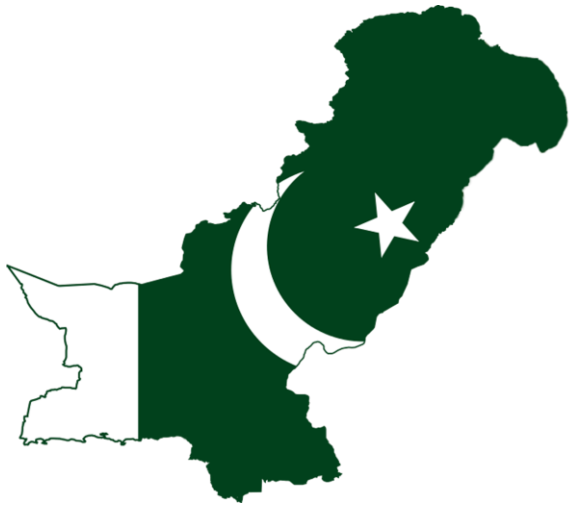
CORVIT
Reaching end to end

PRIME MINISTER'S Kamyab Jawan

کامیاب جوان

NATIONAL YOUTH DEVELOPMENT PROGRAMME

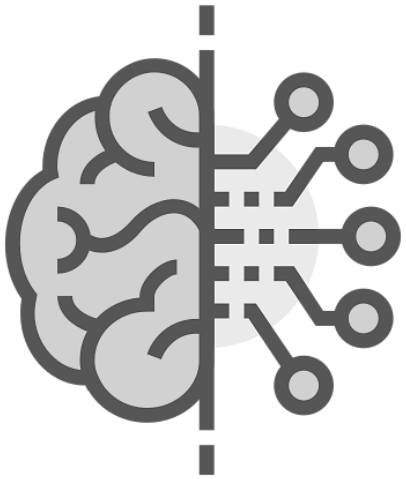
کامیاب جوان کامیاب پاکستان



Skills for all - Batch III 2022

Artificial Intelligence (Robotics)

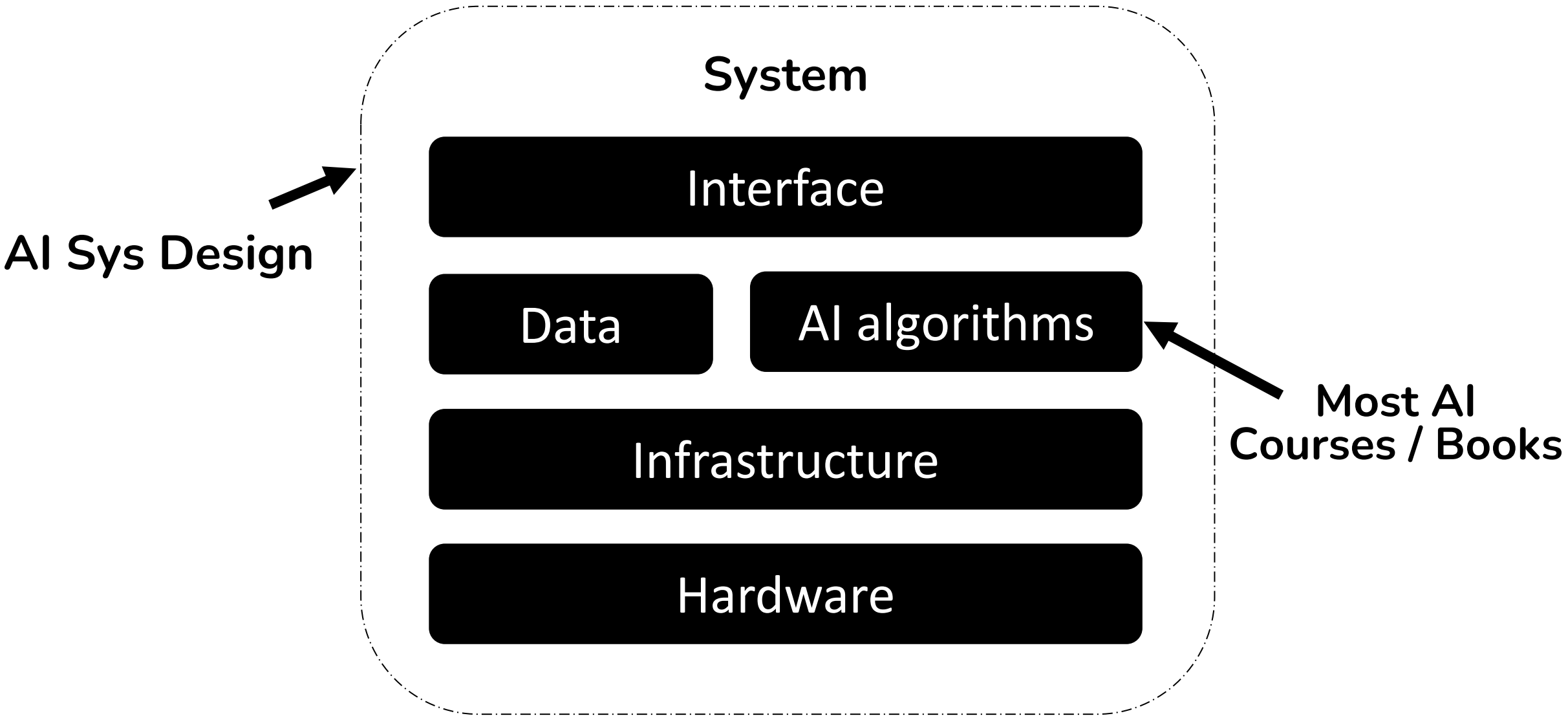
Lecture 1: Introduction



Muhammad Rizwan – Corvit Systems

Artificial Intelligence: Expectations

This course won’t teach you how to do this 😂


System
Interface
Data
AI algorithms
Infrastructure
Hardware
AI Sys Design
Most AI
Courses / Books

