Adoption of AI for Healthcare

Clinical Transformation through AI

Vijay Aski **Partner Director**
**Microsoft AI Platform**

9th Edition
CAHOTECH 2024

Introduction

- Vijay Aski – Partner Director, AI Platform Microsoft
- 22+ years in MSFT, 8+ years in AI Platform
- Moved to India this year to lead the AI efforts in Microsoft India after 25 years in US
- Head the Model Training, Customization, Evaluation and Responsible AI effort
- Lead both the OpenAI (GPT*) and OSS Models (Llama) efforts

Exponential AI advancement vs Human Adaptability – And a feeling of disorientation

"There are two ways to adopt exponential technology:

too early or too late"
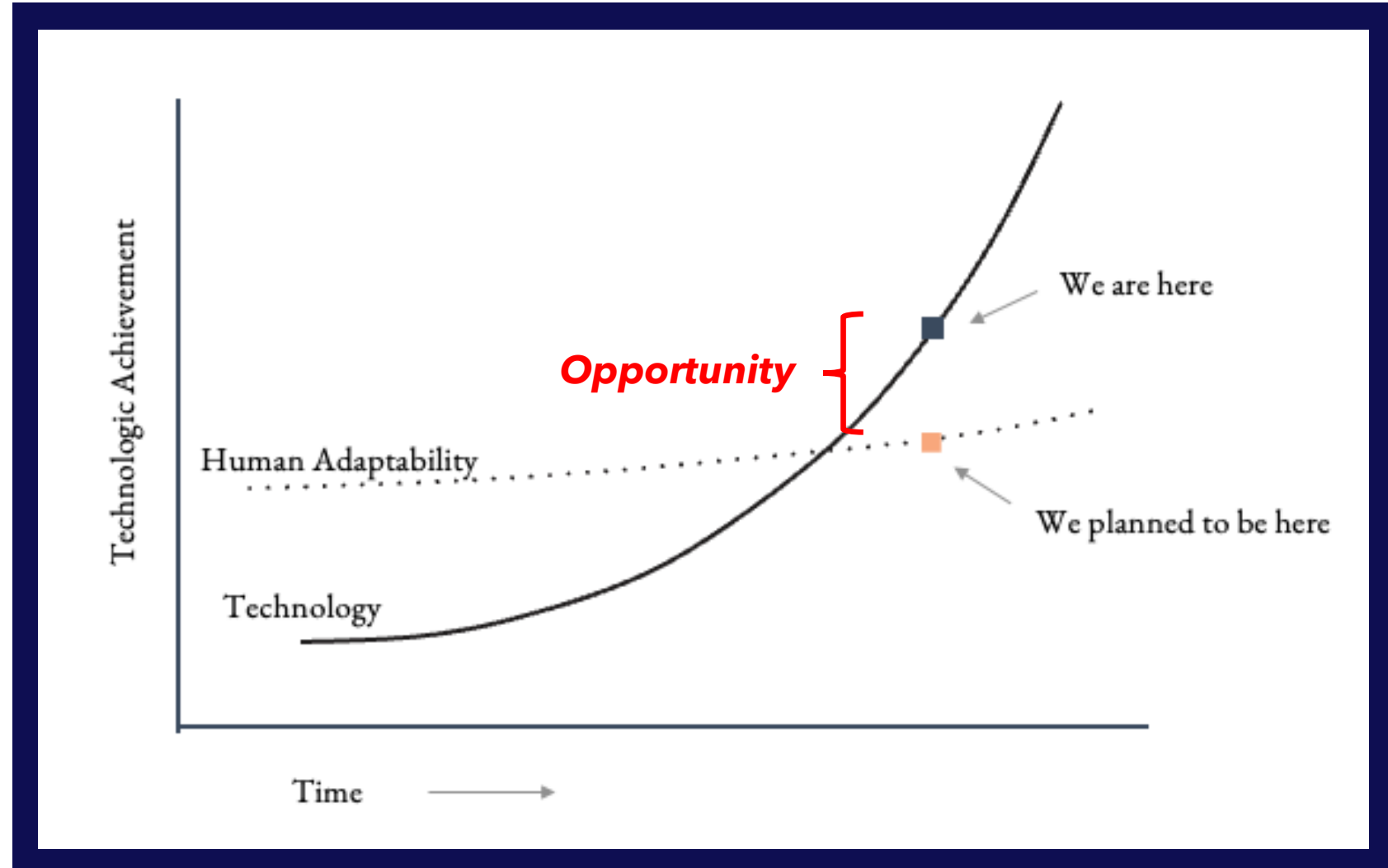
AI is poised to reshape healthcare

IMPROVE PROVIDER AND PATIENT EXPERIENCE

Ambient clinical intelligence,
chart summarization ...

