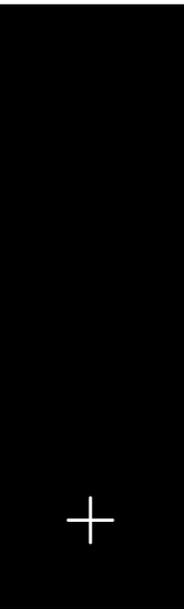




AI輔助工具在 學術研究中的應用

數位與系統組 張素卿

112.11.22



課程大綱

研究者常面臨的問題

AI輔助工具：Semantic Scholar

AI輔助工具：Connected Papers

AI輔助工具：ChatGPT

實作練習



課程大綱

研究者常面臨的問題

AI輔助工具：**Semantic Scholar**

AI輔助工具：**Connected Papers**

AI輔助工具：**ChatGPT**

實作練習



知識爆炸時代的挑戰



資料太多



可用的有限...



亂槍打鳥，見樹不見林！



研究者常面臨的問題

- 文章數量無法掌握
- 文章理解時間不足
- 文章追蹤工具缺乏
- 文章閱讀反覆失焦

課程大綱

研究者常面臨的問題

AI輔助工具：Semantic Scholar

AI輔助工具：Connected Papers

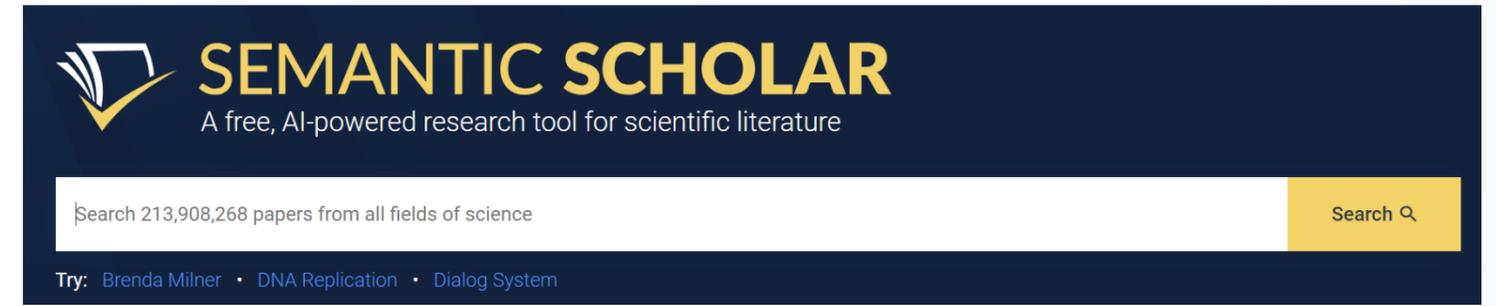
AI輔助工具：ChatGPT

實作練習



Semantic Scholar

<https://www.semanticscholar.org/>



Semantic Scholar 創立於2015年，由微軟創辦人旗下Allen Institute for AI所主持，為一個非營利組織網站。特色為基於人工智慧技術來提供學術搜索，平台上收藏超過兩億篇科學文獻。

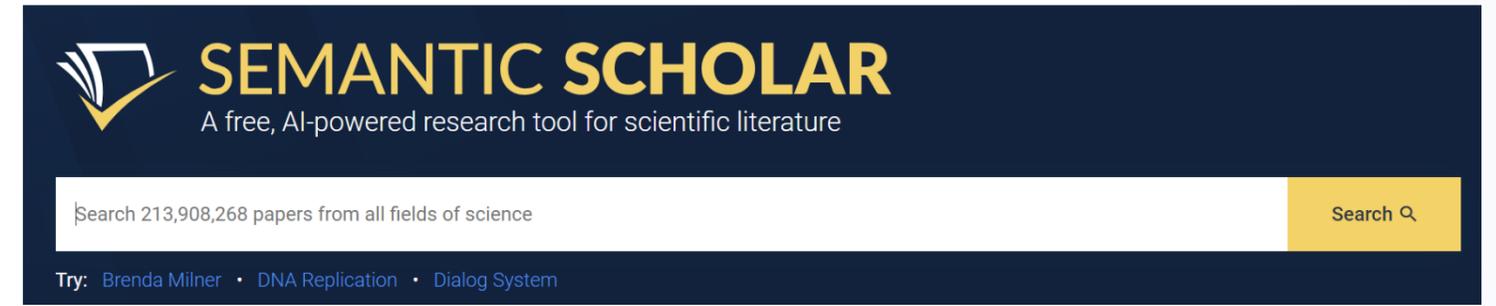
它主要是使用AI驅動的演算法提供更相關和更具影響力的搜尋結果。

利用Semantic Scholar可以非常輕鬆地找到某領域的高引用論文，然後接著找到這篇高引論文的高引論文，從而順藤摸瓜發現領域的後續發展脈絡。

- ✓ 協助分析文章影響力。
- ✓ 自動產生文章摘要。
- ✓ 登入後，根據用戶收藏的論文，可根據AI算法進行相關論文推薦，提供自適應研究推送（adaptive research feed），方便用戶查看相關研究領域之最新或重要論文。
- ✓ 可自動顯示並提供連結至相關引用文獻。

Semantic Scholar

<https://www.semanticscholar.org/>



Mapping Research Trajectories

[Bastian Schafermeier](#), [Gerd Stumme](#), [Tom Hanika](#) · Computer Science · ArXiv · 2022

TLDR This work proposes a principled approach for mapping research trajectories, applicable to all kinds of scientific entities that can be represented by sets of published papers, and transfers ideas and principles from the geographic visualization domain, specifically trajectory maps and interactive geographic maps. [Expand](#)

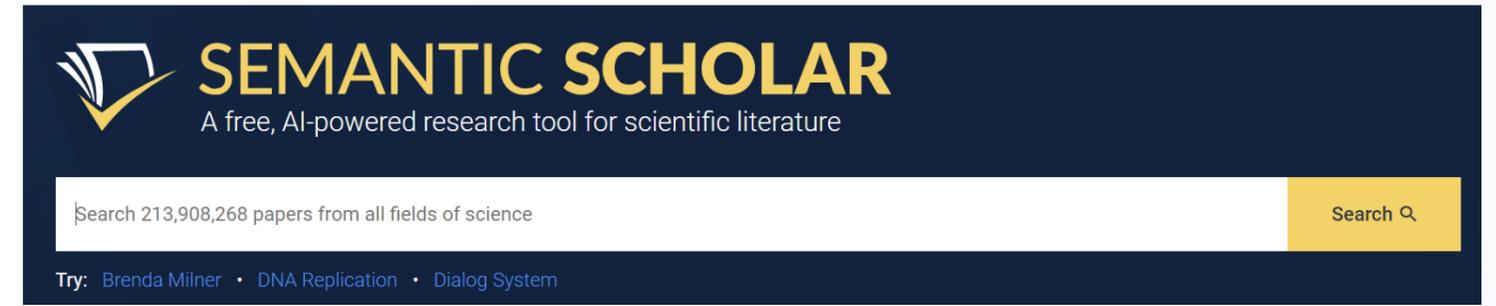
Highly Influenced

PDF

View 7 excerpts, cites results, background and methods Save Alert

Semantic Scholar

<https://www.semanticscholar.org/>



Some Salient Aspects of Machine Learning Research: A Bibliometric Analysis

S. Bhattacharya · Computer Science · J. Sci. Res. · 2019

TLDR The paper tries to capture the intellectual structure of this field and research trends from quantitative and statistical analysis of research publications and develops a conceptual framework for the study. [Expand](#)

