

BMI 510/610 Unit 8 References – William Hersh, MD

- Abalkina, A., Aquarius, R., Bik, E., Bimler, D., Bishop, D., Byrne, J., Cabanac, G., Day, A., Labbé, C., Wise, N., 2025. “Stamp out paper mills” - science sleuths on how to fight fake research. *Nature* 637, 1047–1050. <https://doi.org/10.1038/d41586-025-00212-1>
- Afful-Dadzie, E., Afful-Dadzie, A., Egala, S.B., 2023. Social media in health communication: A literature review of information quality. *Health Inf Manag* 52, 3–17. <https://doi.org/10.1177/1833358321992683>
- AI is killing the web. Can anything save it?, 2025. . *The Economist*.
- Allen, M., 2020. “Immune to Evidence”: How Dangerous Coronavirus Conspiracies Spread. ProPublica. URL <https://www.propublica.org/article/immune-to-evidence-how-dangerous-coronavirus-conspiracies-spread?token=YxUntiDH12MOMHz5OD1yLBRth4wTzdEG> (accessed 8.20.20).
- Allen, M.R., Desai, N., Namazi, A., Leas, E., Dredze, M., Smith, D.M., Ayers, J.W., 2024. Characteristics of X (Formerly Twitter) Community Notes Addressing COVID-19 Vaccine Misinformation. *JAMA* e244800. <https://doi.org/10.1001/jama.2024.4800>
- Alonso, O., Baeza-Yates, R. (Eds.), 2024. *Information Retrieval: Advanced Topics and Techniques*, 1st ed. Association for Computing Machinery, New York, NY, USA.
- Alperin, J.P., 2022. Why I think ending article-processing charges will save open access. *Nature* 610, 233. <https://doi.org/10.1038/d41586-022-03201-w>
- Amugongo, L.M., Mascheroni, P., Brooks, S., Doering, S., Seidel, J., 2025. Retrieval augmented generation for large language models in healthcare: A systematic review. *PLOS Digit Health* 4, e0000877. <https://doi.org/10.1371/journal.pdig.0000877>
- Anderson, D., 2023. Quality rater and algorithmic evaluation systems: Are major changes coming? [WWW Document]. Search Engine Land. URL <https://searchengineland.com/quality-rater-algorithmic-evaluation-systems-changes-434895> (accessed 12.10.23).
- Arora, V.M., Bloomgarden, E., Jain, S., 2022. Supporting Health Care Workers to Address Misinformation on Social Media. *N Engl J Med* 386, 1683–1685. <https://doi.org/10.1056/NEJMp2117180>
- Avissar-Whiting, M., 2022. Downstream retraction of preprinted research in the life and medical sciences. *PLoS One* 17, e0267971. <https://doi.org/10.1371/journal.pone.0267971>
- Azzam, A., Bresler, D., Leon, A., Maggio, L., Whitaker, E., Heilman, J., Orlovitz, J., Swisher, V., Rasberry, L., Otoide, K., Trotter, F., Ross, W., McCue, J.D., 2017. Why Medical Schools Should Embrace Wikipedia: Final-Year Medical Student Contributions to Wikipedia Articles for Academic Credit at One School. *Acad Med* 92, 194–200. <https://doi.org/10.1097/ACM.0000000000001381>
- Bachrach, C.A., Charen, T., 1978. Selection of MEDLINE contents, the development of its thesaurus, and the indexing process. *Med Inform (Lond)* 3, 237–254. <https://doi.org/10.3109/14639237809014183>
- Baraldi, J.H., Picozzo, S.A., Arnold, J.C., Volarich, K., Gionfriddo, M.R., Piper, B.J., 2022. A cross-sectional examination of conflict-of-interest disclosures of physician-authors publishing in high-impact US medical journals. *BMJ Open* 12, e057598. <https://doi.org/10.1136/bmjopen-2021-057598>
- Baron, R.J., Coleman, C.H., 2023. Protecting the Legitimacy of Medical Expertise. *N Engl J Med* 388, 676–678. <https://doi.org/10.1056/NEJMp2214120>

- Baron, R.J., Ejnes, Y.D., 2022. Physicians Spreading Misinformation on Social Media - Do Right and Wrong Answers Still Exist in Medicine? *N Engl J Med* 387, 1–3. <https://doi.org/10.1056/NEJMp2204813>
- Bastian, H., Glasziou, P., Chalmers, I., 2010. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? *PLoS Med* 7, e1000326. <https://doi.org/10.1371/journal.pmed.1000326>
- Bauchner, H., McDermott, M.M., Butte, A.J., 2023. Data Sharing Enters a New Era. *Ann Intern Med* 176, 400–401. <https://doi.org/10.7326/M22-3479>
- Bautista, J.R., Zhang, Y., Gwizdka, J., 2021. Healthcare professionals' acts of correcting health misinformation on social media. *Int J Med Inform* 148, 104375. <https://doi.org/10.1016/j.ijmedinf.2021.104375>
- Baxter, S.L., Lander, L., Clay, B., Bell, J., Hansen, K., Walker, A., Tai-Seale, M., 2022. Comparing the Use of DynaMed and UpToDate by Physician Trainees in Clinical Decision-Making: A Randomized Crossover Trial. *Appl Clin Inform* 13, 139–147. <https://doi.org/10.1055/s-0041-1742216>
- Bero, L., 2025. Systematic Reviewers Have an Obligation to Promote Research Integrity. *J Law Med Ethics* 1–6. <https://doi.org/10.1017/jme.2025.34>
- Blau, W., Cerf, V.G., Enriquez, J., Francisco, J.S., Gasser, U., Gray, M.L., Greaves, M., Grosz, B.J., Jamieson, K.H., Haug, G.H., Hennessy, J.L., Horvitz, E., Kaiser, D.I., London, A.J., Lovell-Badge, R., McNutt, M.K., Minow, M., Mitchell, T.M., Ness, S., Parthasarathy, S., Perlmutter, S., Press, W.H., Wing, J.M., Witherell, M., 2024. Protecting scientific integrity in an age of generative AI. *Proc Natl Acad Sci U S A* 121, e2407886121. <https://doi.org/10.1073/pnas.2407886121>
- Borrego, Á., 2023. Article processing charges for open access journal publishing: A review. *Learned Publishing* 36, 359–378. <https://doi.org/10.1002/leap.1558>
- Boyd, D., Golebiewski, M., 2019. Data Voids [WWW Document]. *Data & Society*. URL <https://datasociety.net/library/data-voids/> (accessed 4.22.22).
- Boyle, P., 2022. Widespread distrust in science: Is the way we communicate to blame? [WWW Document]. *AAMC*. URL <https://www.aamc.org/news-insights/widespread-distrust-science-way-we-communicate-blame> (accessed 4.27.23).
- Bradley-Ridout, G., Nekolaichuk, E., Jamieson, T., Jones, C., Morson, N., Chuang, R., Springall, E., 2021. UpToDate versus DynaMed: a cross-sectional study comparing the speed and accuracy of two point-of-care information tools. *J Med Libr Assoc* 109, 382–387. <https://doi.org/10.5195/jmla.2021.1176>
- Brainard, J., 2024. Open for business. *Science* 385, 486–489. <https://doi.org/10.1126/science.ads1013>
- Brin, S., Page, L., 1998. The anatomy of a large-scale hypertextual Web search engine, in: *Proceedings of the Seventh International Conference on World Wide Web 7, WWW7*. Elsevier Science Publishers B. V., NLD, pp. 107–117.
- Brock, J., 2020. Rapid Registered Reports initiative aims to stop coronavirus researchers following false leads. *Nature Index*. URL <https://www.natureindex.com/news-blog/rapid-registered-report-coronavirus-aims-to-stop-researchers-following-false-research-leads> (accessed 4.25.21).
- Broder, A., 2002. A taxonomy of web search. *SIGIR Forum* 36, 3–10. <https://doi.org/10.1145/792550.792552>

