



PHUSE US Connect 2024

**What Role Does
ChatGPT Have in
your Clinical
Programming?**

**potential role of AI bots and chat bots in the
creation of statistical analyses for regulatory review
advantages and limitations**

Abstract

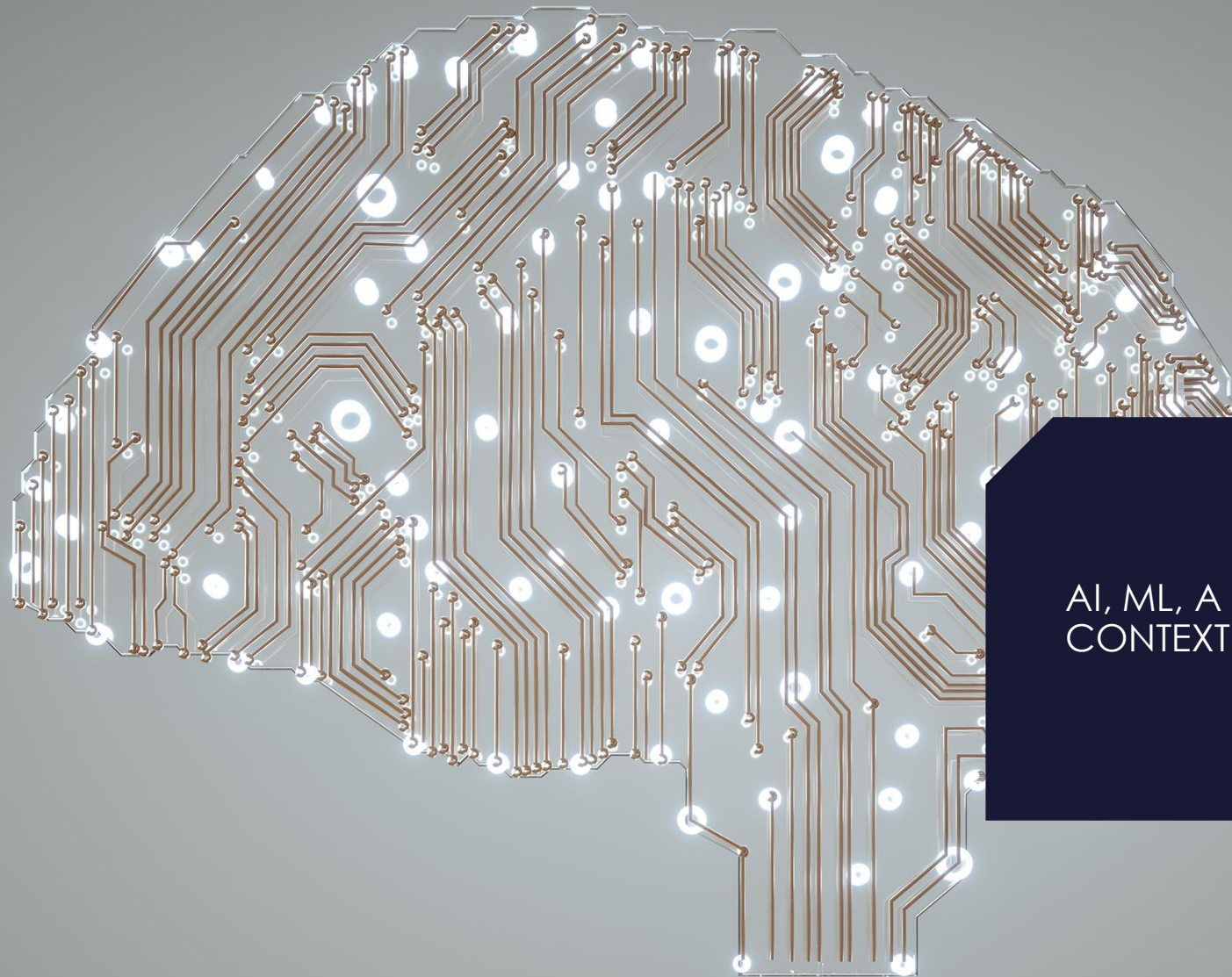
AGENDA

