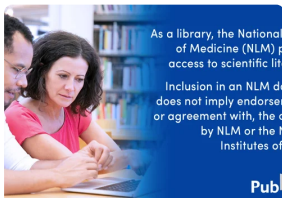


Veille Santé Num

28 Items

Veille en Santé numérique et documentation médicale - DéSaN CHU Rouen

https://www.zotero.org/groups/5886551/veille_scientifique_dsan/library



A scoping review: how evaluation methods shape our understanding of ChatGPT's effectiveness in healthcare - PubMed

[nih.gov](#)

Int J Med Inform / To address methodological heterogeneity and the lack of standardization, this study recommends multi-method cross-validation strategies and a risk-stratified, standardized evaluation framework. These



APMnews - OpenAI lance ChatGPT Santé, un dossier médical boosté à l'IA

[apmnews.com](#)

SAN FRANCISCO, 8 janvier 2026 (APMnews) - OpenAI a annoncé jeudi sur son site le lancement de ChatGPT Santé, un "espace distinct" au sein de ChatGPT permettant aux patients de réunir l'ensemble de leurs



Public Sector


Accelerate Medical Discovery with PubMed in BigQuery | Google Cloud Blog

Google

Accelerate medical research by accessing PubMed data in BigQuery. Learn how to use Vertex AI to perform semantic search for drug discovery and oncology.



ZebraMap: A Multimodal Rare Disease Knowledge Map with Automated Data Aggregation & LLM-Enriched Information Extraction Pipeline - PubMed

 nih.gov

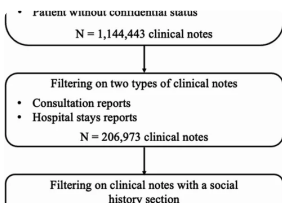
Diagnostics / Background: Rare diseases often lead to delayed diagnosis because clinical knowledge is fragmented across unstructured research, individual case reports, and heterogeneous data formats. This study



LLM-augmented semantic embeddings enable Cross-Lingual mapping of medical procedure terms - Scientific Reports

 nature.com

Scientific reports / Cross-lingual information retrieval limits global exchange of data because of the high diversity in the methods to classify, document and encode medical proced...



Improving social determinants of health documentation in French electronic health records using large language models - PMC #EDS

 nih.gov

Scientific reports / Social determinants of health (SDoH) significantly influence health outcomes, shaping disease progression, treatment adherence, and health disparities. However, their documentation in



A Unified Multi-Domain Framework for Hallucination Detection and Reliability Evaluation in Large Language Models #preprint

 researchsquare.com

Large Language Models which consist of models like GPT5 from OpenAI, Claude Sonnet from Anthropic, Google's Gemini are known for their reasoning capability, ability to summarize, cybersecurity analysis and

Implications for:			
s	Health administrators	Policy makers and regulators	Research
Implications for:	<ul style="list-style-type: none"> Clinical practice Performance evaluation Policy making Biomedical research 	<ul style="list-style-type: none"> Optimized hospital operations. Optimized staffing. Optimized resources. 	<ul style="list-style-type: none"> Developing evidence-based health policies and guidelines. Identifying health trends and outbreaks.

Building a healthcare data warehouse: considerations, opportunities, and challenges - PMC #EDS

[nih.gov](#)

Front Digit Health / The increasing digitalization of health systems is accelerating the transition towards a new era of data-driven, evidence-based care. This has profound implications for clinical practice,



Vulnerability of Large Language Models to Prompt Injection When Providing Medical Advice

[doi.org](#)

JAMA Network Open / Question Can commercial medical large language models (LLMs) be manipulated through prompt-injection attacks (ie, maliciously crafted inputs that manipulate an LLM's behavior) to

Editors
Soren Auer
Dirk Betz
Claudia Binossek
Anna Jaczyn
Azam Jomekong
Hannah Kim
Hande Kuzuk McGinty
Keith O. Mills
Allard Oelen
Golan Rabby
Enayat Rajabi



Hallucinations in Scholarly LLMs: A Conceptual Overview and Practical Implications

[tib-op.org](#)

Open Conference Proceedings / The issue of large language models (LLMs) is gradually infiltrating the academic workflow, but it also presents one significant problem: hallucination. The hallucinations involve invented



Scientific production in the era of large language models

[science.org](#)

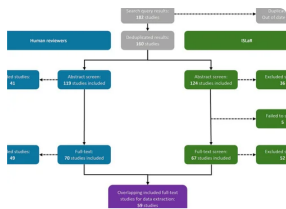
Science / Despite growing excitement (and concern) about the fast adoption of generative artificial intelligence (Gen AI) across all academic disciplines, empirical evidence remains fragmented, and systematic



Two-stage prompting framework with predefined verification steps for evaluating diagnostic reasoning tasks on two datasets - npj Digital Medicine

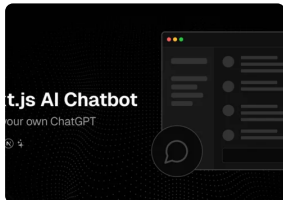
[nature.com](#)

npj Digital Medicine / Despite their growing use in medicine, large language models (LLMs) demonstrate limited diagnostic reasoning. We evaluated a two-stage prompting framework with predefined verification



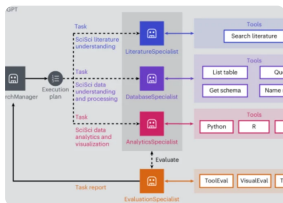
Analysis of article screening and data extraction performance by an AI systematic literature review platform - PMC #SR

Front Artif Intell / Systematic literature reviews (SLRs) are critical to health research and decision-making but are often time- and labor-intensive. Artificial intelligence (AI) tools like large language models (LLMs)



SciSciGPT - SciSciGPT

An AI-powered data analyst for Science of Science.



