

The Impact of Large Language Models (LLMs) on Quality Management Systems in Biotech, Medtech and Pharma Industries

Elahe Haghsheno, Piper O'Keeffe, Luke Shi, Connor White, Hongyi Jiang



2023 PDA Asia Pacific Regulatory Conference





The Process



- Group of 5 Master of Biotechnology students from the University of Melbourne
- Completing an industry project for SeerPharma
- Minimal prior knowledge about QMS and Large Language Models
- Conducted 28 interviews with industry professionals from Australia, New Zealand and Singapore
- Produced a comprehensive report (62 pages) based on interviewee answers

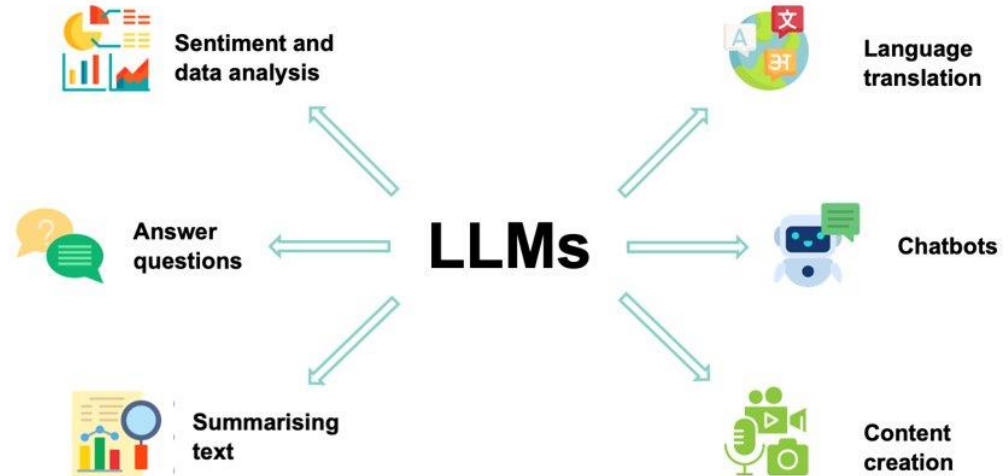
The Aim



The project will provide insights for all industry stakeholders into the broader implications of merging technology with industry practice in the areas of GMP, consulting, and quality systems. These insights will help raise awareness and potentially help with the optimisation of the current systems.

Understanding Large Language Models

- LLM = Large Language Model
- Conversational AI models that can understand and generate human-like text based on input prompts.
- Recognise patterns within language to generate relevant and coherent responses to the prompt they receive





