Imperial College London

Department of Earth Science and Engineering

MSc in Applied Computational Science and Engineering

Independent Research Project
Project Plan

Current Content Discovery for Module Teaching

by

Guanyuming He

Email: guanyuming.he24@imperial.ac.uk
GitHub username: esemsc-gh124
Repository: https://github.com/ese-ada-lovelace-2024/irp-gh124

Supervisors:

Sean O'Grady
Rhodri Nelson

4th June 2025

Abstract

TBD. **Keywords:** content discovery, information retrieval, LLM, search engines, case method, Business school,

1 Introduction

1.1 Problem background

Since the emergence of the first Business schools in the late 19th century [40, 33], several distinct pedagogical teaching strategies have been applied. First institutionalized at Harvard Business School [14, 4] in the early 20th century, a method about teaching students with real world business cases (will be called *case method* in the rest of the thesis) has been found more effective and engaging [42, 3, 23] than many traditional, for example, big lecture based, teaching methods. The case method is valued as a form of *active learning* methods for students, for its ability to expose students to complex, context-specific problems that lack clear-cut solutions. As such, it has found wide adoption across the world [41, 9].

However, despite its aforementioned adoption and performance in business school teaching, the case method faces a significant constraint: the availability and collection of timely and relevant case material [36, 28]. In particular, Christensen has identified in his classical article that instructors would have to conduct "extensive preparation" [6] for case method. Another factor contributing to this constraint of case method is the ever-evolving business world and the necessity of the latest information: Clark argues that, because learned skill will lose value quickly in five years, it is critical for students to be up to date to remain relevant in the business world [7]; McFarlane emphasises the importance of updated cases, as otherwise students could be disengaged or discouraged [24].

1.2 Past advancements in information retrieval

During the past two centuries, a number of key developments have profoundly expanded an individual's capacity to retrieve information about the world. In the early 19th century, the transmission of information was still traditional — carried by person on paper or simply remembered. This fundamentally limited both the speed and geographic reach of information retrieval. The invention of telegraph in the 1840s by Morse, Cornell, and Henry [34, 22], notably with Morse's first telegraph message, "*What hath God wrought?*" in 1844 [26], marked a paradigm shift by enabling very fast transmission of encoded information over relatively short distances (w.r.t. the earth) via wired networks. This technological innovation was considerably improved by the invention of the telephone by Bell in 1876 [44, 11], enabling communication directly by human voice, instead of encoded Morse code.

Wired communication is critically constrianed by geographical features on where the wires were laid. Around the late 19th century, Marconi's experiments with wireless telegraphy [10] and the first successful transatlantic signal in 1901 [2] introduced electromagnetic wave-based wireless communication, eventually accumulating into the world's first voice broadcast by radio in 1906 [39]. These milestones collectively redefined the temporal and spatial boundaries of information access.

The next many decades have seen people improving on the serious limitations of wireless communication: signal strength, interferences and attenuation, carrying capacity, and even deliberate sabotage during war times [13, 45, 1]. Theoretically, Hartley observed a logritham pattern of information capacity [15] and then Shannon expanded on it to first define *bits* and *entropy*, giving a formal mathematical theory of information [37] in 1948. Meanwhile, engineers

were experimenting with alternative modulation techniques, notably frequency modulation (FM), and the concepts were formalized in the 1930s [16].

As the theoretical understanding of information progressed, people began to have the idea of *searching* for information based on content and by relevance, instead of by unique identifier [31]. Actually, the term *information retrieval* was not invented until Mooers coined it in 1950 [25]. Since then, information retrieval systems have quickly evolved, and went through four phases before 2000: "(1) manual and mechanical devices; (2) offline computing; (3) online computing, vendor access; (4) distributed, networked, and mass computing." [12], with the last three substantially contributed to by the invention of the Internet [21], and consequently the emergence of search engines in the 1990s [35, 27]. For the next two decades, search engines greatly expanded in speed and coverage and has been significantly impacting the society's information for at least a decade [5, 18].

LLMs have undoubtedly entered and transformed many areas, including daily life, business and the industry, and research [8]. One strong appeal of LLMs is that they could process user's natural language input and generate natural language output in return with remarkable resemblence to what a human would say [47, 20]. However, because of the inherent limit of neural networks, some argue that they could not formally reason about what they output [32, 19, 30]. Indeed, hallucination [17, 29] and other forms of distortion of facts, is a big problem of LLMs. On the other hand, although search engines rely on determinstic algorithms that give precise reference to searched result, they could not compare with LLMs' ability to process search prompt and summarize results.

