ACM India Summer School on Responsible & Safe AI – IIT Madras

Operationalizing AI

Dr. Anand Rao

Distinguished Service Professor of Applied Analytics and Artificial Intelligence

Heinz College of Information Systems and Public Policy

Carnegie Mellon University

Email: anandsrinivasarao@gmail.com

LinkedIn: https://www.linkedin.com/in/anandsrao/

Instructor Bio

Education:

- PhD Artificial Intelligence, University of Sydney
- MBA, Melbourne Business School
- MSc(Tech), Computer Science, BITS Pilani, India

Current:

- Distinguished Service Professor of Applied Analytics and Artificial Intelligence at Carnegie Mellon University
- **Courses:** Operationalizing AI, Responsible AI & Applications of LLM, Agent-Based Models and Digital Twins
- Venture Partner at Golden Sparrow, an early-stage VC firm; Advisor at Innospark Ventures; Advisor/Director for 4 startups;
- Advisory council of Oxford's Institute for AI Ethics, Chair for ISDM's Center for Decision Science and Social Impact, Northwestern's MBAi program

Previous:

- Global AI Leader, Partner in Data, Analytics, and AI, Innovation lead for AI and Emerging Technology, PwC
- Partner, Diamond Management & Technology Consultants
- Chief Research Scientist, Australian Artificial Intelligence Institute

Agenda

- Foundations of Operationalizing AI
- Mindsets: Data Scientists are from Mars & Software Developers are from Venus
- End-to-end AI Systems Lifecycle
  - Value Scoping and ROI of AI
  - Value Delivery and Scaling AI Systems
  - Value Stewardship and Continuous Monitoring
- Operationalizing AI Responsibly
- Operationalizing Generative AI

FOUNDATIONS OF OPERATIONALIZING AI – TURNING INSIGHTS TO OUTCOMES

Agenda – Foundations of Operationalizing AI

- Scenario: The GAMMOA Corp
- What is Operationalizing AI?
- Why is Operationalizing AI important?
- When do you Operationalize AI?
- Who should be involved in Operationalizing AI?
- How do you Operationalize AI?
- Case study: Success and failure stories of AI operationalization

GAMMOA Corp's AI Odyssey: From Innovation to Ethical Dilemmas

Discussion Question:

1. What does operationalizing AI mean for you? What key components or stages do you associate with this process?
2. Why is operationalizing AI important for modern businesses like GAMMOA?
3. Who should be involved in operationalizing AI? What new roles emerge, and what roles do cross-functional teams play? Is there a need for a dedicated oversight role or team?
4. How do you operationalize AI? What best practices and strategies should organizations adopt to ensure smooth AI model deployment, integration, and ongoing relevance in business operations?
5. At what stage of your organizational maturity should you consider operationalizing AI?

GAMMOA Corp's AI Odyssey: From Innovation to Ethical Dilemmas

Discussion Question:

6. How can organizations ensure that critical ethical warnings are not sidelined, especially when there's pressure to launch a product?
7. Discuss the implications of using data without proper lineage or intellectual property rights. How can negative media attention impact a company's reputation and customer trust?
8. What steps can organizations take during the model development and validation phases to identify and rectify biases? How can companies ensure their AI models are both accurate and fair?

What is “Operationalizing AI”?

“Operationalizing AI is a framework for scaling AI models to AI systems to realize the full business value of AI. It encompasses the strategic, ethical, sustainable, financial, technological, operational, people, organizational, risk, governance, and compliance dimensions of AI.”

What is “Operationalizing AI”?

1. **Strategy:** AI goals and roadmap aligned with business
2. **Value:** Discovering, realizing, and maintaining AI’s financial and strategic impact
3. **Operations:** Workflow, lifecycle, and process management across, data, models and code
4. **Technology:** Tools, techniques, and infrastructure for development, deployment, and operations
5. **Responsibility:** Governance, risk management, compliance, and ethical consideration of AI
6. **People & Organizational:** Skills, roles, and team structure
7. **Culture**: Ethical, societal, and workplace culture and trust

Focus

PRIMARY

Business Decision Making

- Industry sector
- Functional areas
- Business model
- Innovation
- Talent & Skills
- Risk Management
- Governance
- Process

What business decisions need to be made when designing, building, deploying, and operating AI systems? Why do we make these decision?

Technical Decision Making

- Applications
- Architectures
- Build vs Buy
- Data
- Models
- Software
- Compute
- Compliance
- Tools
- Techniques

What technical decisions need to be made when designing, building, deploying, and operating AI systems? Why do we make these decision?

Tools, Frameworks, Programming

- Languages – Python, R
- Frameworks – Tensorflow, PyTorch
- Cloud platforms
- Edge devices
- Data manipulation
- LLMs
- Adaptation
- Prompt Eng.

How do we implement AI systems end-to-end and what kinds of tools, frameworks, languages should practitioners know?

Mathematical Foundations

- Probabilities & Statistics
- Linear algebra
- Operations research
- Calculus
- Information theory
- Discrete mathematics

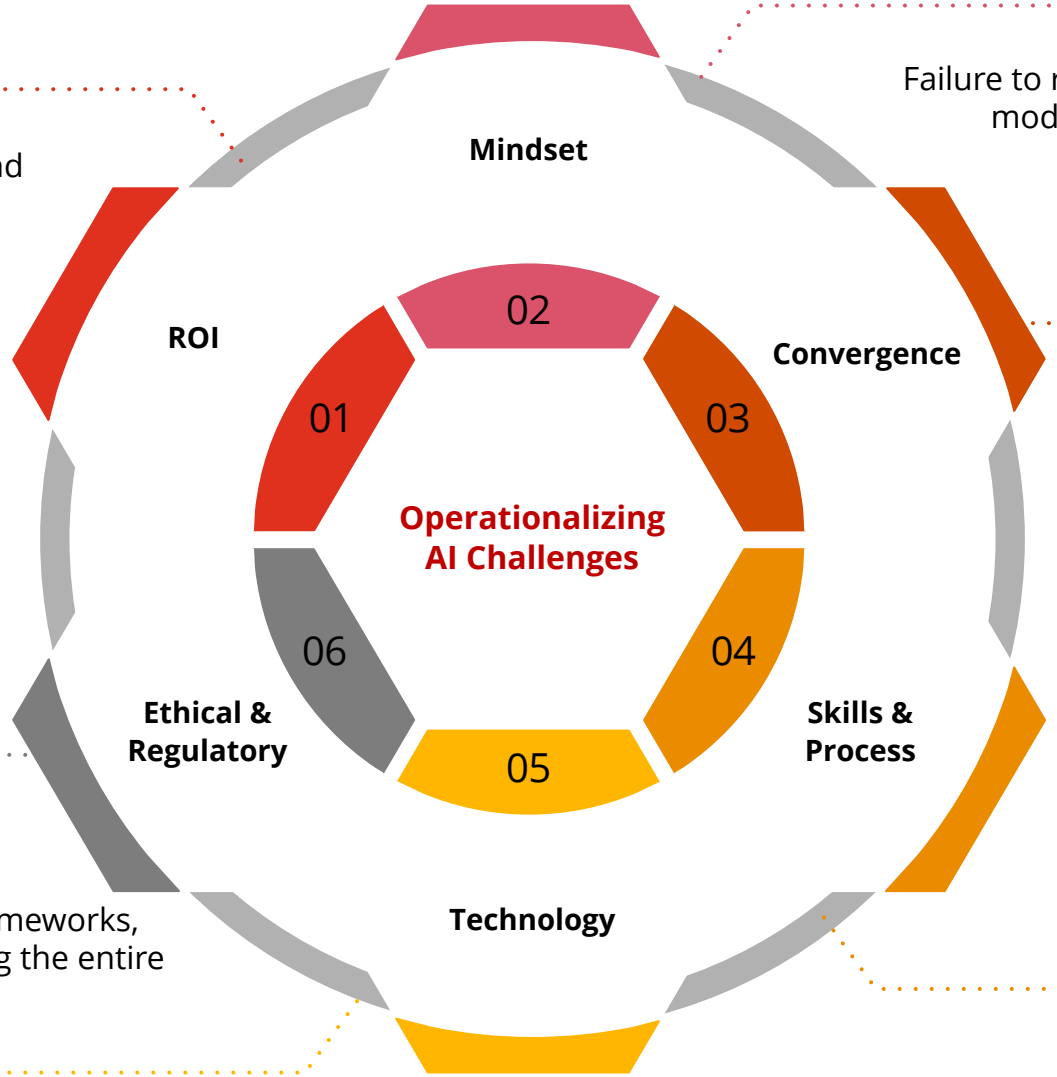
What are the fundamental mathematical foundations to design, build, and operate AI systems better?

Why is operationalizing AI challenging?

ROI Challenge
Lack of clarity on measuring, maximizing, and maintaining ROI in AI initiatives

Ethical & Regulatory Challenge
Uncertainty in cost to build, maintain, and comply with emerging principles, policies, and regulations

Technology Challenge
Constant innovation and proliferation of frameworks, tools, techniques, and vendors for managing the entire IT stack and end-to-end AI system life cycle



Mindset Challenge
Failure to recognize that software development and AI model development are fundamentally different

Convergence Challenge
Convergence of data, model, and code across the different phases of scope, design, develop, and operate are resulting in multiple, iterative, and interacting feedback loops

Skills & Process Challenge
Continuously evolving best practices and skills to address the end-to-end life cycle of data/(AI) model/code (referred to as "AI system") leading to shortage of these skills

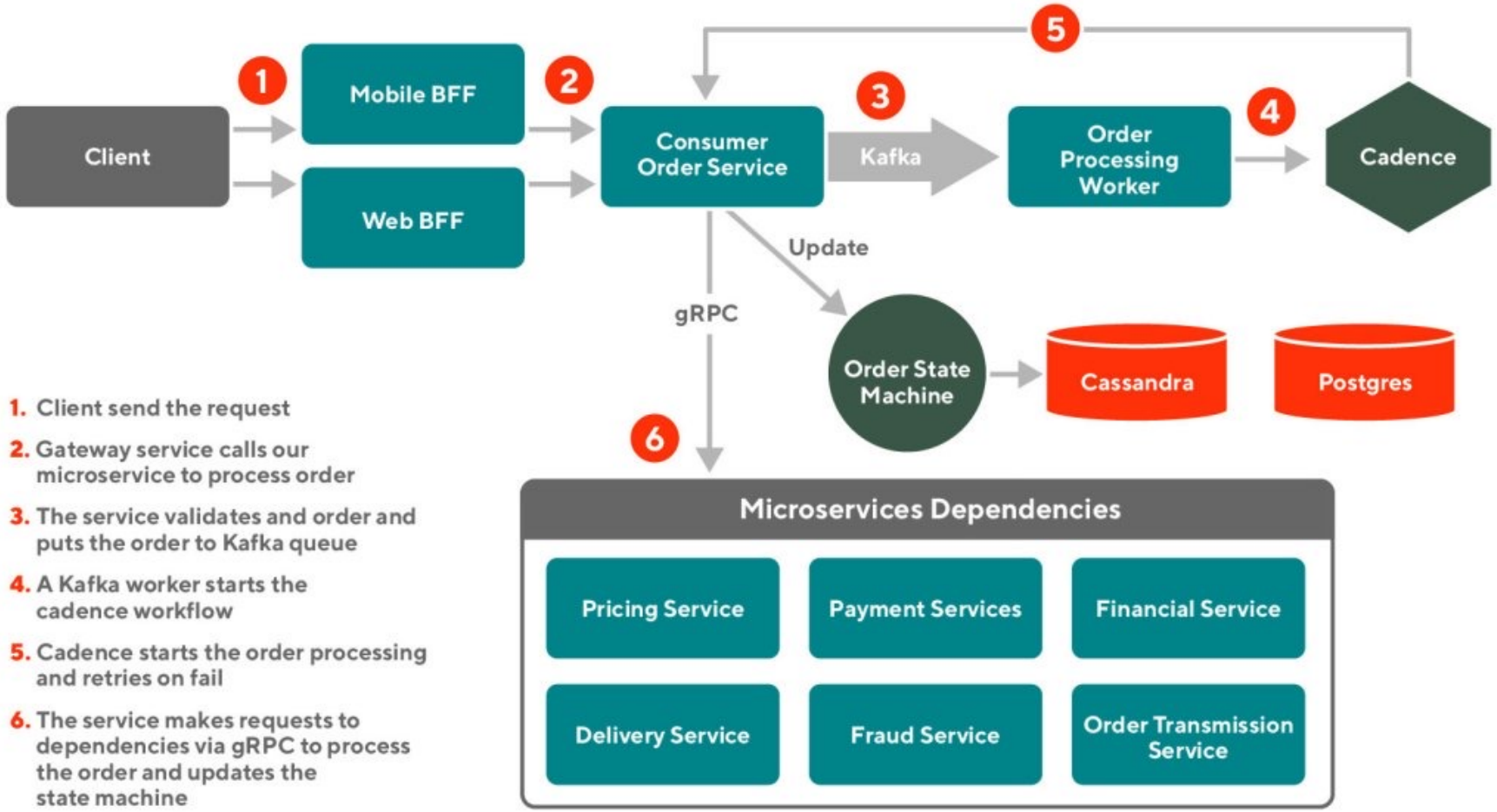
Source: Data Scientists are from Mars and Software Developers are from Venus (Part 1), Rao, A., Towards Data Science, August 29, 2020

Why is Operationalizing AI important?

- Automation of tasks – manual or cognitive; periodic or repetitive
- Increase in productivity, reduced headcount, and cost savings
- Better decision-making, enhanced experience, stickiness, and more revenue
- Consistency in decision-making, better quality and improved risk-management

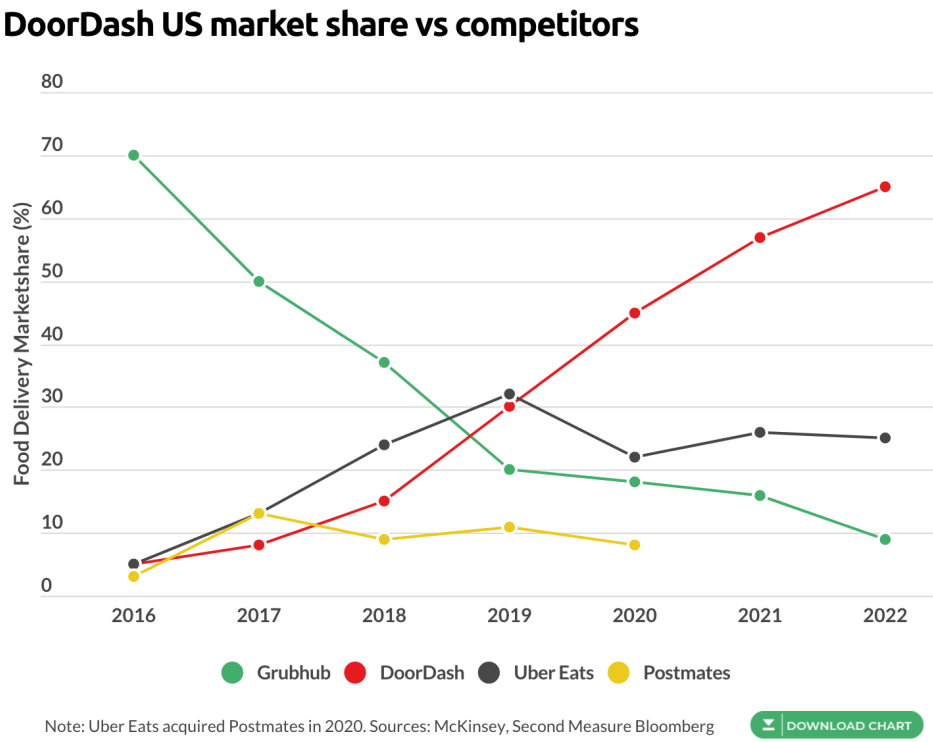
AI systems are much more than AI models

Building Reliable Checkout Service at Scale - DoorDash



Source: [Building a More Reliable Checkout Service at Scale with Kotlin](#) by Yimin Wei, Zhengli Sun, Amiraj Dhawan, Doordash Engineering, February 2, 2021.

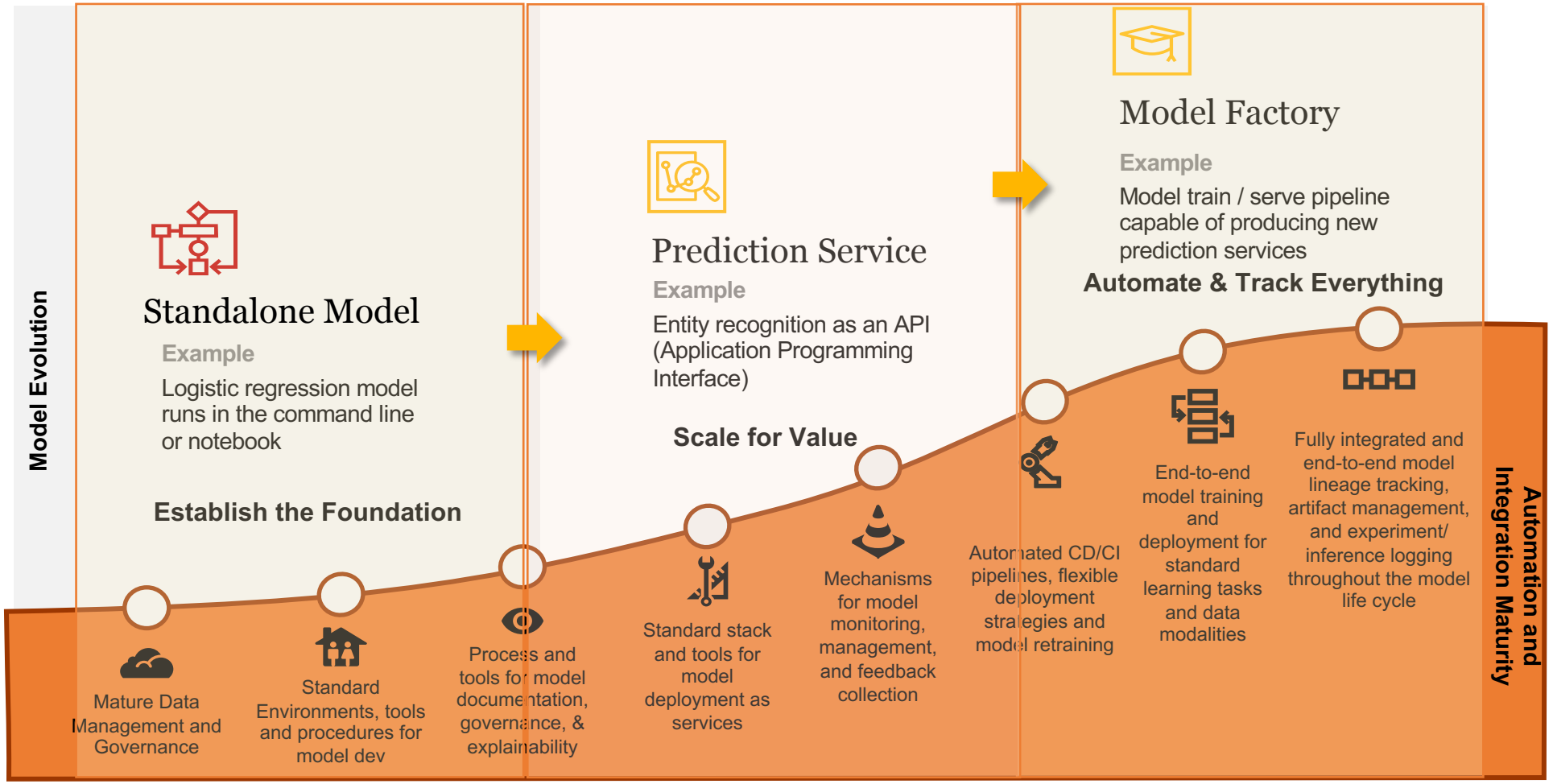
Engineering AI Systems provides significant value



	Mar 2020	June 2020	Aug 2020	Sep 2020	Oct 2020	Dec 2020	Jan 2021
# of models	2	16	20	24	28	44	38
Peak predictions/sec	1k	15k	130k	130k+	1M	2M	6.8M

Source: Scaling Online ML Predictions to Meet DoorDash Growth, apply() Conference 2021; *DoorDash Target Market Segmentation and Marketing Strategy – Audience Demographics & Competitors*, Start.io, August 3, 2022

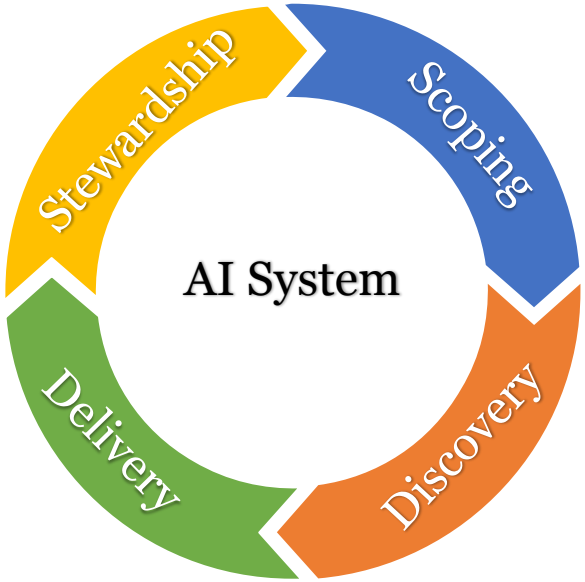
When do you Operationalize AI?



Source: Model Evolution: From standalone models to model factory. Rao, A., Towards Data Science, September 13, 2020

End-to-end AI System Life Cycle

AI System Development



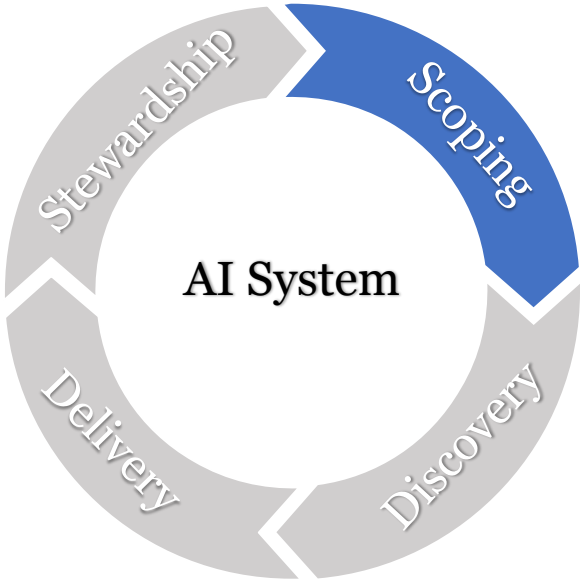
AI System is any software system that uses an AI or ML model with the data.