REDUCE MEDICAL ERRORS

AI oversight on meds,
procedures, discharge ...

IMPROVE ACCESS TO CARE

Universal access to medical knowledge,
overcome language barriers ...

REDUCE ADMIN COSTS

Medical coding and billing,
quality reporting ...

IMPROVE MEDICAL STAFF SKILLS

AI co-pilots guiding providers,
nurses, MAs...

PERSONALIZATION OF CARE

Personalized treatments based
on real-world evidence

Opportunity Ikigai – Golden Intersection

99 of medical data is unstructured and multi-modal. 50 Petabytes produced each year. 97% goes unused

Multi-modal models' capability is increasing exponentially

Models reaching PhD-level reasoning & agentic workflows can do multi-step tasks effortlessly

Cost of AI is decreasing exponentially – 2 order magnitude in 2 years!

AI is powering fundamental & clinical research at expedited pace

Affordability, accessibility and healthcare needs are increasing

Technology adoption and diffusion is standardizing across globe

Medical costs, medical errors, legal liability, provider burnout- are trending towards unsustainable rates

Regulatory bodies, Government, Institutions, Tech companies and Investors are all leaned in and deeply invested

Massive revenue opportunity by optimizing patient care

Healthcare Financial Trends for 2022

Do not judge scaling laws based on progress GPT-4o

Large models are trained every c.1.5-2 years and are optimised between large training cycles
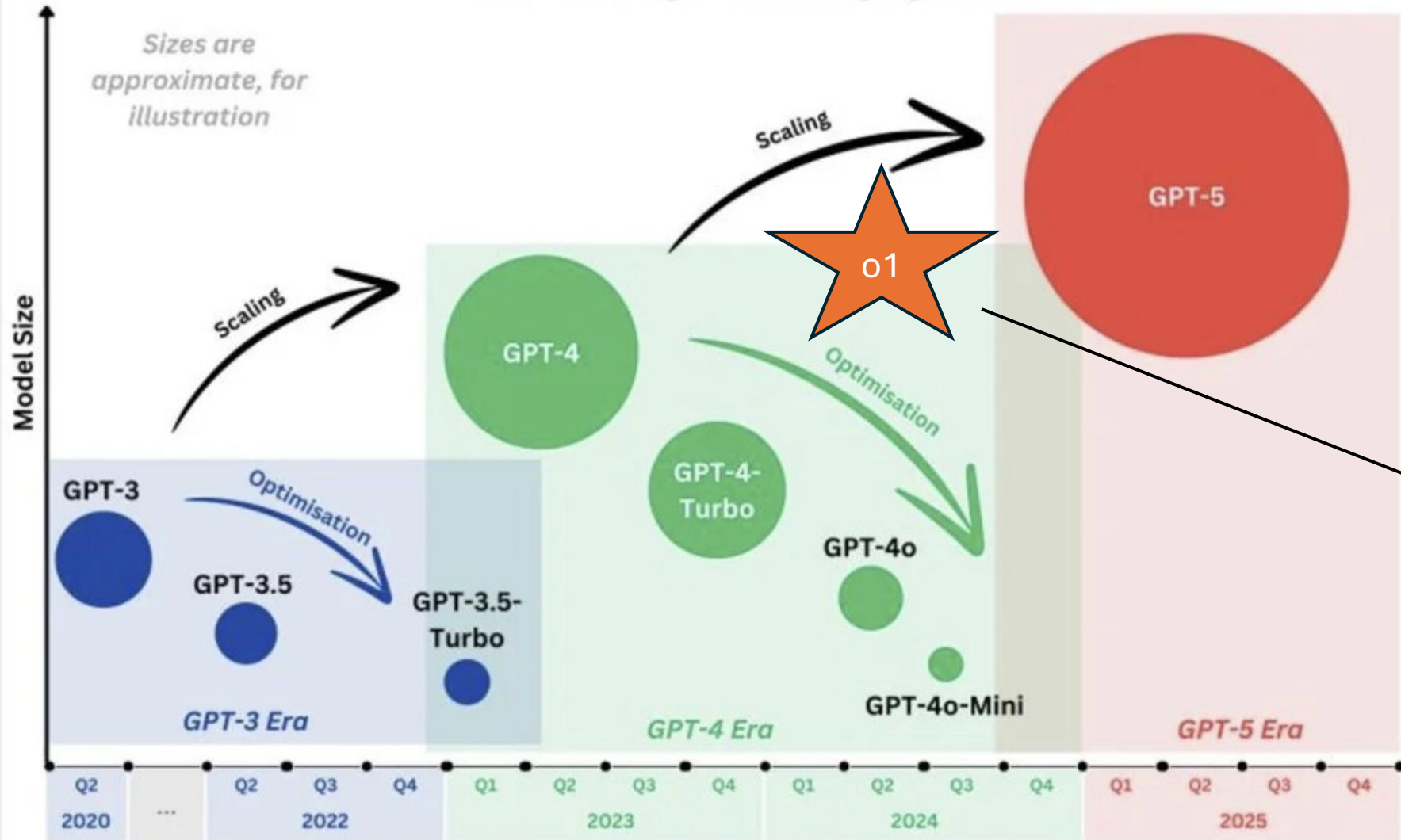
Model quality is getting extremely good