- Cabanac, G., 2024. Chain retraction: how to stop bad science propagating through the literature. *Nature* 632, 977–979. <https://doi.org/10.1038/d41586-024-02747-1>
- Carlisle, J.B., 2021. False individual patient data and zombie randomised controlled trials submitted to *Anaesthesia*. *Anaesthesia* 76, 472–479. <https://doi.org/10.1111/anae.15263>
- Castillo, C., Davison, B.D., 2011. Adversarial Web Search. *Found. Trends Inf. Retr.* 4, 377–486. <https://doi.org/10.1561/1500000021>
- Caulfield, T., 2020. Pseudoscience and COVID-19 — we’ve had enough already. *Nature*. <https://doi.org/10.1038/d41586-020-01266-z>
- Chalmers, I., 1990. Underreporting research is scientific misconduct. *JAMA* 263, 1405–1408.
- Chamberlin, S.R., Bedrick, S.D., Cohen, A.M., Wang, Y., Wen, A., Liu, S., Liu, H., Hersh, W.R., 2020. Evaluation of patient-level retrieval from electronic health record data for a cohort discovery task. *JAMIA Open* 3, 395–404. <https://doi.org/10.1093/jamiaopen/ooaa026>
- Chambers, C., 2019. The registered reports revolution - Lessons in cultural reform [WWW Document]. *Significance*. URL <https://rss.onlinelibrary.wiley.com/doi/10.1111/j.1740-9713.2019.01299.x> (accessed 4.25.21).
- Chauhan, C., Currie, G., 2024. The Impact of Generative Artificial Intelligence on Research Integrity in Scholarly Publishing. *Am J Pathol* 194, 2234–2238. <https://doi.org/10.1016/j.ajpath.2024.10.001>
- Chawla, D.S., 2024. How reliable is this research? Tool flags papers discussed on PubPeer. *Nature*. <https://doi.org/10.1038/d41586-024-01247-6>
- Cheah, P.Y., Piasecki, J., 2022. Should peer reviewers be paid to review academic papers? *Lancet* 399, 1601. [https://doi.org/10.1016/S0140-6736\(21\)02804-X](https://doi.org/10.1016/S0140-6736(21)02804-X)
- Chen, A., Chen, D.O., 2023. Accuracy of Chatbots in Citing Journal Articles. *JAMA Netw Open* 6, e2327647. <https://doi.org/10.1001/jamanetworkopen.2023.27647>
- Chen, Z., Chen, C., Yang, G., He, X., Chi, X., Zeng, Z., Chen, X., 2024. Research integrity in the era of artificial intelligence: Challenges and responses. *Medicine (Baltimore)* 103, e38811. <https://doi.org/10.1097/MD.00000000000038811>
- Cheng, K., Sun, Z., Liu, X., Wu, H., Li, C., 2024. Generative artificial intelligence is infiltrating peer review process. *Crit Care* 28, 149. <https://doi.org/10.1186/s13054-024-04933-z>
- Chiarelli, A., Johnson, R., Pinfield, S., Richens, E., 2019. Accelerating scholarly communication: The transformative role of preprints. *Zenodo*. <https://doi.org/10.5281/zenodo.3357727>
- Cimino, J.J., 2006. Use, usability, usefulness, and impact of an infobutton manager. *AMIA Annu Symp Proc* 151–155.
- Cimino, J.J., 1996. Linking patient information systems to bibliographic resources. *Methods Inf Med* 35, 122–126.
- Clark, J., 2023. How covid-19 bolstered an already perverse publishing system. *BMJ* 380, 689. <https://doi.org/10.1136/bmj.p689>
- Clarke, C.L.A., Dietz, L., 2024. LLM-based relevance assessment still can’t replace human relevance assessment. <https://doi.org/10.48550/arXiv.2412.17156>
- Coletti, M.H., Bleich, H.L., 2001. Medical subject headings used to search the biomedical literature. *J Am Med Inform Assoc* 8, 317–323. <https://doi.org/10.1136/jamia.2001.0080317>
- Dance, A., 2023. Stop the peer-review treadmill. I want to get off. *Nature* 614, 581–583. <https://doi.org/10.1038/d41586-023-00403-8>
- Danchev, V., Min, Y., Borghi, J., Baiocchi, M., Ioannidis, J.P.A., 2021. Evaluation of Data Sharing After Implementation of the International Committee of Medical Journal Editors

- Data Sharing Statement Requirement. *JAMA Netw Open* 4, e2033972.
<https://doi.org/10.1001/jamanetworkopen.2020.33972>
- Day, A., 2022. Exploratory analysis of text duplication in peer-review reveals peer-review fraud and paper mills. arXiv:2202.03310 [cs].
- Deangelis, C.D., Drazen, J.M., Frizelle, F.A., Haug, C., Hoey, J., Horton, R., Kotzin, S., Laine, C., Marusic, A., Overbeke, A.J.P.M., Schroeder, T.V., Sox, H.C., Van Der Weyden, M.B., International Committee of Medical Journal Editors, 2005. Is this clinical trial fully registered? A statement from the International Committee of Medical Journal Editors. *JAMA* 293, 2927–2929. <https://doi.org/10.1001/jama.293.23.jed50037>
- Deda, Y., 2026. Health AI Overviews Trust YouTube Over Medical Platforms. SE Ranking Blog. URL <https://seranking.com/blog/health-ai-overviews-youtube-vs-medical-sites/> (accessed 2.20.26).
- Del Fiol, G., Huser, V., Strasberg, H.R., Maviglia, S.M., Curtis, C., Cimino, J.J., 2012. Implementations of the HL7 Context-Aware Knowledge Retrieval (“Infobutton”) Standard: challenges, strengths, limitations, and uptake. *J Biomed Inform* 45, 726–735.
<https://doi.org/10.1016/j.jbi.2011.12.006>
- Dobbins, N.J., Han, B., Zhou, W., Lan, K.F., Kim, H.N., Harrington, R., Uzuner, Ö., Yetisgen, M., 2023. LeafAI: query generator for clinical cohort discovery rivaling a human programmer. *J Am Med Inform Assoc* 30, 1954–1964.
<https://doi.org/10.1093/jamia/ocad149>
- Dolgin, E., 2020. Core Concept: The pandemic is prompting widespread use—and misuse—of real-world data. *PNAS* 117, 27754–27758. <https://doi.org/10.1073/pnas.2020930117>
- Dorn, S., 2024. How AI Is Changing The Way Doctors Access Medical Knowledge [WWW Document]. *Forbes*. URL <https://www.forbes.com/sites/spencerdorn/2024/12/13/how-ai-is-changing-how-doctors-access-medical-knowledge/> (accessed 4.28.25).
- Ebrahim, S., Sohani, Z.N., Montoya, L., Agarwal, A., Thorlund, K., Mills, E.J., Ioannidis, J.P.A., 2014. Reanalyses of randomized clinical trial data. *JAMA* 312, 1024–1032.
<https://doi.org/10.1001/jama.2014.9646>
- Eddy, K., Lipka, M., Matsa, K.E., St. Aubin, C., Shearer, E., 2026. What does it mean to ‘do your own research,’ and how often do Americans do it? Pew Research Center. URL <https://www.pewresearch.org/journalism/2026/02/11/what-does-it-mean-to-do-your-own-research-and-how-often-do-americans-do-it/> (accessed 2.16.26).
- Eysenbach, G., Powell, J., Kuss, O., Sa, E.-R., 2002. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. *JAMA* 287, 2691–2700. <https://doi.org/10.1001/jama.287.20.2691>
- Faggioli, G., Dietz, L., Clarke, C.L.A., Demartini, G., Hagen, M., Hauff, C., Kando, N., Kanoulas, E., Potthast, M., Stein, B., Wachsmuth, H., 2024. Who Determines What Is Relevant? Humans or AI? Why Not Both? *Commun. ACM* 67, 31–34.
<https://doi.org/10.1145/3624730>
- Faggioli, G., Dietz, L., Clarke, C.L.A., Demartini, G., Hagen, M., Hauff, C., Kando, N., Kanoulas, E., Potthast, M., Stein, B., Wachsmuth, H., 2023. Perspectives on Large Language Models for Relevance Judgment, in: *Proceedings of the 2023 ACM SIGIR International Conference on Theory of Information Retrieval, ICTIR ’23*. Association for Computing Machinery, New York, NY, USA, pp. 39–50.
<https://doi.org/10.1145/3578337.3605136>

- Ferguson, T., 2002. From patients to end users. *BMJ* 324, 555–556.
<https://doi.org/10.1136/bmj.324.7337.555>
- Fernández-Pichel, M., Pichel, J.C., Losada, D.E., 2025. Evaluating search engines and large language models for answering health questions. *NPJ Digit Med* 8, 153.
<https://doi.org/10.1038/s41746-025-01546-w>
- Fiorini, N., Canese, K., Starchenko, G., Kireev, E., Kim, W., Miller, V., Osipov, M., Kholodov, M., Ismagilov, R., Mohan, S., Ostell, J., Lu, Z., 2018. Best Match: New relevance search for PubMed. *PLoS Biol* 16, e2005343. <https://doi.org/10.1371/journal.pbio.2005343>
- Flanagin, A., Fontanarosa, P.B., Bauchner, H., 2020. Preprints Involving Medical Research—Do the Benefits Outweigh the Challenges? *JAMA* 324, 1840–1843.
<https://doi.org/10.1001/jama.2020.20674>
- Flier, J.S., 2021. Misconduct in Bioscience Research: a 40-year perspective. *Perspect Biol Med* 64, 437–456. <https://doi.org/10.1353/pbm.2021.0035>
- Flintoft, L., MacCallum, C.J., Streeter, M., Flanagan, D., Ferguson, L., 2023. Tackling publication manipulation at scale: Hindawi’s journey and lessons for academic publishing [WWW Document]. Wiley. URL <https://www.wiley.com/en-us/network/publishing/research-publishing/open-access/hindawi-publication-manipulation-whitepaper> (accessed 5.5.24).
- Fox, S., 2011. Health Topics. Pew Research Center: Internet, Science & Tech. URL <https://www.pewresearch.org/internet/2011/02/01/health-topics-2/> (accessed 4.25.21).
- Fox, S., Duggan, M., 2013. Health Online 2013. Pew Research Center: Internet, Science & Tech. URL <https://www.pewresearch.org/internet/2013/01/15/health-online-2013/> (accessed 9.22.20).
- Frampton, G., Woods, L., Scott, D.A., 2021. Inconsistent and incomplete retraction of published research: A cross-sectional study on Covid-19 retractions and recommendations to mitigate risks for research, policy and practice. *PLoS One* 16, e0258935.
<https://doi.org/10.1371/journal.pone.0258935>
- Frank, M., 2013. Open but not free—publishing in the 21st century. *N Engl J Med* 368, 787–789.
<https://doi.org/10.1056/NEJMp1211259>
- Fraser, N., Brierley, L., Dey, G., Polka, J.K., Pálffy, M., Nanni, F., Coates, J.A., 2021. The evolving role of preprints in the dissemination of COVID-19 research and their impact on the science communication landscape. *PLoS Biol* 19, e3000959.
<https://doi.org/10.1371/journal.pbio.3000959>
- Funk, M.E., Reid, C.A., 1983. Indexing consistency in MEDLINE. *Bull Med Libr Assoc* 71, 176–183.
- Ganjavi, C., Eppler, M.B., Pekcan, A., Biedermann, B., Abreu, A., Collins, G.S., Gill, I.S., Cacciamani, G.E., 2024. Publishers’ and journals’ instructions to authors on use of generative artificial intelligence in academic and scientific publishing: bibliometric analysis. *BMJ* 384, e077192. <https://doi.org/10.1136/bmj-2023-077192>
- Garisto, D., 2022. ArXiv.org Reaches a Milestone and a Reckoning. *Scientific American*. URL <https://www.scientificamerican.com/article/arxiv-org-reaches-a-milestone-and-a-reckoning/> (accessed 4.22.22).
- Gienapp, L., Scells, H., Deckers, N., Bevendorff, J., Wang, S., Kiesel, J., Syed, S., Fröbe, M., Zuccon, G., Stein, B., Hagen, M., Potthast, M., 2014. Evaluating Generative Ad Hoc Information Retrieval, in: *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*.