AI System development has four phases

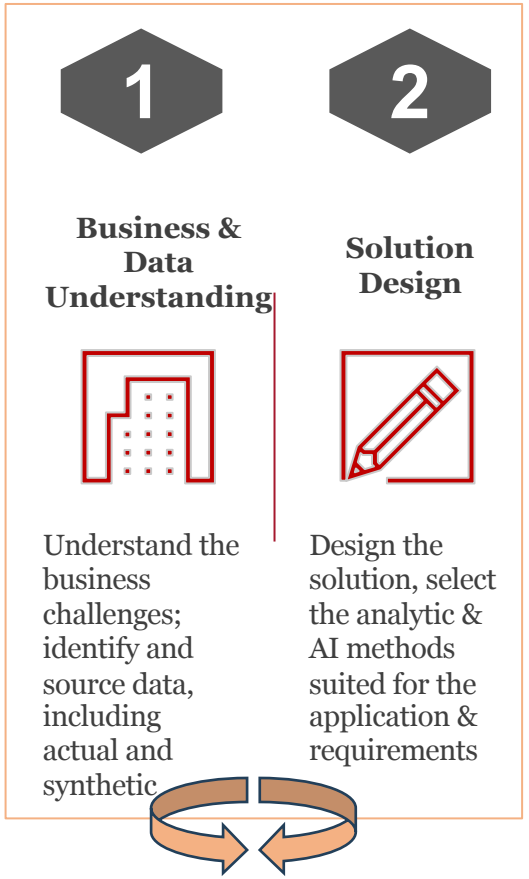
- Scoping
- Discovery
- Delivery
- Stewardship

Value Scoping

AI System Development

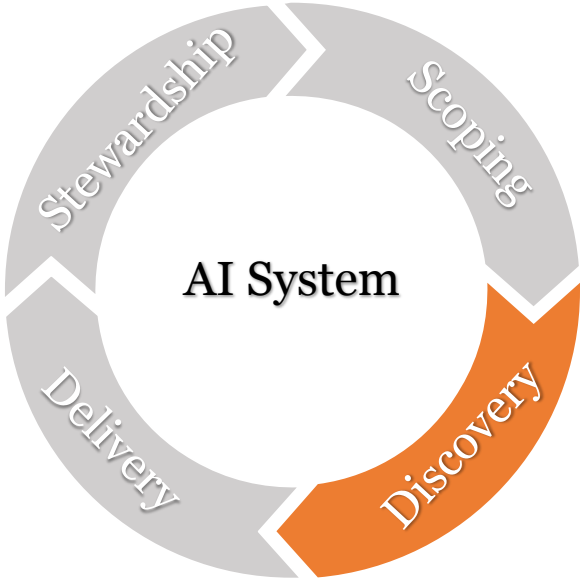


Value Scoping

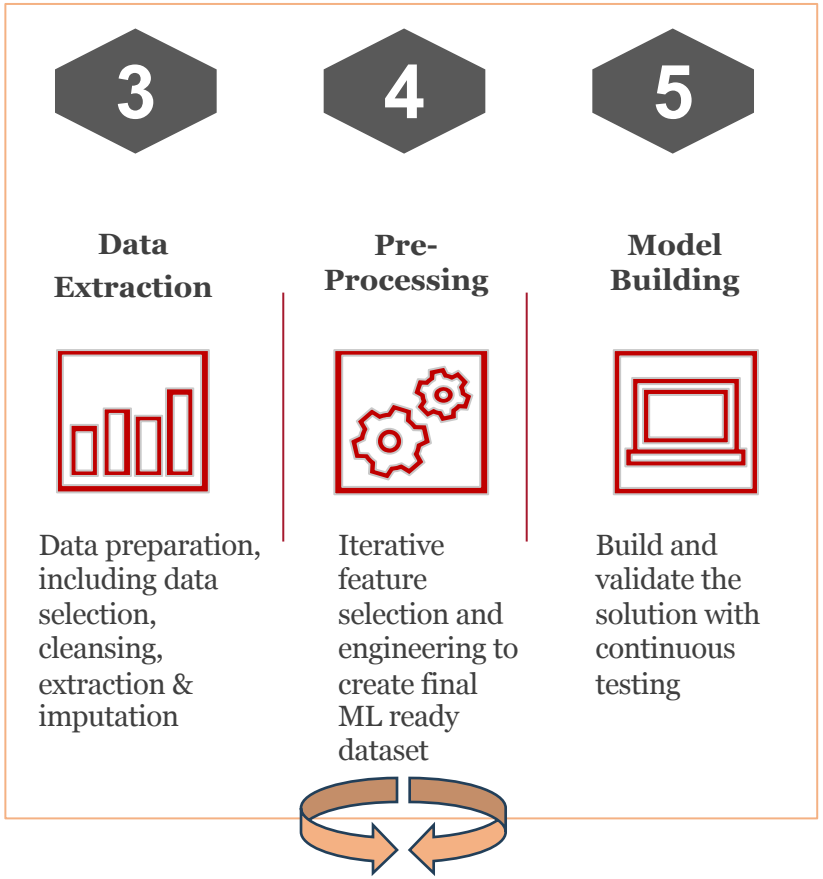


Value Discovery

AI System Development

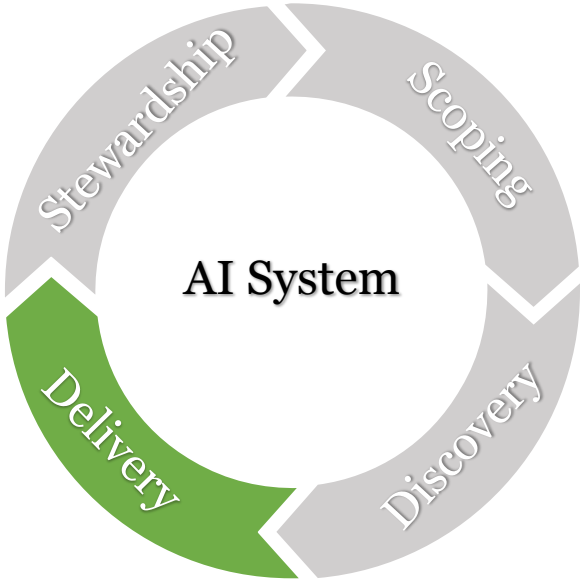


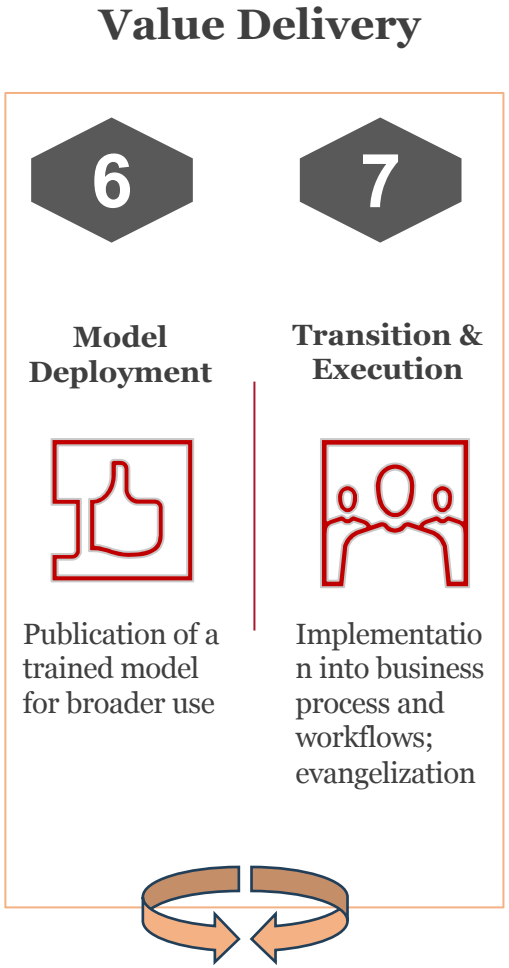
Value Discovery



Value Delivery

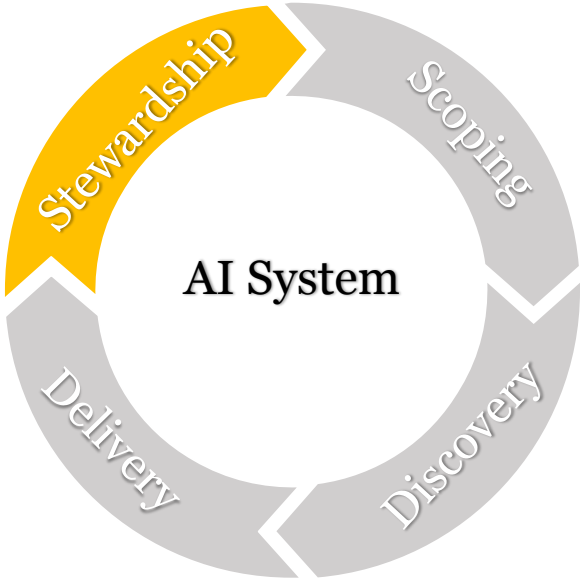
AI System Development



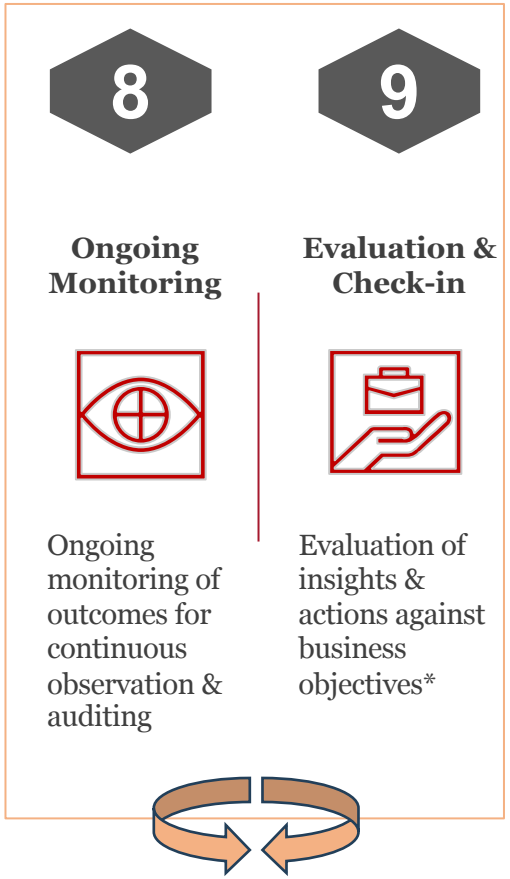


Value Stewardship

AI System Development



Value Stewardship



End-to-end AI Life Cycle: Linear View

Value Scoping

1 Business & Data Understanding

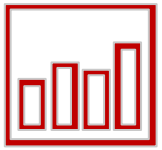
Understand the business challenges; identify and source data, including actual and synthetic

2 Solution Design

Design the solution, select the analytic & AI methods suited for the application & requirements

Value Discovery

3 Data Extraction

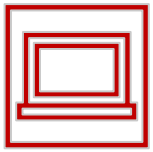


Data preparation, including data selection, cleansing, extraction & imputation

4 Pre-Processing

Iterative feature selection and engineering to create final ML ready dataset

5 Model Building



Build and validate the solution with continuous testing

Value Delivery

6 Model Deployment

Publication of a trained model for broader use

7 Transition & Execution

Implementation into business process and workflows; evangelization

Value Stewardship

8 Ongoing Monitoring

Ongoing monitoring of outcomes for continuous observation & auditing

9 Evaluation & Check-in

Evaluation of insights & actions against business objectives

Trust Management

People & Organization

MLOps Tools Resources

List of tools for MLOps_v2_Dec 2020

File Edit View Insert Format Data Tools Extensions Help

Share

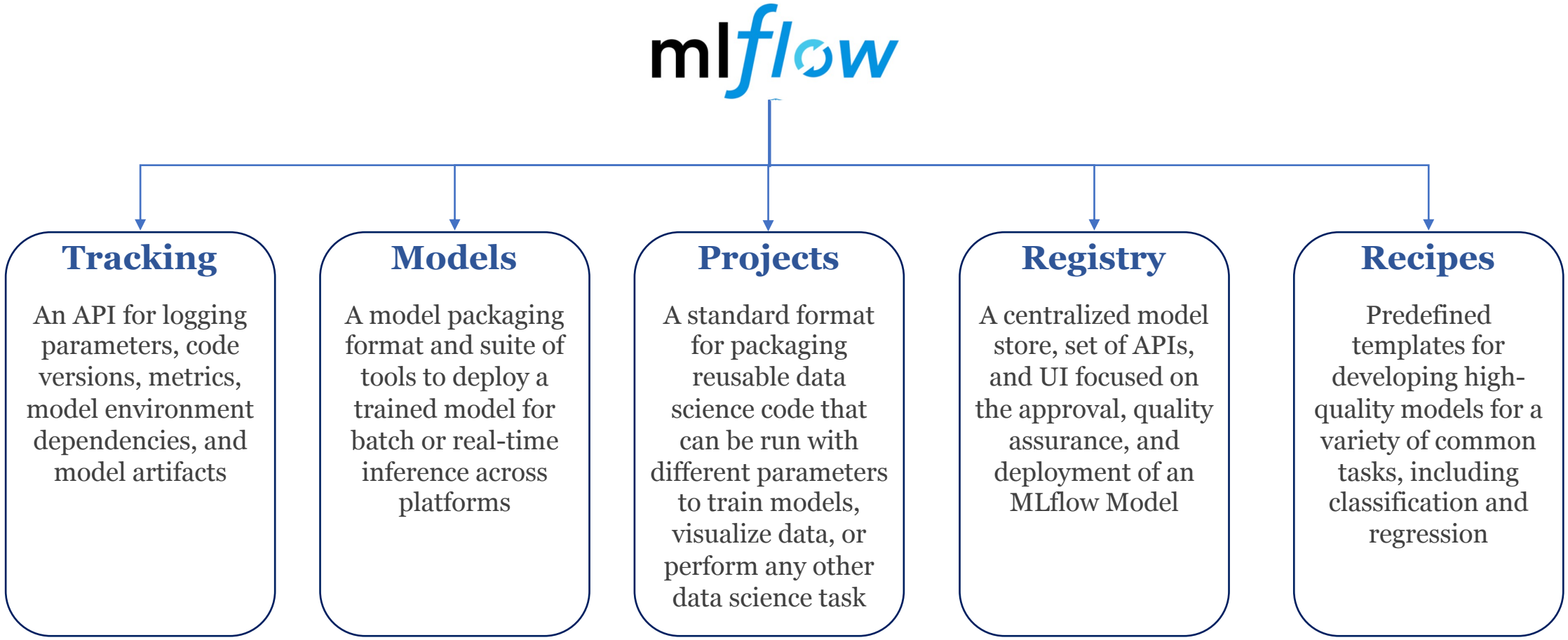
Menus 100% View only

A1 Name

	A	B	C	D	E	F	G	H	I	J	K	L
1	Name	Cat	SubCat	Series	$$$ (M)	Started	HQ	OSS	Change in 2020	Website	Description	IF ACQ
2	Peltarion	All-in-one	AI Apps platform	A	36.8	2005	Sweden			https://peltaric	A single AI platform, for real world deployments, without code. Fast & Efficient Production of AI Applications. Rich da	
3	DataRobot	All-in-one	AI Apps platform	F	750.6	2012	Boston		Raised 320M	https://www.di	DataRobot combines a trusted enterprise AI platform and a trusted AI-native strategic partnership for global enterpris	
4	H2O	All-in-one	AI Apps platform	D	151.1	2012	Bay Area	OSS		https://www.h/	H2O.ai is the creator of H2O the leading open source machine learning and artificial intelligence platform trusted by	
5	Dataiku	All-in-one	AI Apps platform	D	246.8	2013	NYC		Raised 100M	https://www.di	Dataiku's single, collaborative platform powers both self-service analytics and the operationalization of machine learr	
6	Iguazio	All-in-one	AI Apps platform	C	72	2014	Israel			https://www.ig	The Iguazio Data Science Platform automates your machine learning pipeline, transforming AI projects into real-worl	
7	Xpanse AI	All-in-one	AI Apps platform			2015	Ireland			https://xpanse	The power of AI at the click of a button. Xpanse AI brings easy to use and lightning fast analytics to your business.	
8	Stradigi AI	All-in-one	AI apps platform	A	40	2017	Canada		Raised 40M	https://www.st	Stradigi AI's powerful AI business platform, Kepler, fuels tangible results for enterprises. No AI or machine learning e	
9	Cubonacci	All-in-one	AI Apps platform		0	2018	Netherlands			https://www.cu	Machine learning lifecycle management Cubonacci enables organizations to focus on developing custom machine le	
10	Obliviously AI	All-in-one	AI Apps platform	Seed	0	2018	Bay Area		Raised unknown	https://www.ol	The entire process of running Data Science - building Machine Learning algorithm, explaining results and predicting	
11	Snorkel AI	All-in-one	AI Apps platform	A	15.3	2019	Bay Area	OSS	Raised 15M	https://snorkel	Programmatically Building and Managing Training Data	
12	kedro	All-in-one	AI Apps platform	McKinsey		2019	UK	OSS			Kedro is an open source development workflow tool that helps structure reproducible, scaleable, deployable, robust	
13	Abacus AI	All-in-one	AutoML	B	53.3	2019	Bay Area		Raised 48M	https://abacus	Abacus.AI makes it effortless to create large-scale customizable deep learning systems. Accurate predictions genera	

Source: *Machine Learning Tools Landscape v2 (+84 new tools),* Chip Huyen, December 30, 2020. *Spreadsheet of tools*.

MLflow 2.0 Components

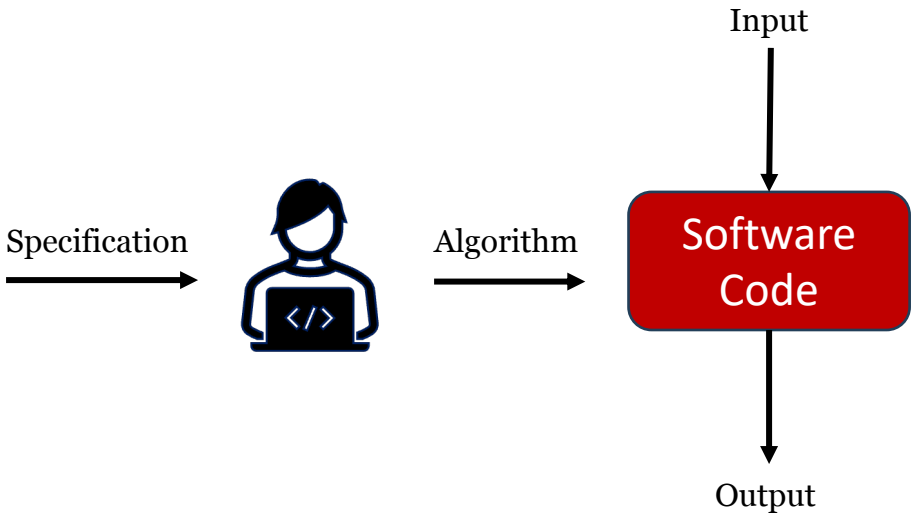


Source: What is Mlflow? Mlflow Documentation (As of October 21, 2023).

MINDSETS - DATA SCIENTISTS ARE FROM MARS AND SOFTWARE DEVELOPERS ARE FROM VENUS

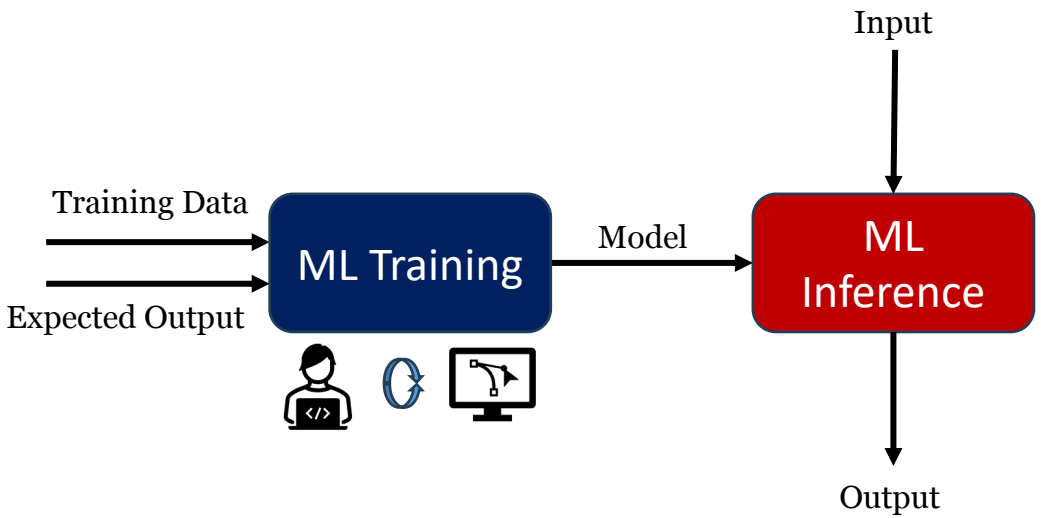
Software vs Machine Learning Models

Software Development



- Humans write specifications that are turned into algorithms embedded within a larger software system (Building software)
- Given new input the software code produces the output (Running software)

Machine Learning (ML) Model Development



- **Training Phase:** Humans train an ML model that can recognize patterns to make predictions
- **Inference Phase:** The trained model is applied to the new input to generate the output

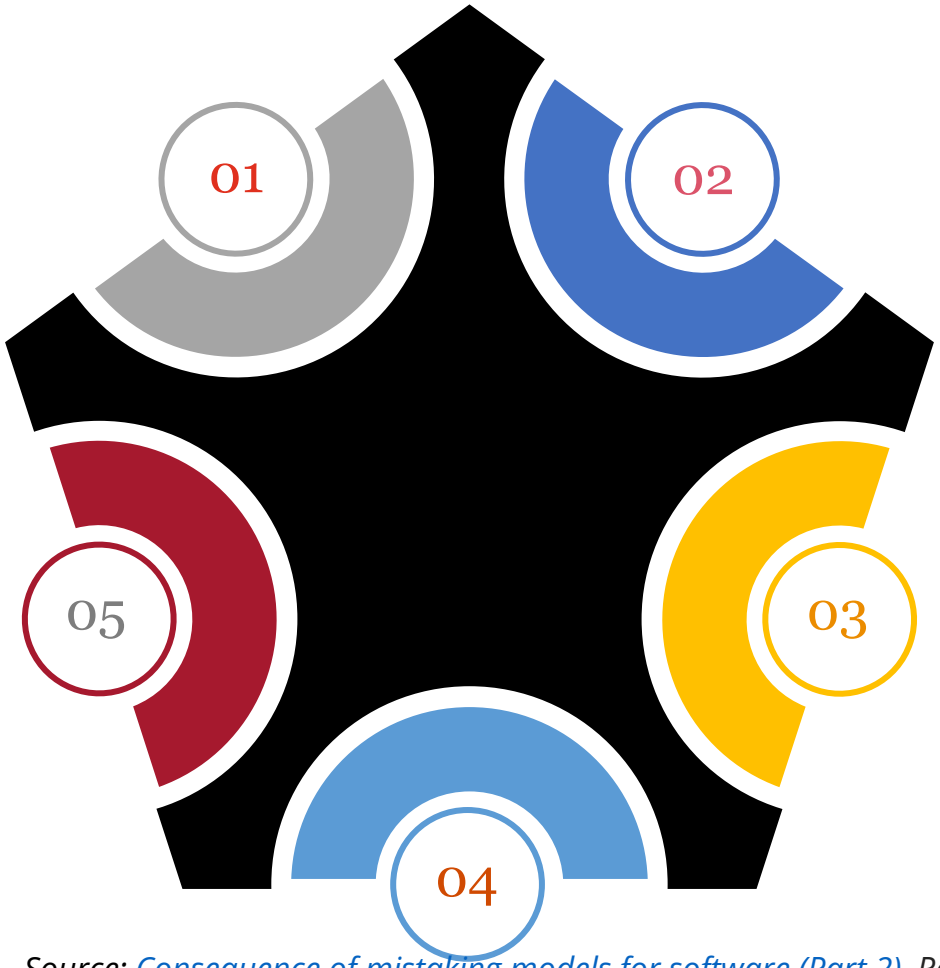
Source: *MLOps: Machine Learning Life Cycle*, ML4Devs

Software vs Models

	Software	VS ML Models
Output	*Deterministic*	*Probabilistic*
Decision space	*Static*	*Dynamic and Ambiguous*
Inference	*Deduction through Code*	*Induction through data*
Development process	*Agile (linear and iterative)*	*Experimentation (test and learn)*
Mindset	*Engineering*	*Scientific*
Development ecosystem	*Mature*	*Evolving*

Source: Data Scientists are from Mars and Software Developers are from Venus (Part 1), Rao, A., Towards Data Science, August 29, 2020

Consequences of mistaking models for code



01 **Data Traps:** Lack of adequate, good quality, labeled data that is unbiased

02 **Scoping Traps:** Time and accuracy estimation before building the model, during training, deployment, and operations

03 **Return Traps:** Estimation and realization of efficiencies and effectiveness of return

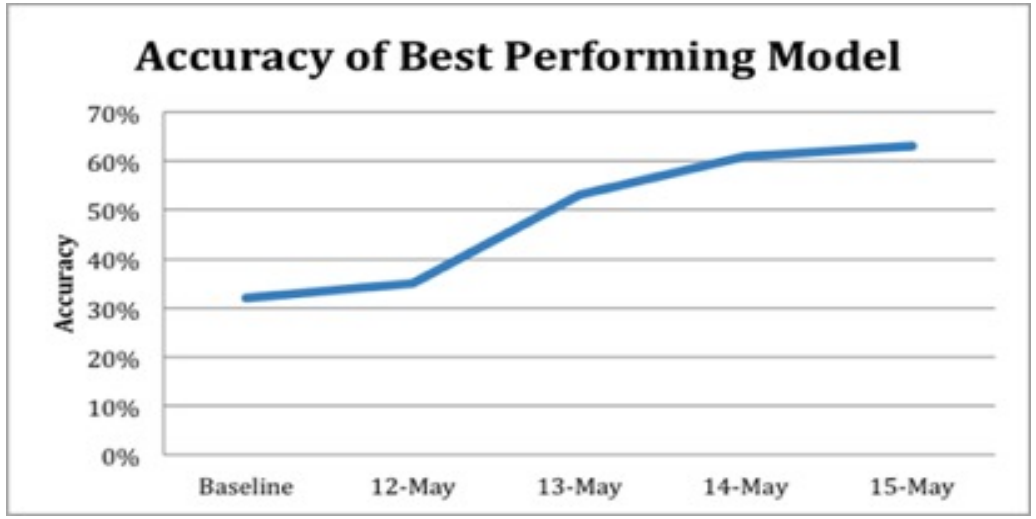
04 **Bias Traps:** Bias in models due to data collection, data cleaning, feature engineering, data split, model training, and model testing

05 **Decay Traps:** Even the most robust models suffer from performance degradation (over time and/or with additional data) in unpredictable ways

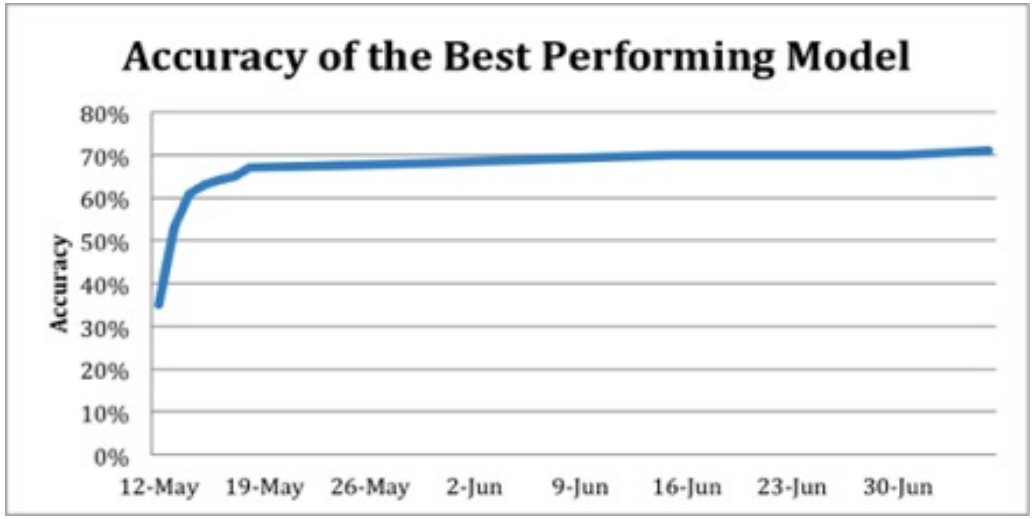
Source: Consequence of mistaking models for software (Part 2), Rao, A., Towards Data Science, September 6, 2020

