



**Hochschule für Technik  
und Wirtschaft Berlin**

**University of Applied Sciences**

# Project Study (M25) Automatize Machine Learning Processes

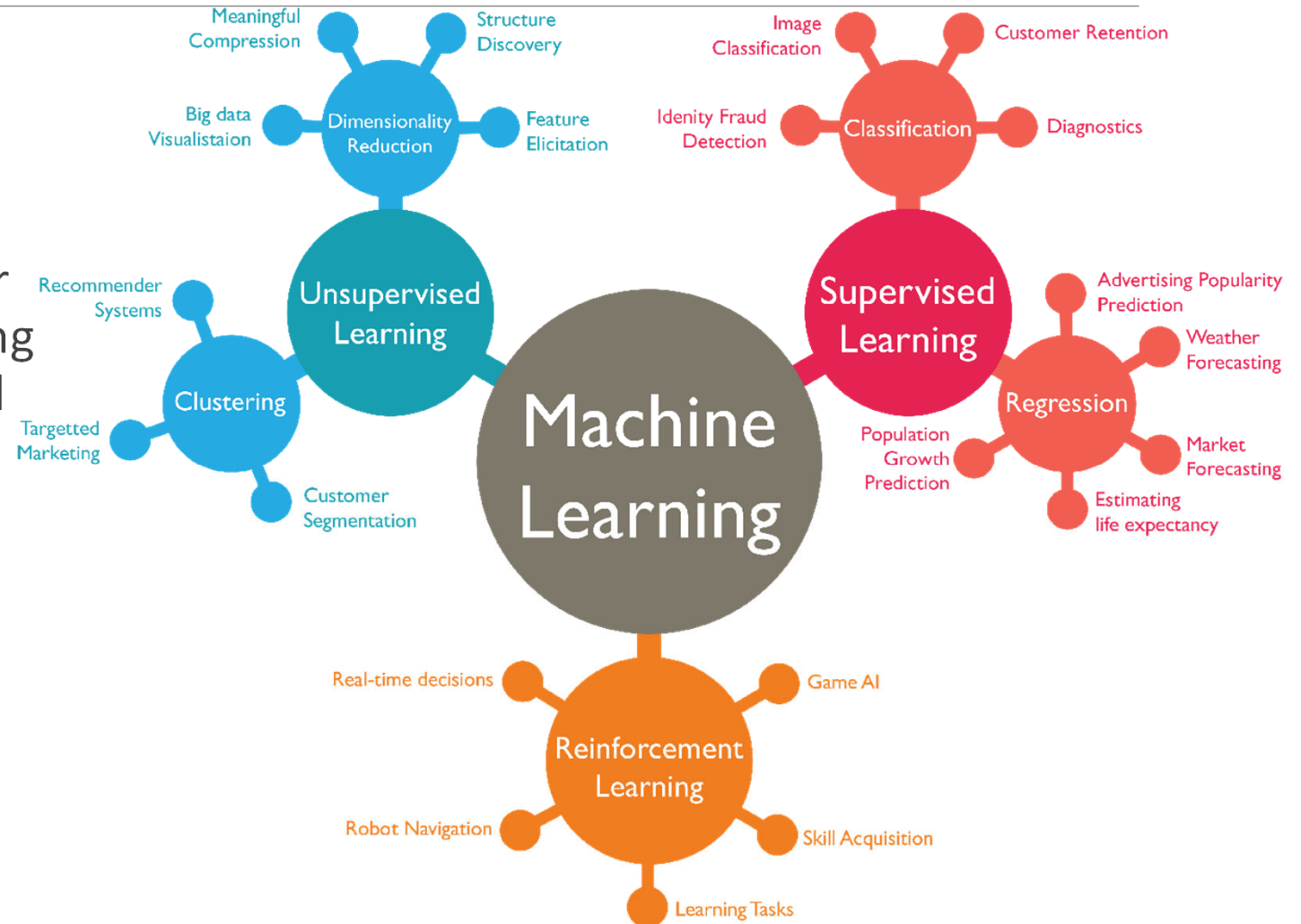
---

PROF. DR. FRANK FUCHS-KITTOWSKI

PAUL SCHULZE, M. SC.

