Al Akhawayn University

Start

AI-POWERED VISION FOR THE BLIND AND VISUALLY IMPAIRED

Topic : THESIS DEFENSE | Date : 09/02/2024 | Submit by : AYOUB MAIMMADI

Supervised by :

Dr. Asmaa Mourhir

Other Examiners :

Dr. Omar Iraqi, Al Akhawayn University
Pr. Abdelouahed Sabri, Sidi Mohamed Ben Abdellah University

INTRODUCTION

OVERVIEW

01 MOTIVATION

02 OBJECTIVES

03 LITERARY PREVIEW

04 SYSTEM DESIGN

05 IMPLEMENTATION

06 RESULTS

07 DISCUSSION

08 CONCLUSION

MOTIVATION

+150,000

As of 2020, an estimated 3.7 million people in Morocco were living with some form of vision loss. Within this population, approximately more than 150,000 individuals were classified as blind, encompassing those with complete blindness as well as those experiencing severe visual impairments, such as extremely blurry vision that significantly impacts their daily activities.

These figures are based on data from the Vision Loss Expert Group (VLEG) and the Global Burden of Disease (GBD) 2020 model

LITERATURE REVIEW

Traditional hybrid architectures use techniques like SIFT and BoVW for robust object detection and scene description in varying conditions.

VQA combines vision and language processing, using attention mechanisms to answer contextually relevant questions about visual scenes.

Description generation techniques like REG provide precise, context-aware descriptions enhancing scene understanding.

Multimodal and vision-language models like CLIP and BLIP laid the groundwork for Audieyes' real-time, accurate scene description capabilities.

LITERATURE GAP

GAP 01

Existing technologies lack **real-time**, **contextually relevant** assistance, often providing generic descriptions unsuitable for dynamic environments faced by BVI users.

GAP 02

Insufficient support for **Moroccan culture and Darija language** in AI models limits effectiveness for local visually impaired users.

GAP 03

Limited **practical end to end implementation** and scalability of advanced AI in assistive technologies restrict widespread adoption and long-term usability.

OBJECTIVES

OBJECTIVE 01

Leverage advanced **Large Language Models (LLMs)** for real-time, culturally relevant scene descriptions tailored to the needs of Moroccan BVI users

OBJECTIVE 02

Ensure robust support for **Moroccan Darija**, enhancing accessibility and user experience for visually impaired individuals in Morocco.

OBJECTIVES

OBJECTIVE 03

Enhance **autonomy and mobility** of BVI users through accurate, context-sensitive environmental descriptions in real-time.

OBJECTIVE 04

Develop a **fault tolerant**, **cost-effective** system seamlessly integrating into BVI users' daily lives, maintaining high performance and reliability.

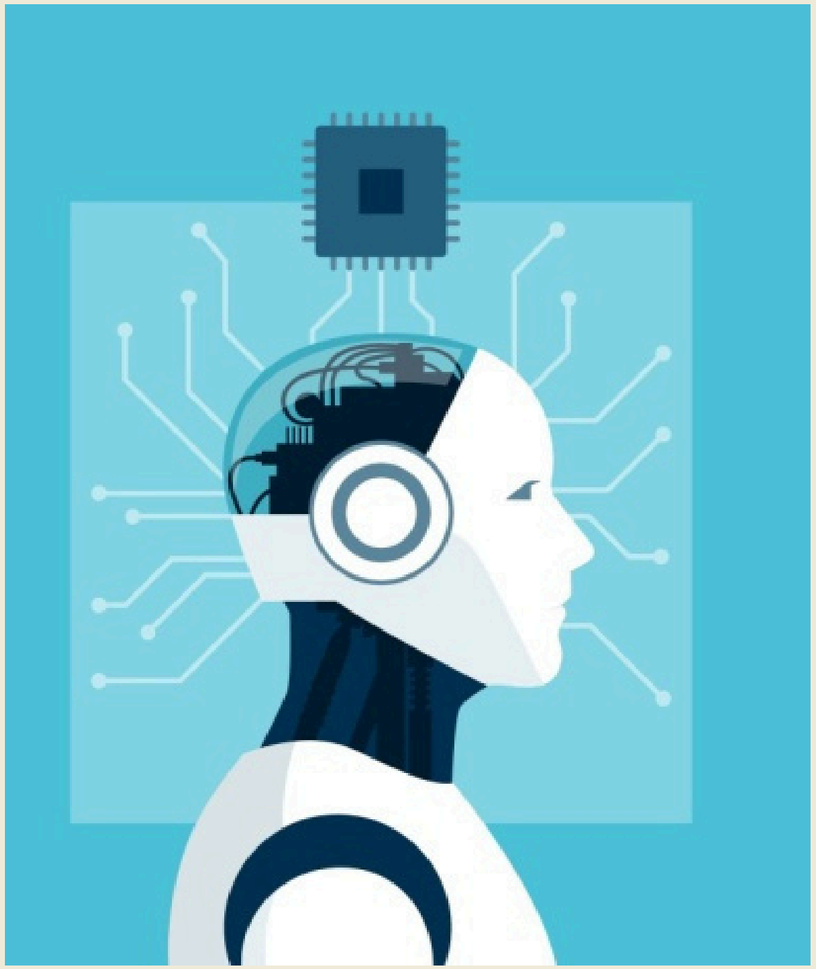
CHALLENGES
OF AI POWERED APPS

CHALLENGE 01

Choosing the **best** AI involves balancing **performance**, **accuracy**, and **compatibility** with the specific needs of the application and its users.

CHALLENGE 02

Ensuring the **safest** and most **cost-effective** AI requires robust security measures and careful **cost analysis** without sacrificing quality or reliability.

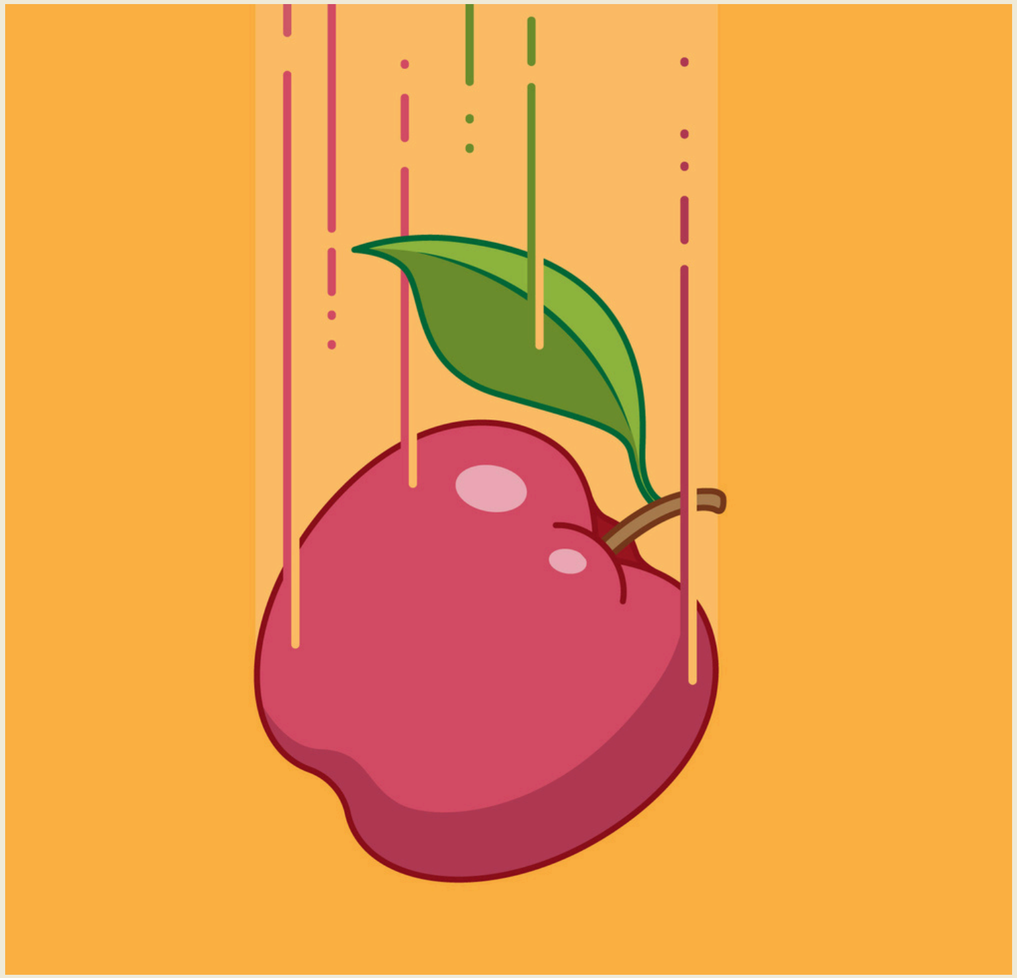


CHOOSING
THE BEST AI

After thorough feasibility analysis, **GPT-4o-mini** was chosen for its optimal performance, language support, and cost-effectiveness.

