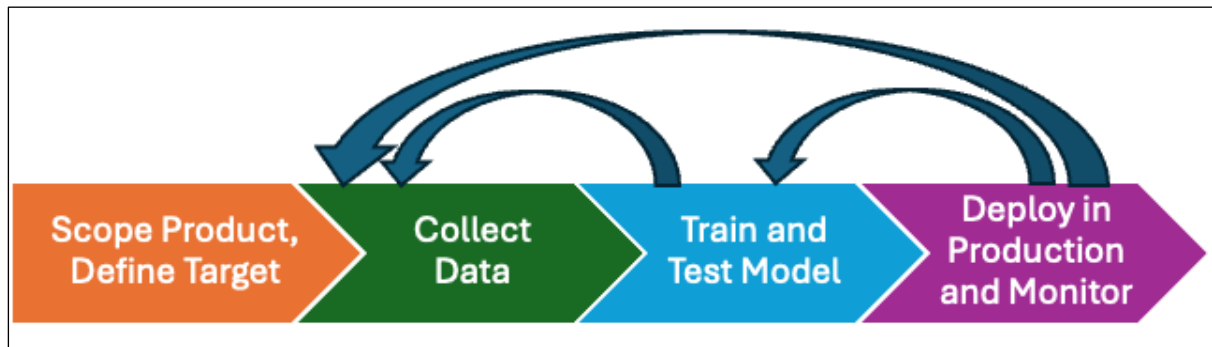
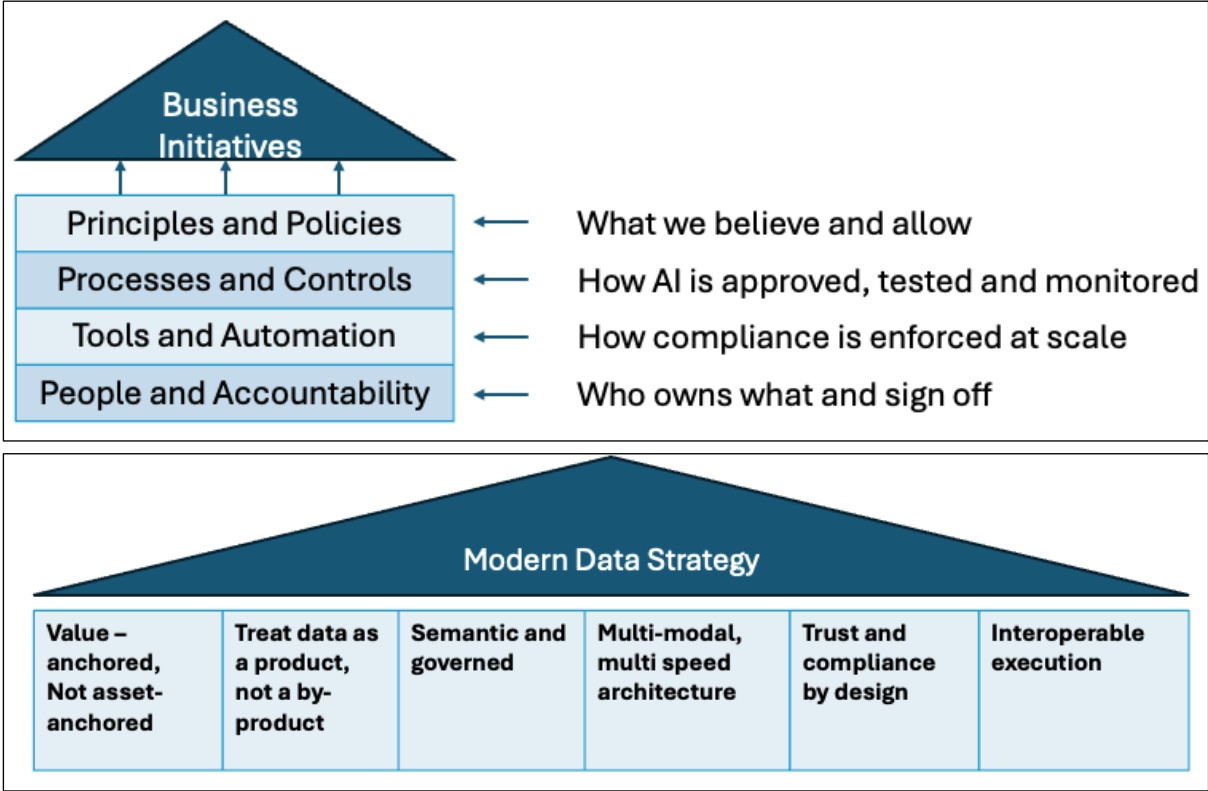


