Artificial Intelligence in Biomedicine and Health

What is Biomedical and Health Informatics? - http://informatics.health/
William Hersh

Oregon Health & Science University

1

Definitions

- Artificial intelligence
- Machine learning
- Data science
  - Data analytics
- Big Data
- Data mining
  - Text mining
- Other terms

2

Artificial intelligence (AI)

- AI – information systems and algorithms capable of performing tasks associated with human intelligence (Rajpurkar, 2022; Sahni, 2023)
  - Origin attributed to summer workshop held at Dartmouth College in 1956 (McCarthy, 2007)
- Some classify AI into two broad categories (Khare, 2023)
  - Predictive AI – use of data and algorithms to predict some output, e.g., diagnosis, treatment recommendation, prognosis, etc.
  - Generative AI – generates new output based on prompts, e.g., text, images, etc.
- Initial focus starting in 1950s on engineering knowledge representations and algorithms (Shortliffe, 1987; Lea, 2023), which have now given way to approaches involving machine learning (ML) (Shortliffe, 2019)
  - Original methods have transformed to uses such as clinical decision supportin the electronic health record (Greenes, 2018; Greenes, 2023)

Machine learning (ML)

- Term originally attributed to Arthur Samuel in 1959: “Field of study that gives computers the ability to learn without being explicitly programmed” (McCarthy, 1990)
- Derived from
  - Mathematics/statistics – learning relationships from data
  - Computer science – emphasis on efficient algorithms, especially those involving large amounts of data
- Much recent success from deep learning (DL) (Shah, 2022)
  - ML associated with use of neural networks that have deep layers requiring substantial processing (Sevilla, 2022)
  - Initial success in predictive AI, in areas such as image classification, including in medicine (Esteva, 2021)
  - Also has facilitated large language models (LLMs) that drive generative AI (Raschka, 2023)

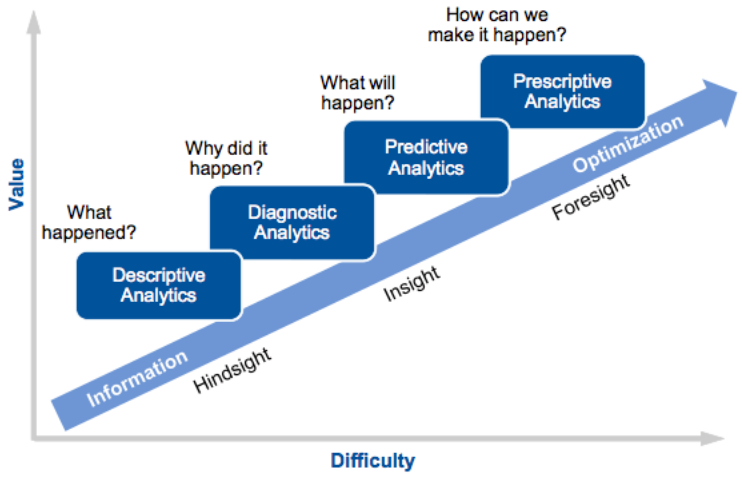
Data science

- "The science of learning from data; it studies the methods involved in the analysis and processing of data and proposes technology to improve methods in an evidence-based manner" (Donoho, 2017)
- "Data science encompasses a set of principles, problem definitions, algorithms, and processes for extracting nonobvious and useful patterns from large data sets" (Kelleher, 2018)
- Applications in biomedicine (Hoyt, 2019; Topol, 2019; Patrishkoff, 2023)
- Technical but non-mathematical overview (Spector, 2022)

Data analytics

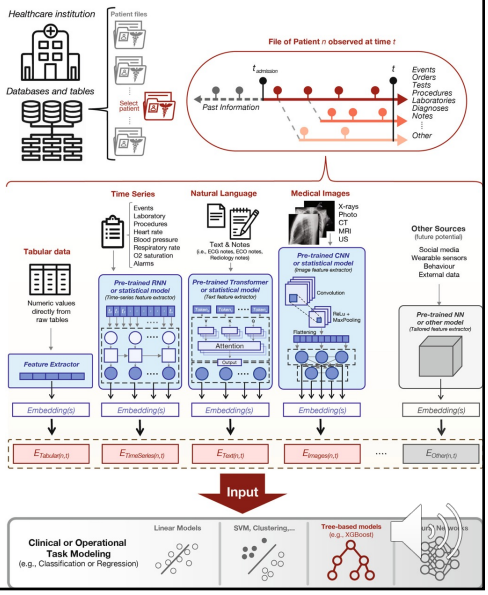
- "The extensive use of data, statistical and quantitative analysis, explanatory and predictive models, and fact-based management to drive decisions and actions" (Davenport, 2017)
- "The systematic use of data and related business insights developed through applied analytical disciplines (e.g. statistical, contextual, quantitative, predictive, cognitive, other [including emerging] models) to drive fact-based decision-making for planning, management, measurement and learning" (IBM, 2012)