AI, ML, a brief history, and context for generative models

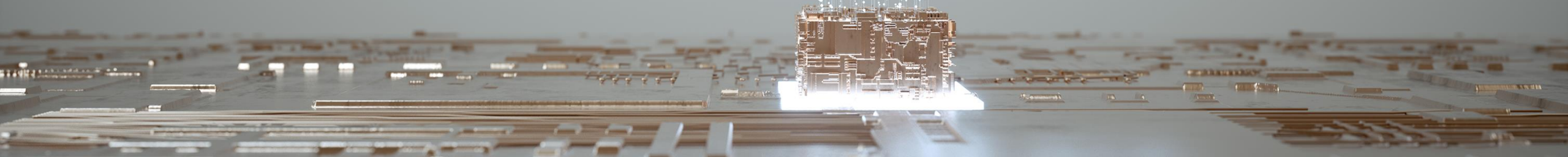
What is ChatGPT?

What are the opportunities and benefits?

Conclusions and next steps

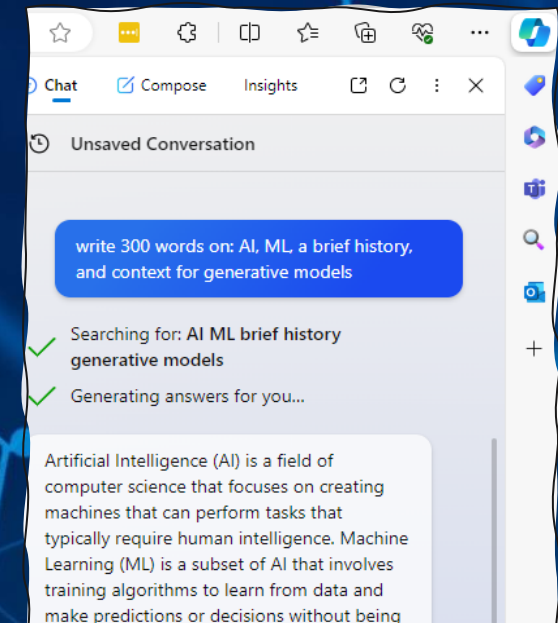
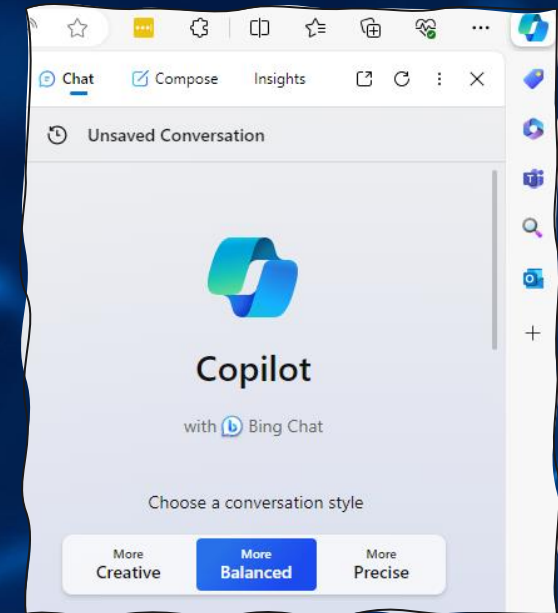
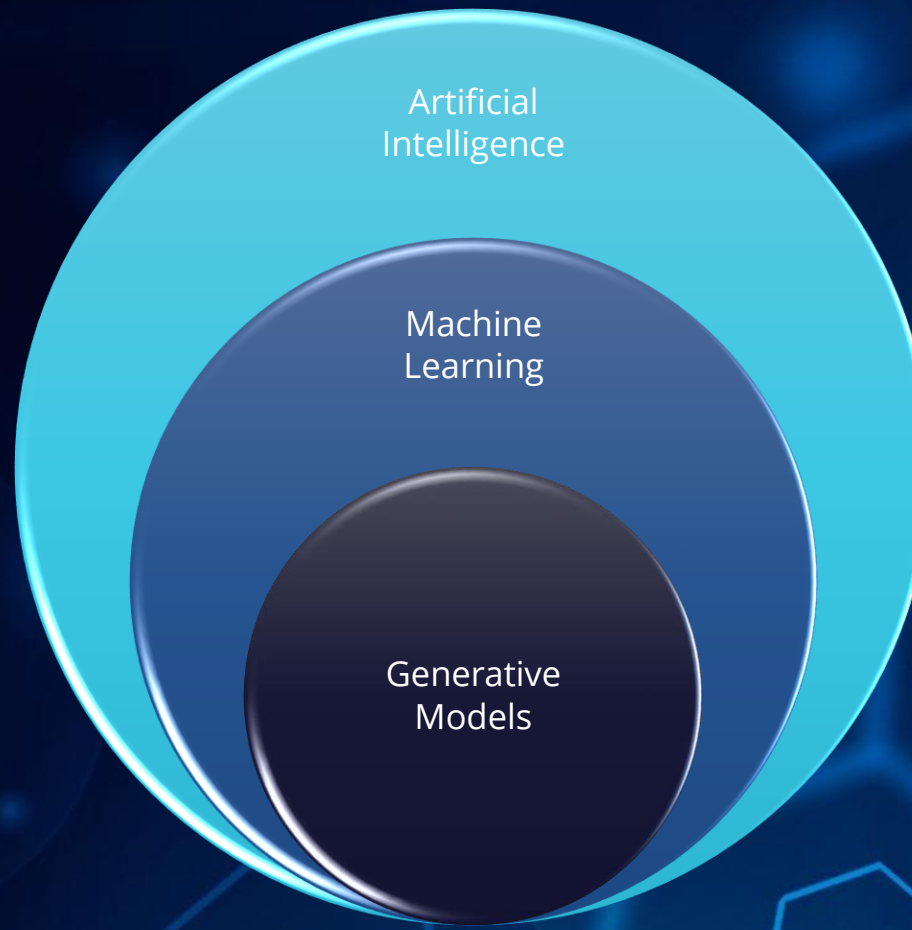