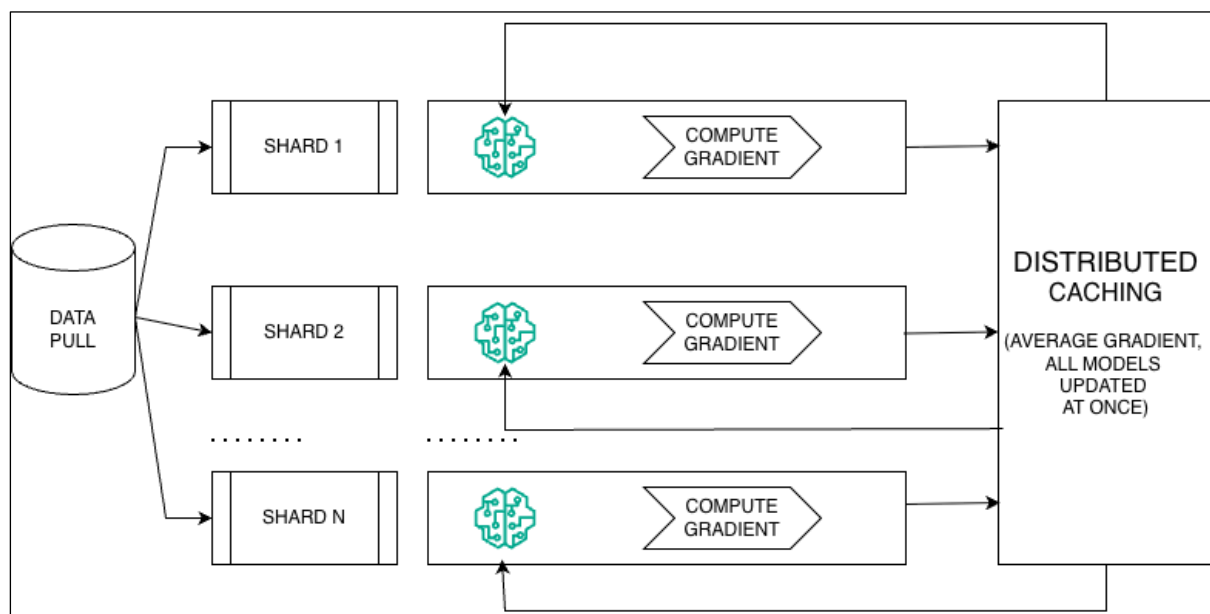
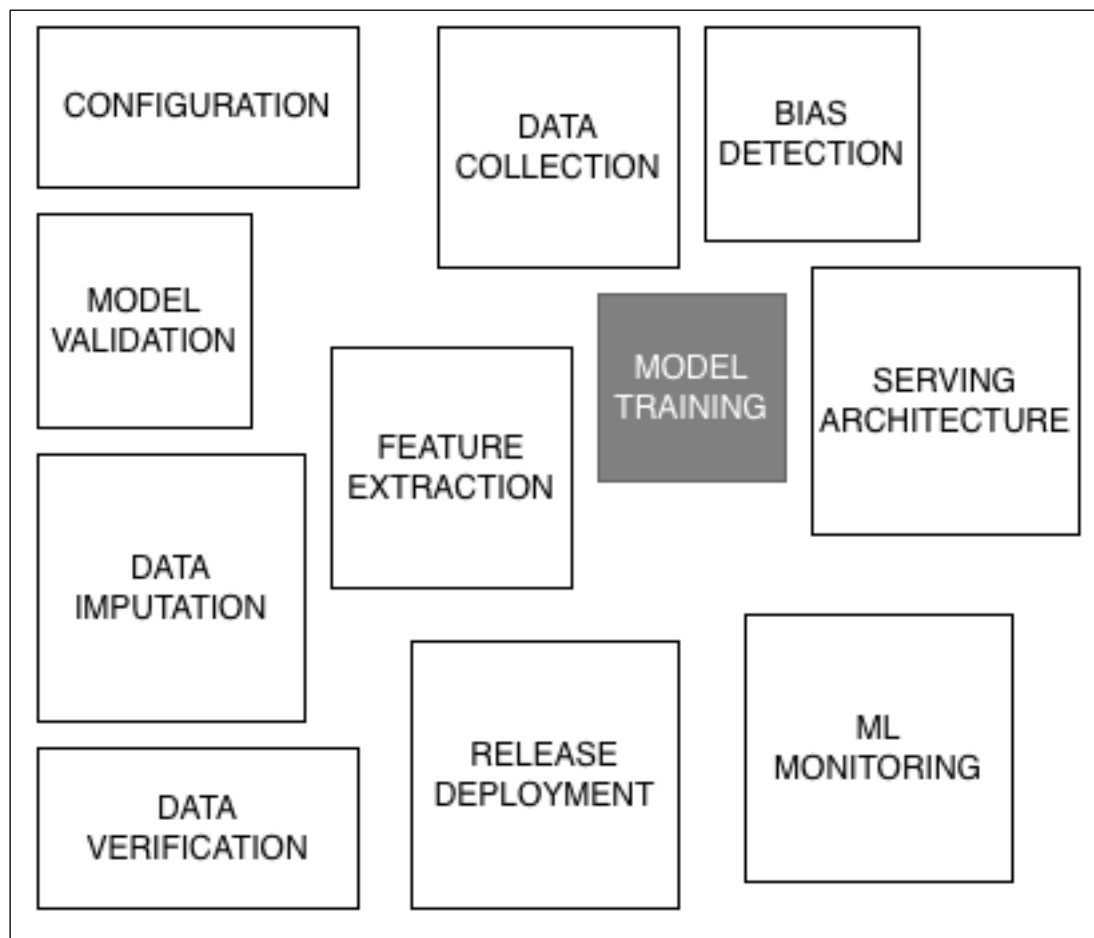
Chapter 1: Understanding the Perils of AI Products



