

# GENERACIÓN DE ESCENAS EN REALIDAD EXTENDIDA MEDIANTE VOZ

*Rodrigo Alonso*

# Contenidos

**1**

Introducción

**2**

Tecnologías utilizadas

**3**

Componentes

**4**

Desarrollo del proyecto

**5**

Experimentación

**6**

Conclusiones y trabajos  
futuros

**7**

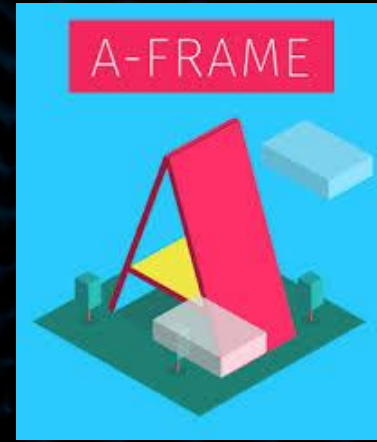
DEMO

# Introducción



# Tecnologías utilizadas

OpenRouter



# Componentes



**Lectura de teclado**  
**text-input**



**Identificación de comandos**  
**command-router**



**Consulta LLM**  
**llm-client**



**Estado de la escena**  
**scene-orchestrator**



**Reconocimiento de voz**

**voice-input**

**voice-input-speechapi**

**voice-input-vosk**

**voice-input-groq**



**Creación de escenas**