- Graña Possamai, C., Cabanac, G., Perrodeau, E., Ghosn, L., Ravaud, P., Boutron, I., 2025. Inclusion of Retracted Studies in Systematic Reviews and Meta-Analyses of Interventions: A Systematic Review and Meta-Analysis. *JAMA Intern Med* e250256. <https://doi.org/10.1001/jamainternmed.2025.0256>
- Gray, A., 2024. ChatGPT “contamination”: estimating the prevalence of LLMs in the scholarly literature. <https://doi.org/10.48550/arXiv.2403.16887>
- Grundtvig, A., Munsie, M., 2025. Stem cells in your YouTube feed: The complex landscape and unexpected mediators of online information. *Stem Cell Reports* 20, 102719. <https://doi.org/10.1016/j.stemcr.2025.102719>
- Gu, J., Wang, X., Li, C., Zhao, J., Fu, W., Liang, G., Qiu, J., 2022. AI-enabled image fraud in scientific publications. *Patterns (N Y)* 3, 100511. <https://doi.org/10.1016/j.patter.2022.100511>
- Gupta, D., Demner-Fushman, D., Hersh, W., Bedrick, S., Roberts, K., 2024. Overview of TREC 2024 Biomedical Generative Retrieval (BioGen) Track. <https://doi.org/10.48550/arXiv.2411.18069>
- Haider, J., Söderström, K.R., Ekström, B., Rödl, M., 2024. GPT-fabricated scientific papers on Google Scholar: Key features, spread, and implications for preempting evidence manipulation. *Harvard Kennedy School Misinformation Review*. <https://doi.org/10.37016/mr-2020-156>
- Hanauer, D.A., Barnholtz-Sloan, J.S., Beno, M.F., Del Fiol, G., Durbin, E.B., Gologorskaya, O., Harris, D., Harnett, B., Kawamoto, K., May, B., Meeks, E., Pfaff, E., Weiss, J., Zheng, K., 2020. Electronic Medical Record Search Engine (EMERSE): An Information Retrieval Tool for Supporting Cancer Research. *JCO Clin Cancer Inform* 4, 454–463. <https://doi.org/10.1200/CCI.19.00134>
- Hanson, M.A., Barreiro, P.G., Crosetto, P., Brockington, D., 2024. The strain on scientific publishing. *Quantitative Science Studies* 5, 823–843. https://doi.org/10.1162/qss_a_00327
- Haug, C., 2013. The downside of open-access publishing. *N Engl J Med* 368, 791–793. <https://doi.org/10.1056/NEJMp1214750>
- Haynes, R.B., McKibbon, K.A., Walker, C.J., Ryan, N., Fitzgerald, D., Ramsden, M.F., 1990. Online access to MEDLINE in clinical settings. A study of use and usefulness. *Ann Intern Med* 112, 78–84. <https://doi.org/10.7326/0003-4819-112-1-78>
- Heilman, J., 2013. Online encyclopedia provides free health info for all. Interview by Fiona Fleck. *Bull World Health Organ* 91, 8–9. <https://doi.org/10.2471/BLT.13.030113>
- Hepler, R., 2025. The Perils of Insta-Research. *CollaborAltion*. URL <https://reedhepler.substack.com/p/the-perils-of-insta-research> (accessed 2.16.26).
- Hersh, W., 2021. Information Retrieval, in: Shortliffe, E.H., Cimino, J. (Eds.), *Biomedical Informatics: Computer Applications in Health Care and Biomedicine*. Springer International Publishing, pp. 761–800. <https://doi.org/10.1007/978-3-030-58721-5>
- Hersh, W., 2020. *Information Retrieval: A Biomedical and Health Perspective*, 4th ed, Health Informatics. Springer International Publishing.
- Hersh, W., 2018. Estrategias Para la Recuperación de la Información, in: Sánchez Mendiola, M., Martínez Franco, A.I. (Eds.), *Informática biomédica*. Elsevier.
- Hersh, W., 1994. Relevance and retrieval evaluation: Perspectives from medicine. *Journal of the American Society for Information Science* 45, 201–206. [https://doi.org/10.1002/\(SICI\)1097-4571\(199404\)45:3%3C201::AID-ASI9%3E3.0.CO;2-W](https://doi.org/10.1002/(SICI)1097-4571(199404)45:3%3C201::AID-ASI9%3E3.0.CO;2-W)

- Hersh, W., Buckley, C., Leone, T.J., Hickam, D., 1994. OHSUMED: an interactive retrieval evaluation and new large test collection for research, in: Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '94. Springer-Verlag, Berlin, Heidelberg, pp. 192–201.
- Hersh, W., Müller, H., Kalpathy-Cramer, J., 2009. The ImageCLEFmed medical image retrieval task test collection. *J Digit Imaging* 22, 648–655. <https://doi.org/10.1007/s10278-008-9154-8>
- Hersh, W., Voorhees, E., 2009. TREC genomics special issue overview. *Inf Retrieval* 12, 1–15. <https://doi.org/10.1007/s10791-008-9076-6>
- Hersh, W.R., 2018. Information Retrieval From Medical Knowledge Resources, in: Hoyt, R.E., Hersh, W.R. (Eds.), *Health Informatics: Practical Guide Seventh Edition*. Informatics Education, pp. 307–326.
- Hersh, W.R., Bhupatiraju, R.T., Greene, P., Smothers, V., Cohen, C., 2006a. Adopting e-learning standards in health care: competency-based learning in the medical informatics domain. *AMIA Annu Symp Proc* 334–338.
- Hersh, W.R., Bhupatiraju, R.T., Ross, L., Roberts, P., Cohen, A.M., Kraemer, D.F., 2006b. Enhancing access to the Bibliome: the TREC 2004 Genomics Track. *J Biomed Discov Collab* 1, 3. <https://doi.org/10.1186/1747-5333-1-3>
- Hersh, W.R., Crabtree, M.K., Hickam, D.H., Sacherek, L., Friedman, C.P., Tidmarsh, P., Mosbaek, C., Kraemer, D., 2002. Factors associated with success in searching MEDLINE and applying evidence to answer clinical questions. *J Am Med Inform Assoc* 9, 283–293. <https://doi.org/10.1197/jamia.m0996>
- Hersh, W.R., Crabtree, M.K., Hickam, D.H., Sacherek, L., Rose, L., Friedman, C.P., 2000. Factors associated with successful answering of clinical questions using an information retrieval system. *Bull Med Libr Assoc* 88, 323–331.
- Hersh, W.R., Greenes, R.A., 1990. SAPHIRE--an information retrieval system featuring concept matching, automatic indexing, probabilistic retrieval, and hierarchical relationships. *Comput. Biomed. Res.* 23, 410–425. [https://doi.org/10.1016/0010-4809\(90\)90031-7](https://doi.org/10.1016/0010-4809(90)90031-7)
- Hersh, W.R., Hickam, D.H., 1998. How well do physicians use electronic information retrieval systems? A framework for investigation and systematic review. *JAMA* 280, 1347–1352. <https://doi.org/10.1001/jama.280.15.1347>
- Hersh, W.R., Müller, H., Jensen, J.R., Yang, J., Gorman, P.N., Ruch, P., 2006c. Advancing biomedical image retrieval: development and analysis of a test collection. *J Am Med Inform Assoc* 13, 488–496. <https://doi.org/10.1197/jamia.M2082>
- Hersh, W.R., Rindfleisch, T.C., 2000. Electronic publishing of scholarly communication in the biomedical sciences. *J Am Med Inform Assoc* 7, 324–325. <https://doi.org/10.1136/jamia.2000.0070324>
- Himelein-Wachowiak, M., Giorgi, S., Devoto, A., Rahman, M., Ungar, L., Schwartz, H.A., Epstein, D.H., Leggio, L., Curtis, B., 2021. Bots and Misinformation Spread on Social Media: Implications for COVID-19. *J Med Internet Res* 23, e26933. <https://doi.org/10.2196/26933>
- Honan, M., 2025. AI means the end of internet search as we've known it [WWW Document]. MIT Technology Review. URL <https://www.technologyreview.com/2025/01/06/1108679/ai-generative-search-internet-breakthroughs/> (accessed 4.28.25).

