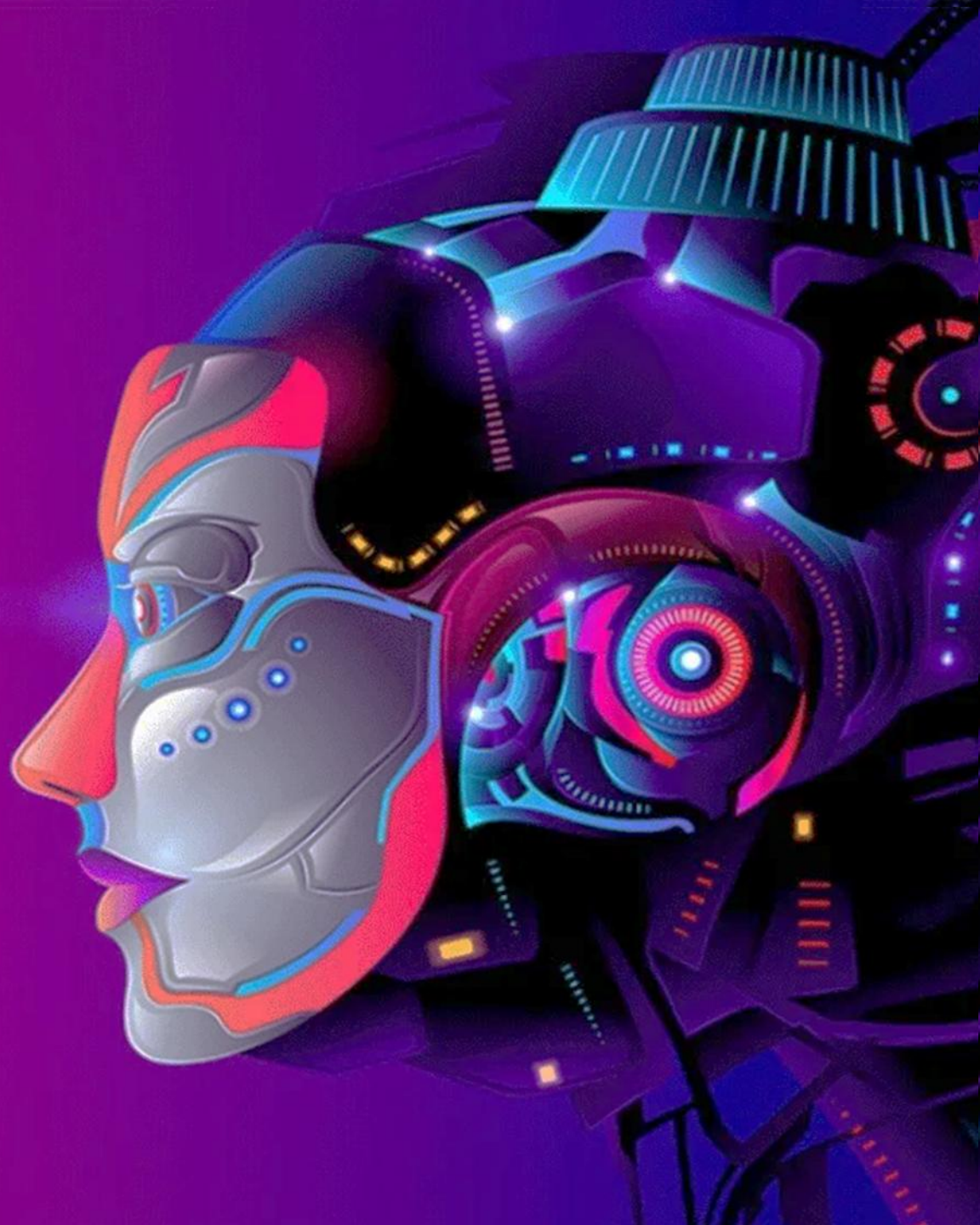
ΔY

Ayita

AGENT-BASE VIRTUAL ASSISTANT

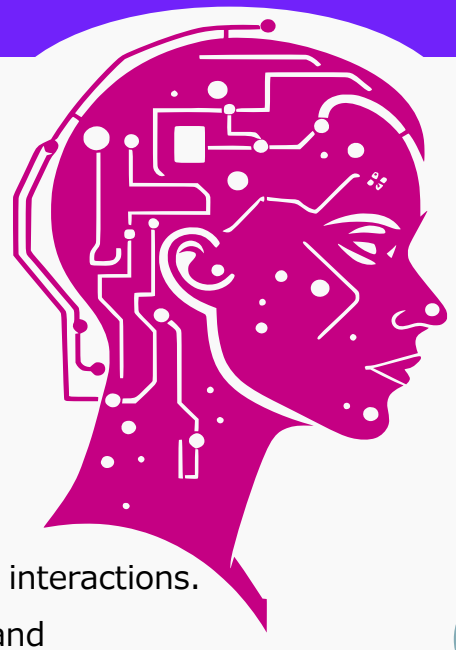


WHAT IS AYITA?

AYITA is an adaptive, privacy-focused AI virtual assistant designed for both personal and professional productivity. Unlike centralized models like REPLIKA, AYITA emphasizes confidentiality and integrates seamlessly with local and external tools, offering proactive support and hyper-personalization..

PRIVACY:

- Processes user data locally, ensuring confidentiality.
- Includes differential privacy measures for tasks like email management and task prioritization.



