

# Collaborative for Human Factors

Introduction to AI: ChatGPT & applicable research

Erkin Ötleş  
April 2023





# Hi, I'm Erkin Ötles!

Medical Scientist Training Program Fellow

MD: x2024

Engineering PhD: 2022

ML Dev & Implementation Lead

Previously:

Healthcare Data & Decision Science Manager

Epic Ambulatory Solutions Engineer

COI:

Patent pending: AI prediction of health outcomes in patients with occupational injuries.

Small amount of IRA stock in various technology & healthcare companies.

Provide AI advising for several startups.



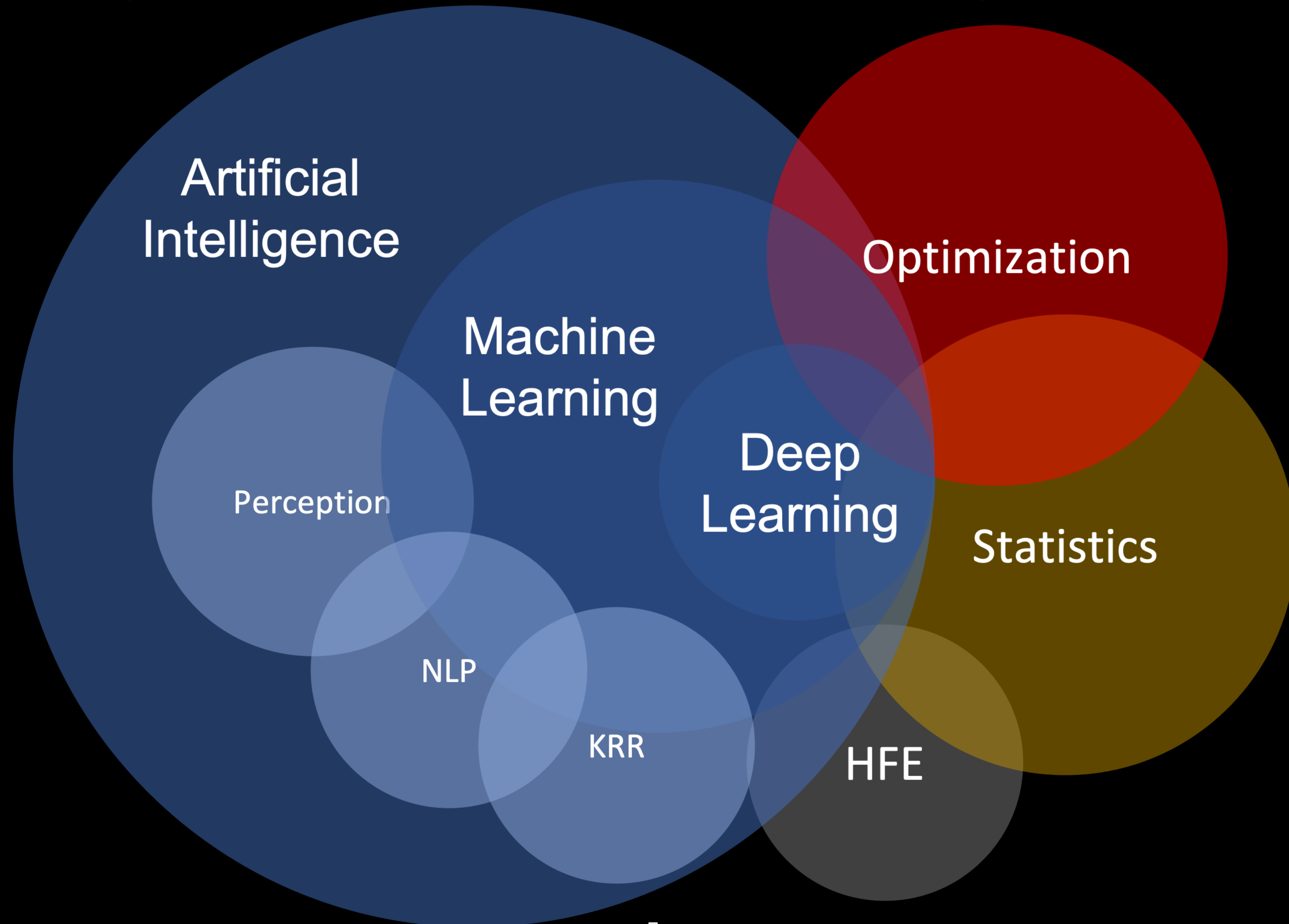
# Introduction to AI

# First, some definitions

**Artificial Intelligence (AI):** *intelligence* (perceiving, synthesizing, and inferring information) demonstrated by machines

