

PHUSE Day @ Boston Seaport

PwC Team

May 7 2024





Today's Agenda

1. Introductions
2. AI & CSV in Pharma
3. Contextual GenAI Demos
 - a. SAP Generator
 - b. Research Buddy

Automation & Validation in Pharma



So... What Is AI?



Artificial Intelligence

AI

is a class of computer programs that are able to learn from data and **simulate intelligent human behavior**.

Machine Learning

Machine Learning

subset of AI that enables machines to learn from existing data and improve upon that data **to make recommendations or predictions**.

Deep Learning

Deep Learning

a machine learning technique in which layers of neural networks are used to process data and **simulate human decision making**.

Generative AI

Generative AI

is a class of computer algorithms that try to **simulate content generation/basic reasoning capabilities in humans**.

Governance and Responsible AI is critical to driving long term Gen AI sustainability

Responsible AI at its core is simply good data science, governed by key guiding principles and made operational from strategy to execution.

Strategy

Data & AI Ethics

Consider the moral implication of uses of data and AI and codify them into your organization's values.

Policy & Regulation

Anticipate and understand key public policy and regulatory trends to align compliance processes.



Control

Governance

Enable oversight of systems across the three lines of defense.

Compliance

Comply with regulation, organizational policies, and industry standards.

Risk Management

Expand transitional risk detection and mitigation practices to address risks and harms unique to AI.



Responsible Practices

Interpretability & Explainability

Enable transparent model decision-making.

Sustainability

Reduce negative environmental impact.

Robustness

Enable high performing and reliable systems.

Bias & Fairness

Define and measure fairness and test systems against standards.



Security

Enhance the cybersecurity of systems.

Privacy

Develop systems that preserve data privacy.

Safety

Design and test systems to prevent physical harm.

Core Practices

Use Cases & Problem Formulation

Identify the concrete problem you are solving for and whether it warrants an AI/ML solution.

Standards

Follow industry standards and leading practices.

Validation

Evaluate model performance and continue to iterate on design and development to improve metrics.

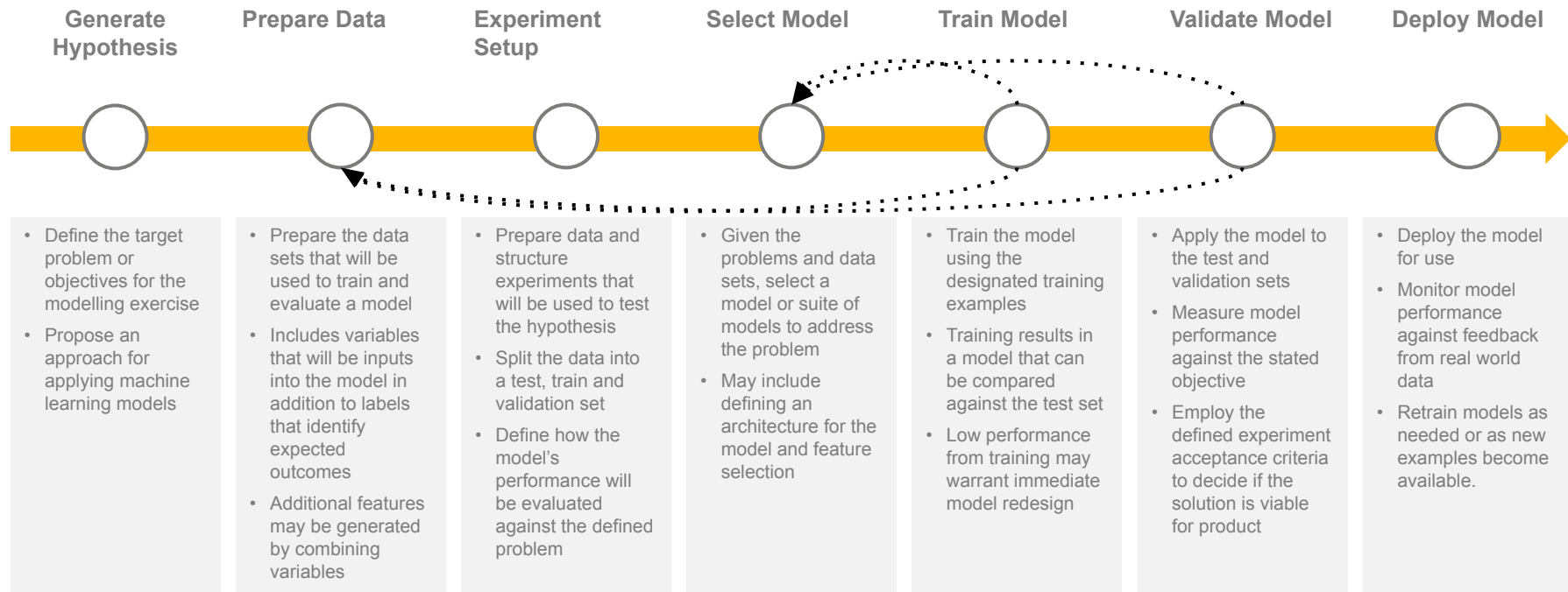
Monitoring

Implement continuous monitoring to identify drift and risks.