Increasing in size and capability

While reducing in cost and optimizations

Customizability and steerability

- PhD-level reasoning
- Can think before responding
- Complex coding tasks
- Instruction following and workflow management

Source: Peter Gostev (https://www.linkedin.com/in/peter-gostev/)

Hours per week (by specialty) spent on paperwork and administrative tasks:

1. Physical medicine and rehabilitation: 19 hours
T-2. Critical care: 18 hours
T-2. Internal medicine: 18 hours
T-2. Nephrology: 18 hours
T-2. Neurology: 18 hours
T-2. Oncology: 18 hours
7. Family medicine: 17 hours
T-8. Cardiology: 16 hours
T-8. Psychiatry: 16 hours
T-8. Diabetes and endocrinology: 16 hours
T-11. OB-GYN: 15 hours
T-11. Pediatrics: 15 hours
T-11. General surgery: 15 hours
T-14. Orthopedics: 14 hours
T-14. Urology: 14 hours
T-14. Otolaryngology: 14 hours
T-17. Emergency medicine: 13 hours
T-17. Gastroenterology: 13 hours
T-19. Dermatology: 11 hours
T-19. Plastic surgery: 11 hours
T-19. Radiology: 11 hours
22. Ophthalmology: 10 hours
23. Anesthesiology: 9 hours

Annals of Internal Medicine®

LATEST   ISSUES   IN THE CLINIC   JOURNAL CLUB   MULTIMEDIA   SPECIALTY COLLECTIONS   CME / MOC   AUTHORS / SUBMI

Original Research | 4 February 2020

Physician Time Spent Using the Electronic Health Record During Outpatient Encounters
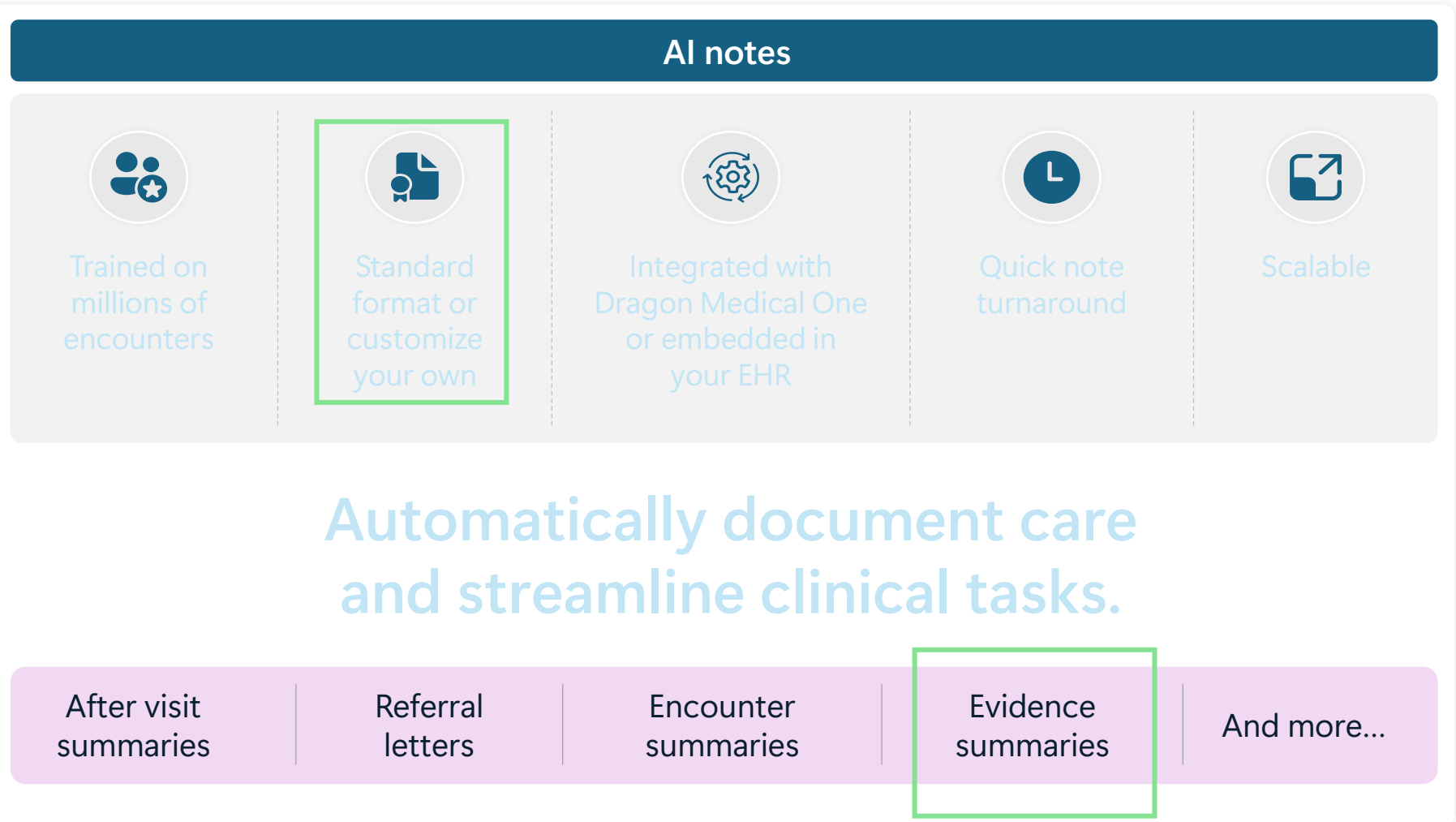
A Descriptive Study