Our Project

QMS

1

CAPA and Deviations

Elahe

2

Auditing and Change Management

Piper

3

Document Control and Training

Hongyi

4

Supplier & Material Qualification,
Calibration & Maintenance

Connor

5

Enterprise Risk and Customer Complaints

Luke

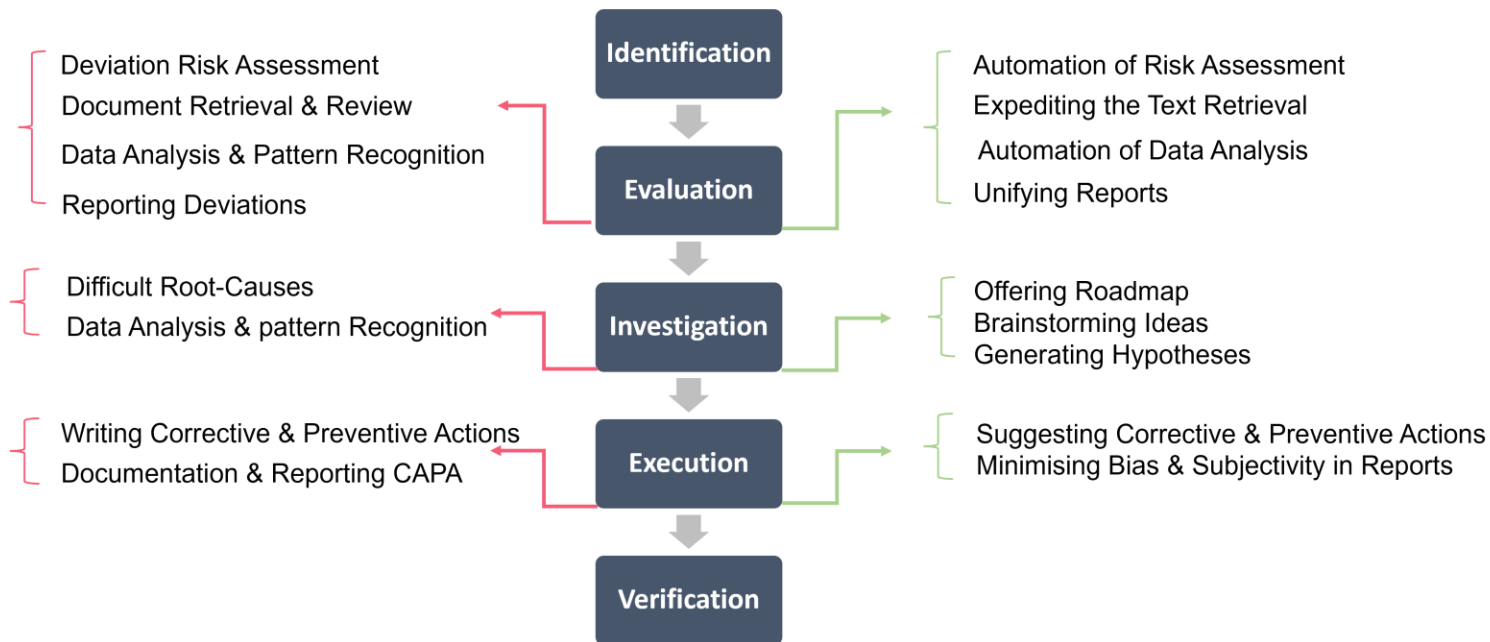


CAPA Challenges and Leveraging LLMs

Challenges

CAPA

Solutions using LLMs



LLMs in Auditing

01

Planning



- Risk-based audit schedules
- Auditor education

02

Desktop Audit



- Deviation trend analysis
- Past Performance

03

On-Site Audit



- Action checklist
- Interview questions

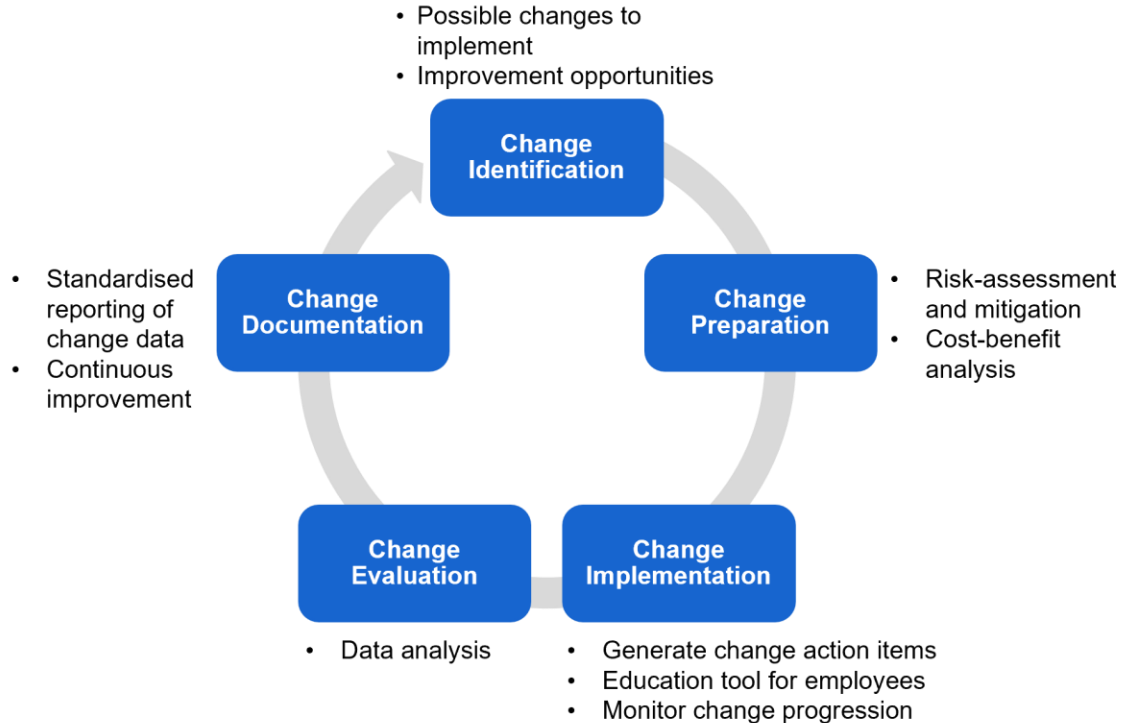
04

Evaluation & Reporting



- Audit data analysis
- Audit report
- Remove human bias

LLMs in Change Management



Hello, I've heard there's been an introduction of an LLM in our company. Can you provide more information about it?



