**Objective:** Identify and develop solutions for technical, policy, security, cultural, and analytic challenges to using enterprise architectures for machine learning.

**Design Criteria**
- **Modular:** Promotes incremental scientific development that manages technical debt in the complex system.
- **Secure:** Provides system-level features for common services to support legal, compliance, security, oversight, and counter intelligence considerations.
- **Robust:** Builds mechanisms for model-driven adaptability through confidence and accuracy related triggers.

**Machine Learning Challenges in the IC**
Ad-hoc approach creates challenges to:
- **Training Data:** Where did the data come from for initial labeling and training? How do we get more to update training?
- **Labeling:** How was the data labeled? By whom? How can others use these labels?
- **Model Validation:** How to ensure continued system integrity and model confidence, not just at initial build but throughout an extended lifecycle?
- **Model Updates:** How to account for both data drift and evolving technical capabilities?
- **Security and Compliance:** What components of data, labels, models need traceability and assurance?

**Labeled Data Registry**
Reliable Storage for Training / Validation Data
- Modular security and data API
- Many labels from many systems – annotation and instrumentation

**Attribute Based Sampling**
Repeatable, compliant and intentional data collection service
- Multi-purpose Selection:
  - Ground truth and validation
- Sampling Framework:
  - Enterprise compliance filter
  - Sampling methods and rates
- **Current focus:** active learning to optimize human effort in labeling

**Labeling Service**
Common, secure, compliant service to annotate data in an analytic workflow
- **Front-End Interface:**
  - Create new labels
  - Assess model results
- **Future Considerations:**
  - Gamification
  - Inter-rater reliability
  - Labeling context
  - Semi-supervised approaches
- **Current focus:** rapid, efficient labeling

**Model Registry**
Versioning and change management service for system components
- Domain relevant reliability and validity assessments
- Model metadata management and transparency
- Enhance trust in model development and sustainment through transparency

**Validation Service**
Service to maximize analytic rigor between model builders and customers
- **Feedback on Model Performance:**
  - Impacts decision making for the lifespan of machine models
  - Enables assessments of the drift of the data’s feature space for retraining
- **Feedback on Compute Use:**
  - Volume over time
  - Compute resources consumed
- **Current focus:** Data pipeline unit tests; hyperparameter optimization.

**Use Cases**
- **Machine Translation**
  - Machine Learning project to develop models for North Korean text translation
  - Used scaffolding to for model assessments and for reliability measures

- **Hyperparameter Optimization**
  - Research partnership with Dr. Tim Menzies and Suvadeep Majumder
  - Developing sampling methods to support active learning
  - Developing hyperparameter optimization routines to support validation

- **News Source Veracity**
  - Machine Learning project to infer international news source veracity
  - Used scaffolding to generate and manage ground truth labeled data