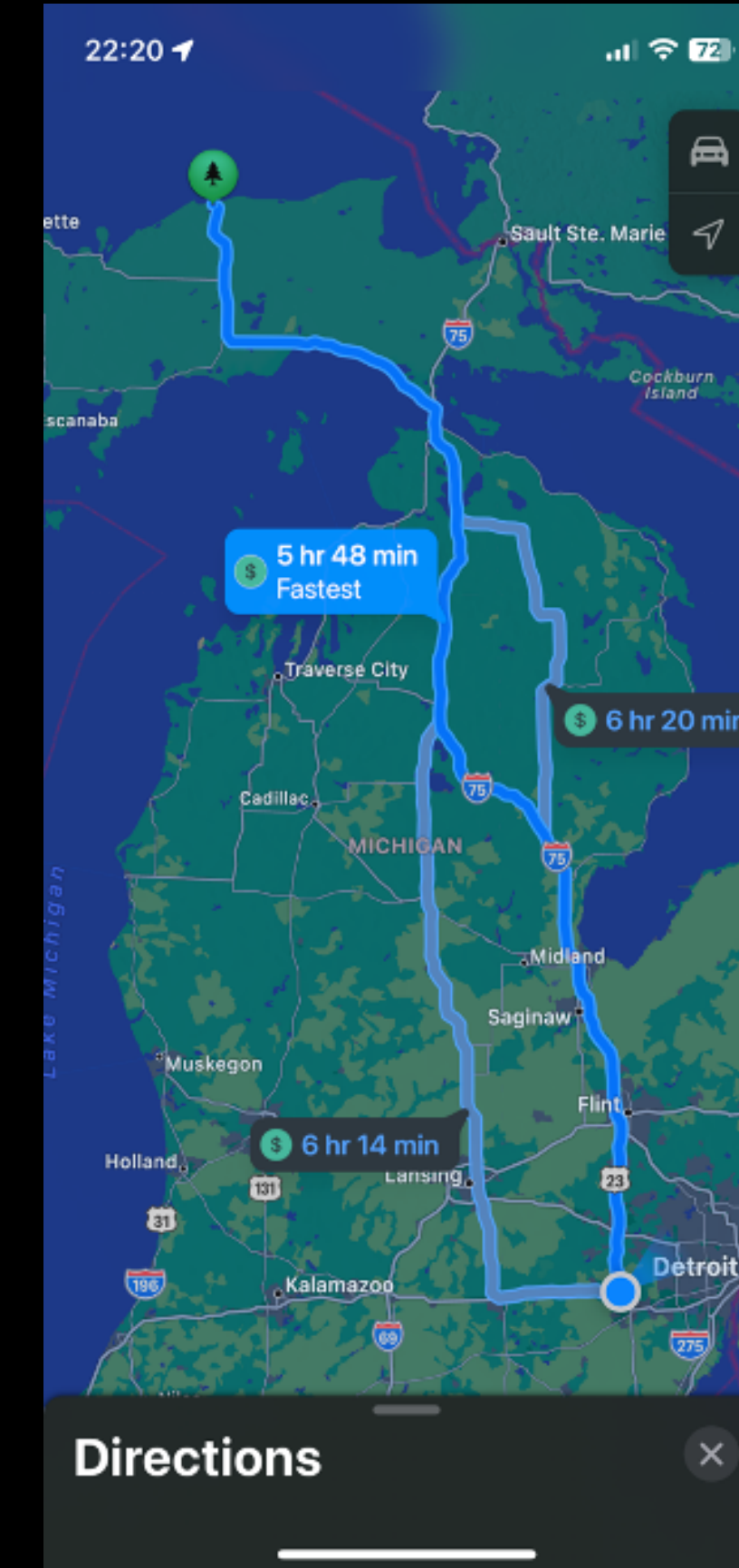
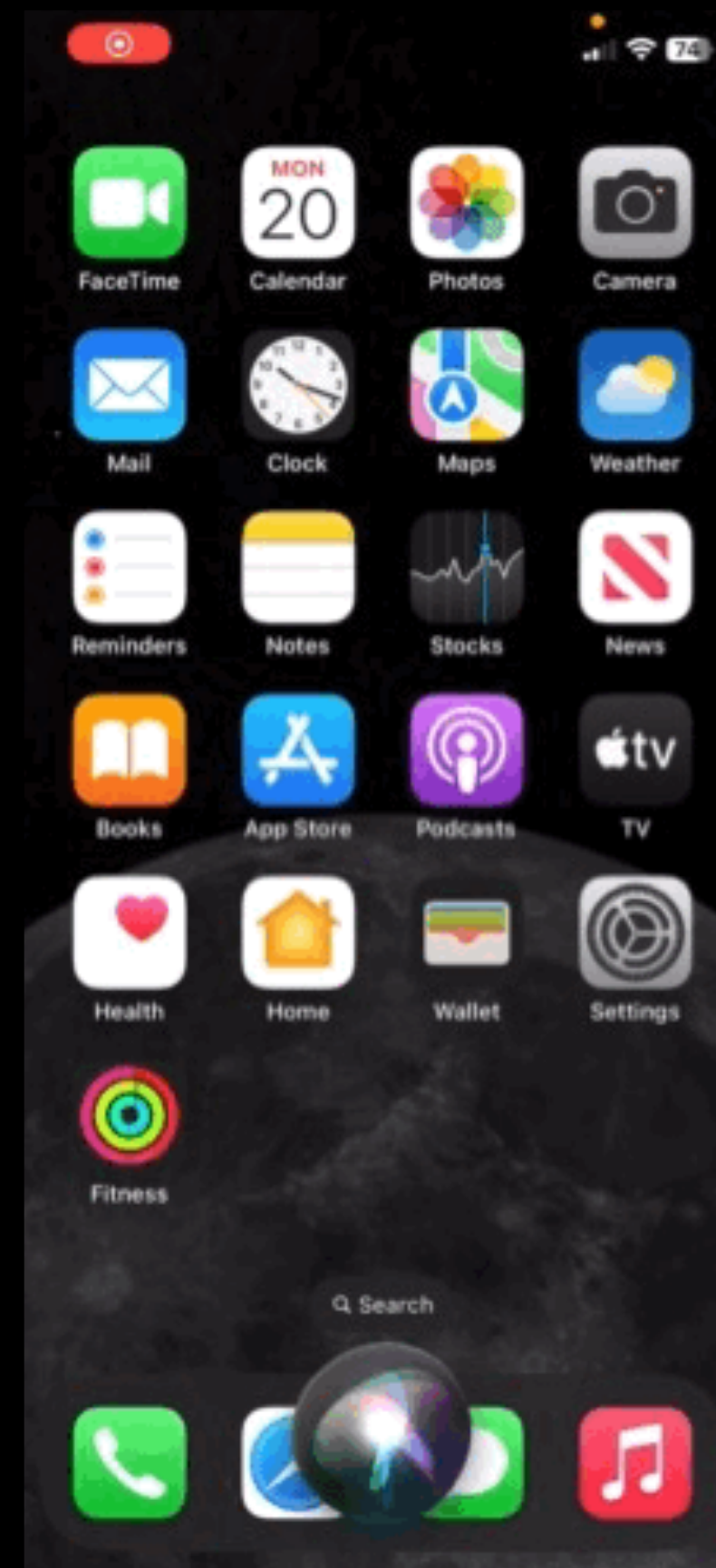
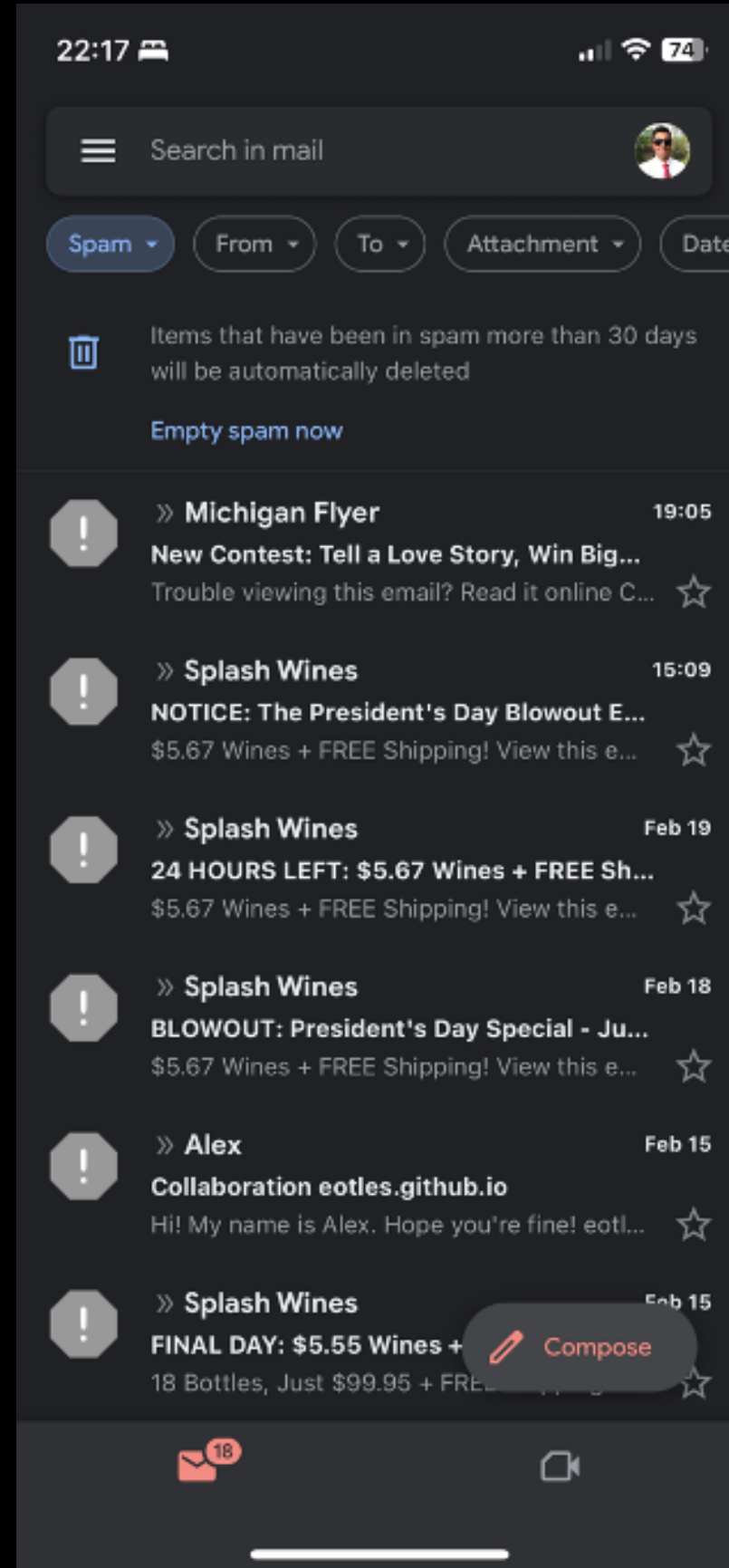
**Machine Learning (ML):** field of inquiry devoted to understanding and building methods that *learn* (use data to improve performance on a task).

