

CBINSIGHTS
AI
100
2023

AI development
AI chips & processors
Vector database tech
ML development deployment

Cross-industry applications

Synthetic voice iElevenLabs	Image & text generation Jasper MidJourney	Privacy & security HiddenLayer Protect AI RELEVANCE AI	Code generation diffblue replit	Digital twins blackshark.ai
Productivity tools mem	Warehouse & logistics covariant	Content moderation Logically unitary	Smell tech osmo	

General-purpose humanoids
FIGURE
SANCTUARY AI
Image processing
Metaspectral
VISIONARY

Industry-specific

Materials & manufacturing Cradle MakindRocks	Gaming convai MODULATE	Fashion & retail refiberd seel	Energy phaidra We
Defense VANNEVAR Labs	Finance arteria Boosted.ai	Agriculture	
Education ello	Media & entertainment character.ai descript runway wonder		

Note: Companies are private as of 6/20/23.

HOW MANY ARE INVESTING IN AI?

HOW MANY WILL SURVIVE?

2021 Data Science & AI: Startups & Scale-ups Rotterdam - The Hague area

General AI Services ASSETHUB INTERACTIVE ROBOTICS GILO Zoi Meet splendo cevinio mavim GEOPHY VORTECH Birds.ai mafinnora Orchest EDGE landscapE supplai LIBROSO Breeze DOMAINS starTXT.ai BREEZE doculayer	Manufacturing fastree 3D s[&t AERIALTRONICS ReliaSol UReason MAINBLADES PARAFY VIBES BIOCKBAX Somatics OptiPort Teaplay ROADEO flyingfish BARGE MASTER FLUMENSIS TECHNOLOGIES Teleport	Life science and health primagen toxys BRISE VitaSense Medis Quantib Ocell OmniGen Genome Scan sommox envison speakee BlueWalk Genome Scan sommox envison speakee Holland AI Ncardia NightWatch
---	--	--

AI LANDSCAPE 2022 AUSTRIA

SEMIOTIC LABS
S[&t
widgeTbrain
S2E
GERON
POLARIKS
Disdro
oddbot

INDIAN AI STARTUP LANDSCAPE

getzio
glana
Lexplora
NORNA
GORDIAN
homecraft
infobateen
Linkura
NORTH LINK
embed
GOAVA
LOGGAMERA
tvinn
happyr
iMetrics
monok
SUSTAINLAB
KENSE VISION
iPercept
HOPSWORKS
LUCERO
medotemic
monok
NUTRIE
TENFIFTY
Ultimefeed
SENSATIVE
REBASE
Therr
Sciling
modulai
tur

SAND
TERLU rek.ai
home
MAIGON
Pandora AI
Tendium
SENTIAN AI
Spotscale
pirr
UNIVERSES
OXIDE.AI
SYNDATA
VOXO
SA
VERAI
ZENTOSA
Weatherlabs

Amberscript
BRAINIAL
doculayer
CIPHIX
yarado
supplai
freeday
klippa
triple
WORKFLOW
AsiArms
LIBOS
Hemisphere
Storify
OPT/NET
SYNTHO
FINTECH
floryn
whayle
Owlin
fourthline
BRIDGE FUND
biller
Atama.AI
sentinels
Federatō
MARKETING & SALES
symson
visualfabriq
Textmetrics
Zoi Meet
WeGain
ROBOTICS & HARDWARE
loop robots
SMART CITIES & SUSTAINABILITY
Spotr.ai
overstory
soblott
CamenAI
KAIOS AI
PLANLOGIC
jungle
CO2.ai
GEOPHY
HAL 24K

© 2023 Endent Capital. Feel free to share. Questions, praise or your company added? Please mail to remy@ai.nl



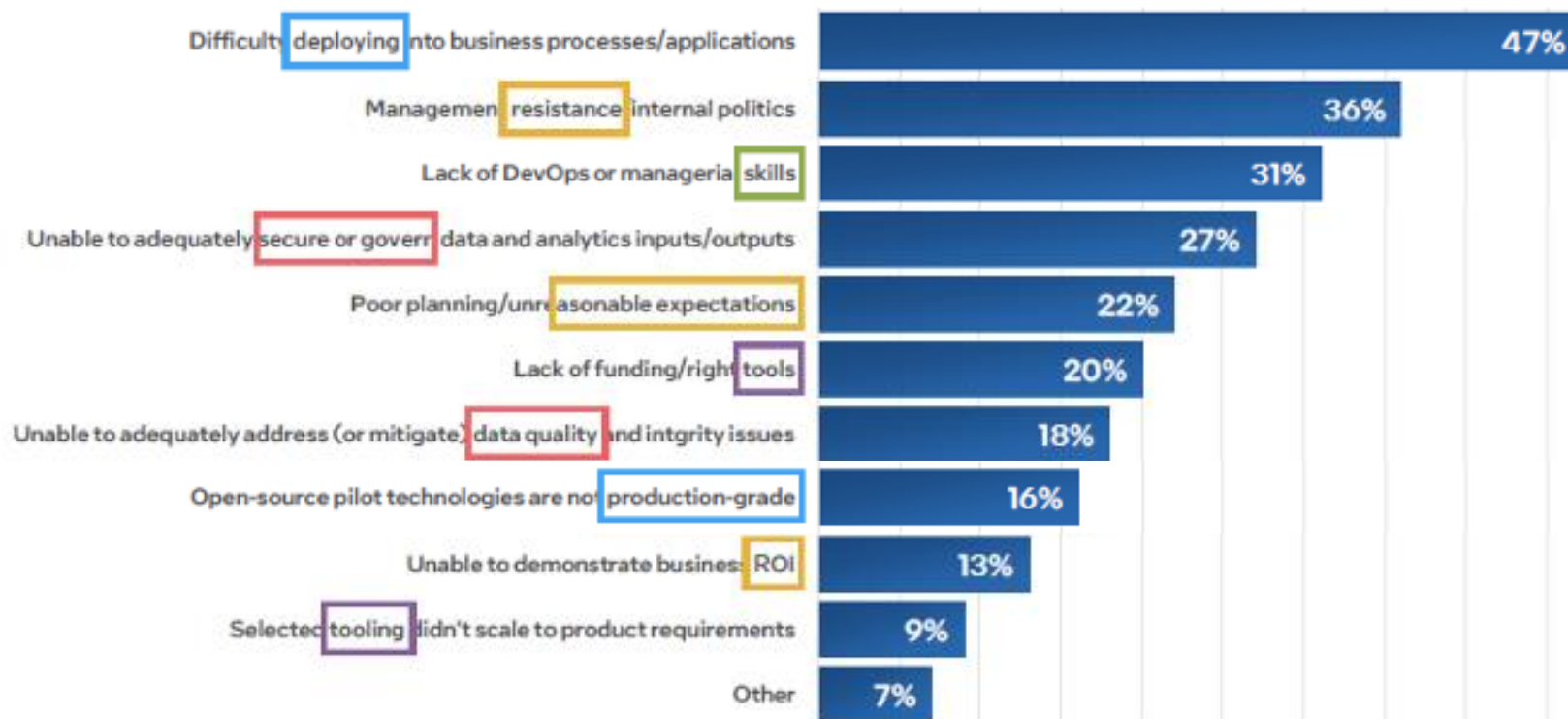
More than 80% of AI projects get stuck in the lab or produce partial success.