PERSONALIZATION:

- Fine-tuned to user preferences and context through local LLM models.
- Avatar customization for personalized, visual engagement.

UTILITY:

- **Multimodal communication**: Supports text and voice interactions.
- **Task & Note Management**: Organizes tasks, to-dos, and reminders efficiently.
- **Calendar Integration**: Syncs with schedules for proactive updates.
- **Email Organization**: Sorts, searches, and drafts emails securely.
- **Marketplace for Skills**: Users can add custom features for AYITA's expansion (e.g., industry-specific tasks).

EVENT-BASED APPROACH:

- Detects user needs and suggests actions (e.g., sending reminders, optimizing workflows).
- Integrates with external APIs to provide dynamic responses to events

WHY AYITA?

AYITA addresses the growing challenges of fragmented tools, limited personalization, and data privacy, offering a unified virtual assistant tailored to diverse user needs.

The Problem:

- Fragmented tools create inefficiencies in workflows and task management.
- Limited personalization fails to meet unique user or departmental requirements.
- Data privacy concerns deter users from fully adopting AI solutions

- **Agent-based AI** brings together a **local conversational model** and a large centralized model, complemented by Retrieval-Augmented Generation (RAG) and specialized models. This architecture enables more flexible and context-aware interactions, allowing each agent to leverage both localized insights and global intelligence.
- AYITA enables **seamless integration with existing services** for data management, task coordination, calendars, and enterprise resources, ensuring centralized access to critical information. The platform also features **a developer marketplace**, allowing for the creation of custom solutions and intelligent agents. This approach empowers organizations to tailor the virtual assistant to specific business needs, enhancing automation and system interoperability.

- AYITA features built-in **Fine-Tuning tools** and local model management, enabling precise adaptation to user needs.

All conversations are secured with end-to-end encryption, ensuring full protection against data leaks and unauthorized access. Additionally, the platform supports **workgroup-level configurations**, allowing enterprises to define access controls and role-based settings for seamless collaboration within teams.

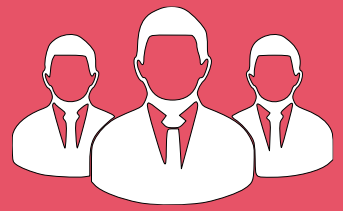
- AYITA offers advanced **LoreBook management**, allowing structured storage and retrieval of contextual memory, including key facts about the user and assistant. Additionally, the platform features an interactive avatar with a unique name and backstory, enhancing engagement through a personalized AI presence that evolves with user interactions..

FROM ENTERPRISES TO INDIVIDUALS

AYITA is a flexible virtual assistant platform that meets the unique needs of enterprises, SMBs, and individuals by combining hyper-personalization, robust privacy, and seamless integration with existing workflows.

FOR CONSUMERS

- **Hyper-personalized** support for creative and professional tasks.
- **Privacy-first design** for handling sensitive client data.
- **Advanced tools for content generation** and time management.



FOR CORPORATE TEAMS

- **Department-specific AI agents** for HR, IT, marketing, and analytics.
- **Secure integration** with internal systems and databases.
- **Automation** of routine processes to boost team efficiency.

FOR DEVELOPERS

- **API access** for building custom integrations and extensions
- **Marketplace opportunities** to create and monetize new skills.
- **Open platform** for contributing to AYITA's growing ecosystem.