# Nesting and overlapping concepts





# AI is ubiquitous in everyday life



# Many industries depend on AI

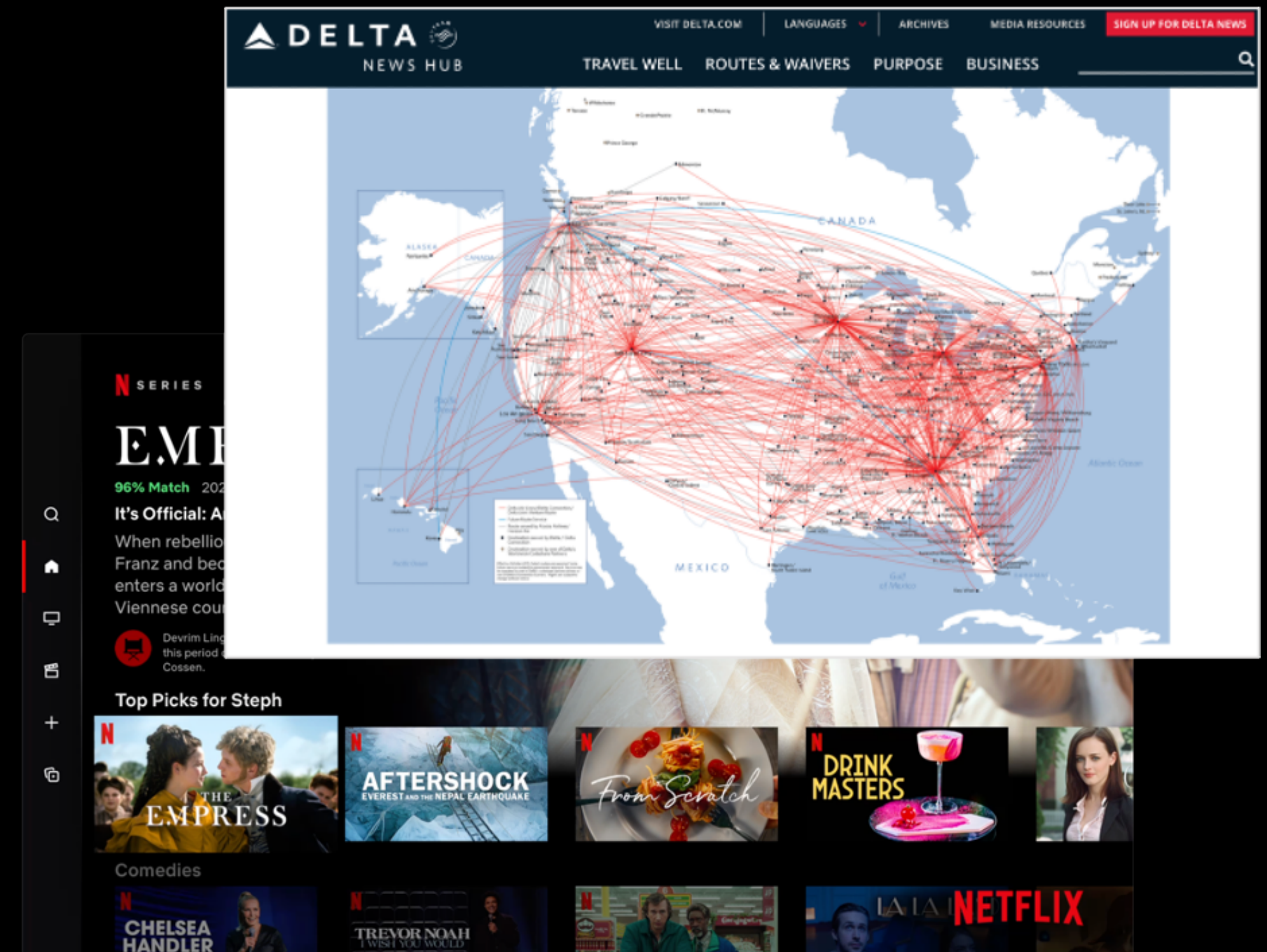
What routes should we fly?

When should we service our planes?

How should we price a product?

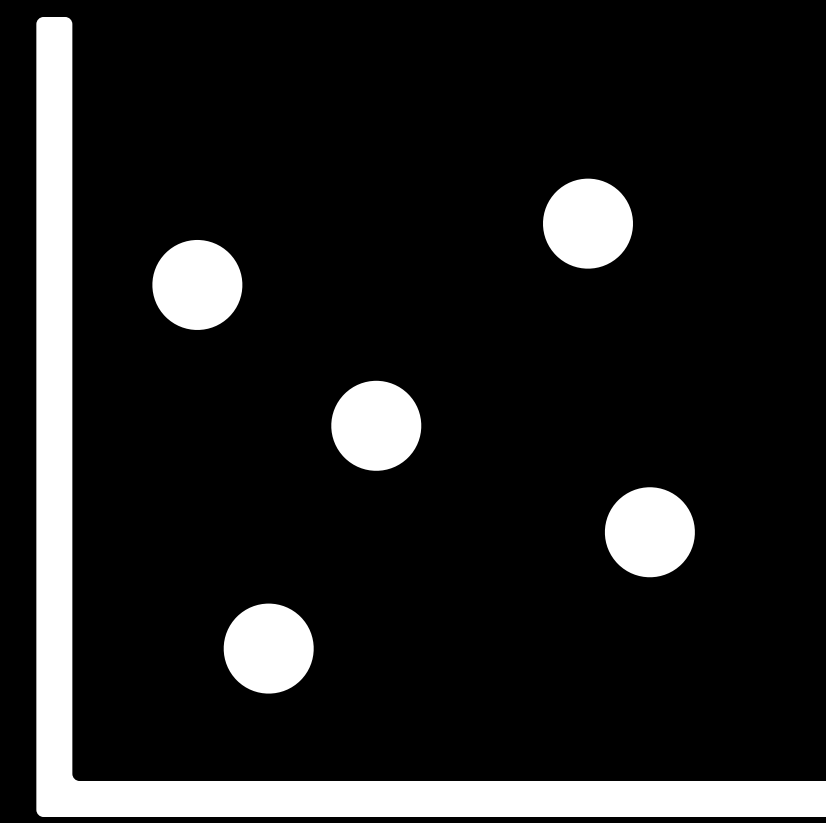
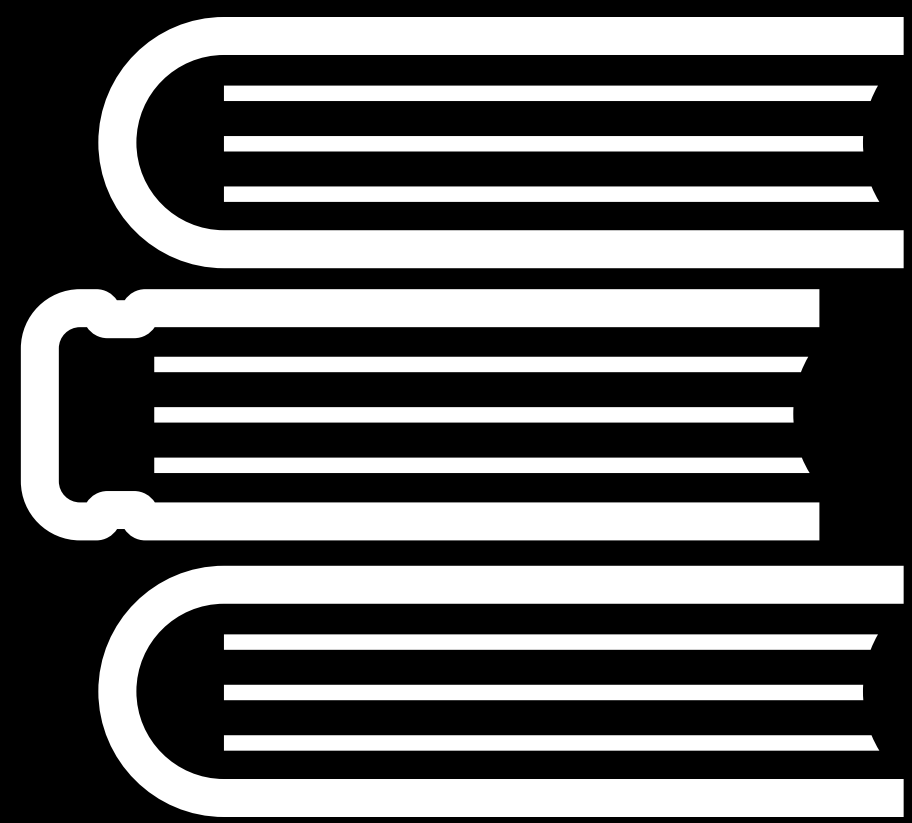
What content should we serve?