13 View 2 excerpts, references methods Save Alert

✔ 可自動顯示並提供連結至相關引用文獻。

Table Of Contents

- 1 Introduction
- 2 Topic Space Trajectories
 - 2.1 Publication Corpus
 - 2.2 Document Representation in Topic Space
 - 2.3 Non-Negative Matrix Factorization
 - 2.4 Venue Representation and Trajectories in Topic Space
 - 2.5 Calculating an Optimal Topic Number
- 3 Case Study on Machine Learning Publications
 - 3.1 Dataset & Filtering
 - 3.2 Preprocessing
 - 3.3 Matrix Factorization
 - 3.4 Optimal Topic Number
 - 3.5 Abstract Length & Term Distribution
- 4 Interpreting Topic Space Trajectories
 - 4.1 Found Topics
 - 4.2 Topic Similarities
 - 4.3 Overall Historical Interest in Topics

We may note that *papers*, *venues*, and *year* are just names. Hence, a publication corpus can be substituted by any structure bearing the same construction, e.g., news articles in certain newspapers on certain days.

2.2 Document Representation in Topic Space

Say the number of papers (documents) in a publication corpus is d and the set of all words from all these documents is of size n . There are various methods for embedding these documents into a topic space (commonly referred to as *topic models*). Most of them employ a real-valued *word-document* matrix as starting point. Hence, we need to represent all elements $p \in P$ as elements of $\mathbb{R}_{\geq 0}^n$. Therefore we consider in the following p to be an element of $\mathbb{R}_{\geq 0}^n$. In this representation each vector component denotes a word-weight. Such a weight for a particular word can be, for example, the word-frequency (*term-frequency*) or a more sophisticated approach like *tf-idf* (Ramos et al., 2003). Having the vector representations of all publications we can construct a word-document matrix V simply by juxtaposing all those vectors. A representation of that word-document matrix V as a matrix of factors in the topic space, foremost the topics and some representing the documents, is possible. These factors represent particular structural information in this space, i.e., as a linear combination of *topic vectors*. Here is that the number of topics t is substantially smaller than n .

A very prominent topic model, called *Latent Semantic Analysis* (LDA) (Blei et al., 1999), is based on the singular value decomposition of the word-document matrix. This method has been proven to work well for natural language processing tasks, e.g., in Steinberger and Blei (2006) or for recommendations of scientific publications (Yang (2006)). Nonetheless, we decided against its application for this is that the results are difficult to comprehend. This problem arises from the fact that topics in LSA are not ordered and negatively to the topical document representation. A better method is *latent Dirichlet allocation* (LDA) from Blei et al. (2003) on a probabilistic model where documents are assumed to be a mixture distribution over topics, which themselves are distributed over words. LDA generally achieves very good results, e.g., in text classification (Blei et al., 2003) it is known to fall short for small documents like papers (Blei et al., 2003) for example been shown empirically on Twitter posts (Blei et al., 2003). Since we do prefer for our modeling a method that works well with small documents we discard using LDA. The authors have investigated methods based on *doc2vec* (Le and Mikolajczyk (2014)) where documents are embedded together with words in a reduced dimensionality. This is done such that related words are mapped closely together and unrelated ones far from each other according to some metric in that vector space. This method, however, is not as well understood as LDA.

This paper is cited by a paper saved to your library

Using TF-IDF to Determine Word Relevance in Document Queries

J. E. Ramos · 2003

TLDR This paper examines the results of applying Term Frequency Inverse Document Frequency to determine what words in a corpus of documents might be more favorable to use in a query and provides evidence that this simple algorithm efficiently categorizes relevant words that can enhance query retrieval. [Expand](#)

1,814 citations 192 likes

Something not right? [Contact Us](#)

上線操作

How can schools make the best use of information technology in the classroom?

「教室中使用的資訊科技」到底有哪些？做背景引入概括資訊科技的最佳用法。也就是直接寫各種資訊科技在教室的最佳用法並作出論證。還可以進一步概述這些資訊科技以及在教室裡是怎麼進行應用的。

關鍵字：computers and education

《Exploring four decades of research in Computers & Education》

《Educational Technology Research Patterns in the Realm of the Digital Knowledge Age》

《Use of information communication technology in education: an analysis of Indian perspective》

《The application of information and communication technology in Nigerian secondary schools》



課程大綱

研究者常面臨的問題

AI輔助工具：**Semantic Scholar**

AI輔助工具：**Connected Papers**

AI輔助工具：**ChatGPT**

實作練習



Connected Papers

<https://www.connectedpapers.com/>



尋找相關文獻，通常是寫論文的一個難點之一，現在有愈來愈多的工具，甚至是可視化的視覺工具，協助研究者在最短的時間，找到相關的文獻。

Connected Papers 強大的功能，就在於能從一篇關鍵的文章，找出與這篇文章主要關聯的文章，並制作出來一個文章之間的網路圖，讓研究者可以了解文章之間的關係。

CONNECTED PAPERS

Share Follow About Pricing Sponsors Log in

Explore academic papers in a visual graph

Search by keywords, paper title, DOI or another identifier

Build a graph

Connected Papers

<https://www.connectedpapers.com/>



尋找相關文獻，通常是寫論文的一個難點之一，現在有愈來愈多的工具，甚至是可視化的視覺工具，協助研究者在最短的時間，找到相關的文獻。

Connected Papers 強大的功能，就在於能從一篇關鍵的文章，找出與這篇文章主要關聯的文章，並制作出來一個文章之間的網路圖，讓研究者可以了解文章之間的關係。

The screenshot displays the Connected Papers website interface. At the top, there is a search bar with the text "Search for a paper..." and navigation links for "Share", "Follow", "About", "Pricing", "Sponsors", and "Log in". Below the search bar, the title of the selected paper is "LANGUAGE LEARNING THROUGH SOCIAL NETWORKS: PERCEPTIONS AND REALITY". To the right of the title are buttons for "Prior works" and "Derivative works", and a note "Sponsored by Az".

The main content area features a network graph on the right, where nodes represent papers and lines represent connections between them. The nodes are colored in shades of blue and green, and their size varies. A timeline at the bottom of the graph indicates the creation date, "Created on Aug 26, 2023", and a range from 1966 to 2020.

