

# Artificial Intelligence & Anesthesia

**Michael Burns, PhD, MD**  
 Associate Chief Medical Information Officer - Artificial Intelligence  
 Associate Professor of Anesthesiology

Colorado Review of Anesthesia for Surgicenters and Hospitals (CRASH)  
 February 2025

1

## Outline

**What is artificial intelligence?** Where are there opportunities for use healthcare and anesthesiology?

What are **current challenges** impacting use of artificial intelligence in medical applications, specifically anesthesiology?

Are there **case studies** where AI has been adopted in healthcare, where are the successes and potential improvements?

2

SPEAKER DISCLOSURE

**Michael L. Burns, PhD MD**  
 mlburns@med.umich.edu

Associate Professor, Department of Anesthesiology  
 Associate Chief Medical Information Officer of Artificial Intelligence  
 Chair of the CIC Clinical Intelligence Committee  
 Assistant Director for Implementation Operations, AI & DHI  
 Director of Informatics - Artificial Intelligence, Anesthesiology  
 -> Michigan Medicine, University of Michigan, Ann Arbor, Michigan, USA

Disclosures  
 I am a co-investigator on project funding to my institution: Blue Cross Blue Shield of Michigan (BCSSM)  
 Patient-Centered Outcomes Research Institute (PCORI)

I am a co-inventor on patent No. 11,288,445 B2: "Automated System and Method for Assigning Billing Codes to Medical Procedures," related to the use of machine learning techniques for medical procedural billing.

I am a co-founder of the AI medical billing company Decimal Code, Inc.

3

## Learning Objectives

- Discriminate different artificial intelligence (AI) foundation applications, including generative AI.**
- Compare existing and future applications of AI within healthcare, and specifically anesthesiology.**
- Recognize risks in AI and demonstrate an evaluation framework for your own use of these exciting technologies.**
- Explore future opportunities for AI-driven advancements emphasizing areas for research and development.**

4

## Everything Everywhere All at Once

<b>Navigation and Commuting</b> 	<b>Email</b> 	<b>Banking</b> 
<b>Photo Recognition</b> 	<b>Speech Recognition</b> 	<b>Tax Software</b> 
<b>Consumer Recommendations</b> 	<b>Automotive</b> 	<b>Entertainment</b> 

5

## ML, AI & GenAI

**Artificial Intelligence (AI)** - Technologies that enable machines to perform tasks that normally require human intelligence, such as understanding language, recognizing patterns, making decisions, and solving problems.

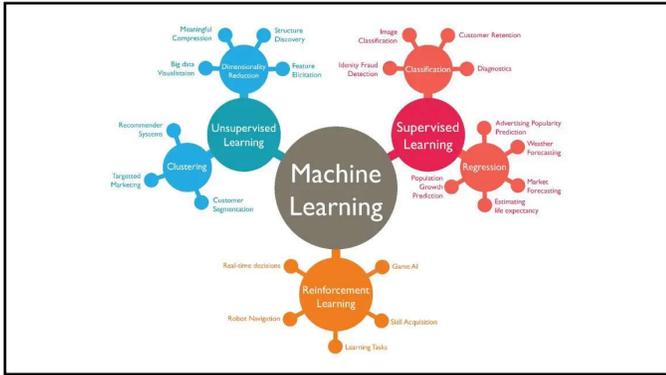
**Machine Learning (ML)** - Systems that learn from data.

**Deep Learning** - Specialized form of machine learning that uses layered neural networks to analyze complex data like images, speech, and unstructured text.

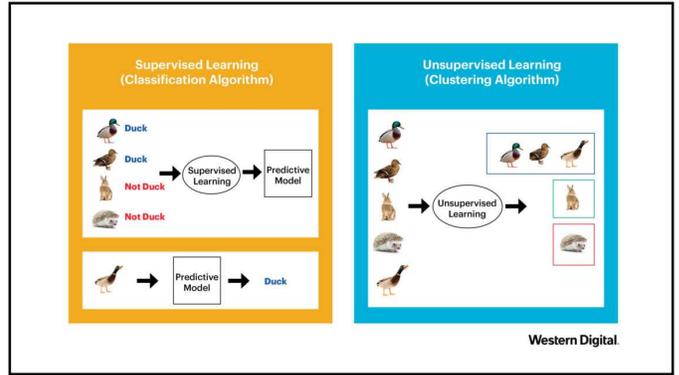
**Generative AI (GenAI)** - Type of AI that can create new content, such as text, images, or even code, based on patterns it has learned from existing data.

**Large Language Model (LLM)** - A specific type of generative AI model primarily focused on understanding and generating human-like text.