- MMMU: **59.1%**
- MathVista: **56.1%**
- AI2D: **88.9%**
- ChartQA: **82.3%**
- DocVQA: **90.3%**
- Input Tokens Cost: **$0.150 (per Million Tokens)**
- Output Tokens Cost : **$0.600 (per Million Tokens)**

GPT-4o mini



SYSTEM DESIGN

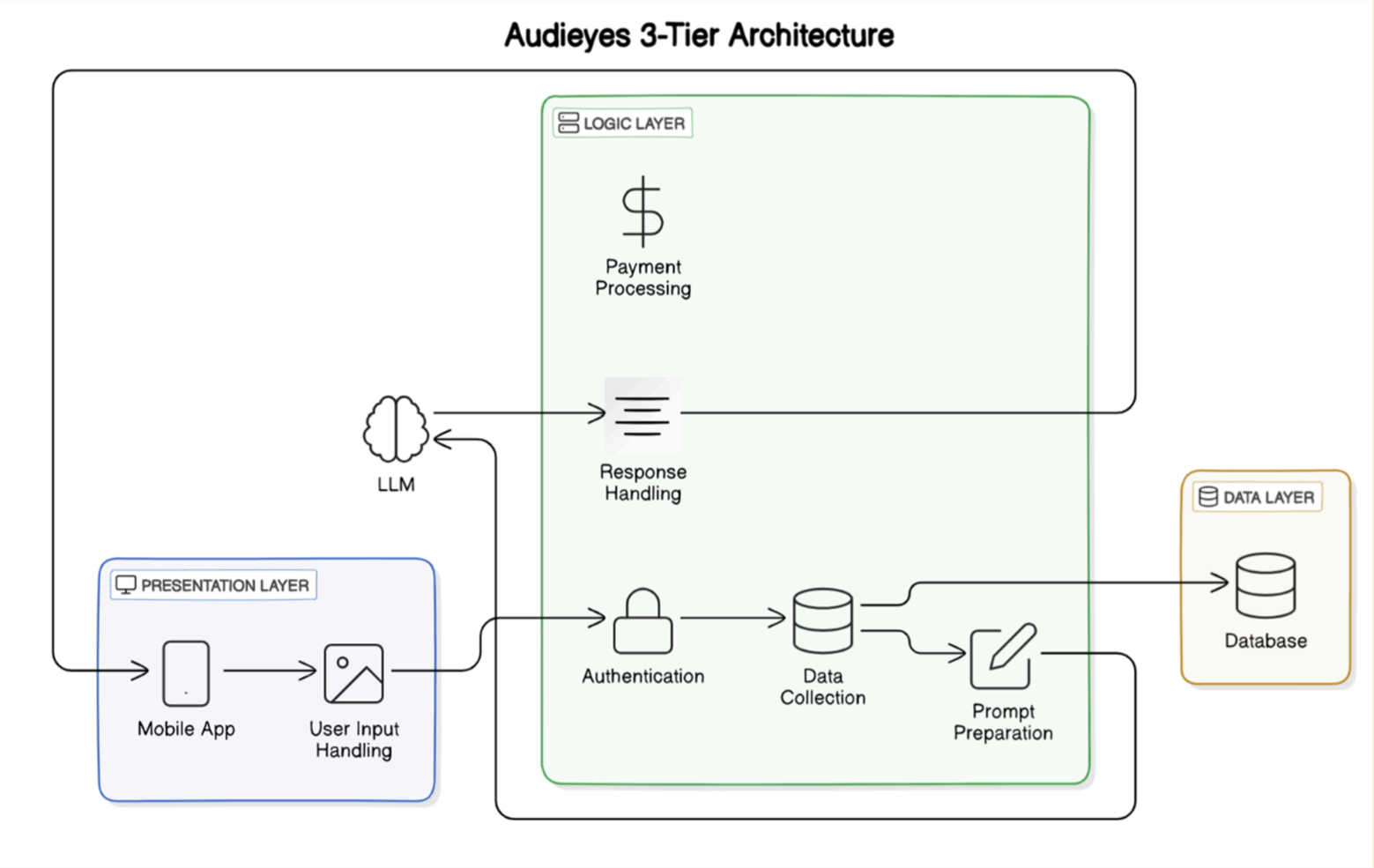
MODULAR AND SCALABLE ARCHITECTURE

The modular and scalable architecture supports easy integration of new features and adapts to evolving user needs and size.

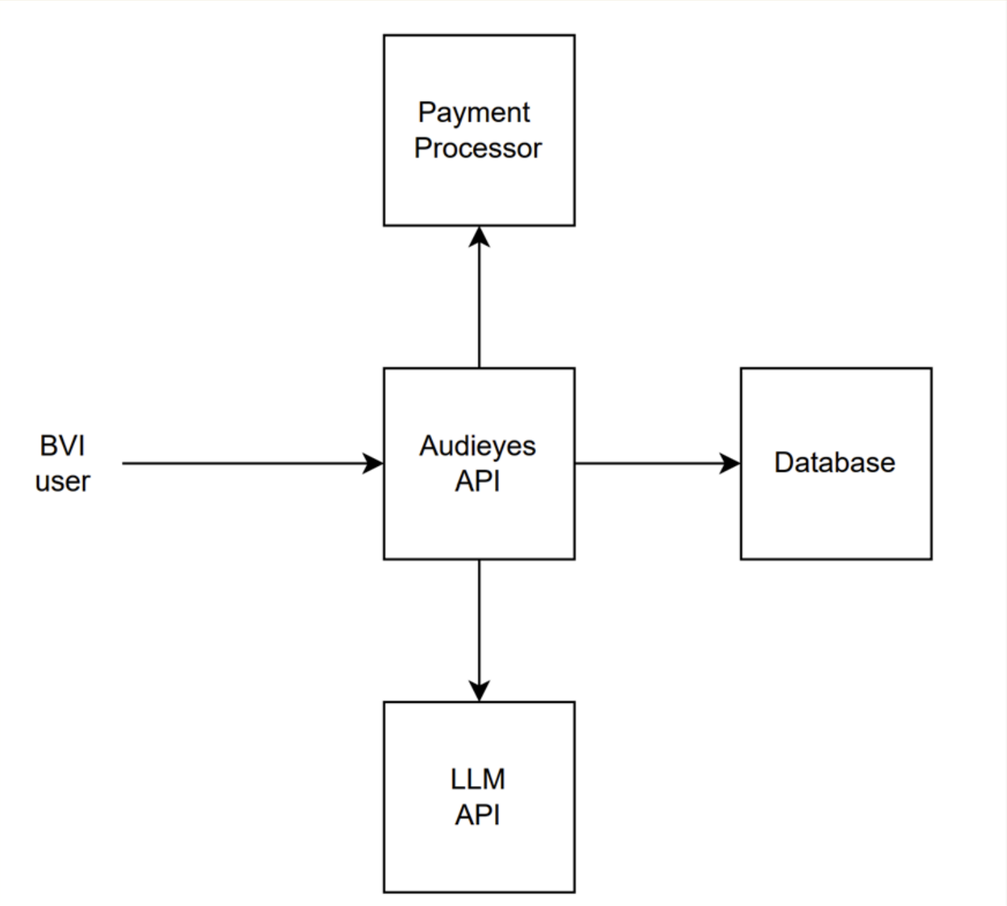
FAULT TOLERANCE & HIGH AVAILABILITY

The system ensures fault tolerance and high availability through API redundancy, database replication, load balancing, and more.

SYSTEM ARCHITECTURE



STARTING SYSTEM DESIGN



WEAKNESS 01

Dependency on an **external service** like the LLM API and payment processor makes the system vulnerable to their potential downtimes.

