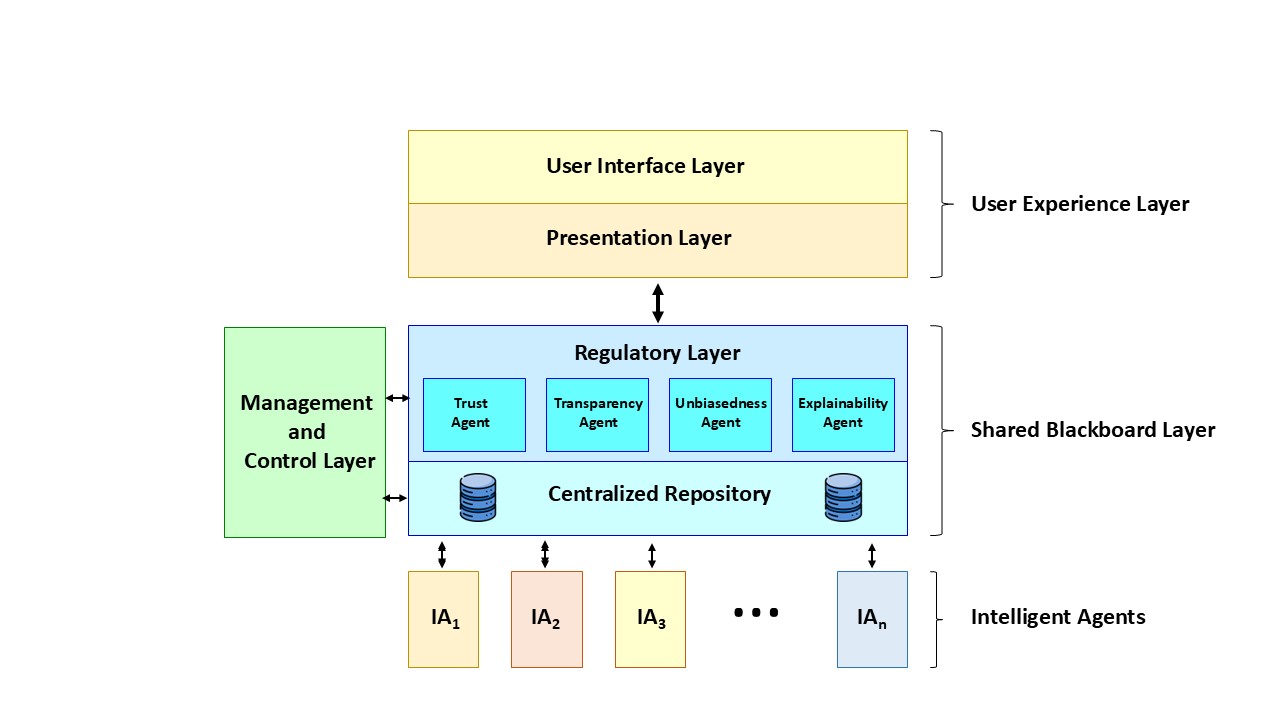
**Blackboard Based Reliable Agentic AI Framework**

The proposed platform for developing reliable and ethical AI applications is based on the blackboard architecture design pattern. The blackboard architecture is a powerful expert system architecture and model of cooperative, distributed problem-solving. It can deal with large amounts ofdiverse, erroneous, and incomplete knowledge sources and expertise to solve problems. It brings together software engineering principles of modularity, AI ideas of pattern-directed invocation, and hardware ideas of shared memory multiprocessors.

The architecture of the proposed platform is shown in Figure 1 below. As can be seen, there are four main components of the platform, namely.

1. User Experience Layer
2. Management and Control Layer
3. Shared Blackboard Layer
4. Intelligent Agents

The various components of the platform are described in the following paragraphs in some detail.



**Figure 1: Blackboard Architecture for Responsible AI Framework**

**1. User Experience Layer**

This layer is responsible for interaction between users and the platform. It consists of a User Interface Layer and a Presentation Layer. The user Interface is responsible to interact with user. It accepts the user inputs and presents the output generated by the platform. The presentation layer is responsible for output/data transformation so that it can be presented in the format as per user preferences. It supports multimodal interaction (text, audio, or video) interaction between a user and the platform.

**2. Management and Control Layer**

The Management and Control Layer handles the interaction among the AI agents and provides serializability. It evaluates the inputted information in terms of its ability to solve the problem at hand. Based on that it prioritizes agents accordingly and makes sure that those pieces of information are used in the right order. This helps in improving the speed of execution and mitigateswastage of unnecessary computation thereby resulting in energy saving and supporting sustainability.

The layer manages the sequence and timing of agent activations. If different agents propose conflicting solutions, it employs predefined strategies (which can be domain-specific or enterprise-specific) to reconcile these conflicts, ensuring coherent decision-making. It monitors the problem-solving process continuously and adjusts its strategies in real time to accommodate changes in the problem context or environment, maintaining flexibility and adaptability.

**3. Shared Blackboard Layer**

The Blackboard layer serves as a centralized repository of the platform where the agents can seamlessly exchange information, facilitating improved collaboration and decision-making.Objects that are placed on the blackboard repository layer could be input data,partial results, hypotheses, or alternatives. Each entry can have an associated certainty factor, this is one way to address uncertainty in knowledge.The repository holds these intermediatory solutions/objects till the time various agents converge on a specific solution to the problem.

