Talentathon

Institute Of Product Leadership

By Asit Piri
Date: 20th June 2022

About Me

Experiences

My name is **Asit Piri**, I am a seasoned technologist turned AI Product Manager with **20+ years** industry experience and a purpose to build products that advance the way people live and work.

Educations

- Executive MBAs in **Product Management** & **Business Analytics**
- Certified AI Product Manager
- Certified Deep Learning & **MLOps** Professional from DeepLearning.AI.
- Certified Google **UX Design Professional** from Coursera.

Linkedin | Website

Skills

1. **Product Management**
2. User Experience Research & Design
3. **Agile Dual-Track**
4. Product Marketing
5. Negotiation & Conflict Management
6. **Data Architecture & Management**
7. **Data Science, AI & MLOps**
8. **Hybrid Cloud Computing**

Competencies

1. **Business Acumen & Strategy**
2. **Design Thinking & System Design**
3. **Critical Problem Solving**
4. **Cross-Functional Collaboration**
5. Business Story Telling
6. **Customer Centricity**
7. Negotiation & Conflict Management
8. Leadership Intelligence

About Target

CEO

Brian C. Cornell

Customer

58 %
Of American
(Between 18 to 64 years)

No of Stores

1931

No of Employees

409,000

TARGET

Revenue

US$ 106 Billion

Net Income

US$ 4.368 Billion

Operating Income

US$ 6.439 Billion

Total Assets

US$ 51.248 Billion

Source: Wiki | Financial Summary

Use Case

Company ABC Inc. wants to improve its **operational efficiency** of its **infrastructure** for applications hosted on their platform on **cloud** and **data centre**.

- What are some of the **features** that you as a product owner will propose for the above use case?
- What will be your **product roadmap** & your **onboarding strategy** for your stakeholders?

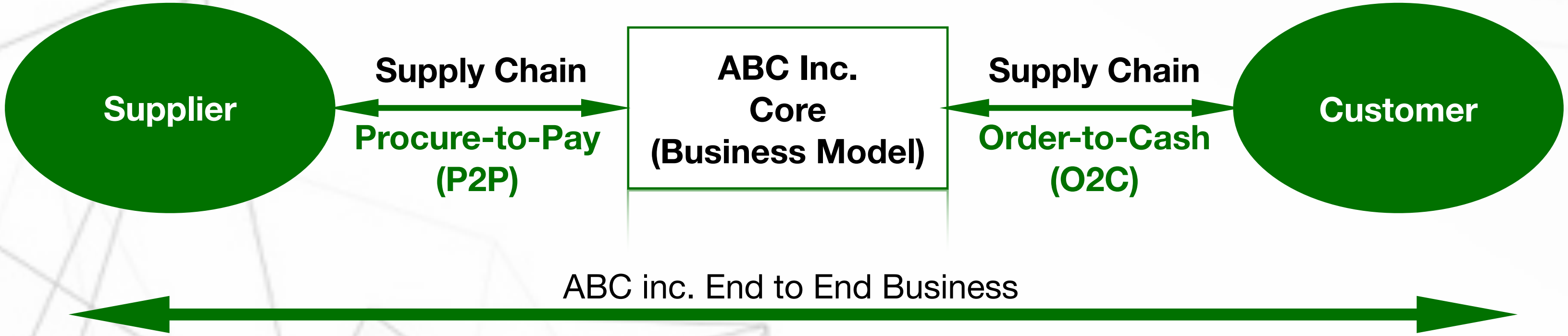
Context:

You represent the "voice of the customer" for the **Performance & Efficiency** team and the needs of stakeholders. You will work with SRE, Technology operations and portfolio Engineers to build and optimise mission-critical infrastructure, tools, and processes that will ensure the highest level of performance and scalability of our platforms and services. As a senior member of the team, you will be expected to work with peers, team members and guests to define and implement the technical vision of the team. This team builds and is responsible for the delivery of Application profilers that help several Product teams (application and platform) analyse and compare code performance at any given time in any environment, including production, with negligible overhead thus improving reliability, enhancing user experience, and reducing overall infrastructure costs.

Assumptions

- **ABC Inc. planning to host AI and Computer Vision application (user interface on cloud and planning for CICD (Continuous Integration, Continuous Deployment & Continuous Training) MLOps deployment strategy.**
- **Efficient Organisation = Efficient Man (Employees) + Efficient Machine (Product)**
- **Operational efficiency of any organisation is directly proportional to the operational efficiency of employee+product.**
- ABC Inc have the similar retail business like Target (1931 stores, 409,000 employees across US locations).
- ABC Inc already having the all applications running at 1931 stores across US and want to migrate to hybrid cloud platform.
- ABC Inc employees running daily multiple simultaneous analytical (machine learning & computer vision) applications from various stores like, video conferencing, competitive price predictions, real time promotions, IoT based automatic product cart addition, updation and deletion including automatic payment from registered credit or debit card at each store for smart shopping experience which required high processing (GPU) machines.
- ABC Inc currently not using Agile Development Methodology and Agile Dual-Track.
- ABC Inc currently not using DevOps / MLOps for Continuous Delivery, Continuous Deployment & Continuous Training

Setting Context



Benefits
(Assumptions)

- **Reduce the system downtime to zero.**
- **Lowering transaction and service costs, decrease by 5%.**
- **Accelerating the receiving process, increased by 10%.**
- **Improve Employee Efficiency and productivity by 25%.**
- **Improved Net Promoter Score (NPS) by 20%.**
- **Delighted Customer, repeat buy by 20%.**

Problem Statement

How might we improve the **operational efficiency** of ABC Inc. by hosting **their software-based applications (infrastructure) to their platform** (**public cloud**, **private data centre**, or **hybrid cloud**) so that they can improve operational efficiency of their **employees** and **product (Infrastructure and code)** to deliver new **services and value to their customer** more quickly and efficiently, thereby increasing their:

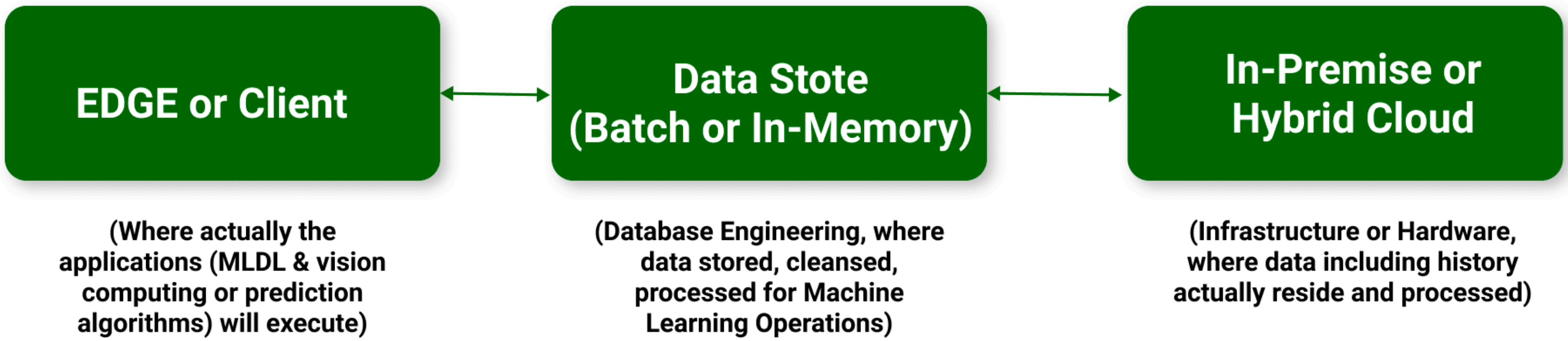
- **Top line (revenue) and**
- **Bottom line (profit)**.

Operational Efficiency = Profit = Fuel for Growth

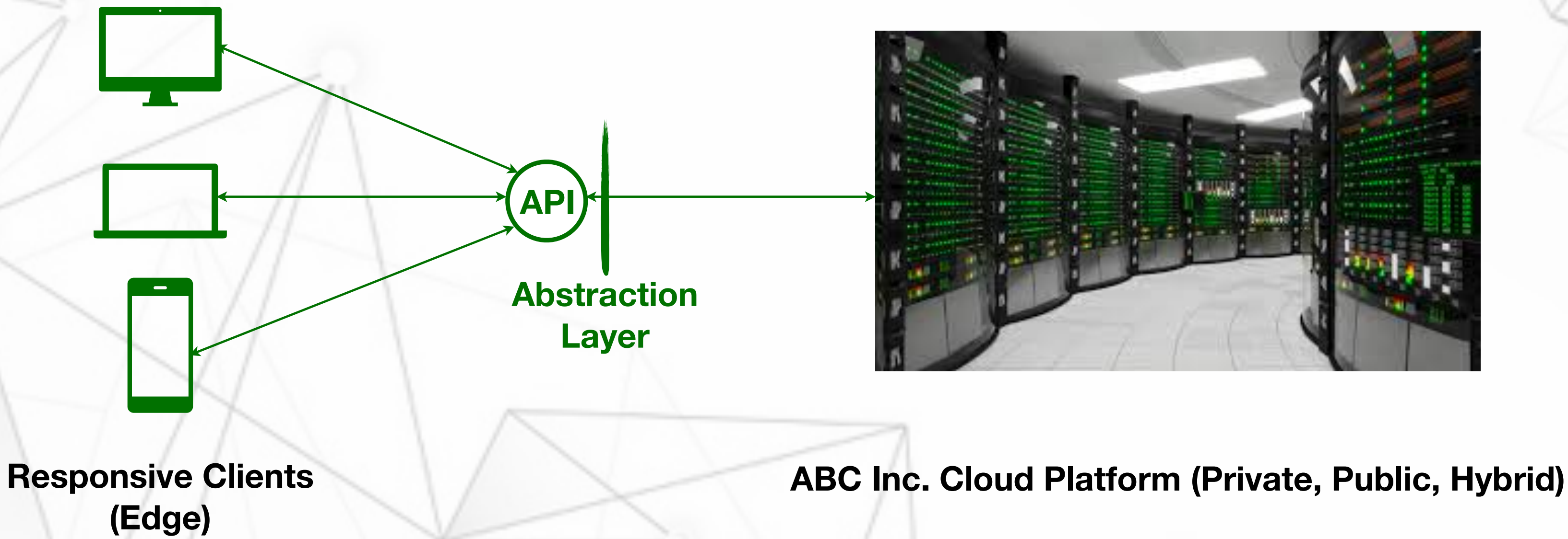
Personas: Cross Domain

Buyer or Influencer Personas	Core User Personas	Cross Domain Personas
• Chief Executive Officer (CEO) • Chief Technology Officer (CT)O • Chief Information Officer (CIO) • Chief Finance Officer (CFO) • Enterprise Architect (EA) • Business Units (Department) Head	• Data Scientist • Data Engineer • ML Engineer • MLOps Engineer • DevOps Engineer • IT Admin • Database Administrator • System Administrator	• Subject Matter (Domain) Expert • Enterprise Data Architect • Product Manager • Engineering Manager • UX Design Manager • Customer

System Design (High Level)



Aspects for System Design



Quality Attributes

- Reliability
- Scalability
- Efficiency
- Cost

Architectural Pattern

- High level universal scope.
- How components are going to be assembled and organised.

API & Interface

- Low level scope.
- How components are going to be build

Challenges

Infrastructure

Build an integrated ML system and continuously operate it in production with a vast array of the surrounding infrastructure:

- Web Applications & User Interface
- ML Codes
- Application Programming Interface (APIs)
- Configuration
- Automation
- Data Collection
- Data Verification
- Feature Engineering
- Testing and Debugging
- Resource Management
- Model Analysis
- Process Management
- Metadata Management
- Serving Infrastructure
- Monitoring

Dependencies

The below changes are dependencies in addition to code must be controlled and integrated into the software delivery process:

- Data Dependency
- Model Complexity
- Reproducibility
- Testing
- Monitoring etc.

People

To automate the process from beginning to end while managing various stakeholders:

- Cross-functional teams
- Using different technologies
- Follow different routines.

Also make them:

- Auditable
- Reproducible

Aspects for Data Management Function and Scope

01 Data Management

02 Data Handling Ethics

03 Data Governance

04 Data Architecture

05 Data Modelling and Design

06 Data Storage and Operations

07 Data Security

08 Data Integration and Interoperability

09 Document and Content Management

10 Reference and Master Data

11 Data Warehousing and Business Intelligence

12 Metadata Management

13 Data Quality

14 Big Data and Data Science

15 Data Management Maturity Assessment

16 Data Management Organisation and Role Expectation

17 Data Management and Organisation Change management

Source: Data Management Body of Knowledge (DMBoK)

Aspects for Agile Methodology (Adaptive Process)

Waterfall Methodology

1. Requirement Study
2. System Design
3. Development
4. Testing
5. Deployment
6. Maintenance

Drawbacks:

- Unidirectional
- It need clear-cut requirements before development
- Rigid or not adaptable in nature
- Not user-centric

Agile (Lean) Methodology

1. Viability/Feasibility/Desirability Study
2. Planning
3. System Analysis
4. System Design
5. Development
6. Testing
7. Deployment
8. Maintenance & Monitoring