Certainly, I'd be happy to provide details. Could you specify which aspects of the introduction you're interested in? We've recently rolled out several changes in this area.



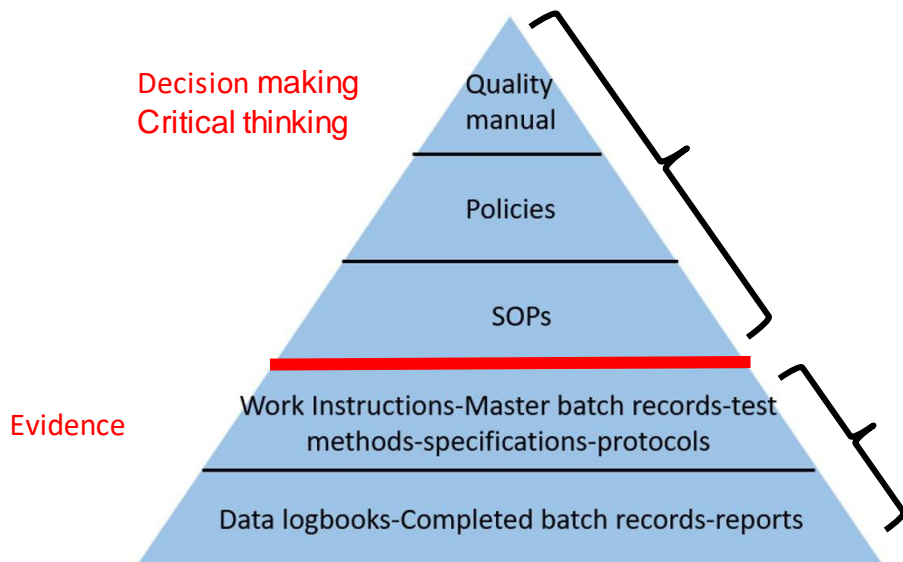
I'm particularly interested in the introduction of the LLM for auditing and change management. What can you tell me about this new development?



The introduction of the LLM for auditing and change management is a significant upgrade. This LLM is designed to improve the efficiency and accuracy of our auditing processes. It can analyse and extract insights from vast amounts of textual data, making it an invaluable tool for ensuring compliance and managing change effectively.



LLMs in Document Control



Current Limitation: General distrust in decision making

- LLM: Provide data support (e.g. CAPA result)
- Strengthen connection between high and low level documents

Current Limitation: Compiling document is repetitive & time consuming

- LLM: Auto-filling required information utilizing language process capabilities
- Closed source LLM to interact with Industrial software
- Role of employee transferred into “supervisor” of LLM

LLMs in Training

Knowledge Based Training



Tailored training method



Personalised content push for learning resources



LLM as chatbots: provide answer for trainee



Exams -> competency profile

Skill Based Training



- **Limitations:** Difficult to manage on paper & limited use of LLM
- **Opportunities:** Integration with AR & VR technology



Significance of Effective Prompting

A poorly written prompt:

H

Write a commercialization report about aerosol delivery of siRNA particles.

A well written prompt:

H

Write a 2500-word commercialization report about aerosol delivery of siRNA particles, relevant information is in company database (XXXXXX). The report should cover these subsections: the invention; intellectual property; financing; regulation; manufacturing and market. Proper diagram should be used. Provide rationale and reasoning for your decisions and recommendations and ensure that these are grounded in reality and where possible, backed by evidence.

- Know the Model's Capabilities and Limitations.
- Ask Focused, Concise Questions.
- Illustrate Intent with Examples.
- Specify Expected Response Format.