**room-renderer**

**aframe-lounge**

**primitives-manager**

**objects-catalog**

**star-sky**



**Persistencia remota**  
**push-to-github**



**Ayuda**

**help-panel**

**error-panel**

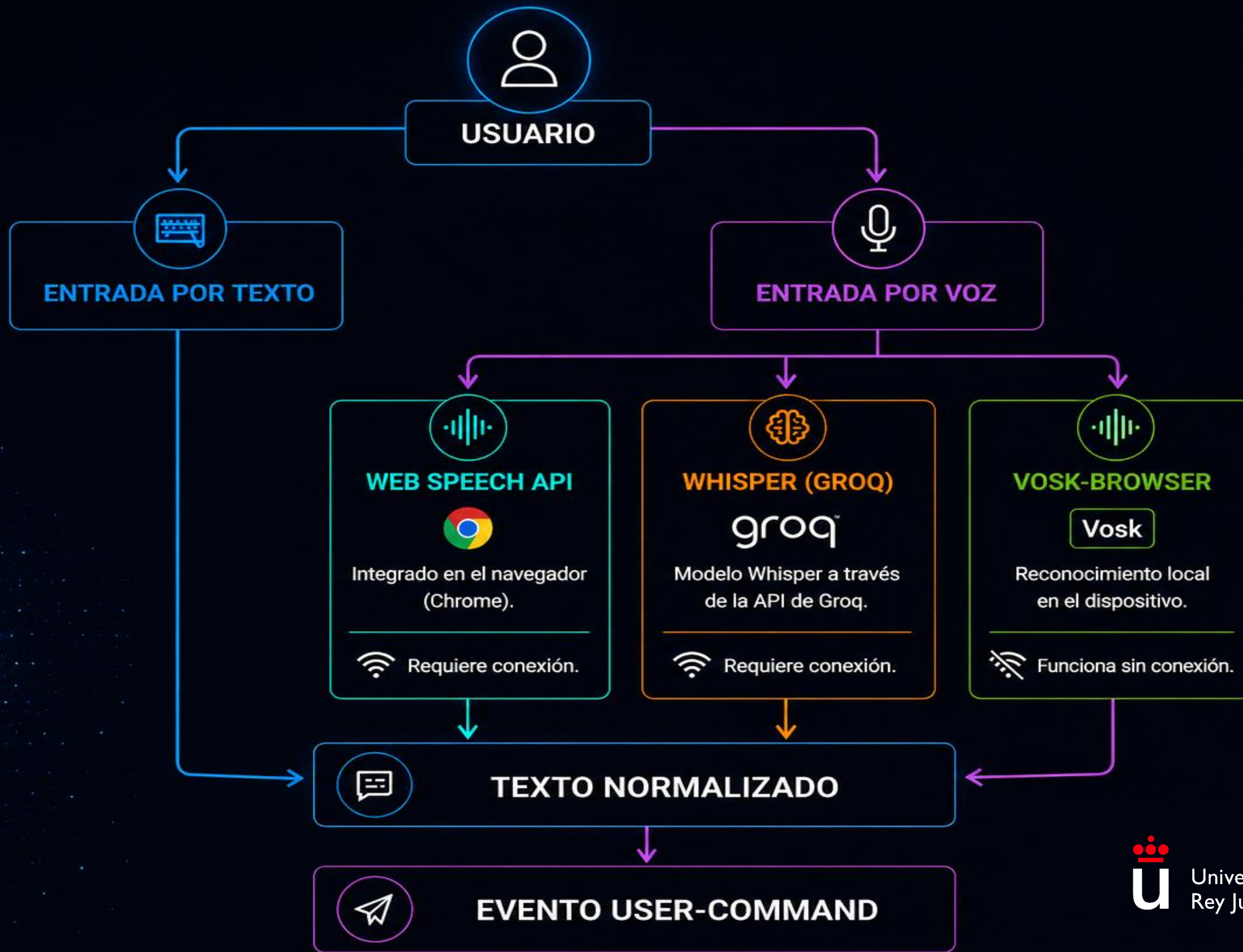
**success-push-panel**

**error-push-panel**

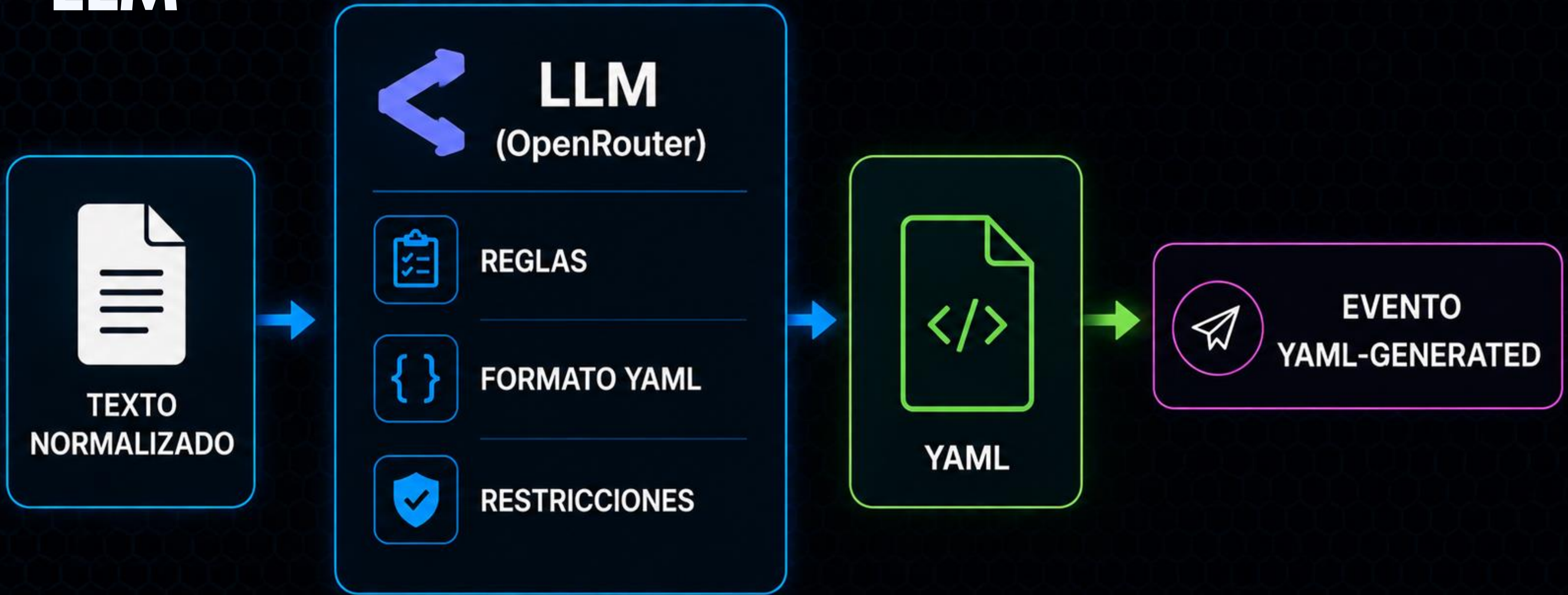


Universidad  
Rey Juan Carlos

# INTERFAZ DE USUARIO



# LLM



# GENERACIÓN DE ESCENAS

Escenario guardado correctamente en GitHub.

ERROR: Instrucción no válida. Di 'que puedo hacer' o 'muestra un panel de ayuda' para ver opciones.

BIENVENIDO AL SISTEMA DE CREACION DE ESCENAS XR CON LENGUAJE NATURAL

Este sistema permite crear y editar escenas XR usando lenguaje natural.