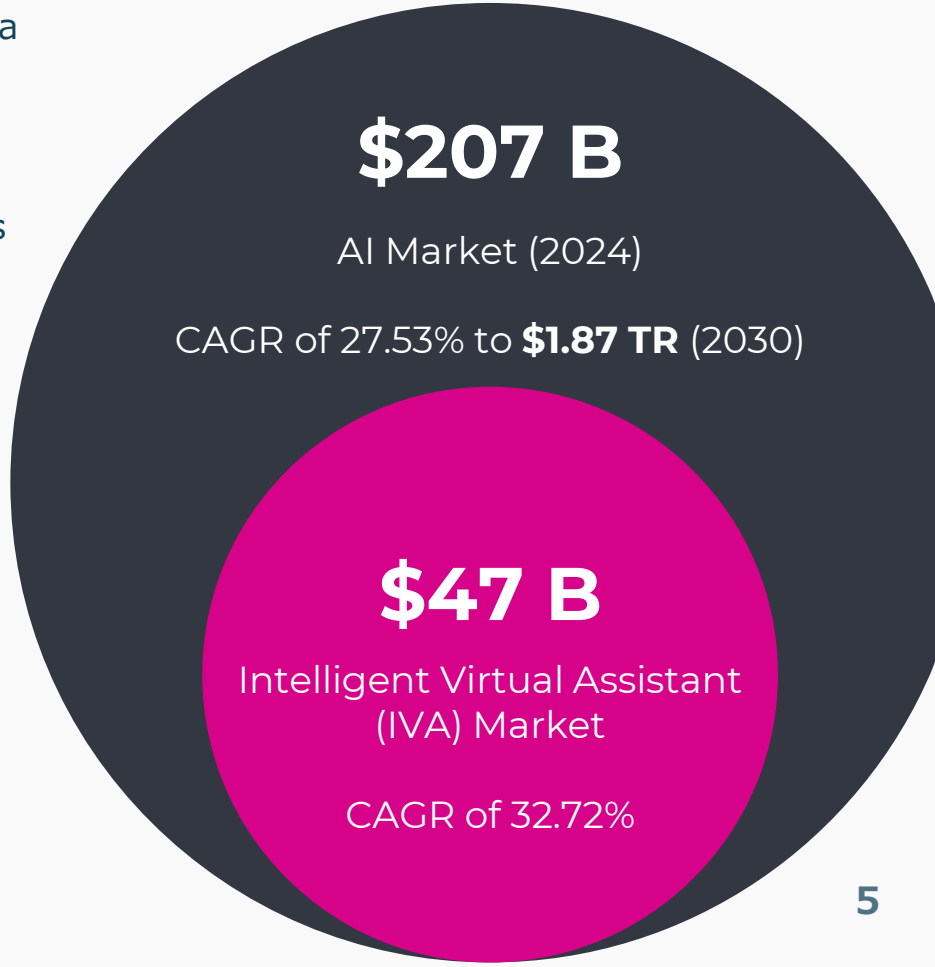
EXPLORING THE LANDSCAPE

Before delving into AYITA's development, it's important to understand the current landscape of virtual assistants and identify opportunities for differentiation. The market lacks a solution that truly bridges personal productivity, professional needs, and privacy-first design. AYITA fills this gap by combining these critical elements into one cohesive product.

The Market Gap. Current virtual assistants excel in specific areas but often fail to deliver:

- **A privacy-first design**: Most rely on cloud-based architectures, raising concerns about data security.
- **Proactive and event-driven utilities**: Existing solutions lack seamless adaptation to user workflows and events.
- **Integration across personal and professional domains**: Assistants rarely address needs spanning both work and life.

Category	Examples	How AYITA Stands Out
Companion AI	REPLIKA, Woebot	Privacy-focused, task-centric design with robust utility.
Task Managers	X.AI, Microsoft Copilot	Greater personalization and proactive suggestions.
Corporate Assistants	Google Assistant, Siri, Alexa	Local LLM processing and secure corporate task handling.
General AI Agents	OpenAI GPT Agents, Hugging Face Apps	Marketplace for customization and dynamic skill-building.



BUILDING AYITA'S CORE

Many of the technologies needed for AYITA, such as local LLMs, fine-tuning, and RAG, are already available. The innovation lies in integrating these components seamlessly into a cohesive pipeline. Haystack serves as the glue, connecting LLMs, RAG, and AI Agents into a modular, scalable system that powers personalized and secure user interactions.

MVP Scope and Features

Feature	Description	Technology
Avatar	Pre-generated, customizable avatars for enhanced engagement.	React UI, Stable Diffusion
Text Chat & UI	Intuitive interface supporting natural language interactions.	React UI, Haystack
RAG (Memory)	Dynamic memory system for long-term context adaptation.	RAG, Haystack, Local DB
Task/Notes APIs	Built-in APIs for task and note management integrated into workflows.	Python, Haystack API
Fine-Tuning	Tailored models for user-specific needs in personal and professional domains.	Local LLMs, Fine-Tuning Pipelines