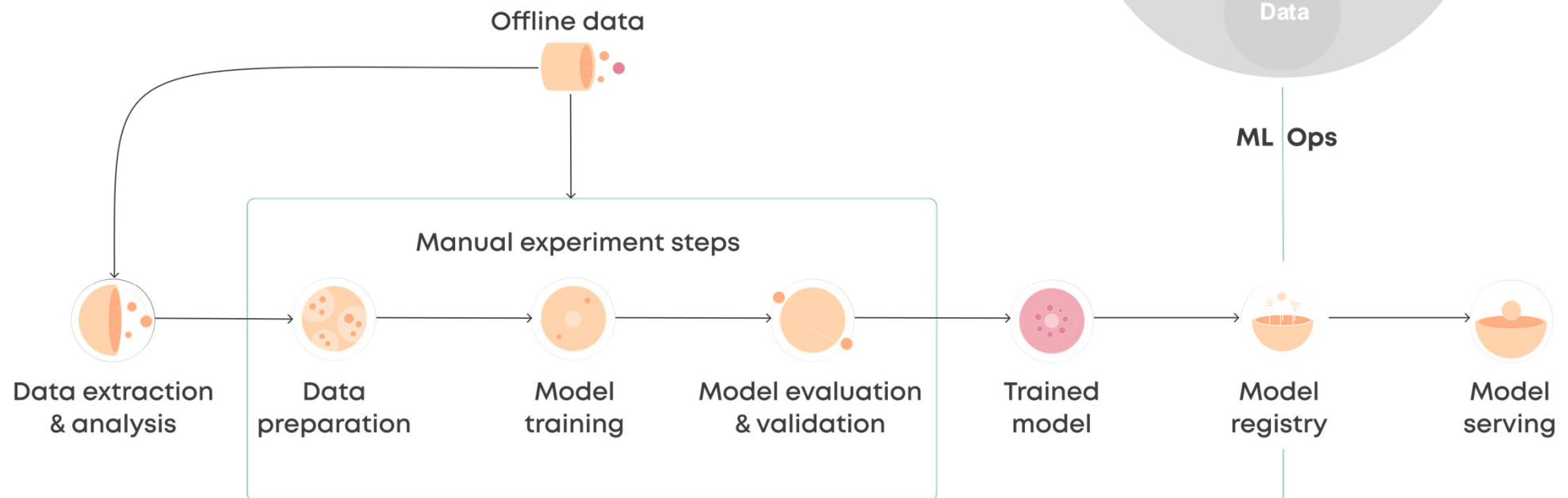
# What is Machine Learning?

- Understand and build models that let machines „learn“
- to make predictions or decisions without being explicitly programmed to do so
- Various use cases
- Various software



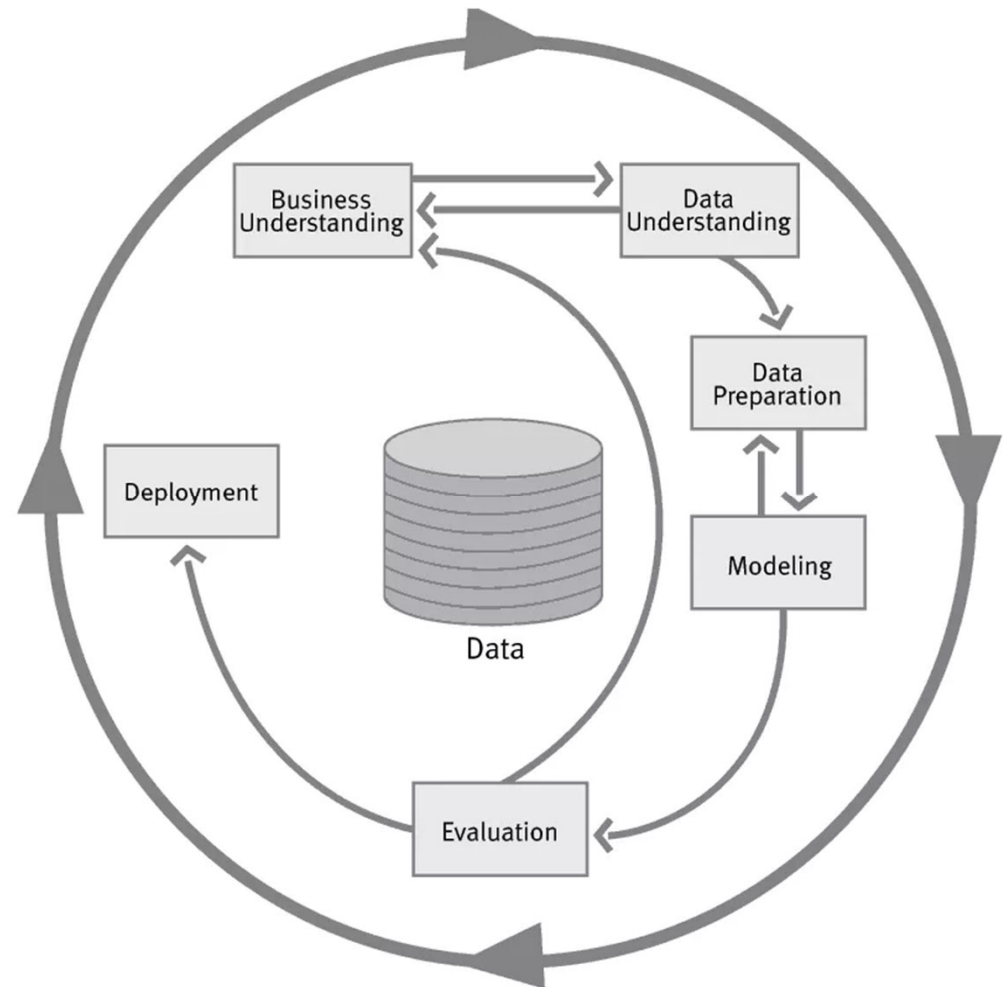
# What is AutoML?