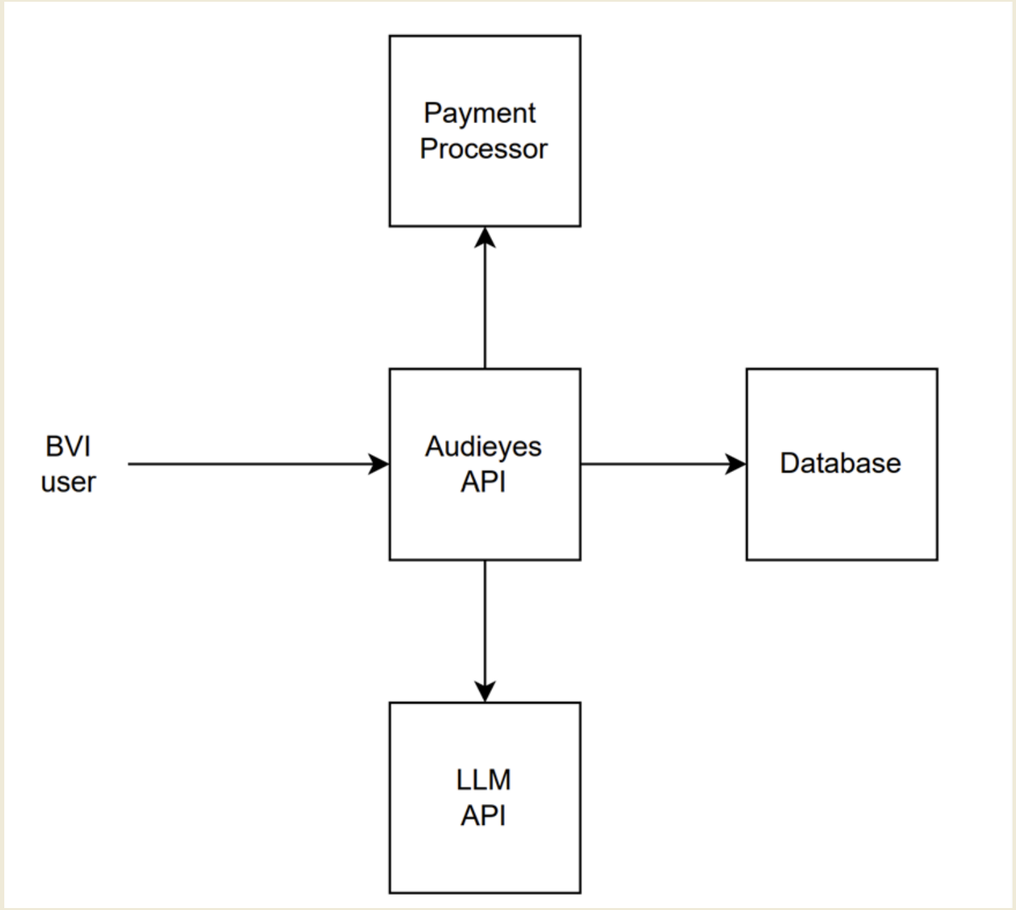
WEAKNESS 02

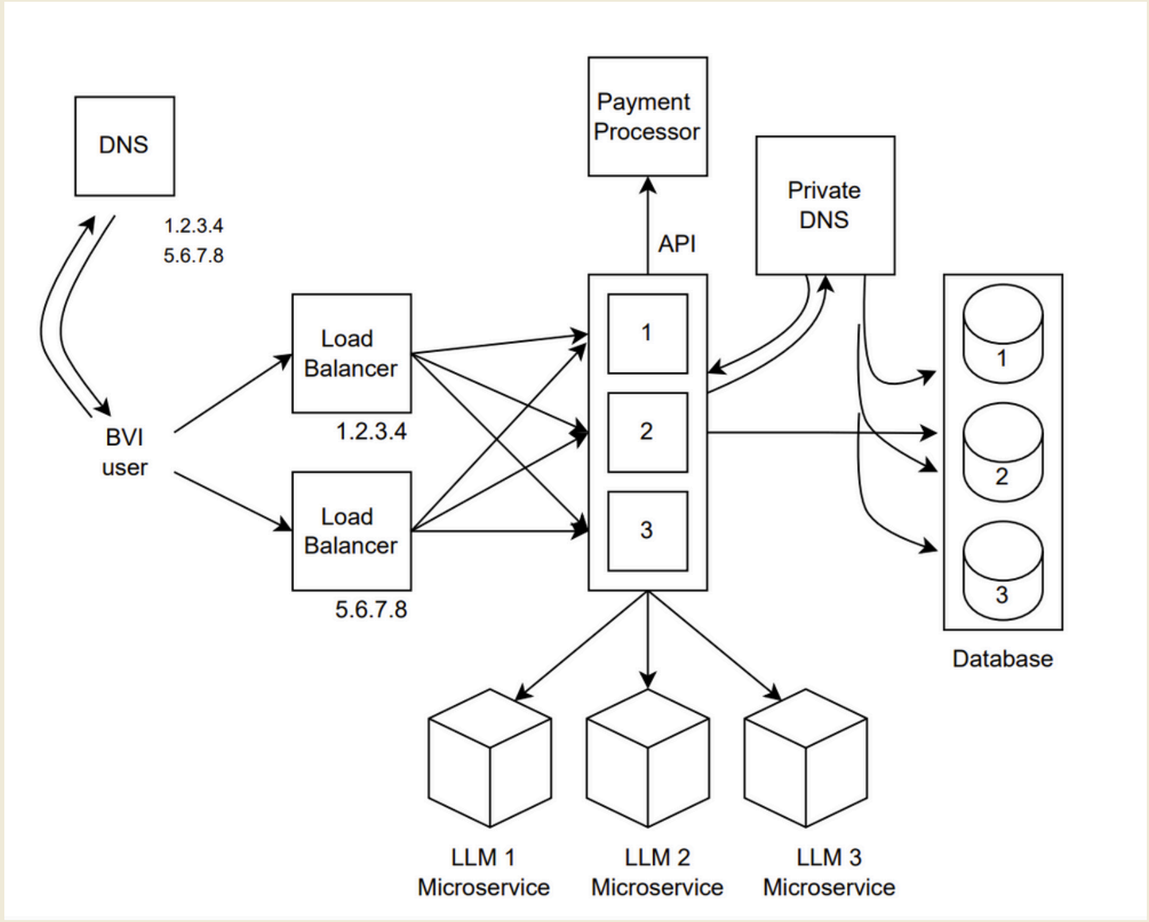
Database issues such as **connection failures** or **slow query** performance can significantly disrupt the user experience and system functionality.

WEAKNESS 03

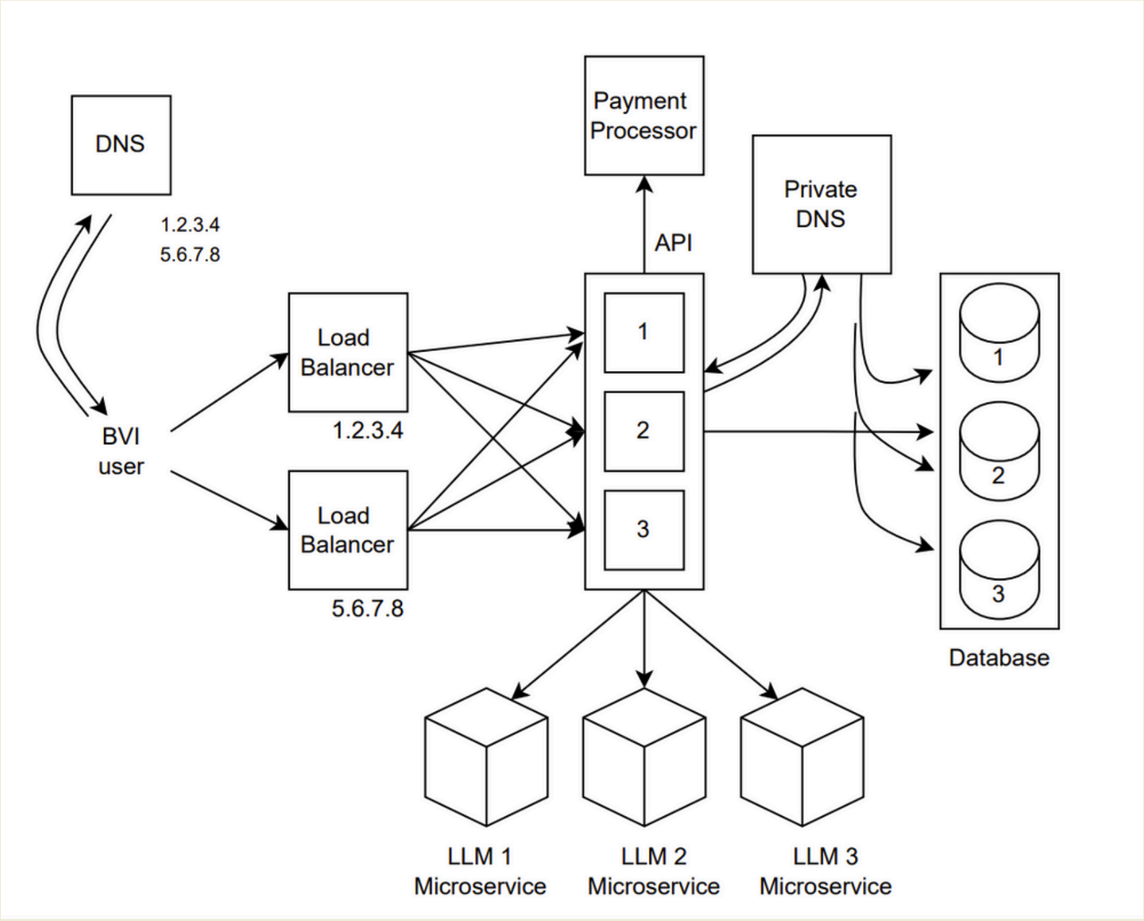
The system has many **Single Points of Failure**, making it highly susceptible to complete outages if one fails.