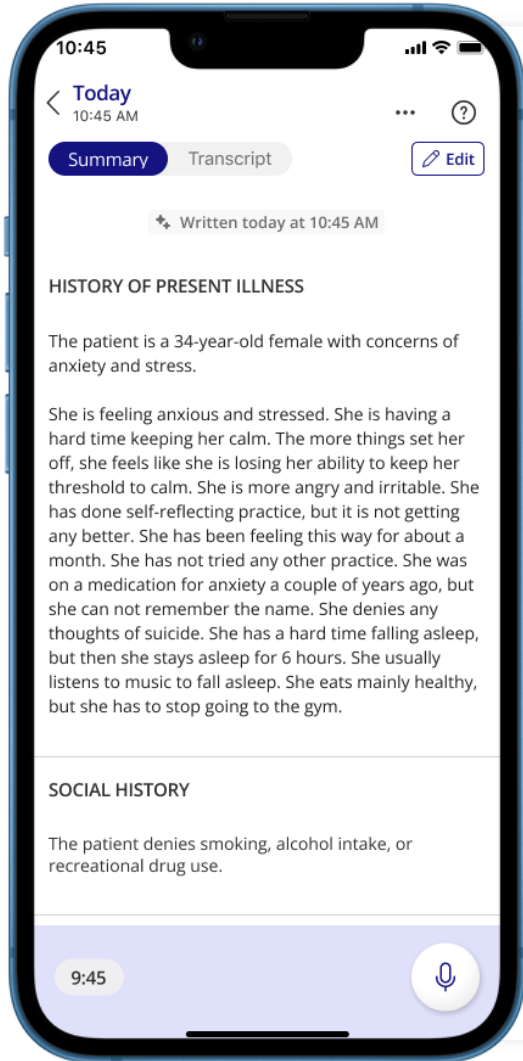- Physicians spend one-third of their EHR time on chart review

- Physicians average 100 inbox messages daily during working hours and another 50 each weekday evening