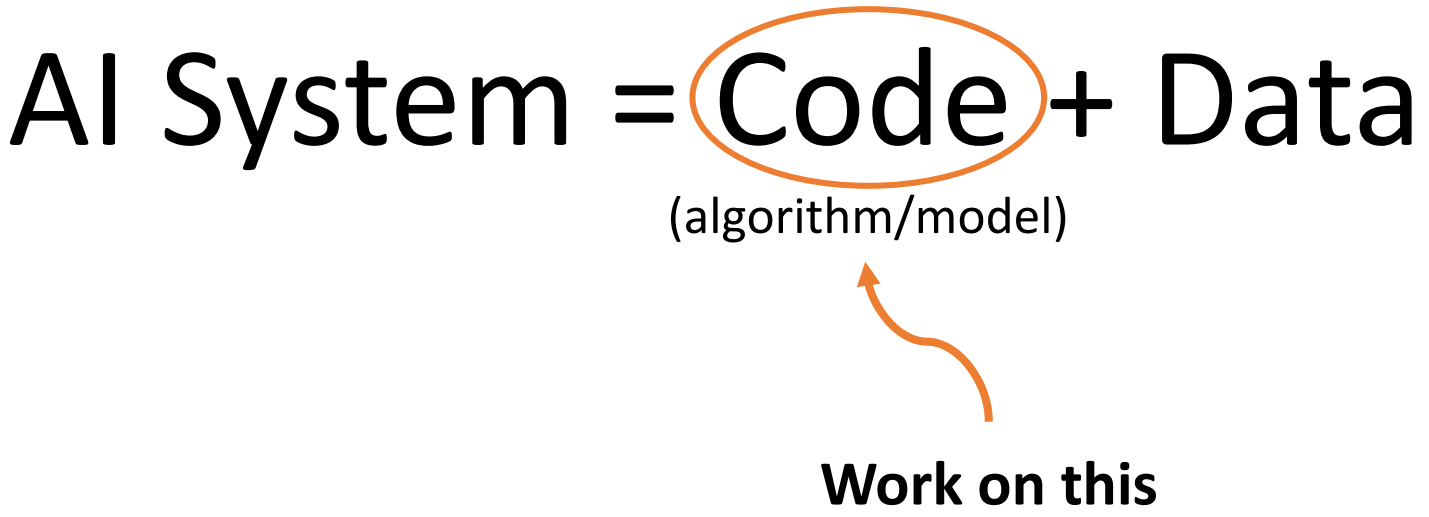
Data-Centric AI

Data-Centric AI

- Major Movement of AI
- Most recent developments shared by multiple experts in AI field
- Latest cutting edge techniques
- Build successful Machine Learning Systems (ML Sys)

Data-centric AI *is*
the discipline of systematically engineering
the data used to build an AI system.

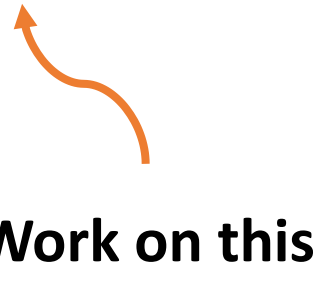
Model-centric approach



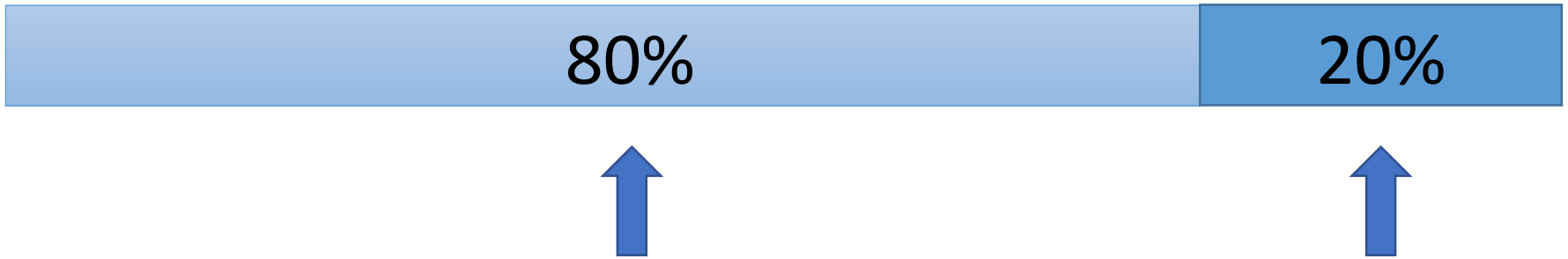
Data-centric approach

AI System = Code + Data

(algorithm/model)