FAULT TOLERANT SYSTEM DESIGN





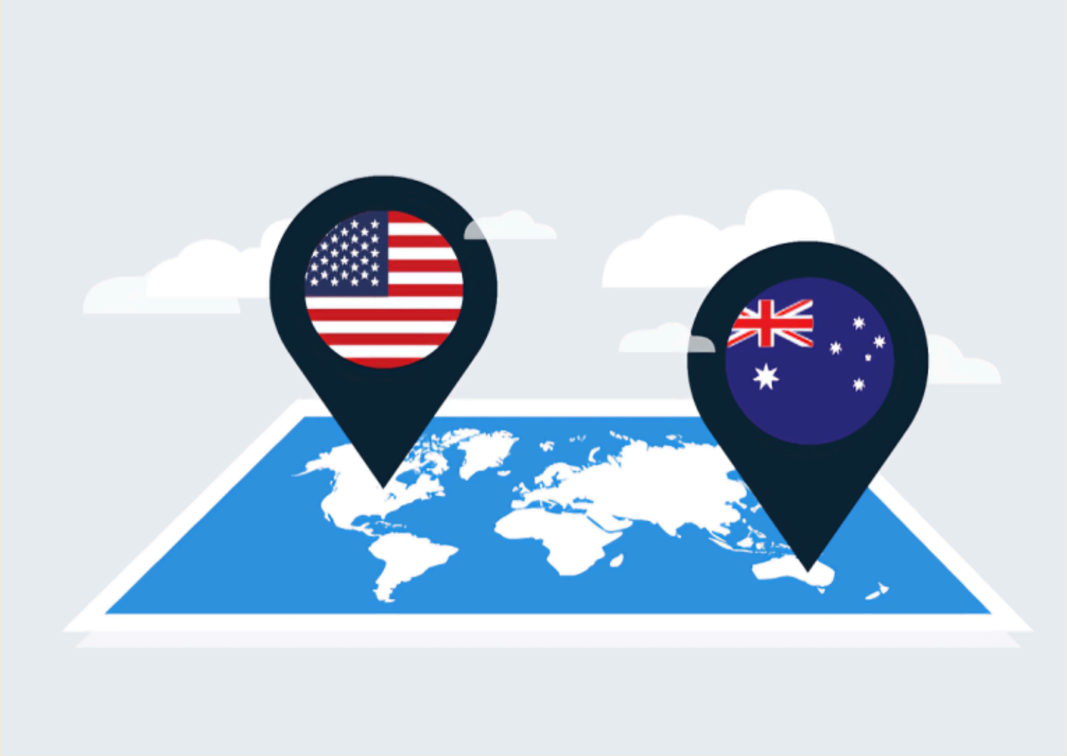
FAULT TOLERANT SYSTEM DESIGN



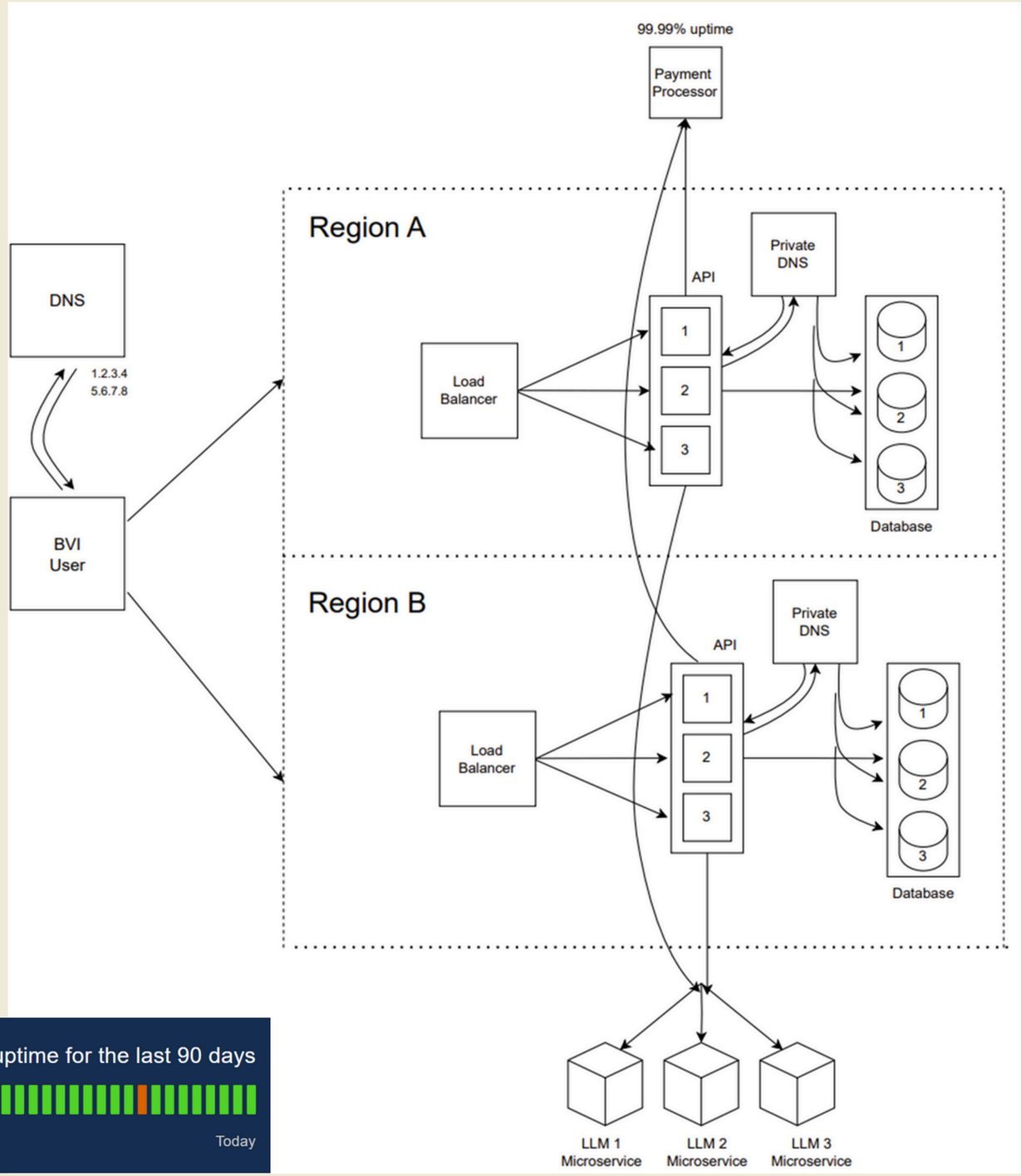
STRENGTHS

- Redundancy and Load Balancing
- Scalability
- Improved Fault Isolation
- Private DNS for Internal Resolution
- High Availability of Database
- Flexible Microservices Architecture
- Resilience Against External Service Failures

MULTI-REGION SYSTEM DESIGN



https://status.stripe.com/



IMPLEMENTATION

TECHNOLOGY ENABLERS

Utilized advanced LLMs like GPT-4-mini, alongside NestJS and React Native, to enable real-time, culturally relevant assistive capabilities.