Chapter 2: Building the Enterprise AI Strategy



Chapter 7: From Model to Market: Operationalizing ML Systems



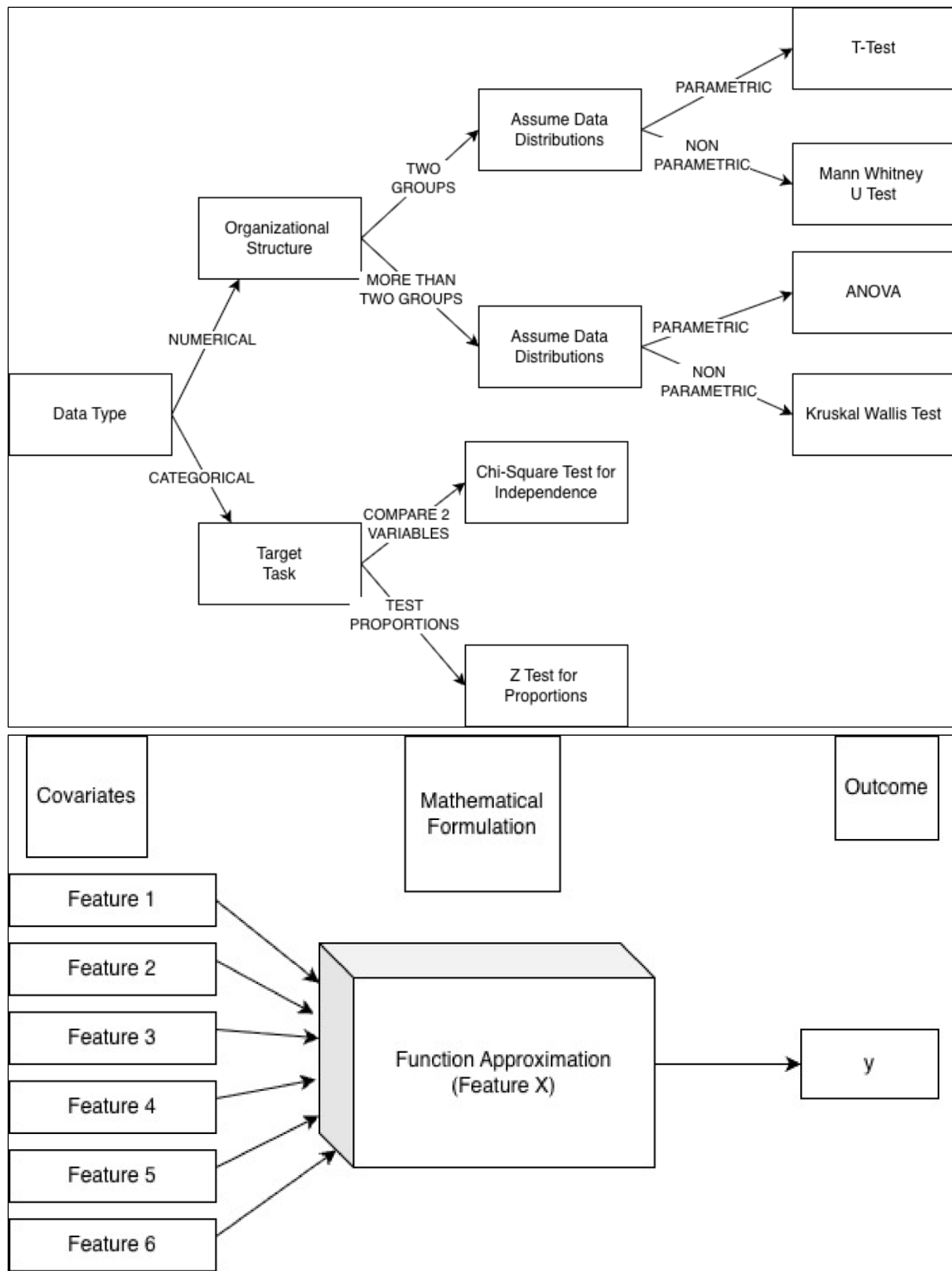
