



# Apollo: An AI Ecosystem for Sports Analytics

An AI-powered, agent-driven intelligence hub built to turn live sports data into actionable insight.

Andrea Chang, Anvii Mishra, Aviel Rosen, Haithem Salmi, Jason Sun, Kaan Eroltu, Kien Le, Maria Romo Nichols, Matthew Gordon, Mirra Klimov, Radomyr Couture  
with the help of Dr. Anton Dahbura and Tad Berkery  
Sports Analytics Research Group



JOHNS HOPKINS  
WHITING SCHOOL  
of ENGINEERING

## Problem Statement

Sports analysis is increasingly driven by **unstructured data**, including news articles, social media, and scouting reports. This makes it difficult to **efficiently search, verify, and synthesize** in real time. Extracting meaningful insights from **fragmented, text-heavy sources** poses a challenge, especially for **lesser-known players** with depth pieces, recent roster changes, and draft prospects. We address this gap by building a system that rapidly **aggregates and verifies** unstructured information, transforming it into **concise, contextual, and decision-ready insights**, enabling users to quickly understand emerging situations that traditional data pipelines often miss.

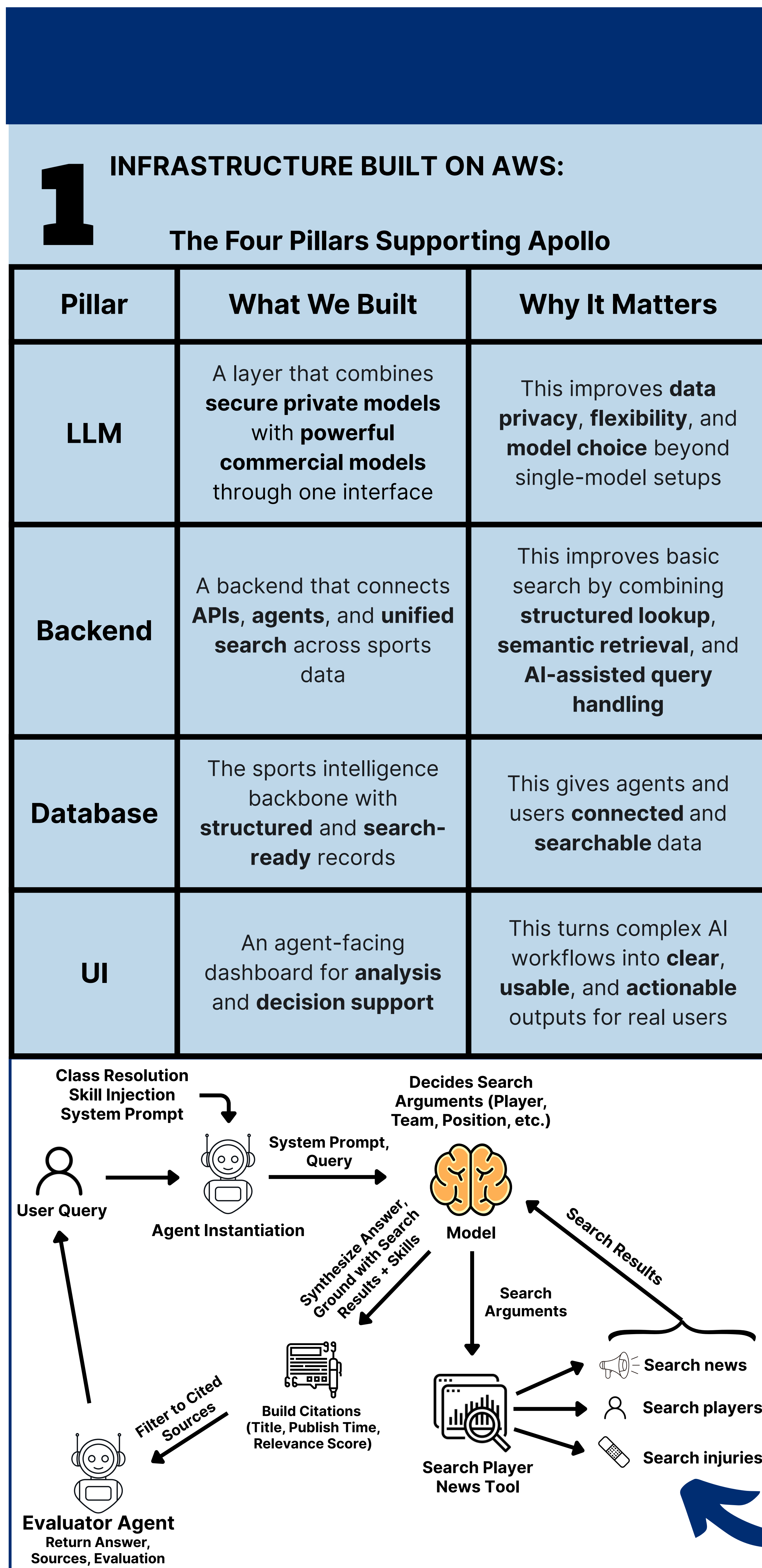
## Our Solution

We developed a platform that...

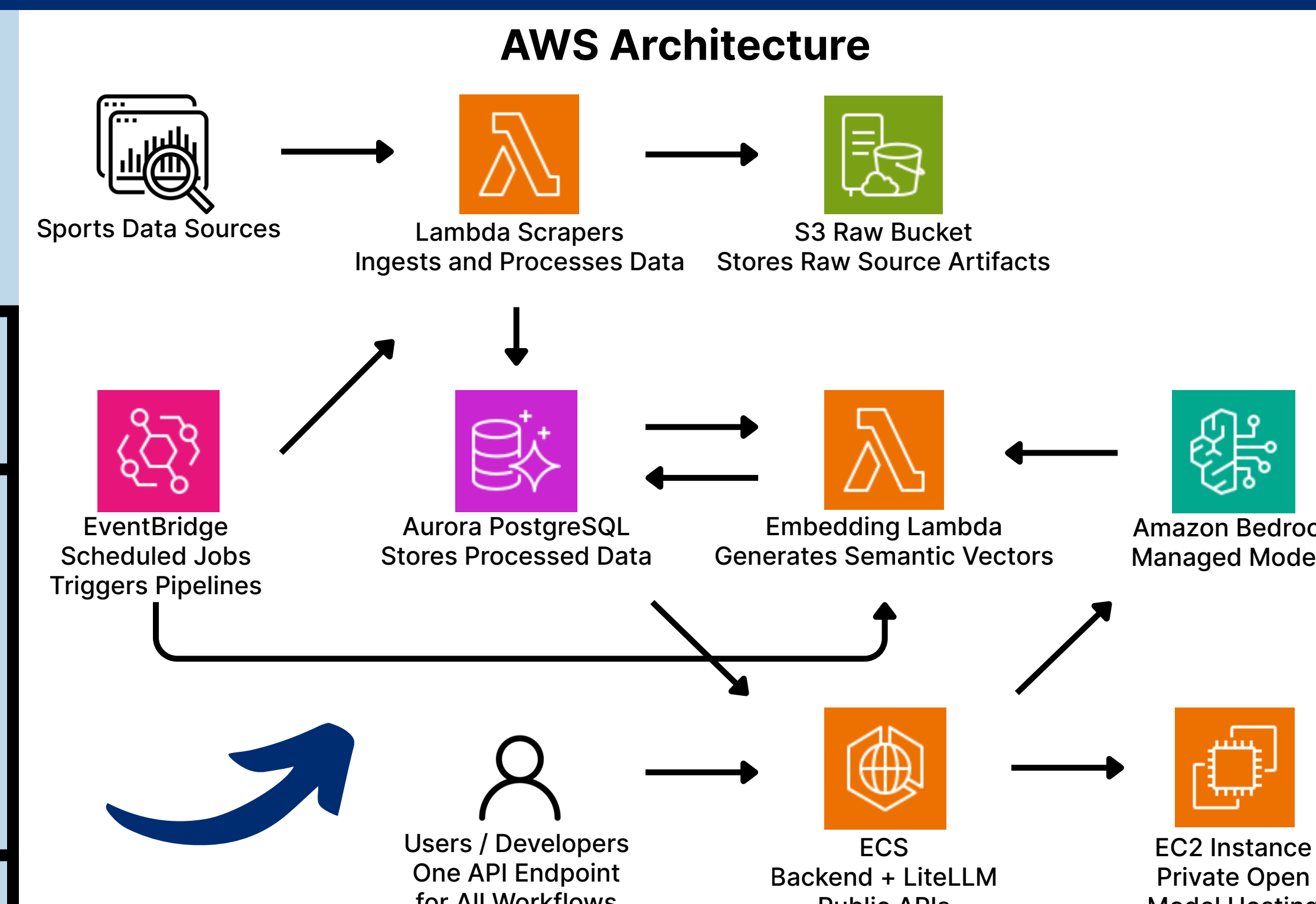
- Offers front offices and **fans and professionals** access to a dynamic set of **agents** with **expert-level** knowledge
- Provides **users** with the ability to **modify agents** to serve their own **specific** sports-related **needs**
- Uses **state-of-the-art methods** and **industry-leading data** to provide insight that would otherwise be **time-consuming** to **generate organically**