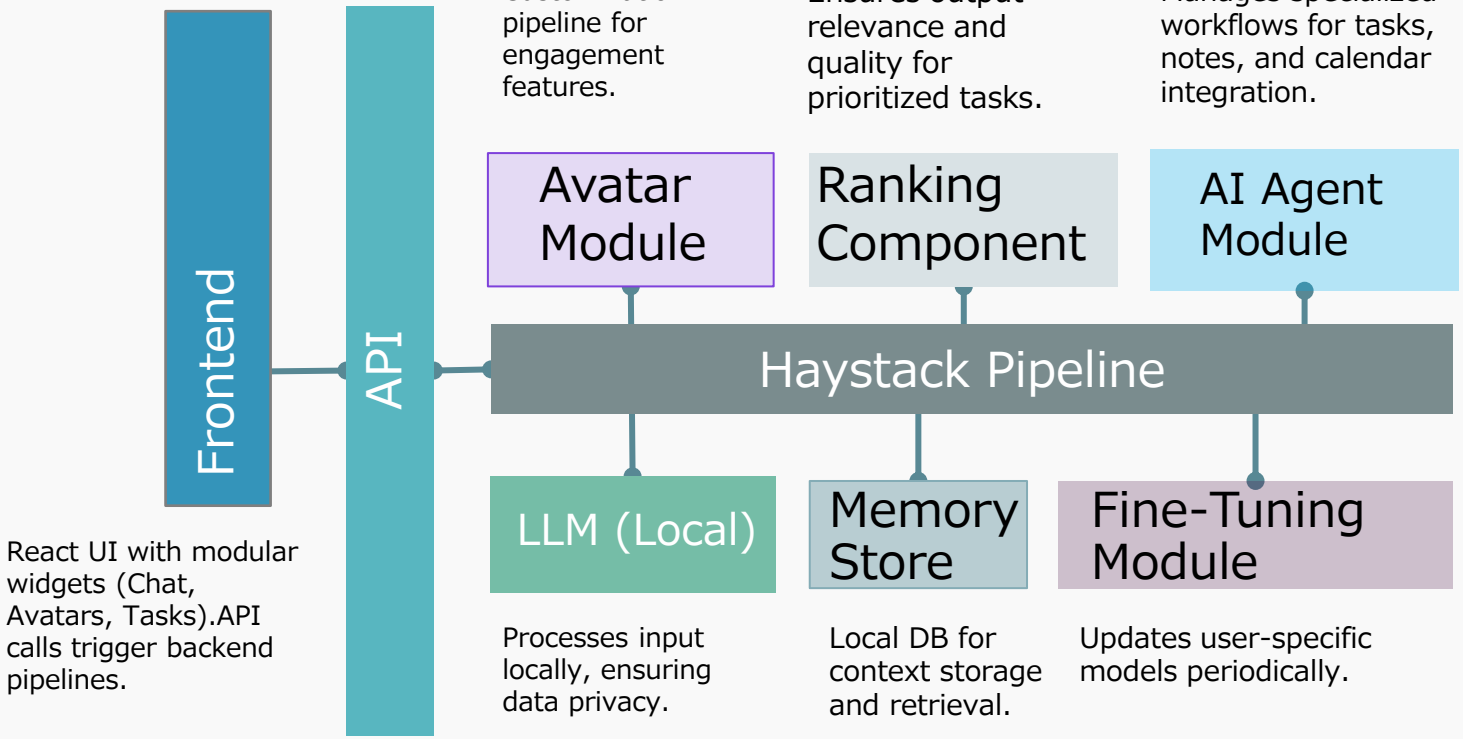
Technological Foundation

- **LLM (Local Language Models)**:Enables personalized, secure interactions with inference performed locally, ensuring data privacy.
- **Fine-Tuning**: Customizes AYITA for specific user needs using curated datasets, enhancing adaptability across personal and professional domains.
- **RAG (Retrieval-Augmented Generation)**:Combines memory and knowledge by integrating local knowledge bases and real-time user inputs, enabling dynamic, context-aware responses.
- **AI Agents**: Orchestrates task-specific functionality like handling emails, calendar scheduling, or retrieving notes through API-driven workflows.
- **Haystack Framework**: Serves as the backbone for integrating LLMs, RAG, and agents into pipelines, providing modular extensibility and smooth user interactions.

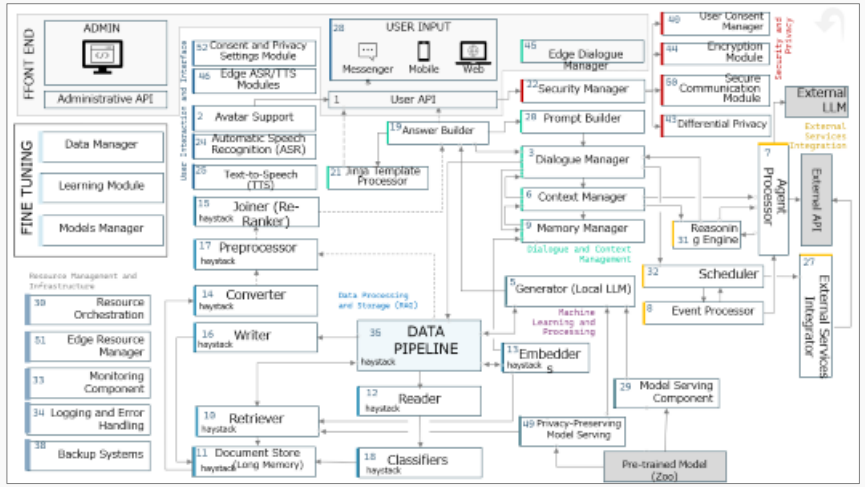
ARCHITECTURE OVERVIEW

AYITA's architecture is built around **Haystack** and **Local LLaMA** 3, enabling seamless integration of memory (via RAG), fine-tuning, and agent-based task management. The architecture prioritizes:

Essential Stack



Full-Scale Architecture



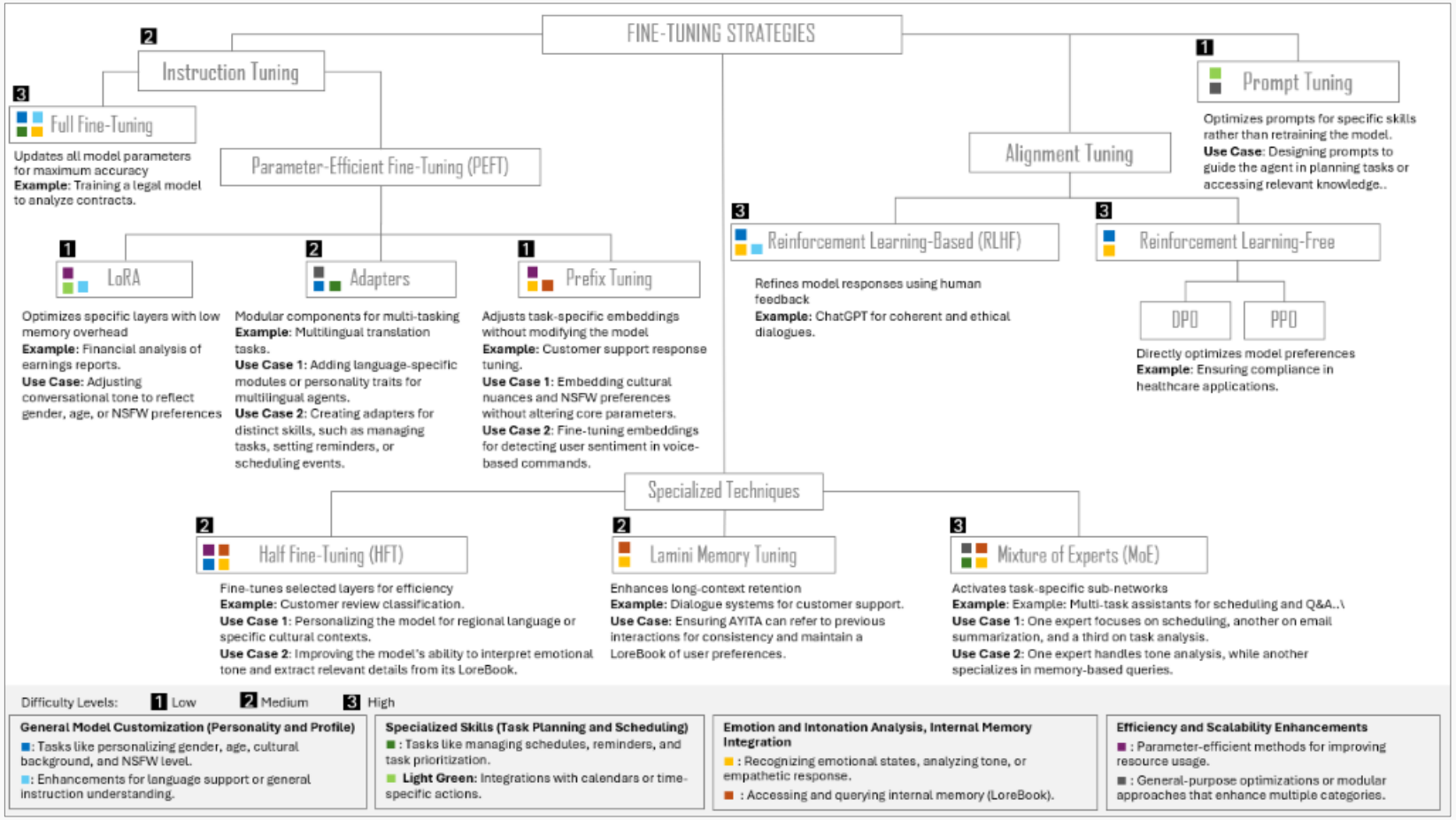
Workflow:

- **User Interaction:** Input through the UI triggers pipelines via Haystack.
- **Processing:** Tasks such as retrieval (RAG), fine-tuning, ranking, and agent actions (e.g., managing tasks or calendar) are orchestrated.
- **Response:** Results are processed by LLaMA 3, prioritized by the ranking component, and displayed in the UI, with context stored in memory.

FINE-TUNING

Fine-tuning is the process of adapting a pre-trained AI model to a specific task or domain by continuing its training on a smaller, task-specific dataset. This technique enhances model accuracy, relevance, and efficiency for specialized applications.

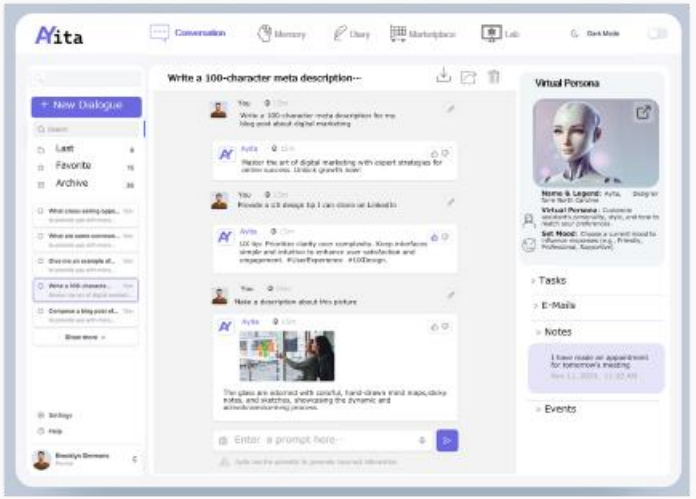
There are different types of fine-tuning, including **full fine-tuning**, which updates all model parameters, **LoRA (Low-Rank Adaptation)** for efficient tuning with minimal changes, and **prefix tuning**, which modifies only input embeddings while keeping the core model intact. Each method balances customization, resource efficiency, and adaptability to different use cases.



UI OVERVIEW

The AYITA interface is designed with a user-centered approach, emphasizing personalization, ease of use, and integration of key functionalities. The platform organizes features like conversations, tasks, notes, and virtual persona management into an intuitive layout. It ensures seamless transitions between personal and professional contexts while maintaining a clean and engaging visual experience.