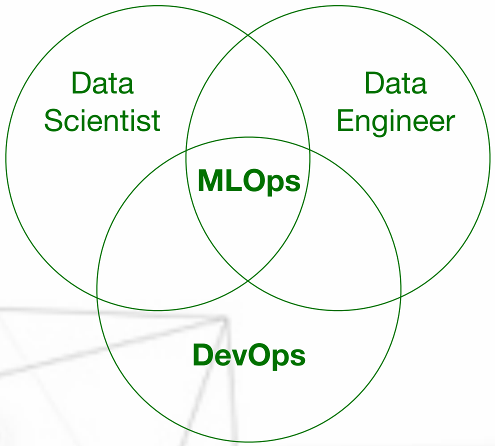
Advantages:
Agile allow business changes to incorporate easy and fast and bi-directional, I.e. business and development closely collaborate through out the process Hence more user-centric and focuses on:

- Adaptability
- Modularity
- Reusability
- Bi-directional

Aspects for DevOps and MLOps

DevOps

Build → Test → Integrate → Deliver → Deploy

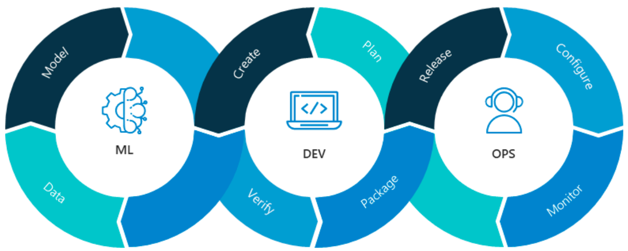


MLOps

Build → Test → Integrate → Train → Deliver → Deploy

Advantage:

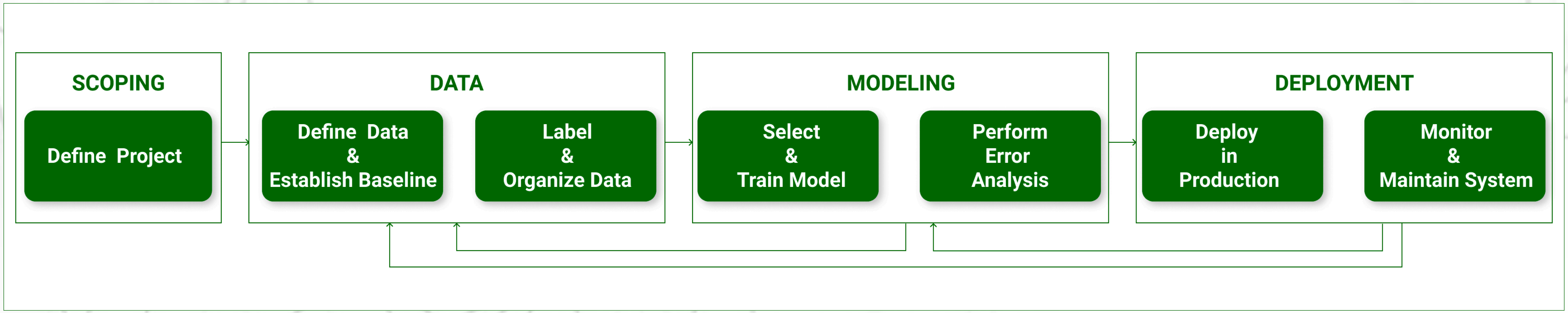
- Continuous Integration (CI)
- Continuous Deployment (CD)



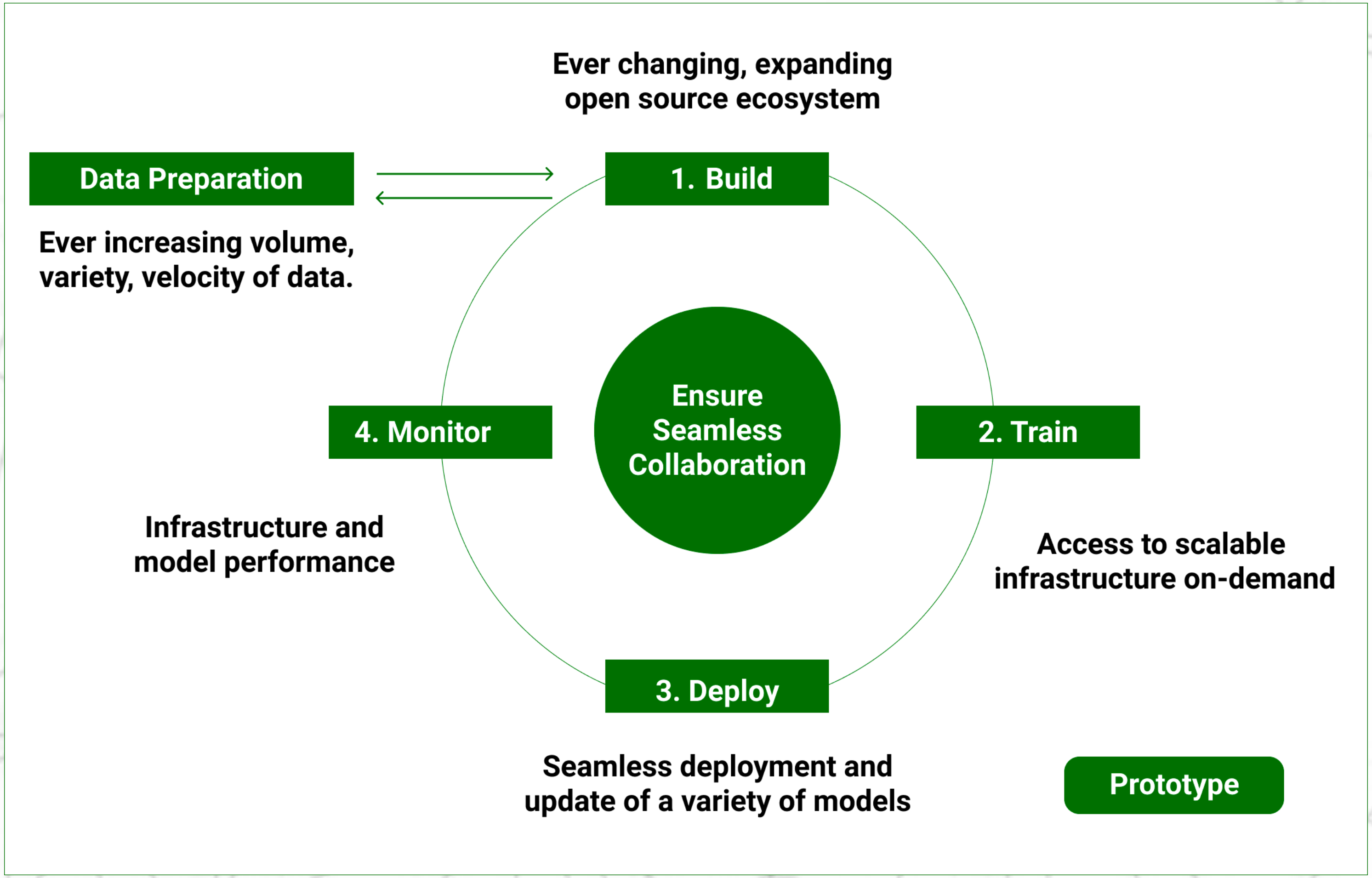
Advantage:

- Continuous Integration (CI)
- Continuous Deployment (CD)
- Continuous Training (CT)

The AI/ML Project Lifecycle



The AI/ML Project Lifecycle



Proposed EPIC (Bunch of Features) to drive operational efficiency

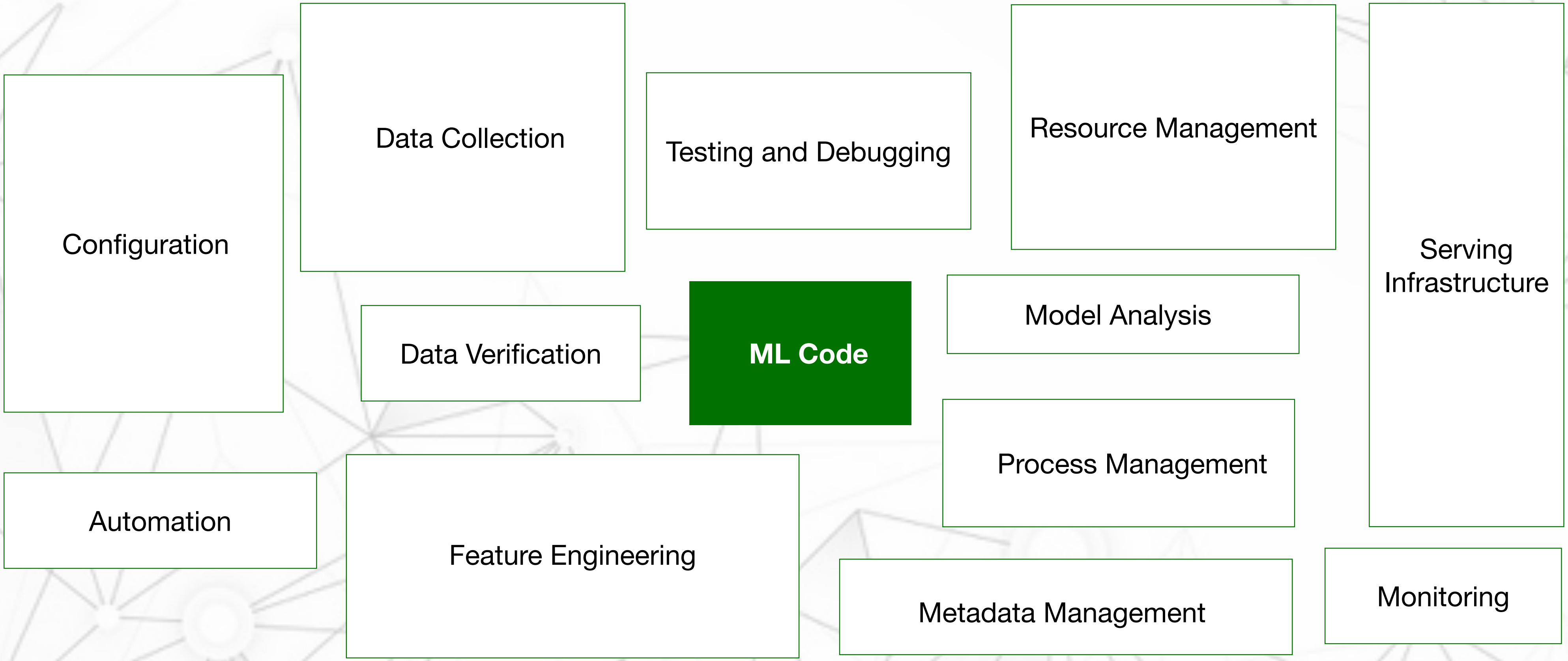
EPICS

- **Automation**
  - **Level 0**
  - **Level 1**
  - **Level 2**
- Scalable capacity
- Advanced self-service
- Agile service creation
- Application execution efficiency
- Efficiency analytics for monitoring
- Predictive analytics for proactive prevention

Key indicators (metrics) to measure operational efficiency

- **The unit cost of service production**
- **The cost of capacity waste**
- **The cost of creating functionality**

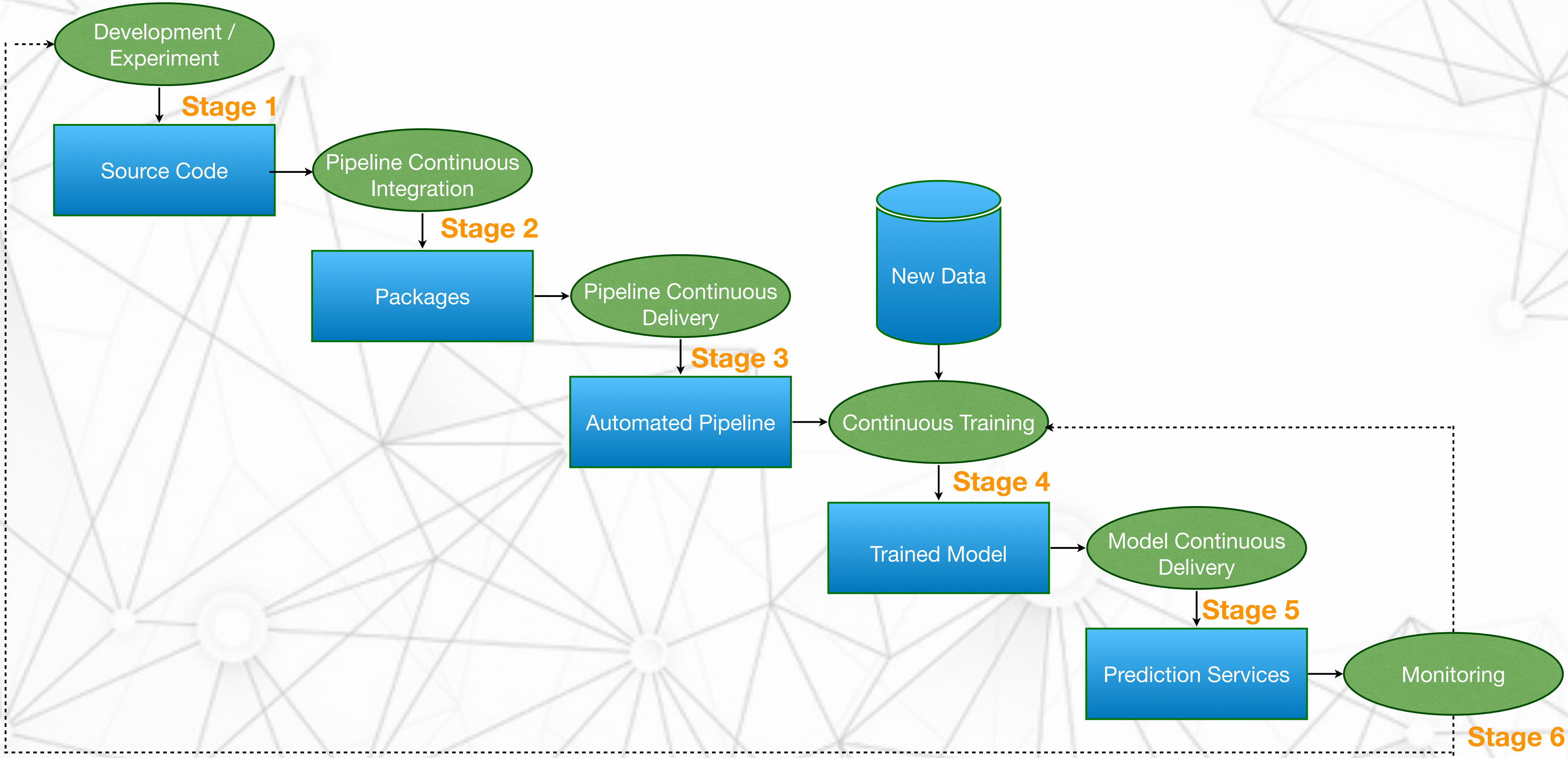
Hidden Technical Debt in Machine Learning System



