



# Human-in-the-loop Learning for Safe Exploration through Anomaly Prediction and Intervention

*Prajit Thazhurazhikath Rajendran, Huascar Espinoza, Chokri Mraidha (CEA, DILS-LSEA),  
Agnes Delaborde (LNE)*

**SafeAI, AAAI 2022, 1 March 2022**



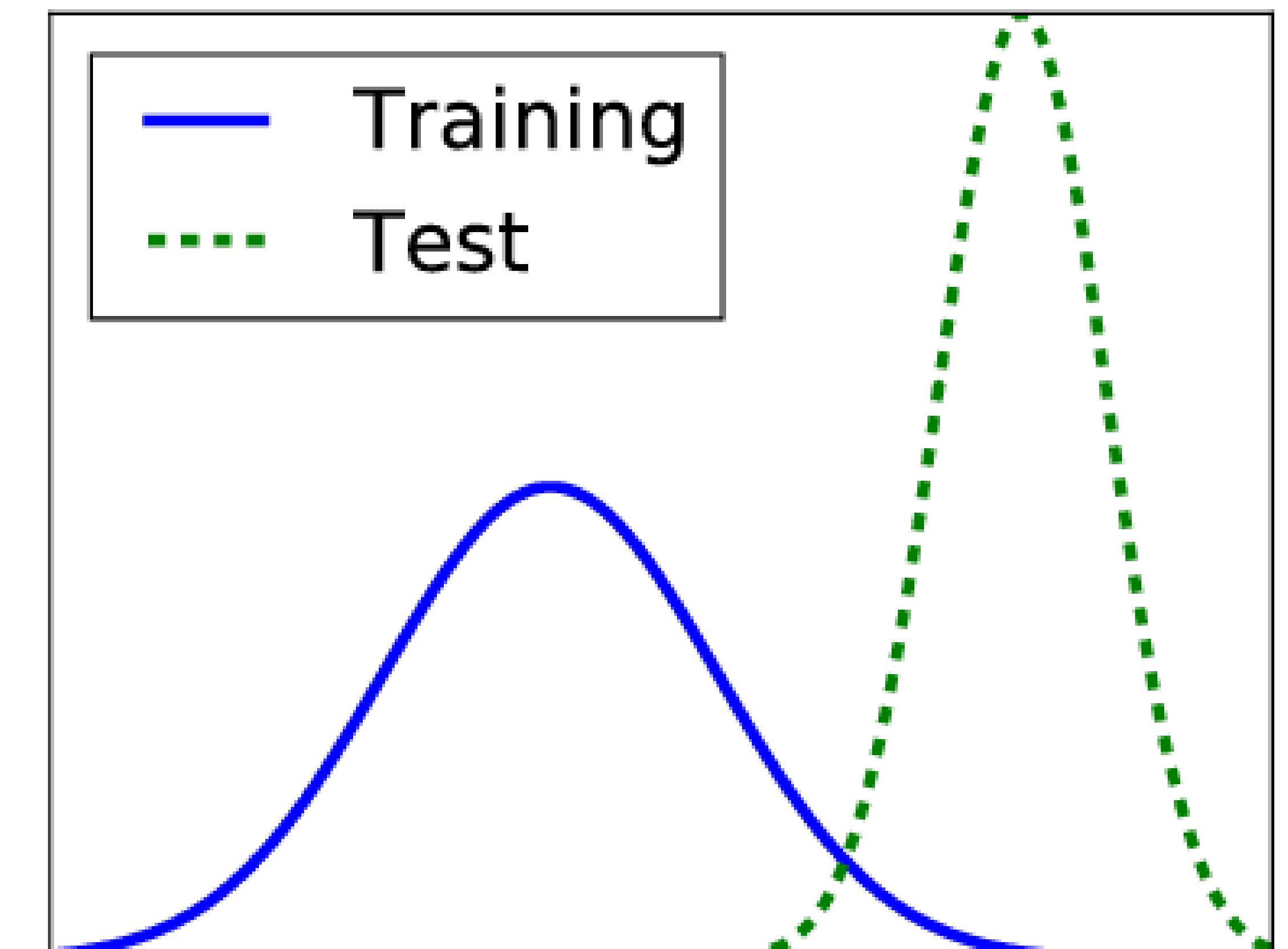
# INTRODUCTION

- **Aim:** Ensure safer exploration in autonomous driving policies, reduce anomalies in data
- **Assumption:** Bad demonstration samples affect safety; Full self-exploration by system is also infeasible
- **Hypothesis:** Human contextual knowledge can help in facilitating safer exploration
- **Methodology:** Unsupervised anomaly detection, human-in-the-loop, reinforcement learning
- **Premise:** Infeasible to start training afresh due to large training time, unsafe exploration