What products should we stock?





# AI has the potential to advance medicine



AI has techniques to rapidly **summarize** information, **predict** outcomes, and **learn** over time

Society has big expectations for AI in medicine



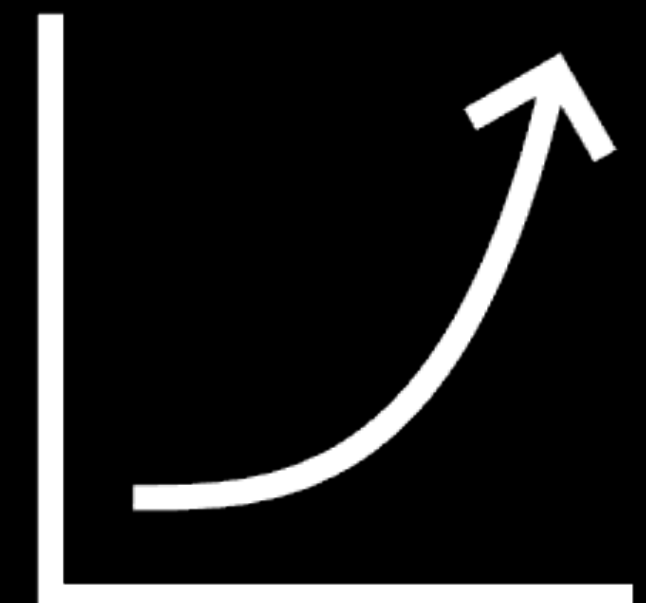
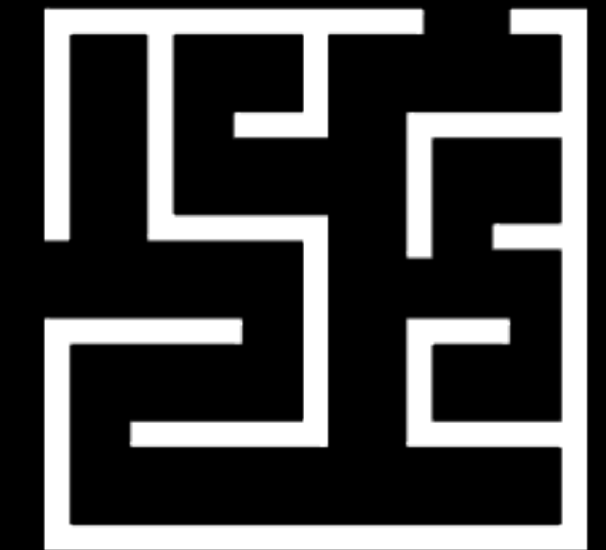
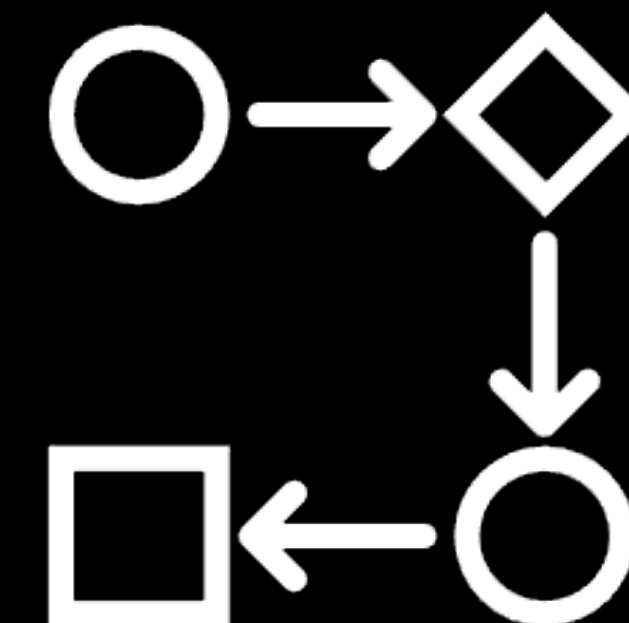
# AI is not a part of medical education

Use of AI in medicine is not straightforward

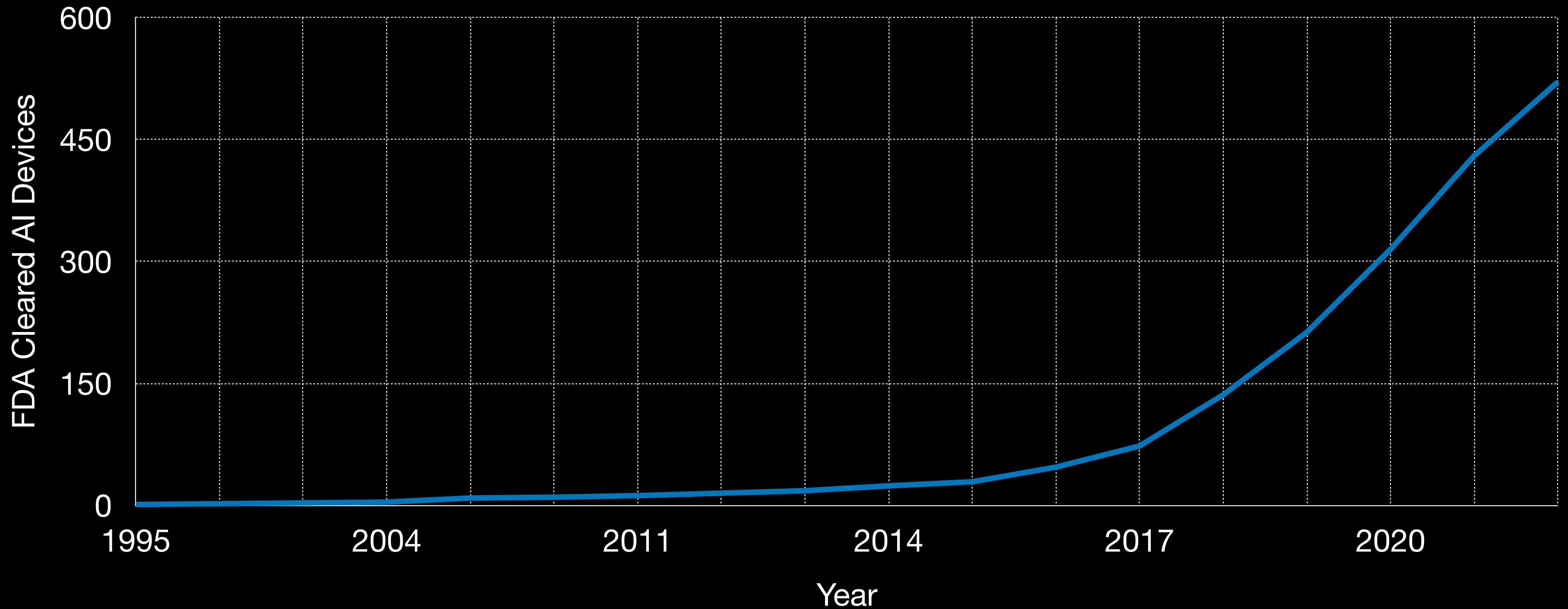
AI tools depend on complicated data and workflows that physicians understand