Productizing ML is one of the biggest challenges in AI practices today.

* Source Gartner

Why AI Projects Fail*

Production is the main barrier
towards delivering business value



Legend

- Skills
- Selection & execution
- Tools
- Governance
- Productization

"Productizing ML is one of the biggest challenges in AI practices today. Many AI projects, more than 80% according to research, get stuck in the lab, produce partial success, or consume far more resources and time than initially planned."

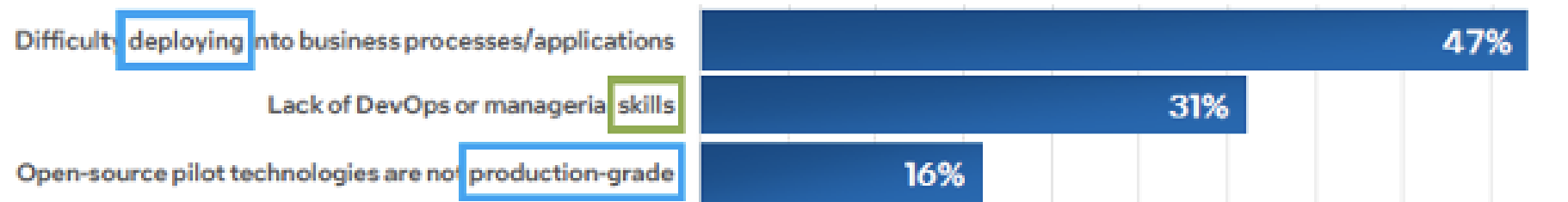
* Source: Gartner

Why AI Projects Fail*

Production is the main barrier
towards delivering business value

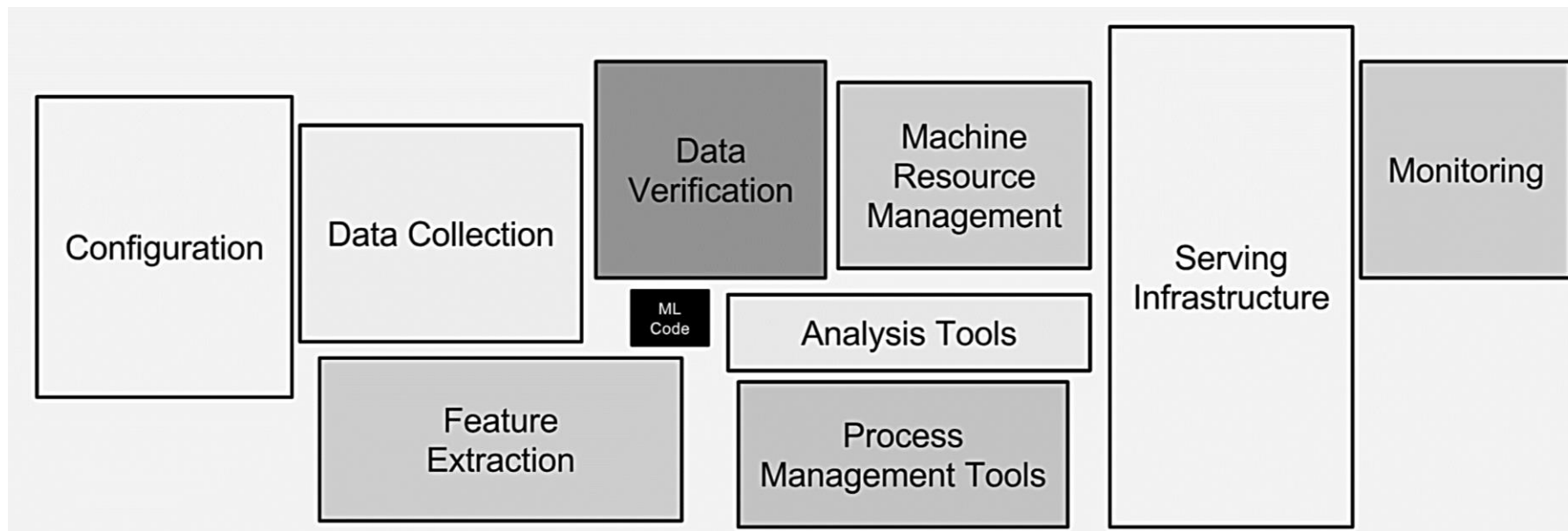
Legend

- Skills
- Selection & execution
- Tools
- Governance
- Productization



* Source: Gartner

“Only a small fraction of real-world ML systems is composed of the ML code” Google, inc.



* Sculley, David, et al. "Hidden technical debt in machine learning systems." *Advances in neural information processing systems* 28 (2015).



Evolutionary Changes in Data Analysis (ECiDA)



MOSTAFA
HADADIAN



ALEXANDER
LAZOVIK



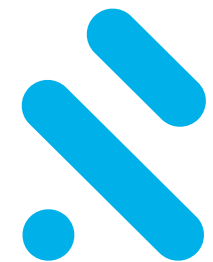
VIKTORIYA
DEGELER



rijksuniversiteit
groningen



UNIVERSITEIT VAN AMSTERDAM



RESEARCHABLE

Mission

- Simplify engineering tasks to let scientists concentrate on coding.
- A **Lifecycle Management System** specifically designed to support **Real-time Data Processing**, with enhanced **Modularity**, using **Microservices**.