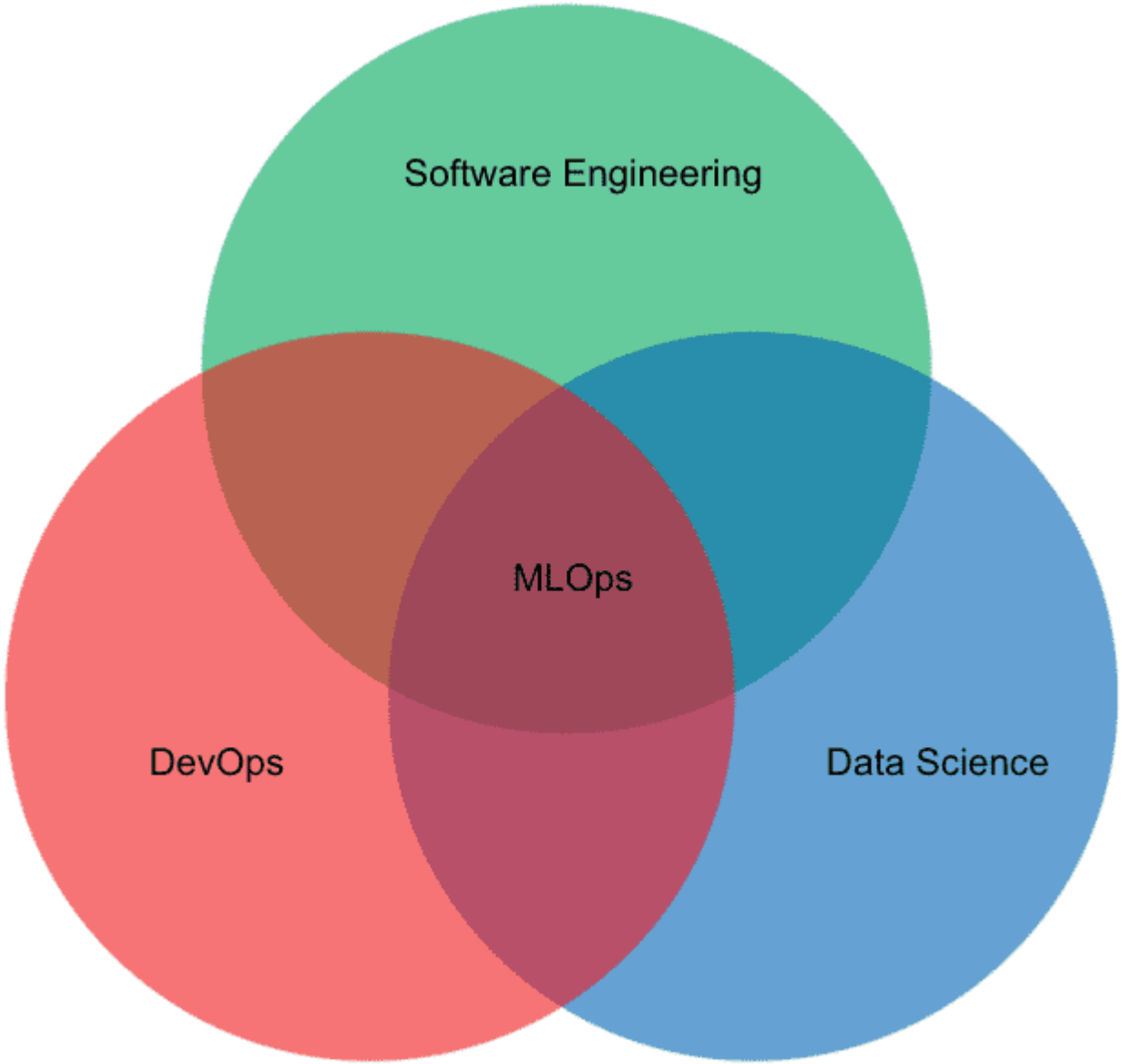
Data is Food of AI



PREP	ACTION
Source and prepare high quality ingredients	Cook a meal
Source and prepare high quality data	**Train a model**