LLMs in Supplier and Material Qualification

Challenges

Lack of efficiency comparing suppliers

Limited ongoing analysis of supplier provided data

Difficulties with supply chain management


Solutions


Set up matrices to automatically provide a risk rating based on supplier data

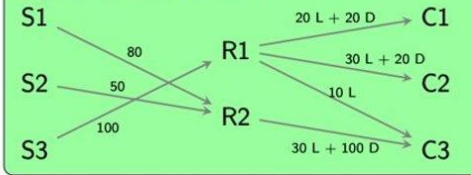
Train LLM to track supplier performance over time and alert any problems

Extract relevant information from supply chain data for specific suppliers and present reports

LLMs in Supply Change Management


 Show me the shipping plan.


 Here is the plan. You can hover your mouse on each node or edge to see more details.



```

    graph LR
      S1 -- 80 --> R1
      S2 -- 50 --> R1
      S2 -- 100 --> R2
      S3 -- 100 --> R2
      R1 -- "20 L + 20 D" --> C1
      R1 -- "30 L + 20 D" --> C2
      R2 -- "10 L" --> C2
      R2 -- "30 L + 100 D" --> C3
  
```

 Is it possible for Roastery 1 to be exclusively used by Cafe 2?

 If we exclusively transport beans from Roastery 1 to Cafe 2, the cost would amount to \$2570, representing an approximate increase of $\frac{2570-2470}{2470} \approx 4\%$ compared to the current plan, which costs \$2470. Would you like to implement this change for future planning purposes?

Type your question here...

- Optiguide by Microsoft
- Trained on supply chain scenarios that are continually updated
- Ask "what if" question
- Outputs both:
 - Textual explanation
 - Visualisations of optimisation solution

Optiguide Interface (Li et al., 2023).



LLMs in Calibration and Maintenance

Challenges

Suboptimal calibration and maintenance schedules

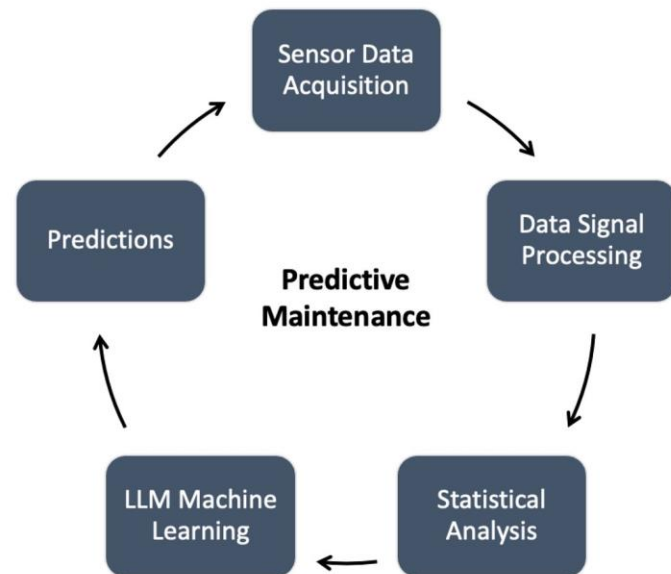
Current systems are reactive to problems

Solutions

- Create optimised schedules based on all previous calibrations and maintenance

- Assist in the integration of predictive maintenance strategies
- Analysis of multivariate time series sensor data

Sensor Integrated LLM for Predictive Maintenance





Customer Complaints and Chatbots

- **24/7 availability:** Chatbots can provide fast and efficient customer service around the clock, eliminating wait times and offering instant support ¹.

- **Personalized experiences:** Chatbots can offer more personalized experiences, based on the data provided by the customers ¹.

- **Multilingual support:** Chatbots can deliver multilingual support, catering to customers from different regions and backgrounds ¹.

- **Consistent support:** Chatbots ensure more consistent support, following the same standards and protocols ¹.

- **Convenient self-service options:** Chatbots offer convenient self-service options, allowing customers to solve their issues without human intervention ¹.

- **Proactive customer service:** Chatbots provide proactive customer service, anticipating customer needs and offering solutions ¹.

- **Omnichannel support:** Chatbots deliver omnichannel support, integrating with different platforms and devices ¹.

Valid benefits but more specifically...