PLATFORM
SOLUTION

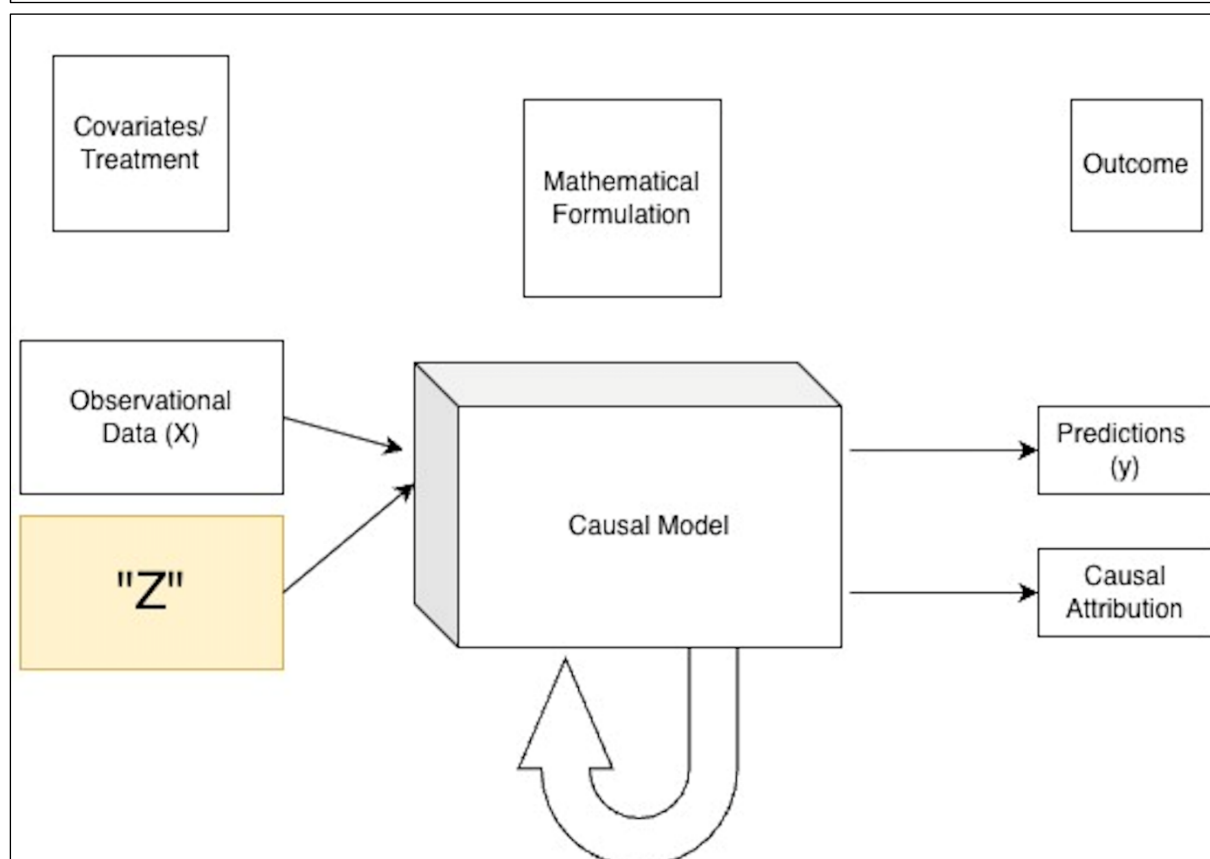
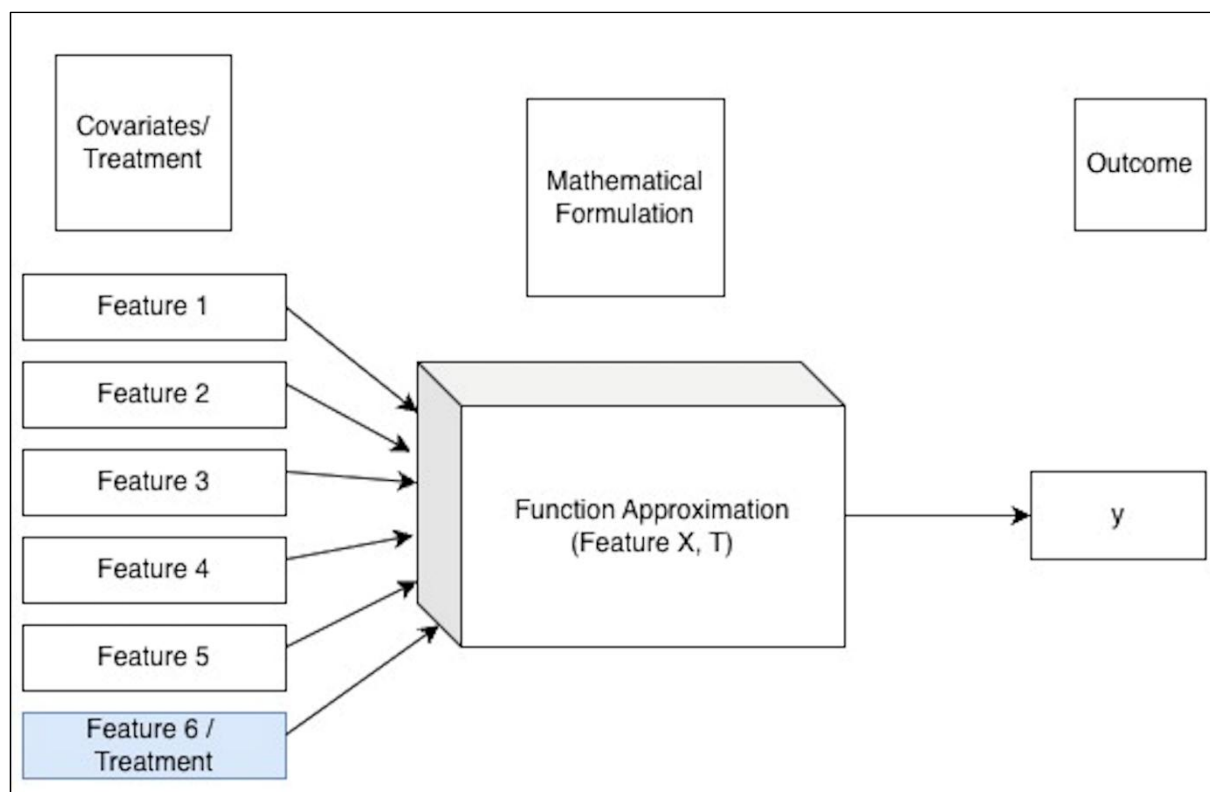
INTEGRATION
COMPONENT