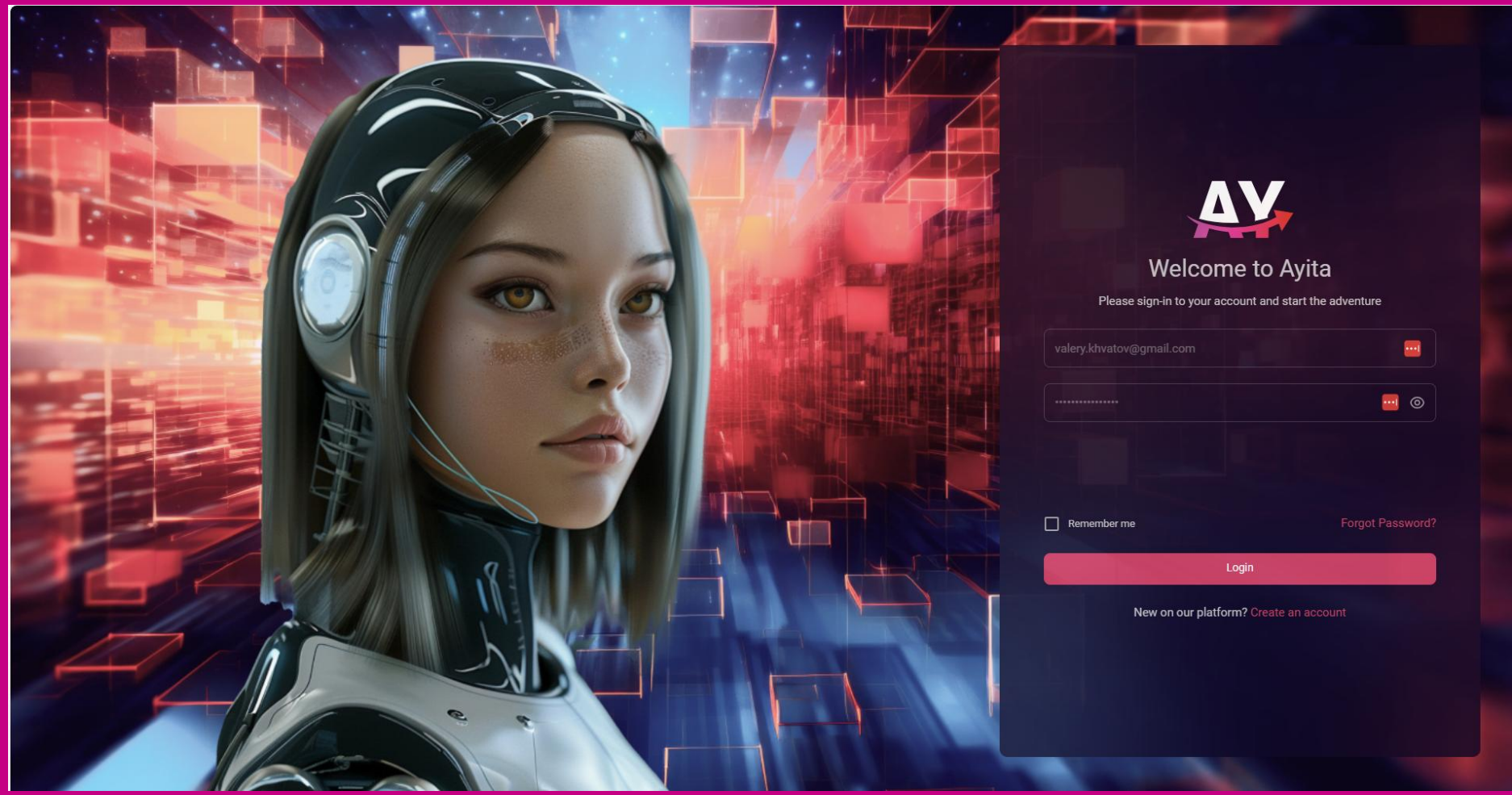
- **Top Navigation Panel**:
  - Features tabs like **Conversation**, **Memory**, **Diary**, **Marketplace**, and Lab (Fine Tuning).
  - Provides quick access to toggle dark mode and user settings.
- **Left Sidebar**: offers navigation for dialogue management, including settings and profile management and options for:
  - New Dialogue creation.
  - Last, Favorite, and Archived conversations.
- **Central Interaction Panel**: Displays the current conversation thread with:
  - User and assistant dialogue in a chat format.
  - Inline tools for editing, downloading, or deleting specific messages.
  - Rich responses, including text, images, and action suggestions.
- **Right Sidebar**:
  - Virtual Persona customization:
    - Displays the avatar, name, and legend of the assistant.
    - Allows users to adjust the personality, mood, and interaction tone.
  - Utility shortcuts for: Tasks, E-Mails, Notes, and Events with quick previews and reminders.
- **Input Area**:
  - A text box for entering prompts, paired with options for attaching files or rich content.
  - Includes a subtle disclaimer about AI-generated information accuracy.



LIVE DEMO

AYITA is now available on the REALM platform as part of Epic 1, showcasing its core capabilities in a Live Show format.

This release introduces conversational AI, local model management, and secure task automation, demonstrating how AYITA enhances productivity with context-aware interactions and privacy-first design. The Live Show provides an interactive experience, highlighting AYITA's adaptability and future expansion within the REALM ecosystem.