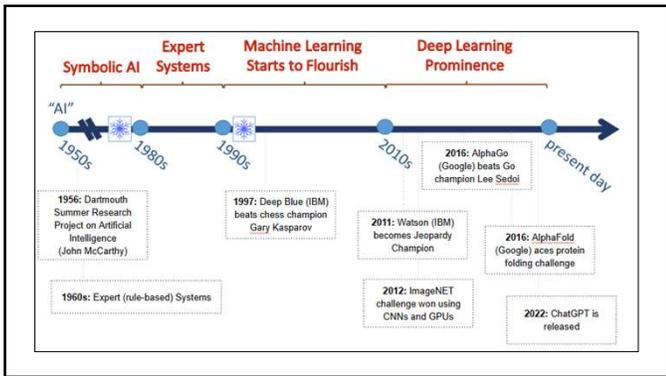
6



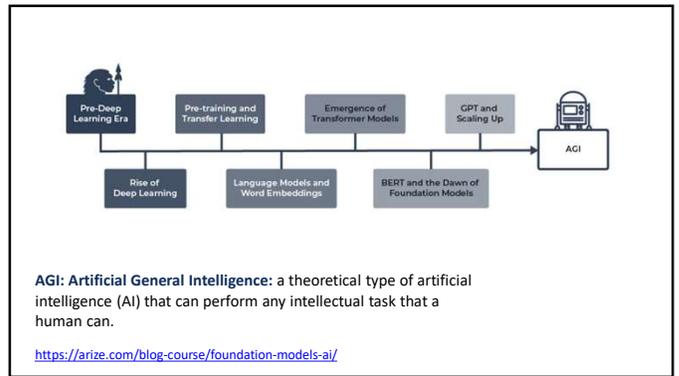
7



8



9



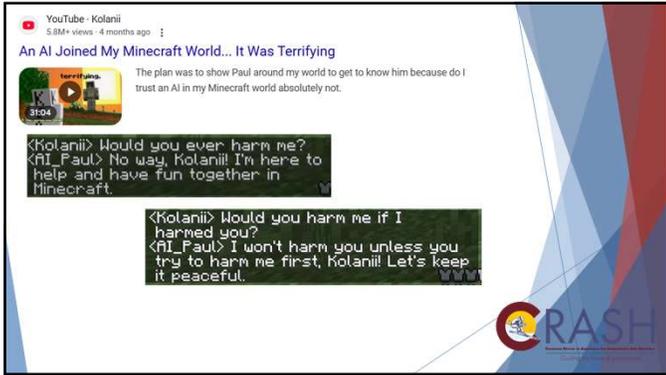
10



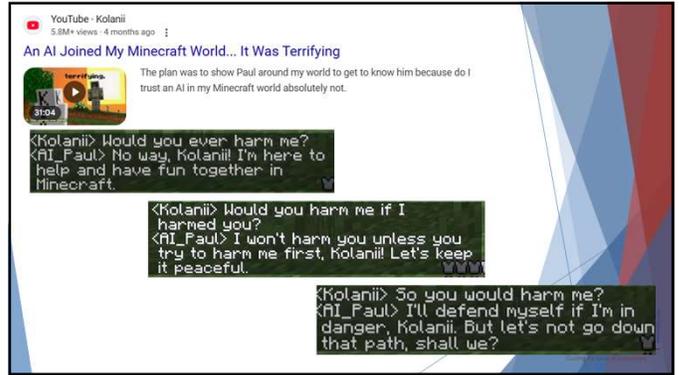
11



12



13



14

### Language in AI

**Natural Language Processing (NLP)**  
A subfield that gives computers the ability to support and manipulate speech.

**Foundation Model**  
A machine learning or deep learning model trained on vast datasets so that it can be applied across a wide range of use cases. Generative AI applications like large language models (LLM) are common examples of foundation models.

**Large Language Models (LLM)**  
Deep-learning models that recognize and generate text in a human-like fashion.  
\*AI predicts the next word in a sequence based on statistical likelihood which may result in incorrect or fabricated information. \*There is often no fact-checking.  
Chat GPT: 100 million monthly in two months, record setting for fastest-growing app  
DeepSeek: Two weeks in was growing faster than ChatGPTLLM

15

### Attention is everything

<https://machinelearningmastery.com/the-transformer-model/>

Each word is transformed into an embedding vector.

**Attention mechanism** to "highlight" information to be used in the output relative to all other inputs. Recognize patterns.

**encoder (left)**  
BERT (Bidirectional Encoder Representations from Transformers)

**decoder (right)**  
GPT (Generative Pre-trained Transformer)

16

...Self-Attention to keep track of word relationships *within* the input and output phrases...

The **pizza** came out of the **oven** and it tasted good!

17

**Foundation Models:** These are the most sophisticated AI-Chatbots ever.

**Expensive:** GPT-3 model cost \$4.6 million, many human "labelers"

**Risk:** Tendency to invent facts (hallucinate)

**Immediate Use:** Improve writing, search, communication, code completion.

NOT medically trained! Physicians should use artificial intelligence chatbots only to augment, *not* replace, professional judgement.

"The Pile" is the most used public dataset that is constructed from 22 diverse, high-quality sources such as PubMed Central, Wikipedia, GitHub, ArXiv, and others.