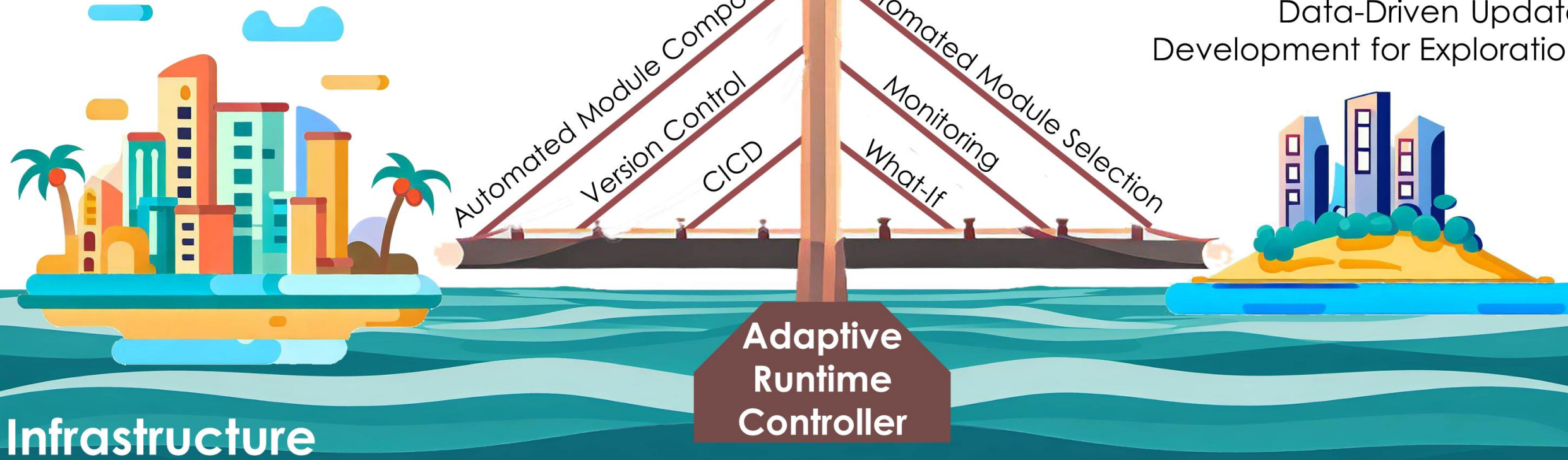


Service-Oriented Computing

Mature
Clear Scope
Predefined Logic
Development for production

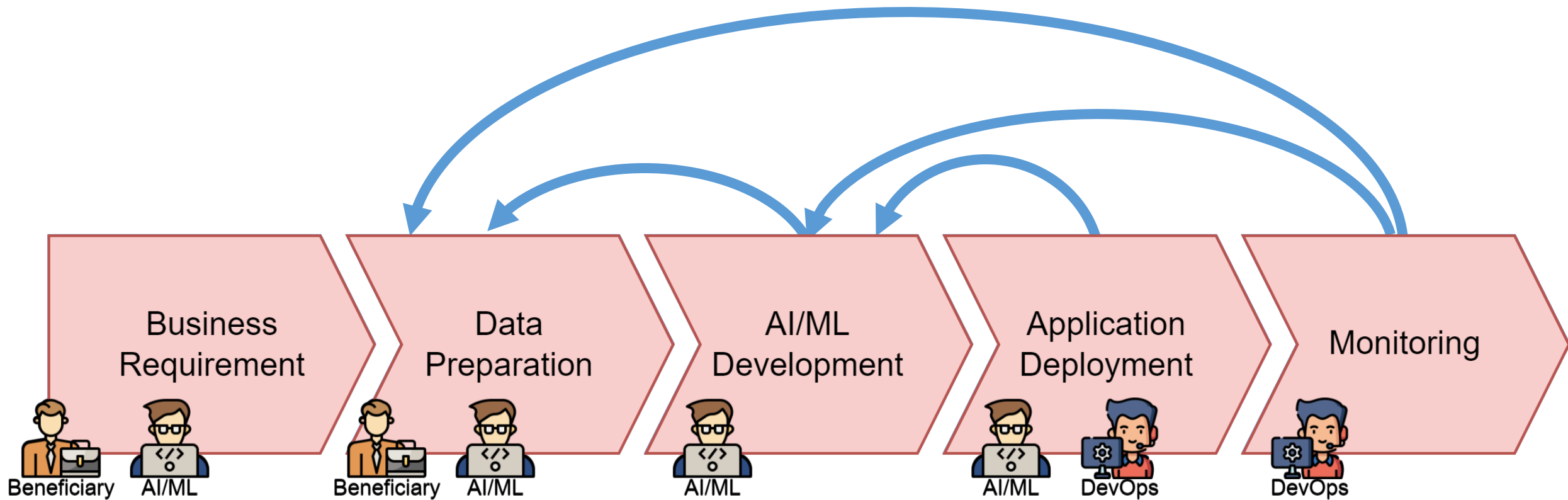
Machine Learning Development

Evolving
Probabilistic
More Iterative
Evolving Scope
Data-Driven Update
Development for Exploration



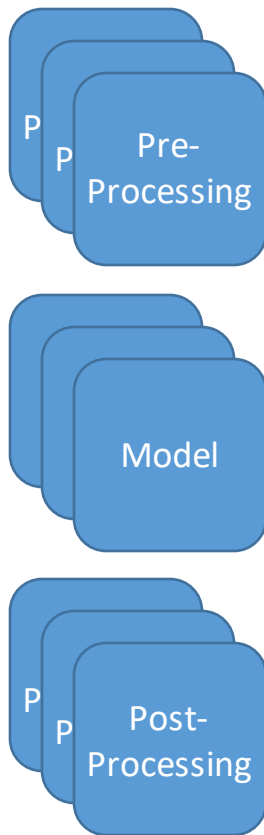
* Hadadian Nejad Yousefi, Mostafa, et. al. "Empowering Machine Learning Development with Service-Oriented Computing Principles." *Symposium and Summer School on Service-Oriented Computing*. Cham: Springer Nature Switzerland, 2023.