- Use of automated tools and techniques to automate various stages of the machine learning pipeline



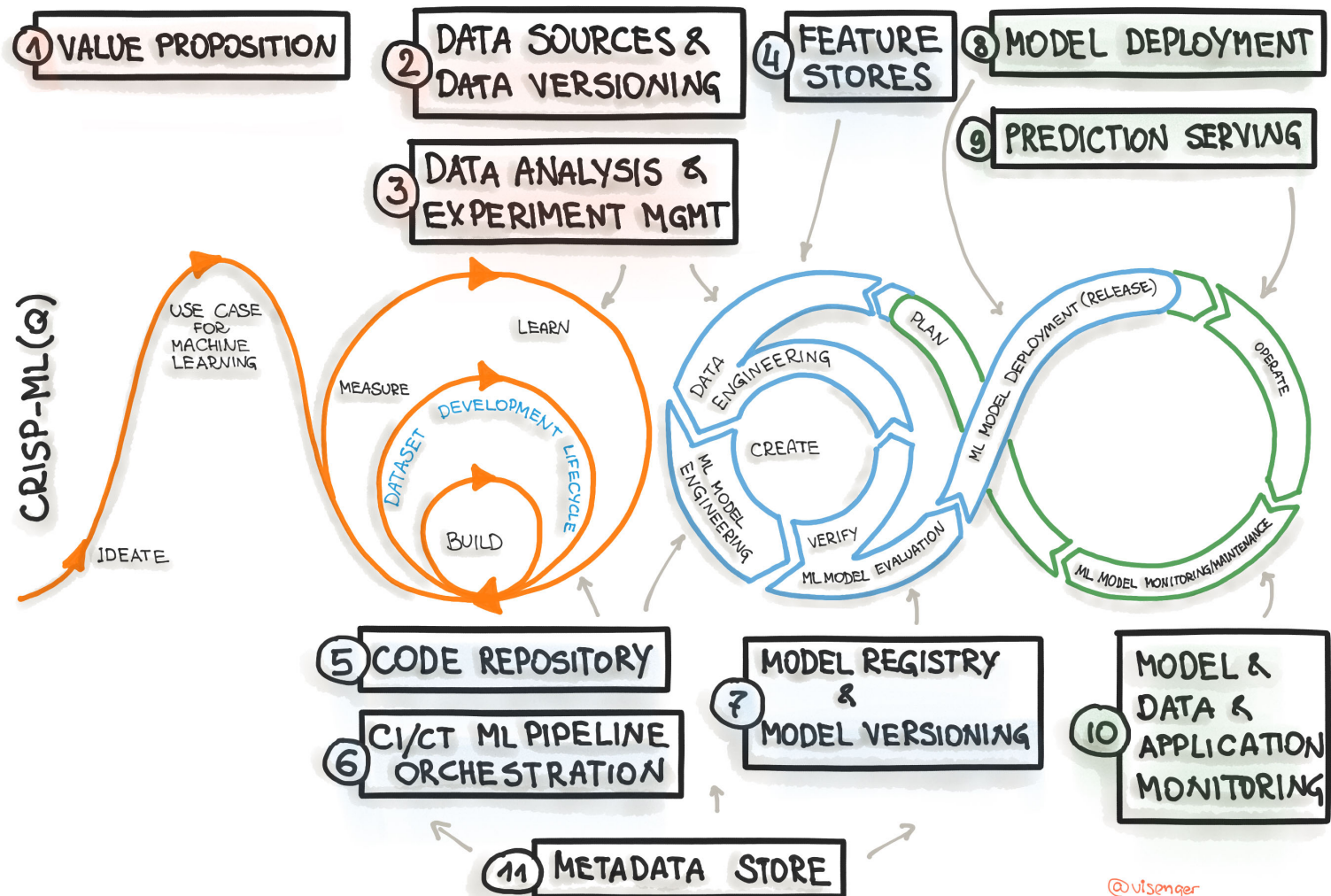
# CRISP-DM

- Establishment of a uniform process and procedure model for data mining projects
- Cross-sectoral use in different industries
- Guidance and blueprint for data mining in 6 steps