- Hopkins, A.M., Logan, J.M., Kichenadasse, G., Sorich, M.J., 2023. Artificial intelligence chatbots will revolutionize how cancer patients access information: ChatGPT represents a paradigm-shift. *JNCI Cancer Spectr* 7, pkad010. <https://doi.org/10.1093/jncics/pkad010>
- Huber, J., Inoua, S., Kerschbamer, R., König-Kersting, C., Palan, S., Smith, V.L., 2022. Nobel and novice: Author prominence affects peer review. *Proc Natl Acad Sci U S A* 119, e2205779119. <https://doi.org/10.1073/pnas.2205779119>
- Huesch, M.D., 2013. Privacy threats when seeking online health information. *JAMA Intern Med* 173, 1838–1839. <https://doi.org/10.1001/jamainternmed.2013.7795>
- Ibrahim, H., Liu, F., Zaki, Y., Rahwan, T., 2025. Citation manipulation through citation mills and pre-print servers. *Sci Rep* 15, 5480. <https://doi.org/10.1038/s41598-025-88709-7>
- Iliadis, A., Acker, A., Stevens, W., Kavakli, S.B., n.d. One schema to rule them all: How Schema.org models the world of search. *Journal of the Association for Information Science and Technology* n/a. <https://doi.org/10.1002/asi.24744>
- Inglefinger, F., 1969. Definition of “sole contribution.” *N Engl J Med* 281, 676–677. <https://doi.org/10.1056/NEJM196909182811208>
- Insel, T.R., Volkow, N.D., Li, T.-K., Battey, J.F., Landis, S.C., 2003. Neuroscience networks: data-sharing in an information age. *PLoS Biol* 1, E17. <https://doi.org/10.1371/journal.pbio.0000017>
- Janda, G., Khetpal, V., Shi, X., Ross, J.S., Wallach, J.D., 2022. Comparison of Clinical Study Results Reported in medRxiv Preprints vs Peer-reviewed Journal Articles. *JAMA Netw Open* 5, e2245847. <https://doi.org/10.1001/jamanetworkopen.2022.45847>
- Jarry, J., 2024. The Science Journals That Will Publish Anything [WWW Document]. Office for Science and Society. URL <https://www.mcgill.ca/oss/article/critical-thinking-general-science/science-journals-will-publish-anything> (accessed 5.5.24).
- Jażwińska, K., Chandrasekar, A., 2025. AI Search Has A Citation Problem [WWW Document]. *Columbia Journalism Review*. URL https://www.cjr.org/tow_center/we-compared-eight-ai-search-engines-theyre-all-bad-at-citing-news.php (accessed 4.28.25).
- Jiang, S., Shen, Z., Agrawal, M., Lam, B., Kurtzman, N., Horng, S., Karger, D., Sontag, D., 2023. Conceptualizing Machine Learning for Dynamic Information Retrieval of Electronic Health Record Notes, in: *Machine Learning for Healthcare*.
- Jin, Q., Kim, W., Chen, Q., Comeau, D.C., Yeganova, L., Wilbur, W.J., Lu, Z., 2023. MedCPT: Contrastive Pre-trained Transformers with large-scale PubMed search logs for zero-shot biomedical information retrieval. *Bioinformatics* 39, btad651. <https://doi.org/10.1093/bioinformatics/btad651>
- Jin, Q., Wang, Z., Floudas, C.S., Chen, F., Gong, C., Bracken-Clarke, D., Xue, E., Yang, Y., Sun, J., Lu, Z., 2024. Matching patients to clinical trials with large language models. *Nat Commun* 15, 9074. <https://doi.org/10.1038/s41467-024-53081-z>
- Jones, N., 2026. Leading preprint server clamps down on “AI slop.” *Science* 391, 432–433. <https://doi.org/10.1126/science.aef8896>
- Kataoka, Y., Banno, M., Tsujimoto, Y., Ariie, T., Taito, S., Suzuki, T., Oide, S., Furukawa, T.A., 2022. Retracted randomized controlled trials were cited and not corrected in systematic reviews and clinical practice guidelines. *J Clin Epidemiol* 150, 90–97. <https://doi.org/10.1016/j.jclinepi.2022.06.015>
- Kember, S., Brand, A., 2023. Opinion | The Corporate Capture of Open-Access Publishing [WWW Document]. *The Chronicle of Higher Education*. URL

- <https://www.chronicle.com/article/the-corporate-capture-of-open-access-publishing> (accessed 5.5.24).
- Khan, H., Vieira Armond, A.C., Ghannad, M., Moher, D., 2022. Disseminating Biomedical Research: Predatory Journals and Practices. *Indian Journal of Rheumatology* 17, S328. <https://doi.org/10.4103/0973-3698.364675>
- Khan, N., 2024. Retrieval Techniques - Sparse, Dense, and Hybrid Representations [WWW Document]. LinkedIn. URL <https://www.linkedin.com/pulse/retrieval-techniques-sparse-dense-hybrid-najeeb-khan-ph-d--wmtpc/> (accessed 4.28.25).
- Khullar, D., 2022. Social Media and Medical Misinformation: Confronting New Variants of an Old Problem. *JAMA* 328, 1393–1394. <https://doi.org/10.1001/jama.2022.17191>
- Kim, H., Sohn, J., Gilson, A., Cochran-Caggiano, N., Applebaum, S., Jin, H., Park, S., Park, Y., Park, J., Choi, S., Contreras, B.A.H., Huang, T., Yun, J., Wei, E.F., Jiang, R., Colucci, L., Lai, E., Dave, A., Guo, T., Singer, M.B., Koo, Y., Adelman, R.A., Zou, J., Taylor, A., Cohan, A., Xu, H., Chen, Q., 2025. Rethinking Retrieval-Augmented Generation for Medicine: A Large-Scale, Systematic Expert Evaluation and Practical Insights. <https://doi.org/10.48550/arXiv.2511.06738>
- Kimothi, A., 2025. A Simple Guide to Retrieval Augmented Generation.
- Kincaid, M.L., Fleisher, L.A., Neuman, M.D., 2015. Presentation on US hospital websites of risks and benefits of transcatheter aortic valve replacement procedures. *JAMA Intern Med* 175, 440–441. <https://doi.org/10.1001/jamainternmed.2014.7392>
- Kodvanj, I., Homolak, J., Virag, D., Trkulja, V., 2022. Publishing of COVID-19 preprints in peer-reviewed journals, preprinting trends, public discussion and quality issues. *Scientometrics* 1–14. <https://doi.org/10.1007/s11192-021-04249-7>
- Kohl, C.B.S., Faggion, C.M., 2023. A comprehensive overview of studies that assessed article retractions within the biomedical sciences. *Account Res* 1–19. <https://doi.org/10.1080/08989621.2022.2154660>
- Koo, M., Lin, S.-C., 2024. Retracted articles in scientific literature: A bibliometric analysis from 2003 to 2022 using the Web of Science. *Heliyon* 10, e38620. <https://doi.org/10.1016/j.heliyon.2024.e38620>
- Koopman, B., Zuccon, G., 2023. Dr ChatGPT tell me what I want to hear: How different prompts impact health answer correctness, in: Bouamor, H., Pino, J., Bali, K. (Eds.), *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*. Presented at the EMNLP 2023, Association for Computational Linguistics, Singapore, pp. 15012–15022. <https://doi.org/10.18653/v1/2023.emnlp-main.928>
- Kozlov, M., 2022. NIH issues a seismic mandate: share data publicly. *Nature* 602, 558–559. <https://doi.org/10.1038/d41586-022-00402-1>
- Kusa, W., Mendoza, Ó.E., Knoth, P., Pasi, G., Hanbury, A., 2023. Effective matching of patients to clinical trials using entity extraction and neural re-ranking. *J Biomed Inform* 144, 104444. <https://doi.org/10.1016/j.jbi.2023.104444>
- Kwon, D., 2025. Science sleuths flag hundreds of papers that use AI without disclosing it. *Nature*. <https://doi.org/10.1038/d41586-025-01180-2>
- Laine, C., Babski, D., Bachelet, V.C., Bärnighausen, T.W., Baethge, C., Bibbins-Domingo, K., Frizelle, F., Gollogly, L., Kleinert, S., Loder, E., Monteiro, J., Rubin, E.J., Sahni, P., Wee, C.C., Yoo, J.-H., Zakhama, L., 2025. Predatory Journals: What Can We Do to Protect Their Prey? *Ann Intern Med* 178, 294–296. <https://doi.org/10.7326/ANNALS-24-03636>