On the left side of the interface, there is a list of related papers. The first entry is the "Origin paper": "LANGUAGE LEARNING THROUGH SOCIAL NETWORKS: PERCEPTIONS AND REALITY" by C. Lin, M. Warschauer, R. Blake, published in 2016. Below it, under the heading "Editorial Board Continued", is a paper by Nina Vyatkina from 2003. Other entries include "Type and Amount of Input-Based Practice in CALI: The Revelations of a..." by Luis Cerezo (2016) and "Social networking and language learning" by Lara L. Lomicka, G. Lord (2016).

On the right side, there is a sidebar with the title "LANGUAGE LEARNING THROUGH SOCIAL NETWORKS: PERCEPTIONS AND REALITY" and the authors "C. Lin, M. Warschauer, R. Blake". It also shows "104 Citations" and a section for "Open in:" with icons for a PDF viewer and a search engine.

課程大綱

研究者常面臨的問題

AI輔助工具：**Semantic Scholar**

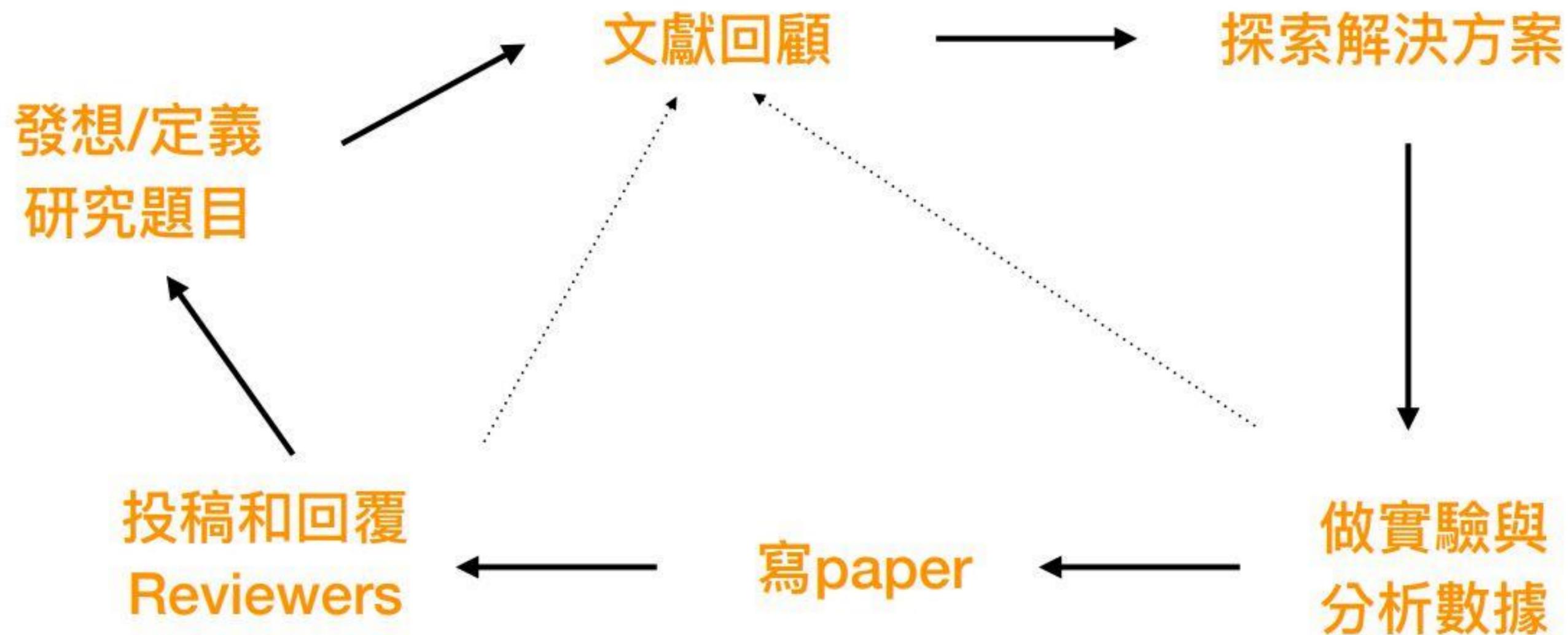
AI輔助工具：**Connected Papers**

AI輔助工具：**ChatGPT**

實作練習



做研究的輪迴



ChatGPT可以參與幾乎所有研究階段！

1. 使用ChatGPT來發想研究題目

ChatGPT已內化各個領域的思維模式，可以幫助研究者做brainstorming

- 原則：先給context (讓ChatGPT理解提問者的背景)，再問問題
(Recall: 語言模型的transformer架構)
- 建議做法：
 - ➡ 1. 先用1~2句話介紹具體研究領域和背景
 - ➡ 2. 再用1~2句話寫出明確的問題和需求

Case Study: 從圖書館的角度發想 (1/3)

我是一位研究型大學的圖書館館員，我感興趣的議題是如何使用近5年的最新的資通訊與人工智慧技術來協助大學的研究生能更有效率做研究以及從圖書館的角度提供支援。但是我不太熟悉相關領域的最新的進展與前沿研究。

因此，我希望能請你協助提供5個可能的結合最新資通訊與人工智慧技術的圖書館服務，並針對每一項可能的服務項目提供詳細的說明。

1. 先給context (介紹你的專業領域和議題背景)

2. 再問問題 (給定明確的需求)

我是一位研究型大學的圖書館館員，我感興趣的議題是如何使用近5年的最新的資訊通訊與人工智慧技術來協助大學的研究生能更有效率做研究以及從圖書館的角度提供支援。

但是我不太熟悉相關領域的最新的進展與前沿研究。

因此，我希望能請你協助提供5個可能的結合最新資訊通訊與人工智慧技術的圖書館服務，並針對每一項可能的服務項目提供詳細的說明。

I am a librarian at a research university. The topic I am interested in is how to use the latest information communications in the past five years.

Use artificial intelligence technology to assist university researchers to conduct research more efficiently and provide support from a library perspective.

But I am not very familiar with the latest developments and cutting-edge research in related fields.

Therefore, I would like to ask you to help provide 5 possible library services that combine the latest information communication and artificial intelligence technologies services and provide detailed descriptions for each possible service item.