Puedes interactuar de dos formas:

- Voz: en VR manten pulsado el boton de atras del mando derecho y habla. En escritorio pulsa el boton 'Hablar' desde el panel 'Control de habitacion'.
- Texto: escribe un comando desde el panel 'Control de habitacion'

Puedes pedir:

- Crear escenas completas mediante figuras geometricas.
- Cambiar las dimensiones de la habitacion.
- Insertar objetos en 3D y figuras de A-Frame.
- Modificar propiedades de los objetos.
- Guardar el escenario en GitHub.

Ejemplos:

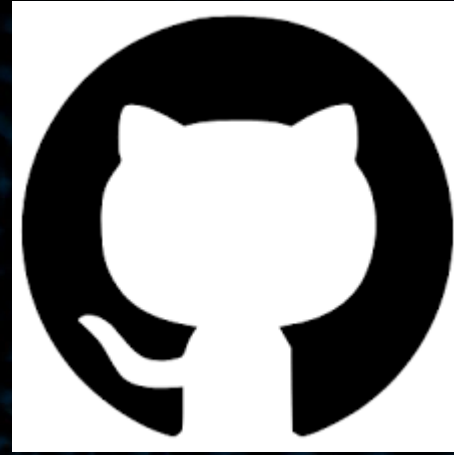
- "Crea una playa con sombrillas"
- "Crea un bosque con arboles y un rio"
- "Pon una esfera roja"
- "Guarda el escenario"

Di "muestra un panel de ayuda" para volver a abrir este panel.