Coarse-Grained AI/ML Lifecycle

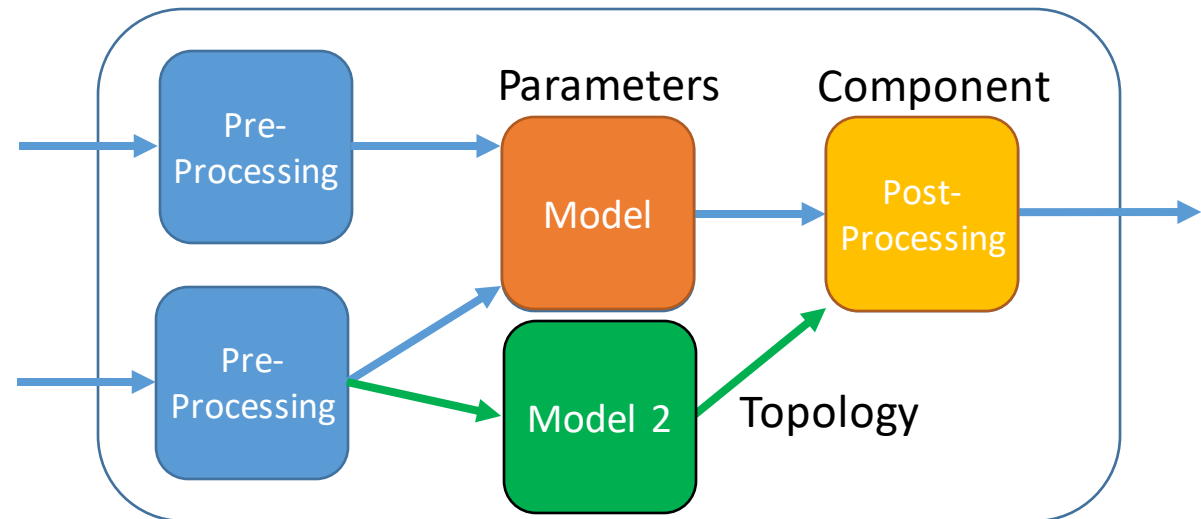


Data Processing Pipeline

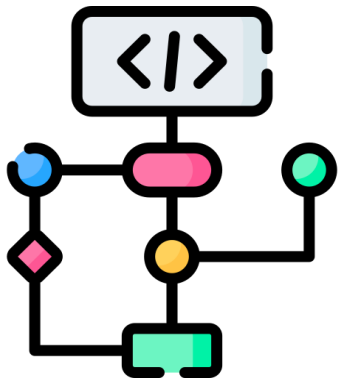
Individual Component



Pipeline

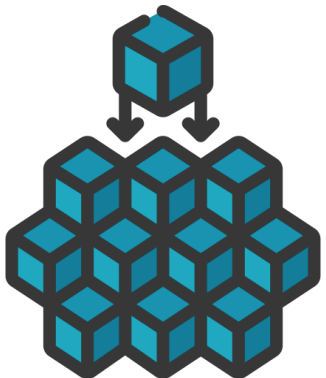


Modularity By Design



Algorithmic Modularity

- Utilization of programming languages or frameworks for the development of machine learning applications



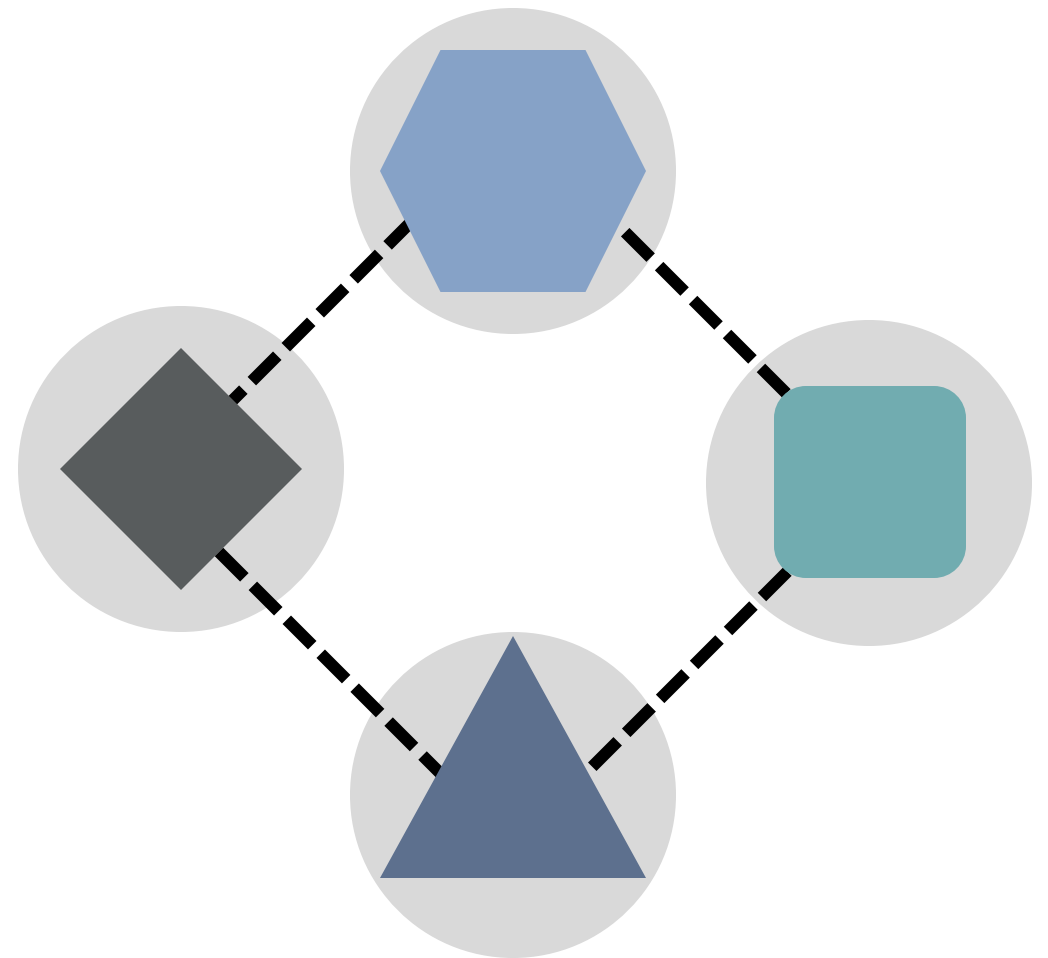
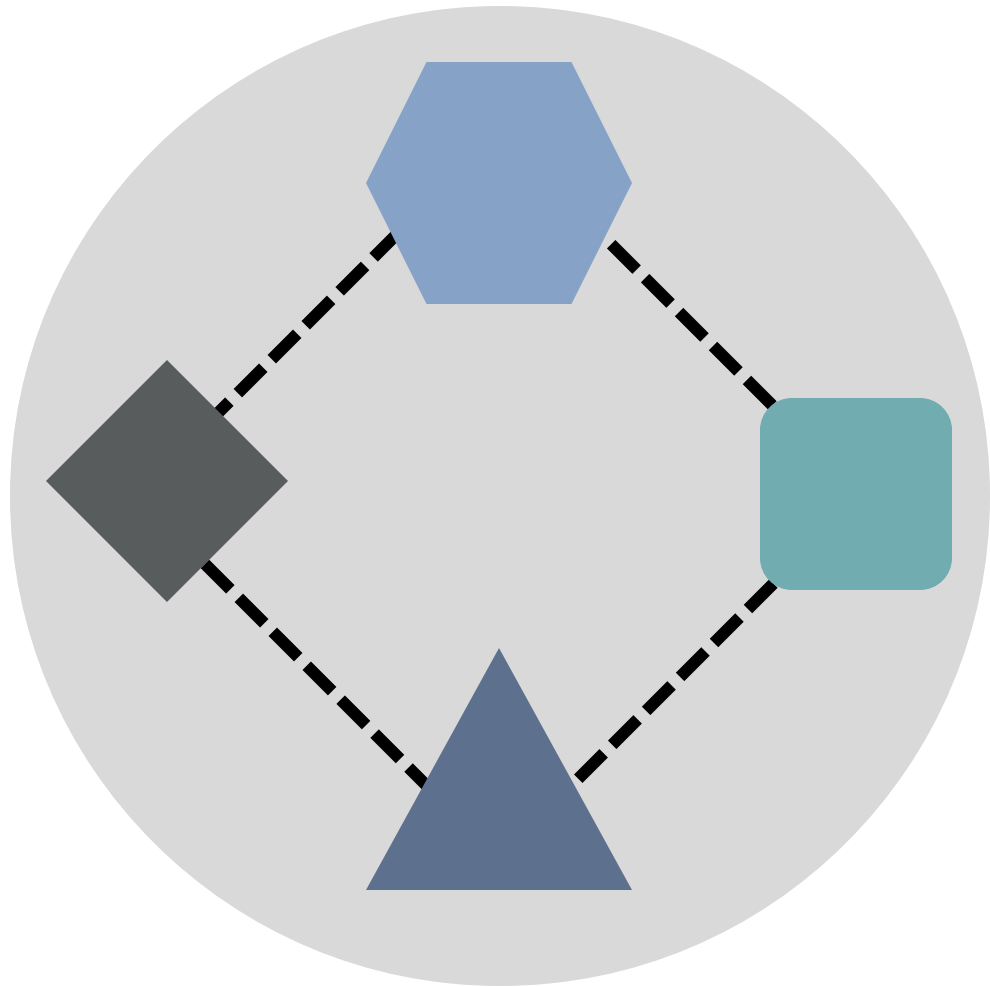
Architectural Modularity