Scoping traps

Accuracy from 35% to 65% in one week…

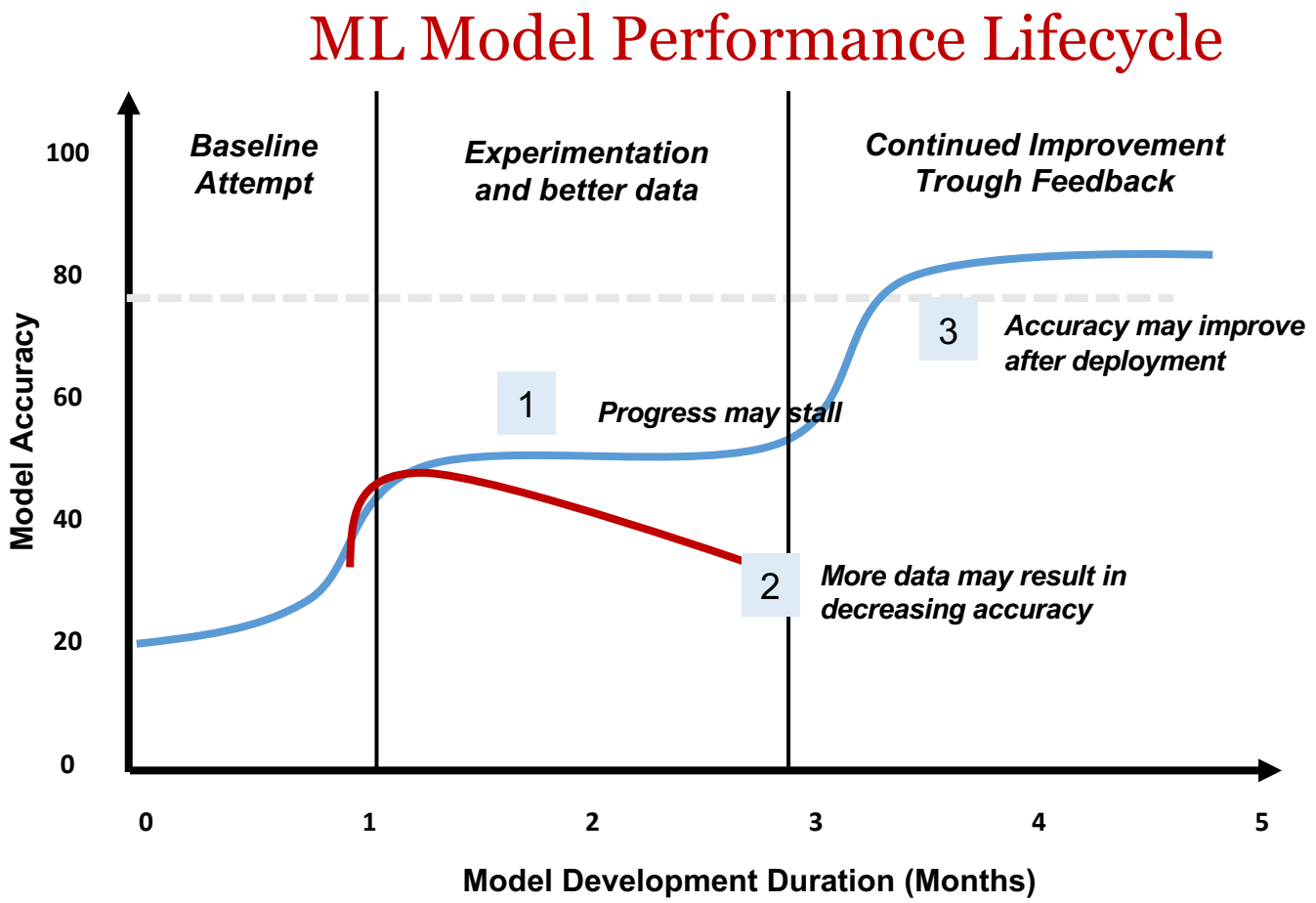


…accuracy didn't move beyond 68% for months



Source: *Why are machine learning projects so hard to manage?, Lukas Biewald, Medium, January 28, 2019*

Scoping traps



Source: *Consequences of mistaking models for software (Part 2).* Anand Rao. Towards Data Science. September 6, 2020.

VALUE SCOPING & ROI OF AI

End-to-end AI Life Cycle: Linear View

Value Scoping

1 Business & Data Understanding

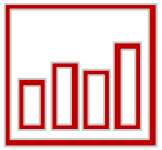
Understand the business challenges; identify and source data, including actual and synthetic

2 Solution Design

Design the solution, select the analytic & AI methods suited for the application & requirements

Value Discovery

3 Data Extraction

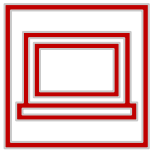


Data preparation, including data selection, cleansing, extraction & imputation

4 Pre-Processing

Iterative feature selection and engineering to create final ML ready dataset

5 Model Building



Build and validate the solution with continuous testing

Value Delivery

6 Model Deployment

Publication of a trained model for broader use

7 Transition & Execution

Implementation into business process and workflows; evangelization

Value Stewardship

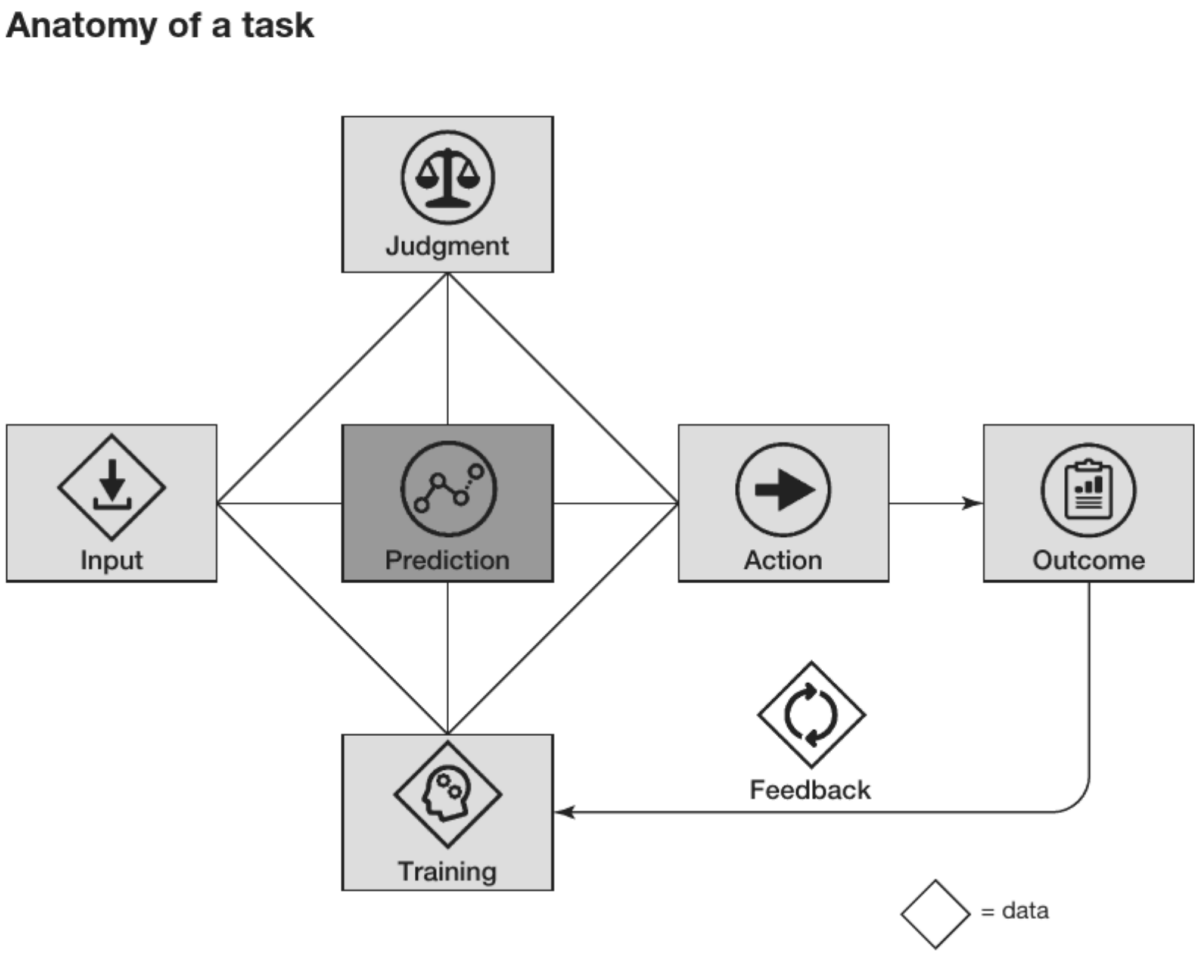
8 Ongoing Monitoring

Ongoing monitoring of outcomes for continuous observation & auditing

9 Evaluation & Check-in

Evaluation of insights & actions against business objectives

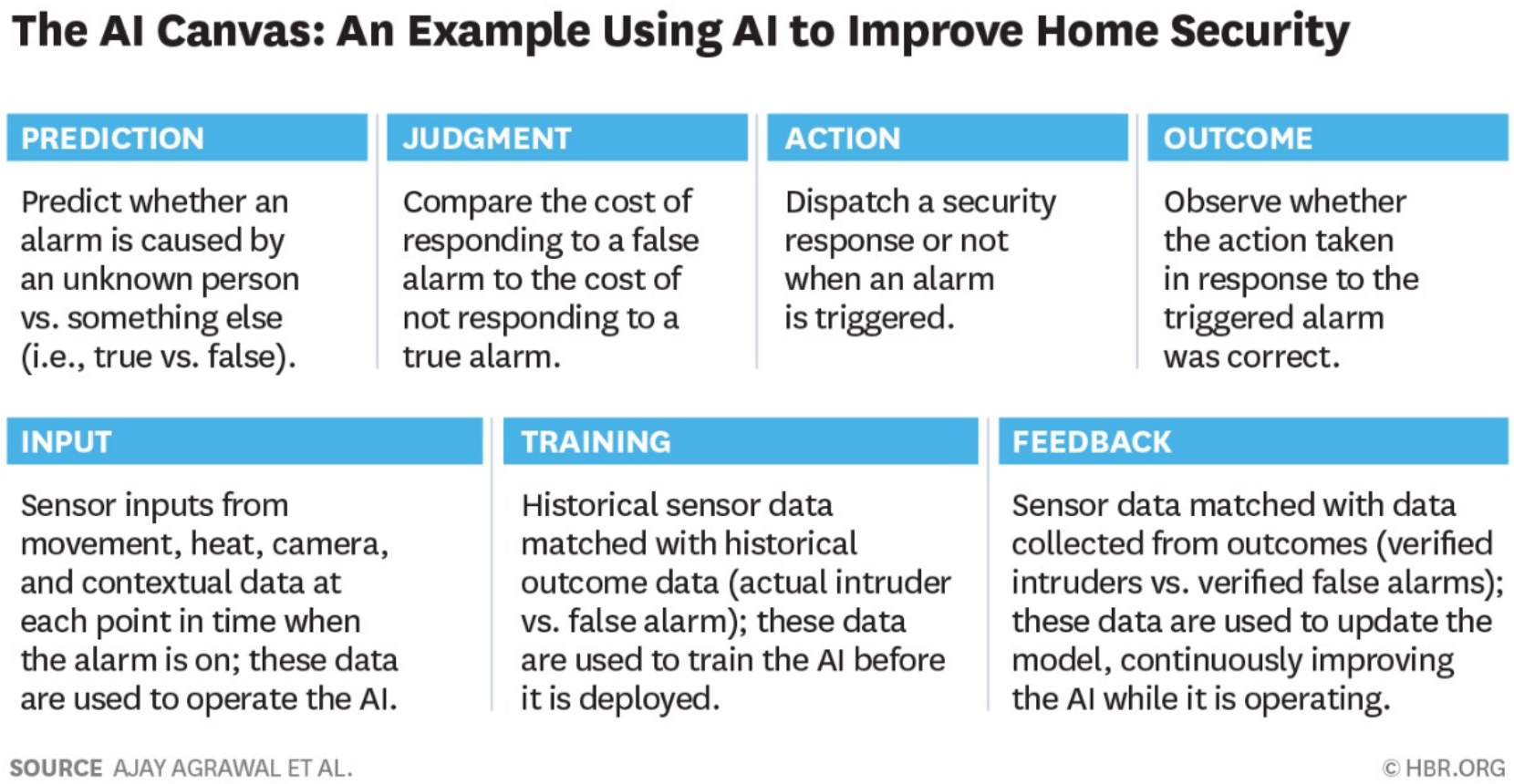
Business Decision Making using AI



- Marginal cost of making predictions are decreasing
- More tasks will be converted into predictions
- Raise the cost of related complements – data, judgement, action
- Diminish the value of substituted – human judgement

Source: Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction machines : The simple economics of artificial intelligence*. Harvard Business Review Press.

AI Canvas: An example

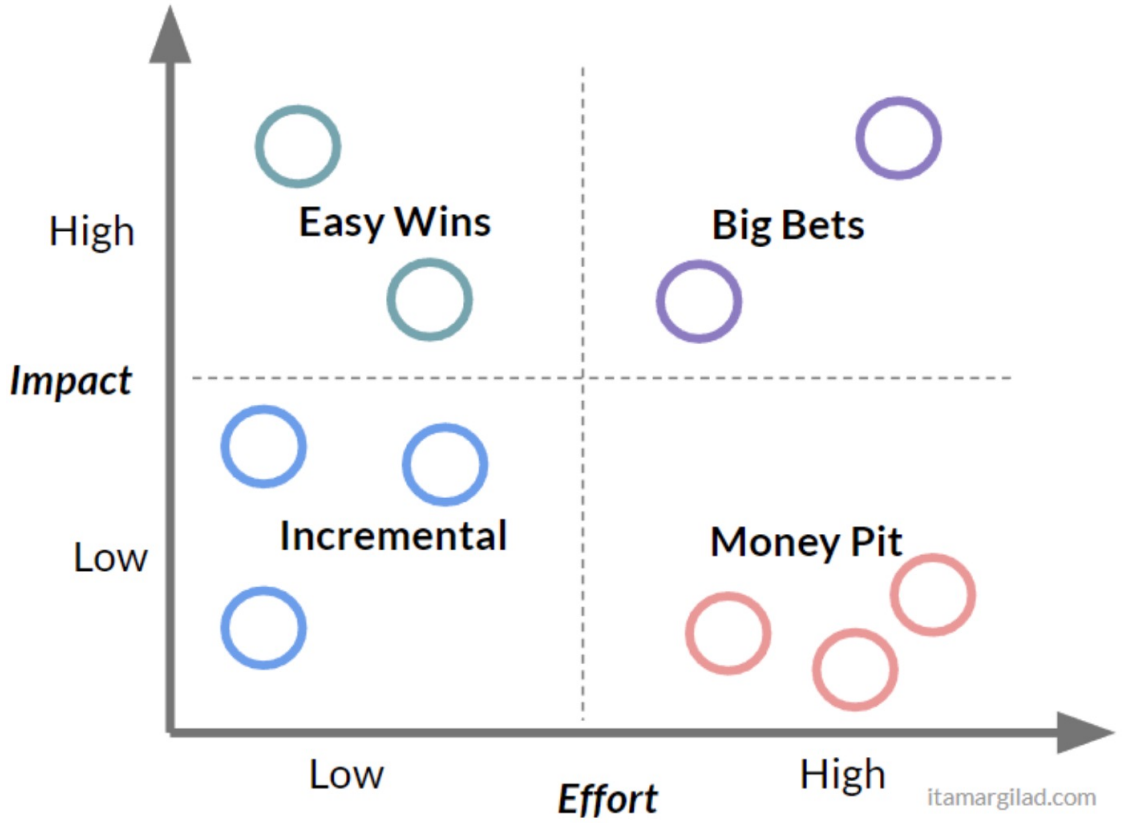


Source: A Simple Tool to Start Making Decisions with the Help of AI by Agrawal, A., Gans, J., & Goldfarb, A. Harvard Business Review, April 2018.

Developing a business case

- Use case approach
  - For each business unit, for each functional area (e.g., marketing, operations etc), list the use cases
  - Prioritize the use cases to develop a roadmap
  - Estimate costs and benefits per use case over different time periods – POC phase; AI system deployment, and ongoing maintenance
- Benefits approach
  - Based on whether AI is automating tasks or making better decisions estimate the benefits and associated costs for those initiatives
- Capability approach
  - Select specific capabilities to build the skills (e.g., image processing) and then deliver multiple use cases for multiple areas

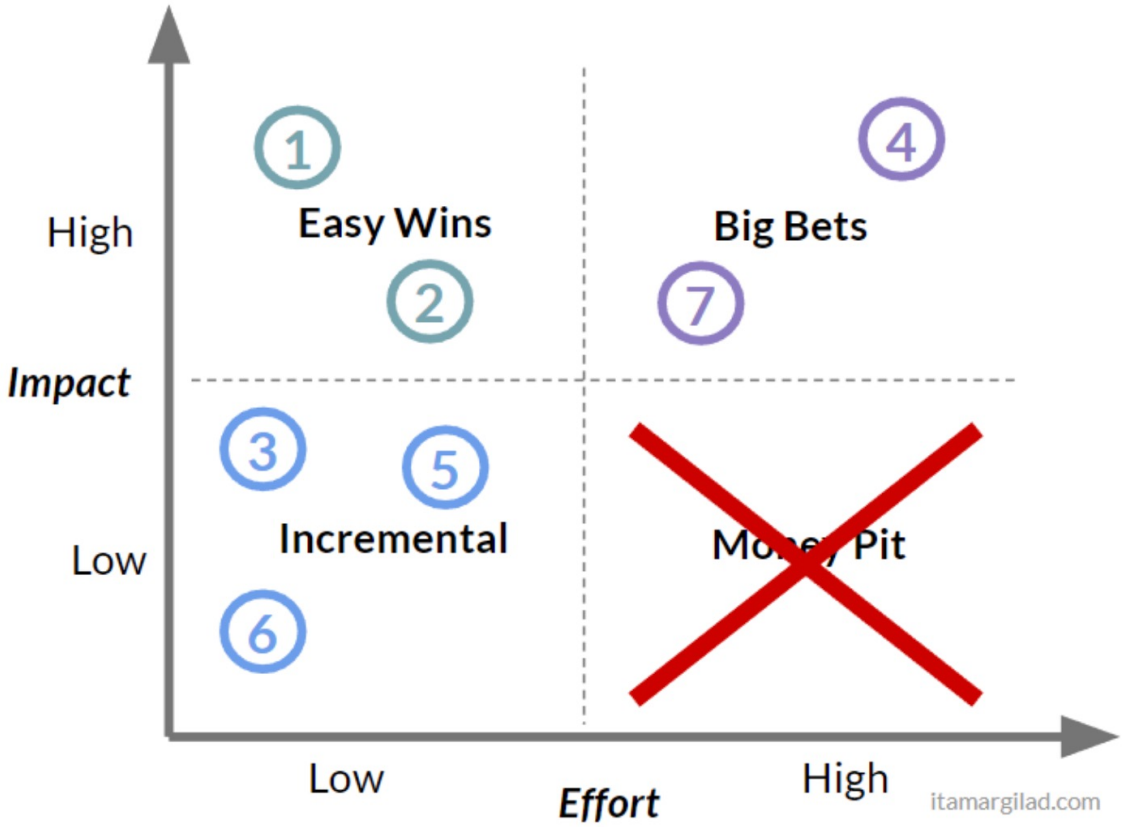
Use Case Analysis – Impact-Effort Matrix



Prioritization with the Impact/Effort Matrix

Source: *Why The Impact Effort Prioritization Matrix Doesn't Work* by Itamar Gilad in High-Impact Product Management.

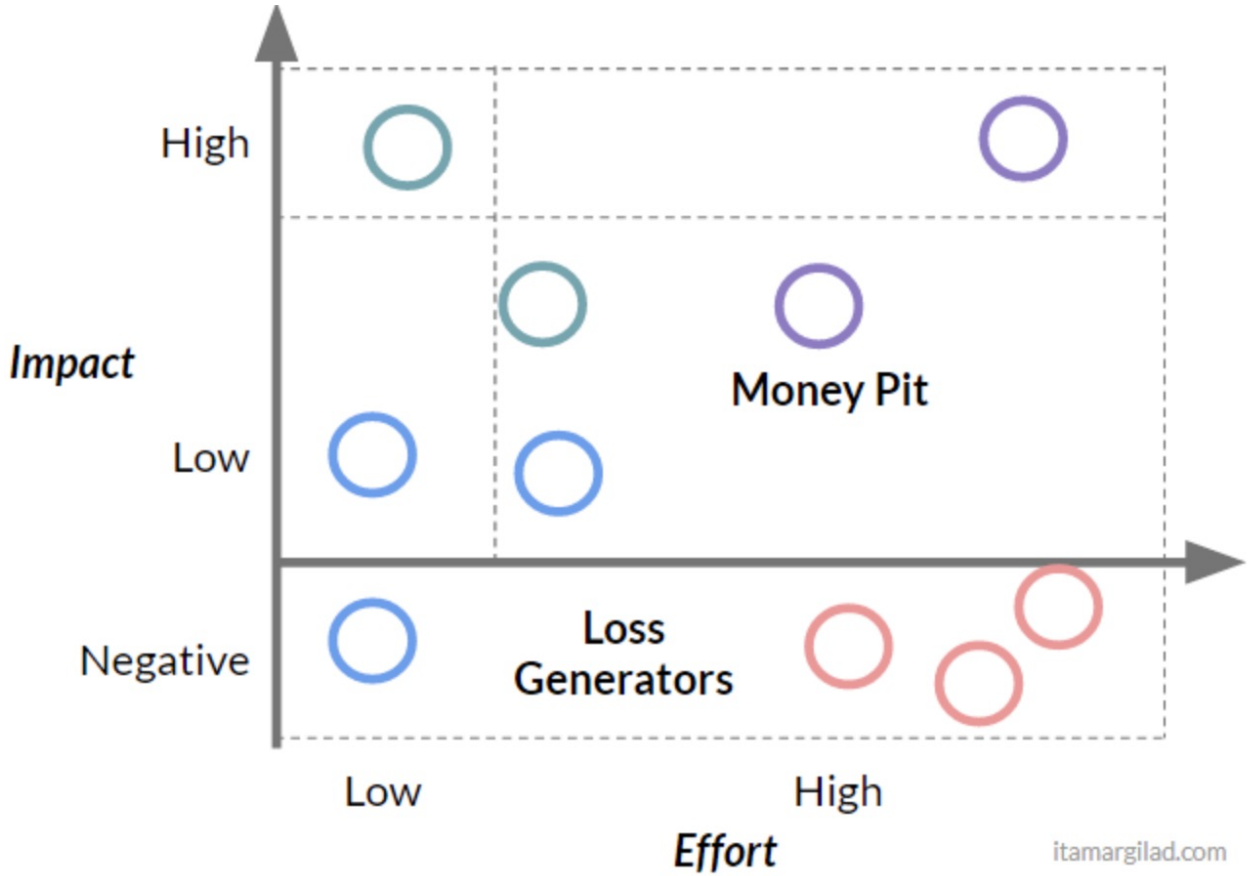
Use Case Analysis – Impact-Effort Matrix



Prioritization with the Impact/Effort Matrix

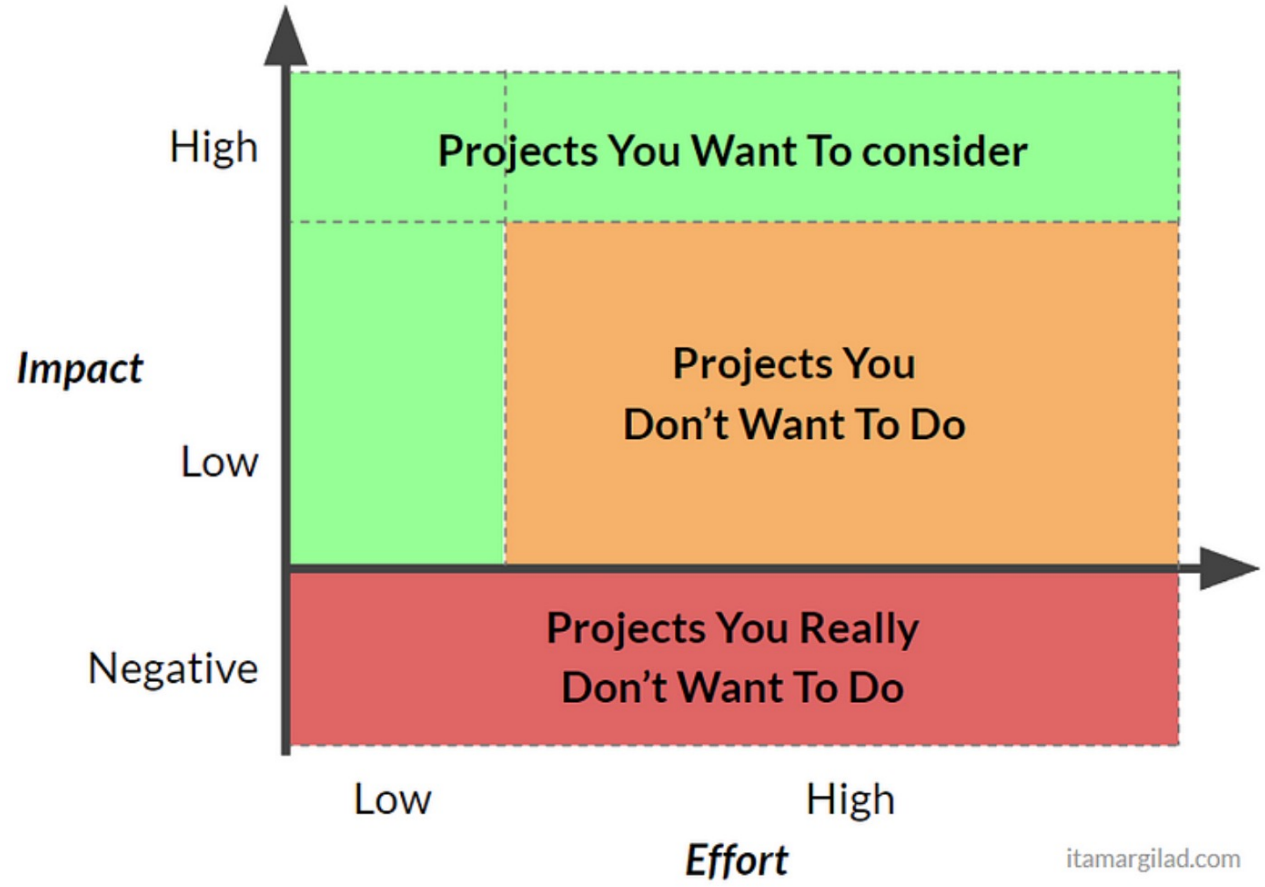
Source: *Why The Impact Effort Prioritization Matrix Doesn't Work* by Itamar Gilad in High-Impact Product Management.