介紹你的專業領域

給定明確的需求

Case Study: 從圖書館的角度發想 (1/3)

我是一位研究型大學的圖書館館員，我感興趣的議題是如何使用近5年的最新的資通訊與人工智慧技術來協助大學的研究生能更有效率做研究以及從圖書館的角度提供支援。但是我不太熟悉相關領域的最新的進展與前沿研究。

因此，我希望能請你協助提供5個可能的結合最新資通訊與人工智慧技術的圖書館服務，並針對每一項可能的服務項目提供詳細的說明。

1. 先給context (介紹你的專業領域和議題背景)

2. 再問問題 (給定明確的需求)



當前，資通訊與人工智慧技術正在圖書館領域帶來許多創新。以下是五個可能的圖書館服務，結合了最新的資通訊與人工智慧技術，可協助大學研究生更有效率地進行研究，並提供圖書館支援：

1. 智能文獻搜索與推薦系統：

- 說明：開發一個智能系統，利用自然語言處理(NLP)和機器學習技術分析研究生的研究需求，提供高度個性化的文獻搜索和推薦服務。這個系統可以根據研究主題，引用歷史，和用戶反饋來優化搜索結果。
- 好處：節省研究生的時間，幫助他們找到更相關的文獻，提高研究效率。

2. 智能參考服務機器人：

- 說明：開發一個基於自然語言處理和機器學習的虛擬助手，可以回答研究生的參考問題、提供圖書館資源的信息、指導使用數據庫等。這個機器人可以24/7提供支援。
- 好處：改進參考服務的可用性，使研究生隨時能夠獲得幫助。

3. 學術寫作支援工具：

- 說明：開發一個智能寫作輔助工具，使用自然語言處理和機器學習來幫助研究生編寫學術論文、報告和摘要。這個工具可以檢查文法、提供結構建議，甚至檢測抄襲。
- 好處：提高研究生的寫作質量，減少錯誤，並教育他們更好地理解學術寫作的要求。

4. 數據管理與分析平台：

- 說明：建立一個集成的平台，幫助研究生有效地管理和分析研究數據。此平台可以使用機器學習來自動處理數據、生成視覺化圖表、執行統計分析等，並提供數據儲存和分享功能。
- 好處：提高數據處理效率，支援研究生在數據驅動的研究中更好地工作。

5. 智能圖書館導覽與定位：

- 說明：開發一個使用室內定位技術和擴增實境的應用程序，幫助研究生在圖書館內快速找到所需的書籍和資源。這個應用程序可以提供路線指導、書架號碼查找等功能。
- 好處：節省研究生的時間，提供更好的圖書館資源訪問體驗。

這些服務可以提高大學研究生的研究效率，並為圖書館提供更現代化的支援，使其成為學術社區的重要資源。要實現這些服務，您可能需要合作開發圖書館內的技術團隊，並不斷更新和優化這些工具，以反映最新的技術趨勢和研究需求。

Q: 從專業圖書館員角度來看，有哪些是短期內可行的？
有哪些是意想不到的？

Case Study: 從圖書館的角度發想 (2/3)

接著可以就你感興趣的議題進一步追問

謝謝你的建議！我對於「智能文獻搜索與推薦系統」特別感興趣，希望能利用自然語言處理(NLP)和機器學習技術分析研究生的研究需求，根據研究主題，引用歷史，和用戶反饋來優化搜索結果，進而提供高度個性化的文獻搜索和推薦服務。

我相信這是能進一步提升大學圖書館在學術研究角色的數位化服務。能否請你提供3~5個具體步驟描述如何讓圖書館能有機會在一年以內推出上述「智能文獻搜索與推薦系統」的服務？

1. 具體描述有興趣的子議題 (包含關鍵字)

2. 給定明確的問題和需求

步驟4：測試和反饋

- 進行內部測試，將系統推出給圖書館內部團隊和一小部分用戶使用，以確保系統穩定性和功能正常。
- 收集用戶反饋，用於改進系統性能和用戶體驗。

步驟5：優化和擴展

具體描述有興趣的子
議題 (包含關鍵字)

謝謝你的建議！我對於「智能文獻搜索與推薦系統」特別感興趣，希望能利用自然語言處理(NLP)和機器學習技術分析研究生的研究需求，根據研究主題，引用歷史，和用戶反饋來優化搜索結果，進而提供高度個性化的文獻搜索和推薦服務。

我相信這是能進一步提升大學圖書館在學術研究角色的數位化服務。

給定明確的問題和需求

能否請你提供3~5個具體步驟描述如何讓圖書館能有機會在一年以內推出上述「智能文獻搜索與推薦系統」的服務？

Thanks for your advice! I am particularly interested in "intelligent literature search and recommendation systems". I hope to use natural language processing (NLP) and machine learning technology to analyze the research needs of graduate students, optimize search results based on research topics, citation history, and user feedback, and then provide Highly personalized literature search and recommendation service.

I believe this is a digital service that can further enhance the role of university libraries in academic research.

Could you please provide 3 to 5 specific steps to describe how the library can have the opportunity to launch the above-mentioned "intelligent document search and recommendation system" service within one year?

Case Study: 從圖書館的角度發想 (2/3)

接著可以就你感興趣的議題進一步追問

謝謝你的建議！我對於「智能文獻搜索與推薦系統」特別感興趣，希望能利用自然語言處理(NLP)和機器學習技術分析研究生的研究需求，根據研究主題，引用歷史，和用戶反饋來優化搜索結果，進而提供高度個性化的文獻搜索和推薦服務。

我相信這是能進一步提升大學圖書館在學術研究角色的數位化服務。

能否請你提供3~5個具體步驟描述如何讓圖書館能有機會在一年以內推出上述「智能文獻搜索與推薦系統」的服務？



推出一個「智能文獻搜索與推薦系統」需要精心策劃和執行。以下是一個可能的步驟序列，可幫助您在一年內實現這項服務：

步驟1：需求分析和計劃

- 確定服務的目標受眾，例如研究生、教師和研究人員。
- 與圖書館內部團隊合作，確定核心功能和用戶期望。
- 制定一個具體的計劃，包括預算、時間表和資源需求。

步驟2：數據收集和整理

- 收集和整理圖書館內部的文獻數據庫，包括書籍、期刊、學術文章等。
- 整理引用數據，以建立引用歷史數據庫，這將用於推薦系統的一部分。

步驟3：技術選擇和開發

- 選擇合適的自然語言處理（NLP）和機器學習工具和框架，例如Python中的NLTK、spaCy和Scikit-Learn。
- 開始系統開發，包括用戶界面、搜索引擎、NLP模型和推薦算法的開發。

1. 具體描述有興趣的子議題 (包含關鍵字)

2. 給定明確的問題和需求

步驟4：測試和反饋

- 進行內部測試，將系統推出給圖書館內部團隊和一小部分用戶使用，以確保系統穩定性和功能正常。
- 收集用戶反饋，用於改進系統性能和用戶體驗。

步驟5：優化和擴展

- 根據內部測試和用戶反饋進行系統優化，提高搜索準確性和推薦效果。
- 擴展系統功能，例如加入多語言支持、文獻全文搜索等。

步驟6：上線和宣傳

- 當系統達到預定的穩定性和性能標準時，正式上線服務。
- 制定宣傳計劃，通知圖書館社區新服務的推出，提供培訓和支援。

步驟7：監控和維護

- 持續監控系統運行，處理技術問題和改進用戶體驗。
- 定期更新系統，以反映新的研究趨勢和技術進展。

在這個過程中，要確保有一個跨職能團隊，包括圖書館學專業人員、數據科學家、軟體開發者和用戶體驗設計師，共同合作來實現這個項目。同時，不斷與用戶和研究生保持溝通，以確保服務滿足其需求，並在一年內完成這個重要的數位化服務。