ARCHITECTURE

API

Chapter 8: From Metrics to Measurement: Experimentation and Causal Inference





Chapter 10: Understanding GenAI Operations

The odd numbers in this group add up to an even number: 4, 8, 9, 15, 12, 2, 1.

A: Adding all the odd numbers (9, 15, 1) gives 25. The answer is False.

The odd numbers in this group add up to an even number: 17, 10, 19, 4, 8, 12, 24.

A: Adding all the odd numbers (17, 19) gives 36. The answer is True.

The odd numbers in this group add up to an even number: 16, 11, 14, 4, 8, 13, 24.

A: Adding all the odd numbers (11, 13) gives 24. The answer is True.

The odd numbers in this group add up to an even number: 17, 9, 10, 12, 13, 4, 2.

A: Adding all the odd numbers (17, 9, 13) gives 39. The answer is False.

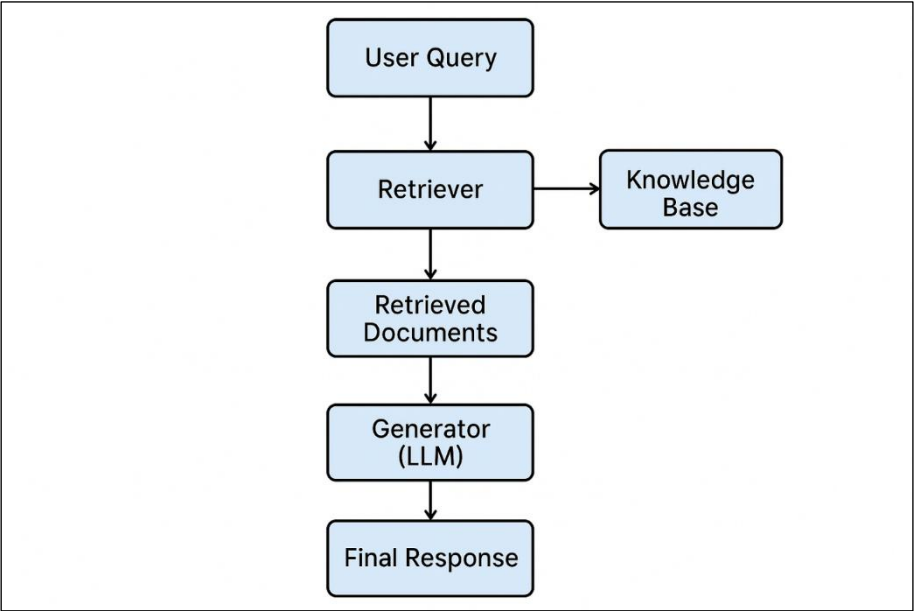
The odd numbers in this group add up to an even number: 15, 32, 5, 13, 82, 7, 1.

