

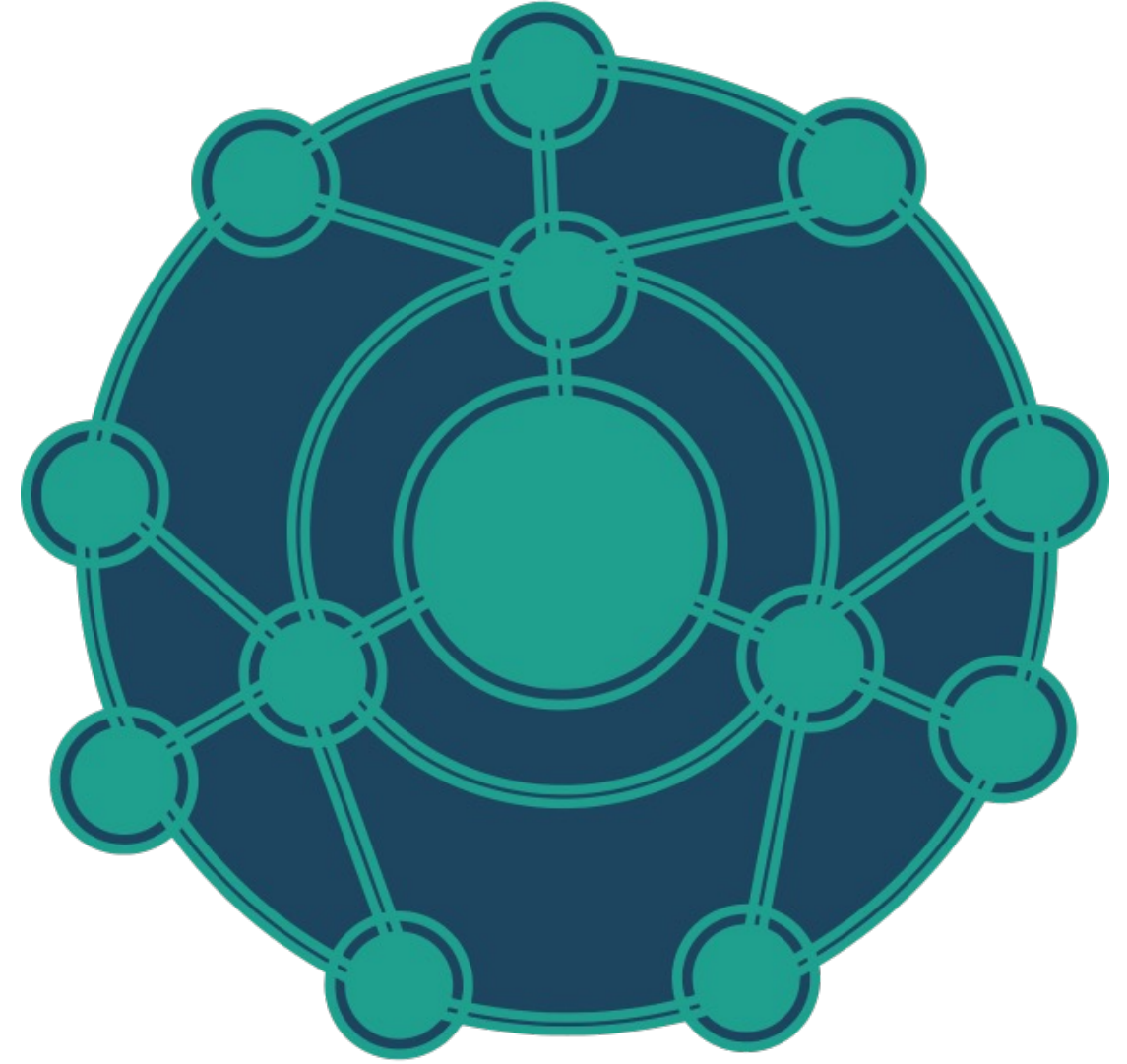


# ISPOR Issue Panel

Section 1 - LLMs in healthcare

Tim Reason

---





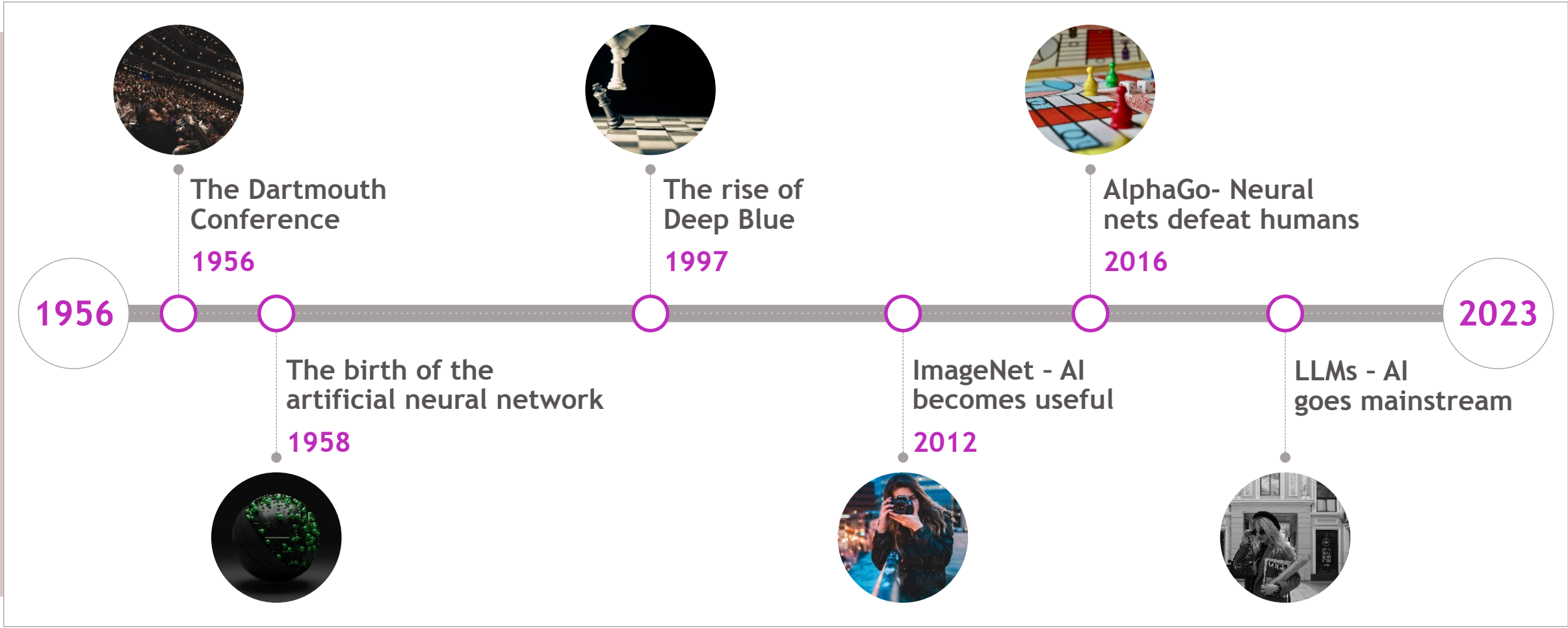
# Agenda



- 01 The AI timeline
- 02 AI, LLMs and key definitions
- 03 The AI debate
- 04 AI in healthcare
- 05 LLMs in healthcare
- 06 LLMs and the “ground truth” model
- 07 AI and HEOR - Can we afford not to?



# A brief history of AI



©2022 Proprietary and Confidential. All Rights Reserved.



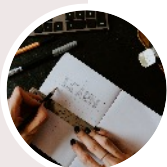
# Key definitions:

AI - What do all these words mean?



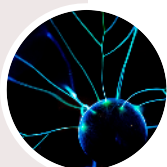
## Artificial intelligence

Branch of computer science concerned with the automation of intelligent behaviour



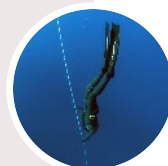
## Machine Learning

A subset of AI that focusses on machines/computers learning to improve at tasks autonomously with experience



## Neural Networks

Computational models inspired by and designed to mimic the functioning of the human brain



## Deep Learning

A multi-layered extension of neural networks that spans multi layers and is designed to mimic more complex behaviour