Use Case Analysis – Impact-Effort Matrix



Source: *Why The Impact Effort Prioritization Matrix Doesn't Work* by Itamar Gilad in High-Impact Product Management.

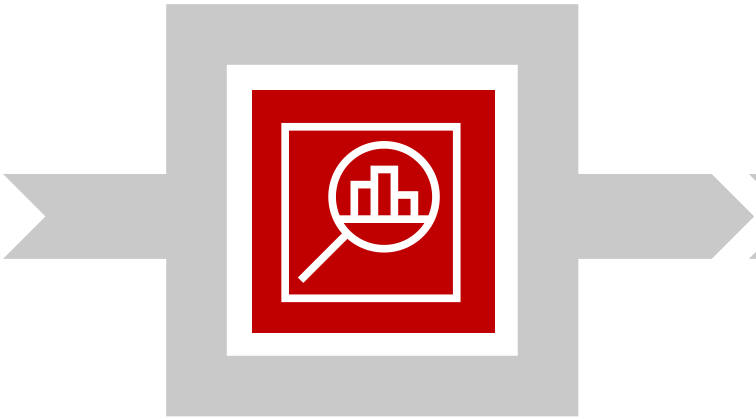
Use Case Analysis – Impact-Effort Matrix



Source: *Why The Impact Effort Prioritization Matrix Doesn't Work* by Itamar Gilad in High-Impact Product Management.

The ROI Challenge of AI

Measurement Challenge



- What do I measure in Returns & Investments?
- How to quantify ROI in AI?

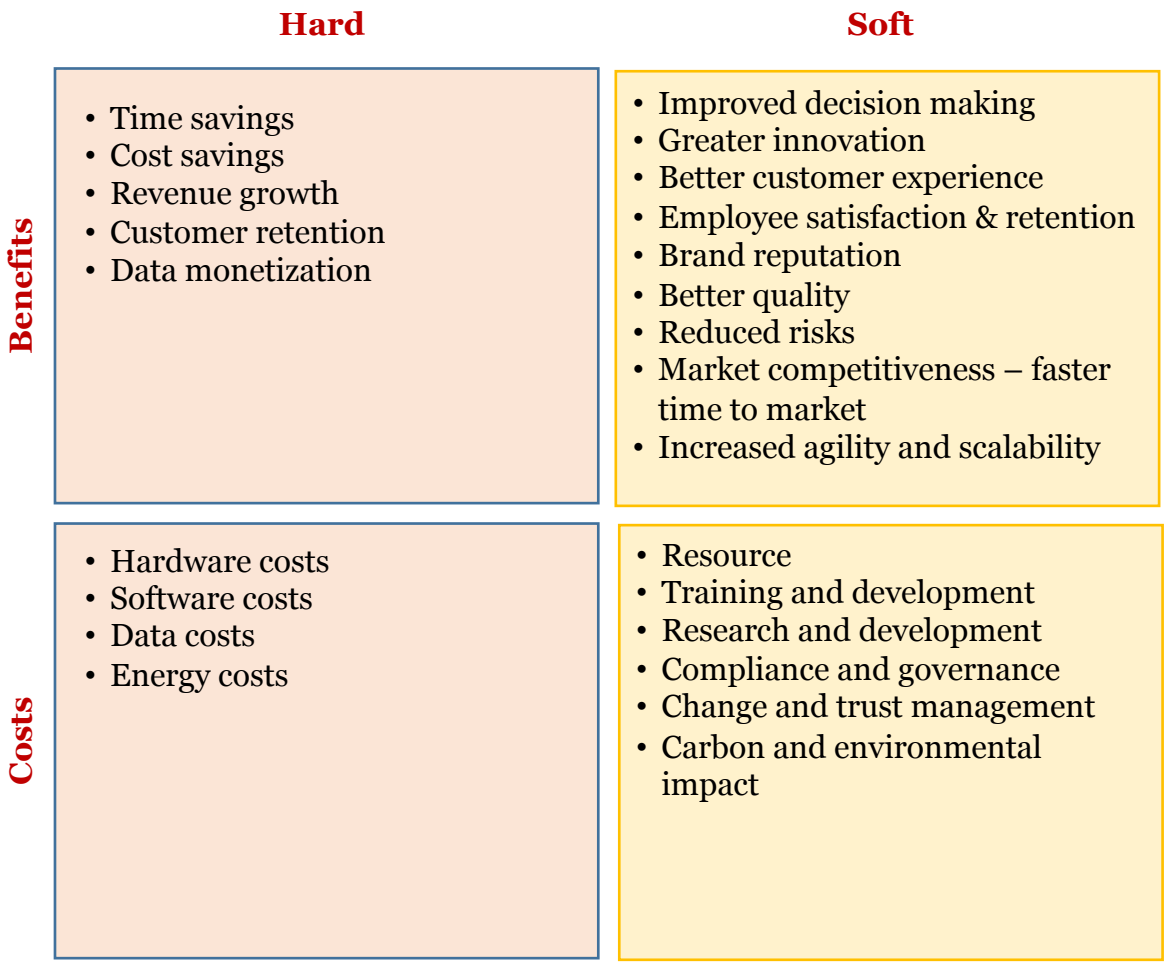
Realization Challenge

- How do I deploy AI across the organization to realize the ROI?

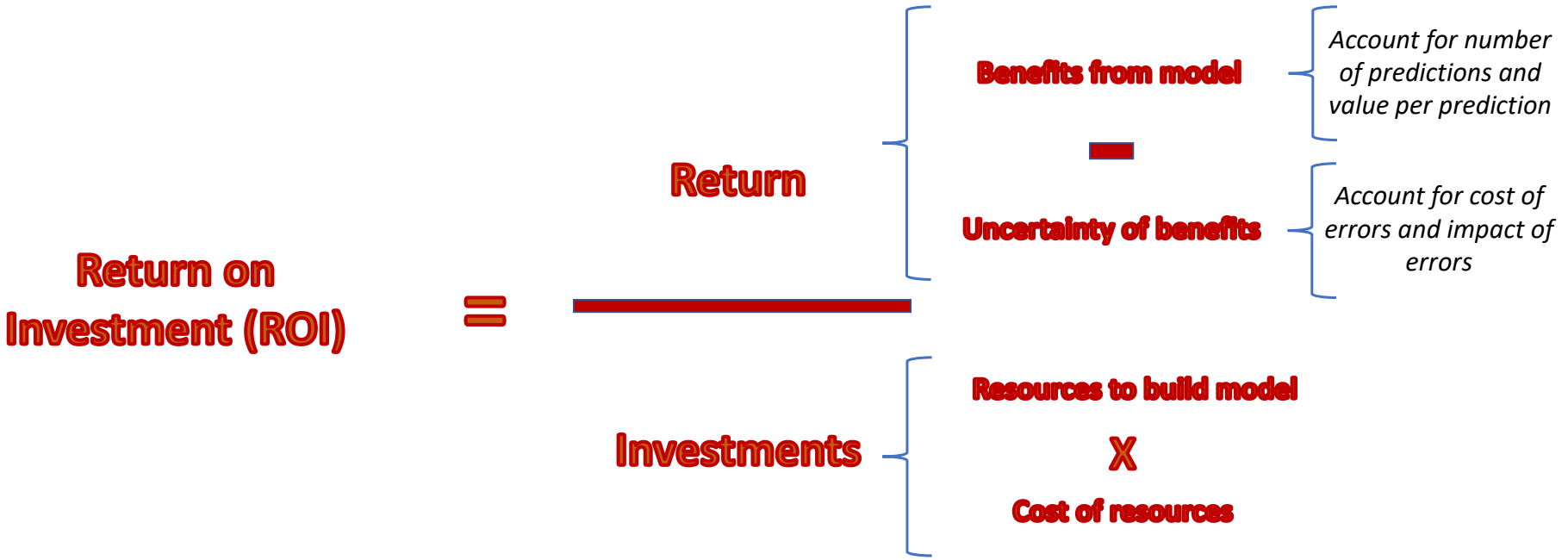
Maintenance Challenge

- How do I ensure that my ROI is sustainable and build long-term trust between AI and humans?

Business Case

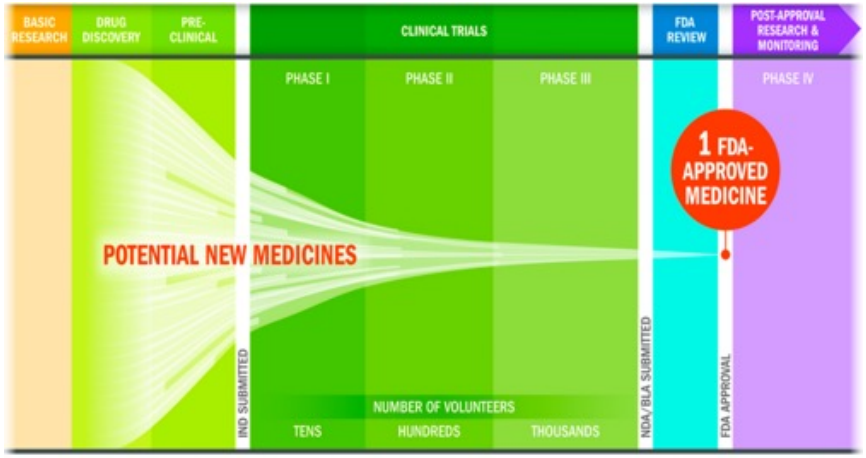


ROI of AI for Predictions

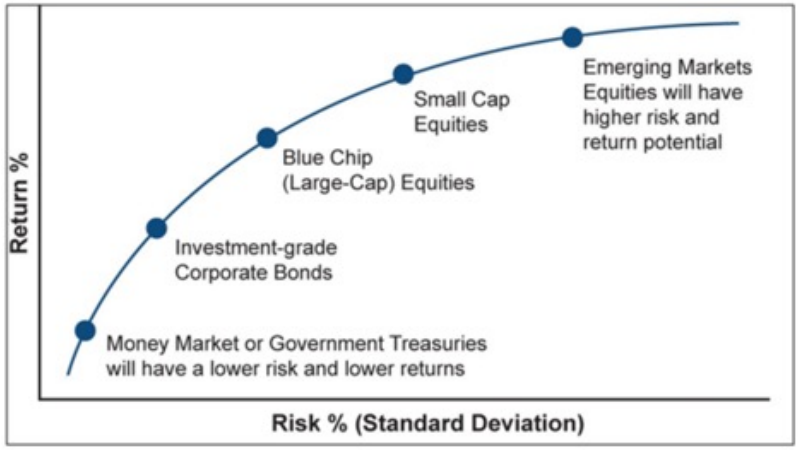


Portfolio Approach to ROI

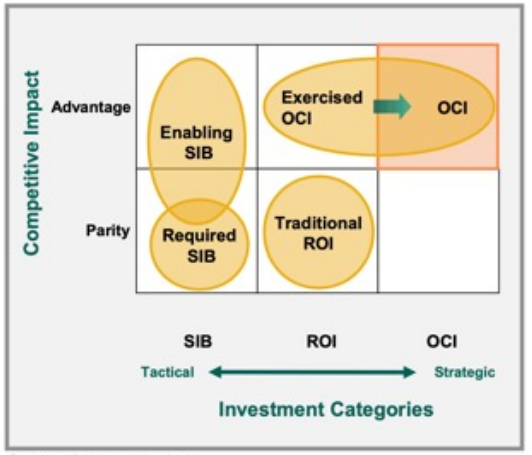
Pharma Portfolio: Drug discovery to FDA approval – 10 years, $2.6bn, 12% success



Financial Portfolio: Efficient frontier and risk-return trade-off



Project Portfolio: Gaining competitive advantage



Source: Diamond Analysis.

- Treat analytics/AI initiatives as test-and-learn experiments
- Consider risk-return trade-off
- Ensure sufficient initiatives in Stay-in-business, ROI, and Option-creating-initiatives

AI Return-on-Investment (ROI)Canvas

Contact: anandr2@Andrew.cmu.edu

Objectives
- Strategic alignment
- Purpose of portfolio

Costs
- Tangible
- Intangible

Portfolio of AI Initiatives
- Stay-in-business
- ROI-generating
- Option creating investments

Benefits
- Tangible
- Intangible

Inputs
- Financial
- People
- Data
- Technology
- Time

Time-line/Roadmap
- Time horizon
- Resources

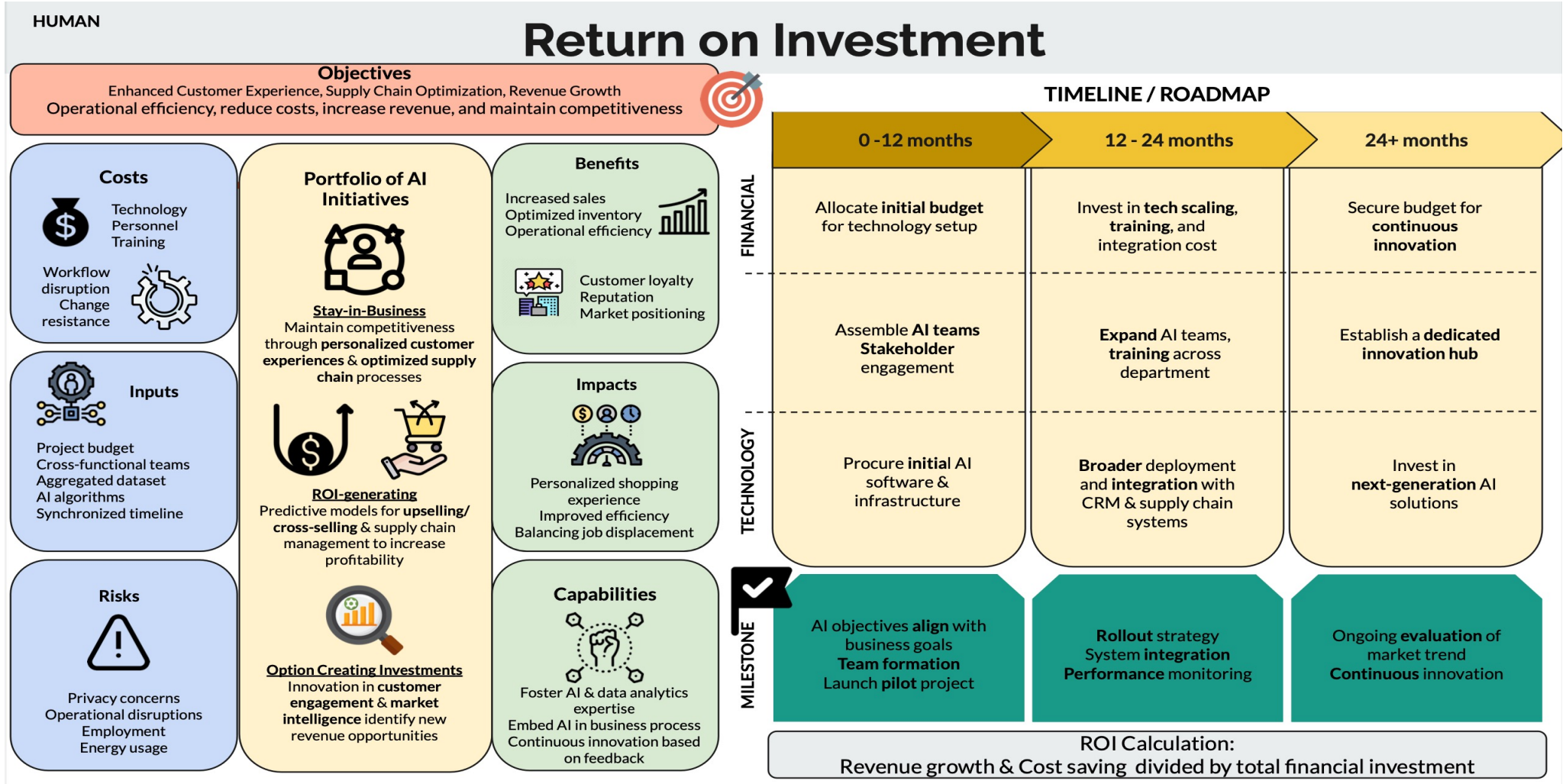
Impact
- Individual
- Enterprise
- Societal

Risks
- Consumer
- Company
- Societal
- Environmental

Capabilities
- Talent
- Data
- Operationalization
- Governance
- Innovation

Portfolio Retun-on-investment
- Assumptions
- Returns

CMU Operationalizing AI – Student Project



VALUE DELIVERY & SCALING AI SYSTEMS

End-to-end AI Life Cycle: Linear View

Value Scoping

1 Business & Data Understanding

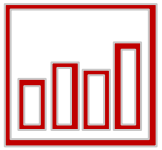
Understand the business challenges; identify and source data, including actual and synthetic

2 Solution Design

Design the solution, select the analytic & AI methods suited for the application & requirements

Value Discovery

3 Data Extraction

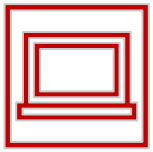


Data preparation, including data selection, cleansing, extraction & imputation

4 Pre-Processing

Iterative feature selection and engineering to create final ML ready dataset

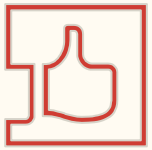
5 Model Building



Build and validate the solution with continuous testing

Value Delivery

6 Model Deployment



Publication of a trained model for broader use

7 Transition & Execution

Implementation into business process and workflows; evangelization

Value Stewardship

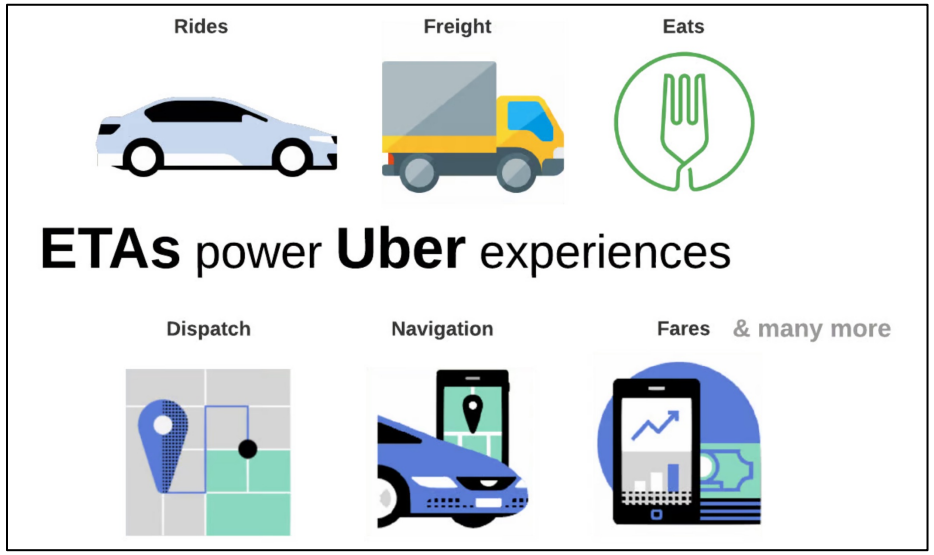
8 Ongoing Monitoring

Ongoing monitoring of outcomes for continuous observation & auditing

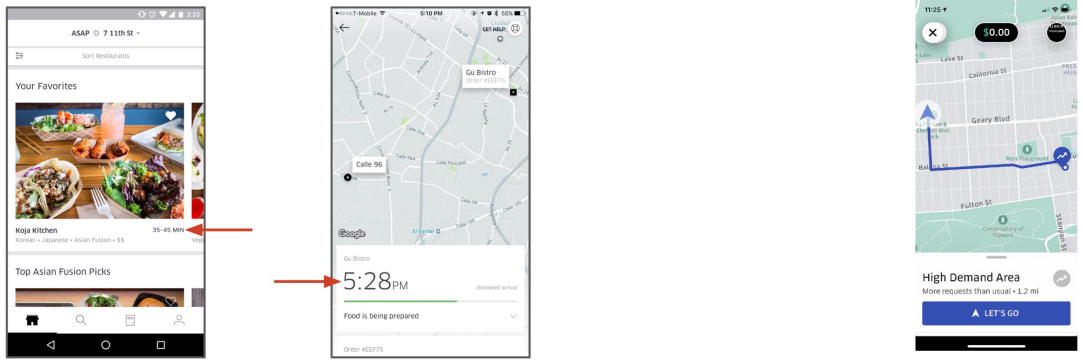
9 Evaluation & Check-in

Evaluation of insights & actions against business objectives

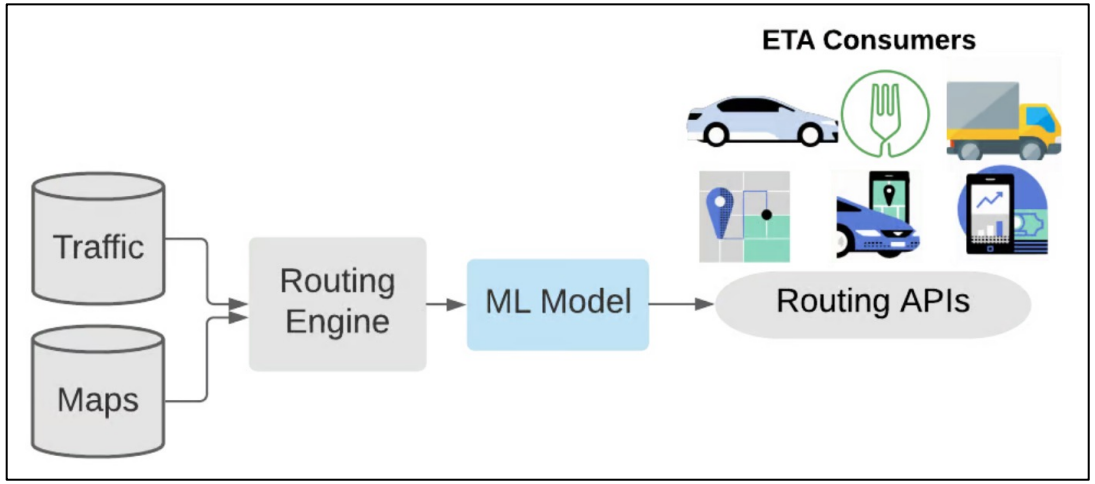
Uber – AI System and AI Models



Interfaces and Systems

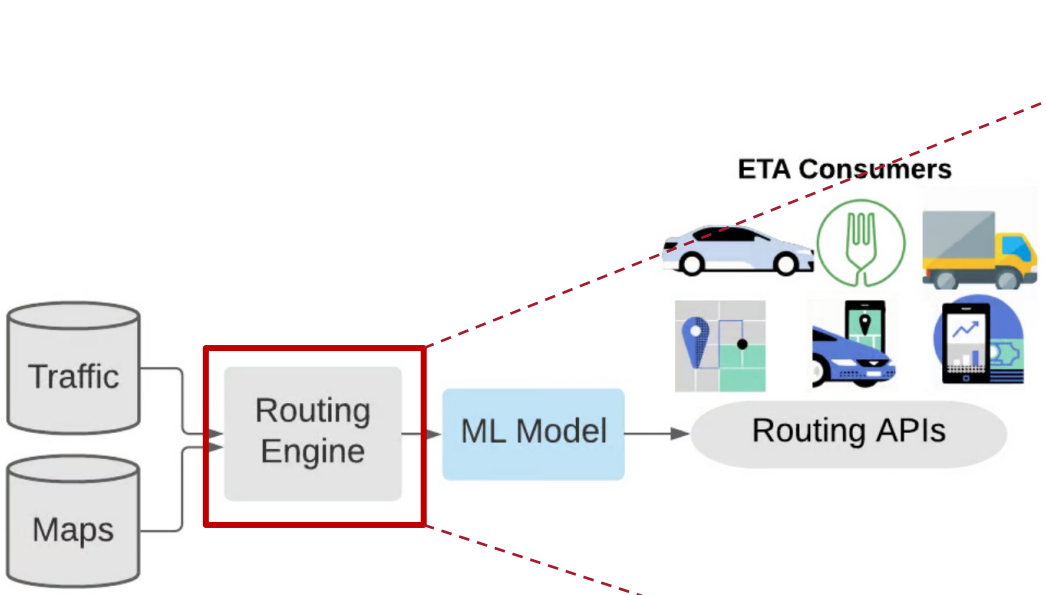


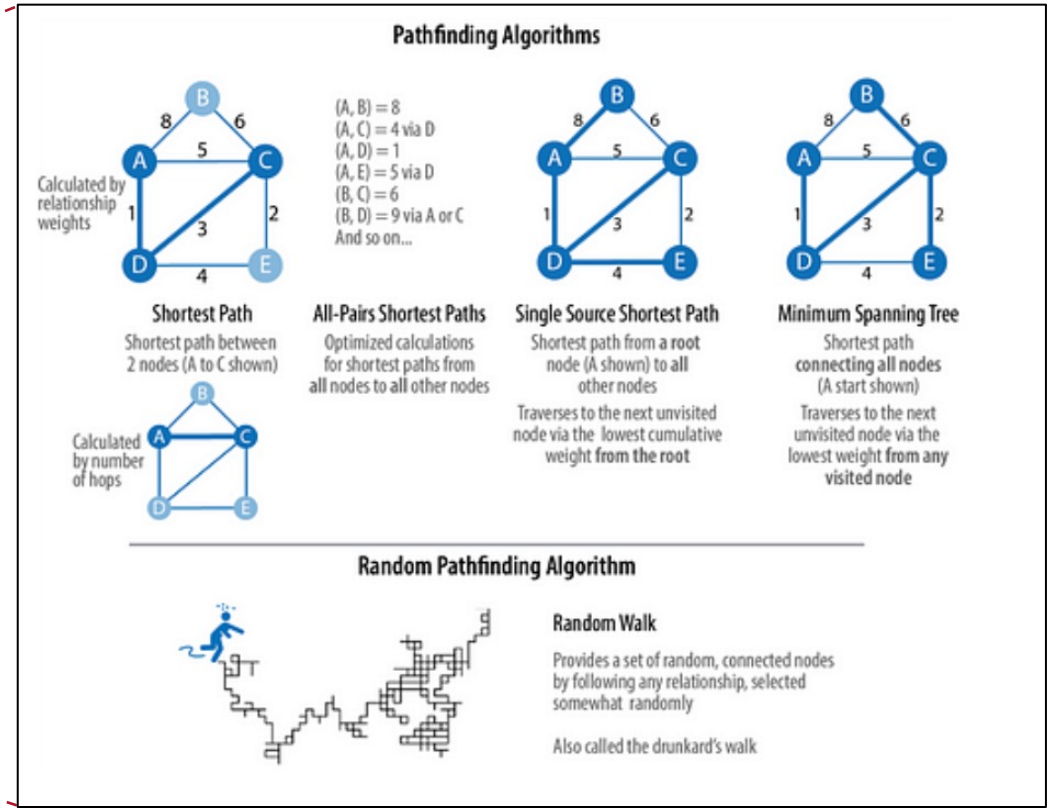
AI System Architecture



Source: [*DeepETA: How Uber Predicts Arrival Times Using Deep Learning*](), Uber Blog, February 10, 2022

