AcademyHealth

AI for HSR: Glossary of Key Terms

Unless otherwise indicated, the following terms are sourced and adapted from the National Institute of Standards and Technology (NIST)'s resource on "The Language of Trustworthy AI: An In-Depth Glossary of Terms."

Section 1: Terms & Concepts Related to Artificial Intelligence (AI)

Algorithm: a set of instructions designed to accomplish a task: taking one or more inputs, running them systematically through a series of steps, and providing one or more outputs. Typically associated with computing, algorithms are an essential element of computer programming. They can be used to accomplish a variety of computational tasks, such as performing calculations or finding information in databases.[1]

Artificial intelligence (AI): generally refers to the capacity of computers or other machines to perform tasks that typically require human intelligence such as reasoning, problem-solving, and decision-making. AI systems use algorithms and computational techniques to process large volumes of data, extract patterns, and make predictions or decisions based on those patterns learning.[2]

Bias (computational): an effect which deprives a statistical result of representativeness by systematically distorting it, as distinct from a random error which may distort on any one occasion but balances out on the average.[3]

Bias (legal or ethical): discrimination against or in favor of particular individuals or groups. In the context of ethics and politics, the question arises whether a particular bias is unjust or unfair[4]

Big data: a massive amount of information on a given topic. This includes information generated, stored, and analyzed on a scale too vast to manage with traditional information storage systems. In health care, big data sources can include patient medical records, hospital records, medical exam results, and information collected by healthcare testing machines.[5]

Data mining: techniques for analyzing large amounts of information to gain insights, spot trends, or uncover substantive patterns. Data mining often involves use of algorithms, especially machine learning, more than traditional statistical methods.[6]

Deep learning: a subset of machine learning that relies on neural networks with many layers of neurons. In so doing, deep learning employs statistics to spot underlying trends or data patterns and applies that knowledge to other layers of analysis. Some have labeled this as a way to "learn by example" and a technique that "perform[s] classification tasks directly from images, text, or sound" and then applies that knowledge independently. Deep learning requires extensive computing power and labeled data.[7]

Generative AI (generative AI or gen AI): a specific subset of AI focused on creating content such as text, images, video, music, and other outputs in response to user input (or prompts). Generative AI models are designed to learn the patterns and structure of their input training data, and to generate new data with similar characteristics. Generative AI tools can quickly and easily generate a wide variety of human-like outputs, so they have the potential to radically transform approaches to content creation across a wide range of domains and industries. However, their use is also subject to numerous important limitations.[8]

Hallucination: generated content that is nonsensical or unfaithful to the provided source content. There are two main types of hallucinations.[9]

- An ***intrinsic hallucination*** is a generated output that contradicts the source content.
- An ***extrinsic hallucination*** is a generated output that cannot be verified from the source content (i.e., output can neither be supported nor contradicted by the source).

Large language models (LLMs): a class of language models using deep learning algorithms and trained on extremely large textual datasets that can be multiple terabytes in size. LLMs can be classed into two types.[10]

- **Discriminatory LLMs** are mainly used for supervised machine learning tasks (e.g., regression). They are designed to classify information by learning the decision boundaries separating different classes, labels, or categories of information within a dataset.[11]
- **Generative LLMs** can be used for semi-supervised tasks, including transforming and translating unstructured data. They have the ability to generate seemingly new realistic data, including by analyzing and integrating vast amounts of unstructured content from different data formats.[12]

Machine learning (ML): the study or the application of computer algorithms that improve automatically through experience. Machine learning algorithms build a model based on training data in order to perform a specific task, like aiding in prediction or decision-making processes, without necessarily being explicitly programmed to do so.[13]

Natural language processing (NLP): a computer's attempt to understand, process, and manipulate human language in spoken or written form. This is often done through machine learning processes that parse vocabulary, grammar, and intent while also allowing for variation in language use.[14]