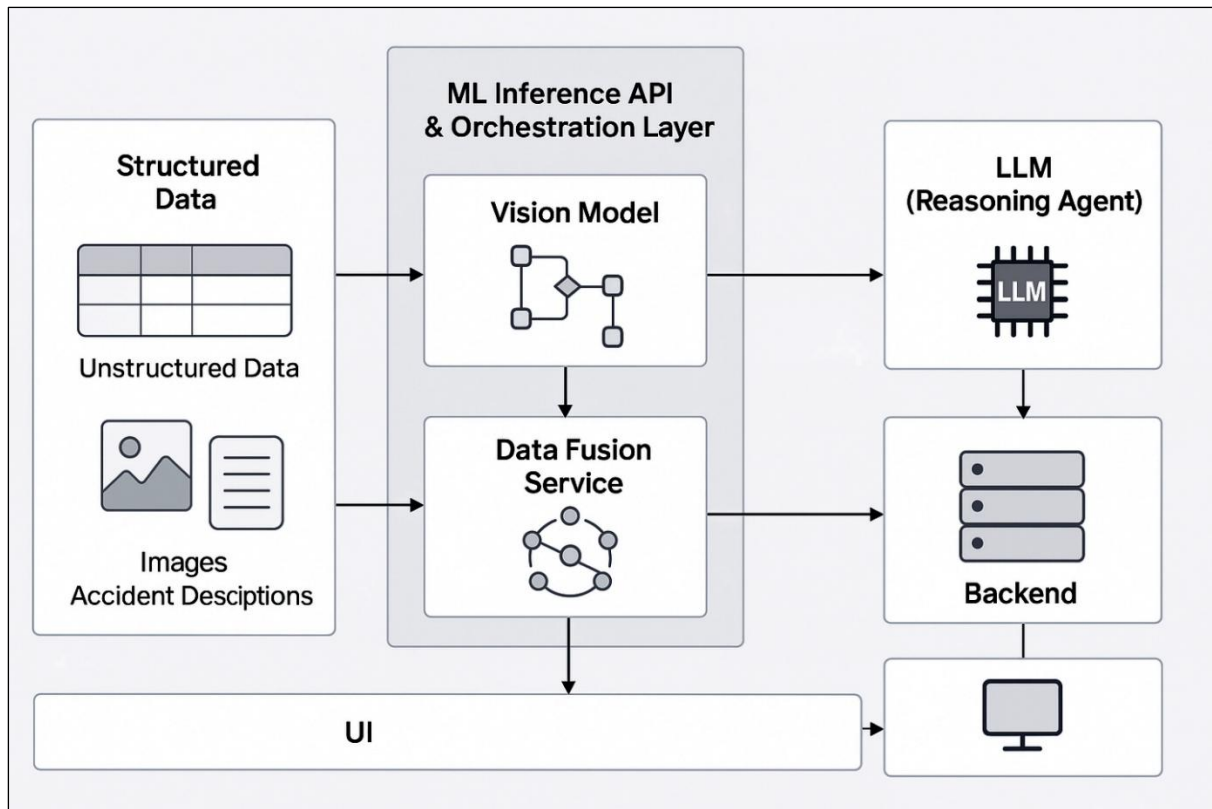
A:



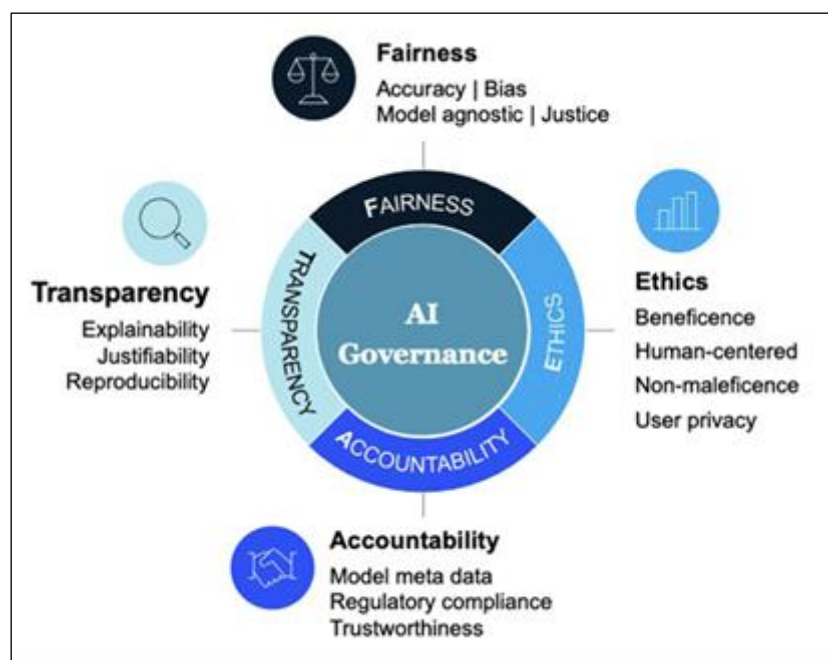
A: Adding all the odd numbers (15, 5, 13, 7, 1) gives 41. The answer is False.