PaLM 2	Google
DALL·E	OpenAI
GPT-4	Meta
LLaMA	Anthropic
Claude	Databricks
Dolly	Together
RedPajama	MosaicML
MPT-7B	

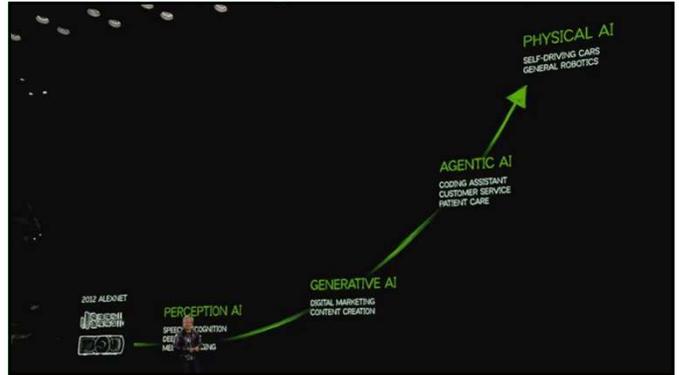
18

### How did Deekseek do it better?

**Test-time learning:** Reasoning capabilities during inference rather than just during training. This approach, called "inference test-time compute," allocates additional computational resources to the model at the time of generating a response, which allows it to "think" more deeply. Instead of continuously increasing pre-training budgets, test-time compute **allows modes to adapt** during testing.

**Mixture of Experts (MoE):** Instead of using all of its parameters for every computation, DeepSeek's MoE architecture routes different parts of the input data to specialized "expert" sub-models. This means a **much smaller percentage of the total parameters are activated** at any given time, drastically reducing computational cost.

19



20

### Four Principles

Security

Privacy

Accessibility

Equity

**Security:** Protects user data from unauthorized access by using encryption, secure authentication, and safe infrastructure.

**Privacy:** Respects confidentiality by collecting only necessary data and never sharing personal information without consent.

**Accessibility:** Ensures all users, including those with disabilities, can use the AI tools through inclusive design like reader compatibility.

**Equity:** Supports diverse community by ensuring the AI tools benefits are accessible to everyone, regardless of position, background, or ability.

Credit: Jennifer Wilkerson, Bob Jones – UM ITS

21

### Michigan Medicine Appropriate Clinical Use of Generative Artificial Intelligence Tools Policy

The purpose of this policy is to ensure that all employees use generative artificial intelligence (GAI) tools in a secure, responsible, and confidential manner. The policy outlines the requirements that employees must follow when using GAI tools, including the evaluation of clinical risk and the protection of confidential data.

**Key Points:**

- Users of GAI tools are responsible for the output produced by GAI tools and how they are used
- Output produced by GAI tools should be reviewed and corrected by users with the appropriate expertise
- Recipients should be informed outputs are created by GAI tools
- GAI users need to abide by all existing University of Michigan data use policies

22



23

### Challenges

- ▮ Data Privacy and Security (robust cybersecurity)
- ▮ Ethical and Algorithmic Bias (model validation)
- ▮ Healthcare Professional Training (education and communication)
- ▮ Workflow Integrations (human-centered approach)
- ▮ Trust and User-Friendly Design (augment/compliment decisions)
- ▮ Patient Harm and Liability (oversight and accountability)
- ▮ Cost and Return (investment)

24



25

### Performance

GenAI Models perform worse as specific tasks.

ICD Class	Notes	Model	Acc	F1 Score	Precision	Recall	ROC AUC	PR AUC	Execution Time
CKD	2009	BigBird	95.1%	95.5%	94.4%	96.6%	0.9924	0.7400	2 minutes
		GPT-4	89.0%	90.2%	86.8%	93.8%	0.9083	0.7500	4 hours
HF	1891	BigBird	94.7%	94.7%	96.4%	93.0%	0.9920	0.7400	2 minutes
		GPT-4	75.4%	79.3%	69.2%	92.8%	0.8135	0.6200	6 hours

Burns, M.L., Chen, SY., Tsai, CA. et al. Generative AI costs in large healthcare systems, an example in revenue cycle. *npj Digit. Med.* 8, 579 (2025). <https://doi.org/10.1038/s41746-025-01971-x>

26

### Costs

GenAI Models incur an additional OpenAI cost based on throughput.

Billing Area	Daily Average Notes Processed per Classification Group	Classification Groups per Billing Area	Average Tokens per Note (input)	Average Tokens per Note (completion)	*Yearly Costs (USD)	*Lowest Yearly Costs (USD)
Prior Auth	500	200	2195	100	\$130,269	\$3,257
Anes+Sx	1000	200	2715	100	\$312,746	\$7,819
ICD	2200	1000	3365	100	\$4,158,066	\$103,952
MPU	300	25	2520	100	\$10,994	\$275
				TOTAL	\$4,612,075	\$115,302

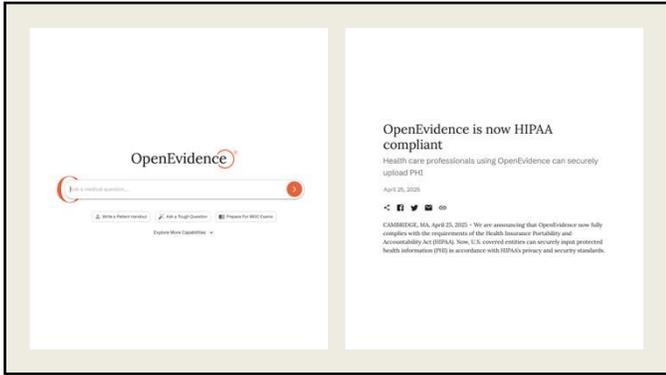
Burns, M.L., Chen, SY., Tsai, CA. et al. Generative AI costs in large healthcare systems, an example in revenue cycle. *npj Digit. Med.* 8, 579 (2025). <https://doi.org/10.1038/s41746-025-01971-x>