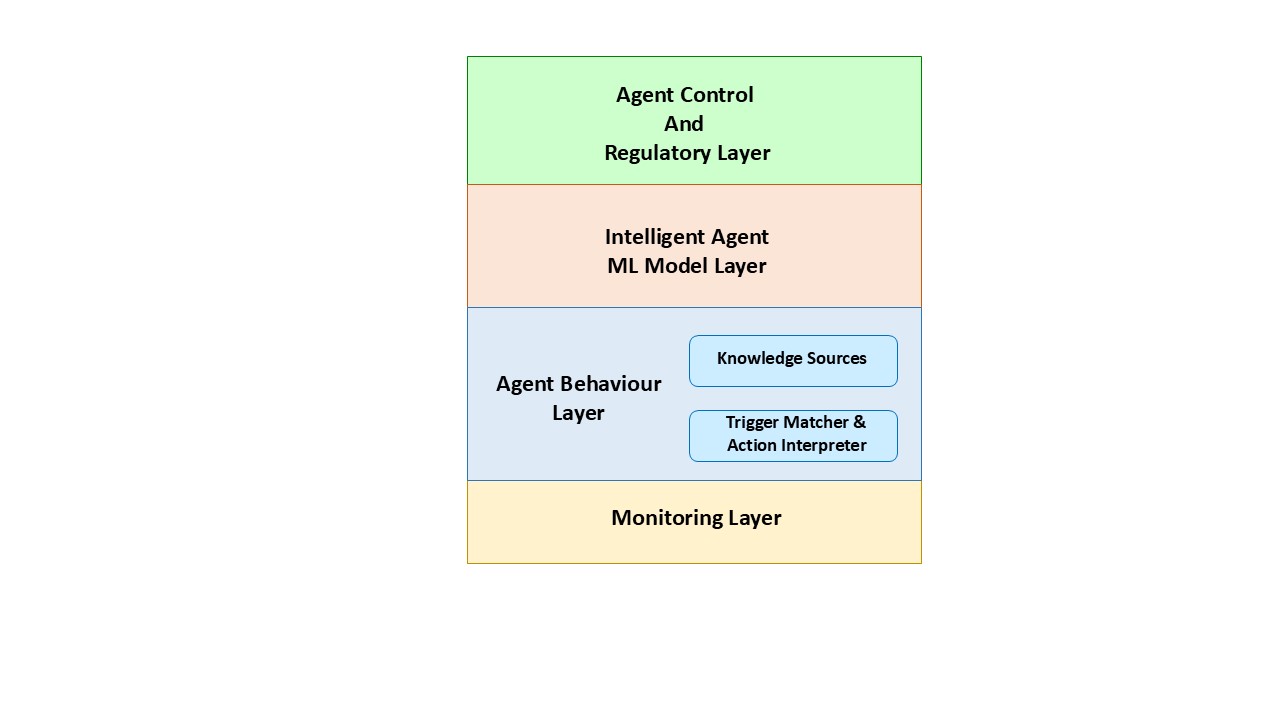
Each agent within the system has specific roles and responsibilities, for example, some can be focused on information retrieval, others on processing queries, and still others on providing personalized assistance through chatbots.

The shared blackboard layer in addition to acting as a central repository also acts as a regulatory layer, where it has agents embedded to ensure trust, transparency, unbiasedness, and explainability. These agents can use the most suitable machine model for performing the required task.

* **Trust Agent:**This agent focuses on establishing trust by implementing mechanisms to verify the reliability and integrity of the AI system. This includes certification processes and auditing mechanisms to ensure that the AI system performs as intended and adheres to ethical standards.
* **Transparency Agent:**This agent is responsible for transparency by making the AI system's decision-making processes understandable and interpretable. This involves providing users with insights into how the AI system works, including the data used for training, the features influencing decisions, and the reasoning behind outcomes.
* **Unbiasedness Agent:**This agent is dedicated to addressing biases within the AI system to ensure fairness and equity in decision-making. This agent employs techniques such as data preprocessing, algorithmic adjustments, and bias detection tools to identify and mitigate biases in the data and algorithms used by the system.
* **Explainability Agent:**This agent focuses on enhancing the explainability of the AI system by providing clear and comprehensible explanations for its decisions and actions. This involves developing methods for generating human-readable explanations, visualizations, and summaries that help users understand the rationale behind the AI system's outputs.

**4. Intelligent Agents**

Intelligent Agents are highly specialized modules with their representation. The typical structure of an Intelligent Agent is shown in Figure 2. As can be seen, each agent is characterized by a set of triggering conditions and an ML Model (an executable code) that retrieves data from the blackboard and contributes to the problem-solving process. The Agent monitors the blackboard continuously for patterns matching their expertise. When an agent identifies a matching hypothesis, it engages in processing and generates new hypotheses.This triggers a chain reaction as new hypotheses are posted in the shared repository layer of the blackboard, potentially invoking other agents. The process continues until no new agent can be invoked, indicating quiescence or a solution.



**Figure 2: Structure of an Intelligent Agent**

These agents are capable of handling multimodalitysuch as text, audio, and video. These agents can be built using any of the machine learning techniques that best meet the purpose of the agent. The agent can be developed using any of the supervised or unsupervised models such as Statistical Methods, Regression, Rule Based, Decision Trees, Random Forest, SVM, KNN, Ensemble, Neural Networks, Clustering, or Convolutional Neural Networks (CNNs) to augment the system's intelligence enabling sophisticated understanding and response capabilities.

We propose to use a wrapper approach where we encapsulate AI models of agents with additional layers of functionality aimed at enhancing their ethical and responsible behavior. Within this framework, each agent responsible for providing features like trust, transparency, unbiasedness, and explainability plays a crucial role in ensuring the overall ethical and responsible functioning of the AI system.