- Packaging each stage into distinct module and deploying these modules into appropriate software environments

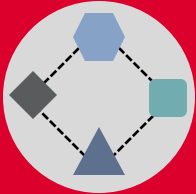
Monolith

vs

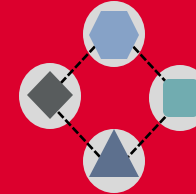
Microservice



Monolith



Microservice



Pros

- Easy to develop
- Easy to deploy
- Less prone to network errors

- Better Modularity and Visibility
- Scalability
- Technology Diversity
- Continuous Updates

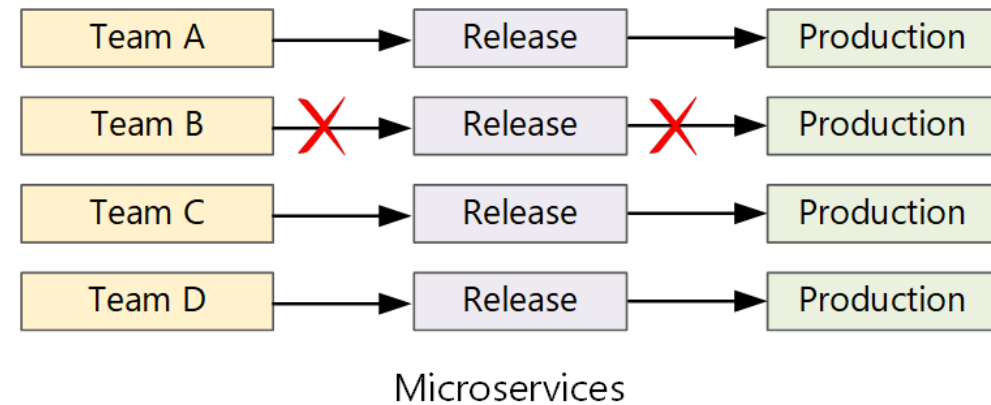
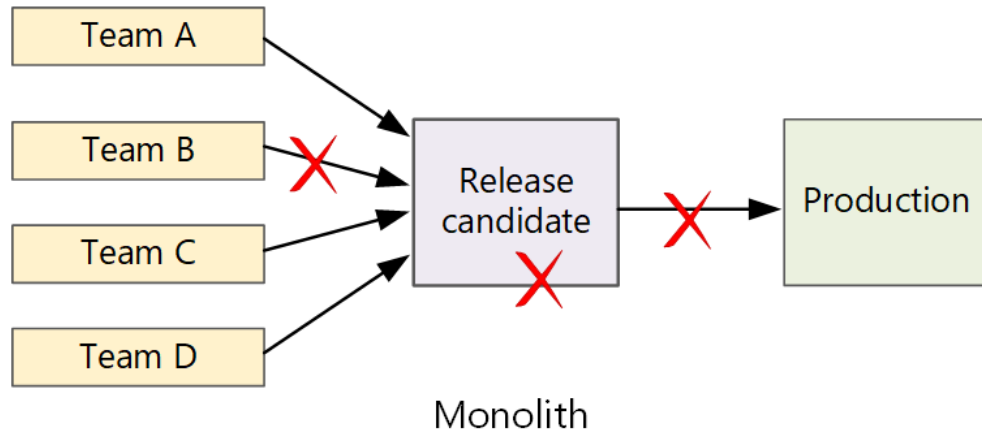
Cons

- Lack of isolation
- Reusability
- Hard to minor updates and patches.

- Operational Overhead
- Hard to debug
- Hard to share code

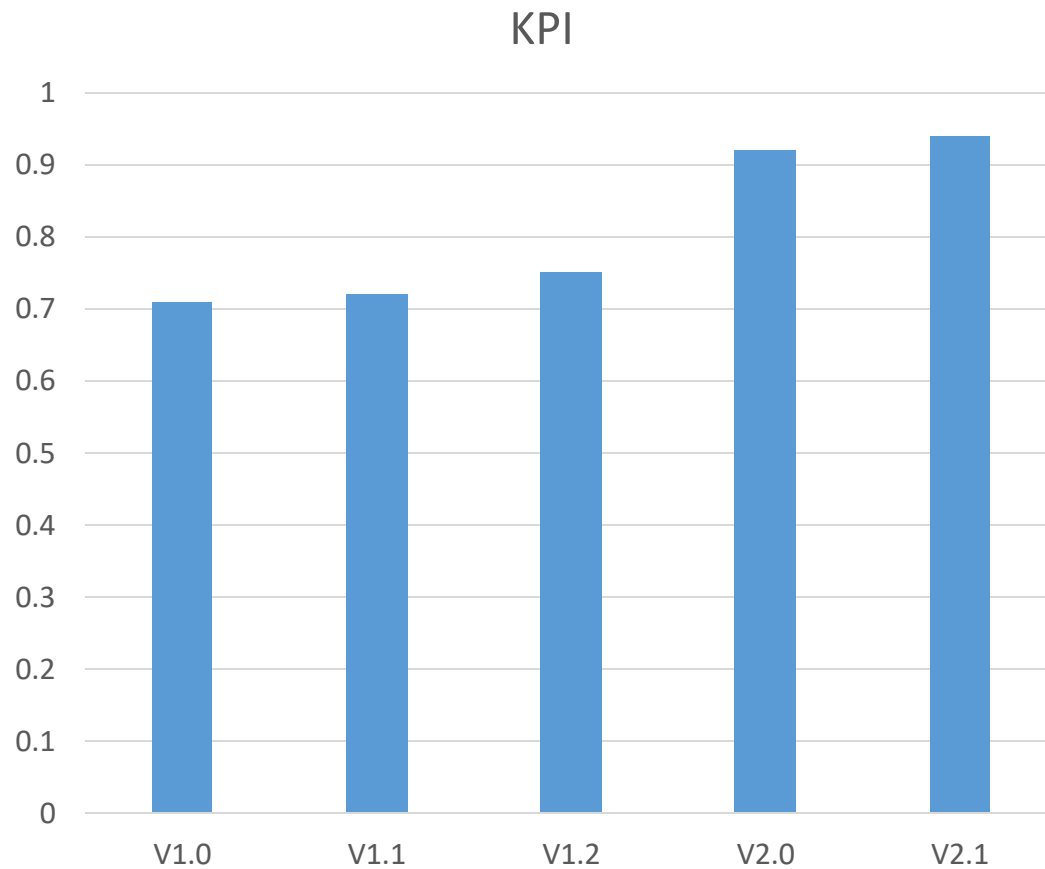
* <https://www.infoq.com/news/2020/04/microservices-back-again/>