Therefore, various attempts have been made to integrate LLMs with search engines [46, 38, 43]. Specifically, Xiong et al. proposes to categorize them into "LLM4Search" and "Search4LLM", where "A4B" means using A to improve B [46]. Here is where my thesis will build upon. Previously, one would need to carefully craft search engine prompts into clearly and precise list of words or a short sentence. Now, with the help with LLMs that can easily process frivolous natural language input, I plan to integrate them together to boost information retrieval further, especially in the area of business school teaching content discovery.

2 Methods

3 Results

4 Discussion

5 Conclusion

References

[1] E. F. W. Alexanderson. Transatlantic radio communication. *Proceedings of the American Institute of Electrical Engineers*, 38(10):1077–1093, 1919.

[2] W. J. G. Beynon. Marconi, radio waves, and the ionosphere. *Radio Science*, 10(7):657–664, 1975.

[3] Kevin M. Bonney. Case study teaching method improves student performance and perceptions of learning gains. *Journal of Microbiology & Biology Education*, 16(1):21–28, 2015.

[4] Todd Bridgman, Stephen Cummings, and Colm McLaughlin. The case method as invented tradition: revisiting harvard's history to reorient management education. In *Academy of Management Proceedings*, volume 2015, page 11637. Academy of Management Briarcliff Manor, NY 10510, 2015.

[5] Junghoo Cho and Sourashis Roy. Impact of search engines on page popularity. In *Proceedings of the 13th International Conference on World Wide Web*, WWW '04, page 20–29, New York, NY, USA, 2004. Association for Computing Machinery.

[6] C Roland Christensen. Teaching and the case method harvard business school. `https://www.hbs.edu/teaching/case-method/Pages/default.aspx`, 1987.

[7] Tom Clark. Case method in the digital age: how might new technologies shape experiential learning and real-life story telling? LSE Impact of Social Sciences, 2016.

[8] Qinxu Ding, Ding Ding, Yue Wang, Chong Guan, and Bosheng Ding. Unraveling the landscape of large language models: a systematic review and future perspectives. *Journal of Electronic Business & Digital Economics*, 3(1):3–19, 2023.

[9] Rebekka Eckhaus. Supporting the adoption of business case studies in esp instruction through technology. *Asian ESP Journal*, 14:280–281, 2018.

[10] Gabriele Falciasecca. Marconi's early experiments in wireless telegraphy, 1895. *IEEE Antennas and Propagation Magazine*, 52(6):220–221, 2010.

[11] J.E. Flood. Alexander graham bell and the invention of the telephone. *Electronics and Power*, 22:159–162, 1976.

[12] J.-M. Griffiths and D.W. King. Us information retrieval system evolution and evaluation (1945-1975). *IEEE Annals of the History of Computing*, 24(3):35–55, 2002.

[13] Brian N. Hall. The british army and wireless communication, 1896–1918. *War in History*, 19(3):290–321, 2012.

[14] John S Hammond. *Learning by the case method*. Harvard Business School Boston, MA, 1980.

[15] R. V. L. Hartley. Transmission of information. *Bell System Technical Journal*, 7(3):535–563, 1928.

[16] Raymond A. Heising. Modulation methods. *Proceedings of the IRE*, 50(5):896–901, 1962.

[17] Lei Huang, Weijiang Yu, Weitao Ma, Weihong Zhong, Zhangyin Feng, Haotian Wang, Qianglong Chen, Weihua Peng, Xiaocheng Feng, Bing Qin, and Ting Liu. A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *ACM Trans. Inf. Syst.*, 43(2), January 2025.

[18] Olagoke Olawale Israel and Olatunji Olusoji Samson. The impact of search engines in the world today. *International Journal of Management, IT and Engineering*, 8(3):10–22, 2018.

[19] Andrei Kucharavy. Fundamental limitations of generative llms. In *Large Language Models in Cybersecurity: Threats, Exposure and Mitigation*, pages 55–64. Springer Nature Switzerland Cham, 2024.

[20] Pranjal Kumar. Large language models (llms): survey, technical frameworks, and future challenges. *Artificial Intelligence Review*, 57(10):260, 2024.

[21] Barry M. Leiner, Vinton G. Cerf, David D. Clark, Robert E. Kahn, Leonard Kleinrock, Daniel C. Lynch, Jon Postel, Larry G. Roberts, and Stephen Wolff. A brief history of the internet. *SIGCOMM Comput. Commun. Rev.*, 39(5):22–31, October 2009.

[22] Kenneth B Lifshitz. *Makers of the Telegraph: Samuel Morse, Ezra Cornell and Joseph Henry*. McFarland, 2017.

[23] Alberto Lusoli. Teaching business as business: The role of the case method in the constitution of management as a science-based profession. *Journal of Management History*, 26(2):277–290, 2020.