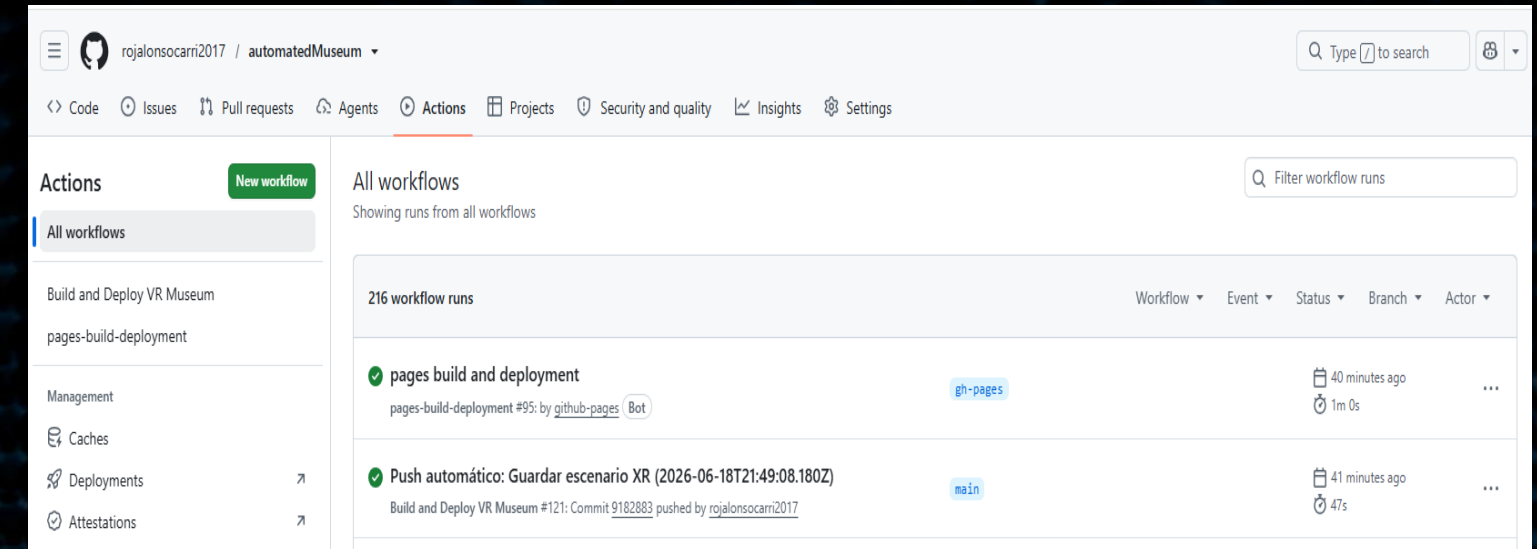
# PERSISTENCIA EN GITHUB

“ GUARDA ESTE ESCENARIO ”



room.yaml

```
rooms:  
  width: 200  
  depth: 300  
  height: 100  
  ceiling: true  
walls:  
  north: barrier  
  east: glass  
  south: wall  
  west: wall  
textures:  
  floor: ./assets/floor-texture.jpg  
  wall: ./assets/brick-texture.jpg  
  ceiling: ./assets/jereqlifico.jpg  
entryPoint:  
  x: -80  
  y: 2  
  z: 20  
environment:  
  skyColor: '#8808F8'  
  stars: true  
lights:  
  - type: ambient  
    color: '#FFFFFF'  
    intensity: 0.85  
  - type: ambient  
    color: '#ebf3e9'  
    intensity: 0.3  
objects: []
```



build  
succeeded 41 minutes ago in 43s

- ✔ Set up job
- ✔ Checkout repository
- ✔ Set up Python
- ✔ Install dependencies
- ✔ Generate HTML from YAML
- ✔ Deploy to GitHub Pages
- ✔ Post Set up Python
- ✔ Post Checkout repository
- ✔ Complete job

Triggered via GitHub Pages 42 minutes ago

Actor	Status	Total duration	Artifacts
github-pages[bot] -> e1c8839 gh-pages	Success	1m 0s	1