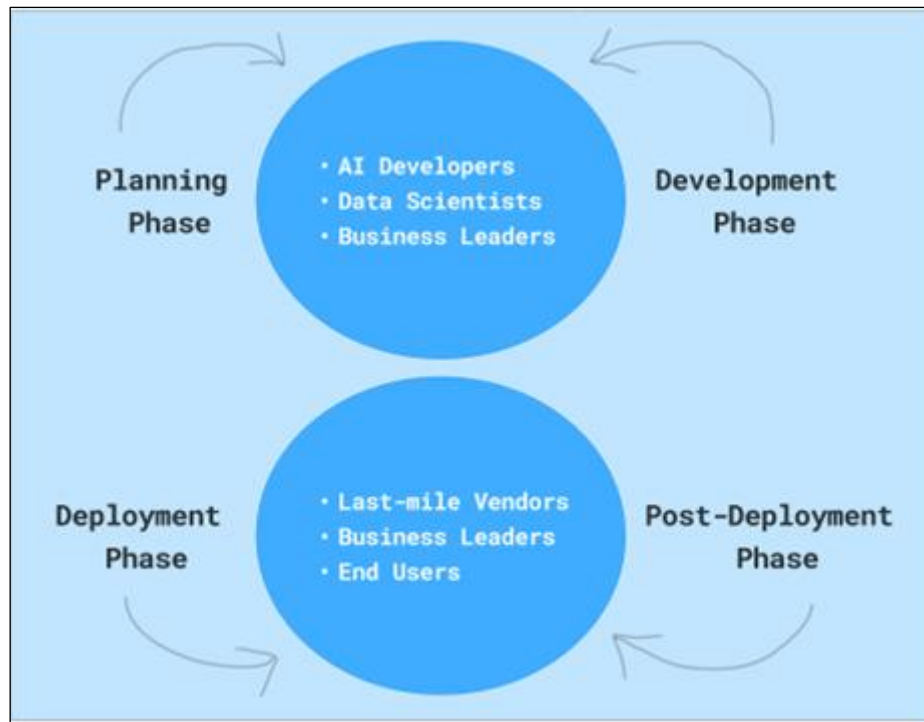




Chapter 12: Introduction to Responsible AI



Fairness 	Ethics 	Accountability 	Transparency 
Accuracy: To determine what level of accuracy makes sense for the business or product. Bias: Given that data always carries bias, it is important to filter the unwanted bias at the stage of data collection. Model Agnostic: Applies to the algorithmic design, use of technology tools, and the business logic associated with the model as how well the model is interoperable among various systems. Justice: Promote equal opportunities for all individuals and groups.	Beneficence: Refers to "Do only good" that firmly underlines the central importance of promoting the well-being of people and the community. Human-Centered: The organization working on an AI model should be for human benefit and document its intentions, as well as underwrite them with standards of certain desirable values such as human rights, transparency, and harm-avoidance. Non-Maleficence: Doing "no-harm" through prevention of both accidental and deliberate damages. User Privacy: Application design often needs to balance competing demands to optimize the accuracy of a system or ensure user privacy for explainability needs.	Model Meta Data: Ability to capture model metrics and meta-data across the entire model lifecycle that demonstrates the model and its delivery process to generate results that are auditable, traceable, and verifiable. Regulatory Compliance: Striking an appropriate and compliant balance between AI beneficence and the drive for innovation and growth. Trustworthiness: Integrity of the AI has been purposefully designed, driven with effective oversight and agile governance, maintained and well- executed.	Explainability – System generates output that can be understood and interpreted by both developers and users. The goal is not to expose the exact, inner technical workings of the algorithms used to get to a certain outcome. Rather, the goal should be to expose why certain standards for the application are met or not met. Justifiability – Soundness in the justification of its use such that one can demonstrate both the design and implementation processes for the result of a particular decision or behavior. Also, the model should be ethically permissible and worthy of public trust. Reproducibility – Developing the infrastructure to enable for a reasonable level of reproducibility across ML system operations.