ML + DevOps = 

MLOps



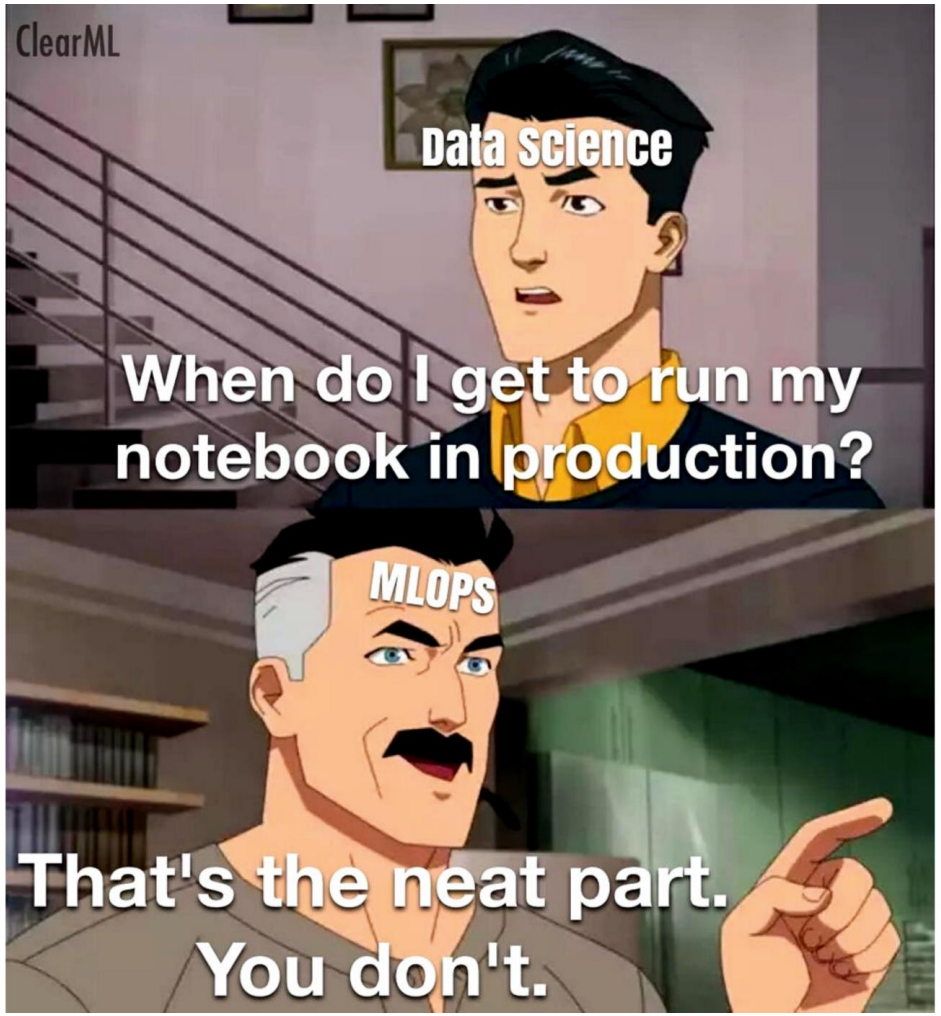
Making it systematic: The rise of **MLOps**

AI Systems	=	Code	+	Data
Creation		Software Engineers		ML Engineers
Quality / Infrastructure		DevOps		MLOps

From Big Data to Good Data

MLOps most important task: ***Ensure consistently high-quality data in all phases of the AI / ML project lifecycle*.**

Machine Learning in Production



ML in production: Expectation

1. Collect data
2. Train model
3. Deploy model

ML in production: Reality

1. Choose a metric to optimize
2. Collect data
3. Train model
4. Realize many labels are wrong -> relabel data
5. Train model
6. Model performs poorly on one class > collect more data for that class
7. Train model
8. Model performs poorly on most recent data -> collect more recent data
9. Train model

ML in production: Reality

10. Deploy model
11. Dream about $$$
12. Wake up at 2am to complaints that model biases against one group
    -> revert to older version
13. Get more data, train more, do more testing
14. Deploy model
15. **Pray**
16. Model performs well but revenue decreasing
17. **Cry**
18. Choose a different metric
19. Start over

Unexpected problems in production

ML Ecosystem

"All-in-one"

Amazon SageMaker · gradient by Paperspace · FLOYD · DOMINO DATA LAB

Data

- Versioning: DVC, Pachyderm, LIQUIDATA
- Labeling: figure eight, scale, Aquarium
- Processing: Spark, DAGSTER
- Exploration: pandas, RAPIDS, dbt
- Data Lake / Warehouse: snowflake, databricks
- Sources: S3, LOG, Parquet