AI, ML, A BRIEF HISTORY, AND CONTEXT FOR GENERATIVE MODELS



AI, ML, A BRIEF HISTORY, AND CONTEXT FOR GENERATIVE MODELS

- Artificial Intelligence
 - Dates back to the early 20th century
 - Manual creation of (software) machines
- Machine Learning
 - Algorithms learn from data
 - Make predictions or decisions
 - No explicit programming



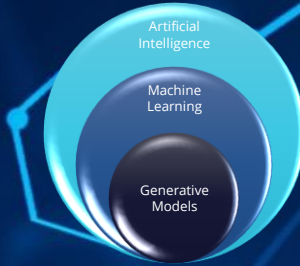
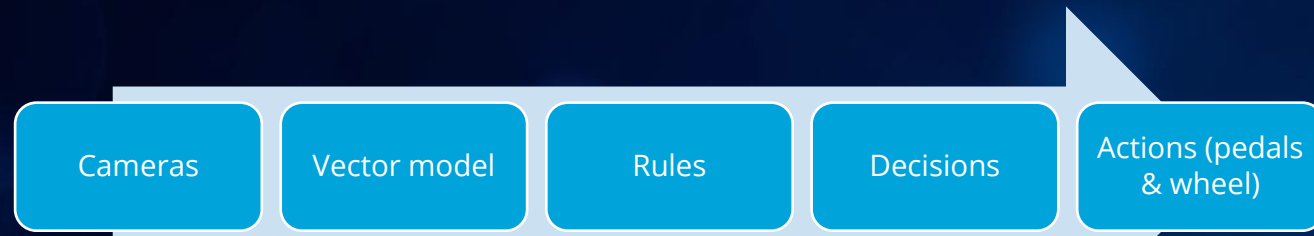
AI, ML, A BRIEF HISTORY, AND CONTEXT FOR GENERATIVE MODELS