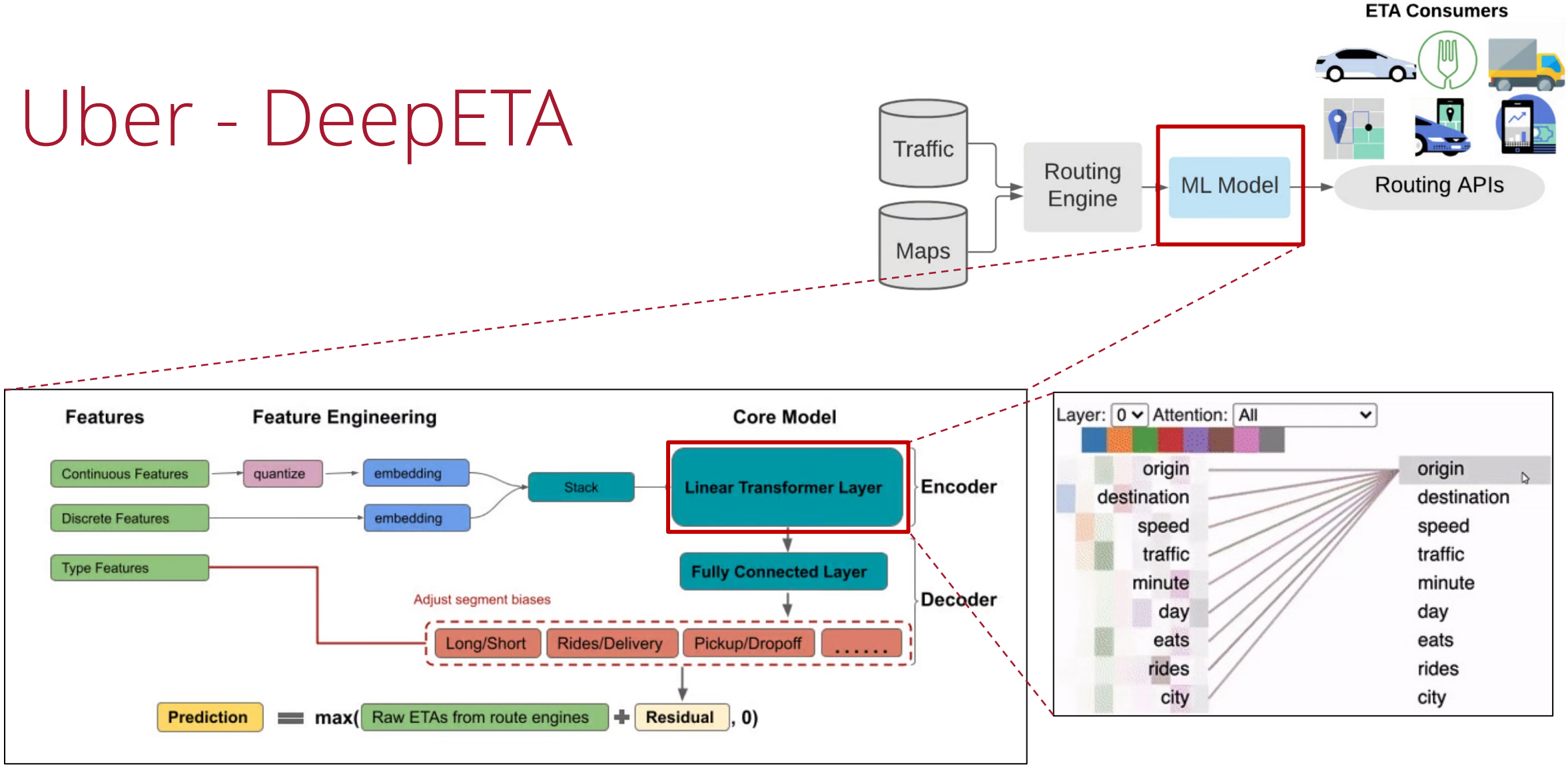
Uber – Routing Engine





Source: [*DeepETA: How Uber Predicts Arrival Times Using Deep Learning*](), Uber Blog, February 10, 2022; [*How Uber uses AI to serve you better*](), Louis Bouchard, May 21, 2022.

Uber - DeepETA



Source: [*DeepETA: How Uber Predicts Arrival Times Using Deep Learning*](), Uber Blog, February 10, 2022

Case Study: DoorDash Journey

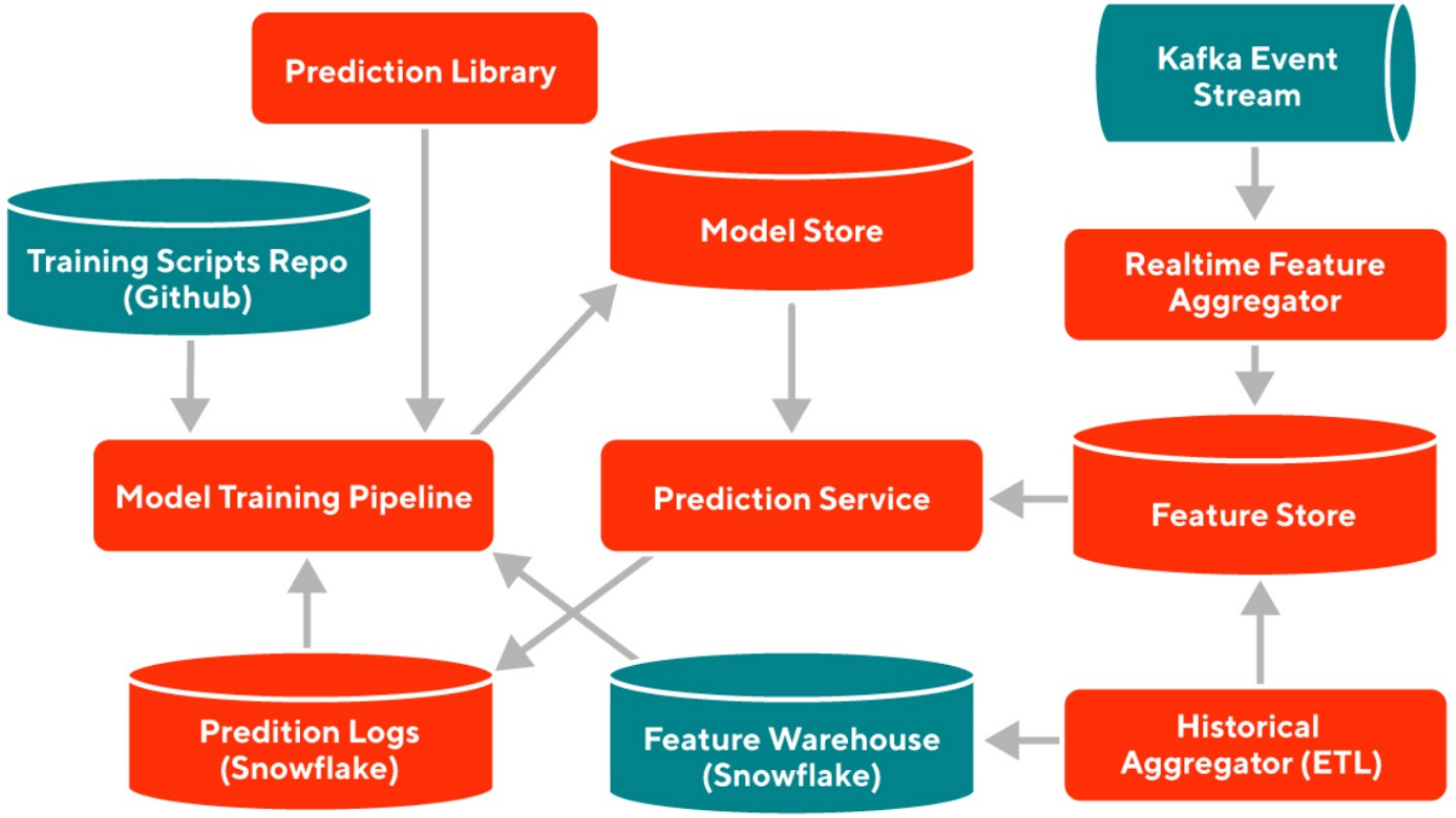
Source: [Scaling Online ML Predictions to Meet DoorDash Growth](), apply() Conference 2021

DoorDash Case Study

	Mar 2020	June 2020	Aug 2020	Sep 2020	Oct 2020	Dec 2020	Jan 2021
# of models	2	16	20	24	28	44	38
Peak predictions per sec	1k	15k	130k	130k+	1M	2M	6.8M

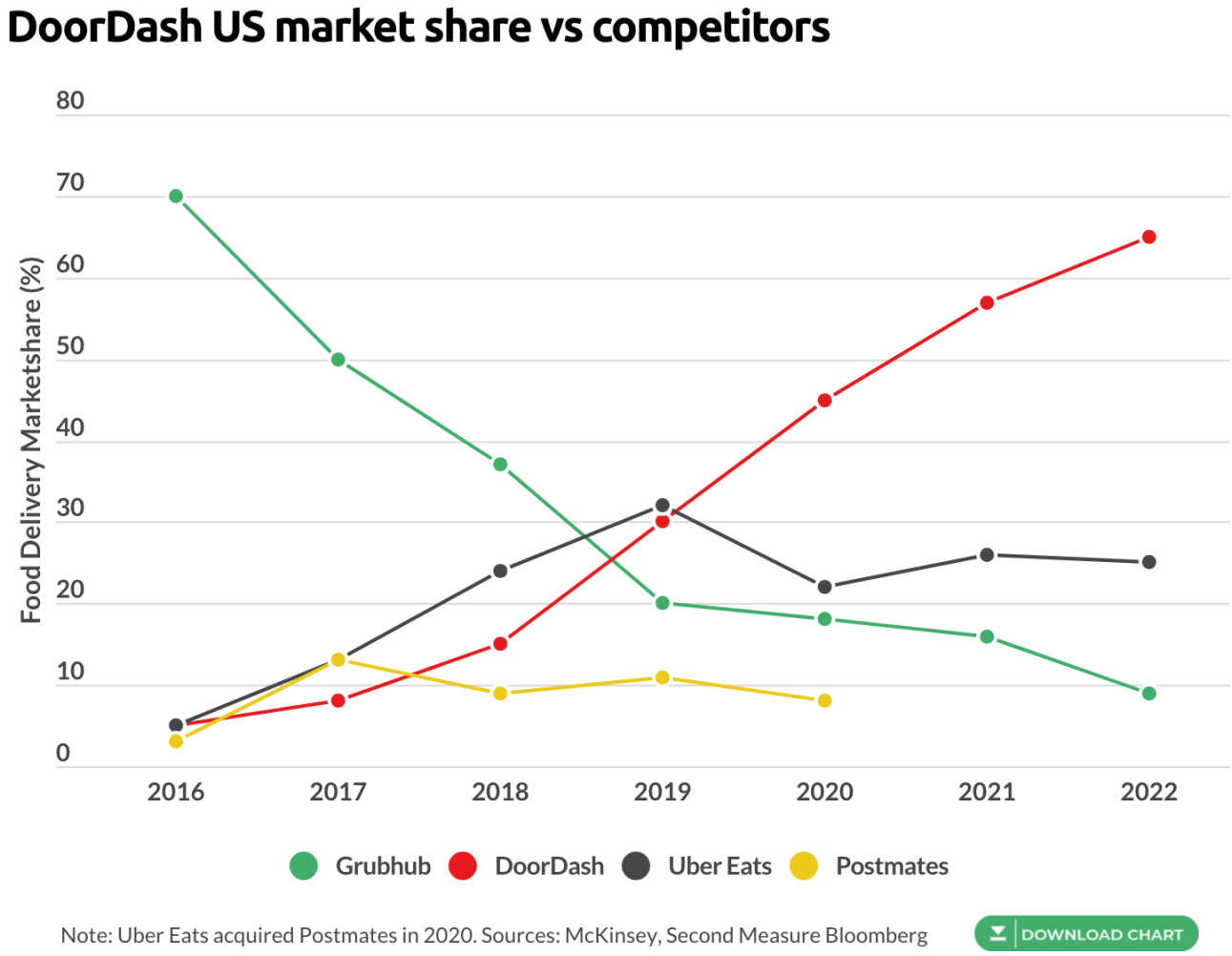
Source: Scaling Online ML Predictions to Meet DoorDash Growth, apply() Conference 2021

DoorDash Case Study



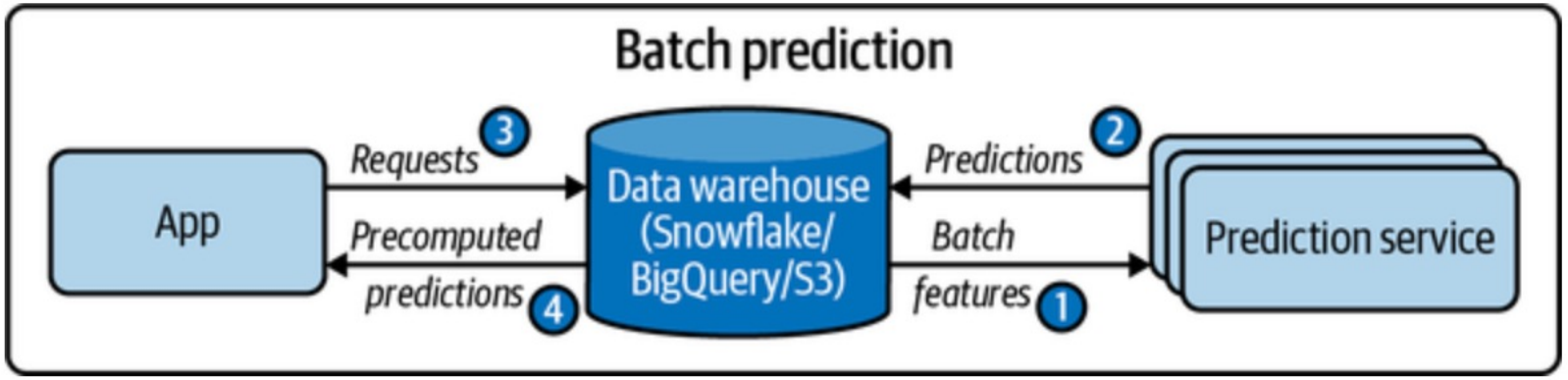
Source: [Scaling Online ML Predictions to Meet DoorDash Growth](), apply() Conference 2021

DoorDash Outcomes



Source: *DoorDash Target Market Segmentation and Marketing Strategy – Audience Demographics & Competitors*, Start.io, August 3, 2022.

Batch prediction with batch features



Uber

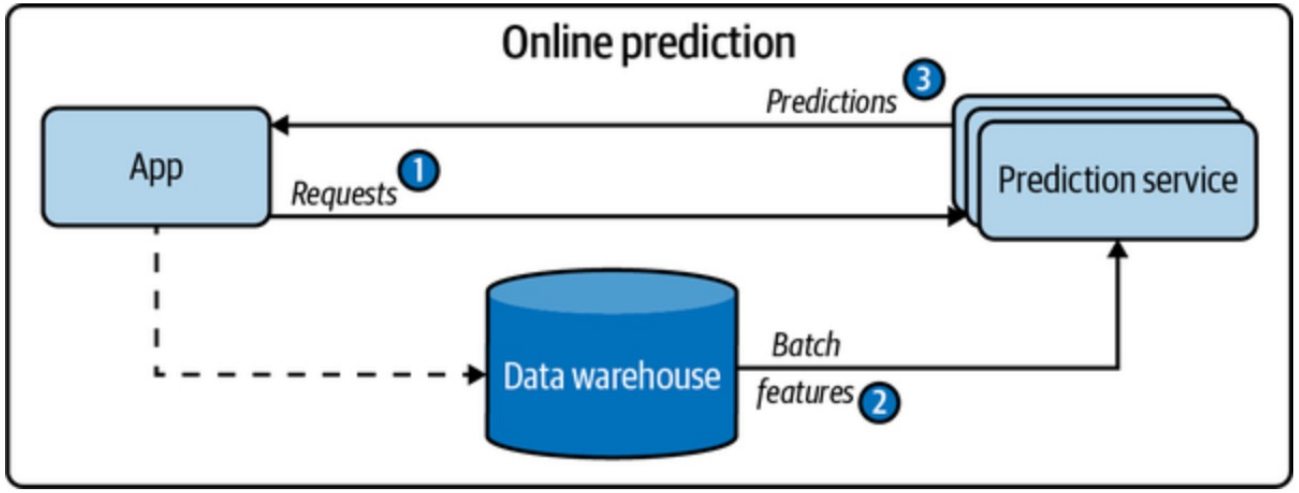
Predicting the demand for rides in a particular city for the next week

NETFLIX

Generating movie recommendations for all users at the start of each day.

Source: [Designing Machine Learning Systems](), Chip Huyen, O'Reilly Media, May 2022.

Online prediction with batch features



Uber

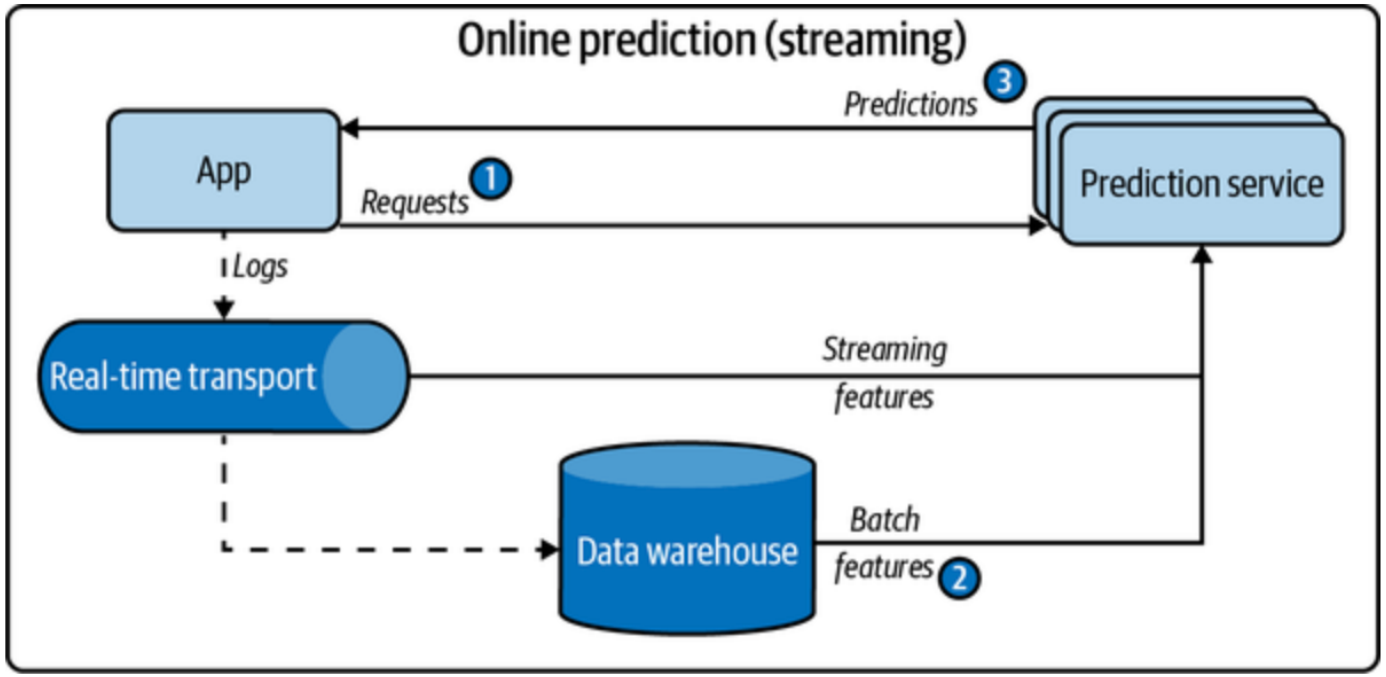
Estimating the price for a ride based on historical pricing data and current demand

NETFLIX

Recommending movies based on a user's historical viewing behavior when they log in

Source: Designing Machine Learning Systems, Chip Huyen, O'Reilly Media, May 2022.

Online prediction with streaming features



Uber

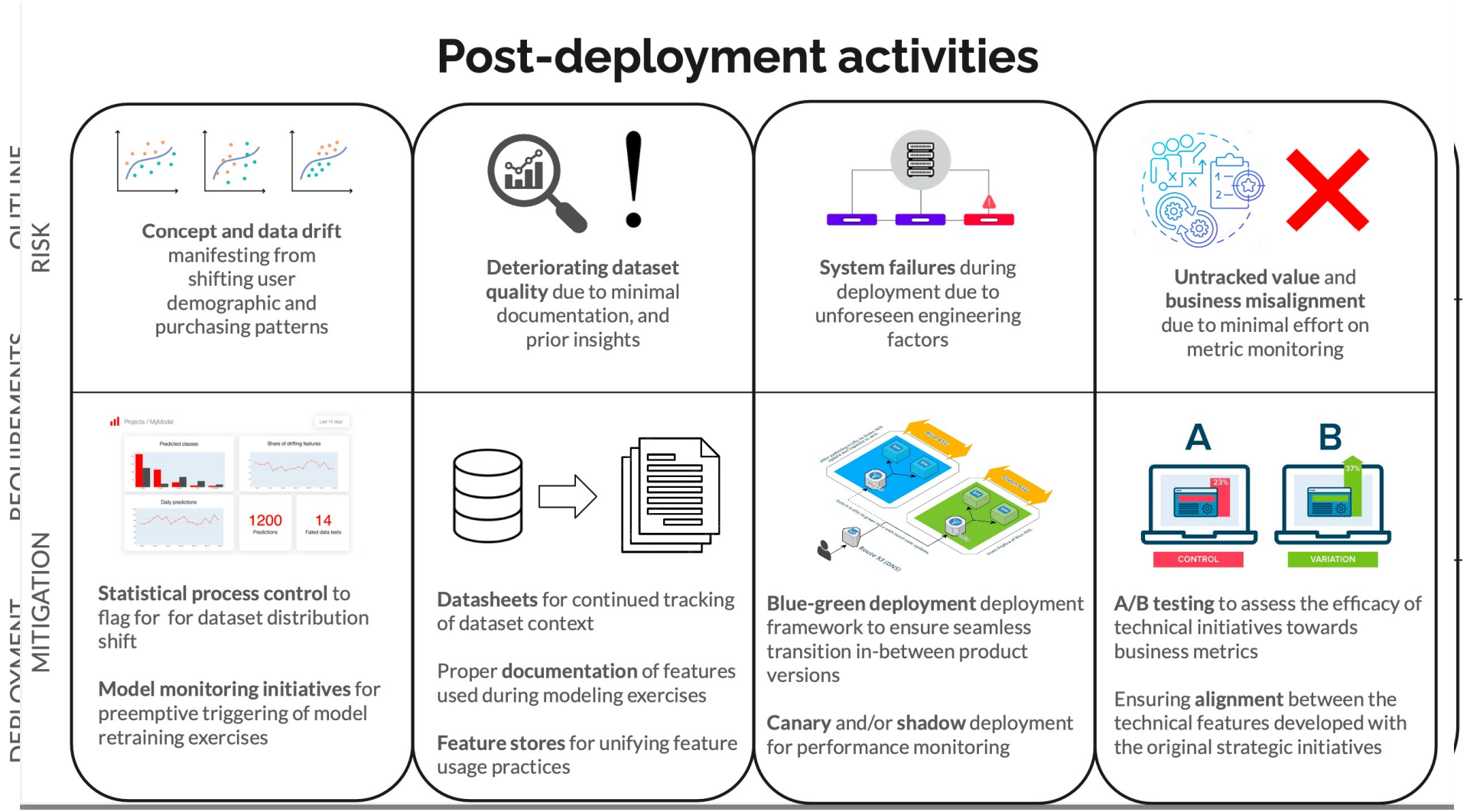
Estimating the arrival time of a ride using real-time traffic data and the driver's current location.