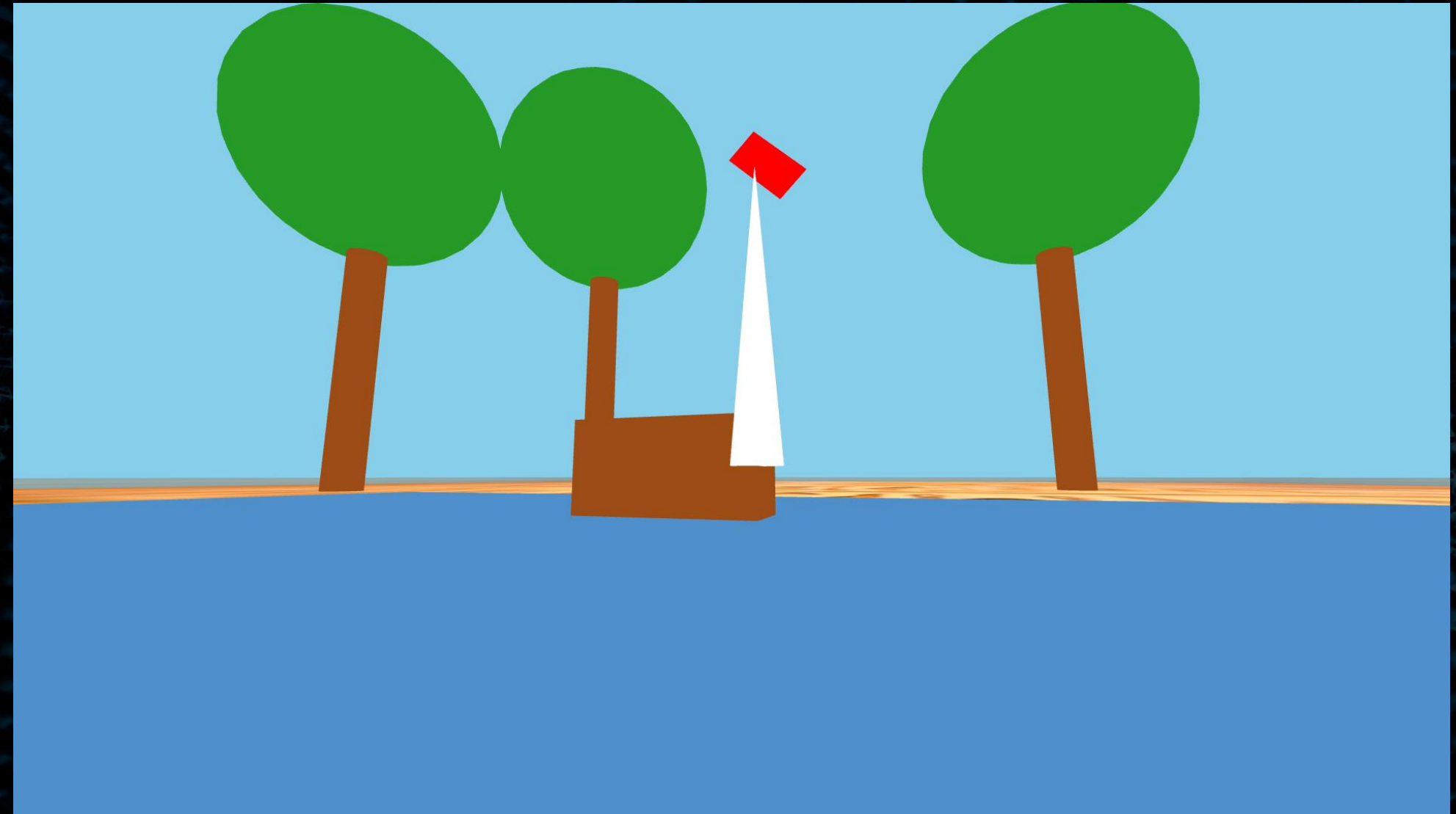
pages-build-deployment  
on: dynamic

```
graph LR; build[✔ build 31s] --> report[✔ report-build-status 6s]; report --> deploy[✔ deploy 21s];
```

https://rojalonsocari2017.github.io/auto...

# RESULTADO

“ QUIERO UN LAGO CON UN  
BARCO VELERO RODEADO DE  
ARBOLES ”