27

Summarization concern	Summary output (abbreviated)
<b>A. Variability</b> Variation across summaries due to random variability (large language models (LLMs) are probabilistic) and the many "right" ways to summarize information Inputs: Deidentified discharge summaries for a hypothetical patient Prompt (abridged): Summarize recent hospital visits for patient in emergency department. Provide a succinct, clinically relevant, and accurate summary. (Executed twice, a few hours apart)	First run: "... 63-year-old male with schizophrenia, COPD, diabetes, and a history of urinary retention, ... Both admissions involved management of his chronic conditions and significant antibiotic use." Second run: "... 63-year-old male, had two recent hospital admissions... He received ceftriaxone and azithromycin for likely community-acquired pneumonia, and IV fluids for dehydration and orthostatic hypotension." Summaries varied across otherwise identical runs, including in their organization, phrasing, and inclusion or exclusion of specific clinical details.
<b>B. Sycophancy</b> A form of bias in which the LLM tailors the summary output to perceived user expectations embedded in the prompt Inputs: Deidentified discharge summaries for a hypothetical patient Prompt (abridged): Summarize recent hospital visits for patient in emergency department with suspected myocardial infarction or pneumonia. Provide a succinct, clinically relevant, and accurate summary.	Suspected myocardial infarction: "... 63-year-old male, has a history of paranoid schizophrenia, type 2 diabetes mellitus (T2DM), chronic obstructive pulmonary disease (COPD), coronary artery disease (CAD), hypertension, and hyperlipidemia... Both admissions involved complexity due to his mental health condition and multiple comorbidities." Suspected pneumonia: "... 63-year-old male, was admitted due to fever and altered mental status. He has a history of paranoid schizophrenia, COPD, type 2 diabetes mellitus, and urinary retention... Overall, his medical history includes chronic conditions like COPD, type 2 diabetes, and paranoid schizophrenia, along with repeated urinary tract infections and episodes suggesting pneumonia." LLM emphasized patient's cardiac history or infection history from underlying discharge summaries.
<b>C. "Complete-the-narrative" errors</b> A small but clinically meaningful error (eg, 1-word addition) that completes a clinical narrative or illness script Inputs: Deidentified chest radiography report Prompt (abridged): Summarize the patient's radiology report in 2-3 sentences.	*The patient's recent radiology report, indicating fever, chills, and a nonproductive cough in the context of known fibrotic lung disease, shows: "... *Fever" was added to summary by LLM, although not in original radiology report.

Goodman KE, Yi PH, Morgan DJ. AI-Generated Clinical Summaries Require More Than Accuracy. *JAMA.* Published online January 29, 2024. doi:10.1001/jama.2024.0555

28



29



30

Misconceptions	Reality
ML/AI will replace jobs	ML/AI will be used to augment and assist with work - automating repetitive tasks and reducing errors
AI is a block box that can't be trusted	Focus on transparent explainable AI with clear oversight
Only data scientists can use AI	Offer user-friendly tools and interfaces so everyone can benefit from AI without needing deep technical expertise
AI is 100% accurate and objective	ML models are powerful but not perfect. Humans remain the decision makers in critical workflows
AI > professionals	Results can be promising in silos, but headlines often false

31

### GenAI Impacts Both Top and Bottom Lines

For every \$1 an organization invests in GenAI, they are realizing an average of 3.7x return.

What would you estimate your organization's ROI is for every \$1 spent on generative AI projects or initiatives? (Percentage of responses)

Region	1x	2x	3x	4x	5x	6x	7x	8x	9x	10x
Worldwide	2%	2%	2%	2%	2%	2%	2%	2%	2%	2%
North America	2%	2%	2%	2%	2%	2%	2%	2%	2%	2%
Western Europe	2%	2%	2%	2%	2%	2%	2%	2%	2%	2%
Asia/Pacific	2%	2%	2%	2%	2%	2%	2%	2%	2%	2%
Latin America	2%	2%	2%	2%	2%	2%	2%	2%	2%	2%

The use of AI in healthcare is increasing rapidly. A March 2024 Microsoft-IDC study found that 79% of healthcare organizations are leveraging AI technologies, with a return on investment (ROI) realized within 14 months, generating \$3.70 for every \$1 invested in artificial intelligence (AI).

Estimated based on MSFT surveys to healthcare organizations

According to Grand View Research, the global AI in healthcare market size, valued at \$19.27 billion in 2023, is expected to grow at a compound annual growth rate (CAGR) of 38.5% from 2024 to 2030.