- Artificial Intelligence

- Dates back to the early 20th century
- Manual creation of (software) machines

- Machine Learning

- Algorithms learn from data
- Make predictions or decisions
- No explicit programming



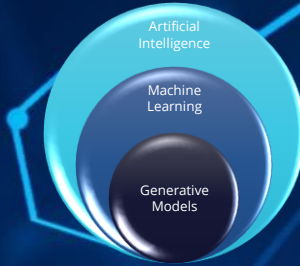
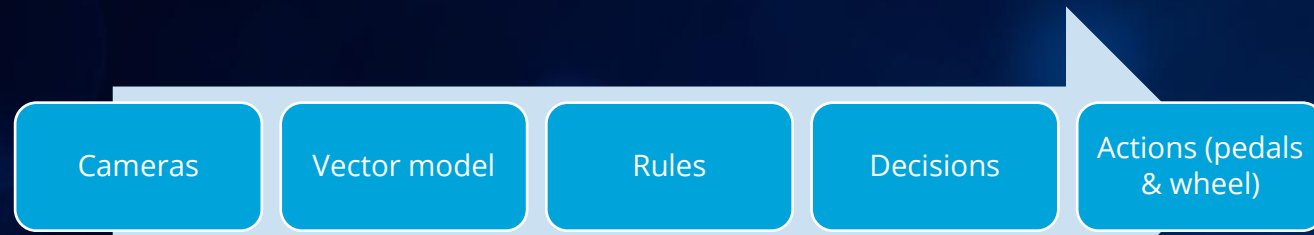
AI, ML, A BRIEF HISTORY, AND CONTEXT FOR GENERATIVE MODELS

- **Artificial Intelligence**

- Dates back to the early 20th century
- Manual creation of (software) machines

- **Machine Learning**

- Algorithms learn from data
- Make predictions or decisions
- No explicit programming



Training	Deployment
• Video: input • Pedals & wheel: output • Model trained to behave like the training materials	• Input: video • Model chooses output behaviour • Output: pedal and wheel movement

Outcomes

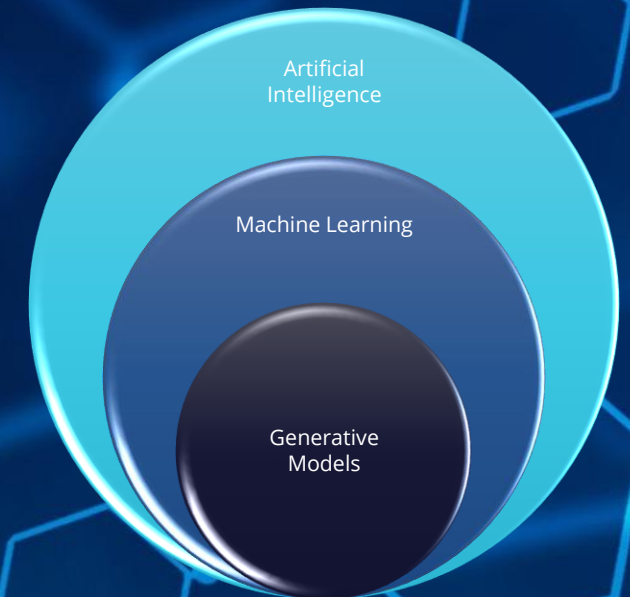
1. Model can cope with wide variety of situations; no reliance on pre-planning
2. Vehicles don't completely stop at stop lines!

AI, ML, A BRIEF HISTORY, AND CONTEXT FOR GENERATIVE MODELS

- Artificial Intelligence
 - Dates back to the early 20th century
 - Manual creation of (software) machines
- Machine Learning
 - Algorithms learn from data
 - Make predictions or decisions
 - No explicit programming

Manual Rules

Training Materials





WHAT IS CHATGPT?

WHAT IS CHATGPT?