- Laine, C., De Angelis, C., Delamothe, T., Drazen, J.M., Frizelle, F.A., Haug, C., Hébert, P.C., Horton, R., Kotzin, S., Marusic, A., Sahni, P., Schroeder, T.V., Sox, H.C., Van der Weyden, M.B., Verheugt, F.W.A., 2007. Clinical trial registration: looking back and moving ahead. *Ann Intern Med* 147, 275–277. <https://doi.org/10.7326/0003-4819-147-4-200708210-00166>
- Laurent, M.R., Vickers, T.J., 2009. Seeking health information online: does Wikipedia matter? *J Am Med Inform Assoc* 16, 471–479. <https://doi.org/10.1197/jamia.M3059>
- LeBlanc, A.G., Barnes, J.D., Saunders, T.J., Tremblay, M.S., Chaput, J.-P., 2023. Scientific sinkhole: estimating the cost of peer review based on survey data with snowball sampling. *Res Integr Peer Rev* 8, 3. <https://doi.org/10.1186/s41073-023-00128-2>
- Lenzer, J., Brownlee, S., 2020. Pandemic Science Out of Control. *Issues in Science and Technology*. URL <https://issues.org/pandemic-science-out-of-control/> (accessed 8.20.20).
- Levin, J.M., Bukowski, L.A., Minson, J.A., Kahn, J.M., 2023. The political polarization of COVID-19 treatments among physicians and laypeople in the United States. *Proc Natl Acad Sci U S A* 120, e2216179120. <https://doi.org/10.1073/pnas.2216179120>
- Liang, W., Izzo, Z., Zhang, Y., Lepp, H., Cao, H., Zhao, X., Chen, L., Ye, H., Liu, S., Huang, Z., McFarland, D.A., Zou, J.Y., 2024. Monitoring AI-Modified Content at Scale: A Case Study on the Impact of ChatGPT on AI Conference Peer Reviews. <https://doi.org/10.48550/arXiv.2403.07183>
- Libert, T., 2015. Privacy implications of health information seeking on the web. *Commun. ACM* 58, 68–77. <https://doi.org/10.1145/2658983>
- Liu, S., McCoy, A.B., Wright, A., 2025. Improving large language model applications in biomedicine with retrieval-augmented generation: a systematic review, meta-analysis, and clinical development guidelines. *J Am Med Inform Assoc* 32, 605–615. <https://doi.org/10.1093/jamia/ocaf008>
- Liu, Y., Moosavi, N.S., Lin, C., 2023. LLMs as Narcissistic Evaluators: When Ego Inflates Evaluation Scores. <https://doi.org/10.48550/arXiv.2311.09766>
- Mager, D., 2022. Types of open access publishing and its benefits – Researcher.Life. Researcher.Life. URL <https://researcher.life/blog/article/types-of-open-access-publications/> (accessed 4.28.23).
- Maggio, L.A., Aakre, C.A., Del Fiore, G., Shellum, J., Cook, D.A., 2019. Impact of Clinicians' Use of Electronic Knowledge Resources on Clinical and Learning Outcomes: Systematic Review and Meta-Analysis. *J Med Internet Res* 21, e13315. <https://doi.org/10.2196/13315>
- Majumder, M.S., Mandl, K.D., 2020. Early in the epidemic: impact of preprints on global discourse about COVID-19 transmissibility. *The Lancet Global Health* 8, e627–e630. [https://doi.org/10.1016/S2214-109X\(20\)30113-3](https://doi.org/10.1016/S2214-109X(20)30113-3)
- Marcetich, J., Rappaport, M., Kotzin, S., 2004. Indexing consistency in MEDLINE, in: *MLA 04 Abstracts*. Presented at the Medical Library Association.
- Marcus, A., Abris, A.J., Oransky, I., 2022. How to Stop the Unknowing Citation of Retracted Papers. *Anesthesiology* 137, 280–282. <https://doi.org/10.1097/ALN.0000000000004333>
- Markoff, J., 2011. Computer Wins on 'Jeopardy!': Trivial, It's Not. *The New York Times*.
- Maslove, D.M., 2018. Medical Preprints-A Debate Worth Having. *JAMA* 319, 443–444. <https://doi.org/10.1001/jama.2017.17566>
- Mastroianni, A., 2022. The rise and fall of peer review [WWW Document]. *Experimental History*. URL <https://www.experimental-history.com/p/the-rise-and-fall-of-peer-review> (accessed 4.24.23).
- McCool, J.H., 2017. Opinion: Why I Published in a Predatory Journal. *The Scientist Magazine*®.

- McLoughlin, K.L., Brady, W.J., Goolsbee, A., Kaiser, B., Klonick, K., Crockett, M.J., 2024. Misinformation exploits outrage to spread online. *Science* 386, 991–996. <https://doi.org/10.1126/science.adl2829>
- Meagher, K., 2021. Introduction: The Politics of Open Access — Decolonizing Research or Corporate Capture? *Development and Change* 52, 340–358. <https://doi.org/10.1111/dech.12630>
- MEDLINE 2022 Initiative: Transition to Automated Indexing, 2021. . *NLM Technical Bulletin* 443, e5.
- Mehregan, M., 2022. Scientific journals must be alert to potential manipulation in citations and referencing. *Research Ethics* 18, 163–168. <https://doi.org/10.1177/17470161211068745>
- Merson, L., Gaye, O., Guerin, P.J., 2016. Avoiding Data Dumpsters--Toward Equitable and Useful Data Sharing. *N Engl J Med* 374, 2414–2415. <https://doi.org/10.1056/NEJMp1605148>
- Meserole, C., 2018. How misinformation spreads on social media—And what to do about it. Brookings. URL <https://www.brookings.edu/blog/order-from-chaos/2018/05/09/how-misinformation-spreads-on-social-media-and-what-to-do-about-it/> (accessed 4.22.22).
- Moher, D., Moher, E., 2016. Stop Predatory Publishers Now: Act Collaboratively. *Ann Intern Med* 164, 616–617. <https://doi.org/10.7326/M15-3015>
- Monigatti, L., 2023. Retrieval-Augmented Generation (RAG): From Theory to LangChain Implementation [WWW Document]. Medium. URL <https://towardsdatascience.com/retrieval-augmented-generation-rag-from-theory-to-langchain-implementation-4e9bd5f6a4f2> (accessed 5.2.24).
- Mulins, M., 2021. Opinion: The Problem with Preprints. *The Scientist Magazine*®. URL <https://www.the-scientist.com/critic-at-large/opinion-the-problem-with-preprints-69309> (accessed 4.22.22).
- Naddaf, M., 2025. AI is transforming peer review - and many scientists are worried. *Nature* 639, 852–854. <https://doi.org/10.1038/d41586-025-00894-7>
- National Academies of Sciences, Engineering, and Medicine, 2018. Open Science by Design: Realizing a Vision for 21st Century Research. <https://doi.org/10.17226/25116>
- Neil, S.J.D., Campbell, E.M., 2020. Fake Science: XMRV, COVID-19, and the Toxic Legacy of Dr. Judy Mikovits. *AIDS Research and Human Retroviruses* 36, 545–549. <https://doi.org/10.1089/aid.2020.0095>
- Nelson, J.T., Tse, T., Pupilampu-Dove, Y., Golfopoulos, E., Zarin, D.A., 2023. Comparison of Availability of Trial Results in ClinicalTrials.gov and PubMed by Data Source and Funder Type. *JAMA* e232351. <https://doi.org/10.1001/jama.2023.2351>
- Nievas, M., Basu, A., Wang, Y., Singh, H., 2024. Distilling large language models for matching patients to clinical trials. *J Am Med Inform Assoc* ocae073. <https://doi.org/10.1093/jamia/ocae073>
- O’Leary, N.A., Cox, E., Holmes, J.B., Anderson, W.R., Falk, R., Hem, V., Tsuchiya, M.T.N., Schuler, G.D., Zhang, X., Torcivia, J., Ketter, A., Breen, L., Cothran, J., Bajwa, H., Tinne, J., Meric, P.A., Hlavina, W., Schneider, V.A., 2024. Exploring and retrieving sequence and metadata for species across the tree of life with NCBI Datasets. *Sci Data* 11, 732. <https://doi.org/10.1038/s41597-024-03571-y>
- Oransky, I., 2022. Retractions are increasing, but not enough. *Nature* 608, 9. <https://doi.org/10.1038/d41586-022-02071-6>