Training images collected in Sunny California



Out of domain images collected in Massachusetts



# PROPOSED METHOD

- Prior historical data or human demonstrations used as a starting point

**Non-exploratory training phase:** Data from the data store is used to train the anomaly predictor and policy learning modules

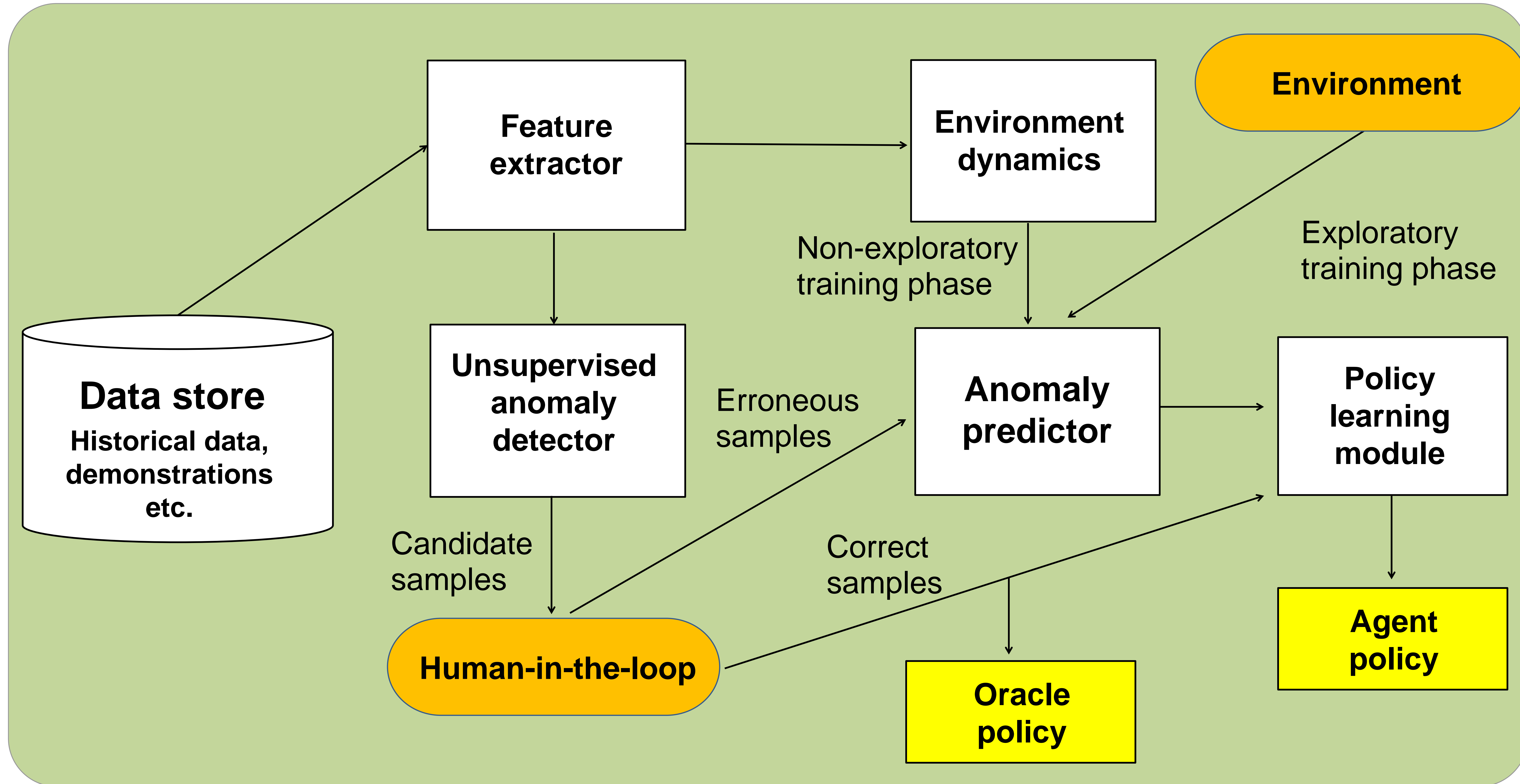
- Can use human-in-the-loop to classify outliers as correct or erroneous
- Correct samples can directly be used for policy training
- Erroneous samples can be used to predict future anomalies/faults by combining with model of environment dynamics