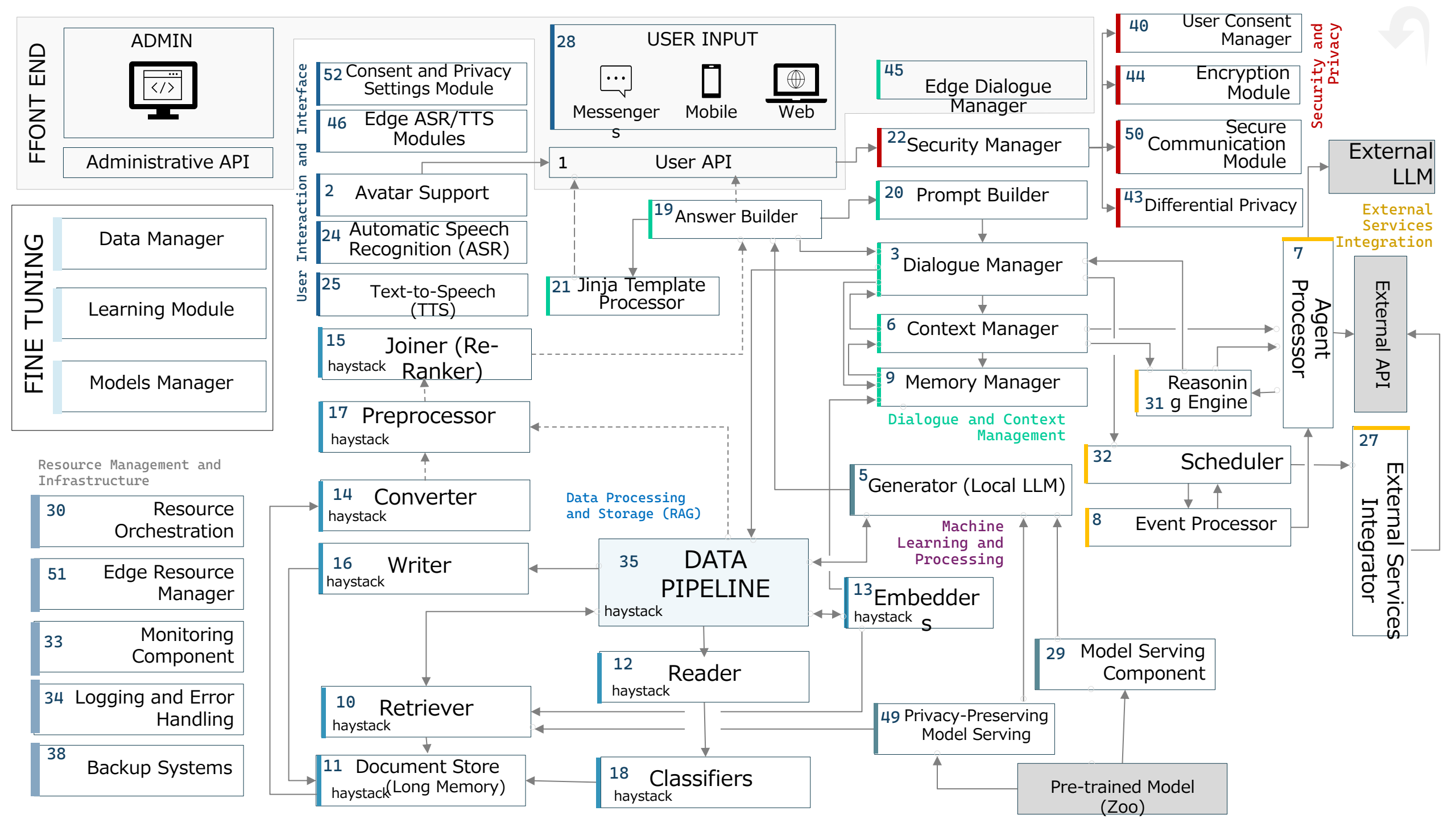
https://ayita.realmdata.io/

CONCLUSIONS

The MVP for AYITA focuses on creating a robust foundation to support privacy-centric, proactive, and adaptive AI interactions. By enabling seamless integration with central and local LLMs, it emphasizes flexibility, privacy, and dynamic task management. The architecture is designed to evolve, supporting multiple LLMs that can adapt to specific user needs and configurations.

Key Focus Areas for MVP	Future Enhancements	Long-Term Objectives
- **Privacy:** Secure and private data processing using local LLMs and Retrieval-Augmented Generation (RAG). - **Proactivity:** Event-driven task management and user interaction. - **Flexibility:** Modular architecture to support multiple LLMs and extendable pipelines. - **Personalization:** Dynamic memory and customizable virtual personas. - **LoreBook (**Memory)	- Integration of advanced **xAI** (Explainable AI) for transparency in recommendations. - Enhanced **multi-modal interactions** (e.g., speech, text, and visuals). - **Advanced Task Agents** with expanded utility (e.g., task ranking, advanced calendaring). - **Marketplace for Skills:** An extensible marketplace to customize AYITA with additional capabilities for specific use cases.	- Establish AYITA as the **leading privacy-first virtual assistant** for personal and professional domains. - Build a scalable ecosystem with **marketplace integrations** for skills and workflows. - Expand into corporate use cases with **secure enterprise-grade solutions**. - Drive **AI innovation** by enabling adaptive frameworks for user-centric task management.

THANK YOU


FFONT END
ADMIN
Administrative API
FINE TUNING
Data Manager
Learning Module
Models Manager
User Interaction and Interface
52 Consent and Privacy Settings Module
46 Edge ASR/TTS Modules
2 Avatar Support
24 Automatic Speech Recognition (ASR)
25 Text-to-Speech (TTS)
28 USER INPUT
Messengers
Mobile
Web
1 User API
19 Answer Builder
21 Jinja Template Processor
15 Joiner (Re-Ranker)
haystack
17 Preprocessor
haystack
14 Converter
haystack
16 Writer
haystack
Data Processing and Storage (RAG)
35 DATA PIPELINE
haystack
12 Reader
haystack
10 Retriever
haystack
11 Document Store
haystack
(Long Memory)
18 Classifiers
haystack
45 Edge Dialogue Manager
22 Security Manager
20 Prompt Builder
3 Dialogue Manager
6 Context Manager
9 Memory Manager
Dialogue and Context Management
5 Generator (Local LLM)
Machine Learning and Processing
13 Embedders
haystack
49 Privacy-Preserving Model Serving
Pre-trained Model (Zoo)
29 Model Serving Component
Security and Privacy
40 User Consent Manager
44 Encryption Module
50 Secure Communication Module
43 Differential Privacy
External LLM
External Services Integration
7 Agent Processor
31 Reasoning Engine
32 Scheduler
8 Event Processor
External API
27 External Services Integrator
Resource Management and Infrastructure
30 Resource Orchestration
51 Edge Resource Manager
33 Monitoring Component
34 Logging and Error Handling
38 Backup Systems