Medical AI adoption increasing

**Learners unprepared to use, assess, and develop AI tools**

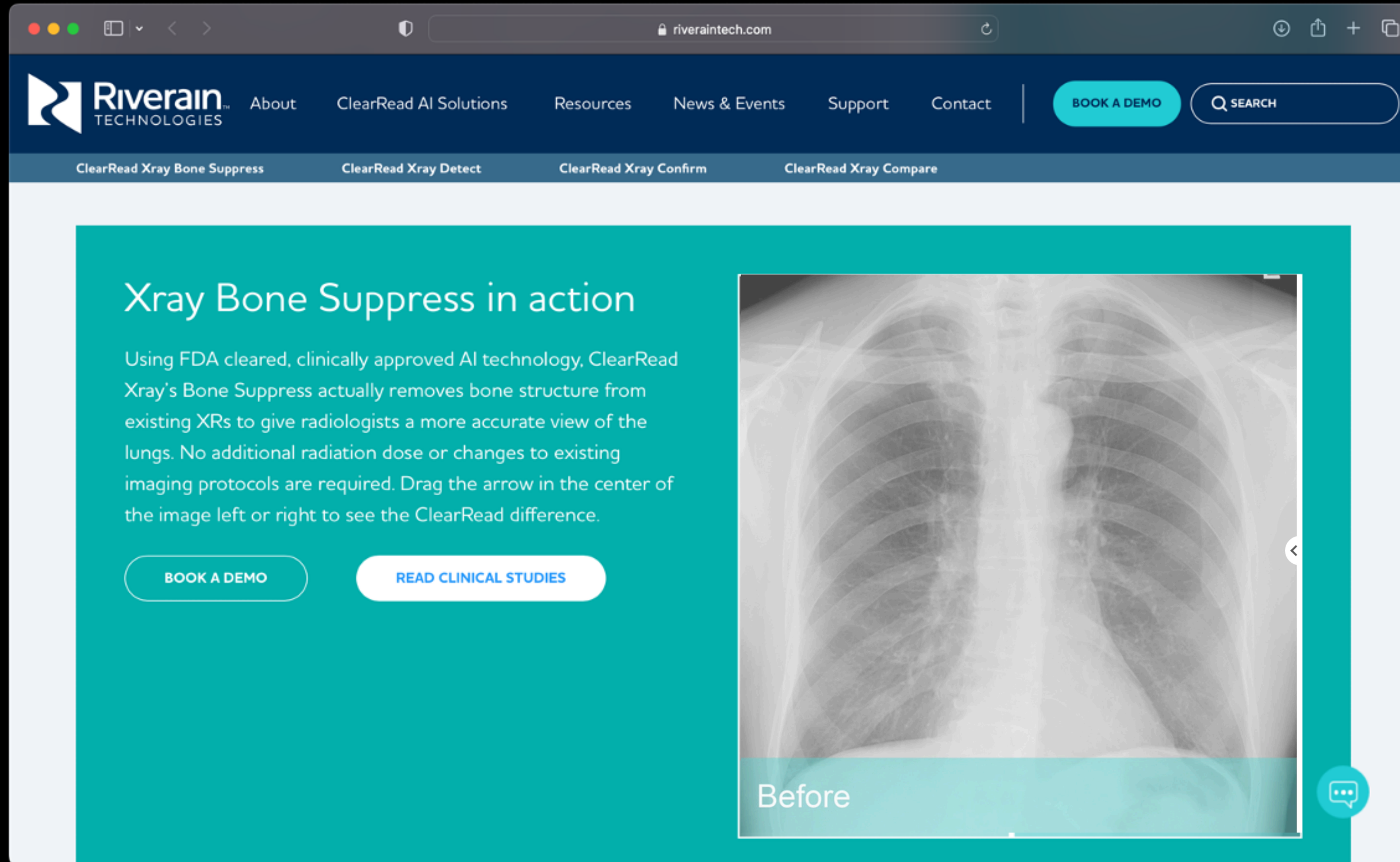


# Increasing prevalence of medical AI





# AI in use at Michigan Medicine



The screenshot shows a web browser window displaying the Riverain Technologies website. The browser's address bar shows 'riveraintech.com'. The website's navigation bar includes the Riverain Technologies logo, links for 'About', 'ClearRead AI Solutions', 'Resources', 'News & Events', 'Support', and 'Contact', a 'BOOK A DEMO' button, and a search bar. Below the navigation bar, there are four tabs: 'ClearRead Xray Bone Suppress', 'ClearRead Xray Detect', 'ClearRead Xray Confirm', and 'ClearRead Xray Compare'. The main content area features a teal background with the heading 'Xray Bone Suppress in action'. The text describes the technology as FDA cleared and clinically approved, explaining that it removes bone structure from X-rays to provide a clearer view of the lungs. It also notes that no additional radiation or changes to imaging protocols are required. Below the text are two buttons: 'BOOK A DEMO' and 'READ CLINICAL STUDIES'. To the right of the text is a large X-ray image of a chest. The image is split vertically, with the left side showing the original X-ray and the right side showing the result after bone suppression. The word 'Before' is written at the bottom left of the image. A chat bubble icon is visible in the bottom right corner of the image area.

Riverain  
TECHNOLOGIES

About ClearRead AI Solutions Resources News & Events Support Contact

BOOK A DEMO SEARCH

ClearRead Xray Bone Suppress ClearRead Xray Detect ClearRead Xray Confirm ClearRead Xray Compare

## Xray Bone Suppress in action

Using FDA cleared, clinically approved AI technology, ClearRead Xray's Bone Suppress actually removes bone structure from existing XRs to give radiologists a more accurate view of the lungs. No additional radiation dose or changes to existing imaging protocols are required. Drag the arrow in the center of the image left or right to see the ClearRead difference.

BOOK A DEMO READ CLINICAL STUDIES

Before

# Michigan AI in use

**THE JOURNAL OF UROLOGY**

**Development and Validation of Models to Predict Pathological Outcomes of Radical Prostatectomy in Regional and National Cohorts**

Erkin Ottes,<sup>1,2</sup> Brian T. Denton,<sup>1,2</sup> Bo Ou,<sup>3</sup> Adharsh Murali,<sup>4,5</sup> Sejin Merdan,<sup>1</sup> Gregory B. Auffenberg,<sup>1</sup> Spencer C. Hiller,<sup>6</sup> Brian R. Lane,<sup>6</sup> Arvin K. George<sup>7</sup> and Karandeep Singh<sup>1,2,3,4,5,6,7,8,9,10,11</sup> for the Michigan Urological Surgery Improvement Collaborative

**Abstract:** Prediction models are recommended by national guidelines to support clinical decision making in prostate cancer. Existing models to predict pathological outcomes of radical prostatectomy (RP), such as the Memorial Sloan-Kettering (MSK) model, Partin tables, and the Briganti nomogram—have been developed using data from tertiary care centers and may not generalize well to other settings.