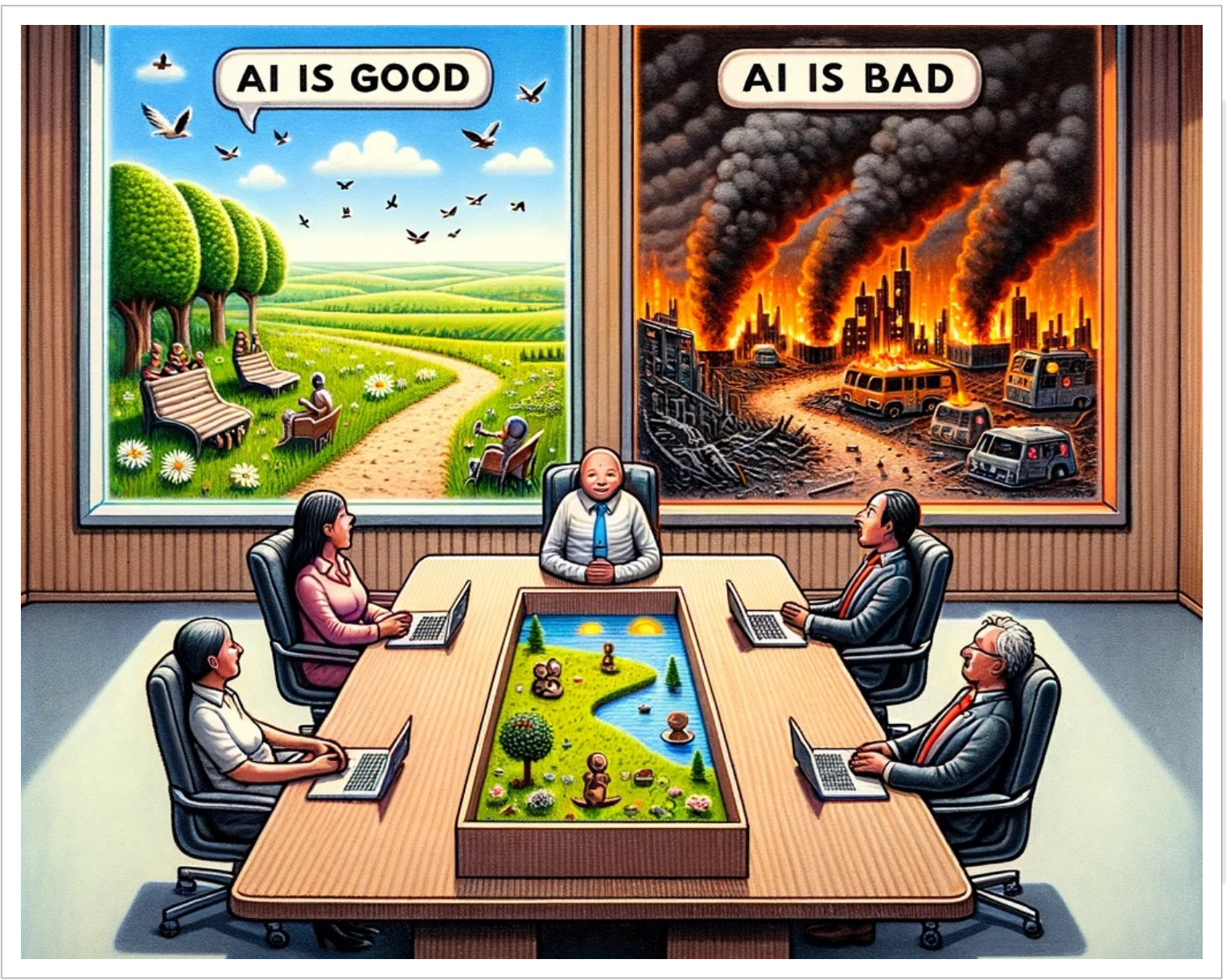
## Natural Language Processing

A field at the intersection of AI, language and computer science concerned with the processing of language by machines



## Large Language Models (LLMs)

A particular type of deep learning neural network trained on large corpuses of text



Where is AI debate right now?

---

What are people saying?

©2022 Proprietary and Confidential. All Rights Reserved.