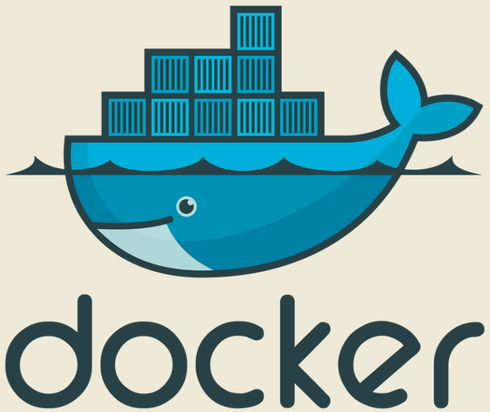
DEVOPS & LLMOPS

Implemented a CI/CD, and applied LLMOps for continuous model optimization, performance monitoring, and secure deployment.

DEPLOYMENT STRATEGY

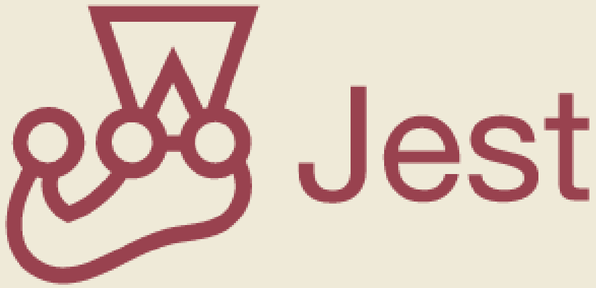
Deployed using multi-region Linode infrastructure with Kubernetes orchestration, ensuring scalability, fault tolerance for real-time services.