**Materials and Methods:** Data from a regional cohort (Michigan Urological Surgery Improvement Collaborative [MUSIC]) were used to develop models to predict extraprostatic extension (EPE), seminal vesicle invasion (SVI), lymph node invasion (LNI), and recurrence-free survival (RFS) in patients undergoing RP. The MUSIC models were compared against the MSK model, Partin tables, and Briganti nomogram (for LNI) using data from a national cohort (Surveillance, Epidemiology, and End Results [SEER] registry).

**Results:** We identified 7,401 eligible patients in the SEER registry. The MUSIC model had good discrimination (SEER AUC EPE: 0.77; SVI: 0.86; LNI: 0.83; NOCDI: 0.77) and was well calibrated. While the MSK models had similar discrimination to the MUSIC models (SEER AUC EPE: 0.74; SVI: 0.86; LNI: 0.84; NOCDI: 0.76), they overestimated the risk of EPE, LNI, and NOCDI. The Partin tables had inferior discrimination (SEER AUC EPE: 0.67; SVI: 0.76; LNI: 0.76).

**Conclusion:** Prediction models are recommended by national guidelines to support clinical decision making in prostate cancer. Existing models to predict pathological outcomes of radical prostatectomy (RP), such as the Memorial Sloan-Kettering (MSK) model, Partin tables, and the Briganti nomogram—have been developed using data from tertiary care centers and may not generalize well to other settings.

**Keywords:** prostate cancer; prediction model; radical prostatectomy; surveillance, epidemiology, and end results; validation.

**Abbreviations and Acronyms:** AUC = area under the receiver operating characteristic curve; EPE = extraprostatic extension; LNI = lymph node invasion; MSK = Memorial Sloan-Kettering; MUSIC = Michigan Urological Surgery Improvement Collaborative; NOCDI = no recurrence-free survival; PSA = prostate-specific antigen; RP = radical prostatectomy; SEER = Surveillance, Epidemiology, and End Results Program; SVI = seminal vesicle invasion.

**Editorial:** This article is the third of 5 published in this issue for which category 1 CME credits can be earned. Instructions for obtaining credits are given with the questions on page 351 and 352.

**DOI:** 10.1097/JU.0000000000000130

**THE JOURNAL OF UROLOGY** Vol. 202, No. 2, February 2020

**www.urologyjournal.org**

Proceedings of Machine Learning Research 136:1–26, 2021

Machine Learning for Healthcare

**Mind the Performance Gap: Examining Dataset Shift During Prospective Validation**

Erkin Ottes<sup>1,2,3</sup>

**HHS Public Access**  
Author manuscript  
Infect Control Hosp Epidemiol. Author manuscript; available in PMC 2019 March 17.

Published in final edited form as:  
Infect Control Hosp Epidemiol. 2018 April ; 139(4): 425–433. doi:10.1017/S0950268817000048.

**A Generalizable, Data-Driven Approach to Predict Daily Risk of Clostridium difficile Infection at Two Large Academic Health Centers**

Jeebekh Oh, MS<sup>1,4</sup>, Maggie Meeker, MS<sup>1,4</sup>, Christopher Fusco, BS<sup>1</sup>, Robert McCaffrey, BS<sup>1</sup>, Krishna Rao, MD, MS<sup>1,4</sup>, Erin E. Ryan, MPH, CRP<sup>5,6</sup>, Laraine Washer, MD<sup>1,7</sup>, Lauren R. West, MPH<sup>1,8</sup>, Vincent B. Young, MD, PhD<sup>1,9,10</sup>, John Guttag, PhD<sup>1</sup>, David C. Hooper, MD<sup>1,11</sup>, Erica S. Shenoy, MD, PhD<sup>1,9,10,12</sup>, and Jenna Wilens, PhD<sup>1,3</sup>

**Abstract:** Clostridium difficile infection (CDI) is a leading cause of hospital-acquired infection. We developed a data-driven approach to predict the daily risk of CDI at two large academic health centers. We used data from a regional cohort (Michigan Urological Surgery Improvement Collaborative [MUSIC]) to develop models to predict CDI risk. The MUSIC models were compared against the MSK model, Partin tables, and Briganti nomogram (for LNI) using data from a national cohort (SEER registry). We identified 7,401 eligible patients in the SEER registry. The MUSIC model had good discrimination (SEER AUC EPE: 0.77; SVI: 0.86; LNI: 0.83; NOCDI: 0.77) and was well calibrated. While the MSK models had similar discrimination to the MUSIC models (SEER AUC EPE: 0.74; SVI: 0.86; LNI: 0.84; NOCDI: 0.76), they overestimated the risk of EPE, LNI, and NOCDI. The Partin tables had inferior discrimination (SEER AUC EPE: 0.67; SVI: 0.76; LNI: 0.76).

**Keywords:** Clostridium difficile infection; data-driven approach; machine learning; prediction model; validation.

**Abbreviations and Acronyms:** AUC = area under the receiver operating characteristic curve; EPE = extraprostatic extension; LNI = lymph node invasion; MSK = Memorial Sloan-Kettering; MUSIC = Michigan Urological Surgery Improvement Collaborative; NOCDI = no recurrence-free survival; PSA = prostate-specific antigen; RP = radical prostatectomy; SEER = Surveillance, Epidemiology, and End Results Program; SVI = seminal vesicle invasion.

RESEARCH SPECIAL PAPER

**Early identification of patients admitted to hospital for covid-19 at risk of clinical deterioration: model development and multisite external validation study**

Fahad Kamran,<sup>1,2</sup> Shengpu Tang,<sup>1,3</sup> Erkin Ottes,<sup>1,3</sup> Dustin S McEvoy,<sup>1,3</sup> Sameeh N Saleh,<sup>1,4</sup> Jim Gong,<sup>1,5</sup> Benjamin Y Li,<sup>1,5</sup> Sajon Dutta,<sup>1,6</sup> Kianan Liu,<sup>1,6</sup> Richard J Medford<sup>1,6</sup>

**Abstract:** The Epic Sepsis Model (ESM), a proprietary sepsis prediction model, is implemented at hundreds of US hospitals. The ESM's ability to identify patients with sepsis has not been adequately evaluated despite widespread use.

**Objective:** To externally validate the ESM in the prediction of sepsis and evaluate its potential clinical utility compared with usual care.

**Design, Setting, and Participants:** This retrospective cohort study was conducted among 27 657 patients aged 18 years or older admitted to Michigan Medicine, the academic health system of the University of Michigan, Ann Arbor, with 38 452 hospitalizations between December 6, 2018, and October 20, 2019.

**Exposures:** The ESM score, calculated every 15 minutes.

**Main Results and Conclusions:** Sepsis, as defined by a composite of 10 Centers for Disease Control and Prevention surveillance criteria and 12 International Statistical Classification of Diseases and Related Health Problems, tenth revision diagnostic codes accompanied by 2 systemic inflammatory response syndrome criteria and organ dysfunction criteria within 6 hours of one another. Model discrimination was assessed using the area under the receiver operating characteristic curve of the hospitalization-level and with prediction horizons of 4, 8, 12, and 24 hours. Model calibration was evaluated with calibration plots. The potential clinical benefit associated with the ESM was assessed by evaluating the added benefit of the ESM score compared with contemporary clinical practice based on timely administration of antibiotics. Mean fatigue was evaluated by comparing the clinical value of different alerting strategies.

**Results:** We identified 27 657 patients who had 38 452 hospitalizations (21 904 women [57%]; median age, 56 years [interquartile range, 25–89 years]) meeting inclusion criteria. Of all sepsis events (n=252) (7%), the ESM had a hospitalization-level area under the receiver operating characteristic curve of 0.63 (95% CI, 0.62–0.64). The ESM identified 183 of 2632 patients with sepsis (7%) who did not receive timely antibiotic administration. Highlighting the low sensitivity of the ESM in comparison with contemporary clinical practice, the ESM also did not identify 1783 patients with sepsis (67%) despite generating alerts for an ESM score of 6 or higher for 60% of all 38 454 hospitalized patients (16%), thus creating a large burden of alert fatigue.