Only a small fraction of a real-world ML system is composed of the ML code. The required surrounding elements are vast and complex.

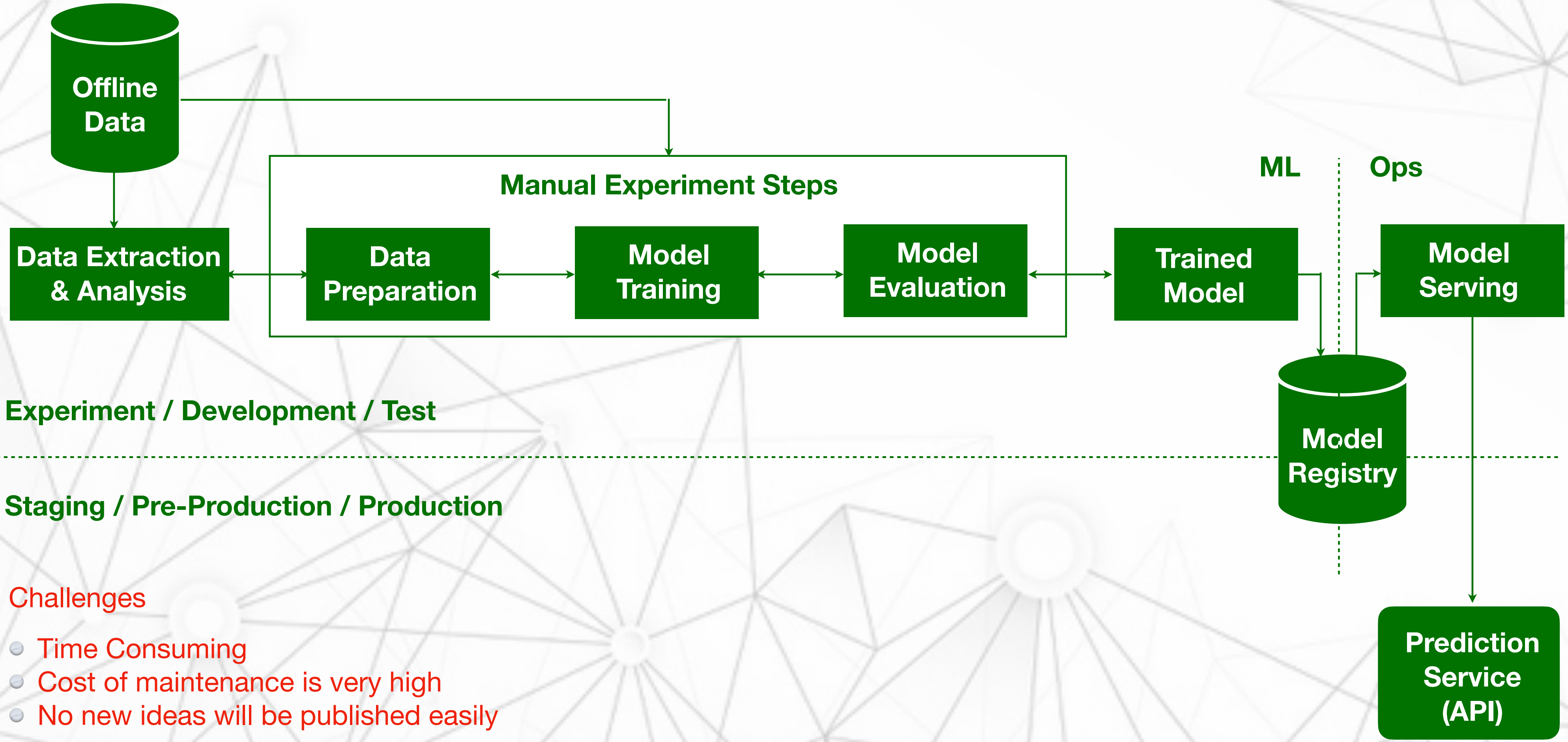
Source: https://papers.nips.cc/paper/2015/file/86df7dcfd896fcaf2674f757a2463eba-Paper.pdf