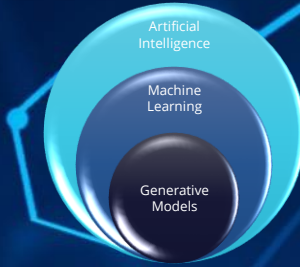
- Chat Generative Pre-trained Transformer
- AI model; specifically, ML
- Engages in conversational dialogue
- Chatbot developed by OpenAI (tech research company)
- Uses deep learning to generate human-like text based on prompts from users
- Predicts the next word in a given text, based on the patterns it has learned from a Massive amount of data during its training process
- Launched on November 30, 2022
- January 2023, fastest-growing consumer software application in history (100+ million users)
- Worldwide news headlines about ability to write essays, blog posts, and promotional copy
- Correct and accurate answers not guaranteed

Training Materials:

“The Whole Internet”

Ref: *What Kind of Mind Does ChatGPT Have?* Cal Newport

<https://www.newyorker.com/science/annals-of-artificial-intelligence/what-kind-of-mind-does-chatgpt-have>



CHATGPT LOOKS COMPELLING

what is phuse
where will phuse usa connect 2025 be held
is bethesda a good place to visit
how many days should i allow for touring bethesda in Maryland
show me a sas program for listing clinical demography data
show me an R program for listing clinical demography data
show me a sas program for listing clinical demography data sorted by subject id
Show me an excel formula for calculating compound interest
Write me a 1000 word paper on: What Role Does ChatGPT Have in your Clinical Programming?

A BANDWAGON THAT EVERYBODY IS ON...



OpenAI	Microsoft	Git Hub	Google	Amazon	Many more...
ChatGPT	Bing AI	Copilot	Bard	CodeWhisperer	Focused use cases...



WHAT ARE THE OPPORTUNITIES AND
BENEFITS?

WHAT ARE THE OPPORTUNITIES AND BENEFITS?

- 1) Natural language interfaces to Reference manuals and User Guides...
...for languages, editors and SCEs
...for regulations

BENEFIT: Faster, quicker access to information; better-informed teams



Only if you
trust the
results

- 2) Generate tables, listings and figures (TLFs) on behalf of statistical programmers

- Mock shells
- Test data
- TLFs ("results")

BENEFIT: Faster, more accurate results;
talented stats programmers move to data science



- Maybe only for 80/20 in early cases
- Peer review (QC) still required
- Model requires training on individual company's output standards

Ref: *LLMs Unleashed: A New Chapter in Clinical Programming?* Karma Dorje Tarap, BMS

<https://www.linkedin.com/pulse/llms-unleashed-new-chapter-clinical-programming-karma-dorje-tarap>

Mark Lambrecht (Senior Director, Health and Life Sciences, SAS)