NETFLIX

Adjusting streaming quality in real-time based on current network conditions and user preferences.

Source: Designing Machine Learning Systems, Chip Huyen, O'Reilly Media, May 2022.

CMU Operationalizing AI – Student Project



VALUE STEWARDSHIP & CONTINUOUS MONITORING

End-to-end AI Life Cycle: Linear View

Value Scoping

1 Business & Data Understanding

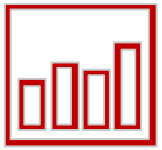
Understand the business challenges; identify and source data, including actual and synthetic

2 Solution Design

Design the solution, select the analytic & AI methods suited for the application & requirements

Value Discovery

3 Data Extraction

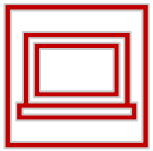


Data preparation, including data selection, cleansing, extraction & imputation

4 Pre-Processing

Iterative feature selection and engineering to create final ML ready dataset

5 Model Building



Build and validate the solution with continuous testing

Value Delivery

6 Model Deployment

Publication of a trained model for broader use

7 Transition & Execution

Implementation into business process and workflows; evangelization

Value Stewardship

8 Ongoing Monitoring

Ongoing monitoring of outcomes for continuous observation & auditing

9 Evaluation & Check-in

Evaluation of insights & actions against business objectives

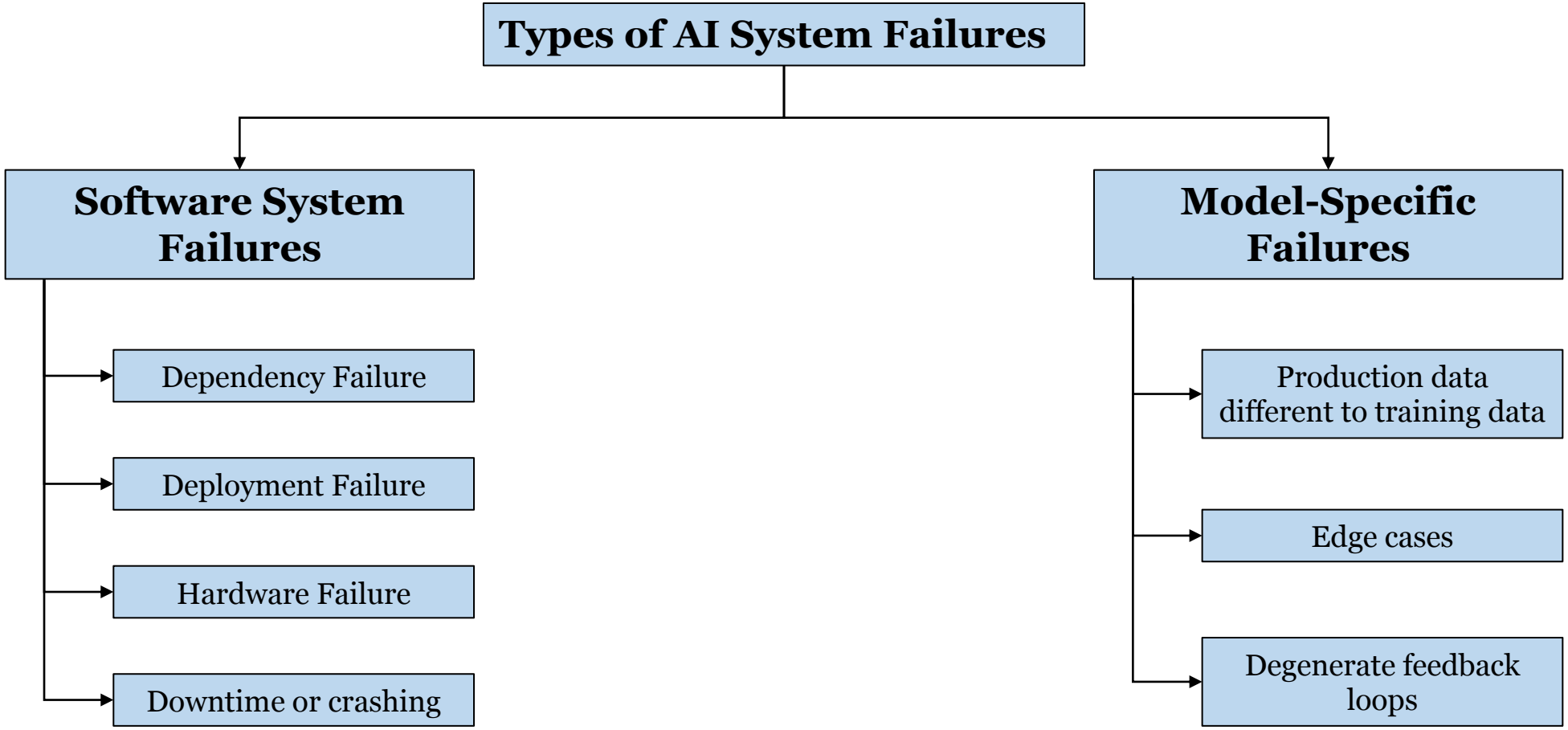
Why monitor?

Systems Fail

Models Decay

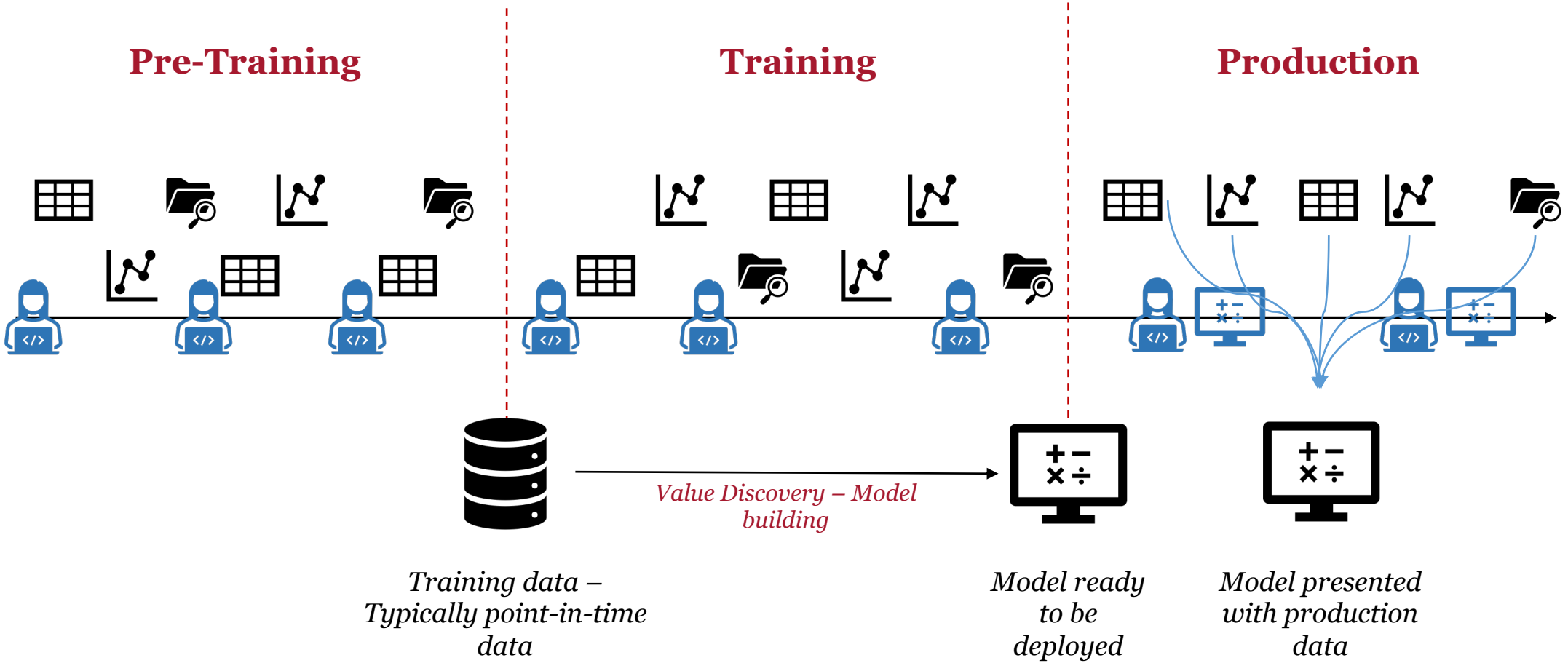
Source: OpenAI. (2023). *ChatGPT* [Large language model]. /g/g-2fkFE8rbu-dall-e

Types of AI System Failures

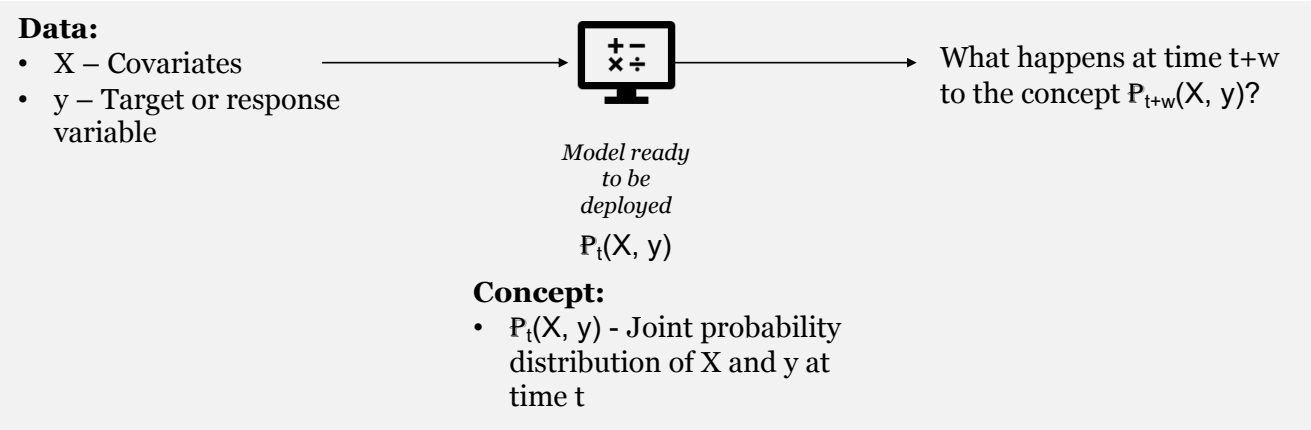


Source: *Designing Machine Learning Systems by Chip Huyen, O'Reilly Media Inc., June 28, 2022.*

Training data vs Production Data



Three Major Classifications of Drift



1. Underlying data distribution changes

$$P_t(X) \neq P_{t+w}(X)$$

2. Underlying concept being learnt changes

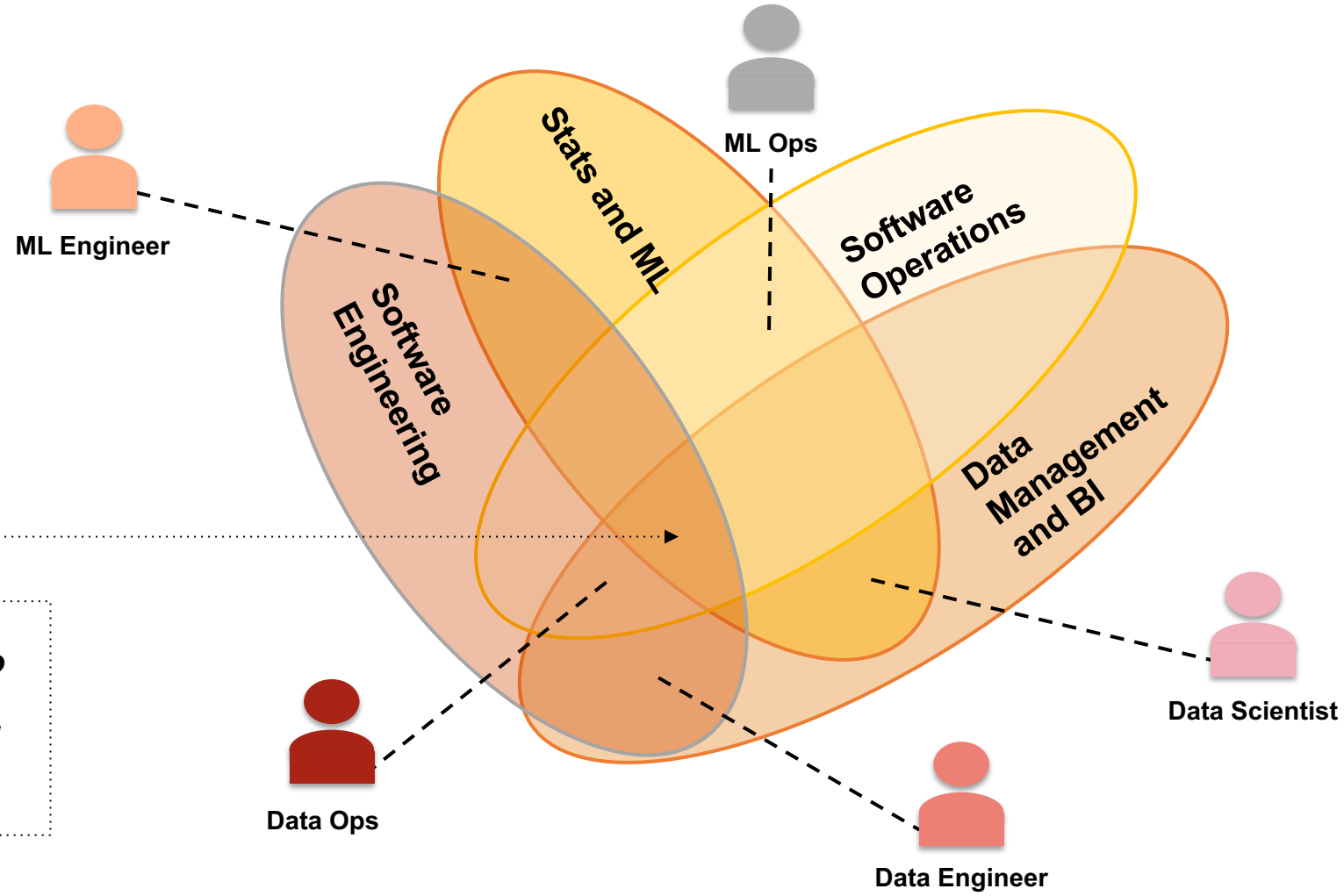
$$P_t(y|X) \neq P_{t+w}(y|X)$$

3. Underlying target or response variable changes

$$P_t(y) \neq P_{t+w}(y)$$

Source: *From concept drift to model degradation: An overview on performance-aware drift detectors*. Bayram., et. Al., Knowledge-Based Systems 245, 2022.

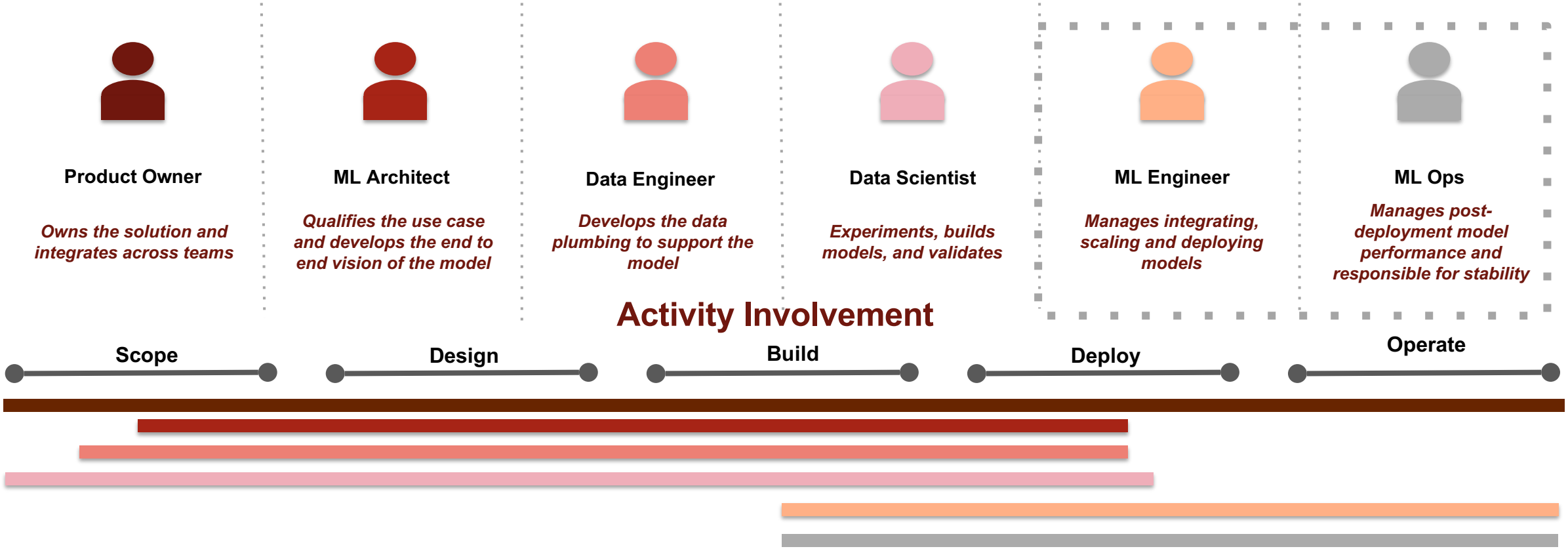
Who should Operationalize AI: New Roles



Superset of skills needed to build and deploy entire production ML training and inference pipeline

Who should Operationalize AI: Process

New roles and skills are required to operationalize the end-to-end DMC Life Cycle



Core Operational Roles

Role/Description	Product Owner	ML Architect	Data Engineer	Data Scientist	ML Engineer	ML Ops
	Owns the solution and integrates across teams	***Qualifies the use case and develops the end to end vision of the model***	***Develops the data plumbing to support the model***	***Experiments, builds models, and validates***	***Manages integrating, scaling and deploying models***	***Manages post-deployment model performance and responsible for stability***
Skills	• Project Management • Communication • Analytical Thinking • Model Development • Domain Knowledge	• Computer science, data science or AI-related background • Enterprise AI / ML development • Programming capabilities in multiple languages (e.g. Python, R, Scala, Octave) • Statistical analytics and modeling	• Data Architecture • Data management • Systems engineering • Distributed Compute • Software development • Cloud compute	• Model experimentation and selection • Data preparation and manipulation tools and DBs • Mathematics, engineering and/or statistics background • Analytically minded	• Software development and engineering • Systems admin and infra • ML and statistics • Model development lifecycle • Distributed and/or GPU compute environments	• Architect, deploy and maintain ML • ML on cloud • Agile methodologies and DevOps • API architecture and container orchestration (Kubernetes) • Auto-scaling ML and big data
Outputs	• User stories • Product backlog and roadmap • Status update templates and reporting • Communications to leadership, development team, and other stakeholders	• Model predictional outputs • ML end-to-end architecture • Model registry/inventory and services library • Model dashboards and graphical analysis outputs	• ETL, integration, collection, and cleansing scripts for generating data sets across multiple data stores • Data model or schema to support ongoing analysis / model • Scalable data pipelines for modelling, BI or analytics related tasks • Visualizations	• PoC model that best serves decisions, forecast, or desired predictions • Documented workbooks outlining modelling process, model selection, experimentation, and logs • Rough sketch pipelines for inference along with retraining and validation • Visualizations and analysis used for analysis and model development	• Production pipelines for train and inferences • Scripts for monitoring production model performance • Retraining pipelines along with test and validation code • Scripts and tools for collecting labeled training data	• Infrastructure and platform to support model development lifecycle • Automation pipelines for CI/CD/CT • Monitoring, logging and alerting of model services in production

Strategy, Policy and Experience Roles

	Chief AI Officer	AI Policy Advisor	AI Strategy Consultant	Human-AI Interaction Designer
Role/Description	**Drives the company's overarching AI vision and strategy; champions the adoption of AI across business functions while ensuring alignment with corporate values.**	**Shapes the organization's AI policy agenda; analyzes and interprets the impact of regulatory environments on AI deployment and usage.**	**Partners with business leaders to integrate AI into corporate strategy; ensures AI initiatives align with long-term business objectives and market dynamics.**	**Designs intuitive and ethical interfaces for human-AI interaction; ensures user experiences are inclusive and enhance human capabilities.**
Skills	• Leadership in the development and execution of organization-wide AI strategy. • Ability to foster innovation and integrate AI into various business units. • Capability to articulate the AI vision to stakeholders and secure buy-in. • Insight into emerging AI trends and their strategic implications.	• Expertise in AI policy, regulation, and ethics. • Ability to navigate complex regulatory landscapes and advise on compliance. • Skills in policy drafting, analysis, and advocacy. • Proficiency in stakeholder engagement and consensus building.	• Strategic planning with a focus on leveraging AI for competitive advantage. • Proficiency in market analysis and identifying AI-driven business opportunities. • Aptitude for aligning AI projects with business goals and customer needs. • Strong consulting and stakeholder management abilities.	• Design thinking with a focus on human-centric AI solutions. • Proficiency in user experience (UX) design for AI-powered applications. • Knowledge of inclusive design principles and accessibility standards. • Ability to prototype and test AI interfaces for usability and ethical considerations.
Outputs	• A comprehensive AI strategy that supports and enhances business objectives. • Governance models for AI across the organization. • Leadership in establishing an AI-centric culture • Frameworks for measuring the performance and impact of AI initiatives.	• Policy briefs and guidelines for responsible AI deployment. • Regulatory analysis reports that inform strategic decision-making. • Advocacy strategies for shaping AI policy discussions. • Engagement with policymakers and industry groups to influence AI regulation.	• Strategic roadmaps for AI adoption and scaling within the business. • Business case analyses for AI investments and initiatives. • Alignment of AI initiatives with market trends and consumer insights. • Recommendations for strategic partnerships and collaborations in the AI space.	• User interface (UI) and UX design frameworks for AI applications. • Prototypes and design specifications that prioritize user trust and ease of use. • Usability testing reports and design iteration plans. • Guidelines for ethical and inclusive design in AI-enabled products and services.

Compliance Roles

	AI Compliance Officer	AI Audit & Assurance Professional	Data & AI Privacy Offier
Role/Description	**Ensures AI systems adhere to legal and regulatory requirements; evaluates AI ethical implications; aligns AI strategies with compliance frameworks.**	**Audits AI systems for adherence to internal and external standards; ensures accountability, fairness, and transparency in AI applications.**	**Oversees compliance with data privacy regulations in AI systems; champions data protection and privacy by design within AI initiatives.**
Skills	• Knowledge of international AI regulatory and compliance standards. • Expertise in ethical AI frameworks and best practices. • Proficiency in AI governance and risk management methodologies. • Familiarity with data privacy laws such as GDPR and CCPA.	• Ability to design and execute audit plans for AI systems. • Competence in assessing the transparency and explainability of AI models. • Proficiency with audit software and AI performance tracking tools. • Strong analytical skills for evaluating AI impact and performance.	• In-depth understanding of data privacy laws like GDPR, CCPA, and HIPAA. • Knowledge of data governance practices and AI ethical standards. • Expertise in privacy impact assessments and data protection strategies. • Skills in communicating privacy concepts to technical and non-technical stakeholders.
Outputs	• Compliance reports and risk assessments for AI projects. • Regulatory compliance frameworks and guidelines for AI deployment. • Training programs for AI ethics and compliance across departments. • Policies and procedures to govern the responsible use of AI.	• Detailed audit reports outlining compliance with AI ethical standards. • Assurance frameworks for AI systems' reliability and performance. • Recommendations for improvements in AI governance structures. • Best practice guides for continuous auditing of AI systems.	• Comprehensive privacy policies tailored for AI data lifecycle management. • Privacy impact assessments for new and existing AI applications. • Data protection training modules for AI development teams. • Incident response plans for potential AI data privacy breaches.

Responsible Practices Roles

	AI Ethicist	AI Risk Manager	AI Trust & Safety Manager
Role/Description	**Guides the ethical development and deployment of AI systems; advises on moral considerations and societal impacts of AI technology.**	**Identifies and mitigates risks in AI deployment; ensures operational integrity and aligns AI initiatives with risk management strategies.**	**Oversees the development and enforcement of trust and safety policies for AI systems; ensures AI applications are transparent, explainable, and socially responsible.**
Skills	• Expert knowledge in ethical theory and its application to AI. • Ability to analyze and interpret the social implications of AI technologies. • Strong communication skills for ethical guideline development and advocacy. • Proficiency in cross-functional collaboration with legal, technical, and business teams.	• Deep understanding of AI technologies and potential risk factors. • Competency in risk assessment methodologies and tools. • Strong analytical skills for evaluating risk in complex AI systems. • Knowledge of industry-specific regulations and standards.	• In-depth knowledge of trust and safety principles in digital and AI contexts. • Expertise in transparency and explainability techniques for complex AI models. • Familiarity with the societal impact of AI and methods for community engagement. • Strong leadership skills for cross-departmental collaboration and policy implementation.
Outputs	• Ethical guidelines and frameworks for AI development and use. • Regular ethical impact assessments for ongoing AI projects. • Educational workshops and training on ethical practices in AI. • Recommendations to leadership on ethical AI decision-making.	• Risk management plans and strategies for AI implementations. • Detailed risk assessments and reports for AI projects. • Development of risk mitigation tools and practices for AI use cases. • Training and guidance on risk-aware AI development practices.	• Trust and safety policy frameworks that guide AI development and user interactions. • Tools and protocols to enhance the explainability of AI decisions and outputs. • Social impact reports that evaluate the wider effects of AI on various communities. • Initiatives and programs that foster public trust in AI technologies.