TECHNOLOGY ENABLERS

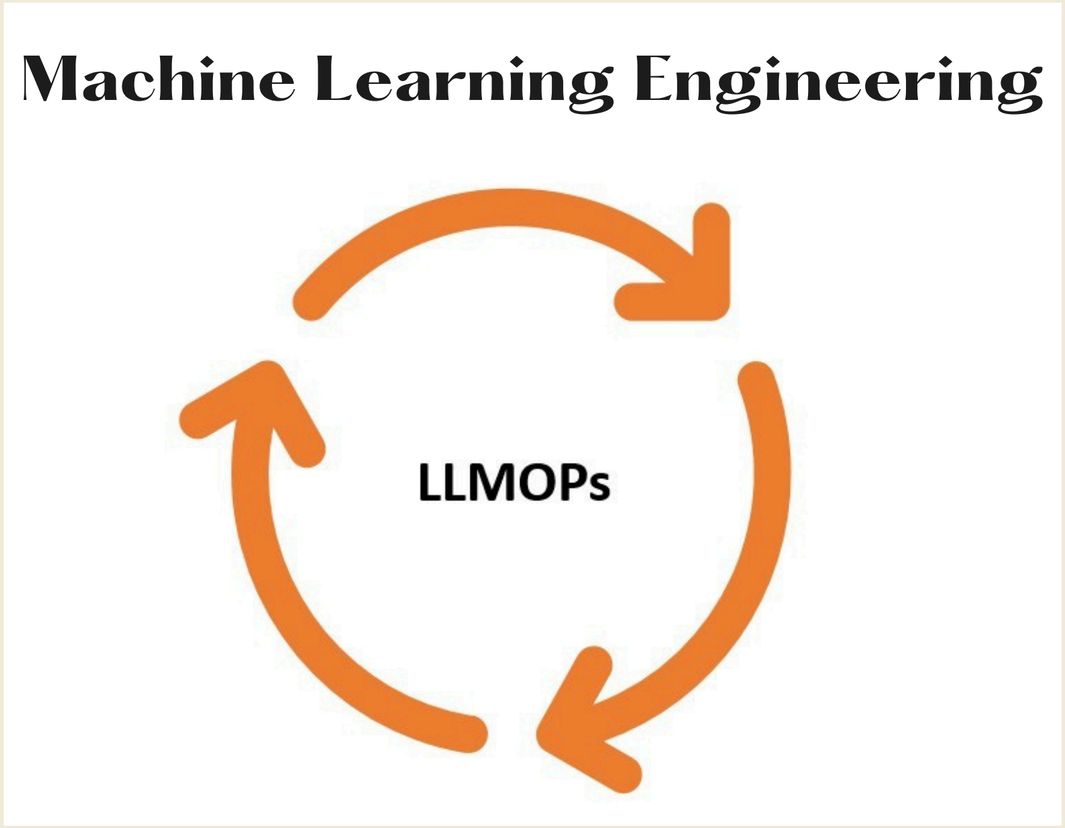


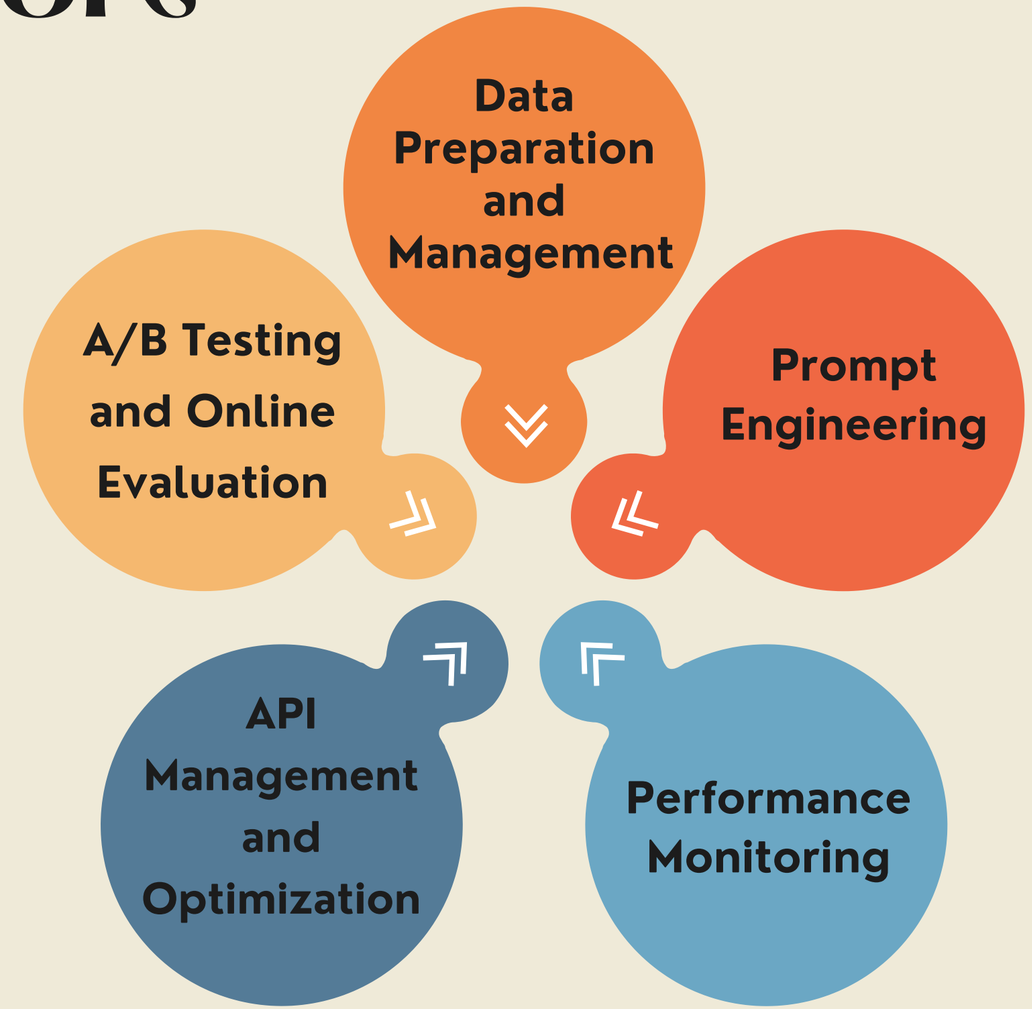
kubernetes

linode

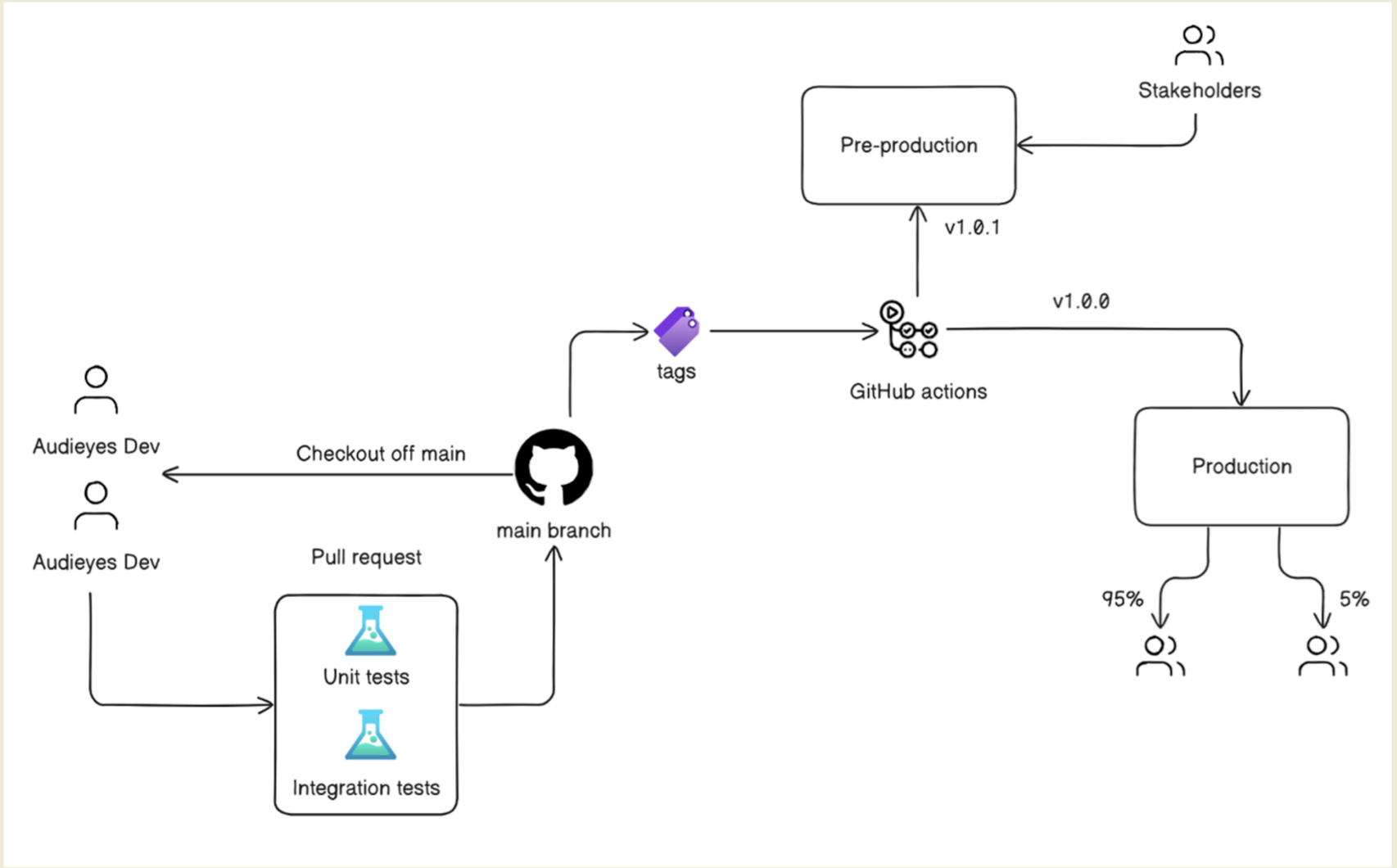


LLMOPS


Machine Learning Engineering
LLMOPs


Data Preparation and Management
Prompt Engineering
Performance Monitoring
API Management and Optimization
A/B Testing and Online Evaluation

DEVOPS


Stakeholders
Pre-production
v1.0.1
v1.0.0
tags
GitHub actions
Audieyes Dev
Checkout off main
main branch
Audieyes Dev
Pull request
Unit tests
Integration tests
Production
95%
5%

Deployment Strategy - Canary Deployment

01

Canary deployment in Audieyes uses Kubernetes to **gradually release** updates, ensuring stability by **testing** changes on a small user subset.

02

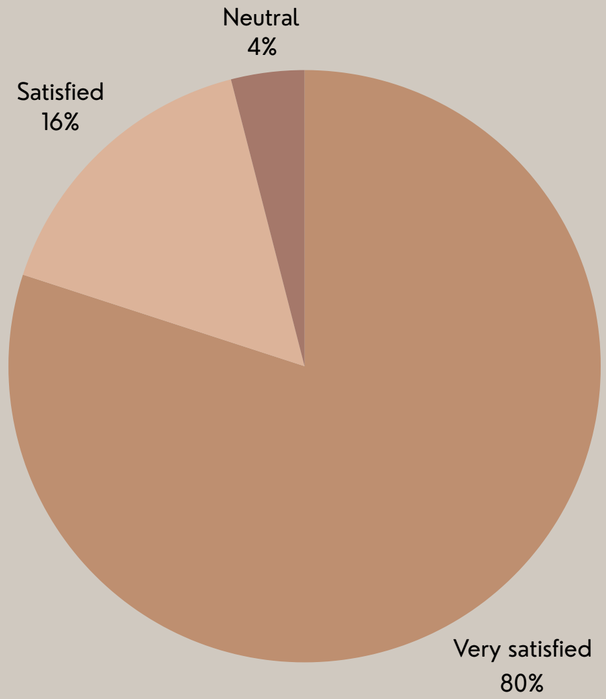
Kubernetes enables dynamic canary deployments, allowing **real-time monitoring** and **quick rollback** of updates to maintain Audieyes' reliability.

03

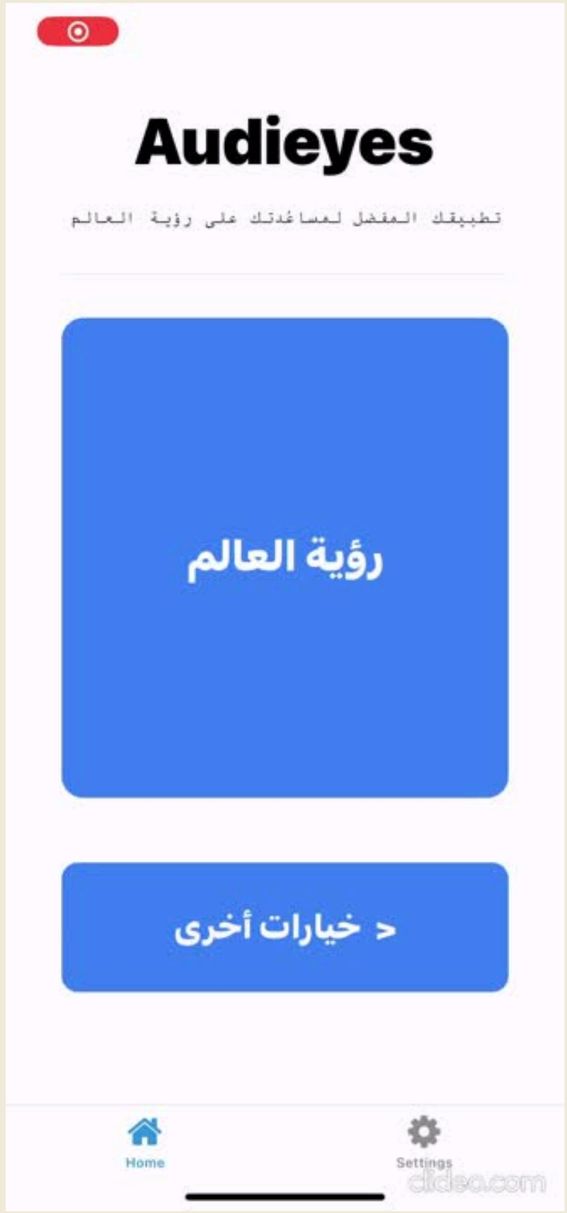
This strategy **minimizes risks**, ensuring new features are **stable** and user feedback is **positive** before full-scale deployment across Audieyes.

RESULT

- Audieyes Demonstration
- Online Testing Results
- User Evaluation



AUDIEYES DEMONSTRATION



ONLINE TESTING RESULTS

RESULT 01

The system achieved a **97% accuracy rate in object classification**, ensuring reliable identification in diverse real-world scenarios.