Levels of data analytics from Gartner (Schaap, 2020)



Big Data

- Various 4 Vs (e.g., Zikopolous, 2011); most recent from NIST (Chang, 2019)
  - Volume
  - Velocity
  - Variety
  - Variability
- Multimodal data (Acosta, 2022; Soenksen, 2022; Topol, 2023)
  - Data from many modalities, from cells and molecules; tissues and organs; and organisms, populations, and the global environment



Data mining

- Processing and modeling of data to discover previously unknown patterns or relationships (Bellazzi, 2008; Zaki, 2020)
- Text mining – applying data mining to unstructured textual data (Aggarwal, 2012; Cohen, 2013)
- Terms have faded from vernacular but methods still used

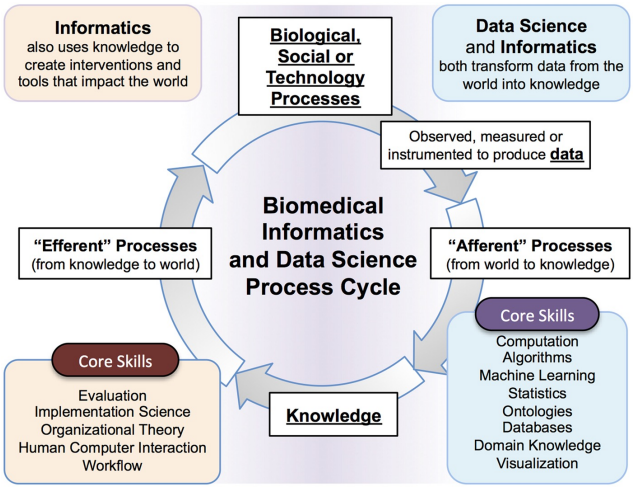
Some other terms

- Data provenance – origin and trustworthiness of data (Doan, 2012)
- Business intelligence – use of data to obtain timely, valuable insights into business and clinical data (Brijs, 2012)
- Precision medicine (IOM, 2011; Collins, 2015; Denny, 2019); formerly personalized (Hamburg, 2010) or computational medicine (Winslow, 2012)
- "Re-use" or "secondary use" of clinical data, especially from the electronic health record (EHR) (Safran, 2007; Meystre, 2017) and other sources (Näher, 2023)

Other terms (cont.)

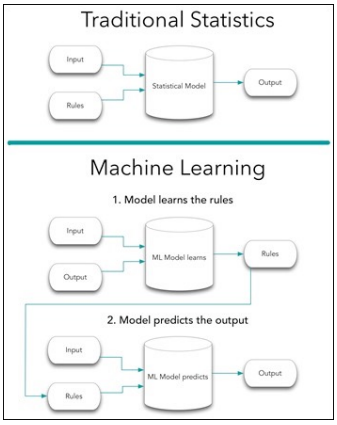
- Metadata – data about data (Riley, 2017)
- Model – representation of data, structure, and relationships for ML (and other tasks) (Shin, 2020)
- Data visualization – applying visual methods to "tell the story" about the data (Wilke, 2019)
- Data wrangling – processing data into format suitable for analytics, learning, visualization, etc.

Where does informatics fit into data science/AI/ML (Payne, 2018)?



How does data science/ML differ from statistics?

- Statistics draws population inferences from samples; ML finds generalizable predictive patterns (Bzdok, 2018)
  - Diagrammed by Scarlat (2019)
  - A false dichotomy? (Finlayson, 2023)
- Data scientist is a "person who is better at statistics than any software engineer and better at software engineering than any statistician" (Donoho, 2017)