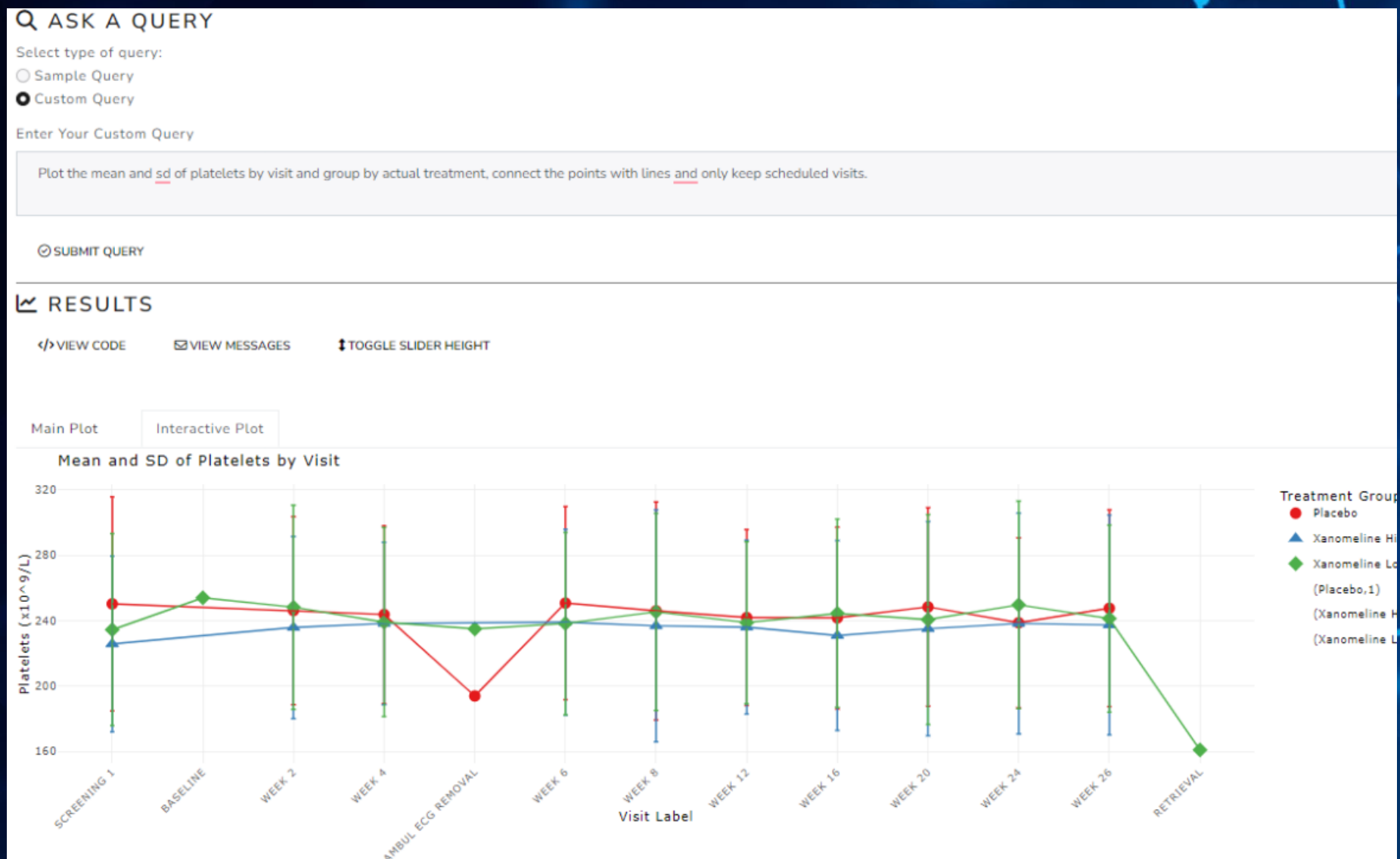
"SAS will soon offer a codepilot capability based on OpenAI LLMs as part of the SAS extension for Visual Studio Code in Q1 of 2024"

https://www.linkedin.com/feed/update/urn:li:ugcPost:7105414031355596800?commentUrn=urn%3Ali%3Acomment%3A%28ugcPost%3A7105414031355596800%2C7112920206452121600%29&dashCommentUrn=urn%3Ali%3Afsd_comment%3A%287112920206452121600%2Curn%3Ali%3AugcPost%3A7105414031355596800%29