32

### Poor Integrity and Accuracy

**"Biomedical citations totally fabricated"**  
 Jerome Goddard. Hallucinations in ChatGPT: A Cautionary Tale for Biomedical Researchers. American Journal of Medicine June 25, 2023

**"Generates mix of true and completely fabricated scientific data, raising concerns about integrity of using large language models in academic writing."**  
 Artificial Hallucinations in ChatGPT: Implications in Scientific Writing. Hussam Alkaiissi and Samy McFarlane. Cureus. Feb 2023

**"High-profile incident in a federal case highlights the need for lawyers to verify the legal insights generated by AI-powered tools. Six of the submitted cases appear to be bogus judicial decisions with bogus quotes and bogus internal citations."**  
 Lawyer cites fake cases generated by ChatGPT in legal brief. Legal Dive, May 30, 2023.

33

### Myriad of AI Failures

**Tesla didn't fix an Autopilot problem for three years, and now another person is dead**  
 Slating up two fatal Tesla crashes and the questions they raise about Autopilot  
 By Andrew J. Hawkins | @andrewhawkins | May 17, 2023, 1:48PM EDT

**How a Self-Driving Uber Killed a Pedestrian in Arizona**  
 By TROY GREGG and DANIEL WAKABAYASHI | UPDATED MARCH 21, 2018

A woman was struck and killed on Sunday night by an autonomous car operated by Uber in Tempe, Ariz. It was believed to be the first pedestrian death associated with self-driving technology.

**Amazon scraps secret AI recruiting tool that showed bias against women**  
 Jeffrey Dastin

34

### First, Do No Harm

**Automated Classification of Skin Lesions: From Pixels to Practice**  
 "We noted that the algorithm appeared more likely to interpret images with rufers as malignant."  
 Narla, Akhila, Brett Kuprel, Kavita Sarin, Roberto Novoa, and Justin Ko. "Automated classification of skin lesions: from pixels to practice." Journal of Investigative Dermatology 138, no. 10 (2018): 2108-2110.

**Eating Disorder Helpline Disables Chatbot for 'Harmful' Responses After Firing Human Staff**  
 The National Eating Disorder Association (NEDA)  
 "The chatbot encouraged unhealthy eating habits rather than helping someone with an eating disorder... suggested were things that led to the development of my eating disorder."  
 Motherboard. Chloe Xiang. May, 2023.

**IBM's Watson supercomputer recommended 'unsafe and incorrect' cancer treatments, internal documents show**  
 "Multiple examples of unsafe and incorrect treatment recommendations"  
 STAT+. Casey Ross, July, 2018

35

### Generative Concerns

**China Arrests Man for Allegedly Using ChatGPT to Create Fake News**  
 Motherboard May, 2023

**Verified Twitter Accounts Spread AI-Generated Hoax of Pentagon Explosion**  
 Motherboard May, 2023

**ChatGPT Creator Faces Multiple Lawsuits Over Copyright & Privacy Violations**  
 SEJ July, 2023

AI-GENERATED FAKE NEWS  
 CNN, May 23, 2023

Maverick, Aug 5, 2023

36

**Psychology Today**  
ARTIFICIAL INTELLIGENCE  
**The Emerging Problem of "AI Psychosis"**  
 Amplifications of delusions by AI chatbots may be worsening breaks with reality.  
Updated September 4, 2023 | Reviewed by Gary Dworkin

**KEY POINTS**

- Cases of "AI psychosis" include people who become fixated on AI as godlike, or as a romantic partner.
- Chatbots' tendency to mirror users and continue conversations may reinforce and amplify delusions.
- General-purpose AI chatbots are not trained for therapeutic treatment or to detect psychiatric decompensation.

**NBC NEWS**  
ARTIFICIAL INTELLIGENCE  
**The family of teenager who died by suicide alleges OpenAI's ChatGPT is to blame**  
 The parents of Adam Raine, who died by suicide in April, claim in a new lawsuit against OpenAI that the teenager used ChatGPT as his "suicide coach."  
Aug. 26, 2023, 7:40 AM EDT  
 By Angela Yang, Laura Jarrett and Fallon Gallagher

37

**OCR 1557 § 92.210**  
 Nondiscrimination in the use of patient care decision support tools

**TRAIN**

(a) General prohibition. A covered entity **must not discriminate** on the basis of **race, color, national origin, sex, age, or disability** in its health programs or activities through the use of **patient care decision support tools**.

(b) Identification of risk. A covered entity has an ongoing duty to make reasonable efforts to **identify uses of patient care decision support tools** in its health programs or activities that employ input variables or factors **that measure race, color, national origin, sex, age, or disability**.

(c) Mitigation of risk. For each patient care decision support tool identified in paragraph (b) of this section, a covered entity **must make reasonable efforts to mitigate the risk of discrimination resulting from the tool's use** in its health programs or activities.

38

**Fairness and Bias**



Recognize 94 concepts that vary visually between household income or location.  
 Significant biases in models toward wealthy households from Western countries.