## What is an Apollo agent?

- **RAG (Retrieval-Augmented Generation):** Combines information retrieval with language generation to produce accurate, context-aware answers
- **Skills:** Define how the agent analyzes queries and formats responses
- **Tools:** Fetch data, run queries, or perform computations.
- **Database:** Aggregates news from multiple sources, then filters, organizes, and indexes it so the agent can quickly retrieve the most relevant and reliable information



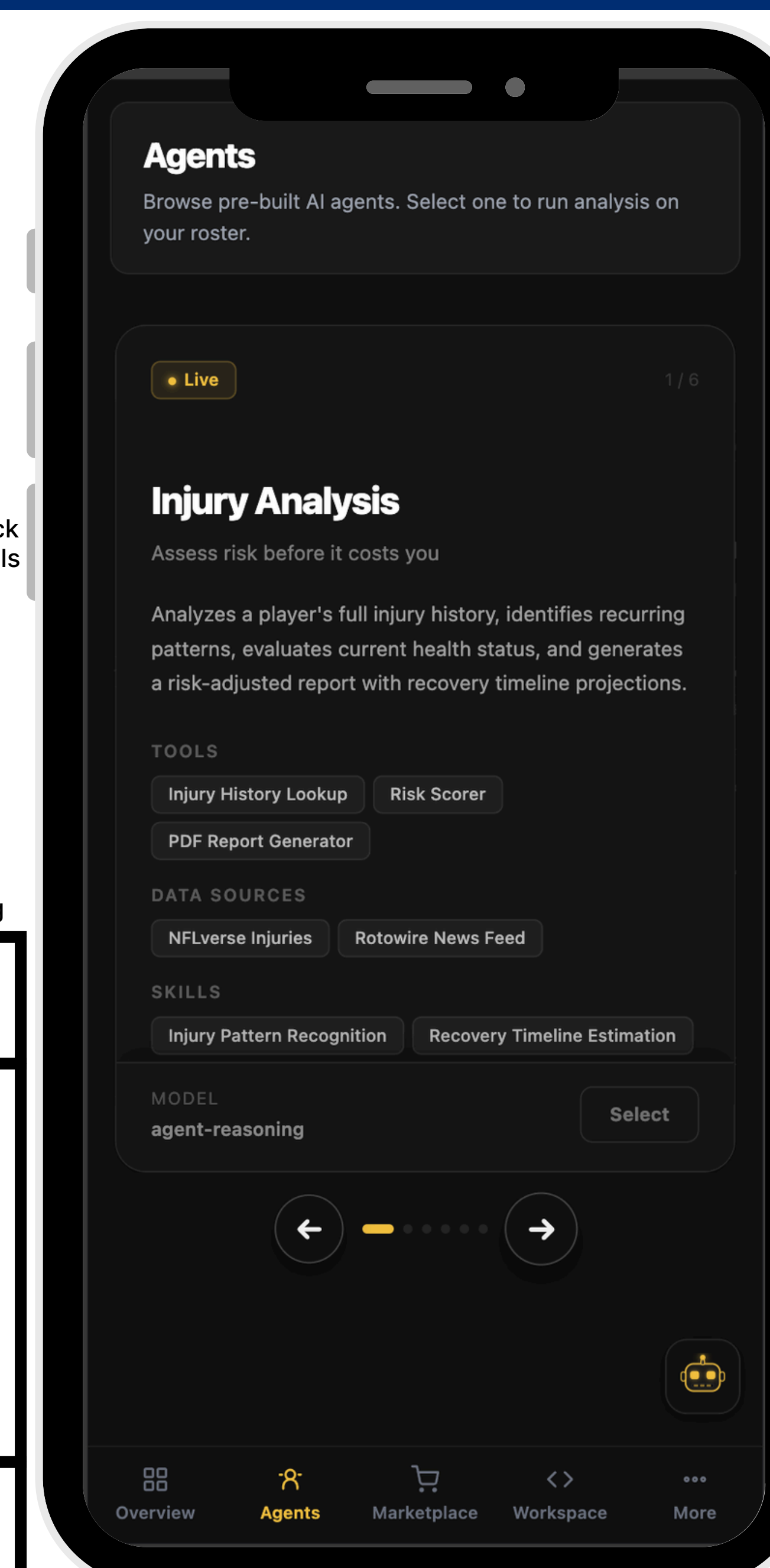
## Implementation & Findings



Player	Agent Focus	Manager Intelligence
A.J. Brown (WR)	Trade Risk & Cap Outlook	<b>\$20M Dead Cap Differential:</b> Aggregates evasive GM quotes with cap data to flag a critical trade decision: moving him before June 1 costs \$20M+ more in immediate dead cap.
A.J. Brown (WR)	On-Field Performance	<b>Efficiency Risk:</b> Notes high-level performance and statistics (1,000 yards, 7 TDs) but highlights efficiency concerns (low yards, costly drops) in critical, high-pressure playoff games.
DeVonta Smith (WR)	Immediate Availability Alert	<b>Thursday Clearance Deadline:</b> Listed as <b>Questionable</b> (chest/shoulder) after missing consecutive practices. Must practice by Thursday to be available for the Friday game.
Dallas Goedert (TE)	Roster Stability & Cap Value	<b>\$20M Cap Savings Confirmed:</b> Validates the 1-year, \$7M contract as a successful cap maneuver that avoided a major \$20M dead cap hit, securing the tight end position.

### 2 AGENT WORKFLOW:

The agent processes a user query by selecting **relevant tools and data sources, retrieving and analyzing** information, and **synthesizing** a cited, evaluated response through coordinated model and agent interactions



Can Provide Intelligence into NFL Teams, like the Philadelphia Eagles

### 3 ACTIONABLE ROSTER INTELLIGENCE:

Our **multi-agent system** transforms the **constant flow of unstructured data** - from GM quotes to injury reports - into **verified, actionable intelligence for managers**. The agents' outputs highlight this **efficiency** by synthesizing **financial risks** (calculating a \$20 million cap differential for A.J. Brown) and **real-time operational clarity** (pinpointing DeVonta Smith's Thursday clearance deadline), allowing staff to make **immediate, high-value roster decisions**.