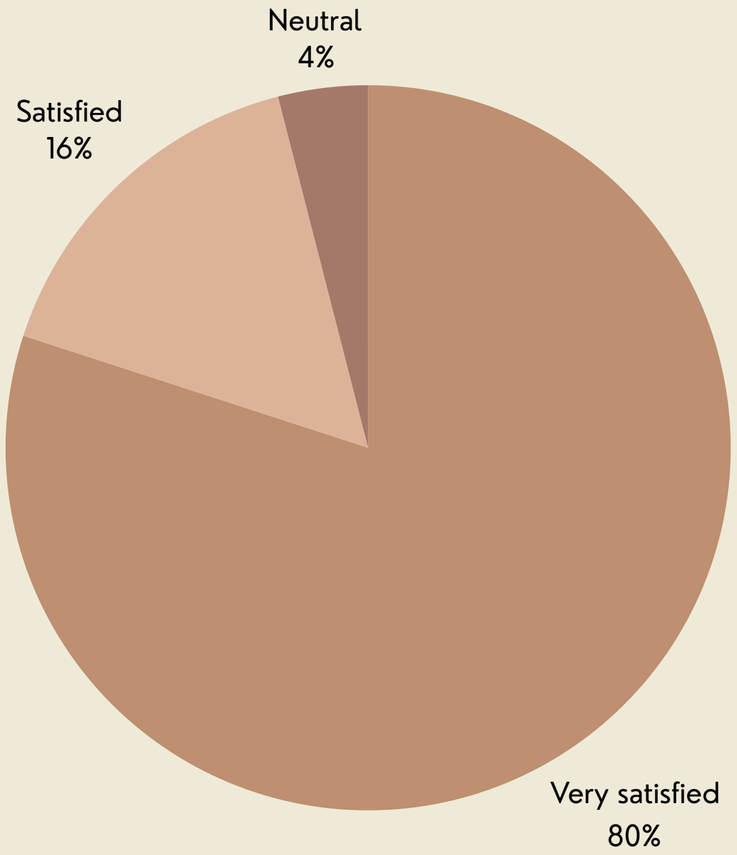
RESULT 02

Audieyes achieved a **BLEU-4 score of 40.12%**, surpassing the **38.6% baseline of the BLIP model**, indicating superior descriptive accuracy.

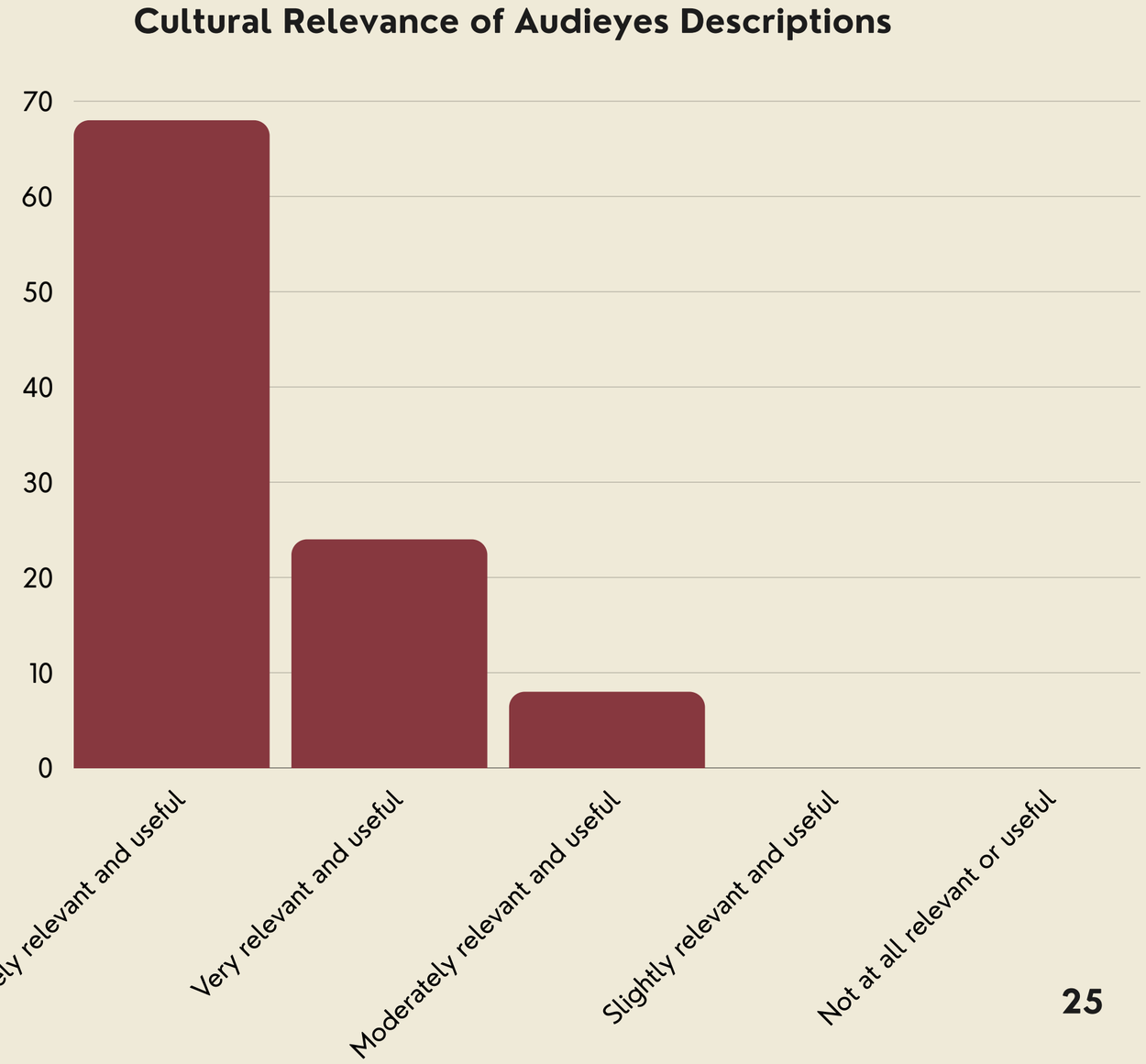
RESULT 03

The system consistently delivered an **average response time of 3.5 seconds**, enabling relatively real-time assistance for users.