CI/CD in Monolith vs Microservice



* <https://docs.microsoft.com/en-us/azure/architecture/microservices/ci-cd>

Classical Software Evolution



- Incremental improvement
- Definite input/output



AI System = Code + Data

Code = Algorithm/Model

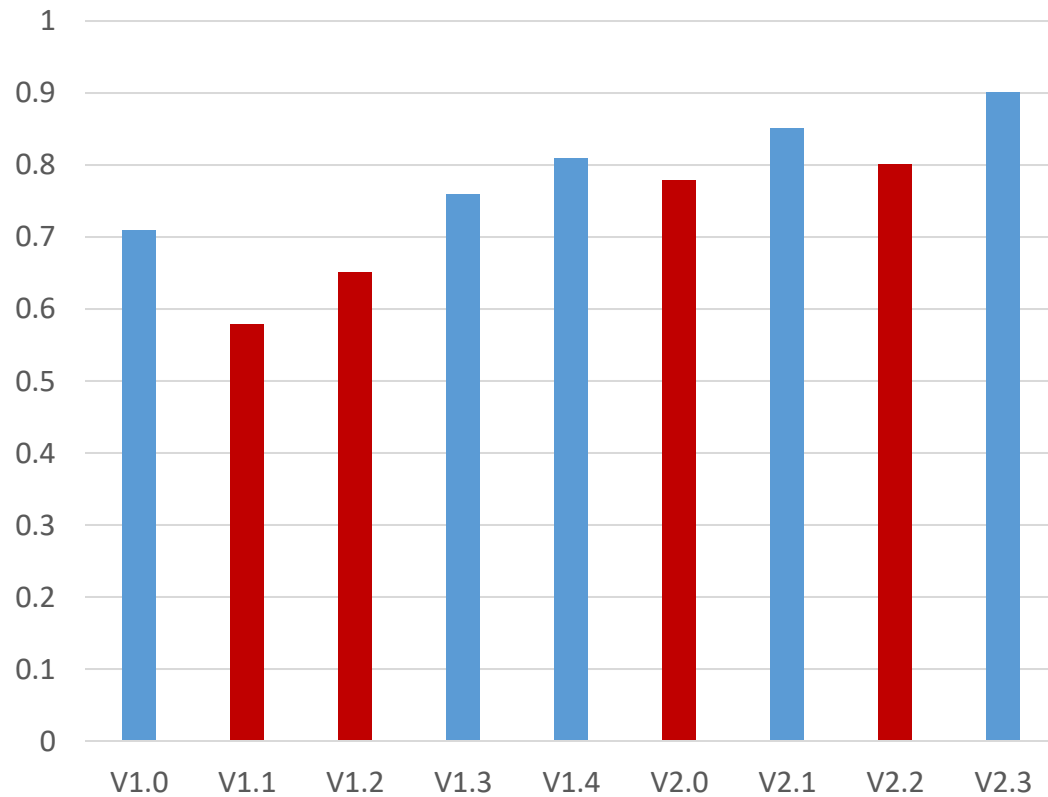


Hyperparameters

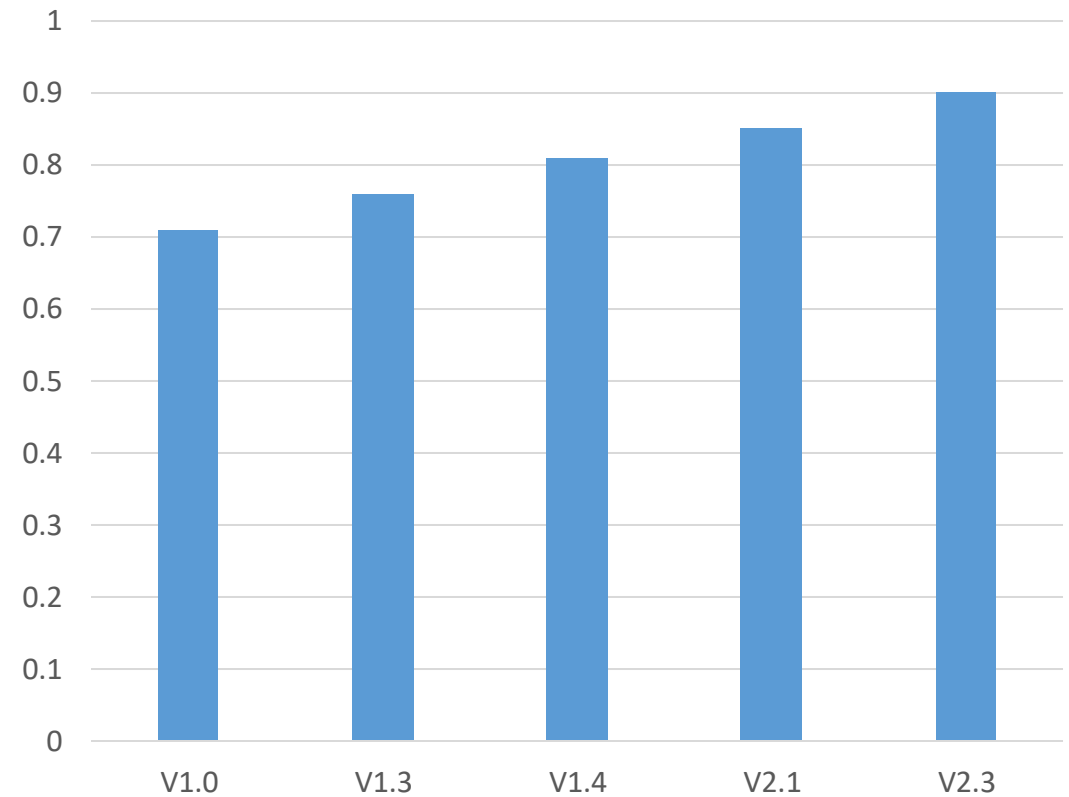


AI System Evolution

KPI



KPI (Production)