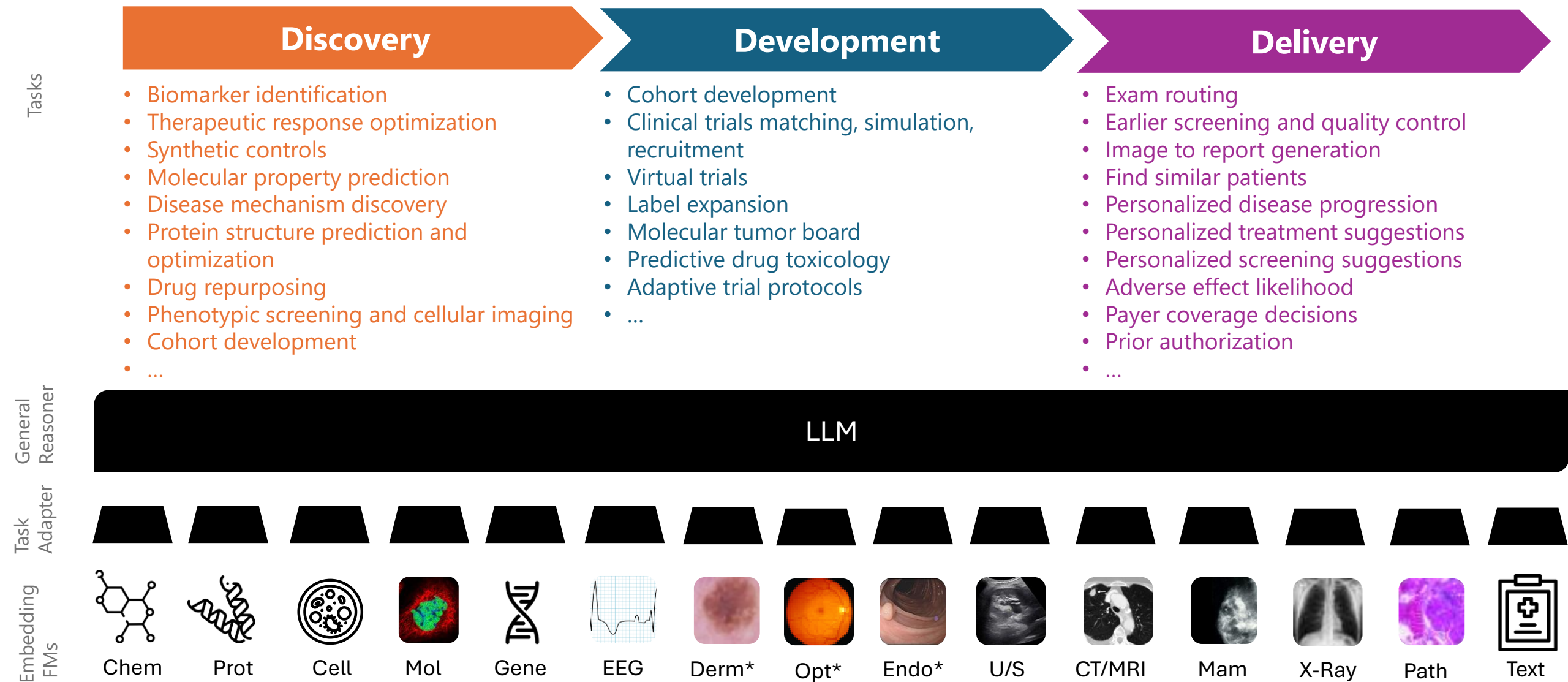The Electronic Health Record Inbox: Recommendations for Relief - PMC (nih.gov)

Physician Time Spent Using the Electronic Health Record During Outpatient Encounters: A Descriptive Study: Annals of Internal Medicine: Vol 172, No 3 (acpjournals.org)

Use case: Dragon Ambient eXperience (DAX) Notes and task automation

10:45

< **Today**
10:45 AM

Summary | Transcript | ✎ Edit

✦ Written today at 10:45 AM

HISTORY OF PRESENT ILLNESS

The patient is a 34-year-old female with concerns of anxiety and stress.

She is feeling anxious and stressed. She is having a hard time keeping her calm. The more things set her off, she feels like she is losing her ability to keep her threshold to calm. She is more angry and irritable. She has done self-reflecting practice, but it is not getting any better. She has been feeling this way for about a month. She has not tried any other practice. She was on a medication for anxiety a couple of years ago, but she can not remember the name. She denies any thoughts of suicide. She has a hard time falling asleep, but then she stays asleep for 6 hours. She usually listens to music to fall asleep. She eats mainly healthy, but she has to stop going to the gym.

SOCIAL HISTORY

The patient denies smoking, alcohol intake, or recreational drug use.

9:45

AI notes

| Trained on millions of encounters | Standard format or customize your own | Integrated with Dragon Medical One or embedded in your EHR | Quick note turnaround | Scalable |

Automatically document care and streamline clinical tasks.

| After visit summaries | Referral letters | Encounter summaries | Evidence summaries | And more... |

Use cases: Multimodal models will accelerate HLS discovery and delivery applications

Tasks

Discovery

- Biomarker identification
- Therapeutic response optimization
- Synthetic controls
- Molecular property prediction
- Disease mechanism discovery
- Protein structure prediction and optimization
- Drug repurposing
- Phenotypic screening and cellular imaging
- Cohort development
- ...

Development

- Cohort development
- Clinical trials matching, simulation, recruitment
- Virtual trials
- Label expansion
- Molecular tumor board
- Predictive drug toxicology
- Adaptive trial protocols
- ...

Delivery

- Exam routing
- Earlier screening and quality control
- Image to report generation
- Find similar patients
- Personalized disease progression
- Personalized treatment suggestions
- Personalized screening suggestions
- Adverse effect likelihood
- Payer coverage decisions
- Prior authorization
- ...