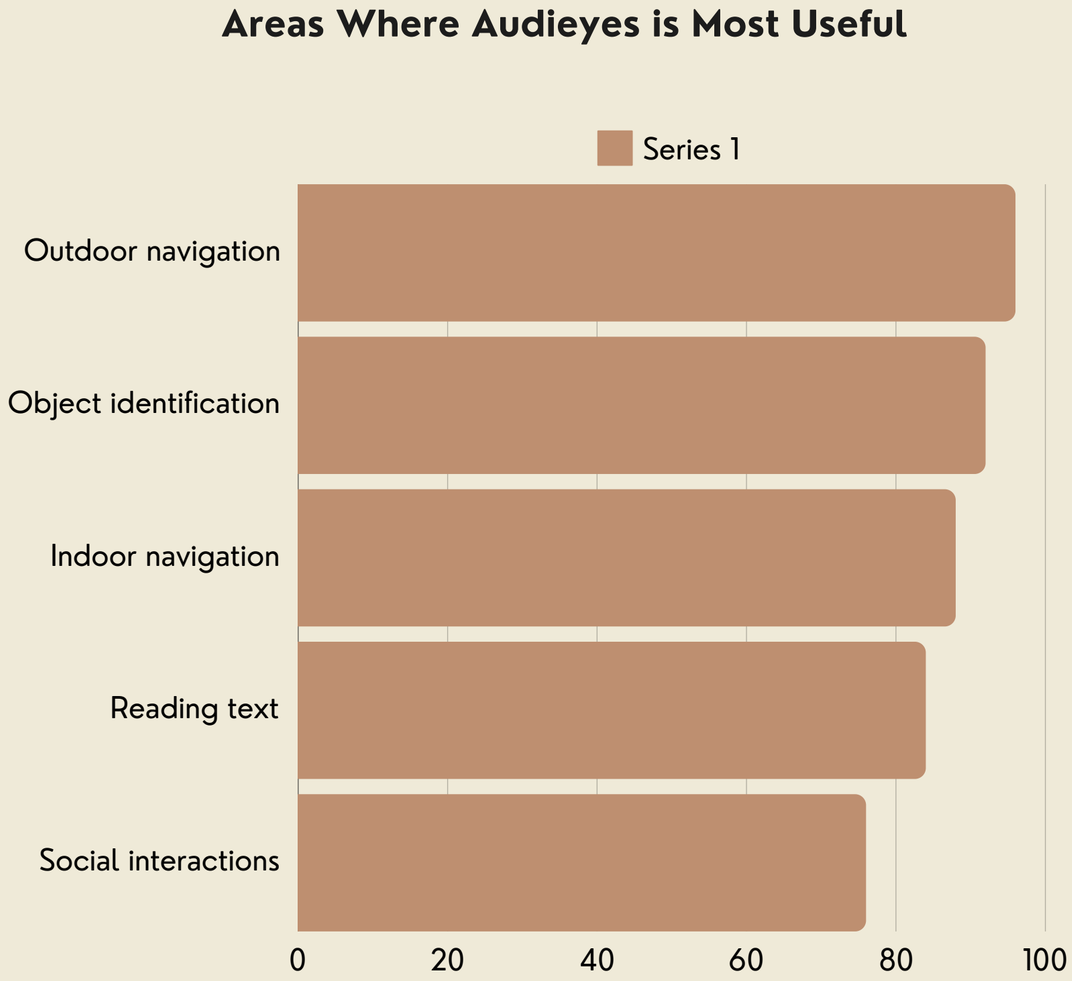
USER EVALUATION

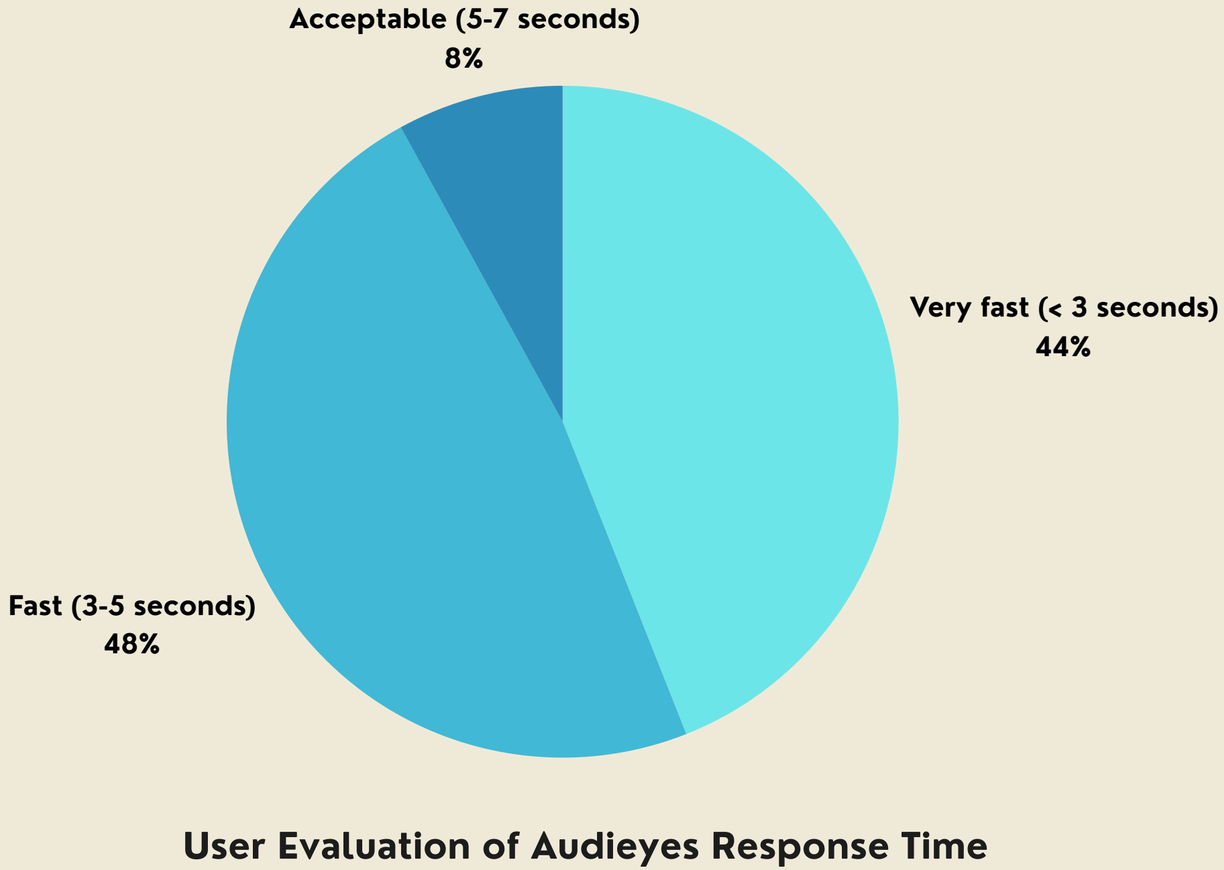


Overall User Satisfaction with Audieyes



USER EVALUATION





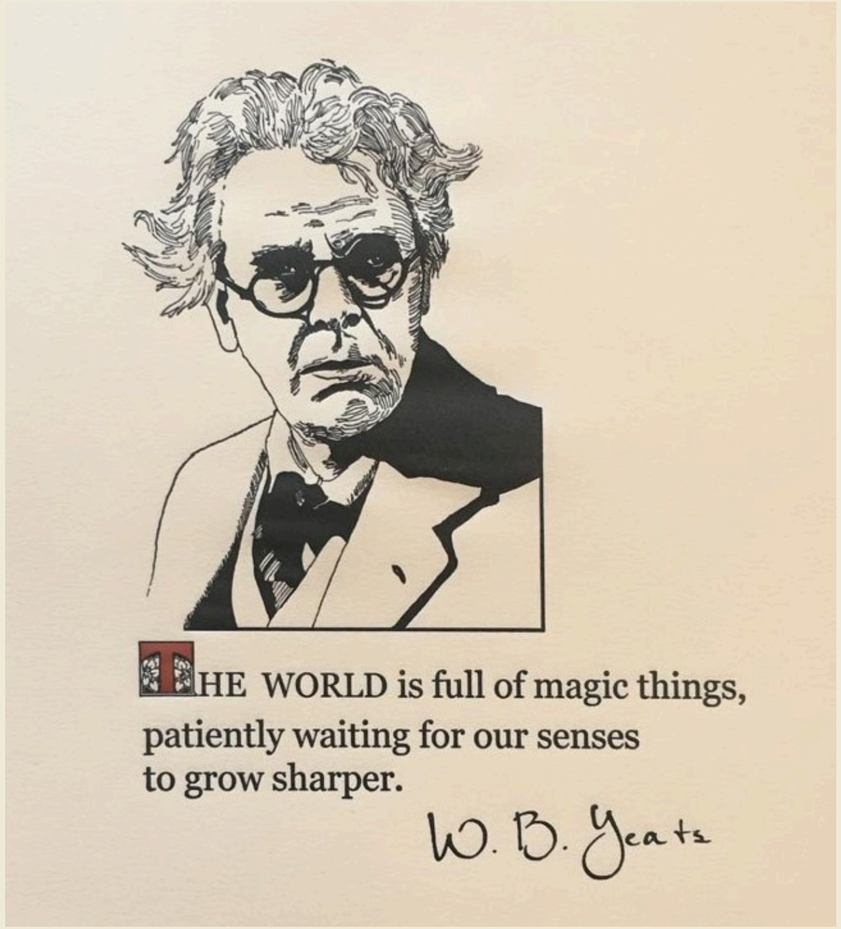
FEEDBACK

FEEDBACK 01

Users suggested incorporating **haptic feedback** into Audieyes, with different vibration patterns indicating navigation directions or proximity to objects, enhancing the tactile experience for visually impaired users during movement.

FEEDBACK 02

Users expressed interest in a feature enabling **real-time assistance from sighted volunteers** via the app, allowing visually impaired users to receive immediate help when navigating challenging environments or tasks in Darija.



CONCLUSION

- Leveraged Advanced LLMs
- Supported Moroccan Darija
- Reliable and Cost-Effective Implementation
- Positive Evaluation Results

500 input tokens
100 Output tokens

15,000
IMAGES

THANK YOU

FOR YOUR ATTENTION

Topic : THESIS DEFENSE

Submit by : AYOUB MAIMMADI

Al Akhawayn University

Start

AI-POWERED VISION FOR THE BLIND AND VISUALLY IMPAIRED

Topic : THESIS DEFENSE | Date : 09/02/2024 | Submit by : AYOUB MAIMMADI

Supervised by :

Dr. Asmaa Mourhir

Other Examiners :

Dr. Omar Iraqi, Al Akhawayn University
Pr. Abdelouahed Sabri, Sidi Mohamed Ben Abdellah University