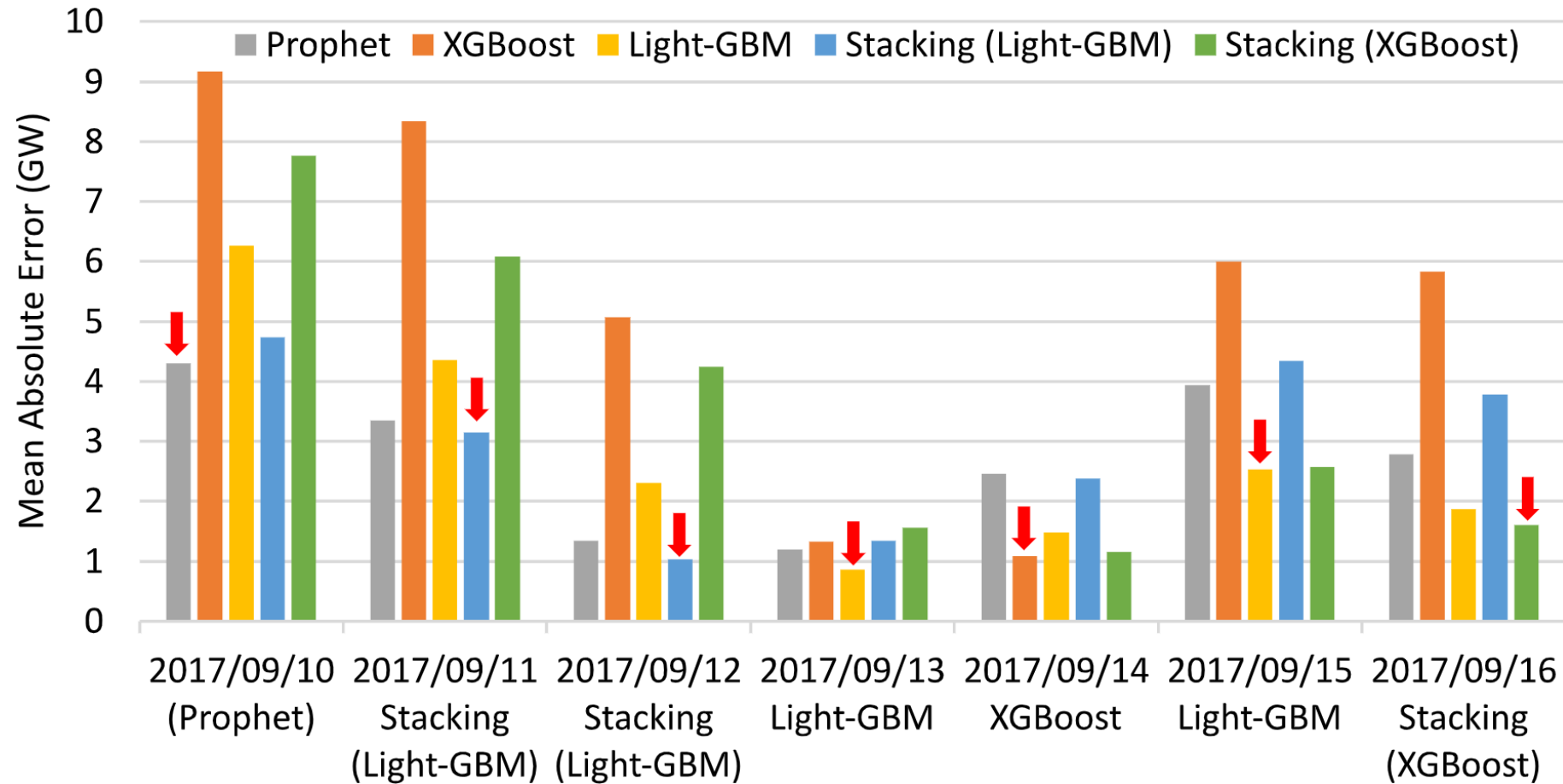




Monitoring Distributed ML Systems

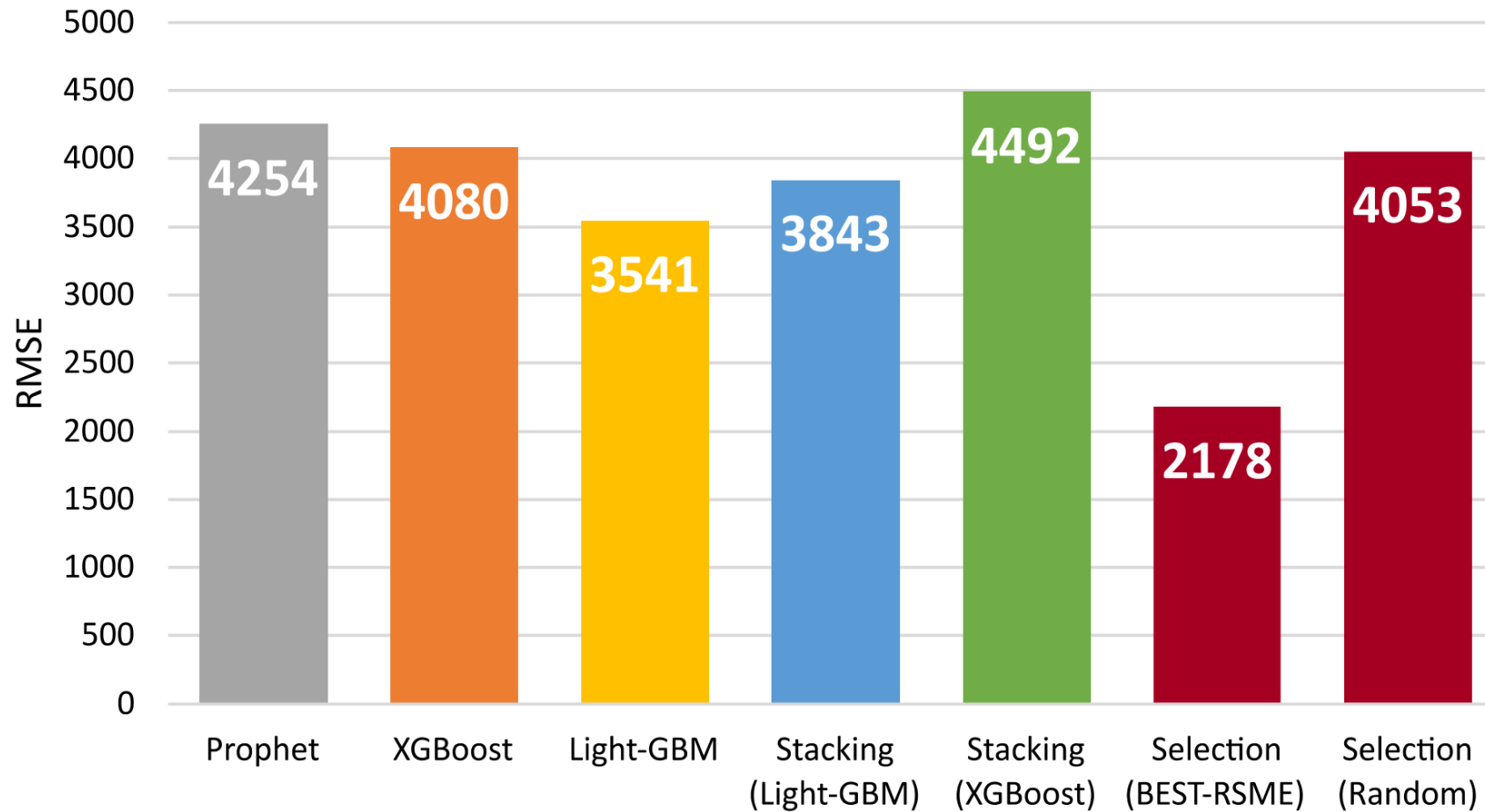
- ML workflow monitoring:
 - Model monitoring
 - Performance Metrics (Accuracy, Precision, Recall)
 - Concept Drift: Change in the relationship of inputs and targets over time
 - Data Monitoring
 - Outlier Detection
 - Data Drift: Statistical Distribution Changes
- Resource usage monitoring
 - CPU, memory, and disk I/O