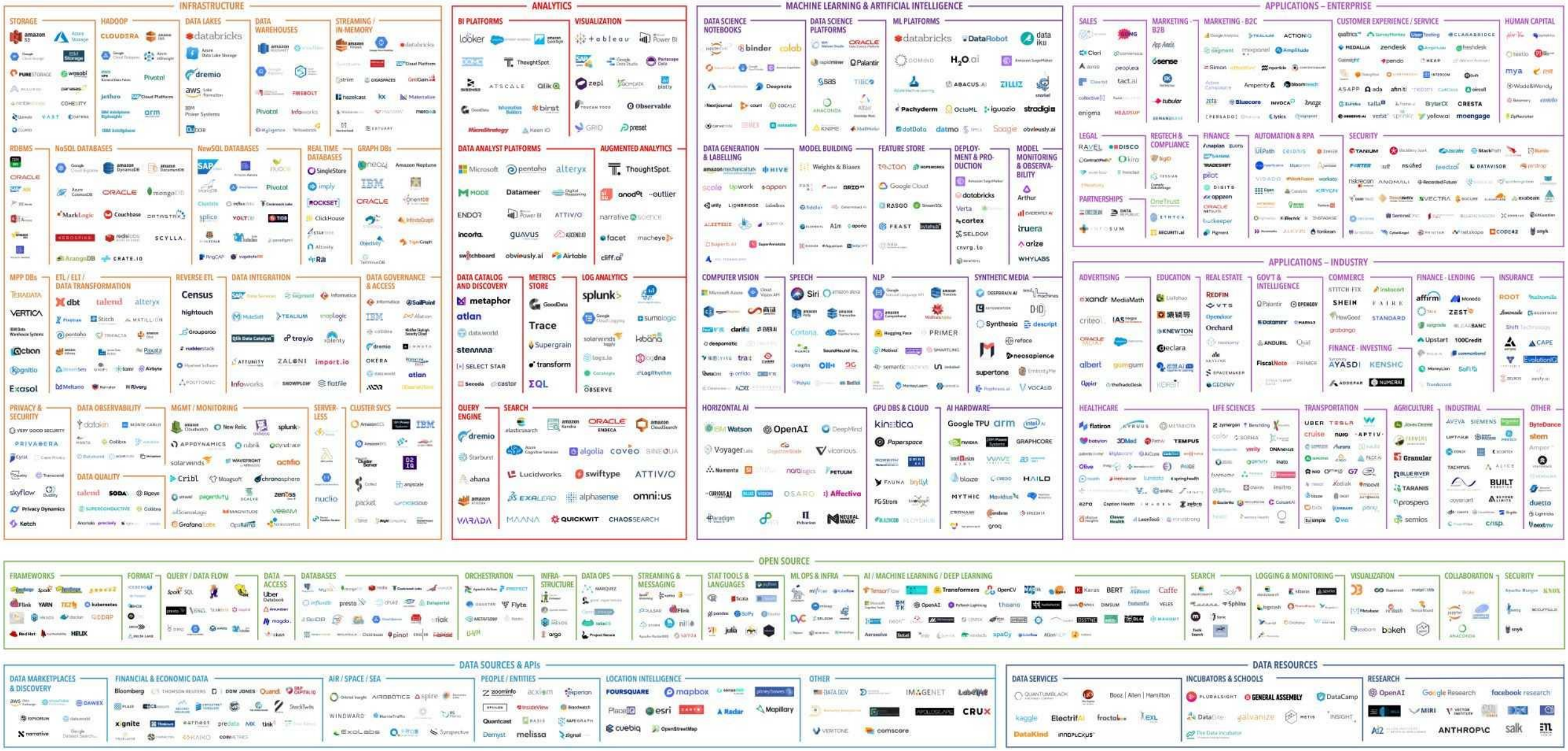
Training/Evaluation

- Frameworks & Distributed Training: fast.ai, PYTORCH, RAY, K, HOROVOD
- Hyperparameter Tuning: SIGOPT, Determined AI, Weights & Biases, tune
- Experiment Management: Weights & Biases, comet, TensorBoard, Neptune, mlflow
- Resource Management: Determined AI, PY, slurm
- Software Engineering: jupyter, Streamlit, git
- Compute: CoreWeave, Lambda

Deployment

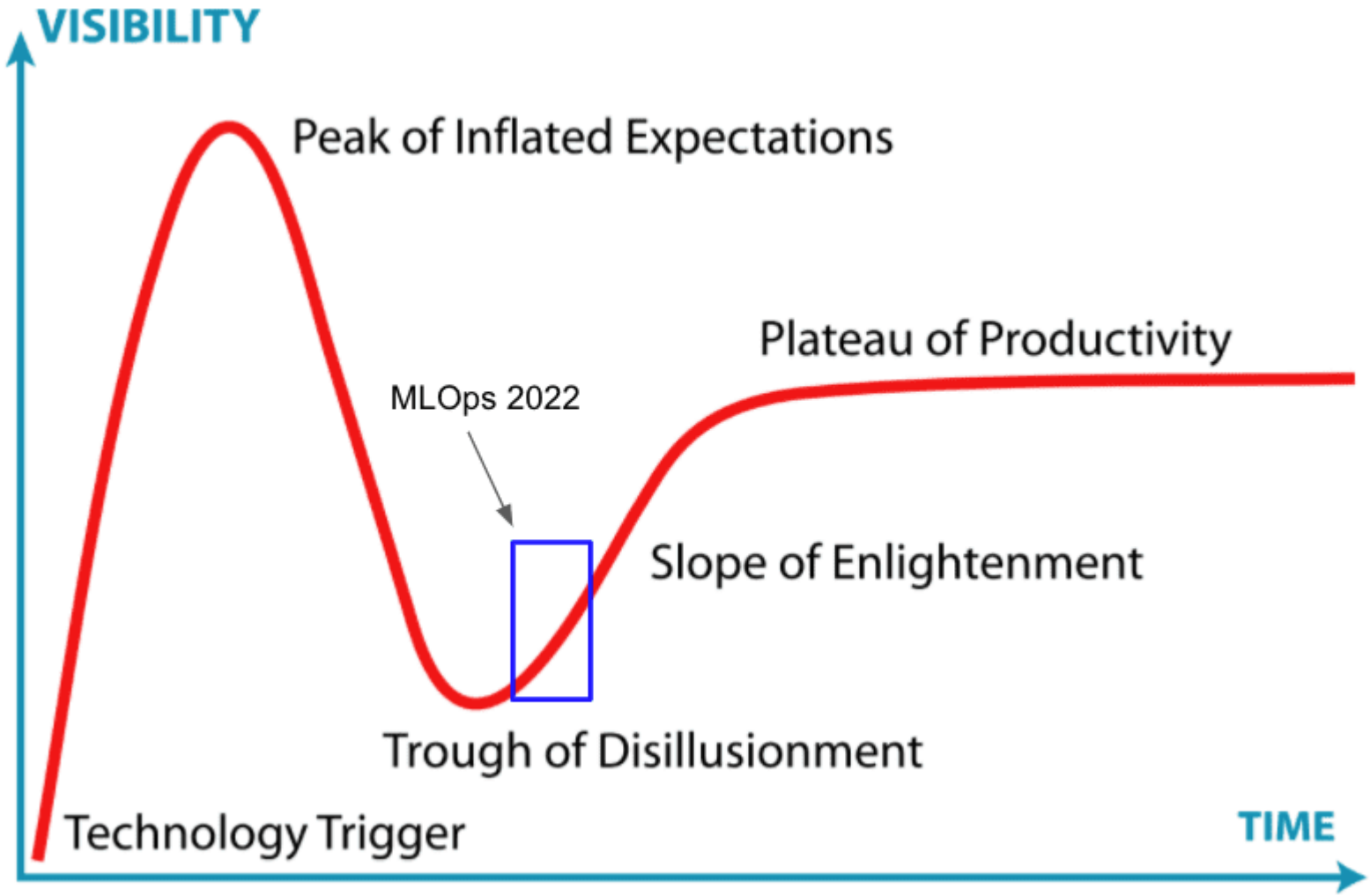
- Feature Store: tecton
- Monitoring: fiddler
- Edge: ML, NVIDIA TensorRT, TensorFlow Lite, ONNX
- Web: ALGORITHMIA, SELDON
- CI / Testing: Buildkite, great_expectations
- or · or