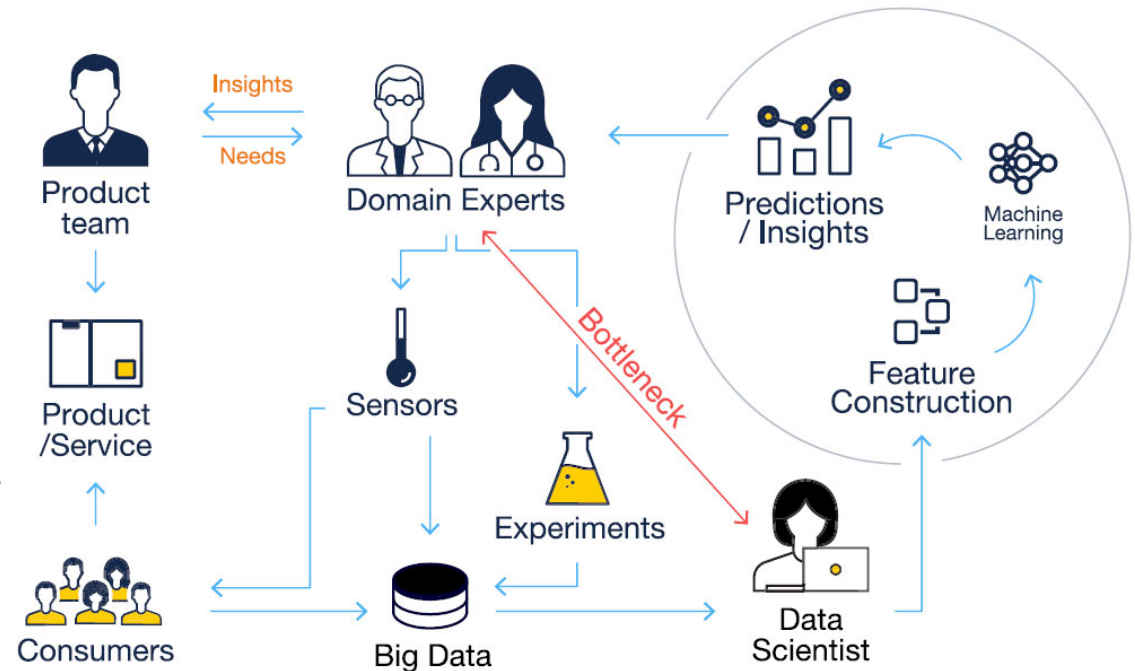
# CRISP-ML

## MLOPS STACK



# What are the challenges to use ML?

- Data quality and quantity
- Model interpretability and explainability (models are black boxes)
- Generalization and robustness (noisy inputs, strange distribution?)
- Scalability and efficiency (big data, training time)
- Continuous learning and adaptation (new data, changing data)
- **Missing knowledge of Domain Experts or Data Scientist**



# What can we automate (AutoML)

	Systems	What is automated?	Access to ML	Efficiency of data scientist
Level 6	???			
Level 5	ComposeML + Level 4 systems			
Level 4	Darpa D3M, MLbazaar, RapidMiner			
Level 3	ATM, Rafiki, Amazon, AutoML, DataRobot, H2O, AUTO-WEKA			
Level 2	Scikit-Learn, Keras, Tensorflow, WEKA, ORANGE, Pytorch			
Level 1	Basic implementation of Decision Tree, KMeans, SVM etc.			
Level 0	Programming languages like python, Java, C++			

Task Formulation  
 Prediction Engineering  
 Feature Engineering  
 Machine Learning  
 Alternative Models Exploration, Testing and Validation  
 Result Summarizing and Recommendation  
 Automated Machine Learning