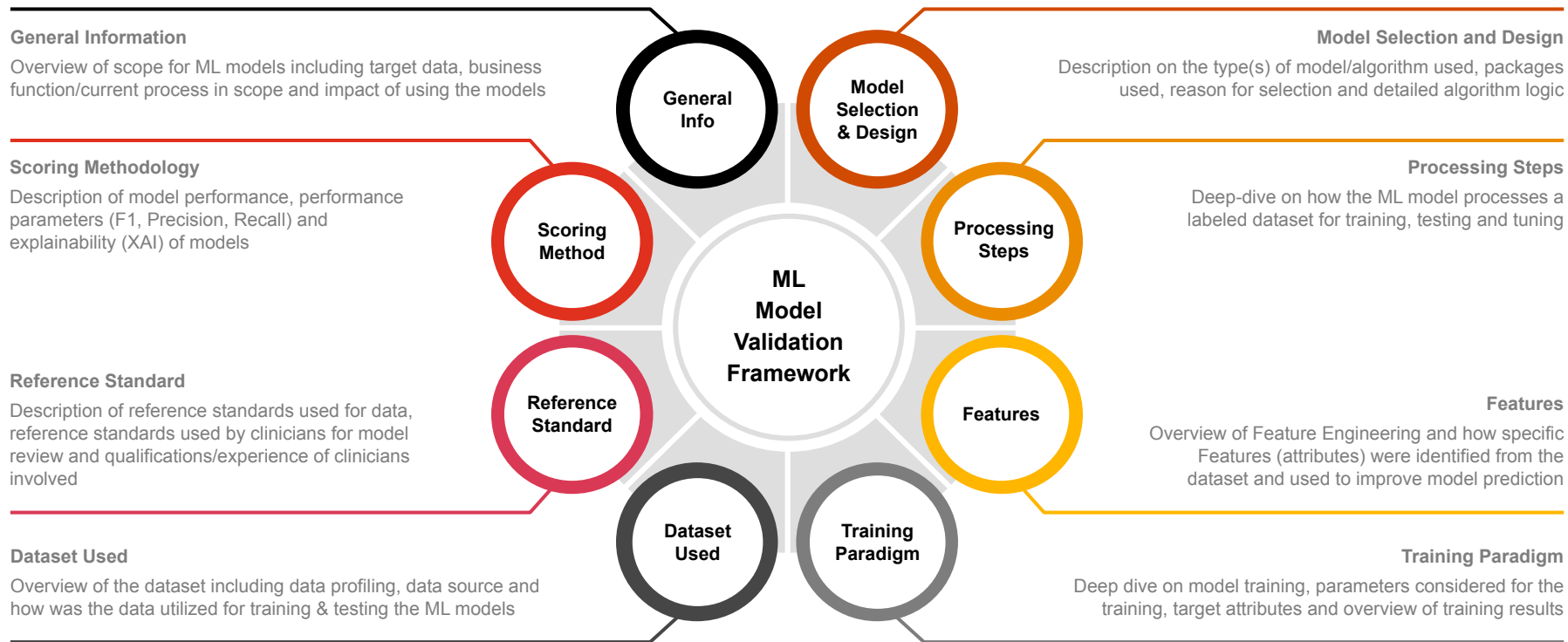
The AI Development Lifecycle is Highly Iterative

The AI development life cycle involves rapid iterations over data set preparation, model selection, testing, and occasionally revisiting objectives as data and algorithmic barriers are identified



ML Model Validation Framework

PwC has successfully deployed GCP compliant ML-based solutions that have been audited by both the FDA and MHRA



How to think differently about GenAI Model Validation

Validating a GenAI model requires additional considerations compared to traditional ML models...

Evaluation Metrics

- **Different metrics from ML** (e.g., ROUGE)
- Centered on the **quality, novelty, coherence, and human perception** of generated data
- Human evaluation plays a crucial role, as experts or human judges often assess the quality and characteristics of the **generated data based on subjective criteria**

Interpretability

- Focuses on **assessing the coherence and understandability** of the generated content
- **Rule-based constraints or semantic techniques** may be leveraged to support the validation process

Ethics

- Emphasis on the **responsible generation and use of novel content**, fairness in generation, and potential impact on individuals and society
- Techniques like **fairness metrics, bias detection, or adversarial testing** can be employed to mitigate biases and ensure fairness

User Experience

- Focuses on assessing user perception, interaction, and customization of the generated content to **create TRUST in the content**
- Involves iterative cycles of generating content, **collecting user feedback, and refining the models** based on the feedback received