# Desarrollo del proyecto



# Desarrollo del proyecto



# Experimentación



SUJETO 1

SUJETO 2

SUJETO 3

SUJETO 4

SUJETO 5

PERFIL TÉCNICO



EXPERIENCIA VR



USO DISPOSITIVOS DIGITALES



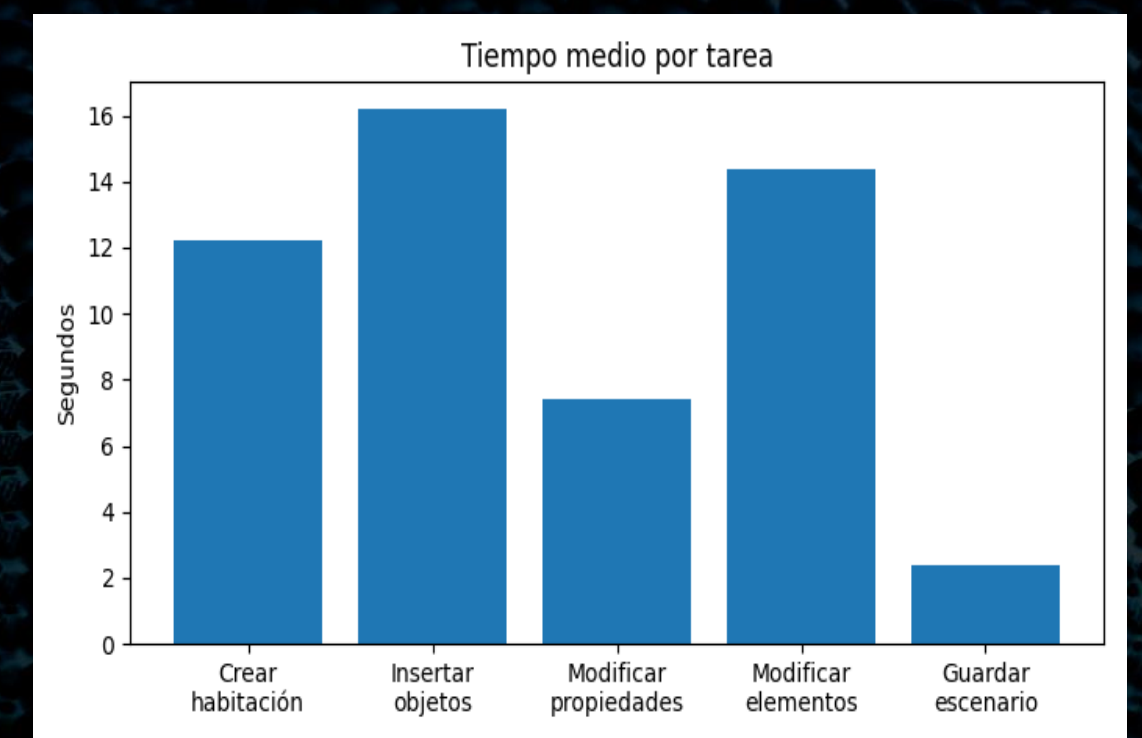
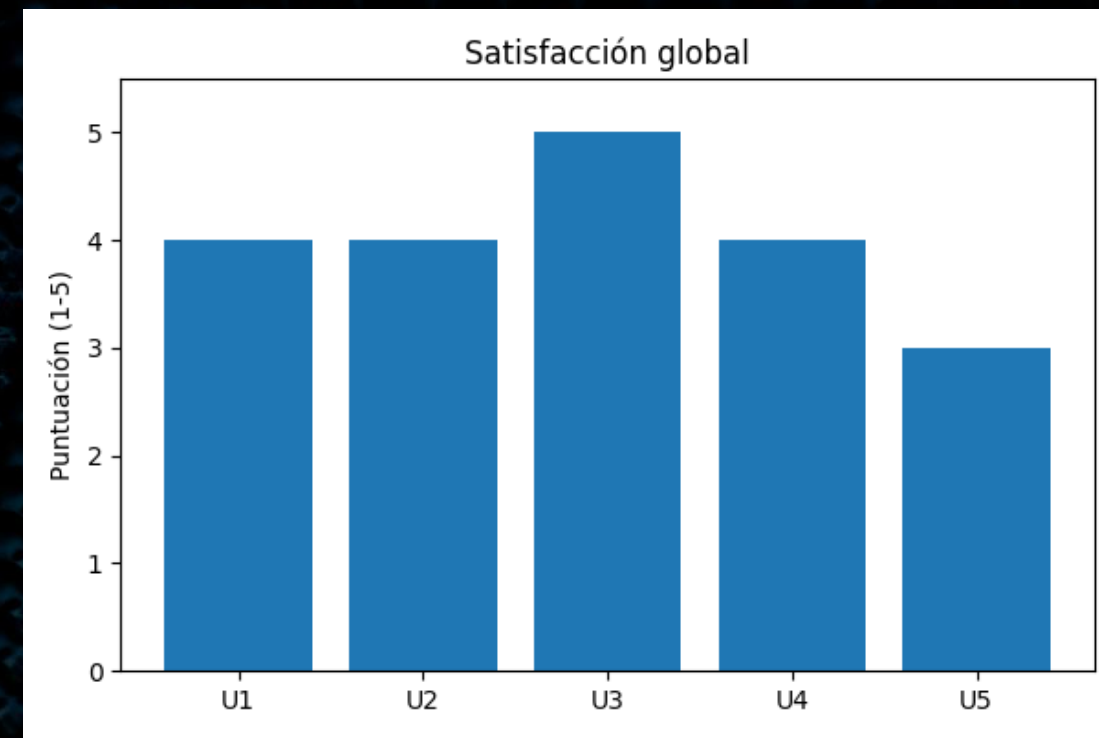
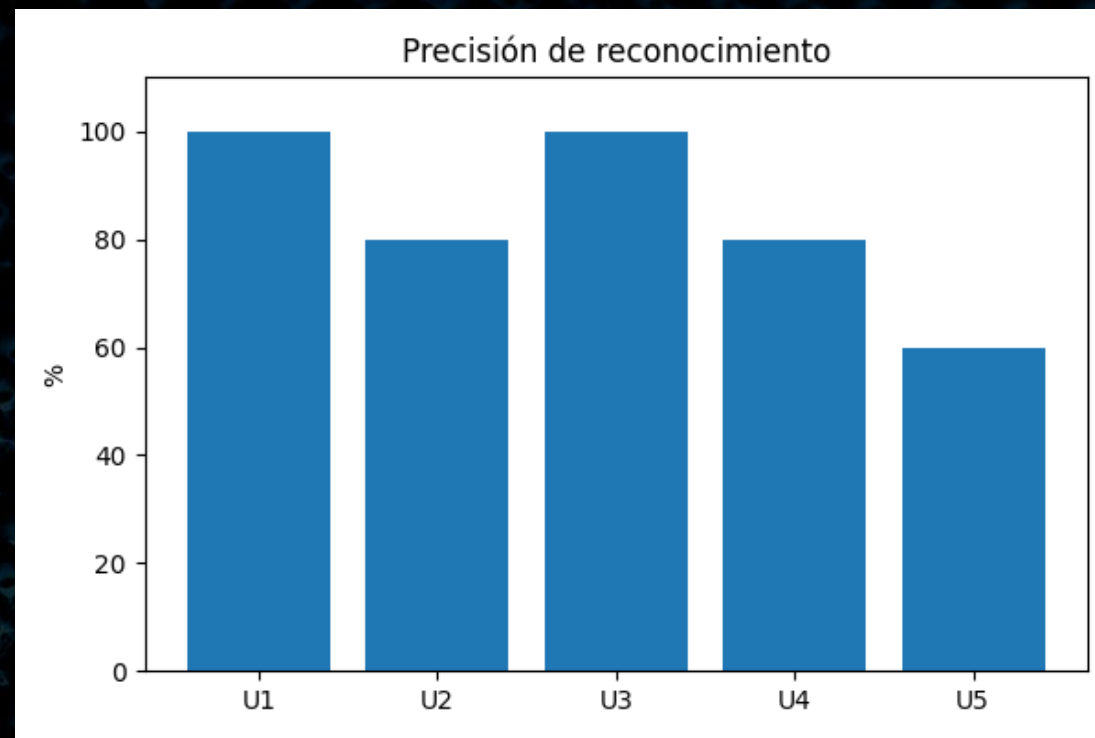
PUSH-TO-TALK

VOSK

5 TAREAS GUIADAS

USO LIBRE

# Experimentación



 84% de precisión media

 4,0/5 de satisfacción global

 Tiempo medio: 10,5 s



# Conclusiones y trabajos futuros

---

La voz es una interfaz viable para la creación de entornos XR

Los LLM permiten transformar lenguaje natural en escenas estructuradas

La arquitectura basada en componentes facilita la escalabilidad

La integración de IA y XR abre nuevas formas de interacción

---

Agentes inteligentes dentro de la escena

Interacción multimodal (voz, manos, mirada y gestos)

Generación automática de modelos 3D realistas

Entornos colaborativos multiusuario

“ La combinación de **Inteligencia Artificial Generativa** y **Realidad Extendida** abre nuevas formas de crear experiencias inmersivas mediante **lenguaje natural**, acercando el desarrollo de entornos 3D a cualquier usuario ”