General Reasoner

LLM

Task Adapter

Embedding FMs

Chem | Prot | Cell | Mol | Gene | EEG | Derm* | Opt* | Endo* | U/S | CT/MRI | Mam | X-Ray | Path | Text

Health playground
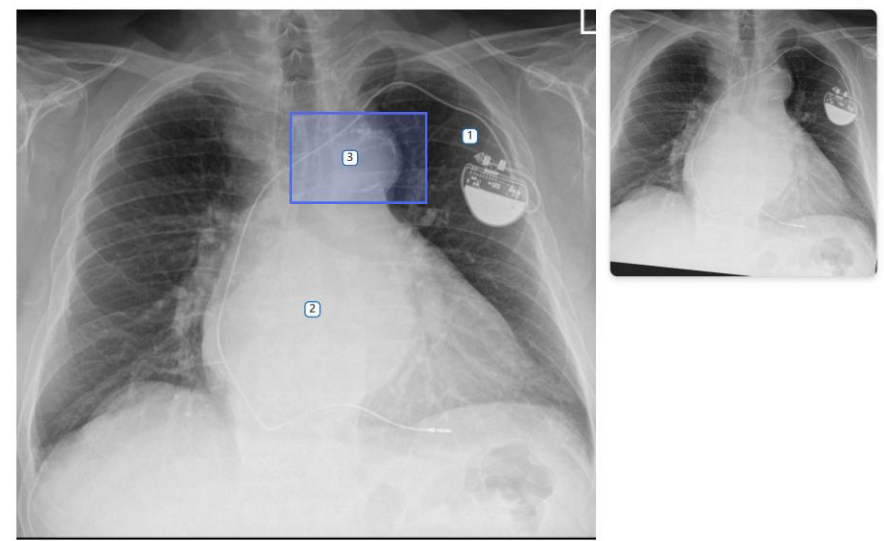
</> View code

Setup   Hide setup

Deployment *

CXRReportGen ▾

Clear chat

Indication: Cough, Smoker prior. Dyspnea, unspecified. Fatigue. Shortness of breath.

Findings:

1.) Left chest wall single lead pacemaker in UNCHANGED position. 2.) STABLE Cardiomegaly. 3.) **Aortic arch atherosclerotic calcifications IS UNCHANGED.** Lungs are clear. No effusions or pneumothorax.

Type your prompt

0/8000 tokens to be sent

Use case: Predict and Prioritize 3D Imaging based on 2D X-Ray

- Model is training using multimodal inputs - Notes, 2D and 3D Images
- Enables differential diagnosis on 2D X-Ray images at low cost and rapid TAT
- Need for expensive CT and MRI 3D images can be prioritized based on 2D image with correlation to 3D image findings

Delivering meaningful outcomes with
Microsoft Cloud for Healthcare
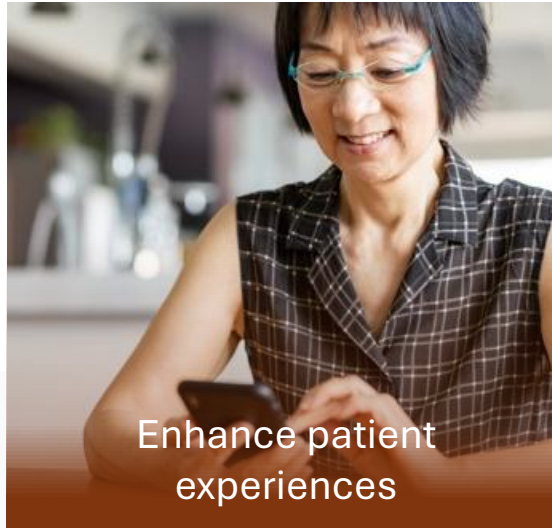
Enhance patient experiences

50% reduction patient wait time[1]

30% improvement in customer experience while helping protect patient privacy[2]

Empower health team collaboration

17.7% increase in improved time to decision through better collaboration[3]

25% reduction in errors due to misinformation[4]

Enhance clinician experiences

70% reduction in feelings of burnout and fatigue[5]

50% less time spent on documentation—7 minutes saved per encounter[6]

Lower cost of care and improve quality

45% better than the national average in O/E mortality score[7]

80% reduction in risk of delayed diagnosis[8]

Improve security and compliance posture
Reduce data breach risks by **60%** [9]
Automate security updates and reduce management costs by **40%**[10]