- Osman, W., Mohamed, F., Elhassan, M., Shoufan, A., 2022. Is YouTube a reliable source of health-related information? A systematic review. *BMC Med Educ* 22, 382. <https://doi.org/10.1186/s12909-022-03446-z>
- Oxman, A.D., Paulsen, E.J., 2019. Who can you trust? A review of free online sources of “trustworthy” information about treatment effects for patients and the public. *BMC Med Inform Decis Mak* 19, 35. <https://doi.org/10.1186/s12911-019-0772-5>
- Parker, L., Boughton, S., Bero, L., Byrne, J.A., 2024. Paper mill challenges: past, present, and future. *J Clin Epidemiol* 176, 111549. <https://doi.org/10.1016/j.jclinepi.2024.111549>
- Pasternak, N., Orsi, C., Mertz, A.F., Firestein, S., 2022. The Attack of Zombie Science [WWW Document]. *Nautilus | Science Connected*. URL <https://nautil.us/the-attack-of-zombie-science-13417/> (accessed 4.22.22).
- Perlis, R.H., Lunz Trujillo, K., Green, J., Safarpour, A., Druckman, J.N., Santillana, M., Ognyanova, K., Lazer, D., 2023. Misinformation, Trust, and Use of Ivermectin and Hydroxychloroquine for COVID-19. *JAMA Health Forum* 4, e233257. <https://doi.org/10.1001/jamahealthforum.2023.3257>
- Perrin, A., Atske, S., 2021a. About three-in-ten U.S. adults say they are ‘almost constantly’ online. *Pew Research Center*. URL <https://www.pewresearch.org/fact-tank/2021/03/26/about-three-in-ten-u-s-adults-say-they-are-almost-constantly-online/> (accessed 4.25.21).
- Perrin, A., Atske, S., 2021b. 7% of Americans don’t use the internet. Who are they? *Pew Research Center*. URL <https://www.pewresearch.org/fact-tank/2021/04/02/7-of-americans-dont-use-the-internet-who-are-they/> (accessed 4.25.21).
- Peterson, C.J., Anderson, C., Nugent, K., 2022. Continued Visibility of COVID-19 Article Removals. *South Med J* 115, 371–373. <https://doi.org/10.14423/SMJ.0000000000001397>
- Pluye, P., Grad, R.M., 2004. How information retrieval technology may impact on physician practice: an organizational case study in family medicine. *J Eval Clin Pract* 10, 413–430. <https://doi.org/10.1111/j.1365-2753.2004.00498.x>
- Pluye, P., Grad, R.M., Dunikowski, L.G., Stephenson, R., 2005. Impact of clinical information-retrieval technology on physicians: a literature review of quantitative, qualitative and mixed methods studies. *Int J Med Inform* 74, 745–768. <https://doi.org/10.1016/j.ijmedinf.2005.05.004>
- Puebla, I., Polka, J., Rieger, O.Y., 2022. Preprints: Their Evolving Role in Science Communication. *Against the Grain (Media), LLC*. <https://doi.org/10.3998/mpub.12412508>
- Rae, A.R., Mork, J.G., Demner-Fushman, D., 2021. A Neural Text Ranking Approach for Automatic MeSH Indexing, in: *CLEF 2021 – Conference and Labs of the Evaluation Forum*. Presented at the CLEF 2021 – Conference and Labs of the Evaluation Forum, pp. 302–312.
- Rahmani, H.A., Siro, C., Aliannejadi, M., Craswell, N., Clarke, C.L.A., Faggioli, G., Mitra, B., Thomas, P., Yilmaz, E., 2025. Judging the Judges: A Collection of LLM-Generated Relevance Judgements. <https://doi.org/10.48550/arXiv.2502.13908>
- Reznick, J.S., Funk, K., Randall, L., Taylor, K.R., Chan, K., Adams, E., 2025. PubMed Central: offering a trove of historical medical journals. *J Med Libr Assoc* 113, 383–386. <https://doi.org/10.5195/jmla.2025.2235>
- Richardson, R.A.K., Hong, S.S., Byrne, J.A., Stoeger, T., Amaral, L.A.N., 2025. The entities enabling scientific fraud at scale are large, resilient, and growing rapidly. *Proc Natl Acad Sci U S A* 122, e2420092122. <https://doi.org/10.1073/pnas.2420092122>

- Rigden, D.J., Fernández, X.M., 2024. The 2024 Nucleic Acids Research database issue and the online molecular biology database collection. *Nucleic Acids Res* 52, D1–D9.
<https://doi.org/10.1093/nar/gkad1173>
- Ritchie, S., 2022. The big idea: should we get rid of the scientific paper? *The Guardian*.
- Rivero-de-Aguilar, A., Pérez-Ríos, M., Ruano-Raviña, A., Candal-Pedreira, C., Puente-Hernandez, M., Ross, J.S., Varela-Lema, L., 2023. Evidence of publication bias in multiple sclerosis clinical trials: a comparative analysis of published and unpublished studies registered in ClinicalTrials.gov. *J Neurol Neurosurg Psychiatry* 94, 597–604.
<https://doi.org/10.1136/jnnp-2023-331132>
- Ro, C., Leeming, J., 2025. Authorship for sale: Nature investigates how paper mills work. *Nature* 642, 823–826. <https://doi.org/10.1038/d41586-025-01824-3>
- Robertson, C.E., Del Rosario, K.S., Van Bavel, J.J., 2024. Inside the funhouse mirror factory: How social media distorts perceptions of norms. *Curr Opin Psychol* 60, 101918.
<https://doi.org/10.1016/j.copsy.2024.101918>
- Robertson, S.E., Walker, S., 1994. Some simple effective approximations to the 2-Poisson model for probabilistic weighted retrieval, in: *Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '94*. Springer-Verlag, Berlin, Heidelberg, pp. 232–241.
- Rodwin, M.A., Abramson, J.D., 2012. Clinical trial data as a public good. *JAMA* 308, 871–872.
<https://doi.org/10.1001/jama.2012.9661>
- Rosenthal, R., 1979. The file drawer problem and tolerance for null results. *Psychological Bulletin* 86, 638–641. <https://doi.org/10.1037/0033-2909.86.3.638>
- Ross, J.S., Krumholz, H.M., 2013. Ushering in a new era of open science through data sharing: the wall must come down. *JAMA* 309, 1355–1356. <https://doi.org/10.1001/jama.2013.1299>
- Ross, J.S., Waldstreicher, J., Krumholz, H.M., 2023. Data Sharing - A New Era for Research Funded by the U.S. Government. *N Engl J Med* 389, 2408–2410.
<https://doi.org/10.1056/NEJMp2308792>
- Royle, J.A., Blythe, J., Potvin, C., Oolup, P., Chan, I.M., 1995. Literature search and retrieval in the workplace. *Comput Nurs* 13, 25–31.
- Rubin, R., 2022. When Physicians Spread Unscientific Information About COVID-19. *JAMA* 327, 904–906. <https://doi.org/10.1001/jama.2022.1083>
- Russell-Rose, T., Chamberlain, J., 2017. Expert Search Strategies: The Information Retrieval Practices of Healthcare Information Professionals. *JMIR Med Inform* 5, e33.
<https://doi.org/10.2196/medinform.7680>
- Russian site peddles paper authorship in reputable journals for up to \$5000 a pop [WWW Document], n.d. URL <https://www.science.org/content/article/russian-website-peddles-authorships-linked-reputable-journals> (accessed 4.22.22).
- Sabel, B.A., Knaack, E., Gigerenzer, G., Bilek, M., 2023. Fake Publications in Biomedical Science: Red-flagging Method Indicates Mass Production.
<https://doi.org/10.1101/2023.05.06.23289563>
- Sage, W.M., Yang, Y.T., 2022. Reducing “COVID-19 Misinformation” While Preserving Free Speech. *JAMA* 327, 1443–1444. <https://doi.org/10.1001/jama.2022.4231>
- Salton, G., 1991. Developments in automatic text retrieval. *Science* 253, 974–980.
<https://doi.org/10.1126/science.253.5023.974>

