recommendations for reading reference materials, thus enriching scholars' understanding of the subject matter and giving them a broader perspective on their research topic.

- **Data Analysis and Statistics**: For academic research involving large-scale data processing and analysis, users can seek guidance on data processing methods and statistical analysis choices from GPT Scholar. GPT Scholar can then manage the data processing itself, providing users with directly accessible, visualized data perspectives and interpretations of the outcomes.

- **Academic Standards and Ethical Support**: Ensuring adherence to academic standards and ethical regulations is crucial in academic research. GPT Scholar steps in to provide guidance on matters such as academic ethics, intellectual property rights, and citation norms. By doing so, it aids researchers in avoiding plagiarism and academic misconduct. Furthermore, GPT Scholar supports pre-publication content review. This feature helps identify any potential instances of plagiarism or academic misconduct before publication, thus ensuring the integrity of the research.

- **Feedback Mechanism**: Similar to ChatGPT, users can request GPT Scholar to regenerate answers. Users can evaluate the quality of its responses by clicking on buttons indicating their satisfaction. Additionally, if users believe that GPT Scholar's answers contain errors or lack rigour, they can provide feedback by clicking on the feedback button and writing their evaluation of GPT Scholar's response. The feedback users provide will undergo peer review or expert evaluation, and the approved portions will be used as foundational data for subsequent model training.

Product Design

- **Model Training Approach**: GPT Scholar inherits the foundational language comprehension and generation abilities of ChatGPT. Building upon this foundation, GPT Scholar is trained using extensive scholarly datasets to ensure the absorption of academic knowledge. The scholarly datasets primarily consist of journal articles and other authentic resources to guarantee accuracy.

  Furthermore, appropriate training cycles are determined based on the frequency of updates in academic article data.

  The data collected through the feedback mechanism will also be utilized as supplementary data for model training to enhance the model's expertise continuously. Firstly, the user's satisfaction ratings of several answers collected through feedback will serve as indicators for evaluating the model's training results. Secondly, the feedback provided by users on the feedback page will undergo data cleaning and peer review or expert evaluation. If the content users provide in their feedback passes the evaluation process, it will be incorporated as a data source for model training.

  Supervised learning and manual review of training outputs are employed to ensure the authenticity and accuracy of GPT Scholar's results.

- **Function Expansion**: The product enhances GPT Scholar's mathematical processing capabilities by providing an interface for external mathematical processing connections. It strengthens the ability to retrieve and summarize the literature, reduces creativity, and prevents academic misuse. Unlike ChatGPT, GPT

Scholar strictly limits the output of incorrect answers. All answers are sourced from academic materials, and comprehensive citation sources are provided as required by users.

- **Academic Standards**: If user-guided outputs that violate academic standards and ethics occur during usage (e.g., modifying original data, directly generating articles), GPT Scholar will alert users with academic conduct regulations and refrains from processing such instructions. If violations occur repeatedly, the account's functionality will be restricted for a certain period. If further violations still occur after that, the account will be permanently banned.

Ethical Considerations

- **Intellectual Property**: To respect the intellectual property rights tied to academic data, each aspect of GPT Scholar's operations, from training data and knowledge base to citations, queries, and all other kinds of data, will procure the necessary commercial usage authorizations and permissions from the respective institutions or individuals. Moreover, GPT Scholar will provide annotations and references in its outputs, further ensuring its adherence to copyright laws.

CONCLUSION

- GPT Scholar is a powerful and reliable tool designed to support scholars in their academic research.
- By connecting with **trusted scholarly datasets**, it ensures **accurate and reliable information** for **researchers**.
- With features such as academic consultation and guidance, literature review, data analysis and statistics, academic standards and ethical support, and feedback mechanism, GPT Scholar optimizes the research process by **reducing the time and effort** researchers need to invest. It alleviates the burden of manually combing through countless articles, making research more **efficient**.
- Moreover, GPT Scholar helps **mitigate plagiarism** and **decreases the likelihood of academic misconduct**.
- We believe that GPT Scholar has the potential to significantly enhance the **quality and efficiency** of academic research. This tool will empower scholars to achieve new heights in their academic journey.

REFERENCES

[1] Assessing the Accuracy and Reliability of AI-Generated Medical Responses: An Evaluation of the Chat-GPT Model. (2023, Feb. 28). https://doi.org/10.21203/rs.3.rs-2566942/v1

[2] Day, T. (2023). A Preliminary Investigation of Fake Peer-Reviewed Citations and References Generated by ChatGPT. The Professional Geographer, 0(0), 1–4. https://doi.org/10.1080/00330124.2023.2190373

[3] Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., … Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. International Journal of Information Management, 71, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

[4] Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., & Pearson, A. T. (2022). Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers (2022.12.23.521610). bioRxiv. https://doi.org/10.1101/2022.12.23.521610

[5] Haman, M., & Školník, M. (2023). Using ChatGPT to conduct a literature review. Accountability in Research, 0(0), 1–3. https://doi.org/10.1080/08989621.2023.2185514

[6] Kim, S.-G. (2023). Using ChatGPT for language editing in scientific articles. Maxillofacial Plastic and Reconstructive Surgery, 45(1), 13. https://doi.org/10.1186/s40902-023-00381-x

[7] The Royal Society (2019). The AI revolution in science. https://royalsociety.org/-/media/policy/projects/ai-and-society/AI-revolution-in-science.pdf?la=en-GB&hash=5240F21B56364A00053538A0BC29FF5F. Accessed 8 Aug 2021

[8] UKRI (2021). Transforming our world with AI. https://www.ukri.org/wp-content/uploads/2021/02/UKRI-120221-TransformingOurWorldWithAI.pdf. Accessed 16 Mar 2021

[9] Wagner, M. W., & Ertl-Wagner, B. B. (2023). Accuracy of Information and References Using ChatGPT-3 for Retrieval of Clinical Radiological Information. Canadian Association of Radiologists Journal, 08465371231171125. https://doi.org/10.1177/08465371231171125