Forbes

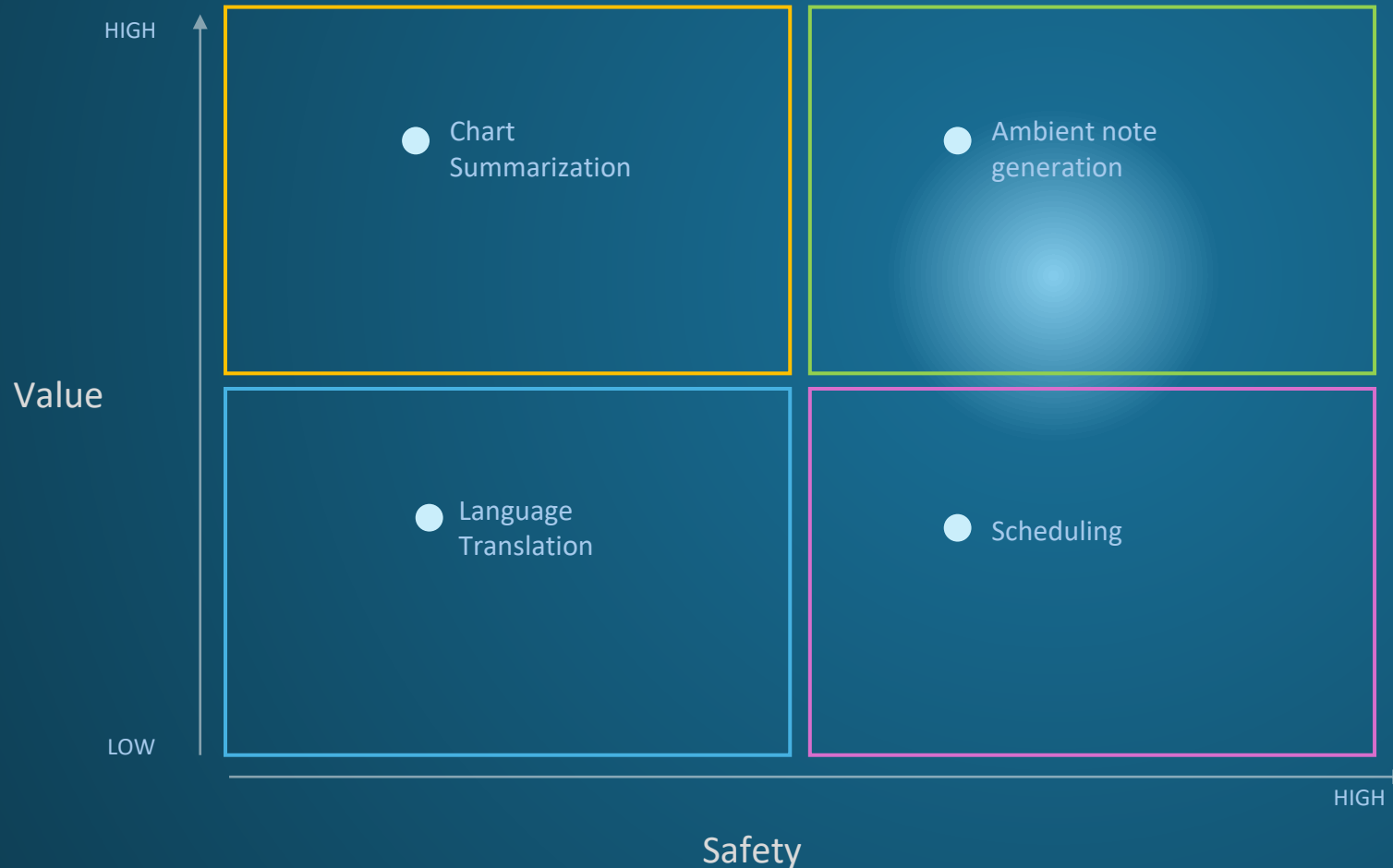## New Survey Shows That Up To 47% Of U.S. Healthcare Workers Plan To Leave Their Positions By 2025

The potential of AI is clear

$1T

Generative AI represents a meaningful new tool that can help **unlock $1 trillion in unrealized improvement potential** in healthcare

Value vs Safety

HIGH

Value

LOW

Chart Summarization

Ambient note generation

Language Translation

Scheduling

Safety

HIGH

How Safe is it?

Is there a human in the loop?