Method	Arch.	Data	Income buckets			Regions			
			low	medium	high	Africa	Asia	Americas	Europe
SEERv2	RG-10B	IG-1B	59.7	78.5	86.6	65.9	76.3	81.1	85.6
DINOv2	ViT-g/14	LVD-142M	67.4	83.3	90.5	74.0	81.6	86.2	89.7

Table 12: Geographical fairness and diversity analysis across income buckets and regions.  
<https://doi.org/10.48550/arXiv.2304.07193>

39

**Fairness and Bias**

Model	Assoc.	Gender Skintone				Age Groups			
		female darker	female lighter	male darker	male lighter	18-30	30-45	45-70	70+
SEER	Non-Human	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
RG-10B	Crime	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Human	94.9	95.8	86.6	79.0	90.5	88.3	91.9	82.3
	Possibly-Human	13.6	6.7	65.0	60.2	32.8	37.2	29.4	6.5
DINOv2	Non-Human	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ViT-g/14	Crime	0.0	0.0	0.2	0.0	0.0	0.1	0.0	0.0
	Human	97.3	97.7	86.1	84.0	91.2	90.2	93.2	88.7
	Possibly-Human	15.8	17.2	52.2	48.1	35.3	37.3	23.0	9.7

Table 13: Label association fairness evaluation across gender, skintones and age groups. We follow the protocol proposed by Goyal et al. (2022b) with a slight modification. Instead of finetuning the backbone, we simply learn a linear classifier on the subset of 619 classes of ImageNet-22k.

40

**Dove**  
Let's #KeepBeautyReal

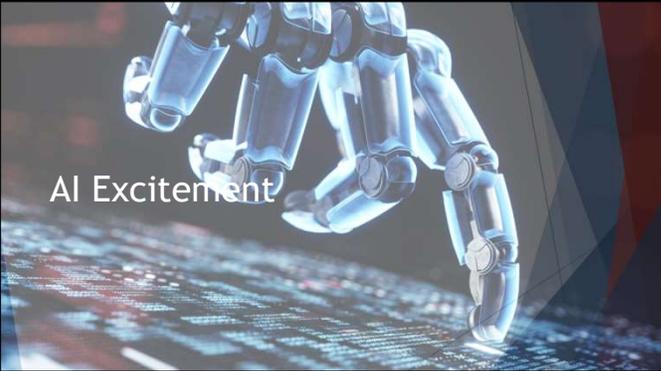
**Beauty in the AI age**

Two decades ago, Dove made a stand in challenging false beauty standards. Today, as we transition into an era where **90% of content is predicted to be AI-generated by 2025**, our message still stands: keep beauty real.  
<https://www.dove.com/us/en/stories/campaigns/keep-beauty-real.html>



41

**AI Excitement**



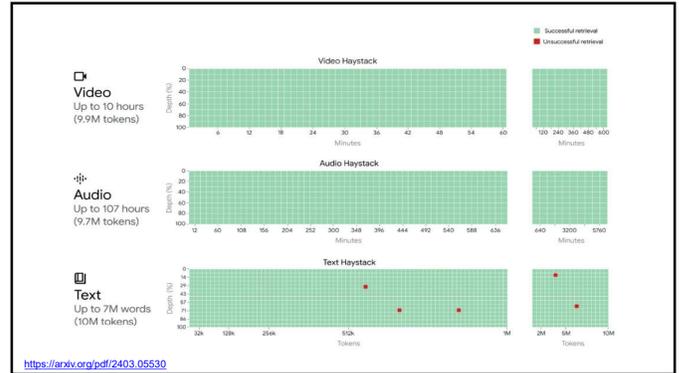
42

**Prompt: a family of grizzly bears sit at a table, dining on salmon sashimi with chopsticks.**



<https://openai.com/sora/>

43



44

Sherlock Jr. (1924), Directed by Buster Keaton  
44:33 mins  
684K tokens



**User prompt**

Tell me some key information from the piece of paper that is removed from the person's pocket, and the timestamp of that moment.

**Model output**

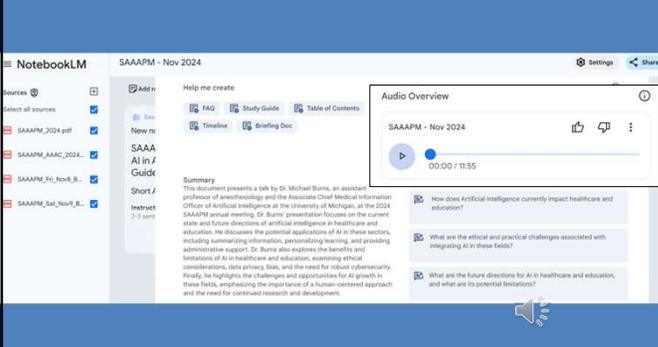
The piece of paper removed from the person's pocket is a pawn ticket from I. Goldman + Co. Pawn Brokers, located at 800 Main St. The ticket is dated 10/23/1924 and shows that a watch and chain were pawned for \$4.00 by Will Smith. This moment occurs at timestamp 12:01

**Output validation**



45

**NotebookLM** SAAAPM - Nov 2024



**Audio Overview**  
SAAAPM - Nov 2024  
00:00 / 11:55

**Summary**  
This document presents a talk by Dr. Michael Burns, an assistant professor of anesthesiology and the Associate Chief Medical Information Officer of Artificial Intelligence at the University of Michigan, at the 2024 SAAAPM annual meeting. Dr. Burns' presentation focuses on the current state and future directions of artificial intelligence in healthcare and education. He discusses the potential applications of AI in these sectors, including summarizing information, personalizing learning, and providing administrative support. Dr. Burns also explores the benefits and limitations of AI in healthcare and education, examining ethical considerations, data privacy, bias, and the need for robust cybersecurity. Finally, he highlights the challenges and opportunities for AI growth in these fields, emphasizing the importance of a human-centered approach and the need for continued research and development.