For Customer Complaints, uses include:

- Customer data collection
- Recording and categorising complaints
- Guide customer through the process of filing complaints
- Tailored customer experience

Limitations:

- Company needs (low # of complaints -> less use of chatbot)
- Potential confidentiality breaches



LLM Validation

Time = $\frac{\text{Genome Size} \times \text{Coverage}}{\text{Sequencing Rate}}$

Time $\approx \frac{33 \text{ gigabases}}{860,160 \text{ bases/second}}$

Time $\approx 38.36 \text{ seconds}$

• Desired coverage: 10x

Actual Answer:

Time = 38,365 seconds or 10.6 hours

Correct Process -> Incorrect answer

- Highlights that AI technology isn't perfect
- Need for:
 - Standards (for all uses)
 - Minimum thresholds (for specific uses)
 - Performance assessment over time

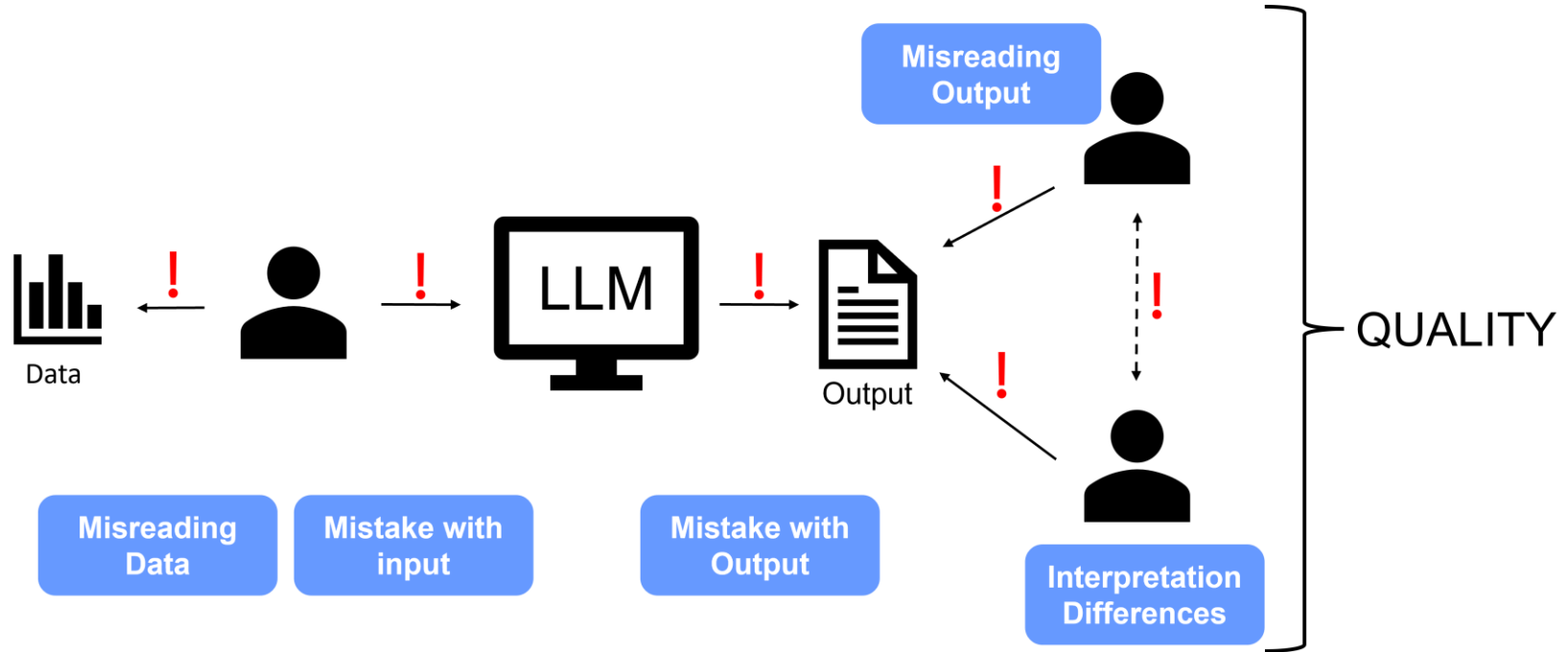
There are **one billion base pairs in a gigabase (Gb)**. The prefix "giga-" denotes a factor of 10^9 , and in the context of genomics, it indicates a unit of one billion base pairs. Therefore, when we refer to a genome size or sequencing output in gigabases, it means that there are one billion base pairs present.

Mathematically, it can be expressed as:

1 gigabase (Gb) = 1 billion base pairs (bp) = 10^9 bp

the red kangaroo genome to a 10-fold coverage using a MinION with 2,048 nanopores.