Predictive analytics: a branch of advanced analytics that makes predictions about future outcomes using historical data combined with statistical modeling, data mining techniques and machine learning. Predictive analytics can be employed to find patterns in data,
to identify risks and opportunities.[15]

Trustworthy AI: generally refers to AI systems that are valid and reliable; safe, secure and resilient; accountable and transparent; explainable and interpretable; privacy-enhanced; and fair with harmful bias managed.[16]

Unstructured data: data that are not readily available in predefined structured (e.g., tabular) formats. These data may be found in formats that lack standardization and need significant preprocessing and feature extraction efforts. Examples may include images or free form text, extracted from social media or medical records.[17]

Section 2: Terms & Concepts Related to Health Services Research (HSR)

Clinical decision support (CDS): provides clinicians, staff, patients or other individuals with knowledge and person-specific information, intelligently filtered or presented at appropriate times, to enhance health and health care. CDS encompasses a variety of tools to enhance decision-making in the clinical workflow. These tools include computerized alerts and reminders to care providers and patients; clinical guidelines; condition-specific order sets; focused patient data reports and summaries; documentation templates; diagnostic support, and contextually relevant reference information, among other tools.[18]

Comparative effectiveness research (CER): examines effectiveness by comparing one or more treatments, procedures, or medications to determine what works best for which patients under real world conditions. This field of study utilizes data generation (through conduct of new studies) and synthesis (through comparisons of existing studies) to build evidence on best practices and policies for improving health care. CER studies are generally conducted after randomized controlled trials (RCTs) have been conducted to determine treatment efficacy (i.e., how well a treatment works under ideal conditions).[19]

Governance: actions to ensure stakeholder needs, conditions, and options are evaluated to determine balanced, agreed-upon enterprise objectives; setting direction through prioritization and decision-making; and monitoring performance and compliance against agreed-upon directions and objectives. AI governance may include policies on the nature of AI applications developed and deployed versus those limited or withheld. This can also refer to the framework of policies, rules, and processes for ensuring direction, management and accountability in undertaking the above-listed actions.[20,21]

Health services research (HSR): a multidisciplinary field of inquiry (basic and applied) that examines access to, and the use, costs, quality, delivery, organization, financing, and outcomes of health care services to produce new knowledge about the structure, processes, and effects of health services for individuals and populations.[22] It generally encompasses the study of what works, for whom, at what cost, and under what circumstances. This encompasses research on how our health system works, how to support patients and providers in choosing the right care, and how to improve health through care delivery.[23]

Interoperability: the ability of different information systems, devices and applications (systems) to access, exchange, integrate and cooperatively use data in a coordinated manner. This includes the exchange of data within and across organizational, regional and national boundaries, to provide timely and seamless portability of information and as needed to optimize the health of individuals and populations.[24]

Mixed methods: a form of research enabling investigators to conceptually and analytically integrate qualitative research and qualitative data (e.g., semi-structured interviews, observations, focus groups) with traditional epidemiological and quantitative methods of research.[25]