Case Study: 從圖書館的角度發想 (3/3)

讓我們繼續追問下去 😎

謝謝你的建議！基於你上述提供的「智能文獻搜索與推薦系統」推展步驟，包含需求分析和計劃、數據收集和整理、技術選擇和開發、測試和反饋、優化和擴展、上線和宣傳、監控和維護，**能否請你提供5個在以上步驟最可能遭遇的困難與挑戰？請就每一個困難與挑戰給予詳細的說明並提示可能的解決方案。**



當推展「智能文獻搜索與推薦系統」時，您可能會面臨以下5個困難和挑戰，以及相應的解決方案：

1. 數據品質和整理挑戰：

- **困難：**文獻數據庫的品質差異，包括不一致的元數據、格式、錯誤和缺失數據，可能使數據整理變得複雜。
- **解決方案：**建立清晰的數據整理流程，使用自動化工具來處理重複性任務，並建立數據品質標準以確保一致性。必要時，考慮合作與出版商合作，獲得更結構化的數據。

2. NLP模型的性能挑戰：

- **困難：**自然語言處理（NLP）模型的性能可能會受到多種語言、學科領域和主題的影響，可能難以滿足所有用戶的需求。
- **解決方案：**持續改進NLP模型，進行領域特定的微調，以提高性能。為用戶提供個性化設置，以選擇不同的NLP模型或參數配置。

3. 用戶隱私和數據安全挑戰：

- **困難：**系統需要處理用戶查詢和閱讀歷史等敏感信息，可能引發隱私和安全問題。
- **解決方案：**實行嚴格的數據隱私政策，使用加密和安全訪問控制，匿名化用戶數據，只收集必要的信息。進行安全審計，確保系統的數據處理遵循最佳實踐。

4. 用戶培訓和採用挑戰：

- **困難：**用戶可能需要時間來適應新系統，可能會有一定的學習曲線。
- **解決方案：**提供定期的用戶培訓，制定易於理解的用戶文檔和教程。建立支援渠道，如幫助中心或客戶支援團隊，以解答用戶的疑問。

5. 技術演進挑戰：

- **困難：**技術領域不斷演進，您的系統需要跟上最新的NLP和機器學習發展。
- **解決方案：**建立一個持續更新和維護的計劃，跟蹤新技術和研究，定期更新系統以擁有最新的功能。與學術界和業界合作，參與研討會和會議，保持技術前沿知識。

通過解決這些挑戰，您可以增加成功實現「智能文獻搜索與推薦系統」的機會，提供高度個性化且功能強大的文獻搜索和推薦服務，以支持大學研究生和學術社區的需求。

Case Study: 從圖書館的角度發想 (4/4)

換個說法或方式問問題，有時能得到更多元的回饋 😎

我是一位研究型大學的圖書館館員，我感興趣的議題是如何使用近5年的最新的資通訊與人工智慧技術來協助大學的研究生能更有效率做研究以及從圖書館的角度提供支援。

但是我不太熟悉相關領域的最新的進展與前沿研究。

因此，我希望能請你協助提供5個結合最新資通訊與人工智慧技術的嶄新圖書館服務，而且這些服務是從未在任何現行圖書館推行過，並針對每一項可能的服務項目提供詳細的說明。

特別強調「嶄新服務」



當前，資通訊和人工智慧技術已經在圖書館領域帶來了許多創新。以下是5個結合最新資通訊和人工智慧技術的嶄新圖書館服務的示例，這些服務在目前的圖書館中尚未推行過：

1. 個人化研究導向建議：

- **服務說明：**這項服務使用機器學習和自然語言處理技術，根據每位研究生的研究方向和需求，提供個人化的研究建議。它分析研究生的閱讀記錄、研究領域的最新文獻以及研究生的研究計劃，然後生成推薦的閱讀材料和研究方法。

2. 實時文獻推送：

- **服務說明：**使用自然語言處理和機器學習，這項服務將實時跟蹤最新的學術出版物、會議報告和研究前沿，並向研究生提供即時的文獻推送。研究生可以訂閱特定主題或關鍵詞，以確保他們不會錯過重要的研究資訊。

3. 虛擬研究合作伙伴：

- **服務說明：**這是一個基於虛擬助手的服務，為研究生提供研究合作伙伴。它使用自然語言處理和對話式AI技術，協助研究生制定研究問題、查找相關文獻、設計實驗和分析數據。這種虛擬合作伙伴可隨時提供幫助，無需實際的合作夥伴。

4. 知識圖書館導覽機器人：

- **服務說明：**這是一個機器人導覽服務，旨在幫助研究生在圖書館內快速找到所需的資源。機器人配備視覺感知技術和自主導航功能，可以回答研究生的問題、提供路線指導，甚至幫助研究生找到特定的書籍或期刊。