Metrics for Roles

Group	Role	Metric 1	Metric 2	Metric 3
Strategy, Policy, Experience Roles	*Chief AI Officer (CAIO)*	ROI of AI Portfolio and AI Impact	Innovation Index	AI Talent Acquisition and Retention
	AI Strategy Consultant	AI Adoption Rate	ROI on AI Initiatives	Strategic Outcome Alignment
	AI Policy Advisor	Policy Implementation Success Rate	Regulatory Impact Score	Stakeholder Engagement Level
	Human-AI Interaction Designer	User Satisfaction Score	Accessibility Compliance Rate	User Engagement Metrics
Responsible AI Practices	*AI Ethicist*	Ethical Compliance Score	Stakeholder Trust Index	Ethical Issue Resolution Time
	AI Risk Manager	Risk Mitigation Effectiveness	Critical Risk Identification Time	Risk Incident Response Time
	AI Trust and Safety Manager	Trust and Safety Policy Violation Rate	Incident Response Time	Social Impact Score
Compliance and Audit	*AI Compliance Officer*	Compliance Rate	Audit Findings Resolution Time	Training Completion Rate
	AI Audit and Assurance Professional	Audit Coverage Ratio	Risk Detection Rate	Audit Recommendation Implementation Rate
	Data and AI Privacy Officer	Data Breach Incident Rate	Privacy Policy Adherence Score	Data Subject Request Fulfillment Time
Operational Roles	*Product Manager*	Product Performance Metrics	Time to Market	Feature Utilization Rate
	Data Engineer	Data Pipeline Reliability	Data Processing Time	Cost Efficiency of Data Operations
	Data Scientist	Model Accuracy Metrics	Insight Generation Speed	Impact of Insights
	ML Engineer	Model Deployment Rate	Model Performance in Production	Post-Deployment Model Tuning Frequency
	ML Ops	Automation Pipeline Efficiency	Model Service Uptime	Incident Response Time

OPERATIONALIZING AI RESPONSIBLY

AI Harm & Risks: BIAS

Apple Card algorithm sparks gender bias allegations against Goldman Sachs

Entrepreneur David Heinemeier Hansson says his credit limit was 20 times that of his wife, even though she has the higher credit score

By Taylor Telford

November 11, 2019 at 10:44 a.m. EST

The Washington Post
Democracy Dies in Darkness

Dave Edwards · Nov 9, 2019
@dedwards93 · **Follow**
Replying to @dhh and @AppleCard

As a former senior Apple employee, this @applecard issue is very disappointing to me. I feel betrayed. Apple is positioned as the good team in tech. And I believe they are. But this is an issue that they have to fix. They have to think different and be better.

Steve Wozniak
@stevewoz · **Follow**

I'm a current Apple employee and founder of the company and the same thing happened to us (10x) despite not having any separate assets or accounts. Some say the blame is on Goldman Sachs but the way Apple is attached, they should share responsibility.

2:06 AM · Nov 10, 2019

2.6K Reply Copy link

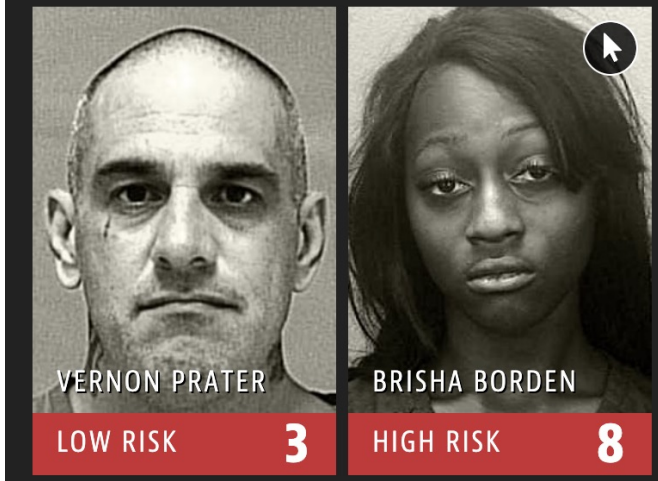
Read 64 replies

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica
May 23, 2016

Two Petty Theft Arrests



Borden was rated high risk for future crime after she and a friend took a kid's bike and scooter that were sitting outside. She did not reoffend.

AI Harm & Risks: Deepfakes

A tweet about a Pentagon explosion was fake. It still went viral.

An apparently AI-generated image sparked a brief dip in stock market. It could have been worse.

By Will Oremus, Drew Harwell and Teo Armus

Updated May 22, 2023 at 6:58 p.m. EDT | Published May 22, 2023 at 5:14 p.m. EDT

The Washington Post
Democracy Dies in Darkness

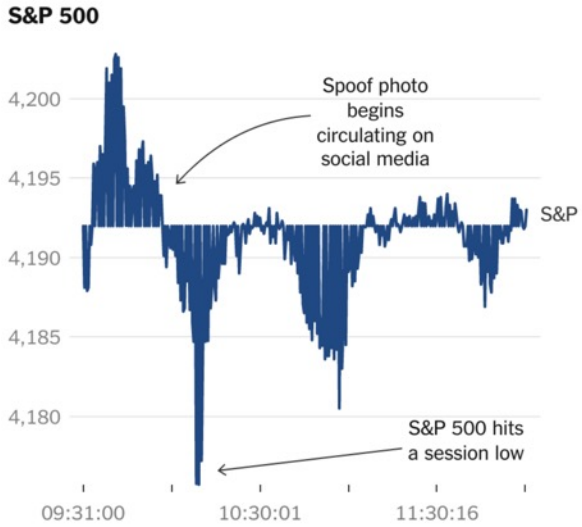
DEALBOOK NEWSLETTER

An A.I.-Generated Spoof Rattles the Markets

The New York Times

By Andrew Ross Sorkin, Bernhard Warner, Sarah Kessler, Michael J. de la Merced, Lauren Hirsch and Ephrat Livni

May 23, 2023



Source: Sentieo/AlphaSense • By The New York Times

AI Harm & Risks: Large Language Models

MORONIC ROBOTIC

A US attorney faces punishment for citing fake cases ChatGPT fed him

The lawyer now regrets trusting the chatbot which misled him

By **Faustine Ngila** Published May 29, 2023

QUARTZ Make business better.™

The New York Times

As an Avianca flight approached Kennedy International Airport in New York, a serving cart collision began a legal saga, prompting the question: Is artificial intelligence so smart? Nicolas Economou/NurPhoto, via Getty Images

Programs to detect AI discriminate against non-native English speakers, shows study

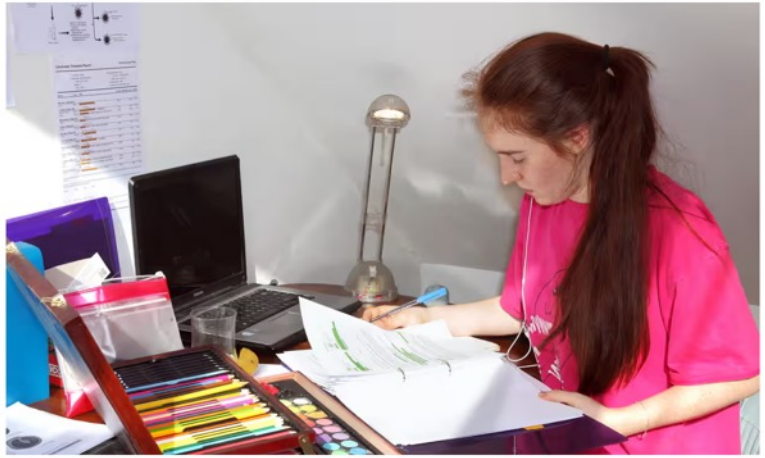
Over half of essays written by people were wrongly flagged as AI-made, with implications for students and job applicants

Ian Sample *Science editor*

@iansample

Mon 10 Jul 2023 11.00 EDT

The Guardian



AI detectors could falsely flag college and job applications and exam essays as GPT-generated, and marginalise non-native English speakers on the internet. Photograph: Peter Phipp/Alamy

History of AI Ethics and Responsible AI

2017 Asilomar Conference
- 23 AI Principles in Research, Ethics & Values, and longer-term issues

2017, IEEE Ethically Aligned Design
- Comprehensive set of guidelines for the ethical implementation of autonomous and intelligent systems
- Emphasizes the prioritization of human well-being in the design and operation of these systems

2018 GDPR
- EU's General Data Protection Rule
- Purpose limitations, lawfulness, data minimization, confidentiality, accuracy, accountability, storage limitations

2018 Toronto Declaration
- Emphasized human rights in the age of AI.
- Advocated for equality and non-discrimination in AI use

2019 OECD AI Principles
- Set guidelines for responsible, human-centric AI development.
- Highlighted inclusive growth, fairness, and accountability

2019 California Consumer Privacy Act(CCPA)
- Enhanced privacy rights and consumer protection in California, USA.
- Influenced how companies handle personal information

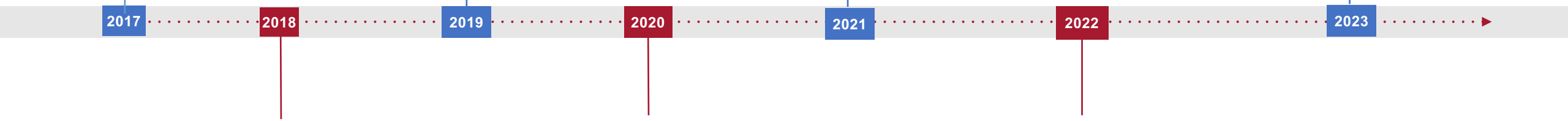
2020, EU AI Act
- EU's proposed regulation focusing on high-risk AI systems.
- Aimed at ensuring safety, transparency, and human rights respect.

2021, Singapore Model AI Governance Framework
- Detailed guidance on responsible AI use.
- Emphasized explainability, transparency, fairness, and human-centricity.

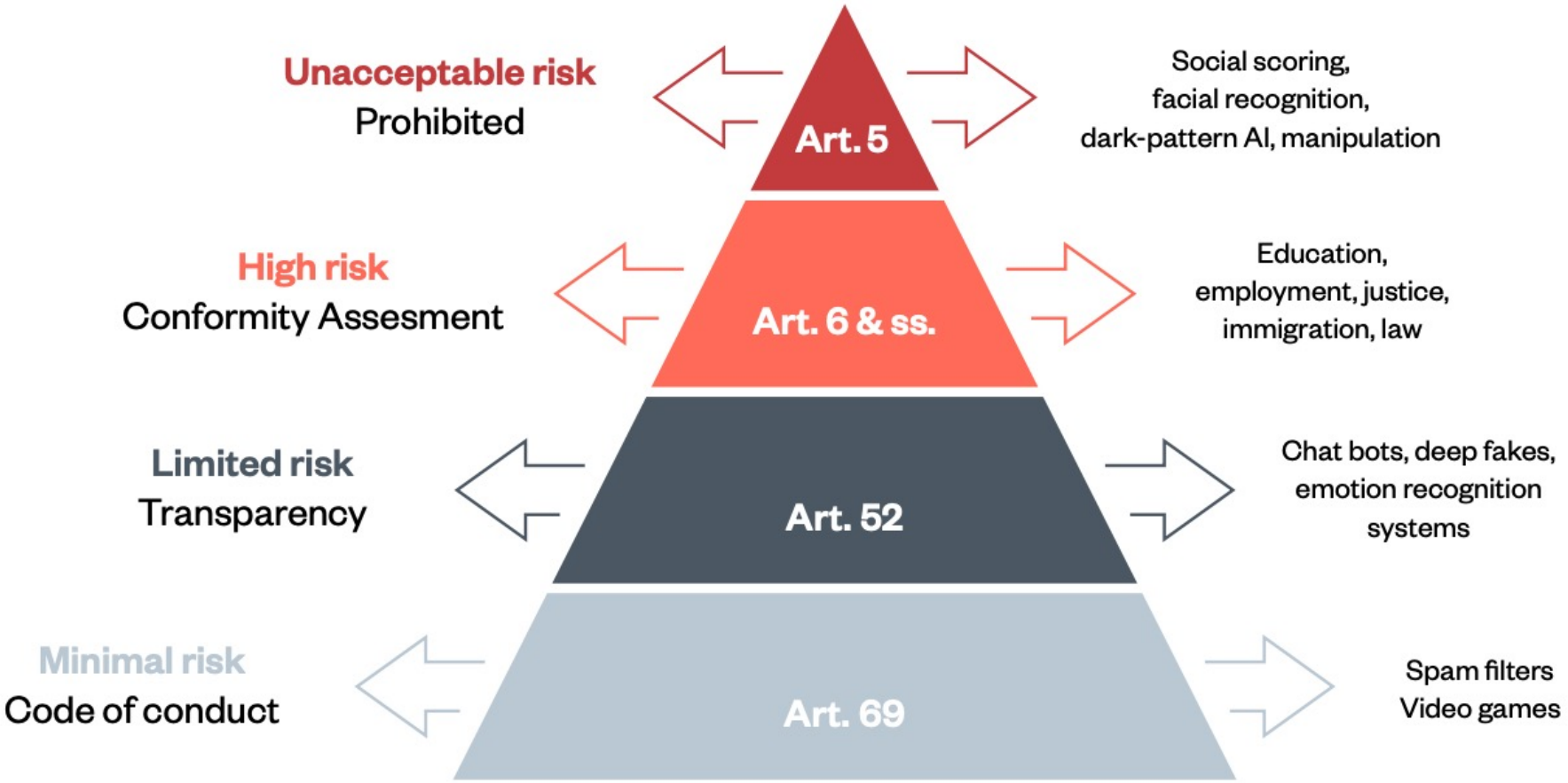
2022, US AI Bill of Rights
- Asserts principles and guidance around equitable access and use of automated, or artificial intelligence systems

2023, Executive Order
- Comprehensive directive aimed at ensuring the safe, secure, and trustworthy development and use of artificial intelligence, with a focus on protecting privacy, advancing equity and civil rights, promoting innovation and competition, and advancing American leadership globally

Algorithmic Accountability Act (2019, 2022, 2023)
- Proposed US legislation for AI system impact assessments.
- Focused on addressing AI bias, discrimination, privacy, and security



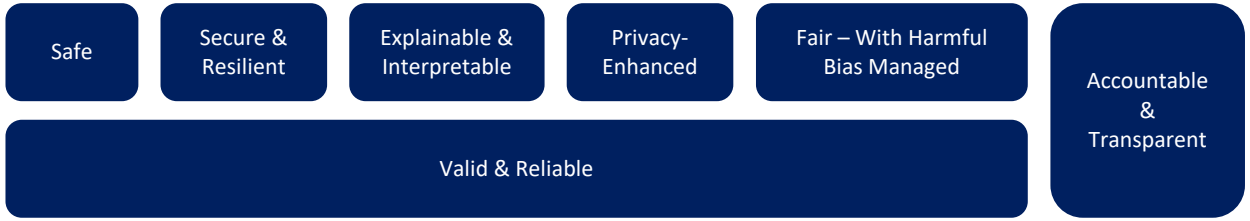
EU AI Act



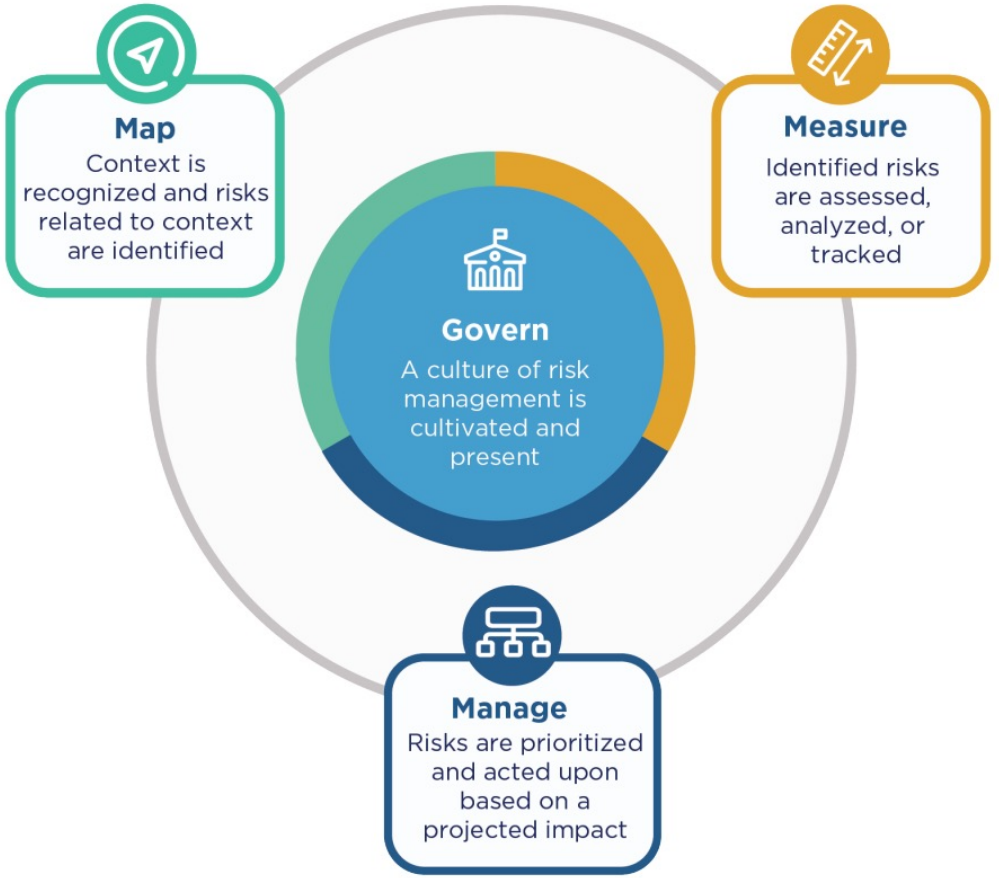
Source: [The EU AI Act: a summary of its significance and scope](#) by Lilian Edward, Ada Lovelace Institute April 8, 2022.

National INSTITUTE of Standards and Technology

Responsible AI Principles



AI Risk Management Framework



Sources: *AI Risk Management Framework (AI RMF 1.0),* NIST, January 2023.

Operationalizing Responsible AI

STRATEGIC

Data & AI Ethics
Extending past "what do we have to do" dictated by compliance to regulation, to the "what we should do" in terms of moral implication of uses of data and AI, role of context and stakeholder impact

Policy & Regulation
Anticipate and understand key public policy and regulatory trends in order to align compliance processes with future regulatory requirements and guidance.

PERFORMANCE & SECURITY

Bias & Fairness
Defining and measuring fairness for intersectional groups and testing systems against defined standards

Interpretability & Explainability
Translating and curating model decision making to different stakeholders based on their needs and uses

Privacy
Utilizing emergent privacy-preserving technologies to train resilient systems on large datasets while respecting data protections

Security
Enhancing the cybersecurity of systems and anticipating malicious attacks, such as adversarial attacks

Robustness
Enabling high performing systems over time, and reducing sensitivity to slight changes.

Safety
Designing and testing model performance in the context of human users to anticipate and remediate potential harms.

CONTROL

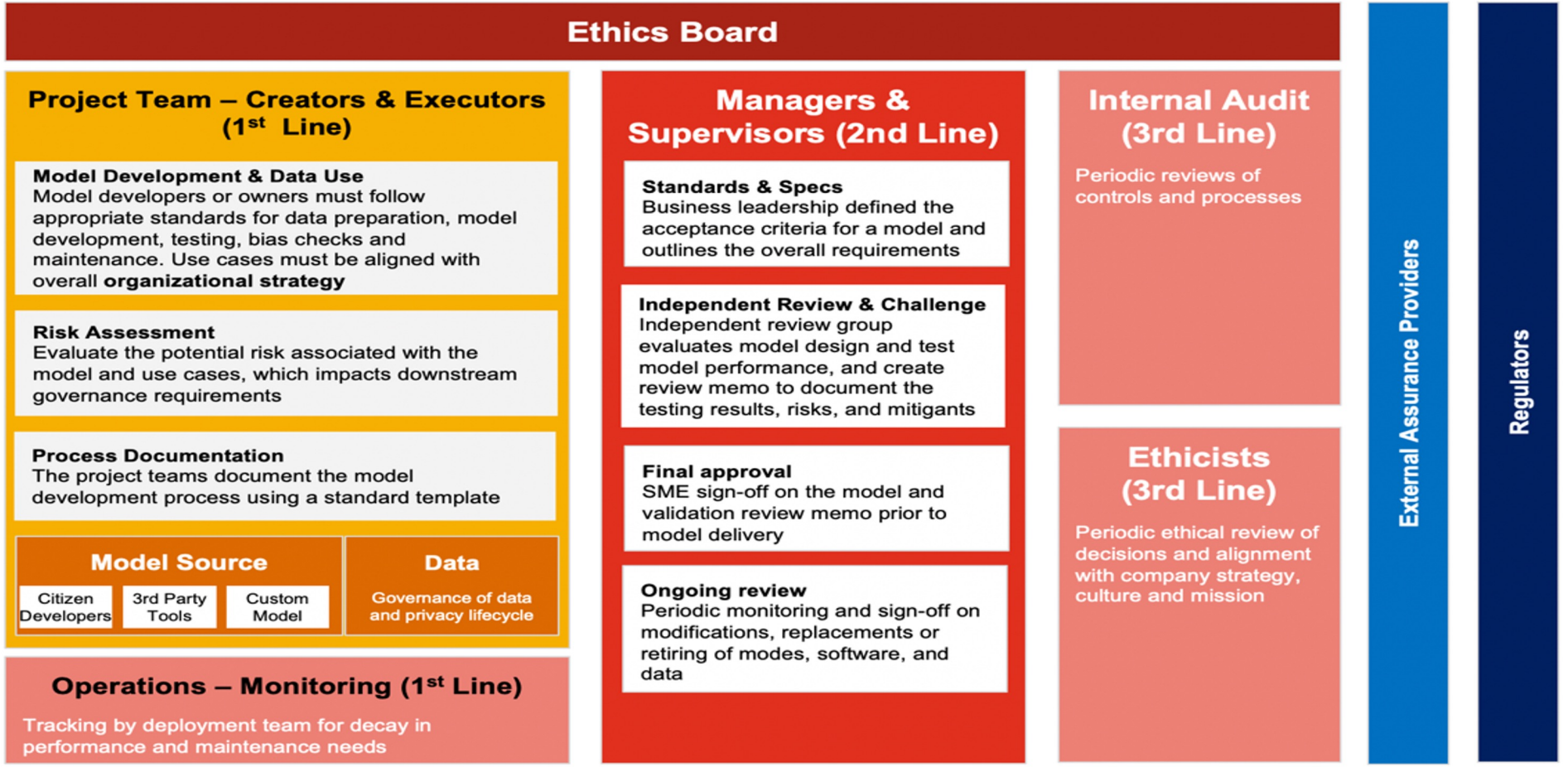
Governance
Enabling oversight with clear roles and responsibilities, articulated requirements across three lines of defense, and mechanisms for traceability and ongoing assessment

Compliance
Complying with data protection and privacy regulation, organizational policies, and industry standards

Risk Management
Expanding risk detection and mitigation practices to address existing and newly identified risks and harms

Source: Ten principles of Responsible AI for corporates (Towards Responsible AI — Part 2)

Operationalizing RAI - Governance



Source: Top-down and end-to-end governance for the responsible use of AI (Towards Responsible AI — Part 3)

Operationalizing RAI - Practice

Stage gates

1. Shall we proceed with the AI solution?
2. Does the model meet our expectations?
3. Do we deploy the model into production?
4. Is the model ready to be transitioned for Business-As-Usual operation?
5. Should the model continue as-is, be retrained, redesigned, or retired?