Additional resources for AI

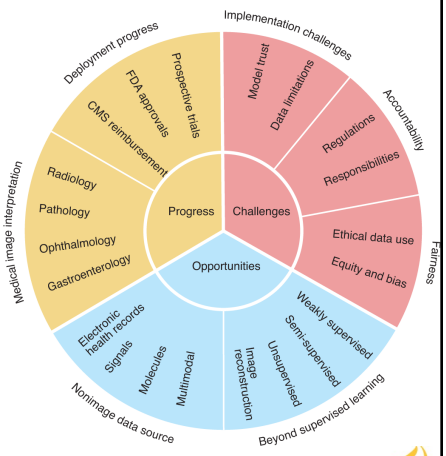
- Books
- Articles
- Other

AI books

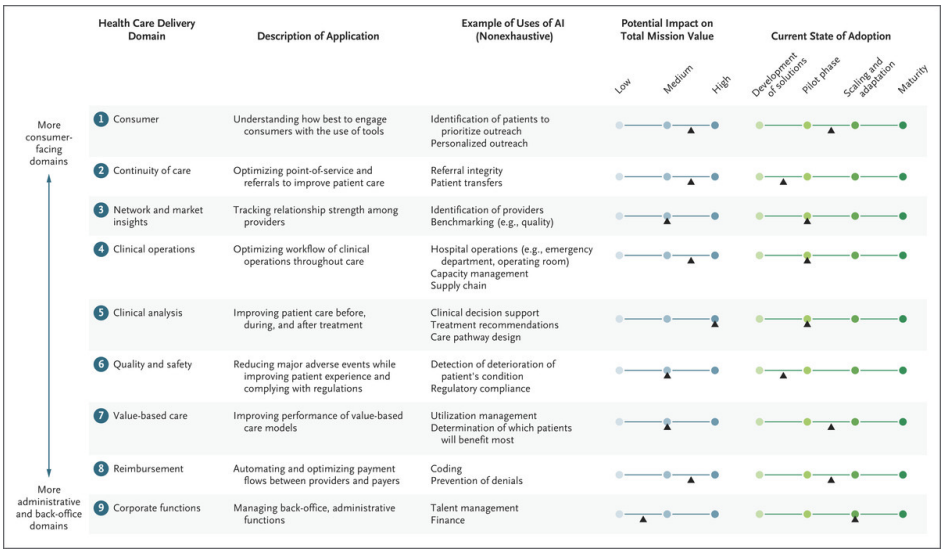
- National Academy of Medicine (Matheny, 2019)
- Deep medicine (Topol, 2019)
- Intelligent systems in medicine and health (Cohen, 2022)
  - Including chapter on history of AI in medicine (Shortliffe, 2022)
- Digitizing diagnosis – early history of AI (Lea, 2023)
- How data happened (Wiggins, 2023)
- AI in healthcare (Davenport, 2022)
- AI revolution in medicine (Lee, 2023)
- Clinical applications of AI in real-world data (Asselbergs, 2023)
- Translational applications in healthcare (Reddy, 2023)
- Co-intelligence – living and working with AI (Mollick, 2024)

AI articles

- Progress, challenges, and opportunities in clinical medicine (Rajpurkar, 2022)
- Scientific discovery in age of AI (Wang, 2023)
- Status of AI in healthcare (Sahni, 2023)
- Growing use of ambient AI (Nahar, 2023; Topol, 2022)



Status of AI in healthcare (Sahni, 2023)



	Health Care Delivery Domain	Description of Application	Example of Uses of AI (Nonexhaustive)	Potential Impact on Total Mission Value	Current State of Adoption
				Low · Medium · High	Development of solutions · Pilot phase · Scaling and adaptation · Maturity
More consumer-facing domains	1 Consumer	Understanding how best to engage consumers with the use of tools	Identification of patients to prioritize outreach Personalized outreach		
	2 Continuity of care	Optimizing point-of-service and referrals to improve patient care	Referral integrity Patient transfers		
	3 Network and market insights	Tracking relationship strength among providers	Identification of providers Benchmarking (e.g., quality)		
	4 Clinical operations	Optimizing workflow of clinical operations throughout care	Hospital operations (e.g., emergency department, operating room) Capacity management Supply chain		
	5 Clinical analysis	Improving patient care before, during, and after treatment	Clinical decision support Treatment recommendations Care pathway design		
	6 Quality and safety	Reducing major adverse events while improving patient experience and complying with regulations	Detection of deterioration of patient's condition Regulatory compliance		
	7 Value-based care	Improving performance of value-based care models	Utilization management Determination of which patients will benefit most		
	8 Reimbursement	Automating and optimizing payment flows between providers and payers	Coding Prevention of denials		
More administrative and back-office domains	9 Corporate functions	Managing back-office, administrative functions	Talent management Finance		

Other AI resources

- AMA Introduction to AI in Health Care
  - https://edhub.ama-assn.org/change-med-ed/interactive/18827029
- Substacks
  - Ground Truths, Eric Topol – https://erictopol.substack.com/
  - One Useful Thing, Ethan Mollick – https://www.oneusefulthing.org/
  - Ahead of AI, Sebastian Raschka – https://magazine.sebastianraschka.com/