# Traditional AI/deep learning

## The process



### Define the objective

- Image recognition
- Text summarising
- Playing games



### Collect LOTS of data

- Typically weeks/months of work
- DL models perform better with more training data
- Data also needs to be classified depending on the objective



### Train the model on the data

- Very expensive process
- Typically requires access to multiple GPUs
- Only really available to very large companies



### Test on unseen samples

- Test whether the model is doing what it's supposed to
- Iterate or abandon



### Deploy (if it works!)

- Deploy if model is working
- Handwriting recognition for post is a good example of a successful training/deployment
- You can only deploy in a setting related to what the model has been trained in



# LLMs

The new frontier - Models are pre-trained and general purpose!



Define the objective

- Summarising clinical trial reports
- Extracting key information from patient records
- Writing code for analytics



Create prompts

- Instead of collecting and training we only need to create prompts
- Prompts are “instructions” to the large language model
- You give the model input and ask for output
- “Summarise this clinical trial publication extracting the following key details...”



Hit the API

- Much cheaper than training
- One simple line of code to hit the API
- No knowledge of AI required!



Test on unseen samples

- Test whether the model is doing what it's supposed to
- Iterate or abandon



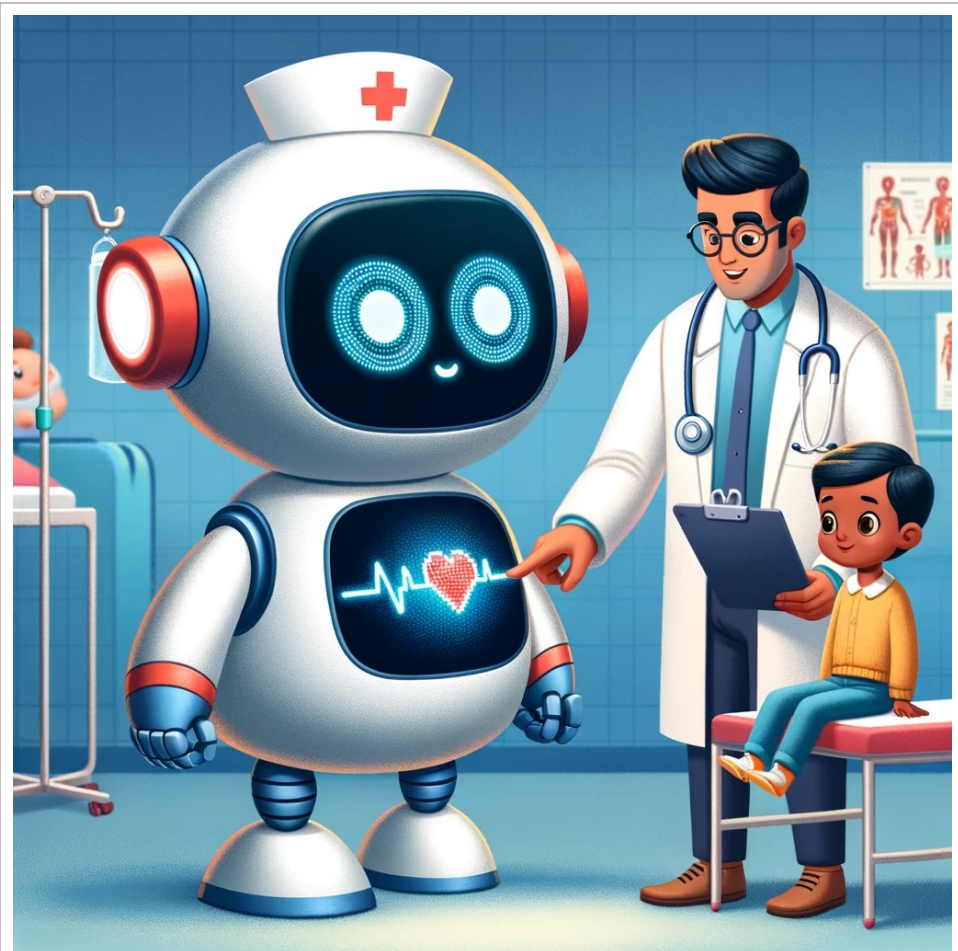
Deploy (if it works!)

- Deploy if model is working
- Models are deployed in a similar way
- You can fine tune models for your own purpose



# AI in healthcare

Traditional AI is already making waves in healthcare



©2022 Proprietary and Confidential. All Rights Reserved.