- Sarwar, N., 2025. Google's AI could lead you into scam support numbers on Search. Digital Trends. URL <https://www.digitaltrends.com/computing/googles-ai-could-lead-you-into-scam-support-numbers-on-search/> (accessed 2.20.26).
- Sayers, E.W., Beck, J., Bolton, E.E., Brister, J.R., Chan, J., Comeau, D.C., Connor, R., DiCuccio, M., Farrell, C.M., Feldgarden, M., Fine, A.M., Funk, K., Hatcher, E., Hoepfner, M., Kane, M., Kannan, S., Katz, K.S., Kelly, C., Klimke, W., Kim, S., Kimchi, A., Landrum, M., Lathrop, S., Lu, Z., Malheiro, A., Marchler-Bauer, A., Murphy, T.D., Phan, L., Prasad, A.B., Pujar, S., Sawyer, A., Schmieder, E., Schneider, V.A., Schoch, C.L., Sharma, S., Thibaud-Nissen, F., Trawick, B.W., Venkatapathi, T., Wang, J., Pruitt, K.D., Sherry, S.T., 2024. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 52, D33–D43. <https://doi.org/10.1093/nar/gkad1044>
- Sayers, E.W., Bolton, E.E., Fine, A.M., Kelly, C., Kim, S., Landrum, M., Lathrop, S., Malheiro, A., Murphy, T.D., Phan, L., Pujar, S., Trawick, B.W., Schneider, V.A., Pruitt, K.D., 2026. Database resources of the National Center for Biotechnology Information in 2026. *Nucleic Acids Res* 54, D20–D27. <https://doi.org/10.1093/nar/gkaf1060>
- Sbaffi, L., Rowley, J., 2017. Trust and Credibility in Web-Based Health Information: A Review and Agenda for Future Research. *J Med Internet Res* 19, e218. <https://doi.org/10.2196/jmir.7579>
- Scaffidi, M.A., Khan, R., Wang, C., Keren, D., Tsui, C., Garg, A., Brar, S., Valoo, K., Bonert, M., de Wolff, J.F., Heilman, J., Grover, S.C., 2017. Comparison of the Impact of Wikipedia, UpToDate, and a Digital Textbook on Short-Term Knowledge Acquisition Among Medical Students: Randomized Controlled Trial of Three Web-Based Resources. *JMIR Med Educ* 3, e20. <https://doi.org/10.2196/mededu.8188>
- Schoolcraft, D., Meltzer, A.C., Sangal, R., Terry, A.T., Robertson, K., Buckland, D., Motalib, S., Genes, N., Vukmir, R., Waseem, T., ACEP AI TASK FORCE, 2026. Health Insurance Portability and Accountability Act Liability in the Age of Generative Artificial Intelligence. *J Am Coll Emerg Physicians Open* 7, 100317. <https://doi.org/10.1016/j.acepjo.2025.100317>
- Sell, T.K., Hosangadi, D., Smith, E., Trotochaud, M., Vasudevan, P., Gronvall, G.K., Rivera, Y., Sutton, J., Ruiz, A., Cicero, A., 2021. National Priorities to Combat Misinformation and Disinformation for COVID-19 and Future Public Health Threats: A Call for a National Strategy. Johns Hopkins Center for Health Security.
- Shafee, T., Masukume, G., Kipersztok, L., Das, D., Häggström, M., Heilman, J., 2017. Evolution of Wikipedia's medical content: past, present and future. *J Epidemiol Community Health* 71, 1122–1129. <https://doi.org/10.1136/jech-2016-208601>
- Shah, C., 2023. AI information retrieval: A search engine researcher explains the promise and peril of letting ChatGPT and its cousins search the web for you [WWW Document]. The Conversation. URL <http://theconversation.com/ai-information-retrieval-a-search-engine-researcher-explains-the-promise-and-peril-of-letting-chatgpt-and-its-cousins-search-the-web-for-you-200875> (accessed 8.22.23).
- Shah, C., Bender, E.M., 2024. Envisioning Information Access Systems: What Makes for Good Tools and a Healthy Web? *ACM Trans. Web* 18, 33:1-33:24. <https://doi.org/10.1145/3649468>
- Sharfstein, J.M., 2022. Battling Falsehoods in the COVID-19 Pandemic. *JAMA Health Forum* 3, e221587. <https://doi.org/10.1001/jamahealthforum.2022.1587>
- Shaw, D., 2025. The digital erosion of intellectual integrity: why misuse of generative AI is worse than plagiarism. *AI & Soc.* <https://doi.org/10.1007/s00146-025-02362-2>

- Silberg, W.M., Lundberg, G.D., Musacchio, R.A., 1997. Assessing, controlling, and assuring the quality of medical information on the Internet: Caveant lector et viewor--Let the reader and viewer beware. *JAMA* 277, 1244–1245.
- Smalley, S., 2021. As Misinformation Grows, Scholars Debate How to Improve Open Access [WWW Document]. Inside Higher Ed. URL <https://www.insidehighered.com/news/2021/11/08/open-access-science-misinformation-era> (accessed 4.22.22).
- Soboroff, I., 2025. Don't Use LLMs to Make Relevance Judgments. *IRRJ* 1, 29–46. <https://doi.org/10.54195/irrj.19625>
- Sox, H.C., 2009. Medical journal editing: who shall pay? *Ann Intern Med* 151, 68–69. <https://doi.org/10.7326/0003-4819-151-1-200907070-00013>
- Spedding, S., 2016. Open Access Publishing of Health Research: Does Open Access Publishing Facilitate the Translation of Research into Health Policy and Practice? *Publications* 4, 2. <https://doi.org/10.3390/publications4010002>
- Strom, B.L., Buyse, M.E., Hughes, J., Knoppers, B.M., 2016. Data Sharing - Is the Juice Worth the Squeeze? *N Engl J Med* 375, 1608–1609. <https://doi.org/10.1056/NEJMp1610336>
- Suchak, T., Aliu, A.E., Harrison, C., Zwiggelaar, R., Geifman, N., Spick, M., 2025. Explosion of formulaic research articles, including inappropriate study designs and false discoveries, based on the NHANES US national health database. *PLoS Biol* 23, e3003152. <https://doi.org/10.1371/journal.pbio.3003152>
- Suelzer, E.M., Deal, J., Hanus, K.L., Ruggeri, B., Sieracki, R., Witkowski, E., 2019. Assessment of Citations of the Retracted Article by Wakefield et al With Fraudulent Claims of an Association Between Vaccination and Autism. *JAMA Netw Open* 2, e1915552. <https://doi.org/10.1001/jamanetworkopen.2019.15552>
- Sun, Y., Zhang, Y., Gwizdka, J., Trace, C.B., 2019. Consumer Evaluation of the Quality of Online Health Information: Systematic Literature Review of Relevant Criteria and Indicators. *J Med Internet Res* 21, e12522. <https://doi.org/10.2196/12522>
- Taichman, D.B., Sahni, P., Pinborg, A., Peiperl, L., Laine, C., James, A., Hong, S.-T., Haileamlak, A., Gollogly, L., Godlee, F., Frizelle, F.A., Florenzano, F., Drazen, J.M., Bauchner, H., Baethge, C., Backus, J., 2017. Data Sharing Statements for Clinical Trials: A Requirement of the International Committee of Medical Journal Editors. *JAMA* 317, 2491–2492. <https://doi.org/10.1001/jama.2017.6514>
- Talley, N.J., Barbour, V., Lapeña, J.F.F., Munk, P.L., Peh, W.C.G., 2024. The rise and rise of predatory journals and the risks to clinical practice, health and careers: the APAME 2024 Sydney declaration on predatory or pseudo journals and publishers. *Med J Aust* 221, 248–250. <https://doi.org/10.5694/mja2.52410>
- Tamber, M.S., Lin, J., 2025. Illusions of Relevance: Using Content Injection Attacks to Deceive Retrievers, Rerankers, and LLM Judges. <https://doi.org/10.48550/arXiv.2501.18536>
- Target, S., 2019. The Rise and Demise of RSS. *Vice*. URL <https://www.vice.com/en/article/a3mm4z/the-rise-and-demise-of-rss> (accessed 4.12.22).
- The Doctor's Digital Path to Treatment [WWW Document], 2012. . Think with Google. URL <https://www.thinkwithgoogle.com/marketing-strategies/search/the-doctors-digital-path-to-treatment/> (accessed 4.25.21).
- Thelwall, M., Lehtisaari, M., Katsirea, I., Holmberg, K., Zheng, E.-T., 2025. Does ChatGPT Ignore Article Retractions and Other Reliability Concerns? *Learned Publishing* 38, e2018. <https://doi.org/10.1002/leap.2018>

- There is a worrying amount of fraud in medical research, 2023. . The Economist.
- Thomas, P., Spielman, S., Craswell, N., Mitra, B., 2023. Large Language Models Can Accurately Predict Searcher Preferences.
- Thorp, H.H., 2026. Resisting AI slop. *Science* 391, 5. <https://doi.org/10.1126/science.aee8267>
- Tomlinson, O.W., 2023. Analysis of predatory emails in early career academia and attempts at prevention. *Learned Publishing* 36, 156–163. <https://doi.org/10.1002/leap.1500>
- Tuason, O., Chen, L., Liu, H., Blake, J.A., Friedman, C., 2004. Biological nomenclatures: a source of lexical knowledge and ambiguity. *Pac Symp Biocomput* 238–249. https://doi.org/10.1142/9789812704856_0023
- Upadhyay, S., Pradeep, R., Thakur, N., Campos, D., Craswell, N., Soboroff, I., Dang, H.T., Lin, J., 2024. A Large-Scale Study of Relevance Assessments with Large Language Models: An Initial Look. <https://doi.org/10.48550/arXiv.2411.08275>
- Van Bulck, L., Moons, P., 2023. What if your patient switches from Dr. Google to Dr. ChatGPT? A vignette-based survey of the trustworthiness, value and danger of ChatGPT-generated responses to health questions. *Eur J Cardiovasc Nurs* zvad038. <https://doi.org/10.1093/eurjcn/zvad038>
- van der Linden, S., 2022. Misinformation: susceptibility, spread, and interventions to immunize the public. *Nat Med* 28, 460–467. <https://doi.org/10.1038/s41591-022-01713-6>
- van der Vegt, A., Zuccon, G., Koopman, B., 2021. Do better search engines really equate to better clinical decisions? If not, why not? *Journal of the Association for Information Science and Technology* 72, 141–155. <https://doi.org/10.1002/asi.24398>
- Van Noorden, R., 2023a. Medicine is plagued by untrustworthy clinical trials. How many studies are faked or flawed? *Nature* 619, 454–458. <https://doi.org/10.1038/d41586-023-02299-w>
- Van Noorden, R., 2023b. How big is science’s fake-paper problem? *Nature* 623, 466–467. <https://doi.org/10.1038/d41586-023-03464-x>
- Van Noorden, R., 2012. 366 days: 2012 in review. *Nature* 492, 324–327. <https://doi.org/10.1038/492324a>
- VanDeMark, S.H., Woloszyn, M.R., Christman, L.A., Gatusky, M.H., Lam, W.S., Tilberry, S.S., Piper, B.J., 2022. Examination of Potential Industry Conflicts of Interest and Disclosures by Contributors to Online Medical Resource Databases. *JAMA Netw Open* 5, e2220155. <https://doi.org/10.1001/jamanetworkopen.2022.20155>
- Vardi, M.Y., 2025. Will AI Destroy the World Wide Web? *Commun. ACM* 68, 5. <https://doi.org/10.1145/3749981>
- Vawdrey, D.K., Hripesak, G., 2013. Publication bias in clinical trials of electronic health records. *J Biomed Inform* 46, 139–141. <https://doi.org/10.1016/j.jbi.2012.08.007>
- Vera, M.A., El-Khoury, J.M., Thorp, H., Tofel, R.J., Ross, J.S., Mandavilli, A., Topol, E., 2022. Public Misinformation and Science Communication in Times of Public Health Crises. *Clin Chem* 68, 1008–1014. <https://doi.org/10.1093/clinchem/hvac088>
- Voorhees, E.M., Harman, D.K. (Eds.), 2005. *TREC: Experiment and Evaluation in Information Retrieval*. The MIT Press, Cambridge, Mass.
- Voorhees, E.M., Hersh, W., 2012. Overview of the TREC 2012 Medical Records Track, in: *The Twenty-First Text REtrieval Conference (TREC 2012) Proceedings*.
- Vosoughi, S., Roy, D., Aral, S., 2018. The spread of true and false news online. *Science* 359, 1146–1151. <https://doi.org/10.1126/science.aap9559>