Enterprise Risk – Data Integrity



Enterprise Risk – Cybersecurity

Open-source LLM

Accessible to public

Sensitive information may be leaked

Uses broad, generalised data

Potential hacking via online means

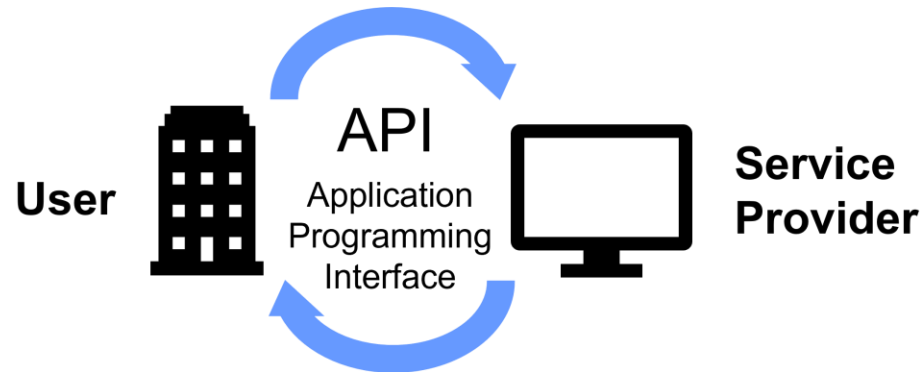
Closed-source LLM

Access limited to organisation

Sensitive information limited to company servers

Specific data aggregated all into one location.

Hacking via phishing/scams to employee emails





Risk Assessment

Risk Factor	Likelihood	Impact	Risk	Mitigation Strategy
Incorrect outputs	Possible	Significant	Med High	Review of output to ensure accuracy.
Incorrect inputs	Possible	Significant	Med High	Review of input to ensure accuracy. Employee training to perform this activity
Cybersecurity breaches	Possible	Severe	High	Employees are still subjected to scams, phishing, etc. Employee training to address these should already be given. Thus, two-factor authentication and more robust cyber protections such as VPNs, firewalls, etc., may be the more favourable mitigation strategy
Reduced Quality of Products	Possible	Significant	Med High	Data training of the tool before implementation should also include quality checks.
Regulatory non-compliance	Possible	Severe	High	Be closely aligned with the regulations of AI technology and have strategies to remain agile and adapt to any changes in regulations.
Poor data quality and integrity	Possible	Significant	Med High	First, consider data processing and cleaning Second, implement ongoing data monitoring and validation Third, make data free of biases, and aligned with regulatory standards.
Incorrect interpretation of outputs	Possible	Significant	Med High	Human review and verify the outputs. Use training modules



SWOT Analysis

STRENGTHS

- Optimised document systems; enhanced data organisation and accessibility
- Automate processes
- Data analysis for more insights
- Efficiency of tasks
- Delegate repetitive and time-consuming tasks to LLM can provide better resource allocation and use.

WEAKNESSES

- Potential for errors and inaccuracies
- Lack of nuanced critical thinking compared to humans
- Necessity of human oversight to mitigate risk
- Use of open-source LLMs may result in potential leakage of sensitive information.
- Performance and function relies on the quality of data.

OPPORTUNITIES

- Development of LLMs for internal use for tailored data analysis
- LLMs trained on company-specific data can provide improved function for the business
- Integrating LLMs across a wide range of applications

THREATS

- Non-compliance with existing and emerging regulations can lead to severe consequences
- Potential public health risk if internal LLMs don't meet standards
- Over-reliance on AI can reduce quality of work by employees
- Potential backlash by customers against the use of AI



Recommendations

- Development of an electronic data pool required before implementation
- Thorough assessment of current QMS systems to understand organisation position
- Conduct risk assessment before integrating LLMs into any system
- Remain agile and adapt to changing regulatory landscape and technology
- Cultural Change



Acknowledgements





Thank you



THE UNIVERSITY OF
MELBOURNE

Session V: Application and Regulation of Artificial Intelligence (AI) in the Pharmaceutical Industry



Micheal Schafferius
Sartorius
Moderator



Lisa Bennett
SeerPharma



Ben Stevens
GSK



Julie Winson
LBT Innovations



Piper O'Keeffe
University of Melbourne



Connor White
University of Melbourne