Which one do you pick?



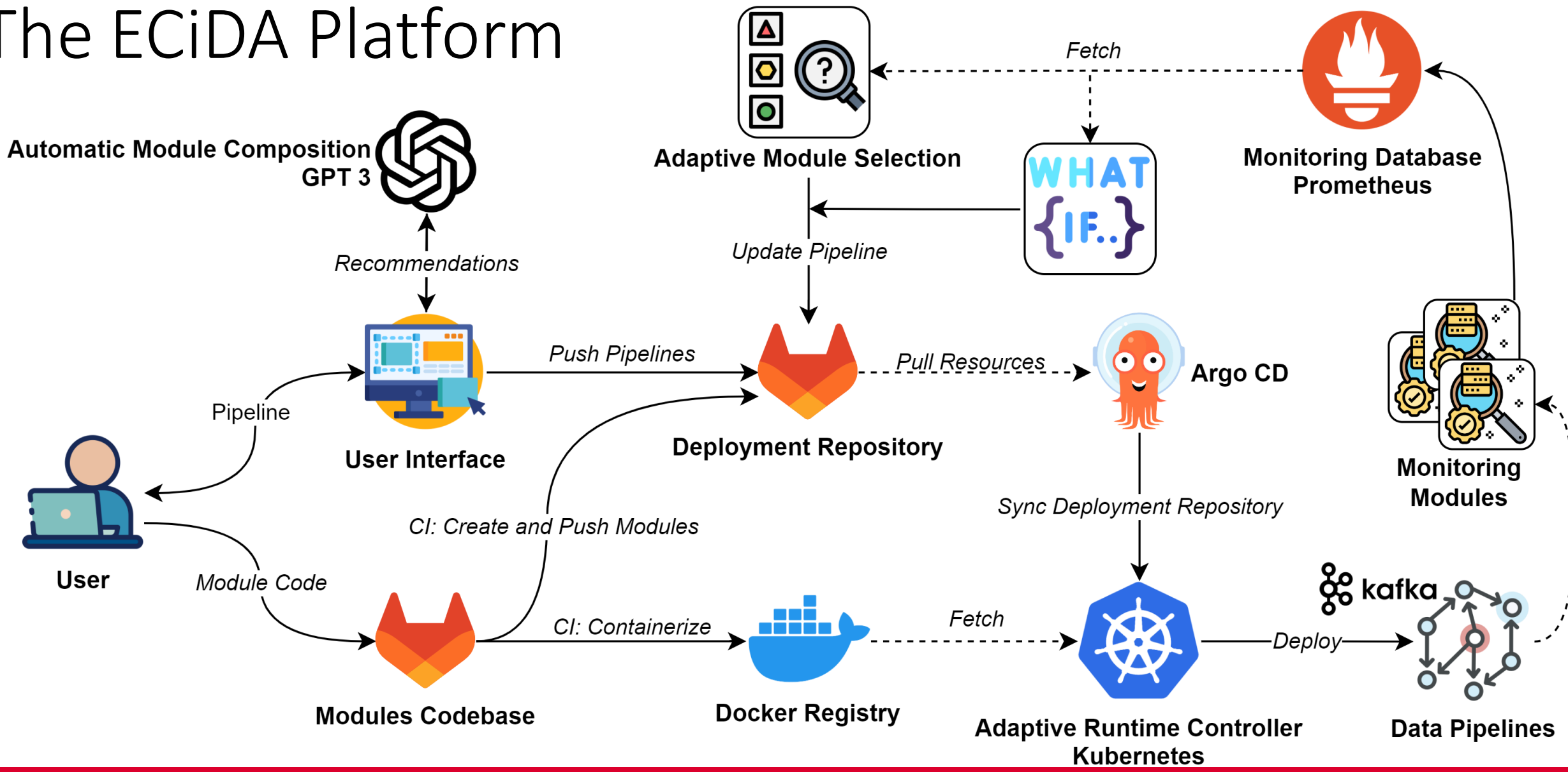


Selection Results





The ECiDA Platform





User Interface

ECiDA > Project Name > Edit Project > consumer-v2 Discard Save

Add modules

Filter Match all

Scaler Monitor env:management Clear

Search for definitions

Implementations (0 selected) Multi-select

- GeneralScaler
- MinMaxScaler **Popover title**
Use the adaptive module selector to find a fitting Scaler
- GeneralMonitor
- Adaptive Scaler
- Adaptive Monitor

Clear Selection Add

Canvas

Component Data Consumer

Module: consumer-v2

Standard implementation of a subscriber in the publisher-subscriber pattern. Provides simple printing of any data sent

Monitor

Suggest

Summary Inputs Outputs +

income	time-series
harmonic_mean	int

new_me|

Get a pipeline suggestion by giving a description of the desired result.

Discard Suggest

Pipeline > Main



Summary

- Code alone doesn't make an AI project.
- Engineering aspects of data science are overlooked
- AI is being used to streamline its own development and deployment.
- There are many specifics for that need careful considerations
 - Pipeline Composition
 - Life Cycle Management
 - Training vs Inference
 - Versioning
 - Continuous Monitoring
 - Experiment Tracking
 - What-if Scenarios



With AI comes great maintaining efforts

use it wisely

Product Launch

 | 

KubeCon | **CloudNativeCon**

Europe 2024

#KUBECON + #CLOUDNATIVECON

19 – 22 MARCH, 2024
PARIS, FRANCE
PARIS EXPO PORTE DE VERSAILLES

University of Groningen
Distributed System Group
Mostafa Hadadian Nejad Yousefi
m.hadadian@rug.nl

 CONNECT