- Wadhwa, R.R., Rasendran, C., Popovic, Z.B., Nissen, S.E., Desai, M.Y., 2021. Temporal Trends, Characteristics, and Citations of Retracted Articles in Cardiovascular Medicine. *JAMA Netw Open* 4, e2118263. <https://doi.org/10.1001/jamanetworkopen.2021.18263>
- Waldrop, M.M., 2023. How to mitigate misinformation. *Proc Natl Acad Sci U S A* 120, e2314143120. <https://doi.org/10.1073/pnas.2314143120>
- Walters, W.H., Wilder, E.I., 2023. Fabrication and errors in the bibliographic citations generated by ChatGPT. *Sci Rep* 13, 14045. <https://doi.org/10.1038/s41598-023-41032-5>
- Wanke, L.A., Hewison, N.S., 1988. Comparative usefulness of MEDLINE searches performed by a drug information pharmacist and by medical librarians. *Am J Hosp Pharm* 45, 2507–2510.
- Wei, C.-H., Allot, A., Lai, P.-T., Leaman, R., Tian, S., Luo, L., Jin, Q., Wang, Z., Chen, Q., Lu, Z., 2024. PubTator 3.0: an AI-powered literature resource for unlocking biomedical knowledge. *Nucleic Acids Res* gkae235. <https://doi.org/10.1093/nar/gkae235>
- Westbrook, J.I., Coiera, E.W., Gosling, A.S., 2005. Do online information retrieval systems help experienced clinicians answer clinical questions? *J Am Med Inform Assoc* 12, 315–321. <https://doi.org/10.1197/jamia.M1717>
- White, R.W., Shah, C. (Eds.), 2025. Information Access in the Era of Generative AI, The Information Retrieval Series. Springer Nature Switzerland, Cham. <https://doi.org/10.1007/978-3-031-73147-1>
- Wilkinson, J., Heal, C., Antoniou, G.A., Flemyng, E., Ahnström, L., Alteri, A., Avenell, A., Barker, T.H., Borg, D.N., Brown, N.J., Buhmann, R., Calvache, J.A., Carlsson, R., Carter, L.-A., Cashin, A.G., Cotterill, S., Färnqvist, K., Ferraro, M.C., Grohmann, S., Gurrin, L.C., Hayden, J.A., Hunter, K.E., Hyltse, N., Jung, L., Krishan, A., Laporte, S., Lasserson, T.J., Laursen, D.R., Lensen, S., Li, W., Li, T., Liu, J., Locher, C., Lu, Z., Lundh, A., Marsden, A., Meyerowitz-Katz, G., Mol, B.W., Munn, Z., Naudet, F., Nunan, D., O’Connell, N.E., Olsson, N., Parker, L., Patetsini, E., Redman, B., Rhodes, S., Richardson, R., Ringsten, M., Rogozińska, E., Seidler, A.L., Sheldrick, K., Stocking, K., Sydenham, E., Thomas, H., Tsokani, S., Vinatier, C., Vorland, C.J., Wang, R., Wattar, B.H.A., Weber, F., Weibel, S., Wely, M. van, Xu, C., Bero, L., Kirkham, J.J., 2024. Assessing the feasibility and impact of clinical trial trustworthiness checks via an application to Cochrane Reviews: Stage 2 of the INSPECT-SR project. <https://doi.org/10.1101/2024.11.25.24316905>
- Wolpert, A.J., 2013. For the sake of inquiry and knowledge--the inevitability of open access. *N Engl J Med* 368, 785–787. <https://doi.org/10.1056/NEJMp1211410>
- Wong, M., 2024. The Death of Search. *The Atlantic*. URL <https://www.theatlantic.com/technology/archive/2024/11/ai-search-engines-curiosity/680594/> (accessed 4.28.25).
- Wu, K., Wu, E., Wei, K., Zhang, A., Casasola, A., Nguyen, T., Riantawan, S., Shi, P., Ho, D., Zou, J., 2025. An automated framework for assessing how well LLMs cite relevant medical references. *Nat Commun* 16, 3615. <https://doi.org/10.1038/s41467-025-58551-6>
- Wu, S., Liu, S., Wang, Y., Timmons, T., Uppili, H., Bedrick, S., Hersh, W., Liu, H., 2017. Intrainstitutional EHR collections for patient-level information retrieval. *Journal of the Association for Information Science and Technology* 68, 2636–2648. <https://doi.org/10.1002/asi.23884>
- Xia, J., 2021. *Predatory Publishing*, 1st edition. ed. Routledge.
- Yandell, M.D., Majoros, W.H., 2002. Genomics and natural language processing. *Nat Rev Genet* 3, 601–610. <https://doi.org/10.1038/nrg861>

- Yang, R., Ning, Y., Keppo, E., Liu, M., Hong, C., Bitterman, D.S., Ong, J.C.L., Ting, D.S.W., Liu, N., 2025. Retrieval-augmented generation for generative artificial intelligence in health care. *npj Health Syst.* 2, 1–5. <https://doi.org/10.1038/s44401-024-00004-1>
- Yavchitz, A., Boutron, I., Bafeta, A., Marroun, I., Charles, P., Mantz, J., Ravaud, P., 2012. Misrepresentation of randomized controlled trials in press releases and news coverage: a cohort study. *PLoS Med* 9, e1001308. <https://doi.org/10.1371/journal.pmed.1001308>
- Zakka, C., Shad, R., Chaurasia, A., Dalal, A.R., Kim, J.L., Moor, M., Fong, R., Phillips, C., Alexander, K., Ashley, E., Boyd, J., Boyd, K., Hirsch, K., Langlotz, C., Lee, R., Melia, J., Nelson, J., Sallam, K., Tullis, S., Vogelsong, M.A., Cunningham, J.P., Hiesinger, W., 2024. Almanac — Retrieval-Augmented Language Models for Clinical Medicine. *NEJM AI* 1, A1oa2300068. <https://doi.org/10.1056/A1oa2300068>
- Zarin, D.A., Fain, K.M., Dobbins, H.D., Tse, T., Williams, R.J., 2019. 10-Year Update on Study Results Submitted to ClinicalTrials.gov. *N Engl J Med* 381, 1966–1974. <https://doi.org/10.1056/NEJMSr1907644>
- Zarin, D.A., Tse, T., Williams, R.J., Carr, S., 2016. Trial Reporting in ClinicalTrials.gov - The Final Rule. *N Engl J Med* 375, 1998–2004. <https://doi.org/10.1056/NEJMSr1611785>
- Zarin, D.A., Tse, T., Williams, R.J., Rajakannan, T., 2017. Update on Trial Registration 11 Years after the ICMJE Policy Was Established. *N Engl J Med* 376, 383–391. <https://doi.org/10.1056/NEJMSr1601330>
- Zerhouni, E.A., 2004. Information access. NIH public access policy. *Science* 306, 1895. <https://doi.org/10.1126/science.1106929>
- Zhu, Y., Yuan, H., Wang, S., Liu, J., Liu, W., Deng, C., Chen, H., Dou, Z., Wen, J.-R., 2024. Large Language Models for Information Retrieval: A Survey. <https://doi.org/10.48550/arXiv.2308.07107>