LLMS UNLEASHED - KARMA DORJE TARAP, BMS

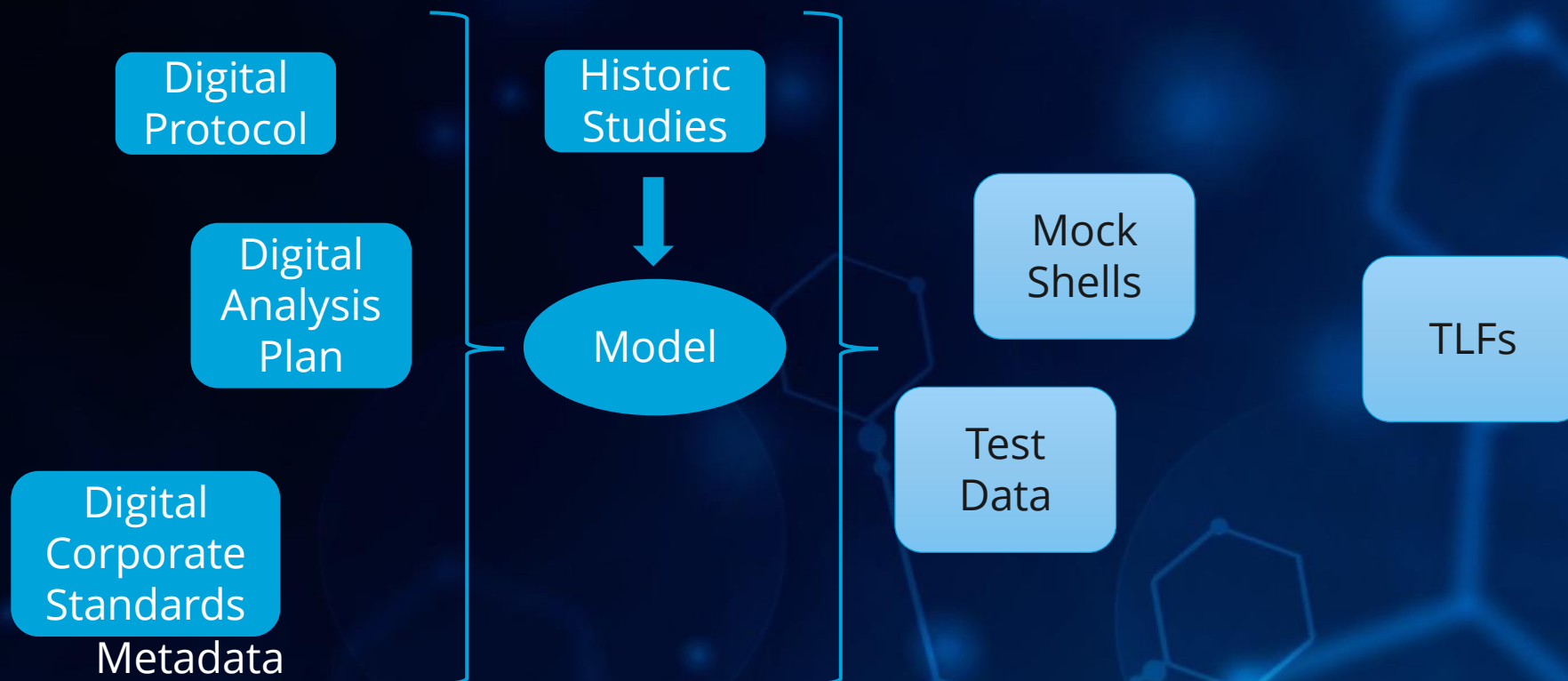
The tool harnesses the power of Large Language Models (LLMs) like GPT-4 to transform natural language queries into interactive, reproducible plots.

Lowens the technical barriers to deriving insights from clinical data.