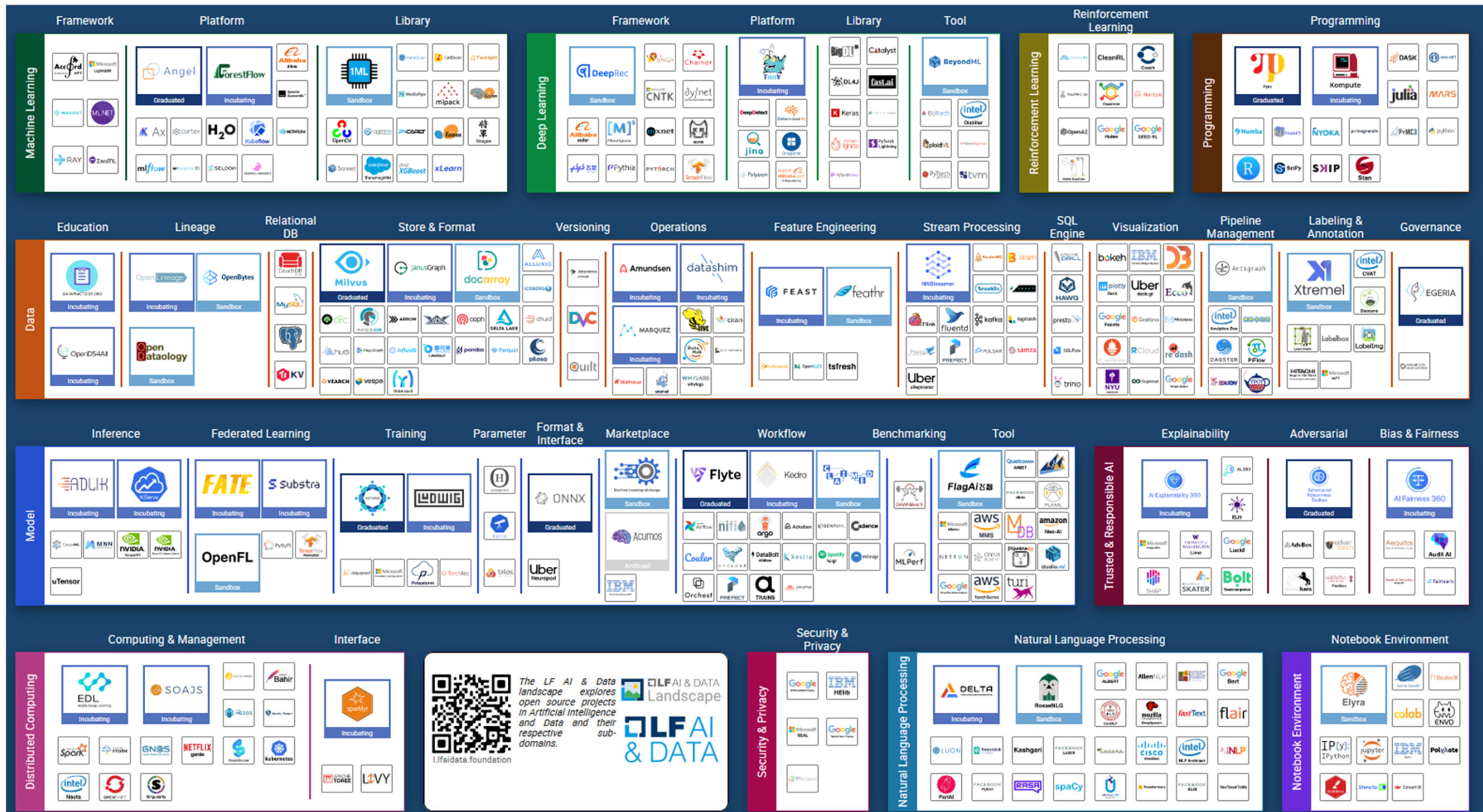
# AutoML Solutions

---

- AutoML Platforms (ready to use end-to-end solutions) like: Google AutoML, Microsoft Azure ML, and DataRobot
- AutoML libraries and framework (libraries offer pre-built functions or classes for automating tasks) like: scikit-learn, H2O.ai, and TPOT
- Integrated AutoML in machine learning frameworks (offer built-in AutoML functionalities) like: TensorFlow, PyTorch, and scikit-learn
- Custom AutoML pipelines (specific machine learning workflows or use cases)
- Cloud-based AutoML services (provide automated machine learning capabilities as a cloud-based service)
- Custom AutoML solutions (DIY)



# AI Landscape



# Possible approach

---

Research-based learning!!!

Potential steps

1. State of the art (literature research)
2. Identification, analysis, evaluation of the problem
3. Technical conception of a solution
4. Technical conception of a solution(Prototypical) implementation
5. Evaluation
6. Report (conference paper)