Stages of the CI/CD Automated ML Pipeline



Reference: GCP

Level 0: Manual Process

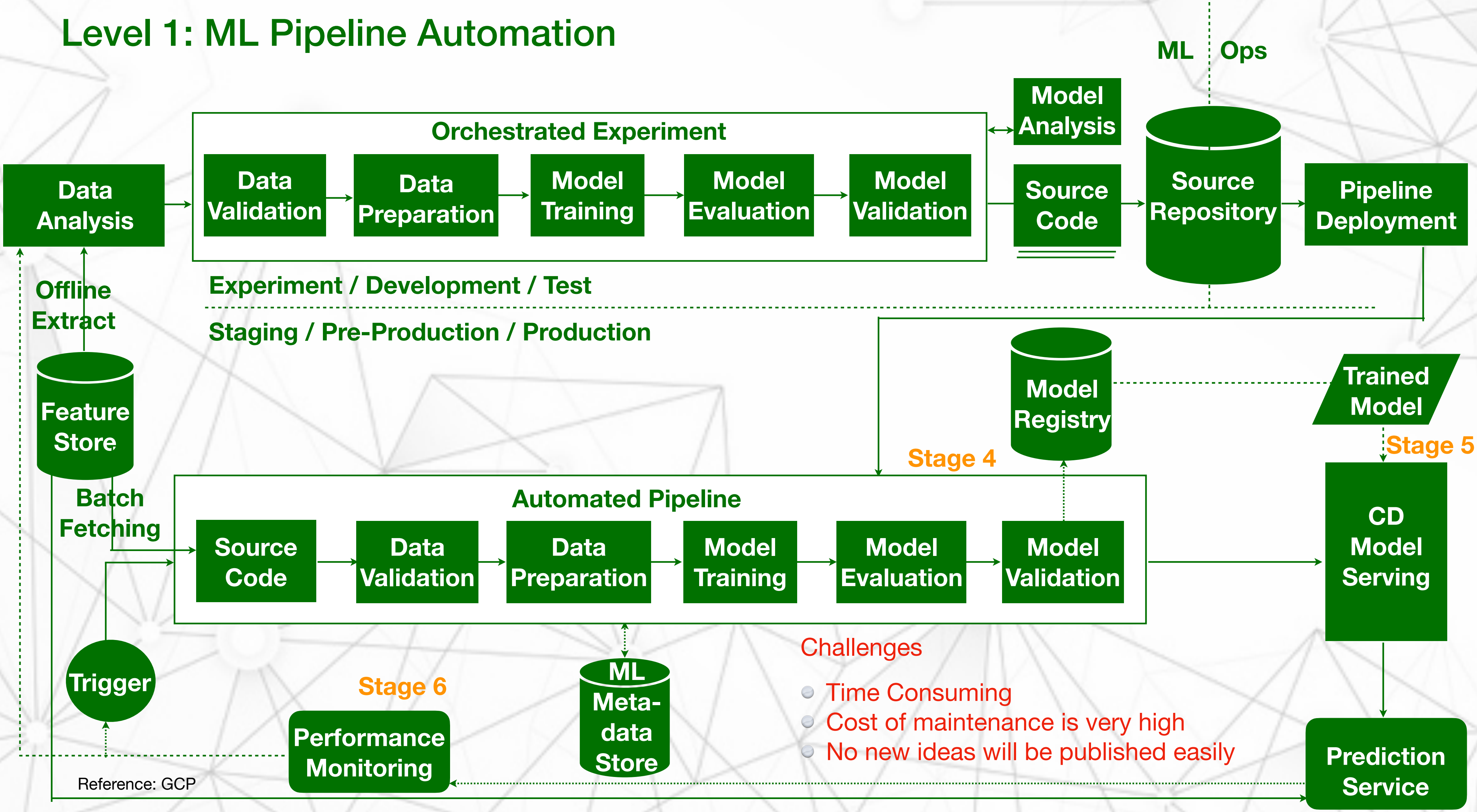


Challenges

- Time Consuming
- Cost of maintenance is very high
- No new ideas will be published easily

Reference: GCP

Level 1: ML Pipeline Automation

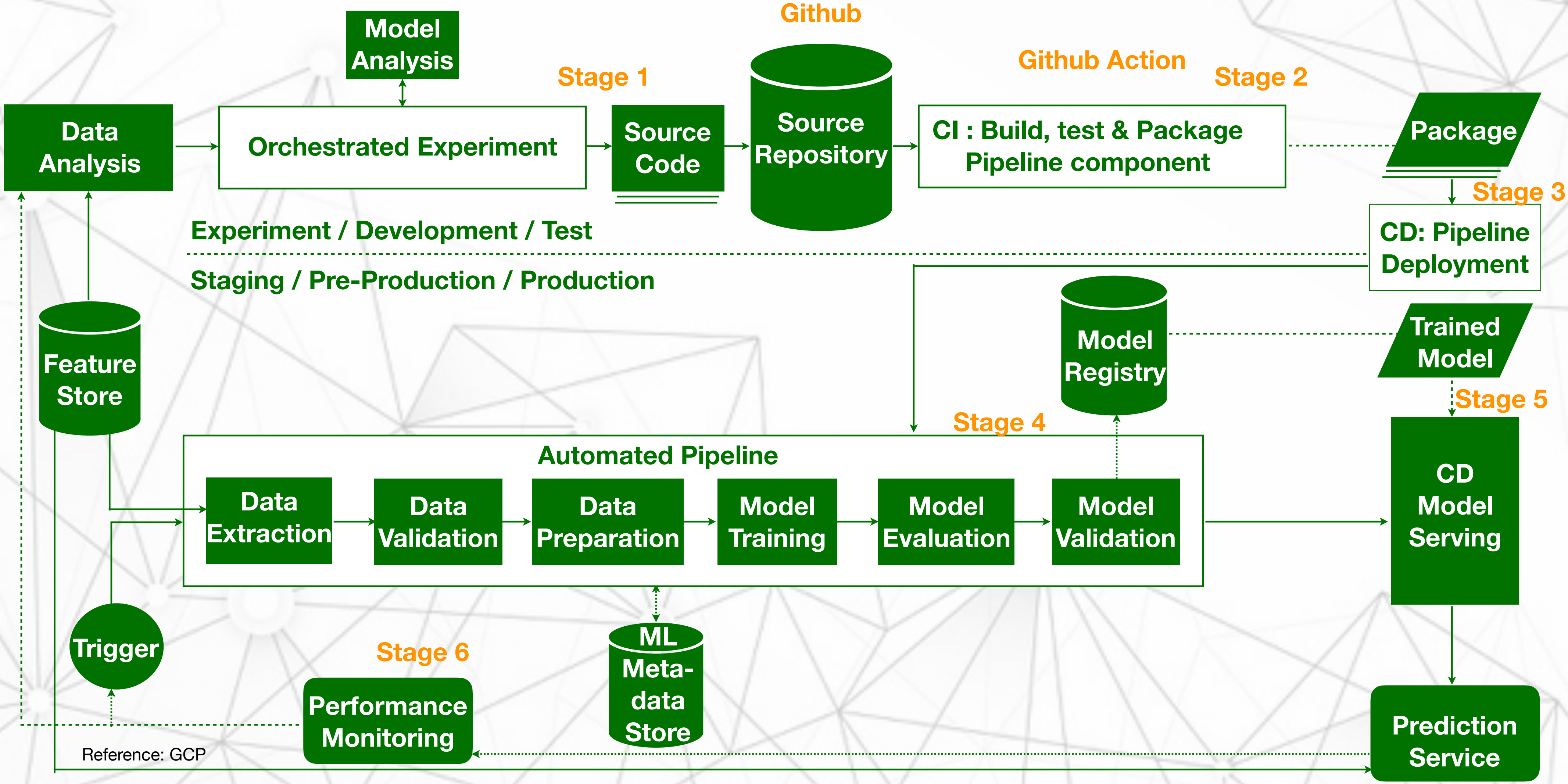


Challenges

- Time Consuming
- Cost of maintenance is very high
- No new ideas will be published easily