**Conclusions and Relevance:** This external validation cohort study suggests that the ESM has poor discrimination and calibration in predicting the onset of sepsis. The widespread adoption of the ESM despite its poor performance raises fundamental concerns about sepsis management on a national level.

**Keywords:** coronavirus disease; deterioration index; prediction model; sensitivity study.

JAMA Internal Medicine | Original Investigation

**External Validation of a Widely Implemented Proprietary Sepsis Prediction Model in Hospitalized Patients**

Andrew Wong, MD, PhD<sup>1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100,101,102,103,104,105,106,107,108,109,110,111,112,113,114,115,116,117,118,119,120,121,122,123,124,125,126,127,128,129,130,131,132,133,134,135,136,137,138,139,140,141,142,143,144,145,146,147,148,149,150,151,152,153,154,155,156,157,158,159,160,161,162,163,164,165,166,167,168,169,170,171,172,173,174,175,176,177,178,179,180,181,182,183,184,185,186,187,188,189,190,191,192,193,194,195,196,197,198,199,200,201,202,203,204,205,206,207,208,209,210,211,212,213,214,215,216,217,218,219,220,221,222,223,224,225,226,227,228,229,230,231,232,233,234,235,236,237,238,239,240,241,242,243,244,245,246,247,248,249,250,251,252,253,254,255,256,257,258,259,260,261,262,263,264,265,266,267,268,269,270,271,272,273,274,275,276,277,278,279,280,281,282,283,284,285,286,287,288,289,290,291,292,293,294,295,296,297,298,299,300,301,302,303,304,305,306,307,308,309,310,311,312,313,314,315,316,317,318,319,320,321,322,323,324,325,326,327,328,329,330,331,332,333,334,335,336,337,338,339,340,341,342,343,344,345,346,347,348,349,350,351,352,353,354,355,356,357,358,359,360,361,362,363,364,365,366,367,368,369,370,371,372,373,374,375,376,377,378,379,380,381,382,383,384,385,386,387,388,389,390,391,392,393,394,395,396,397,398,399,400,401,402,403,404,405,406,407,408,409,410,411,412,413,414,415,416,417,418,419,420,421,422,423,424,425,426,427,428,429,430,431,432,433,434,435,436,437,438,439,440,441,442,443,444,445,446,447,448,449,450,451,452,453,454,455,456,457,458,459,460,461,462,463,464,465,466,467,468,469,470,471,472,473,474,475,476,477,478,479,480,481,482,483,484,485,486,487,488,489,490,491,492,493,494,495,496,497,498,499,500,501,502,503,504,505,506,507,508,509,510,511,512,513,514,515,516,517,518,519,520,521,522,523,524,525,526,527,528,529,530,531,532,533,534,535,536,537,538,539,540,541,542,543,544,545,546,547,548,549,550,551,552,553,554,555,556,557,558,559,560,561,562,563,564,565,566,567,568,569,570,571,572,573,574,575,576,577,578,579,580,581,582,583,584,585,586,587,588,589,590,591,592,593,594,595,596,597,598,599,600,601,602,603,604,605,606,607,608,609,610,611,612,613,614,615,616,617,618,619,620,621,622,623,624,625,626,627,628,629,630,631,632,633,634,635,636,637,638,639,640,641,642,643,644,645,646,647,648,649,650,651,652,653,654,655,656,657,658,659,660,661,662,663,664,665,666,667,668,669,670,671,672,673,674,675,676,677,678,679,680,681,682,683,684,685,686,687,688,689,690,691,692,693,694,695,696,697,698,699,700,701,702,703,704,705,706,707,708,709,710,711,712,713,714,715,716,717,718,719,720,721,722,723,724,725,726,727,728,729,730,731,732,733,734,735,736,737,738,739,740,741,742,743,744,745,746,747,748,749,750,751,752,753,754,755,756,757,758,759,760,761,762,763,764,765,766,767,768,769,770,771,772,773,774,775,776,777,778,779,780,781,782,783,784,785,786,787,788,789,790,791,792,793,794,795,796,797,798,799,800,801,802,803,804,805,806,807,808,809,810,811,812,813,814,815,816,817,818,819,820,821,822,823,824,825,826,827,828,829,830,831,832,833,834,835,836,837,838,839,840,841,842,843,844,845,846,847,848,849,850,851,852,853,854,855,856,857,858,859,860,861,862,863,864,865,866,867,868,869,870,871,872,873,874,875,876,877,878,879,880,881,882,883,884,885,886,887,888,889,890,891,892,893,894,895,896,897,898,899,900,901,902,903,904,905,906,907,908,909,910,911,912,913,914,915,916,917,918,919,920,921,922,923,924,925,926,927,928,929,930,931,932,933,934,935,936,937,938,939,940,941,942,943,944,945,946,947,948,949,950,951,952,953,954,955,956,957,958,959,960,961,962,963,964,965,966,967,968,969,970,971,972,973,974,975,976,977,978,979,980,981,982,983,984,985,986,987,988,989,990,991,992,993,994,995,996,997,998,999,1000.</sup>

**Editorial:** This article is the third of 5 published in this issue for which category 1 CME credits can be earned. Instructions for obtaining credits are given with the questions on page 351 and 352.

**DOI:** 10.1093/ajph/2020.100.0000000000000130

**THE JOURNAL OF UROLOGY** Vol. 202, No. 2, February 2020

**www.urologyjournal.org**

Prostate Cancer Outcomes

In Hospital Infection Risk

Deterioration Risk

In Hospital Sepsis Risk



# We've got to start training physicians on AI fundamentals

Physicians shouldn't just be "users"

Should be actively involved in creating, evaluating, and improving AI

Leadership in AI dependent on:  
**understanding** how it works &  
**partnership** with engineers

**Cell Reports Medicine** **CellPress**  
OPEN ACCESS

**Commentary**  
**Teaching artificial intelligence as a fundamental toolset of medicine**

Erkin Ötles,<sup>1,2,6,7,\*</sup> Cornelius A. James,<sup>3,5</sup> Kimberly D. Lomis,<sup>4</sup> and James O. Woolliscroft<sup>6</sup>

<sup>1</sup>Medical Scientist Training Program, University of Michigan Medical School, Ann Arbor, MI, USA  
<sup>2</sup>Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI, USA  
<sup>3</sup>Department of Pediatrics, University of Michigan, Ann Arbor, MI, USA  
<sup>4</sup>American Medical Association, Chicago, IL, USA  
<sup>5</sup>Departments of Internal Medicine and Learning Health Sciences, University of Michigan, Ann Arbor, MI, USA  
<sup>6</sup>Present address: 1225 Beal Avenue, Ann Arbor, MI 48109, USA  
<sup>7</sup>Twitter: @eotles  
\*Correspondence: eotles@umich.edu  
<https://doi.org/10.1016/j.xcrm.2022.100824>

**Artificial intelligence (AI) is transforming the practice of medicine. Systems assessing chest radiographs, pathology slides, and early warning systems embedded in electronic health records (EHRs) are becoming ubiquitous in medical practice. Despite this, medical students have minimal exposure to the concepts necessary to utilize and evaluate AI systems, leaving them under prepared for future clinical practice. We must work quickly to bolster undergraduate medical education around AI to remedy this. In this commentary, we propose that medical educators treat AI as a critical component of medical practice that is introduced early and integrated with the other core components of medical school curricula. Equipping graduating medical students with this knowledge will ensure they have the skills to solve challenges arising at the confluence of AI and medicine.**