Value Scoping

1. Business & Data Understanding
2. Solution Design

Value Discovery

3. Data Extraction
4. Pre-Processing
5. Model Building

Value Delivery

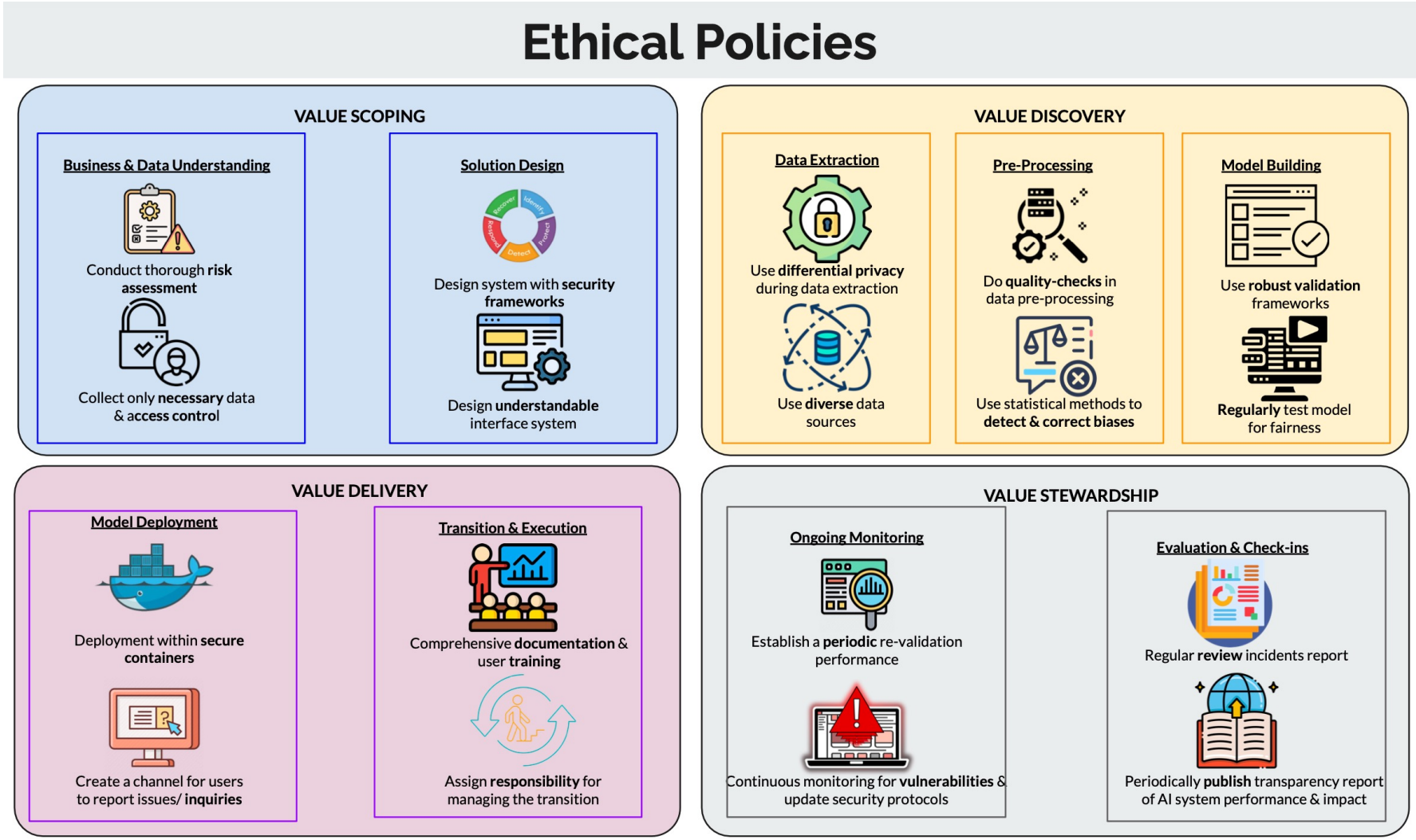
6. Model Deployment
7. Transition & Execution

Value Stewardship

8. Ongoing Monitoring
9. Evaluation & Check-in

Source: Six stage gates to a successful AI Governance(Towards Responsible AI — Part 4)

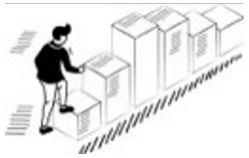
CMU Operationalizing AI – Student Project



Responsible AI: Key Takeaways

1. **Align on AI principles and practices**

2. **Confirm adequate top-down & end-to-end governance**



3. **Design for robustness & safety**

4. **Exercise control and value alignment**

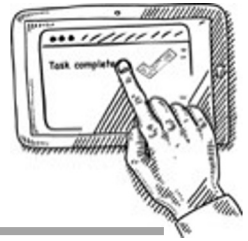
5. **Respect privacy**

6. **Be transparent**

7. **Embed security**

8. **Enable diversity, non-discrimination & fairness**

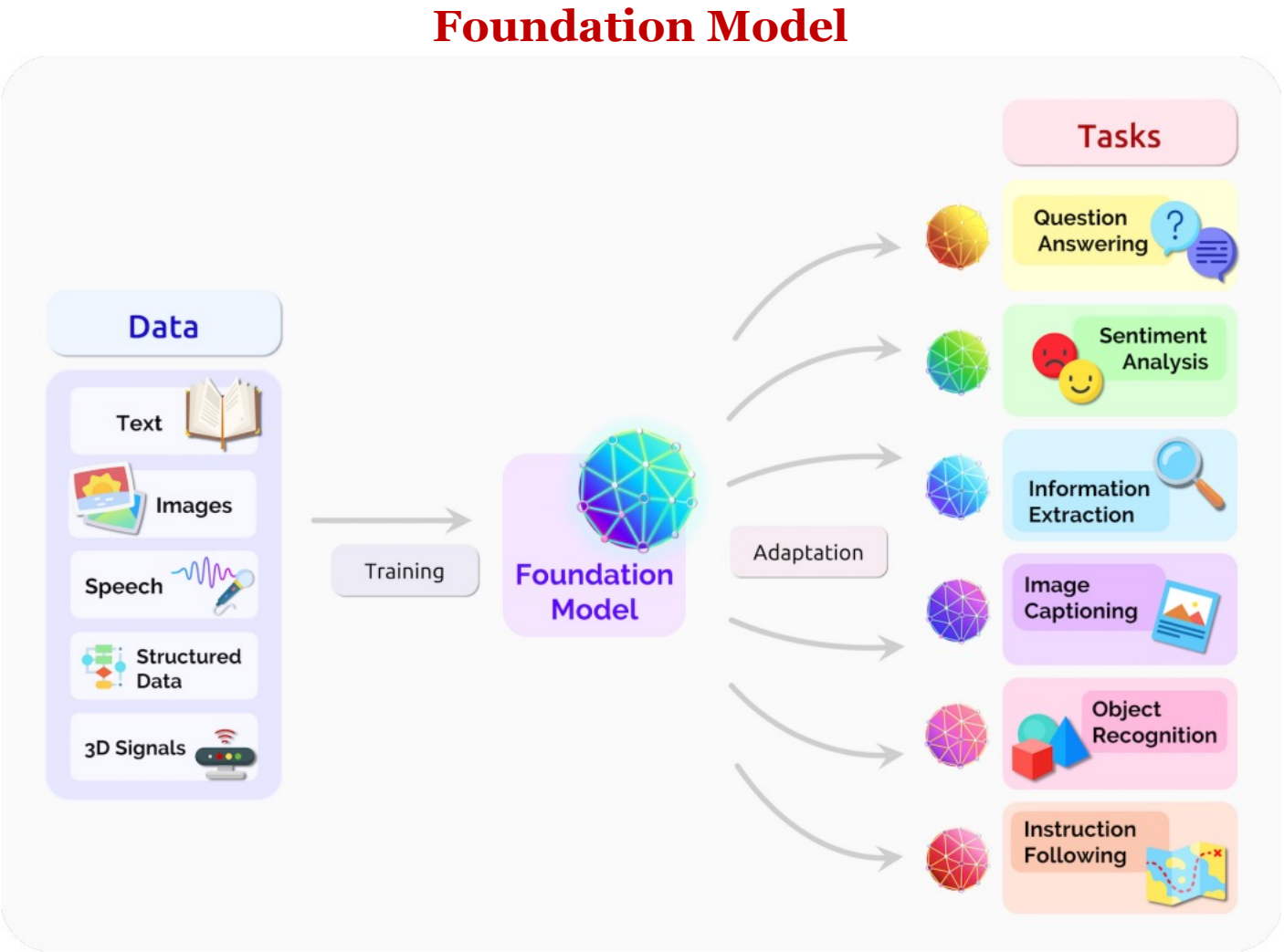
9. **Clarify accountability**



10. **Foster societal and environmental well-being**

OPERATIONALIZING GENERATIVE AI

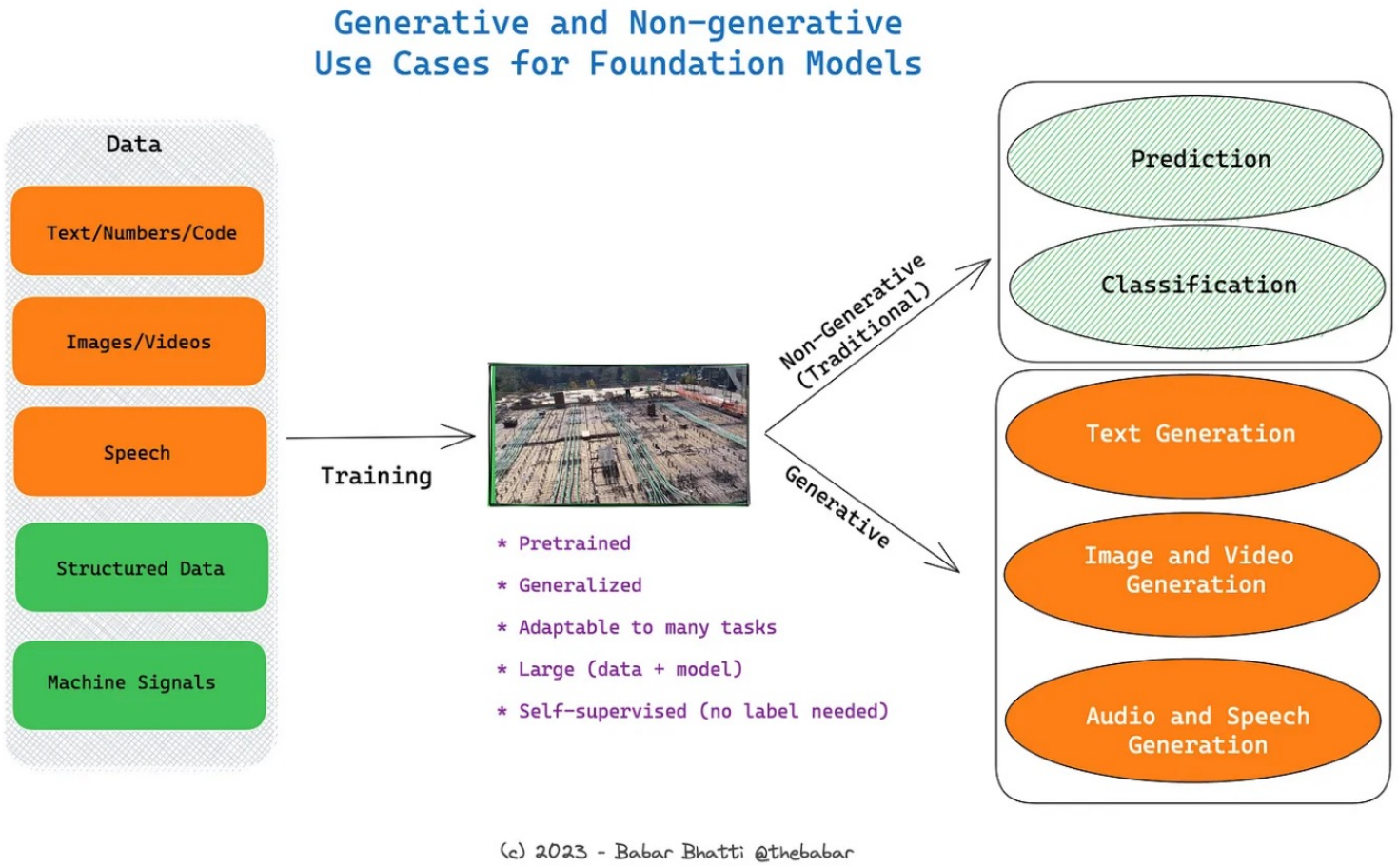
Generative AI: From ‘building’ to ‘adapting’



- **Foundation models** are pre-trained models that exploit *transfer learning* at *scale* to *adapt* the model for multiple tasks
- Foundation models built for natural language text are called **Large Language Models (LLMs)**

Source: Bommasani, R.,.et.al., (2021). *On the Opportunities and Risks of Foundation Models*. *ArXiv, abs/2108.07258*.

Generative AI: From Predictive to Generative



Generative AI is AI that creates content

- Text in any natural language
- Code in any programming language
- Images or videos
- Voice
- Synthesize data
- 3D images

Source: Essential guide to foundation models and large language models by Babar M. Bhatti, Medium, February 5, 2023.

GenAI Model Development Lifecycle

Value Scoping

1 Business & Data Understanding

Understand the business challenges; identify and source data, including actual and synthetic

2 Solution Design

Design the solution, select the analytic & AI methods suited for the application & requirements

Value Discovery

3a Data Extraction

4a Pre-Processing

5a Model Building

PRE-TRAINED FOUNDATIONAL MODEL

3b Model Selection

Select a proprietary or open-source model (foundational, sector, task, or private)

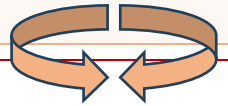
4b Model Adaptation

Adapt model by prompt engineering, fine-tuning, or reinforcement learning

5b Model Evaluation

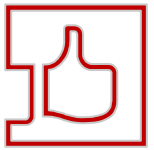
Evaluate model against pre-defined requirements (e.g., truthfulness, creativity) and metrics (e.g., accuracy, bias)

GENAI MODEL ADAPTATION FOR USE



Value Delivery

6 Model Deployment



Packaged model published for broader use

7 Transition & Execution

Implementation into business process and workflows; evangelization

Value Stewardship

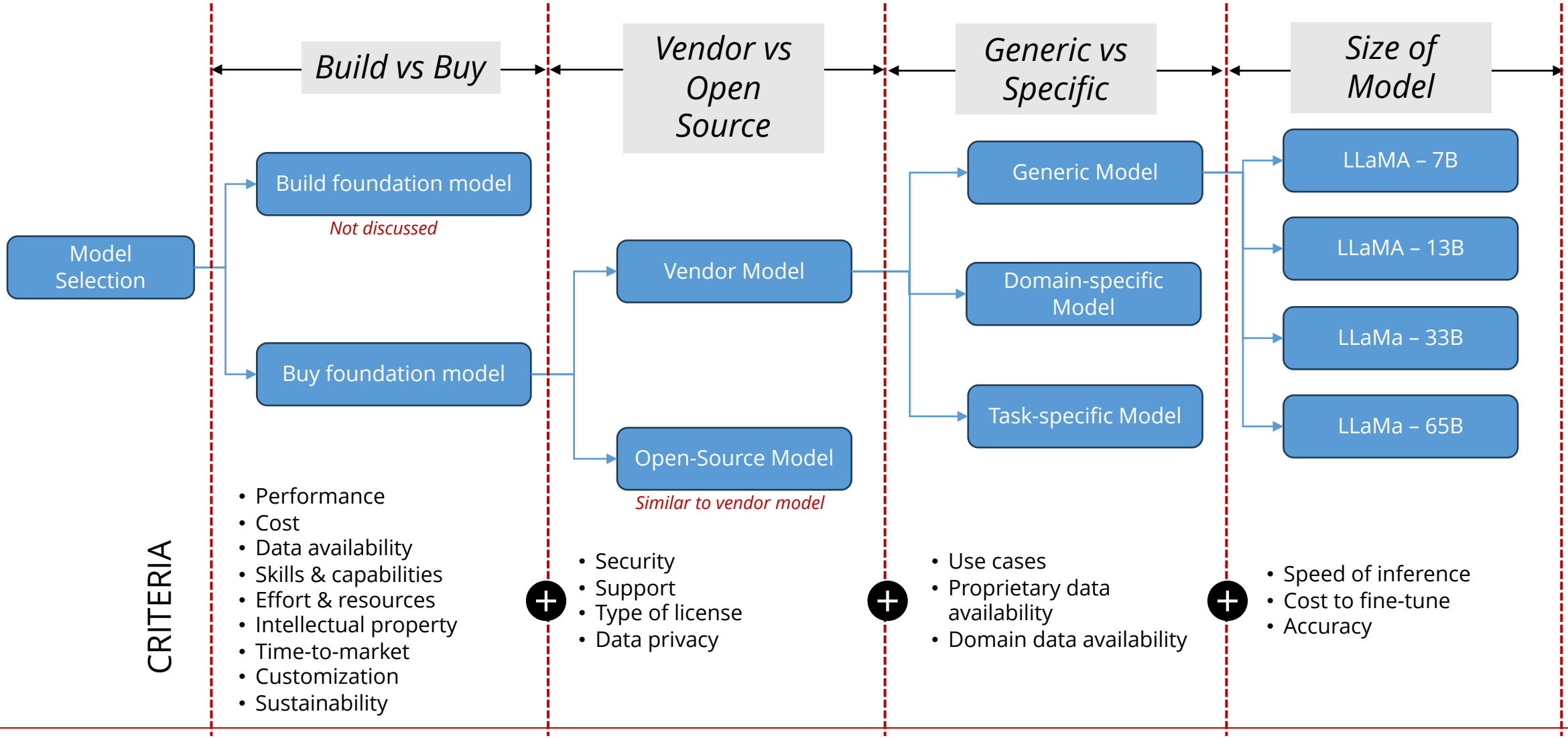
8 Ongoing Monitoring

Ongoing monitoring of outcomes for continuous observation & auditing

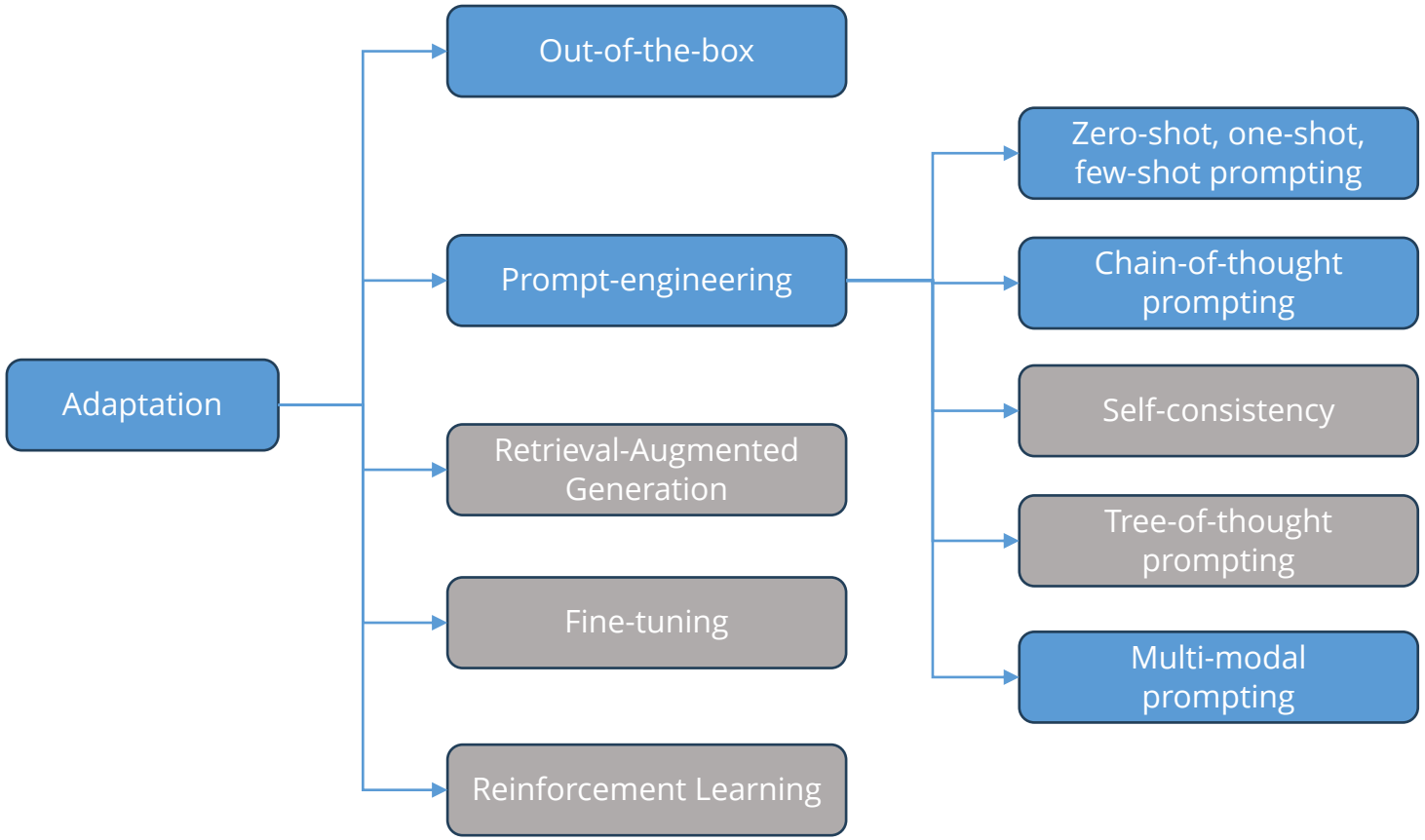
9 Evaluation & Check-in

Evaluation of insights & actions against business objectives*

Model Selection



Model Adaptation

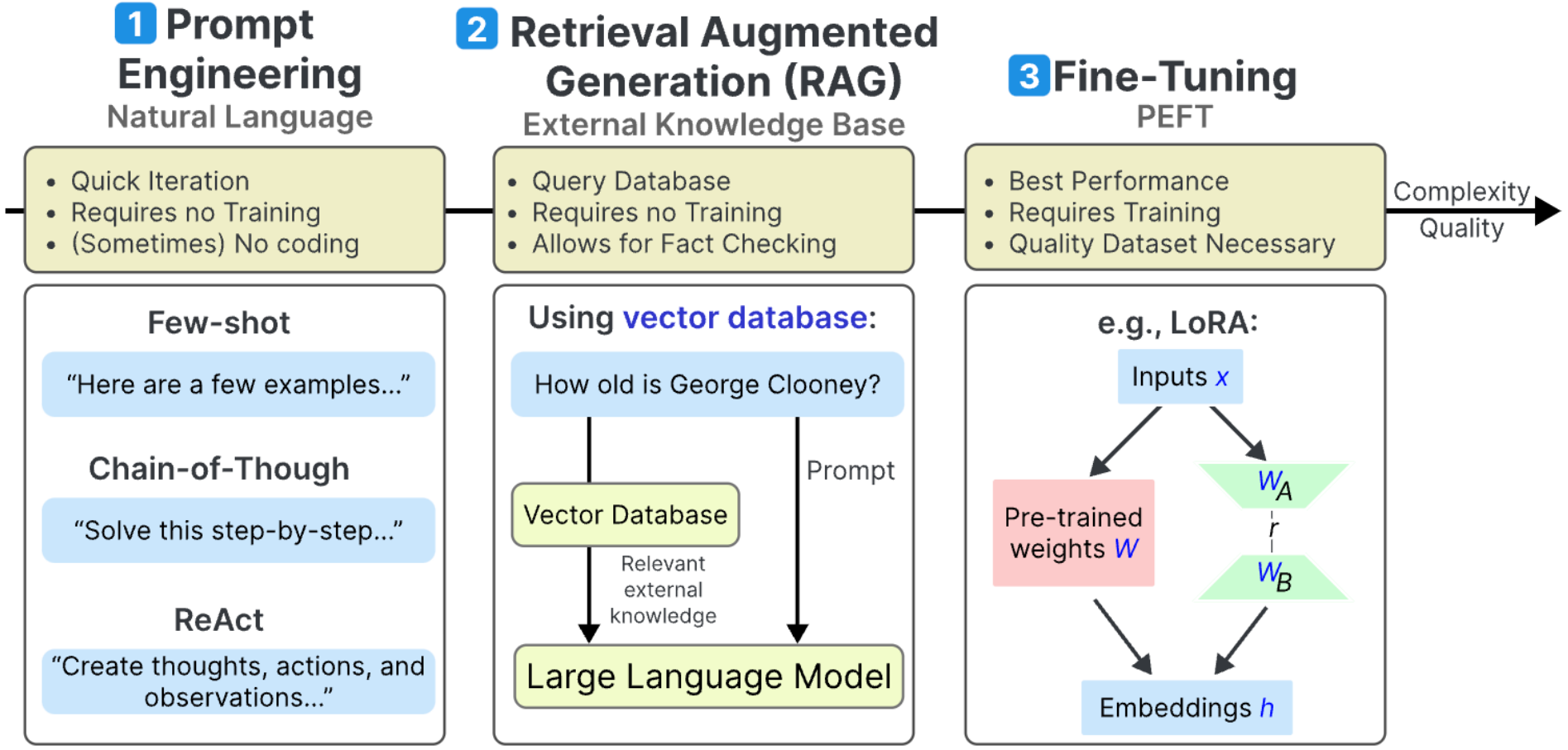


Other prompting methods not covered

- Generate knowledge prompting
- Retrieval augmented generation
- Automatic reasoning and tool-use
- Automatic prompt generation
- Active-prompt
- Directional stimulus prompting
- ReAct (reasoning-acting)
- Graph prompting

Source: *Prompt engineering guide*.

Model Adaptation



Source: Grootendorst, Maarten. "3 Ways to Improve Your Large Language Model." *Maarten Grootendorst*, 11 Sept. 2023.

Evaluation – Benchmarks and empirical evaluation

MODELS

AI21 labs · ANTHROP\C · BigScience

cohere · EleutherAI · Google

Stanford CRFM

Meta · Microsoft · NVIDIA

OpenAI · Yandex · TOGETHER

SCENARIOS

- Question answering
- Information retrieval
- Summarization
- Sentiment analysis
- Toxicity detection
- Text classification
- Languages
- Robustness to contrast sets
- Knowledge
- Reasoning
- Harms
- Efficiency
- Calibration
- Vary # of in context examples
- Vary multiple choice strategy
- Vary prompting

METRICS

- Accuracy
- Calibration
- Robustness
- Fairness
- Bias
- Toxicity
- Efficiency
- General information
- Summarization metrics
- APPS metrics
- BBQ metrics
- Copyright metrics
- Disinformation metrics
- Classification metrics

Source: Liang, P., et.al., *Holistic Evaluation of Language* Models. *Annals of the New York Academy of Sciences, 1525*, 140 – 146, 2022.

Module Takeaways

- **Integrate AI Strategically for Business Impact**: Operational excellence isn't just a buzzword; it's a necessity. Ensure that AI is deeply integrated into your business processes and aligns with your organizational goals. Address challenges proactively to establish best practices.
- **Establish Robust Governance and Ethical Frameworks**: From top-down governance to respecting privacy and fostering societal well-being, having a comprehensive set of principles and practices is crucial. Accountability and transparency should be non-negotiables.
- **Foster Cross-Functional Collaboration and Continuous Oversight**: AI is a team sport. Encourage collaboration across departments for effective deployment and maintenance. Additionally, the AI lifecycle is dynamic—regular updates, monitoring, and feedback loops are essential for long-term success.
- **Adapt and Optimize Generative AI for Personalized Solutions**: The future of AI is in adaptation, not just creation. Understand the nuances of Large Language Model Operations and generative AI. Prioritize rigorous evaluation and holistic approaches, including prompt engineering and reinforcement learning, to ensure ethical and optimal outcomes.

Recommended Resources

- *Operationalizing Artificial Intelligence – Making the Promise a Reality*, Harvard Business Review Analytics Services, Research Report, 2021
- *Data Scientists are from Mars and Software Developers are from Venus (Part 1), Rao, A., Towards Data Science, August 29, 2020.*
- *How these 8 companies implement MLOps*: In-depth guide. Neptune.ao, April 19, 2023.
- *AI Risk Management Framework (AI RMF 1.0),* NIST, January 2023.
- *US Department of Defense Responsible Artificial Intelligence Strategy and Implementation Pathway*, June 2022.