Reference: GCP

Level 2: CI / CD Pipeline Automation



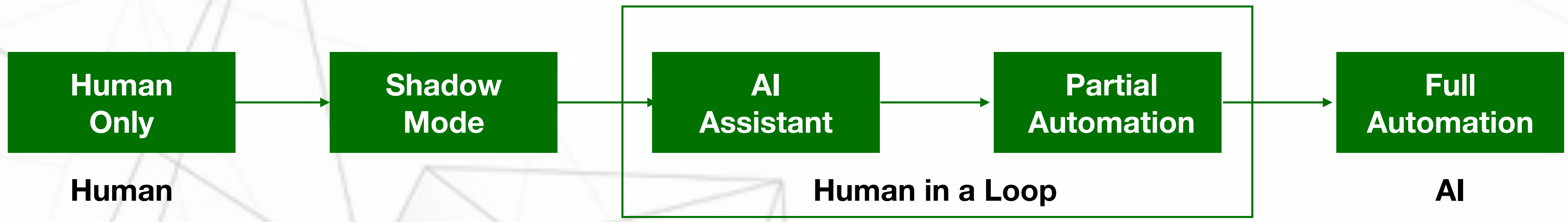
Product Roadmap (18 Months)

	Stage 4 (Q3 2022) Level 1	Stage 5 (Q4 2022) Level 1	Stage 6 (Q1 2023) Level 1	Stage 1 (Q2 2023) Level 2	Stage 2 (Q3 2023) Level 2	Stage 3 (Q4 2023) Level 2
Automation	Continuous Training (CT) & Triggering	Continuous Integration (CI) Pipeline	Contineous Monitoring	Development and Experimentation	Continuous Integration (CI) Pipeline	Continuous Delivery (CD) Pipeline
Input	A deployed ML Pipeline with new ML Ideas	Trained Model		Sample data from feature store	Source code of ML pipeline.	PIPY Package / Docker Container / Desktop App
Process	Triggering the training in the PROD environment based on triggering mechanism (Automated/ Scheduled/Manual)	Pick the suitable model from model registry and integrate with prediction service & prediction pipeline.	Collect the statistics of model performance on new data	Implementation with new ML ideas	Testing	Deployment of the packages in the target environment (Prod / Pre-Prod)
Output	Trained Model, stored in Model Registry server	Working Prediction service	Trigger execution of model pipeline in PROD (less costly) or trigger the new experiment (very costly)	Source code of ML pipeline.	PIPY Package / Docker Container / Desktop App	A deployed ML Pipeline with new ML Ideas

Deployment Strategy

Shadow Mode Deployment	Canary Deployment	Blue green Deployment
ML System shadow the human and run in parallel	Roll out to small fraction (5%) of the traffic initially.	Both old (blue) version and old version up and running connecting through the router. In case of any issue router path change from new to old location
ML system output will not use for any decision making during this phase	Monitor system and ramps trafic gradually.	Easy way to rollback

Degree of Automation



Onboarding Strategy

Awareness
(All Hands Meets, Townhall, Internal Email, Posters)

Interest
(1:1 Meeting with BU or Department Heads, Webinar, Landing Page)

Consideration
(Identifying the stores for canary + Blue green deployment, Trail & Demo to respective employees)

Engagement
(Training, Access to the Environment, Hand-Holding)

Evaluation
(Feedback, Survey)

Purchase
(Start Using The Product & Features)

Key Metrics to Track and Plan Onboarding Strategy

Adoption

- How many users created a itinerary this quarter?
- How many new users requested for Trial this quarter?
- How many times use the trial this quarter?
- How many users request for second trial this quarter?
- How many user started using the product this quarter?

Usage

- Show Demo to 2000 users in every months.
- Improve our average user (employee) rating by one star on major product-review sites
- Monthly/Quarterly Usage Trend
- Frequency
- Sessions

Engagement

- Monthly Active Users
- Product Demo
- Product Sales, Training & Support
- Webinar
- Survey
- Social Media Forum
- Referral/Advocate of our Product

Track

- Increase trial & training adoption by 2000 employee in the every months across location.
- Getting users to use the service often and behave in a way that helps the users and business.
- User (employee) rating by one star or Net Promoter Score (NPS)

My Skill Competency

My name is **Asit Piri,** and I'm passionate about developing digital products and solutions that are technologically feasible and strategically viable, as well as addressing users' actual needs and serving as a product evangelist to inform users about how the product will simplify their lives.

My career **purpose is to develop products and solutions that improve the way people live and work**. My **strength came from two decades of solid information technology experience** that have helped me understand the interdependence of various functions in global technology organisations and businesses when it comes to solving large-scale customer problems. In the last few years, along with my technical skills I've developed a few characteristics, such as passion, rigor, empathy, critical problem solving ability, and a design thinking approach, which I believe are assisting me in becoming an exceptional product manager.

Thank you!

Level 2 Automation: CI/CD Automation Pipeline Stages

Stage 1: Dev and Experimentation Stage
Input: Sample data from feature store
Process: Implementation with new ML ideas.
Output: Source code of ML pipeline.

Stage 2: Pipeline Continuous Integration (CI)
Input: Source code of ML pipeline.
Process: Testing
Output: PIPY Package / Docker Container / Desktop App

Stage 3: Pipeline Continuous Delivery (CD)
Input: PIPY Package / Docker Container / Desktop App
Process: Deployment of the packages in the target environment (Prod / Pre-Prod)
Output: A deployed ML Pipeline with new ML Ideas

Stage 4: Continuous Training (Automated Triggering)
Input: A deployed ML Pipeline with new ML Ideas
Process: Triggering the training in the PROD environment based on triggering mechanism (Automated/Scheduled/Manual).
Output: Trained Model, stored in Model Registry server

Stage 5: Pipeline Continuous Integration (CI)
Input: Trained Model
Process: Pick the suitable model from model registry and integrate with prediction service & prediction pipeline.
Output: Working Prediction service

Stage 6: Monitoring
Process: Collect the statistics of model performance on new data
Output: Trigger execution of model pipeline in PROD (less costly) or trigger the new experiment (very costly)

Level 0 Automation: Characteristics, Challenges, Solutions

Characteristics

- Suitable for MVP or Prototyping.
- Manually build, train, tune & deploy
- Manual script driven
- Manual execution of each step
- Disconnection between ML and Operations.
- It mainly code written in Jupyter Notebook
- Training (n feature to prediction) and serving (n+1 feature to prediction) skew
- Infrequent release of models
- No continuous development of models
- Only few models will be released
- Lack of continuous integration
- Deploy only web app + model, not able to deploy end to end Machine Learning systems.
- Lack of performance monitoring

Challenges

- Time consuming process.
- Cost of maintenance is high.
- No new ML ideas will be published easily.
- Prone to bugs

Solutions

- Active monitoring of model in production.
- Frequently retraining the model (depending on the situation).
- Frequent experiments with new optimised implementation to produce better model (e.g. trying out latest state of the art (SOTA) models.
- To make it more adaptable to the business need.
- Automate as much as possible.