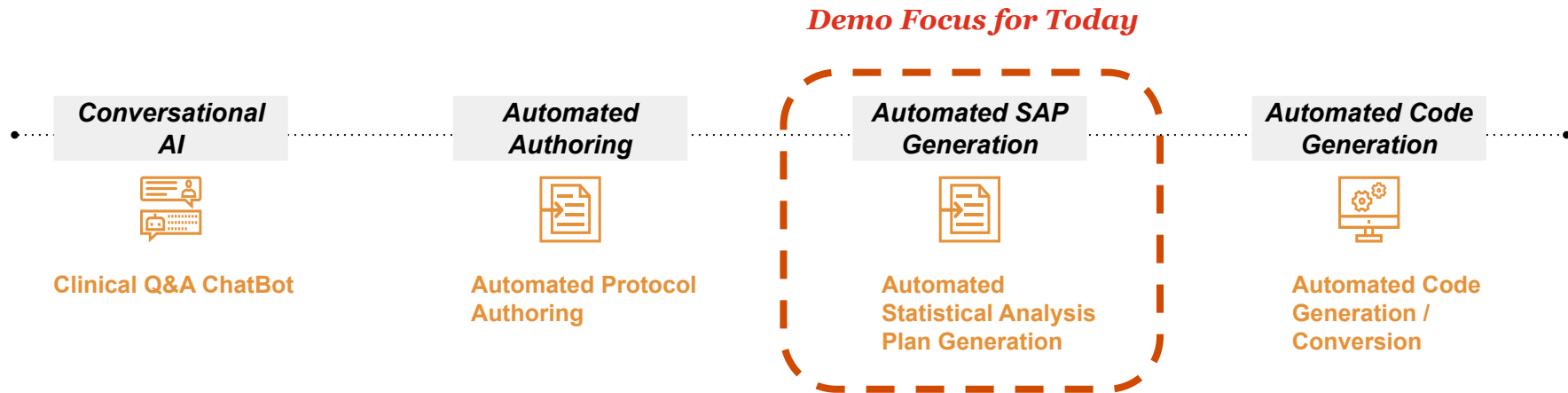
GenAI validation focuses on evaluating the quality and characteristics of generated data, while ML validation primarily assesses the performance and accuracy of predictions on existing data.

AI Capabilities Needed to Enable Monitor Controls

AI technologies should be leveraged in the correct fit-for-purpose scenario, each having their own strengths and weaknesses

<i>Capability</i>		<i>Capability Description</i>
	Monitor and Calculate	Ability to baseline model metric (e.g., precision, recall, f1-score, etc.) and actively compare to actual performance / feedback
	Human “In-the-Loop”	Ability to allow human reviewer to analyze model performance and model predictions to review false positives and false negatives
	Dictionary Management	Ability to manage external dictionary updates (e.g., IME list) and custom model-specific dictionaries (e.g., suppress common false positives)
	Dataset Selection /Generation	Ability to review and assess training data needed, and select or generate dataset for re-training
	Retrain and Validate	Ability to retain models based on user data or dictionary updates, compare performance (baseline vs. updated model), and validate model accuracy
	Model Deployment	Ability to redeploy updated models to production
	Reporting and Notification	Ability to export performance reports and alert relevant stakeholders when model performance is below accepted threshold

We have started to develop a GenAI Co-Pilot for automating Clinical Trial design & analysis...



What You Will See...

Ability to **query clinical trial data** for TrialTrove

- Ability to **automatically generate sections** (e.g., IE, endpoints) **of clinical documents** (e.g., protocol, statistical analysis plan)

Ability to **automatically generate SAS code** and **convert code from SAS to R, Python**

...and we built “Research Buddy” to support Research hypothesis evaluation via GenAI

USE CASES - where could “Research Buddy” help?

Demo Focus for Today

Unmet Medical Need, TA / Disease Identification

Current state of care? Unaddressed issues?

Target Assessment

What genes are associated with a particular disease? To what degree are they associated? What other diseases are impacted by this gene?

Target Evaluation

Function of target? Biomarker ID? Chemical tools? Binding sites (i.e. natural ligands)? Competitive landscape? Toxicity? Post translational modifications? Tissue expression?

Druggability Assessment

What modality (e.g., small molecule, conjugate, antibody)? To what degree could this target be modulated? Potential mechanism of action? Protein structure? Tissue / cell expression?

Drug Candidate Identification

What classes of molecules may modulate the target? To what degree? How do they interact with the target?

Dosing / Therapeutic Window

What known information is there on potential safe dosing ranges? What evidence is there on dose timing?

Repurposing / Secondary Indications

What other ligands might the molecule bind? Are there other indications impacted by the molecule? How is the drug being used off label?

CAPABILITIES - what could “Research Buddy” do?



Structured & unstructured data extraction



Question & answer



Literature review



Experiment ideation



Report generation



Document comparison



Data analysis & visualization



Bioinformatics code generation & review

Demos



1)

[Link to SAP Code Generator](#)

2)

[Link to Research Buddy](#)

Thank You!