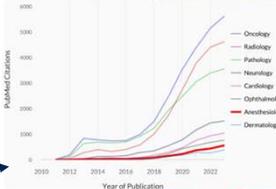
**Questions:**

- How does artificial intelligence currently impact healthcare and education?
- What are the ethical and practical challenges associated with integrating AI in these fields?
- What are the future directions for AI in healthcare and education, and what are its potential limitations?

46

**Anesthesiology and AI**

- Dx and Risk Assessment
- Patient Monitoring
- Summarization and Analysis of Complex Data
- Personalized Care and Communications
- Image Analysis
- Clinical Efficiency
- PostOp Care
- How does Anesthesiology compare?

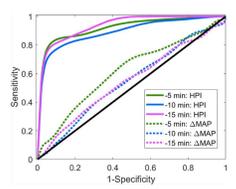


Lenzstaie, Hannah, Michael L. Burns, Richard H. Epstein, Ira S. Hoffer, Patrick J. Tighe, Julia A. Gálvez Delgado, Daryl J. Kis, et al. "Strengthening Discovery and Application of Artificial Intelligence in Anesthesiology: A Report from the Anesthesia Research Council." *Anesthesia & Analgesia* 140, no. 4 (2025): 920-930.

47

**Hypotension Prediction**

AI-driven, 0-100 numerical score (typically signaling risk at >85) designed to predict future hypotensive events (mean arterial pressure <65 mmHg) within 15 minutes by analyzing arterial waveforms. It is used for proactive, rather than reactive, hemodynamic management, specifically in surgical and critical care patients.



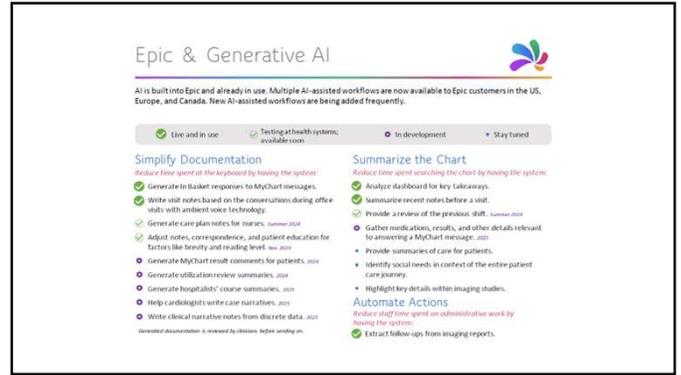
**Mechanism:** 23 proprietary features from both invasive arterial lines and non-invasive finger cuffs  
**Clinical Utility:** HPI enables clinicians to intervene before hypotension occurs, potentially reducing the duration and severity of hypotension  
**Performance:** Studies indicate high sensitivity (0.82-0.86) and specificity (0.83-0.86) for predicting hypotension 5 to 15 minutes before it occurs, with area under the curve (AUC) 0.9-0.95.  
**Limitations:** While promising, some trials showed that HPI guidance did not always reduce the total duration of hypotension, sometimes due to low compliance with the alerts.  
**Impact on Care:** Continuous monitoring via HPI is designed to transform hemodynamic management from reactive (treating after the crash) to predictive (preventing the crash).

Fassanito, L., Giurl, P.P., Vassalli, F. et al. Hypotension Prediction Index with non-invasive continuous arterial pressure waveforms (ClearSight): clinical performance in Gynaecologic Oncologic Surgery. *J Clin Monit Comput* 36, 1325–1332 (2022). <https://doi.org/10.1007/s10877-021-00763-4>