Level 1 Automation: Characteristics, Challenges, Solutions

Characteristics

AIM:
- Rapid experimentation is possible due to orchestration (order execution of steps).
- Continuous Training (CT) of model done in production.
- Experimental - operational symmetry.
- Modularise code (reusable & sharable) component and pipeline.
- Continuous Development (CD) of model
- Pipeline deployment -> deployment of model.
- Frequently used Terms:
  - Pipeline Orchestration
  - Tools: Apache Airflow, Kubeflow

Feature Store:
- Central repository used for both training and serving
- Provide API for high throughput batch serving and low latency real time serving of feature value and supports trying and serving workload.

Metadata Store:
- Record information about each execution of ML Pipeline to help in:
  - Reproducibility
  - Comparison
  - Artifact lineage
  - Debugging and anomalies

Challenges

- Time consuming process.
- No frequent update
- Maintenance cost is high (of course less than level o)
- Prone to bug

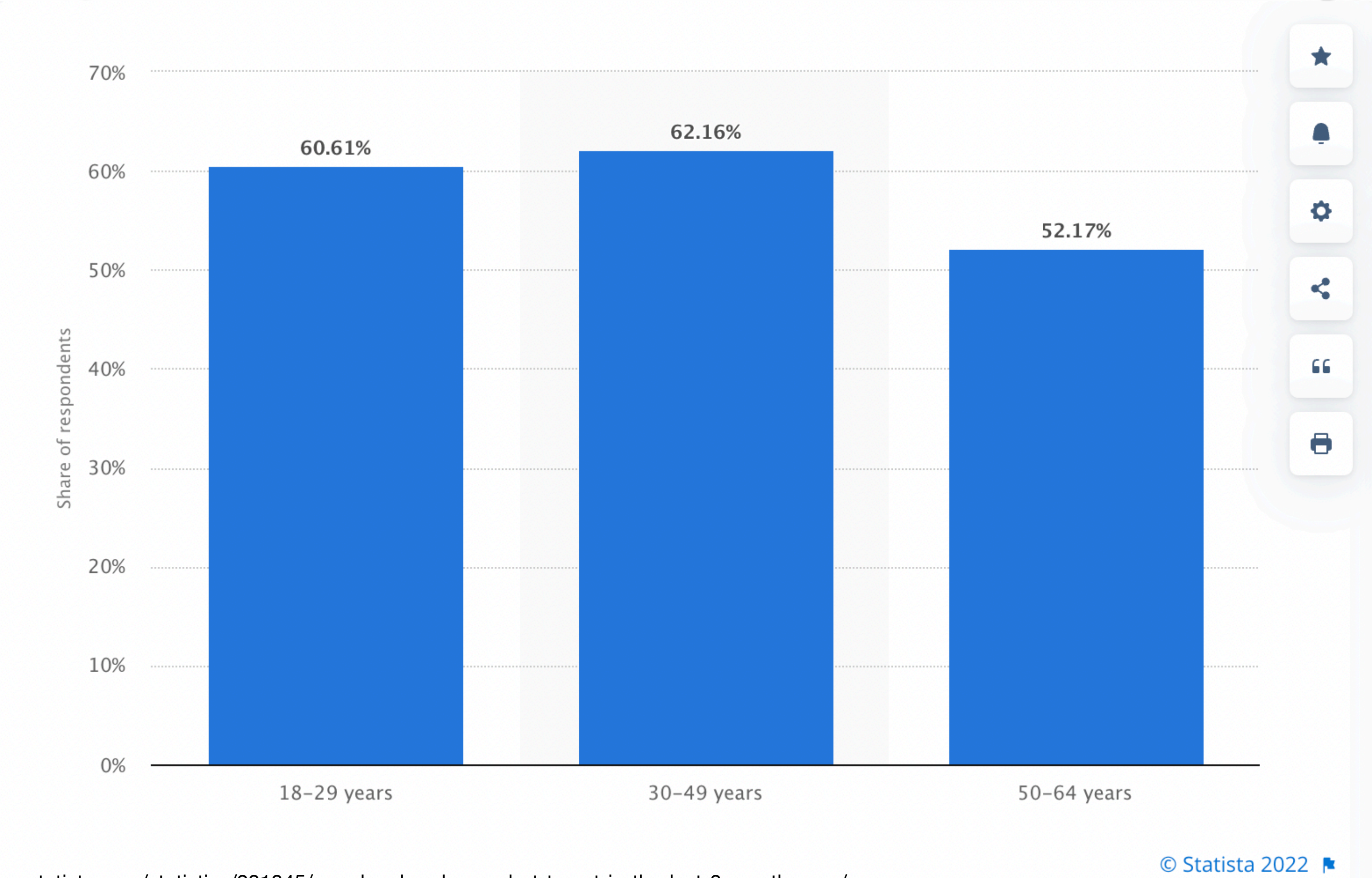
Solutions

CI / CD Pipeline automation set to:
- Build
- Test the ML pipelines
- Deployment

Level 2 Automation: Characteristics, Challenges, Solutions

Characteristics	Challenges	Solutions
Aim: • Rapid & reliable updates of ML pipeline. • Automating build/test/deployment of ML pipeline (beneficial for DS team, can able to focus on new ideas) • New ideas can be brought into production easily and quickly. CICD Stages •	• Time consuming process. • Cost of maintenance is high. • No new ML ideas will be published easily. • Prone to bugs	• Active monitoring of model in production. • Frequently retraining the model (depending on the situation). • Frequent experiments with new optimised implementation to produce better model (e.g. trying out latest state of the art (SOTA) models. • To make it more adaptable to the business need. • Automate as much as possible.

Share of people who shopped at Target Corporation in the United States in 2018, by age



Source: https://www.statista.com/statistics/231345/people-who-shopped-at-target-in-the-last-3-months-usa/