[24] Donovan A McFarlane. Guidelines for using case studies in the teaching-learning process. *College Quarterly*, 18(1):n1, 2015.

[25] Calvin Mooers. The theory of digital handing of non-numerical information and its implications to machine economics. In *Proceedings of the meeting of the Association for Computing Machinery at Rutgers University*, 1950.

[26] Samuel Finley Breese Morse. First telegraph message. Retrieved from the Library of Congress, `https://www.loc.gov/item/mcc.019/`, May 1844.

[27] Rhoda Okunev. *History of the Internet, Search Engines, and More*, pages 9–16. Apress, Berkeley, CA, 2023.

[28] Sandeep Puri. Effective learning through the case method. *Innovations in Education and Teaching International*, 59(2):161–171, 2022.

[29] Vipula Rawte, Amit Sheth, and Amitava Das. A survey of hallucination in large foundation models, 2023.

[30] Walid S Saba. Stochastic llms do not understand language: towards symbolic, explainable and ontologically based llms. In *International conference on conceptual modeling*, pages 3–19. Springer, 2023.

[31] Mark Sanderson and W. Bruce Croft. The history of information retrieval research. *Proceedings of the IEEE*, 100(Special Centennial Issue):1444–1451, 2012.

[32] Erin Sanu, T Keerthi Amudaa, Prasiddha Bhat, Guduru Dinesh, Apoorva Uday Kumar Chate, and Ramakanth Kumar P. Limitations of large language models. In *2024 8th International Conference on Computational System and Information Technology for Sustainable Solutions (CSITSS)*, pages 1–6, 2024.

[33] Steven A Sass. *The pragmatic imagination: A history of the Wharton School, 1881-1981*. University of Pennsylvania Press, 2016.

[34] Mischa Schwartz and David Hochfelder. Two controversies in the early history of the telegraph. *IEEE Communications Magazine*, 48(2):28–32, 2010.

[35] Tom Seymour, Dean Frantsvog, Satheesh Kumar, et al. History of search engines. *International Journal of Management & Information Systems (IJMIS)*, 15(4):47–58, 2011.

[36] Binod Shah. Case method of teaching in management education. *Research Gate*, 2019.

[37] C. E. Shannon. A mathematical theory of communication. *SIGMOBILE Mob. Comput. Commun. Rev.*, 5(1):3–55, January 2001.

[38] Xiang Shi, Jiawei Liu, Yinpeng Liu, Qikai Cheng, and Wei Lu. Know where to go: Make llm a relevant, responsible, and trustworthy searchers. *Decision Support Systems*, 188:114354, 2025.

[39] Elliot N. Sivowitch. A technological survey of broadcasting's "pre-history," 1876–1920. *Journal of Broadcasting*, 15(1):1–20, 1970.

[40] John-Christopher Spender. The business school in america: a century goes by. *The future of business schools: Scenarios and strategies for*, pages 9–18, 2020.

[41] Carlos JO Trejo-Pech and Susan White. The use of case studies in undergraduate business administration. *Revista de administração de empresas*, 57:342–356, 2017.

[42] Klára Vítečková, Tobias Cramer, Matthias Pilz, Janine Tögel, Sascha Albers, Steven van den Oord, and Tomasz Rachwał. Case studies in business education: an investigation of a learner-friendly approach. *Journal of International Education in Business*, 18(2):149–176, 2025.

[43] Tu Vu, Mohit Iyyer, Xuezhi Wang, Noah Constant, Jerry Wei, Jason Wei, Chris Tar, Yun-Hsuan Sung, Denny Zhou, Quoc Le, and Thang Luong. Freshllms: Refreshing large language models with search engine augmentation, 2023.

[44] Thomas A. Watson. How bell invented the telephone. *Proceedings of the American Institute of Electrical Engineers*, 34(8):1503–1513, 1915.

[45] JONATHAN REED WINKLER. Telecommunications in world war i. *Proceedings of the American Philosophical Society*, 159(2):162–168, 2015.

[46] Haoyi Xiong, Jiang Bian, Yuchen Li, Xuhong Li, Mengnan Du, Shuaiqiang Wang, Dawei Yin, and Sumi Helal. When search engine services meet large language models: Visions and challenges. *IEEE Transactions on Services Computing*, 17(6):4558–4577, 2024.

[47] Jingfeng Yang, Hongye Jin, Ruixiang Tang, Xiaotian Han, Qizhang Feng, Haoming Jiang, Shaochen Zhong, Bing Yin, and Xia Hu. Harnessing the power of llms in practice: A survey on chatgpt and beyond. *ACM Trans. Knowl. Discov. Data*, 18(6), April 2024.