## Detection

AI powered smart stethoscopes can detect signs of heart failure correctly in 9/10 patients meaning GPs could use AI as an aid to detecting heart failure without always having to refer to secondary care



## Predicting disease progression

AI has been used to successfully predict whether people will develop wet AMD (AI outperformed 5/6 clinicians at this task) and predicted whether people with UC were about to have a flare (8/10 flares were predicted correctly)



## Diagnosis

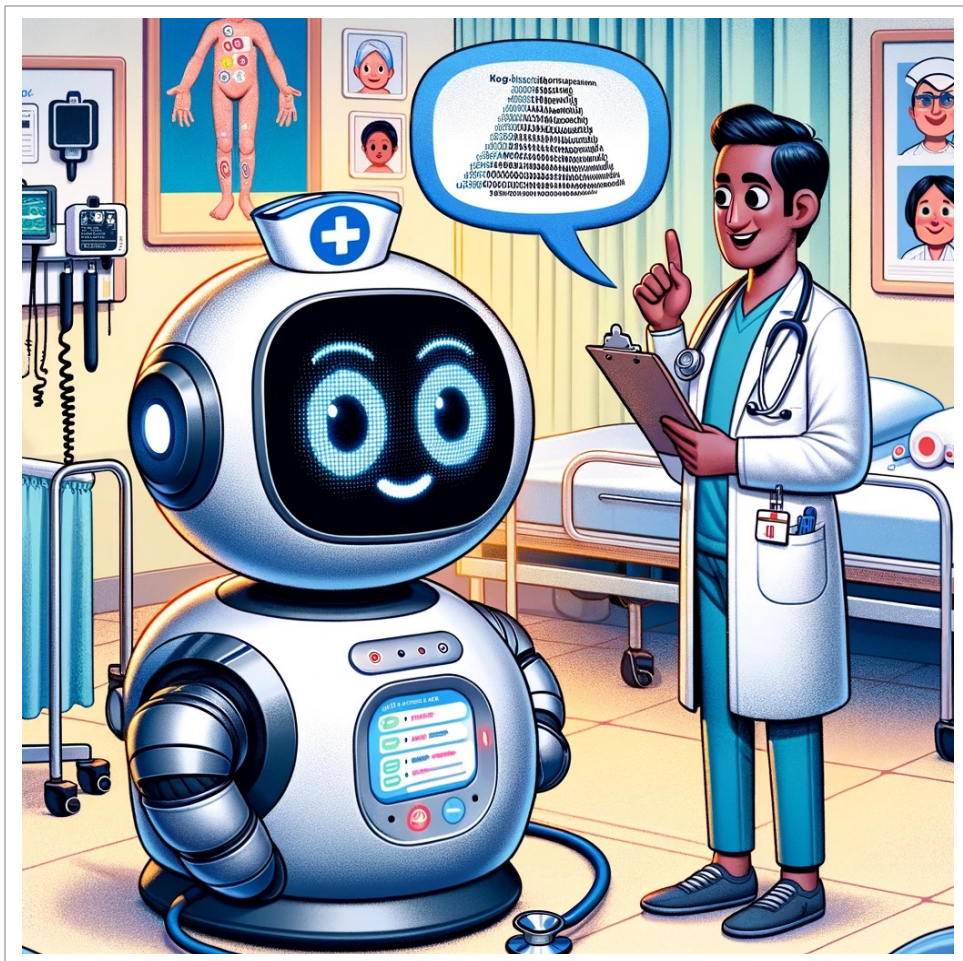
AI has been used successfully to diagnose whether nodules seen on lungs in a CT scan are dangerous (outperforming the Brock score used by clinicians)





# LLMs in healthcare

## How could LLMs revolutionise healthcare



©2022 Proprietary and Confidential. All Rights Reserved.