Chapter 13: Implementing RAI Frameworks, Metrics, and Best Practices

Model Card: Pneumonia Detection System for Healthcare Sector	
Prepared by: Clinical AI Solutions team	
Model Details: - Model name: Pneumonia Detection model - Version: v2.0 Date: Nov 1, 2025 - Developed by: Clinical AI Solutions team (built in-house at the hospital/clinic partner) - Latency: 1-3 seconds. Must be fast enough for use in emergency settings for patient safety. - Contact information: CMO Office / ClinicalAISupport@hospital.org Intended Use: - Primary use case: Assist licensed radiologists and Physicians in rapidly diagnosing acute pneumonia in adult chest X-ray scans for patients aged 18-85. - Primary output: Binary Classification: Pneumonia Detected (Yes/No) with an associated confidence score (0-100%). - Intended users: Licensed Radiologists, Lung Specialists, and Clinical Review teams. - Out-of-Scope use cases: Must not be used for final diagnosis; not validated for pediatric patients (under 18) or pregnant individuals. Factors: - Model inputs: Anonymized medical image data (X-rays), relevant demographic features (age, sex, ethnicity), and historical patient records (co-morbidities). - Model outputs: Risk scores (Severity 1-5), Triage priority flags (High/Medium), and Justification heatmaps. - Demographic factors: Applicable. Performance varies slightly by age and ethnicity. Audited for bias. Metrics: - Accuracy: 94.8%. Technical Performance (Note: "Accuracy" is used here for technical completeness providing high-level view of model correctness. This can be misleading for patients with low disease prevalence). - Precision: 79.5%. Provides the model's efficiency and helps avoid false alarms with unnecessary follow-ups. - Recall/Sensitivity: 92.1%. Measures the ability to not miss a sick patient which is critical for false negative avoidance. - F1 Score: 0.85 - Specificity: 95.5%. Measures the ability to correctly rule out the disease. Quantitative Analysis: Metrics prioritize patient safety. Low False Negative Rate (FNR) is critical for safety and trust so that high-risk cases are not missed, which is essential to minimizing patient harm. - False Negative Rate (FNR): 7.9% (Directly quantifies the rate of missed diagnoses. Type II error) - False Positive Rate (FPR): 4.5% (Quantifies the rate of over-diagnosis. Type I error) - Negative Predictive Value (NPV): 98.2% (High NPV is essential for a physician to trust a "No Pneumonia" result. - Positive Predictive Value (PPV): 79.5% (Indicates the efficiency of the model in a clinical setting to identify how often the alerts are correct) - AUC-ROC: 0.96 (Overall classification performance, independent of the chosen threshold).	Training Data: - Data sources: Historical anonymized patient X-rays, manually labeled by two independent, board-certified radiologists, and structured clinical records. - Sensitive data: The model uses Protected Health Information (PHI). Strict measures (de-identification, encryption, HIPAA/GDPR compliance) are in place to prevent the exposure of any PHI. Evaluation Data & Protocol: - Datasets used: - Training data: Aggregated, de-identified Chest X-ray data from 3 major hospital systems (urban, suburban and a large teaching hospital). - Ground truth labeling: Labels were established by consensus among two board-certified radiologists. - Test size: 10,000 unique Chest X-ray images (stratified by age and sex). - Temporal range: Jan 2015 - May 2025 - Geographic range: Primarily North America (USA). Bias & Error Analysis: - Bias analysis: Regular assessments so that the model does not disproportionately impact any user group. - Error analysis: Analysis of false positives and false negatives to improve model accuracy. Ethical Considerations: - Bias & Fairness: The model's sensitivity for the 70+ age group (89.5%) is lower than for younger adults (93.5%). - Bias mitigation: The system must include a warning flag for patients above 70 advising physicians to exercise higher clinical judgment. - Continuous monitoring and updating of the model to address any biases in the training data is crucial for mitigation. - Transparency: Due to the complexity of the output score, the model is restricted to certified radiologist or emergency physician. A Human-in-the-loop process is mandated; the model's score is a recommendation only and need to be signed-off with a physician for conclusive diagnosis. - Accountability: Deployment is restricted to regions where the local regulatory body (such as FDA) has explicitly signed off on the model card and Intended use. Measures are in place to be respectful of data privacy and security without exposing any PHI. Caveats and Limitations: - The model is validated only for adults aged 18-85. Not validated for pediatric, neonatal populations, and co-morbidity risk. - The model does not screen for other critical conditions (Pneumothorax, cancer) and should not be used as a general screening tool. Recommendations: - Human-in-the-loop: Final diagnosis must be provided by a licensed, supervising physician or radiologist as the model score is only a recommendation. - High-Risk Protocol: For critical decisions in maintaining accuracy and fairness, generate an immediate, high-alert flagging in the system for all cases where the confidence score is < 70% of Pneumonia detected. - Maintenance & Audit: The model's performance must be audited once a quarter against newly obtained data to check for data drift and bias.

@Usha Jagannathan, 2025