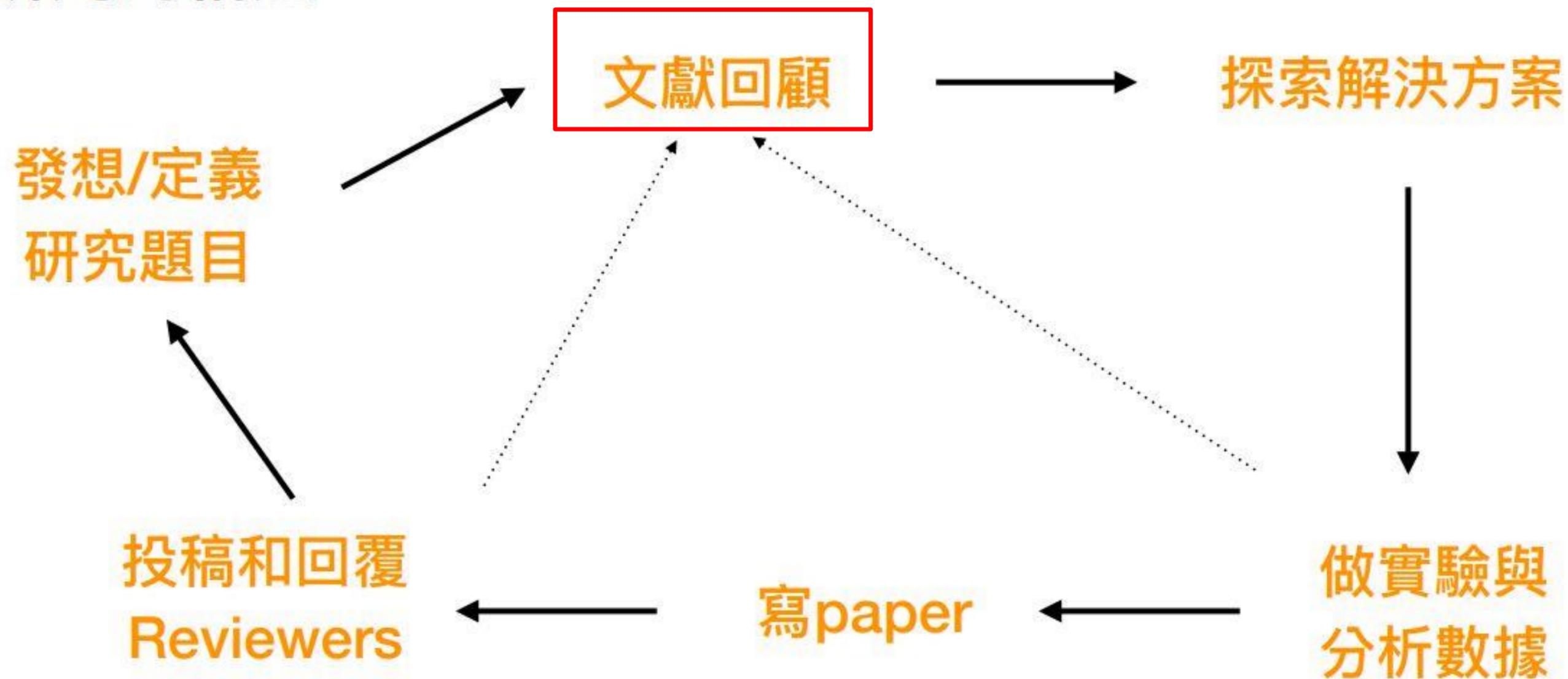
5. 學術寫作輔助：

- **服務說明：**這項服務利用自然語言生成技術，協助研究生改善學術寫作。它可以自動檢查和修正語法、結構和引用格式錯誤，並提供建議以改進寫作風格和邏輯結構。此外，它還能分析類似研究文章，幫助研究生更好地組織和表達他們的研究成果。

這些新服務將資通訊和人工智慧技術應用於圖書館環境，有助於提高研究生的效率和研究品質，同時也提供了更個性化、便捷的支持。實施這些服務需要合適的技術基礎和資源，但它們可以為大學圖書館帶來更多價值，提升其在學術社區中的地位。

Q: 從專業圖書館員角度來看，這些有更「創新」嗎？

做研究的輪迴



ChatGPT可以參與幾乎所有研究階段！

2. 使用ChatGPT寫文獻回顧

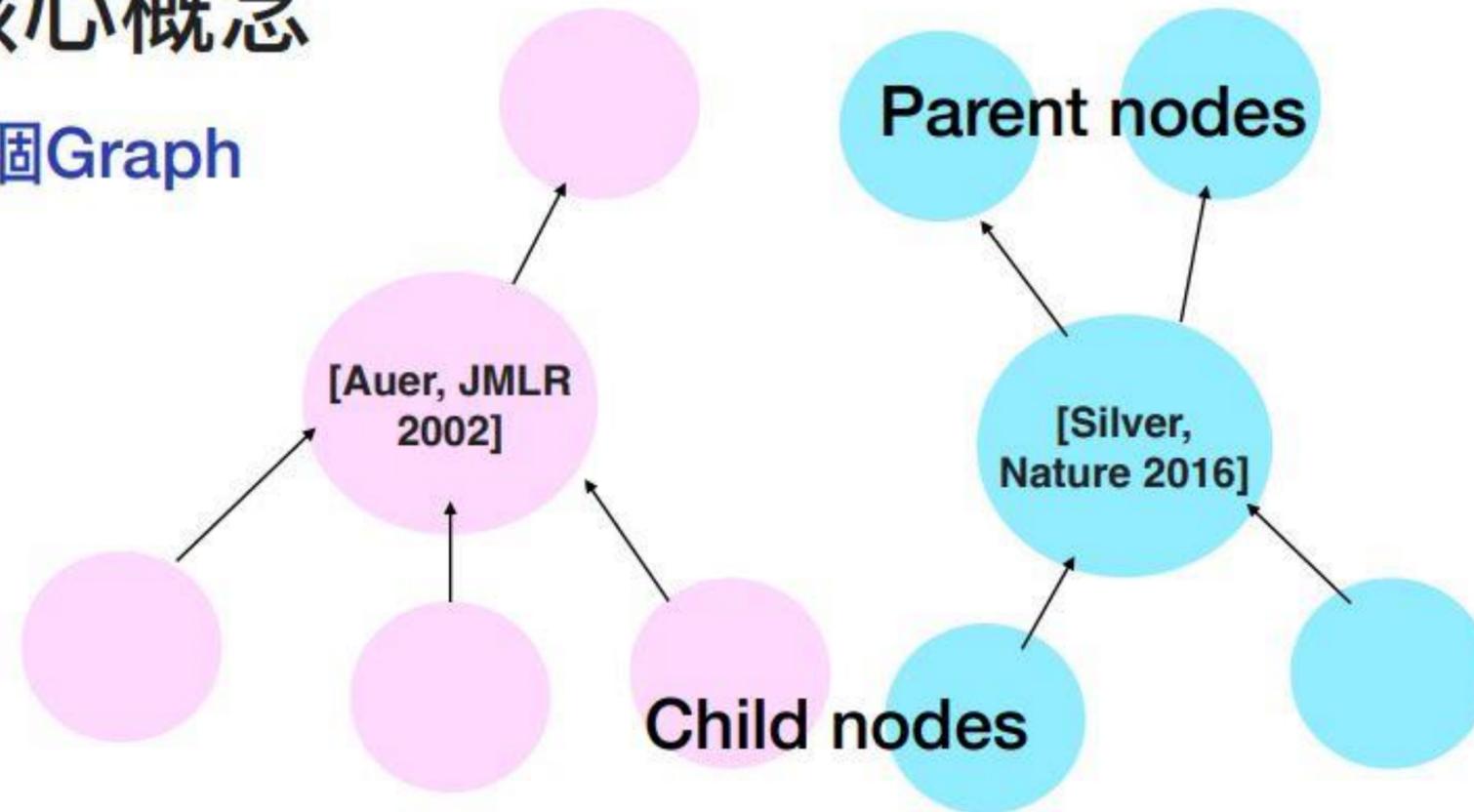
ChatGPT的訓練能implicitly建立關於一個領域的既有文獻的knowledge graph

- 原則：先設計搜尋演算法，再篩選文獻
- 建議做法：
 - ➡ 1. 先用1~2次prompts建立資料庫的搜尋演算法
 - ➡ 2. 再用多次prompts逐步篩選真正關鍵的文獻並取得summary

文獻回顧的兩個核心概念

(A) 所有文獻共同組成一個Graph

Citation network



(B) 「文獻回顧」本身是一個「圖搜尋演算法」

- BFS (Breadth-first search)
- DFS (Depth-first search)
- Hybrid

← 用ChatGPT來設計文獻搜尋演算法

讓我們先簡短地問 🤖

Consider the research question "Reinforcement Learning Algorithms for Portfolio Management With Risk Control."