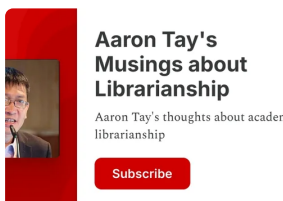
SciSciGPT: advancing human-AI collaboration in the science of science - Nature Computational Science

nature computational science / SciSciGPT is an open-source prototype AI collaborator that explores the use of LLM research tools to automate workflows, support diverse analytical approaches and enhance



Dramatic increases in redundant publications in the Generative AI era - BMC Medicine

BMC Medicine / Background Redundant publication, the practice of submitting the same or substantially overlapping manuscripts multiple times, distorts the scientific record a...



“We’re Good at Search”... Just Not the Kind That the AI era Demands - a Provocation

Aaron Tay's Musings about Librarianship / I might be exaggerating slightly, but if you look at the few new evaluation matrices for AI-powered search circulating, "relevancy" is often just one of several



Automating expert-level medical reasoning evaluation of large language models - npj Digital Medicine

 nature.com

npj Digital Medicine / As large language models (LLMs) become increasingly integrated into clinical decision-making, ensuring trustworthy reasoning is paramount. However, current evaluation strategies of LLMs'



Design and Implementation of a Scalable Clinical Data Warehouse for Resource-Constrained Healthcare Systems #EDS

 arxiv.org

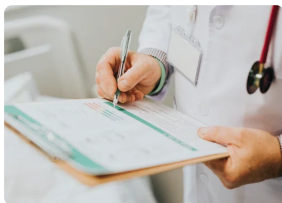
Centralized electronic health record repositories are critical for advancing disease surveillance, public health research, and evidence-based policymaking. How...



Reporting guideline for the use of Generative Artificial intelligence tools in MEDical Research: the GAMER Statement

bmj.com

BMJ Evidence-Based Medicine / Objectives: Generative artificial intelligence (GAI) tools can enhance the quality and efficiency of medical research, but their improper use may result in plagiarism, academic fraud



Knowledge-Practice Performance Gap in Clinical Large Language Models: Systematic Review of 39 Benchmarks

 jmir.org

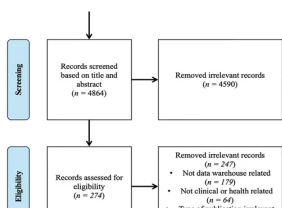
JMIR / Background: The evaluation of large language models (LLMs) in medicine has undergone a shift from knowledge-based testing to practice-based assessment, represe...



Prediction of prolonged length of stay from first 2 days of hospitalization data: a SHAP value-based variable selection method for a simplified model - PubMed #EDS

➤ nih.gov

Int J Qual Health Care / While it is possible to predict prolonged LOS with a large complex set of variables, most models are difficult to use in clinical practice. SHAP-value-based variable selection allowed a reduction



The development and use of data warehousing in clinical settings: a scoping review - PMC #EDS

➤ nih.gov

Front Digit Health / The emergence of data warehousing in clinical settings has greatly enhanced data analysis capabilities, facilitating the accurate and comprehensive extraction of valuable information. This



MedSumGraph: enhancing GraphRAG for medical QA with summarization and optimized prompts - PubMed #RAG

➤ nih.gov

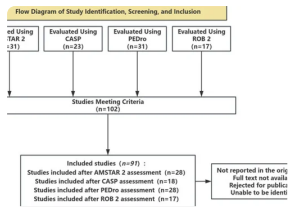
Artif Intell Med / The rapid development of large language models (LLMs) has accelerated research into applying artificial intelligence (AI) to domains such as medical question answering and clinical decision support.



Le NEJM AI a ouvert la voie : reviewers, relecteurs, un métier en voie de disparition (rapide ?) | Rédaction Médicale et Scientifique

redactionmedicale.fr

Blog Rédaction Médicale et Scientifique / Il l'avait annoncé début septembre 2025 à Chicago : le rédacteur en chef du NEJM AI (ISK ou ZAK) avait expliqué que les décisions sur un manuscrit soumis à son journal



Can AI assess literature like experts? An entropy-based comparison of ChatGPT-4o, DeepSeek R1, and human ratings - PMC #SR

nih.gov

Frontiers in research metrics and analytics / Manual quality assessment of systematic reviews is labor-intensive, time-consuming, and subject to reviewer bias. With recent advances in large language models (LLMs), it is



Veille Santé Num (Novembre 2025)

Wakelet Collection

Veille en Santé numérique et documentation médicale - DéSaN CHU Rouen

https://www.zotero.org/groups/5886551/veille_scientifique_dsan/library



Private or Deleted Collection

Wakelet Collection

Veille en Santé numérique et documentation médicale - DéSaN CHU Rouen

https://www.zotero.org/groups/5886551/veille_scientifique_dsan/library