AI-Powered Insights from NBA Data using AWS Cloud

Release- 17/10/2025

Team name / student name—

1) Omoniyi Israel

2) Fulsundar Vishal

3)Anurag Surve

Scope of the Project

Main Objective:	A RAG system answering NBA-related questions using real data.
	Focus on AWS Free Tier–compliant components.
	Provide intelligent, grounded insights through natural language queries.

Features

Automated Data Ingestion — daily NBA stats pulled via API.

Data Storage — stored as JSON in Amazon S3.

Embeddings Creation — Titan embeddings (Bedrock) for semantic search.

Question Answering — powered by Claude 3 Haiku model.

REST API Interface — for user queries via API Gateway.

Cloud Monitoring — with CloudWatch logging and metrics)

The project combines data engineering, machine learning, and **AWS cloud services** to demonstrate a modern Retrieval-Augmented Generation pipeline.

Data Source

Primary Source: **nba- api** from NBA

The dataset consist of the wide range of information, such as **basic player and team demographics, detailed game and season statistics**, and advanced metrics like player efficiency ratings and plus/minus

Key Features of the Dataset:

Player and team data: Access team rosters, player profiles, and historical performance stats.

Game information: Retrieve game schedules, results, and real-time scores.

Statistics: Get detailed statistics on player and team performance.

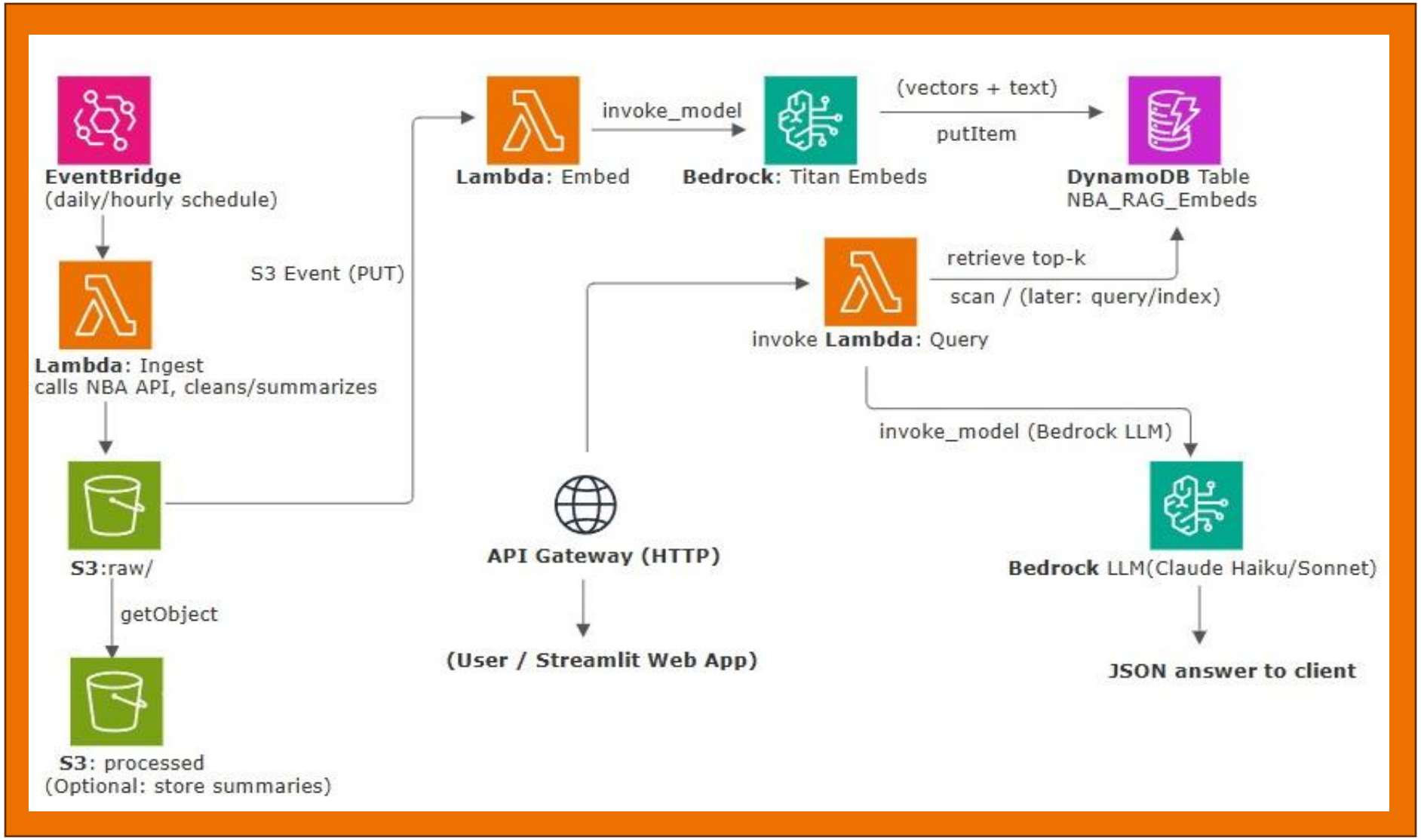
Live data: APIs offer live updates on games in progress.

Historical data:provide access to data from past seasons

Applicable Cloud Services and Data Flow

Layer	AWS Service	Function
Data Ingestion	**AWS Lambda**	Fetch NBA data from API
Storage	**Amazon S3**	Store raw JSON
Data Processing	**AWS Lambda**	Generate text summaries and embeddings
Vector Store	**Amazon DynamoDB**	Store embeddings and context
LLM Inference	**Amazon Bedrock**	Embeddings + Q&A model
API Interface	**Amazon API Gateway**	Expose REST API

AWS Architecture and Data Flow



Expected Outcomes-

1

Technical:

Scalable, low-cost AWS solution.

Working REST API that returns real answers like: LeBron James scored 29 points against the Warriors on Oct 17, 2025.

2

Experience with modern RAG architecture.

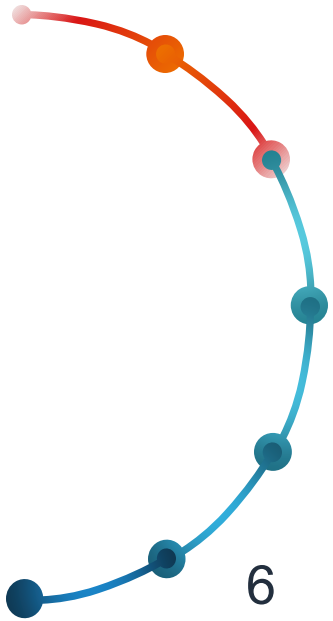
Proficiency in integrating AWS services for AI pipelines.

Cloud cost optimization and security best practices.

3

AWS components we will learn clearly:

- **Lambda, S3,**
- **DynamoDB,**
- **Bedrock,**
- **API Gateway,**



References

https://developer.sportradar.com/basketball/docs/nba-ig-api-basics

Python NBA API - https://pypi.org/project/nba_api/1.1.5/

Bedrock- https://docs.aws.amazon.com/bedrock/

Thank You!