Please provide a search algorithm that could retrieve relevant papers from Web of Science.

1. 先說我們的研究問題 (使用強化學習來做具有控管風險的投資組合管理)

2. 直接請ChatGPT提供搜尋方法

我的研究問題是「具有風險控制的投資組合管理的強化學習演算法」。

請提供一個可以從 Web of Science 檢索相關論文的搜尋演算法。

先說我們的研究問題 (使用強化學習來做具有控管風險的投資組合管理)

直接請ChatGPT提供搜尋方法

Consider the research question "Reinforcement Learning Algorithms for Portfolio Management With Risk Control."

Please provide a search algorithm that could retrieve relevant papers from Web of Science.

Step 1：先設計搜尋演算法 — 以Web of Science為例 (1/4)

讓我們先簡短地問 😎

Consider the research question "Reinforcement Learning Algorithms for Portfolio Management With Risk Control."

Please provide a search algorithm that could retrieve relevant papers from Web of Science.

1. 先說我們的研究問題 (使用強化學習來做具有控管風險的投資組合管理)

2. 直接請ChatGPT提供搜尋方法



To retrieve relevant papers on the research question "Reinforcement Learning Algorithms for Portfolio Management With Risk Control" from Web of Science, you can use a systematic search strategy. Here's a step-by-step approach to help you conduct your search:

1. Define Keywords and Synonyms:

- Start by identifying keywords and synonyms related to your research question. In this case, the main concepts are "Reinforcement Learning," "Portfolio Management," and "Risk Control." You may also want to include variations like "RL," "Investment Portfolio," and "Risk Management."

2. Use Boolean Operators:

- Combine your keywords using Boolean operators (AND, OR, NOT) to create search strings. For example:
 - "Reinforcement Learning" AND "Portfolio Management" AND "Risk Control"

4. Refine Your Search:

- You may receive a large number of results initially. To refine your search, consider adding more specific keywords or narrowing down your search by publication date (if necessary).
- You can also use Boolean operators to exclude irrelevant terms. For example:
 - "Reinforcement Learning" AND "Portfolio Management" AND "Risk Control" NOT "Deep Learning"

5. Filter by Document Type and Publication Date:

- Depending on your research requirements, you can filter results by document type (e.g., research articles, conference papers) and publication date to focus on the most relevant publications.

6. Review and Select Relevant Papers:

- Go through the search results and carefully review the titles, abstracts, and keywords to identify papers that align with your research question.
- Pay attention to the methodology, results, and conclusions mentioned in the abstracts.

Step 1：先設計搜尋演算法 — 以Web of Science為例 (2/4)

讓我們試試更鉅細靡遺的問法 😎

As an [experienced librarian](#) specializing in literature review and information retrieval from bibliographic databases (e.g., Web of Science), your task is to develop a comprehensive, accurate, and tailored [search algorithm for identifying and retrieving relevant articles from Web of Science](#).

To create a search algorithm that accurately reflects the research question, please follow these steps:

1. Break down research question into 2 or more key concepts
2. For each concept, identify closely related keywords
3. Use the Boolean operator “OR” to combine the keywords
4. Finally, use the Boolean operators “AND” to combine the keywords across various concepts

For example, focus on the research question “[Reinforcement Learning Algorithms for Portfolio Management With Risk Control](#).”

First, identify the three key concepts: Reinforcement Learning Algorithms, Portfolio Management, and Risk Control.

Next, develop closely related keywords for each concept.

Relevant keywords on Portfolio Management could be “investment”, “stock market”, “futures”, “foreign exchange”, and “cryptocurrency.”

Finally, please use “OR” to combine the keywords for each concept, and then use “AND” to combine the different concepts.

1. 先告知ChatGPT其角色

2. 列出文獻回顧的步驟

3. 寫出具體的研究問題

4. 更具體引導ChatGPT產生搜尋

Reference: Supercharge Your Research
Productivity with ChatGPT: A Practical Guide
(by Ruopeng An)

52

Step 1：先設計搜尋演算法 — 以Web of Science為例 (2/4)

讓我們試試更鉅細靡遺的問法 🤖

As an [experienced librarian](#) specializing in literature review and information retrieval from bibliographic databases (e.g., Web of Science), your task is to develop a comprehensive, accurate, and tailored [search algorithm for identifying and retrieving relevant articles from Web of Science](#).

To create a search algorithm that accurately reflects the research question, please follow these steps:

1. Break down research question into 2 or more key concepts
2. For each concept, identify closely related keywords
3. Use the Boolean operator "OR" to combine the keywords
4. Finally, use the Boolean operators "AND" to combine the keywords across various concepts

For example, focus on the research question "[Reinforcement Learning Algorithms for Portfolio Management With Risk Control](#)."

First, identify the three key concepts: Reinforcement Learning Algorithms, Portfolio Management, and Risk Control.

Next, develop closely related keywords for each concept.

Relevant keywords on Portfolio Management could be "investment", "stock market", "futures", "foreign exchange", and "cryptocurrency."

Finally, please use "OR" to combine the keywords for each concept, and then use "AND" to combine the different concepts.

1. 先告知ChatGPT其角色

2. 列出文獻回顧的步驟

3. 寫出具體的研究問題

4. 更具體引導ChatGPT產生搜尋

Reference: Supercharge Your Research Productivity with ChatGPT: A Practical Guide (by Ruopeng An)

52

1 先告知ChatGPT其角色

As an experienced librarian specializing in literature review and information retrieval from bibliographic databases (e.g., Web of Science), your task is to develop a comprehensive, accurate, and tailored search algorithm for identifying and retrieving relevant articles from Web of Science.

作為一位經驗豐富的圖書館員，專門從書目數據庫（例如Web of Science）進行文獻回顧和信息檢索，您的任務是開發一個全面、準確且定制的搜索算法，以識別並從Web of Science檢索相關文章。

2 列出文獻回顧的步驟

To create a search algorithm that accurately reflects the research question, please follow these steps:

1. Break down research question into 2 or more key concepts
2. For each concept, identify closely related keywords
3. Use the Boolean operator "OR" to combine the keywords
4. Finally, use the Boolean operators "AND" to combine the keywords across various Concepts

要建立一個能準確反映研究問題的搜索算法，請按照以下步驟進行操作：

1. 將研究問題分解為2個或更多關鍵概念。
2. 對於每個概念，確認密切相關的關鍵詞。
3. 使用布林邏輯運算元 "OR" 來結合這些關鍵詞。
4. 最後，使用布林邏輯運算元 "AND" 來跨不同概念結合這些關鍵詞。

Step 1：先設計搜尋演算法 — 以Web of Science為例 (2/4)

讓我們試試更鉅細靡遺的問法 🤓

As an [experienced librarian](#) specializing in literature review and information retrieval from bibliographic databases (e.g., Web of Science), your task is to develop a comprehensive, accurate, and tailored [search algorithm for identifying and retrieving relevant articles from Web of Science](#).