Ayita

Conversation | Memory | Diary | Marketplace | Lab | Dark Mode

+ New Dialogue

Search

Last 8

Favorite 15

Archive 36

What cross-selling oppo... 15m
to provide you with more...

What are some common... 15m
to provide you with more...

Give me an example of... 15m
to provide you with more...

Write a 100-characte... 15m
Master the art of digital marketi...

Compose a blog post of... 15m
to provide you with more...

Show more

Settings

Help

Brooklyn Simmons
Pro trial

Write a 100-character meta description···

You 15m
Write a 100-character meta description for my blog post about digital marketing

Ayita 15m
Master the art of digital marketing with expert strategies for online success. Unlock growth now!

You 15m
Provide a UX design tip I can share on LinkedIn

Ayita 15m
UX tip: Prioritize clarity over complexity. Keep interfaces simple and intuitive to enhance user satisfaction and engagement. #UserExperience #UXDesign.

You 15m
Make a description about this picture

Ayita 15m

The glass are adorned with colorful, hand-drawn mind maps,sticky notes, and sketches, showcasing the dynamic and activebrainstorming process.

Enter a prompt here···

Ayita has the potential to generate incorrect information

Virtual Persona

Name & Legend: Ayita, Designer form North Caroline

Virtual Persona: Customize assistant's personality, style, and tone to match your preferences.

Set Mood: Choose a current mood to influence responses (e.g., Friendly, Professional, Supportive).

> Tasks

> E-Mails

> Notes

I have made an appointment for tomorrow's meeting
Nov 11, 2024, 11:20 AM

> Events

FINE-TUNING STRATEGIES

Instruction Tuning

3 Full Fine-Tuning

Updates all model parameters for maximum accuracy
Example: Training a legal model to analyze contracts.

Parameter-Efficient Fine-Tuning (PEFT)

1 LoRA

Optimizes specific layers with low memory overhead
Example: Financial analysis of earnings reports.
Use Case: Adjusting conversational tone to reflect gender, age, or NSFW preferences

2 Adapters

Modular components for multi-tasking
Example: Multilingual translation tasks.
Use Case 1: Adding language-specific modules or personality traits for multilingual agents.
Use Case 2: Creating adapters for distinct skills, such as managing tasks, setting reminders, or scheduling events.

1 Prefix Tuning

Adjusts task-specific embeddings without modifying the model
Example: Customer support response tuning.
Use Case 1: Embedding cultural nuances and NSFW preferences without altering core parameters.
Use Case 2: Fine-tuning embeddings for detecting user sentiment in voice-based commands.

Alignment Tuning

1 Prompt Tuning

Optimizes prompts for specific skills rather than retraining the model.
Use Case: Designing prompts to guide the agent in planning tasks or accessing relevant knowledge..

3 Reinforcement Learning-Based (RLHF)

Refines model responses using human feedback
Example: ChatGPT for coherent and ethical dialogues.

3 Reinforcement Learning-Free

DPO

PPO

Directly optimizes model preferences
Example: Ensuring compliance in healthcare applications.

Specialized Techniques

2 Half Fine-Tuning (HFT)

Fine-tunes selected layers for efficiency
Example: Customer review classification.
Use Case 1: Personalizing the model for regional language or specific cultural contexts.
Use Case 2: Improving the model's ability to interpret emotional tone and extract relevant details from its LoreBook.

2 Lamini Memory Tuning

Enhances long-context retention
Example: Dialogue systems for customer support.
Use Case: Ensuring AYITA can refer to previous interactions for consistency and maintain a LoreBook of user preferences.

3 Mixture of Experts (MoE)

Activates task-specific sub-networks
Example: Example: Multi-task assistants for scheduling and Q&A..\
Use Case 1: One expert focuses on scheduling, another on email summarization, and a third on task analysis.
Use Case 2: One expert handles tone analysis, while another specializes in memory-based queries.

Difficulty Levels: 1 Low 2 Medium 3 High

General Model Customization (Personality and Profile)
■: Tasks like personalizing gender, age, cultural background, and NSFW level.
■: Enhancements for language support or general instruction understanding.

Specialized Skills (Task Planning and Scheduling)
■ : Tasks like managing schedules, reminders, and task prioritization.
■ **Light Green**: Integrations with calendars or time-specific actions.

Emotion and Intonation Analysis, Internal Memory Integration
■ : Recognizing emotional states, analyzing tone, or empathetic response.
■ : Accessing and querying internal memory (LoreBook).

Efficiency and Scalability Enhancements
■ : Parameter-efficient methods for improving resource usage.
■ : General-purpose optimizations or modular approaches that enhance multiple categories.