Endnotes

1. *Algorithm*. Network of the National Library of Medicine.(n.d).
2. *Research guides: Artificial intelligence (AI): Introduction*. Introduction - Artificial Intelligence (AI) - Research Guides at RutgersUniversity. (n.d.). https://libguides.rutgers.edu/artificial-intelligence/introduction
3. Organization for Economic Co-Operation and Development. (2019). *Artificial Intelligence in Society*. OECD Publishing.
4. Coeckelbergh, M. (2020). AI Ethics. United Kingdom: MIT Press.
5. Big data in health care and patient outcomes. School of Public Health. (2022, September 15).
6. Darrell M. West, J. R. A., Wheeler, T., Elaine Kamarck, D. M. W., & Aaron Klein, S. L. (2022b, March 9). *The Brookings Glossary of AI and Emerging Technologies*. Brookings. https://www.brookings.edu/articles/the-brookings-glossary-of-ai-and-emerging
7. *Ibid*
8. Research guides: Artificial intelligence (AI): Introduction. Introduction - Artificial Intelligence (AI) - Research Guides at Rutgers University. (n.d.). https://libguides.rutgers.edu/artificial-intelligence/introduction
9. Ziwei Ji, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Ye Jin Bang, Andrea Madotto, and Pascale Fung. 2023. Survey of Hallucination in Natural Language Generation. ACM Comput.Surv. 55, 12, Article 248 (December 2023), 38 pages. https://doi.org/10.1145/3571730
10. AI Assurance: Towards Trustworthy, Explainable, Safe, and Ethical AI. (2022). Netherlands: Elsevier Science.
11. Tortora L. Beyond Discrimination: Generative AI Applications and Ethical Challenges in Forensic Psychiatry. Front Psychiatry. 2024 Mar 8;15:1346059. doi: 10.3389/fpsyt.2024.1346059. PMID: 38525252; PMCID: PMC10958425.
12. *Ibid*
13. *National Security Commission on Artificial Intelligence*. Nscai.gov. (n.d.). https://www.nscai.gov/2021-final-report/
14. Natural language processing. Network of the National Library of Medicine. (n.d.). https://www.nnlm.gov/guides/data-glossary/natural-language-processing
15. IBM. (n.d.). *What is predictive analytics?* https://www.ibm.com/topics/predictive-analytics
16. Tabassi, E. (2023), Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST Trustworthy and Responsible AI, National Institute of Standards and Technology, Gaithersburg, MD, [online], https://doi.org/10.6028/NIST.AI.100-1, https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=936225 (Accessed February 22, 2024)
17. Sedlakova J, Daniore P, Horn Wintsch A, Wolf M, Stanikic M, Haag C, Sieber C, Schneider G, Staub K, Alois Ettlin D, Grübner O, Rinaldi F, von Wyl V; University of Zurich Digital Society Initiative (UZH-DSI) Health Community. Challenges and best practices for digital unstructured data enrichment in health research: A systematic narrative review. PLOS Digit Health. 2023 Oct 11;2(10):e0000347. doi: 10.1371/journal.pdig.0000347. PMID: 37819910; PMCID: PMC10566734.
18. Clinical Decision Support. Healthit.gov. (n.d.). https://www.healthit.gov/topic/safety/clinical-decision-support
19. *Comparative Effectiveness Research*. The Ohio State University Center for Clinical and Translational Science. (n.d.). https://ccts.osu.edu/content/comparative-effectiveness-research
20. *National Security Commission on Artificial Intelligence*. Nscai.gov. (n.d.). https://www.nscai.gov/2021-final-report/
21. Huang, H.-M. (Ed.). (2008, October). Autonomy Levels for Unmanned Systems (ALFUS) Framework Volume I: Terminology Version 2.0 . National Institute of Standards and Technology. https://www.nist.gov/system/files/documents/el/isd/ks/NISTSP1011-I2-0.pdf
22. Institute of Medicine (US) Committee on Health Services Research: Training and Work Force Issues; Thaul S, Lohr KN, Tranquada RE, editors. Health Services Research: Opportunities for an Expanding Field of Inquiry: An Interim Statement. Washington (DC): National Academies Press (US); 1994. A WORKING DEFINITION OF HEALTH SERVICES RESEARCH. https://www.ncbi.nlm.nih.gov/books/NBK231502/
23. *What is health services research?*. AcademyHealth. (n.d.). https://academyhealth.org/sites/default/files/what_is_health_services_research_one-pager_0.pdf
24. HIMSS. (2021, August 25). Interoperability in Healthcare. HIMSS. https://www.himss.org/resources/interoperability-healthcare
25. *Why mixed methods?* Johns Hopkins Bloomberg School of Public Health. (n.d.). https://publichealth.jhu.edu/academics/academicprogram-finder/training-grants/mixed-methods-research-training-program-for-the-health-sciences/