The promise of artificial intelligence (AI) to aid the practice of medicine has long been a topic of discussion.<sup>1</sup> What was once an abstract discussion of the future of medicine is now a clinical reality. Software employing AI is found throughout the clinical care continuum. The US Food and Drug Administration (FDA) has approved over 100 AI software devices.<sup>2</sup> The purposes of these software devices range from measuring pulmonary nodules in chest CT scans to detecting different cell types in peripheral blood smears and screening for diabetic retinopathy using photos taken in primary-care settings. However, not all AI systems require FDA approval. Some of the most widely deployed AI systems are early warning systems that fall outside the FDA's jurisdiction. AI systems for detecting in-hospital deterioration and sepsis are deployed at hundreds of US hospitals.<sup>3</sup> The recent increased interest in medical AI is due to the availability of massive amounts of data, facilitated by widespread adoption of electronic health records (EHRs), and advances in AI techniques, driven by a combination of new hardware and computational methods. Despite the accelerating use of AI in clinical practice, the pace of incorporating AI concepts into medical education has been slow and superficial.<sup>4</sup> Only recently has it been proposed that AI concepts be included in medical education curricula.<sup>5,6</sup> Most suggestions to date have framed training in AI as an added layer to current medical school curricula, hereafter referred to as undergraduate medical education (UME). Recommendations for incorporating AI into UME range widely, covering the gamut from teaching medical students how to code to EHR usage and the ethics surrounding the adoption of AI.<sup>7</sup> However, proposals that treat AI as an additional curricular element or course struggle to gain traction in an overcrowded curriculum. In this commentary, we offer the collective perspective of a medical student, practicing physician, and medical educators. We propose that medical schools view AI as a fundamental component of medical practice and deeply integrate it throughout UME.<sup>8</sup> We believe UME must quickly transition to address AI as a fundamental toolset, meaning that it contains many interrelated techniques that underpin the practice of medicine across specialties and care environments. However, the breadth of AI presents a challenge for medical educators seeking to provide a foundation in UME that can be built upon throughout one's career. AI uses computational methods to process data, from identifying a pattern to generating a prediction or a recommendation. AI can be considered an umbrella term encapsulating many techniques, such as natural language processing and machine learning (ML). Practices from computer science, statistics, decision science, and operations research intersect with AI. These procedures are built upon a foundation of data processing dependent on two types of thinking: computational—being able to provide instructions to computers unambiguously—and statistical—being able to analyze the information derived from processes subject to randomness. To add to the challenge, like the practice of medicine, the practice of AI is a combination of art and science, as AI systems are components of even larger and more complicated socio-technical systems. Therefore, in addition to technical knowledge, applying AI effectively in clinical practice demands careful consideration of the context, patient values and preferences, ethics, policy, and physician user experiences.

Cell Reports Medicine 3, 100824, December 20, 2022 © 2022 The Author(s). 1  
This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

# ChatGPT



# ChatGPT = Chatbot + GPT3

Chatbot: developed by OpenAI  
mix of supervised & reinforcement learning

GPT3: Generative Pre-trained Transformer 3  
type of **large language model** (fancy predictive text)

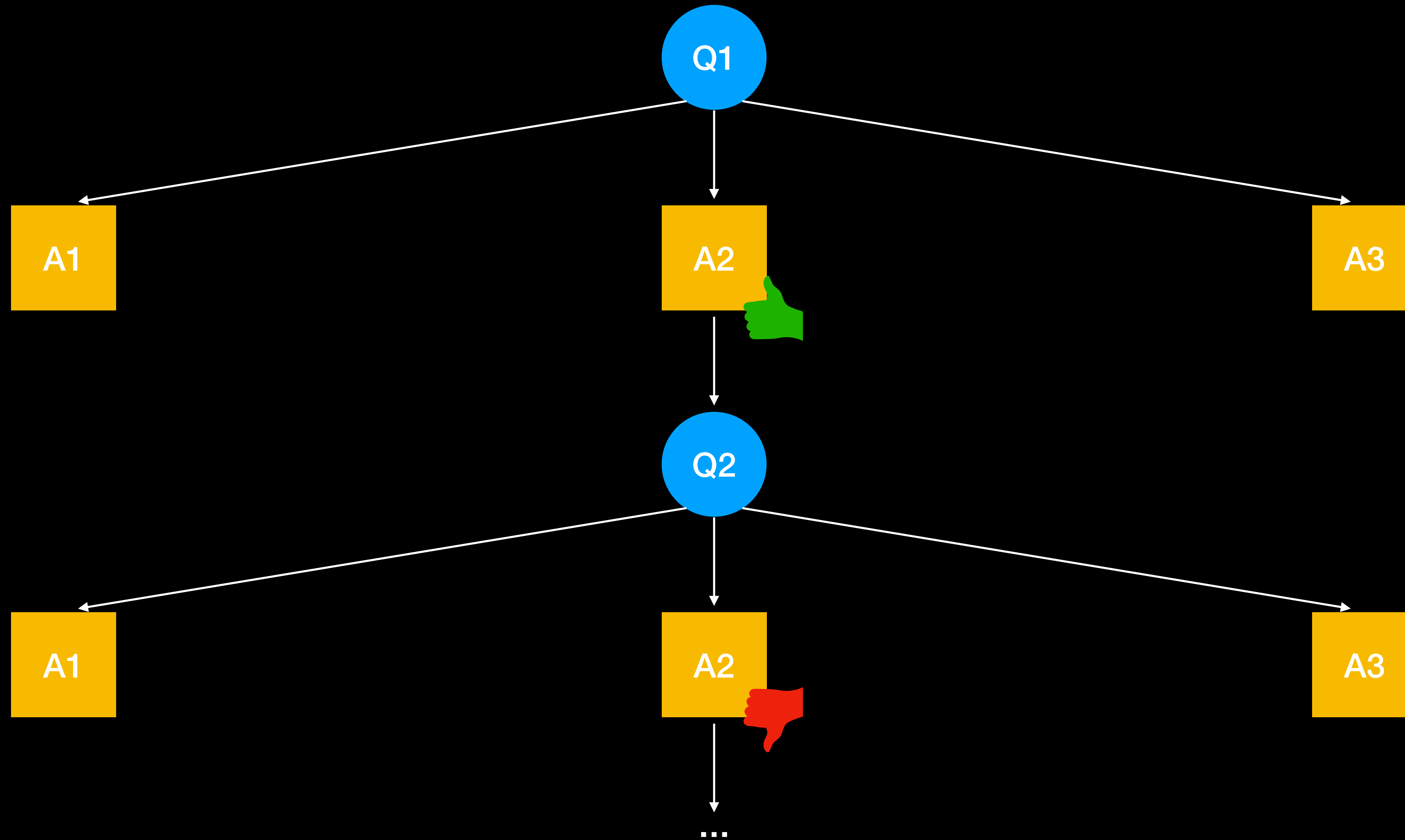
“The quick brown fox jumps over the \_\_\_\_\_”

Lazy	95%
Slow	2%
Fun	1%
...	
Zyzyva	0%

Trained on all available text on the internet



# Chat is a branching tree





# Major issues with large language models

Based on what ever data it was trained on

May not be relevant, accurate, or pleasant

Generative process is inherently stochastic

Response choices and sentence construction depend on sampling distributions randomly

Hard to evaluate and verify

How often will it be right? What is right?

# Recap

AI provides medicine with a set of powerful tools

We need to train physicians to be active leaders of the development and evaluation of these tools

ChatGPT is cool, but problematic

Healthcare AI cycle is dependent on physician-engineer collaboration

Interfaces between physician users and AI tools will need significant human factors engineering



# Questions?

Comments? Concerns? Violent disagreements?

Erkin Ötleş  
Twitter: @eotles  
[eotles.com](http://eotles.com)