To create a search algorithm that accurately reflects the research question, please follow these steps:

1. Break down research question into 2 or more key concepts
2. For each concept, identify closely related keywords
3. Use the Boolean operator "OR" to combine the keywords
4. Finally, use the Boolean operators "AND" to combine the keywords across various concepts

For example, focus on the research question "[Reinforcement Learning Algorithms for Portfolio Management With Risk Control](#)."

First, identify the three key concepts: Reinforcement Learning Algorithms, Portfolio Management, and Risk Control.

Next, develop closely related keywords for each concept.

Relevant keywords on Portfolio Management could be "investment", "stock market", "futures", "foreign exchange", and "cryptocurrency."

Finally, please use "OR" to combine the keywords for each concept, and then use "AND" to combine the different concepts.

1. 先告知ChatGPT其角色

2. 列出文獻回顧的步驟

3. 寫出具體的研究問題

4. 更具體引導ChatGPT產生搜尋

Reference: Supercharge Your Research
Productivity with ChatGPT: A Practical Guide
(by Ruopeng An)

52

3

寫出具體的研究問題

For example, focus on the research question
“Reinforcement Learning Algorithms for Portfolio Management With Risk Control.”

以研究問題 “具有風險控制的投資組合管理的強化學習算法” 為例。

4

具體引導ChatGPT產生搜尋

First, identify the three key concepts: Reinforcement Learning Algorithms, Portfolio Management, and Risk Control.

Next, develop closely related keywords for each concept.

Relevant keywords on Portfolio Management could be “investment”, “stock market”, “futures”, “foreign exchange”, and “cryptocurrency.”

Finally, please use “OR” to combine the keywords for each concept, and then use “AND” to combine the different concepts.

首先，確認三個關鍵概念：強化學習算法、投資組合管理和風險控制。

然後，為每個概念發展相關的關鍵詞。

對於強化學習算法，相關關鍵詞可以包括：“reinforcement learning,” “algorithm,” “machine learning,” “deep learning”等。

對於投資組合管理，相關關鍵詞可以包括：“investment,” “stock market,” “futures,” “foreign exchange,” 和 “cryptocurrency”等。

對於風險控制，相關關鍵詞可以包括：“risk control,” “risk management,” “risk mitigation,” 和 “risk assessment”等。

最後，使用“OR”來結合每個概念的相關關鍵詞，然後使用“AND”來結合不同的概念。

Step 1：先設計搜尋演算法 — 以Web of Science為例 (3/4)



Certainly, I can help you develop a comprehensive search algorithm for the research question "Reinforcement Learning Algorithms for Portfolio Management With Risk Control" using the steps you provided. Let's break it down:

Key Concepts:

1. Reinforcement Learning Algorithms
2. Portfolio Management
3. Risk Control

Keywords for Each Concept:

1. Reinforcement Learning Algorithms:

- "Reinforcement learning"
- "Deep Q-learning"
- "Policy gradient methods"
- "SARSA"
- "Q-learning"
- "Reinforcement learning in finance"
- "Algorithmic trading with reinforcement learning"

2. Portfolio Management:

- "Portfolio optimization"
- "Asset allocation"
- "Portfolio construction"
- "Portfolio rebalancing"
- "Investment strategy"
- "Diversification"
- "Portfolio performance"

3. Risk Control:

- "Risk management"
- "Risk assessment"
- "Risk mitigation"
- "Risk modeling"
- "Risk-adjusted returns"
- "Volatility control"
- "Drawdown management"

Combining Keywords for Each Concept Using "OR":

- (Reinforcement learning) OR ("Deep Q-learning") OR ("Policy gradient methods") OR ("SARSA") OR ("Q-learning") OR ("Reinforcement learning in finance") OR ("Algorithmic trading with reinforcement learning")
- ("Portfolio optimization") OR ("Asset allocation") OR ("Portfolio construction") OR ("Portfolio rebalancing") OR ("Investment strategy") OR ("Diversification") OR ("Portfolio performance")
- ("Risk management") OR ("Risk assessment") OR ("Risk mitigation") OR ("Risk modeling") OR ("Risk-adjusted returns") OR ("Volatility control") OR ("Drawdown management")

Combining Concepts Using "AND":

- (Reinforcement learning OR "Deep Q-learning" OR "Policy gradient methods" OR SARSA OR "Q-learning" OR "Reinforcement learning in finance" OR "Algorithmic trading with reinforcement learning") AND ("Portfolio optimization" OR "Asset allocation" OR "Portfolio construction" OR "Portfolio rebalancing" OR "Investment strategy" OR "Diversification" OR "Portfolio performance") AND ("Risk management" OR "Risk assessment" OR "Risk mitigation" OR "Risk modeling" OR "Risk-adjusted returns" OR "Volatility control" OR "Drawdown management")

Step 1：先設計搜尋演算法 — 以Web of Science為例 (4/4)

Web of Science™ Search Sign In Register

Search > Results for (Reinforcement learning OR "Deep Q-learning" OR "Policy gradi...

10 results from Web of Science Core Collection for:

(Reinforcement learning OR "Deep Q-learning" OR "Policy gradient methods" OR SARSA OR "Q-learning" OR "Reinforcement learning in finance" OR "Algorithmic trading with reinforce...

Analyze Results Citation Report Create Alert

Quick add keywords: < + PORTFOLIO OPTIMIZATION + REINFORCEMENT LEARNING >

Publications You may also like... Copy query link

Refine results

Search within results...

Quick Filters

- Open Access 2
- Enriched Cited References 3

Publication Years

- 2023 1
- 2022 2
- 2021 2
- 2020 1
- 2019 1

See all >

0/10 Add To Marked List Export

Sort by: Relevance < 1 of 1 >

1 Dynamic portfolio rebalancing through reinforcement learning 6 Citations 48 References

Lim, QYE; Cao, Q and Quek, C
May 2022 | Dec 2021 (Early Access) | NEURAL COMPUTING & APPLICATIONS 34 (9) , pp.7125-7139

Enriched Cited References

Portfolio managements in financial markets involve risk management strategies and opportunistic responses to individual trading behaviours. Optimal portfolios constructed aim to have a minimal risk with highest accompanying investment returns, regardless of market conditions. This paper focuses on providing an alternative view in maximising portfolio returns using Reinforcement Learning (RL) by ... Show more

Free Full Text From Publisher

Related records ?

2 Does reinforcement learning outperform deep learning and traditional portfolio optimization models in frontier and developed financial markets? 70 References

Ngo, VM; Nguyen, HH and Van Nguyen, P
Apr 2023 | Mar 2023 (Early Access) | RESEARCH IN INTERNATIONAL BUSINESS AND FINANCE 65

課程大綱

研究者常面臨的問題

AI輔助工具：**Semantic Scholar**

AI輔助工具：**Connected Papers**

AI輔助工具：**ChatGPT**

實作練習



實作練習