MACHINE LEARNING, ARTIFICIAL INTELLIGENCE, AND DATA (MAD) LANDSCAPE 2021

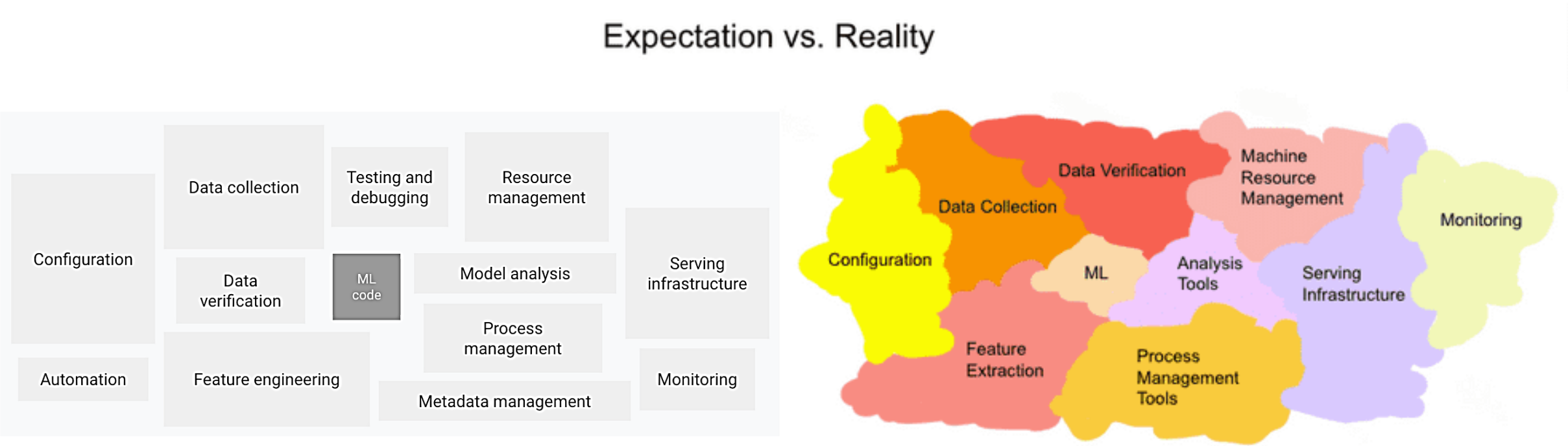


FIRSTMARK
EARLY STAGE VENTURE CAPITAL

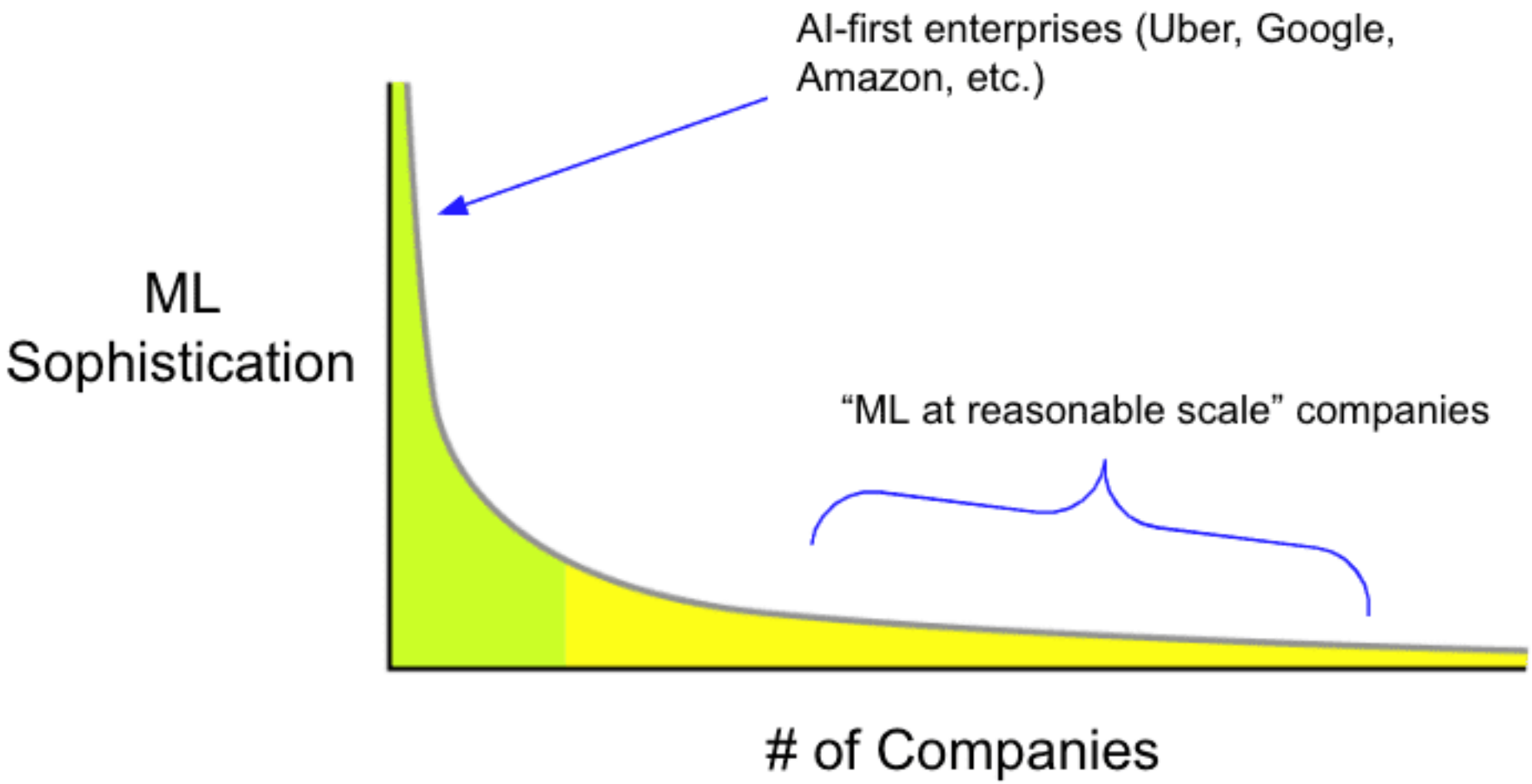
MLOps: Gartner hype cycle paradigm

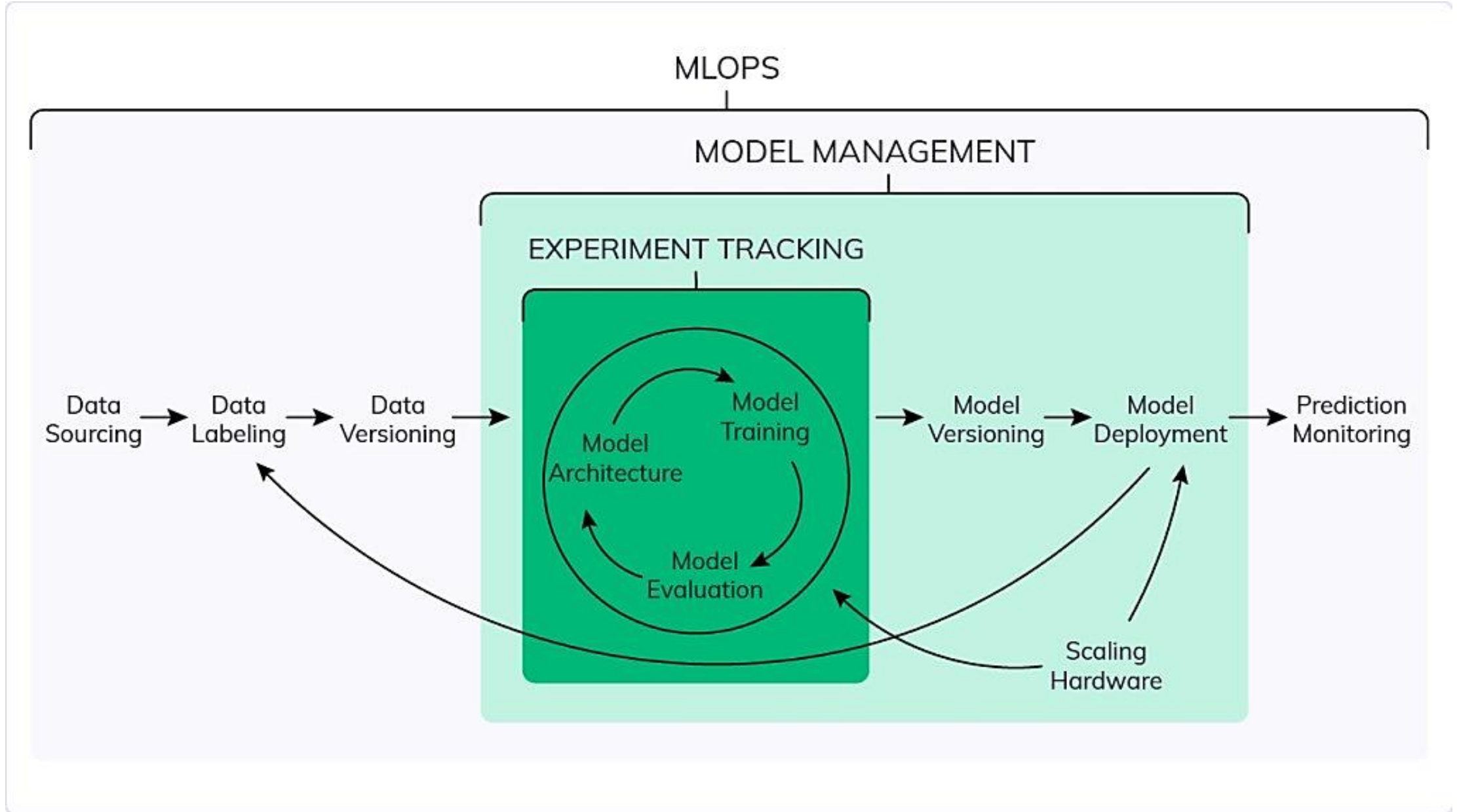


MLOps pipeline: MLOps Amoeba

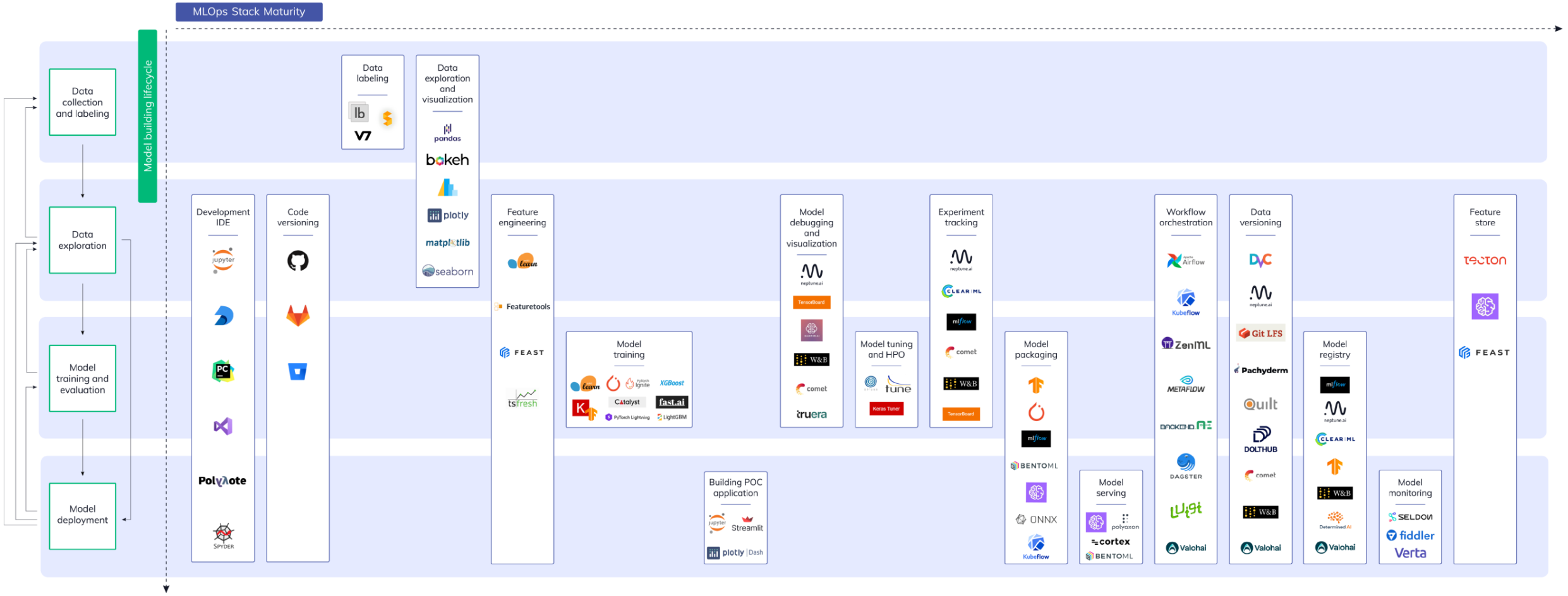


You can start with ML at Reasonable Scale




MLOPS
MODEL MANAGEMENT
EXPERIMENT TRACKING
Data Sourcing
Data Labeling
Data Versioning
Model Architecture
Model Training
Model Evaluation
Model Versioning
Model Deployment
Prediction Monitoring
Scaling Hardware

MLOps tool stack



AI + X: Don’t switch careers, Add AI

Companies need professionals who have domain expertise, augmented with AI skills.

AI + X: Don’t switch careers, Add AI

Many professionals and students want to jump into the world of AI.

But rather than abandoning your current career track to become a ***Data Scientist*** *or a* ***Machine Learning Engineer****, consider developing AI skills to complement your existing subject matter expertise.*