### THIS IS NOT JUST AN LLM WRAPPER:

**Apollo is a unified sports AI ecosystem** that combines **modular agent infrastructure** with a frontend dashboard and backend APIs to create a reusable framework for building and evaluating domain-specific sports agents. Apollo's high value comes through its **specialized, sports-specific** agents that use tools, skills, and data pipelines to answer questions and provide meaningful output in a structured fashion. A key differentiator of Apollo is its **security-first architecture**. The system supports locally hosted or open-weight LLM inference, ensuring that sensitive data remains in controlled infrastructure. To further improve reliability and trust, Apollo incorporates a **verifier agent layer** that grounds model outputs in at least 2-3 pieces of verifiable evidence.

### Fantasy Football Agent Performance

Query: Justin Fields vs Jordan Love (Superflex stash)  
**General LLM Output**  
 "Fields offers higher upside due to rushing, while Love provides a safer and more stable option."  
 → **Broad, high-level comparison with no supporting data or context**

**Apollo Fantasy Football Agent Output**

- Love: 3,381 passing yards, 23 TDs, 6 INTs → proven QB2 production
- Fields: Injury status + potential team change → volatile role outlook
- Scenario analysis: If starting opportunity opens, Fields' rushing gives league-winning upside

→ **Data-backed, context-aware, and scenario-driven reasoning**

## Future Work

- Expand the data pipeline to support **additional sources** such as user-uploaded datasets for more flexible analysis
- Expansion to **college football** and **non-football** sports
- Improve performance by **post-training** and **fine-tuning** self-hosted models for higher-quality outputs
- Develop **new tools and skills** tailored to user needs, and extend the framework to enable collaboration across multiple agents for end-to-end research, retrieval, analysis, and report generation