**Exploratory training phase:** System interacts with the environment but chooses actions based on predicted anomaly score

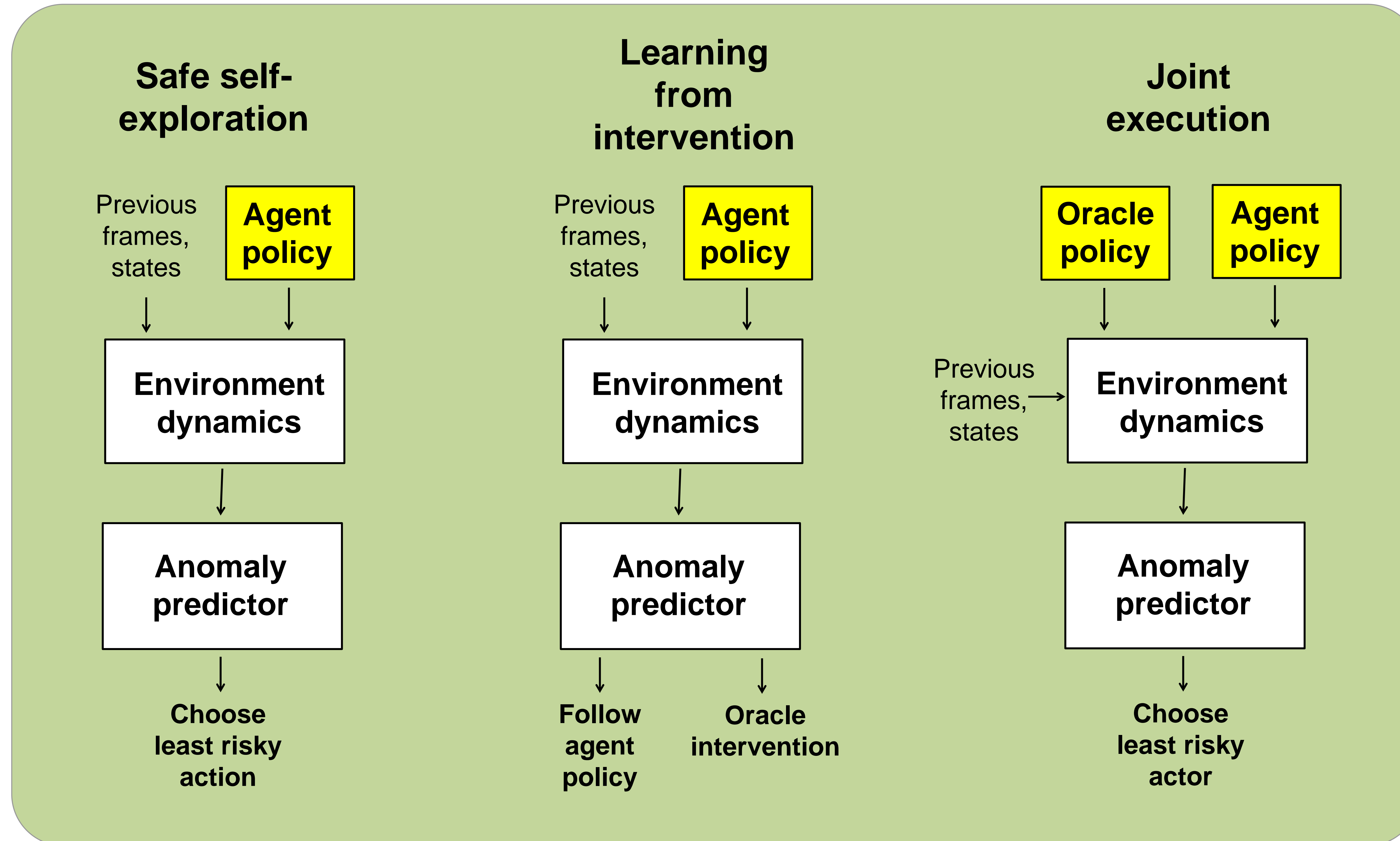
- Facilitates safe exploration by taking previous human feedback into consideration

**Evaluation criteria:** Data quality and quantity, safety, performance, user trust

# BLOCK DIAGRAM



# VARIANTS



***THANK YOU***