### Multi-disciplinary synthesis

LLMs could be used to synthesise information from scans, patient records, treatment guidelines and literature faster and at more scale than any clinical team giving clinicians a comprehensive summary for each patient - This level of care just isn't possible in the current care paradigm



### Diagnosis

LLMs (specifically Google MedPalm) have been studied by Singhal et al to test how LLMs can answer questions pertaining to patient care. The authors found that answers by MedPalm were close to clinicians though slightly inferior (by around 0.2-5% across different domains)



### Research

AI can be used to synthesise, summarise and extract from multiple information sources and medical literature at a far quicker rate and at greater scale than humans



# Could AI be the enabler?

## The “Ground Truth” HEOR model



AI analyses conducted for any disease area



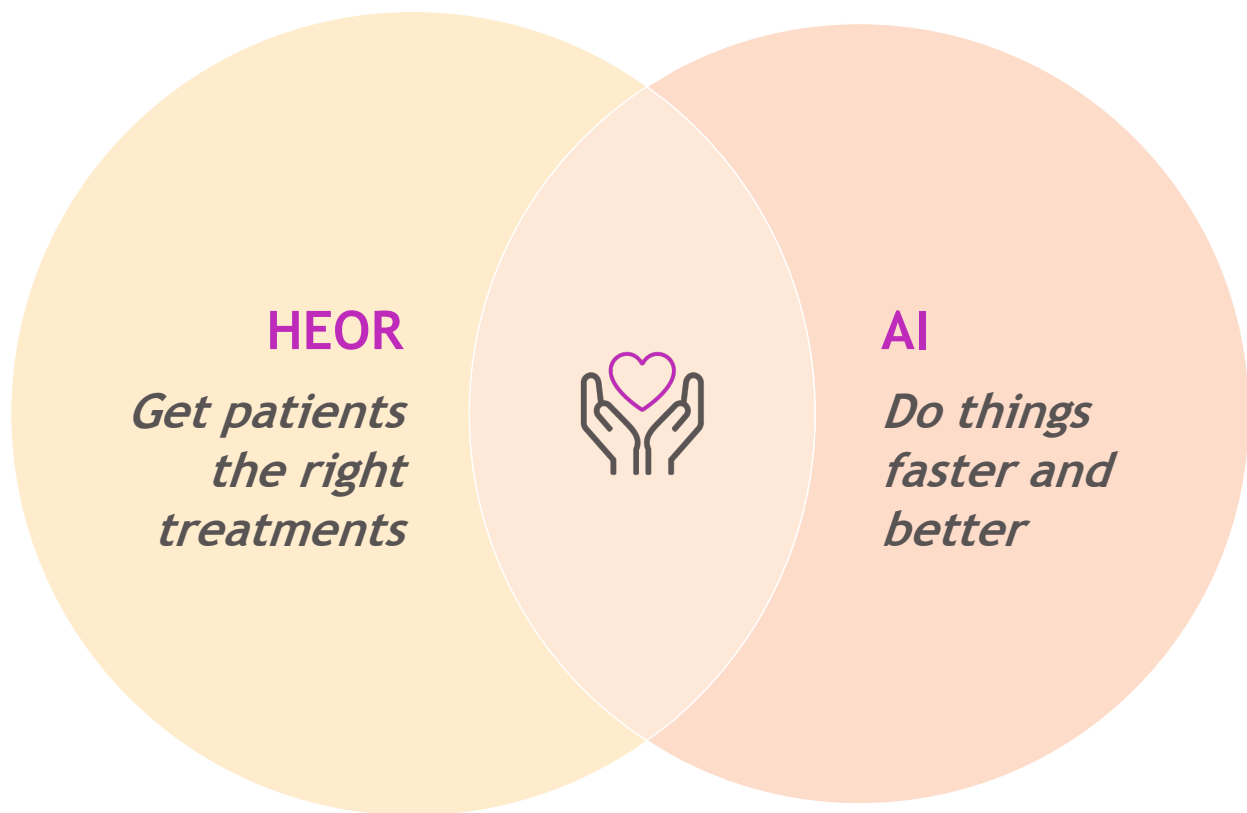
Accessed centrally through a web app by multiple stakeholders



AI generated lay summaries for patients and clinicians



# HEOR and AI - Can we afford not to?



HEOR optimises to get patients the right treatments, AI optimises for speed and accuracy - If AI can 100x speed of patient access to medicines can we really afford not to?



Page 10A The Daily Item — Sumter, S.C. Saturday, April 5, 1986



AP photo

**Elementary school teachers picket against use of calculators in grade school**  
The teachers feel if students use calculators too early, they won't learn math concepts

# Math teachers protest against calculator use

By JILL LAWRENCE

"My older kids don't pay any attention to an answer being absurd," he said. "Teachers are shv."

# Thank you for your time today!

[tim.reason@estima-sci.com](mailto:tim.reason@estima-sci.com)