

## ML System Implementations

Cloud ML Data Centers Training at Scale

Edge ML Local Processing Inference Focus

Mobile ML Personal Devices User Applications

TinyML Embedded Systems Resource Constrained

Data Pipeline Collection – Processing – Deployment

Resource Management Compute – Memory – Energy – Network

System Architecture Models – Hardware – Software

Core System Principles

Optimization & Efficiency Model – Hardware – Energy

Operational Aspects Deployment – Monitoring – Updates

Trustworthy AI Security – Privacy – Reliability

System Considerations