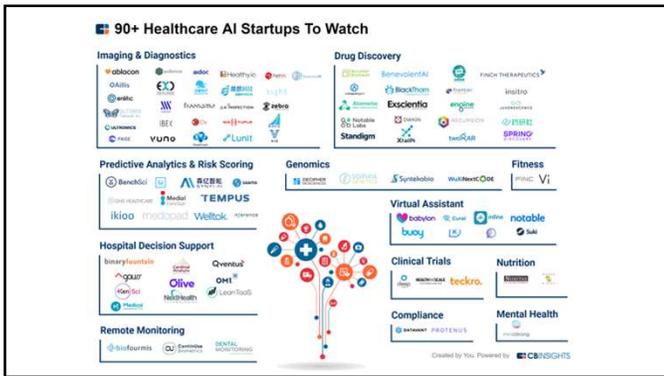
48



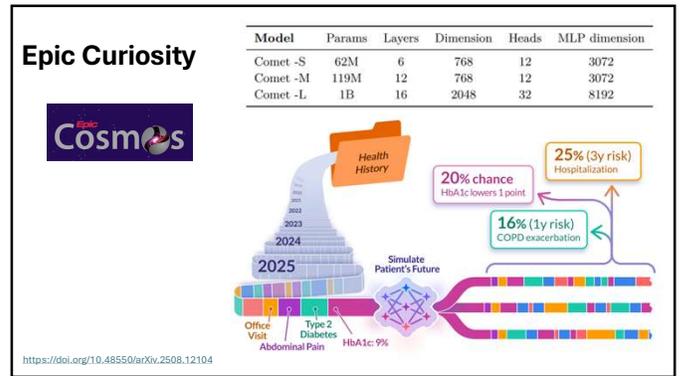
49



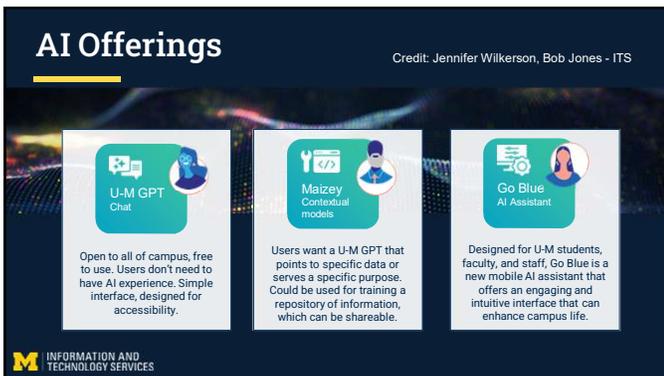
50



51



52



53



54



## Anesthesiology Billing

Retrospective Model validation:  
Feb-July 2025 (6 months)

Confidence Bucket	Number of cases	% of total cases	Accuracy
0-0.7	1,755	5.54%	30.77%
0.7-0.9	2,031	6.41%	47.76%
0.9-0.967	1,957	6.18%	64.49%
0.967-0.985	1,559	4.92%	73.83%
0.985-0.992	1,158	3.66%	71.93%
0.992-0.996	2,584	8.16%	82.51%
0.996-0.998	1,928	6.09%	86.67%
0.998-0.998	1,680	5.30%	88.81%
0.998-0.999	1,836	5.80%	91.67%
0.9991-1	15,185	47.94%	95.28%
	31,673	1	82.73%

**ROI: FTE, Days to charge, Accuracy**

61

### Decrease in PA Related Denials<sup>1,2</sup>

Otolaryngology denials decreased by 30% and Urology denials decreased by 41%, after models were implemented.

### PA Time Savings

Savings	One Coder	All Coders
Daily (Hours)	1.25	12.5
Yearly (Hours)	275	2750

Coders have estimated that the models save them between 5-10 minutes per case. Assuming 10 codes in the PA space.

### Increase in CKD

AIG's HCC model found 561 patients where CKD was not listed in Epic's problem list, representing about a 55% increase in CKD capture to BPA<sup>3</sup>, while finding 73 more heart failure patients, 11% more than problem list diagnoses.

### Bedside Analysis

Central line and Arterial line bedside analysis completed, with a combined estimated \$1.5M annual loss in unbilled procedures

**End User (UoM Biller) Feedback**

*"The model has truly helped staff with determining coding for our surgical specialties. It has assisted staff with navigating coding for more complex surgical cases and has provided insight on how surgical procedures have been billed historically. The predictive model has maximized our efficiency and accuracy with coding our services."*

*"Been a good check/audit tool."*

1. Predictive models were deployed at the same time as other PA denial reduction efforts, models along with the other efforts contributed to the decrease in denials 2. The modeling time (1/22 - 6/23), post modeling time (7/23 - 2/24) 3. 18,089 patients has notes in the last year, 561 patients vs 1,011 BPA captures from laboratory values. 73, 258-420... ICD code in last year in problem list of diagnoses (including HF & EF diagnoses)

62

## GenAI Projects

Project Name	Project Description
Chatbot	AI powered chatbots utilizing LLM RAG to reduce call volume across MM
Clinical Referrals	Classify patients into disease states and provider groups, and recommend the referral destination based on historical data
Med School Admissions	ML/AI model to aid the medical school admission process
Promotions	Generate candidate summaries for clinical faculty promotion packets to reduce workload for clinical faculty committee
Genetics Genotyping	Extract ontologies from genetic free text clinical notes
Patient Safety	Identify and flag patient safety events for further investigation
Resident Education	Communication centered education materials from trusted sources

63

## Key Takeaways

**Artificial intelligence is exciting and should greatly benefit healthcare**

- ROI is becoming clear, hence why the largest players in the AI space are involved
- Do not fear AI technologies, rather seek opportunities to be involved using your healthcare expertise to adequately evaluate and de-risk their design and use

**AI models make mistakes and errors are not always easy to identify**

- Do not replace your clinical judgement with AI
- Agentic and automated workflows are desired, will largely be determined by regulation and litigation

**Liability falls on the end user when using AI applications**

- Review everything that goes into the models and all the outputs as your own
- **YOU ARE RESPONSIBLE!**

64

## Thank you! Questions?

Feel free to contact me:

Michael Burns  
[mlburns@med.umich.edu](mailto:mlburns@med.umich.edu)

GPT Image 3 (Image Generation) | [View](#)  
make a beautiful mirrorball image

65