OPPORTUNITY: MOCKS, TEST DATA, TLFS

Context and training are key



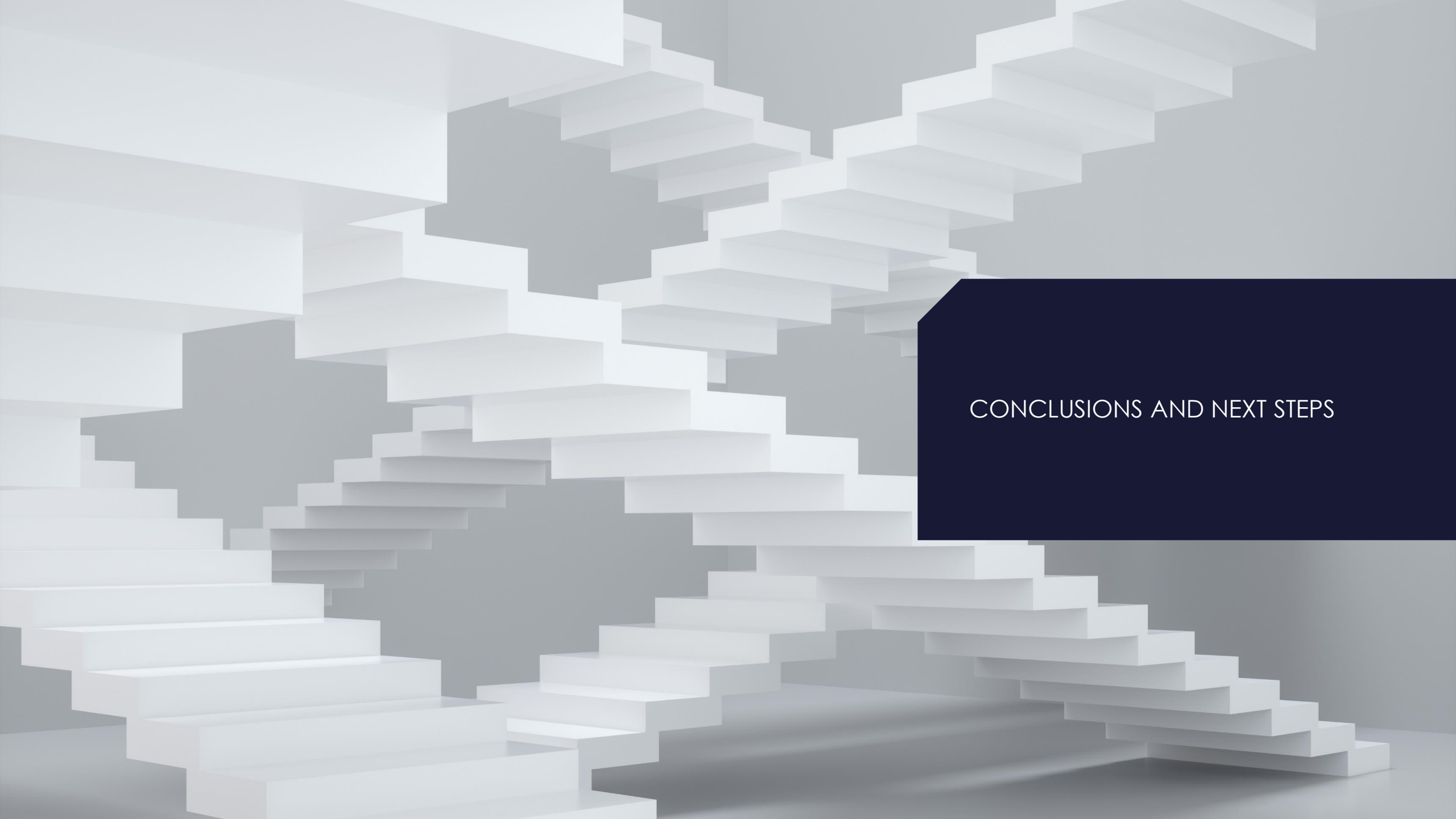
HOT AIR, OR COOKING ON GAS?

- Several industry initiatives around creation and use of digital documents
 - **TransCelerate, Digital Data Flow**

“This initiative aims to move the drug development process from a current state of manual study start-up asset creation (i.e., Case Report Forms, Procedure Manuals, Statistical Analysis Plans, and Schedule of Activities) to a future state of fully automated dynamic, study start-up readiness via an open-sourced, vendor agnostic technical solution that will reduce cycle times and improve data quality for sponsors, third-party providers, sites and regulators.”
 - **CDISC, Digital Data Flow**

“CDISC is collaborating with TransCelerate as a part of TransCelerate's Digital Data Flow Project to develop a Study Definition Reference Architecture that will serve as a standard model for the development of conformant study definition technologies, including the TransCelerate-developed Study Definitions Repository Reference Implementation”





CONCLUSIONS AND NEXT STEPS

CONCLUSIONS

- Significant number of people believe that a beneficial outcome is possible
- Combination of i) digital documents, along with ii) contextual metadata plus iii) copious historical trial data is required to facilitate a capable model
- Progress will be stepwise
- Capability of the modelling software and associated techniques are advancing at pace

Thank You!

Andy Ratcliffe
Product Manager, Aspire
Instem plc

+44 07714 719206
Andrew.Ratcliffe@instem.com
[linkedin.com/in/andrewratcliffeuk](https://www.linkedin.com/in/andrewratcliffeuk)

Andy Ratcliffe



Andrew Ratcliffe
Solution Owner, Aspire at Instem