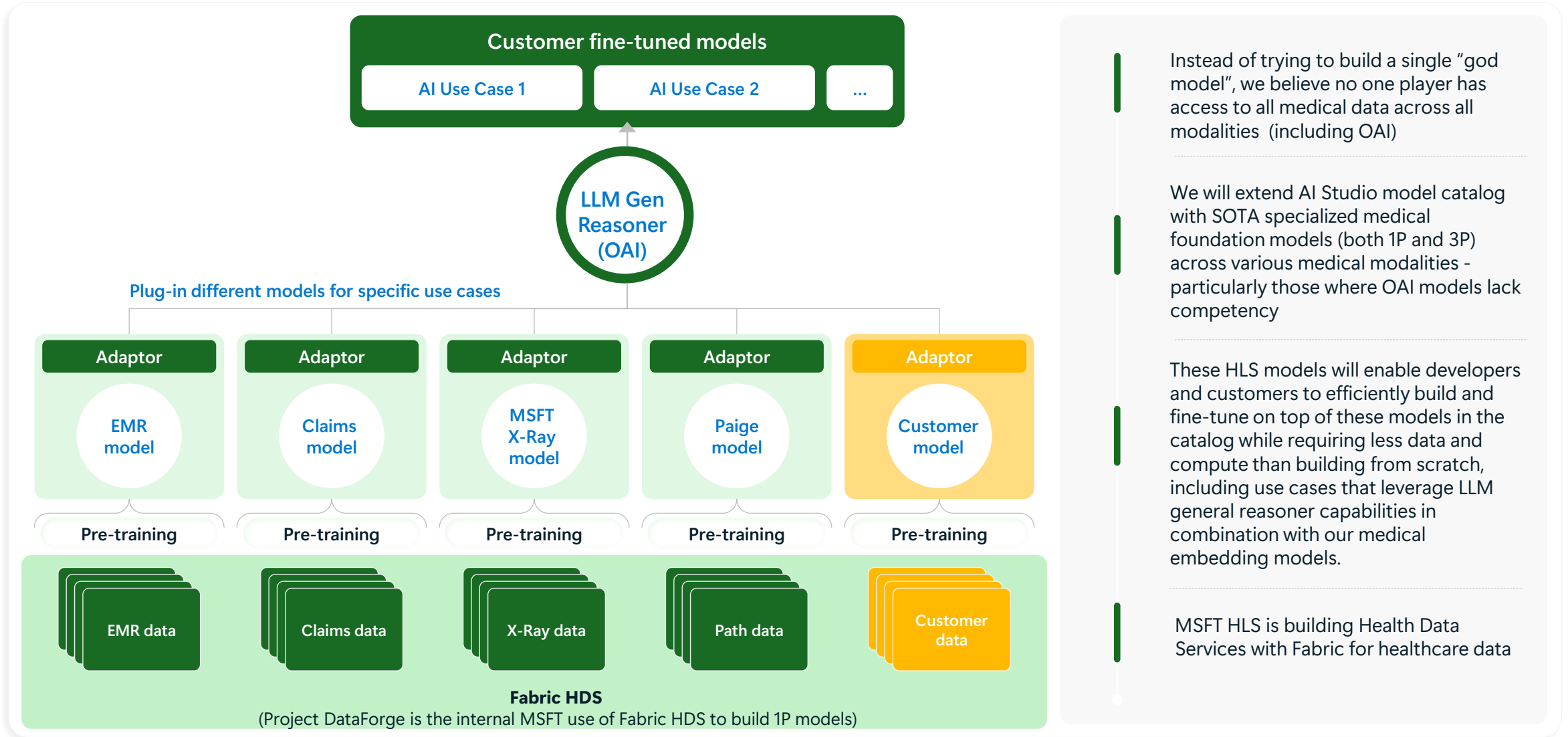Is the user qualified to interpret the results?

What happens in the worst case if the system fabricates information?

Can the system be well grounded in real data and provide citations?

Is there danger of procedural over-reliance?

Microsoft Health Gen AI Platform

Customer fine-tuned models

| AI Use Case 1 | AI Use Case 2 | ... |

LLM Gen Reasoner (OAI)

Plug-in different models for specific use cases

Adaptor	Adaptor	Adaptor	Adaptor	Adaptor
EMR model	Claims model	MSFT X-Ray model	Paige model	Customer model
Pre-training	Pre-training	Pre-training	Pre-training	Pre-training
EMR data	Claims data	X-Ray data	Path data	Customer data

Fabric HDS
(Project DataForge is the internal MSFT use of Fabric HDS to build 1P models)

Instead of trying to build a single "god model", we believe no one player has access to all medical data across all modalities (including OAI)

We will extend AI Studio model catalog with SOTA specialized medical foundation models (both 1P and 3P) across various medical modalities - particularly those where OAI models lack competency

These HLS models will enable developers and customers to efficiently build and fine-tune on top of these models in the catalog while requiring less data and compute than building from scratch, including use cases that leverage LLM general reasoner capabilities in combination with our medical embedding models.

MSFT HLS is building Health Data Services with Fabric for healthcare data

What now?

The AI train has left the station!

- You need to get on-board

- There is need, opportunity, potential, consensus and solutions

- New technologies require an updated architecture. We can't successfully build new on top of old!

- The foundation must be able to support the future.