

Corso di Laurea in Informatica

Progettazione e sviluppo di un assistente virtuale di cucina in realtà mista

Giacomo Paolocci

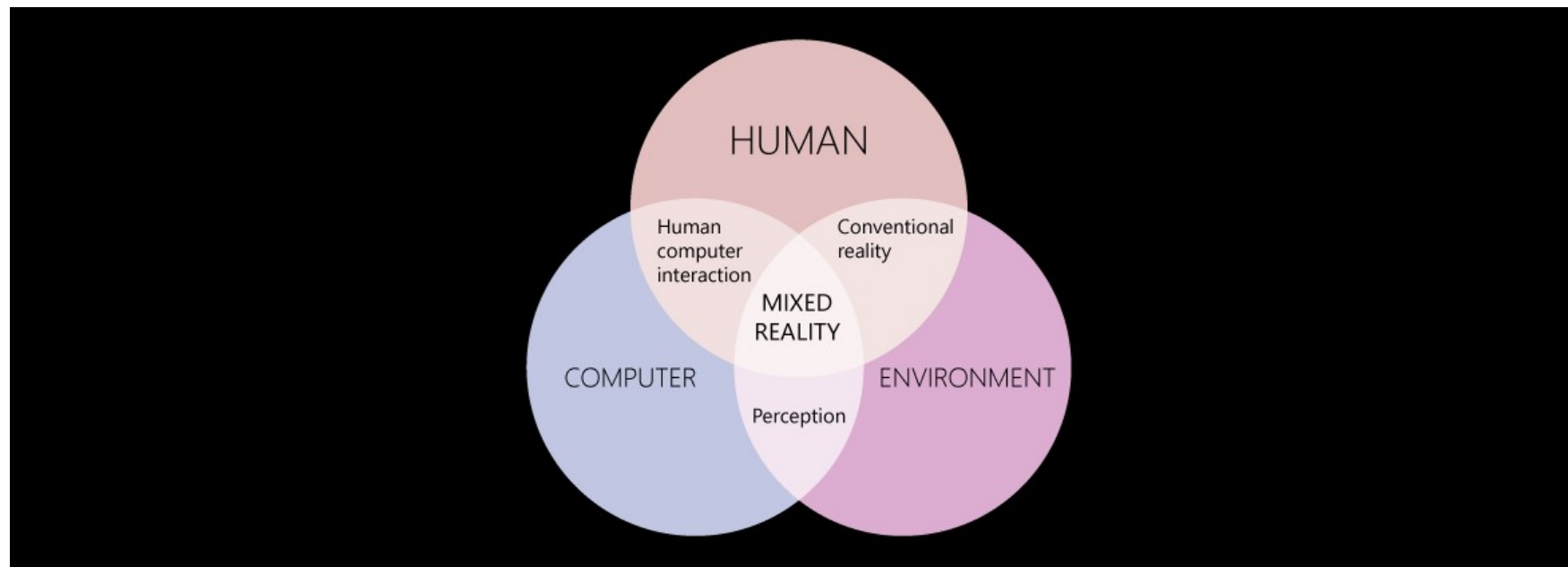
Relatore: Prof. Francesco Tarquini

Correlatore: Dr. Leonardo D'Errico

- **Tesi:** utilizzo della Realtà Mista in ambito culinario
- **Piattaforma:** visore **HoloLens 2**
- **Obiettivo:** creare ricette con ingredienti disponibili, riducendo lo spreco alimentare
- **Tecnologie:** Unity e Mixed Reality Toolkit (MRTK)
- **Integrazione con Google Gemini:** riconoscimento degli ingredienti e generazione delle ricette
- **Funzionalità:** avendo un ambiente virtuale si hanno le mani libere



- La realtà mista **sovrappone** al mondo fisico elementi tridimensionali con cui l'utente può interagire.
- Uso di un **dispositivo**, che tramite appositi sensori, riconosce e mappa tridimensionalmente l'ambiente circostante.
- Inserimento di oggetti tridimensionali, con i quali l'utente interagisce grazie alla comprensione dei movimenti di mani, occhi e degli input vocali.



› HoloLens 2

Visore per la realtà mista sviluppato da Microsoft



› Unity

motore grafico multiplatforma utilizzato per la creazione di videogiochi e applicazioni in 2D e 3D.



› MRTK

(Mixed Reality Toolkit) è un progetto Microsoft che fornisce strumenti e librerie per **sviluppare applicazioni di realtà mista**.

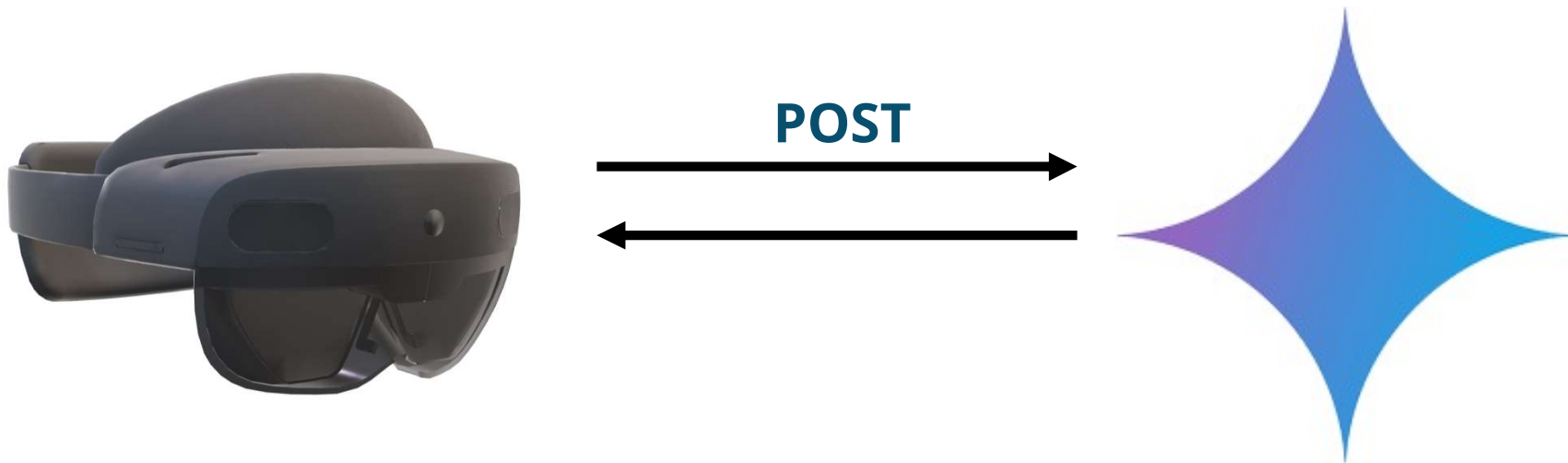


› Google Gemini

Google Gemini, è una **LLM** in grado di comprendere testo, immagini, audio e documenti.



- L'intera applicazione è perlopiù **standalone** sul visore, quindi non necessita di un server per il funzionamento.
- Le uniche operazioni e calcoli che avvengono al di fuori del visore sono quelle che coinvolgono **Google Gemini**, che avvengono grazie a delle chiamate API in **POST**



- La persistenza dei dati avviene tramite la creazione di due file **JSON** che contengono la lista degli ingredienti e le tre ricette generate
- La **foto** viene salvata nella stessa directory dei file JSON
- Viene utilizzato il **persistentDataPath**
- Il file di **Config** presente nella cartella **StreamingAssets** fornisce il nome dei files

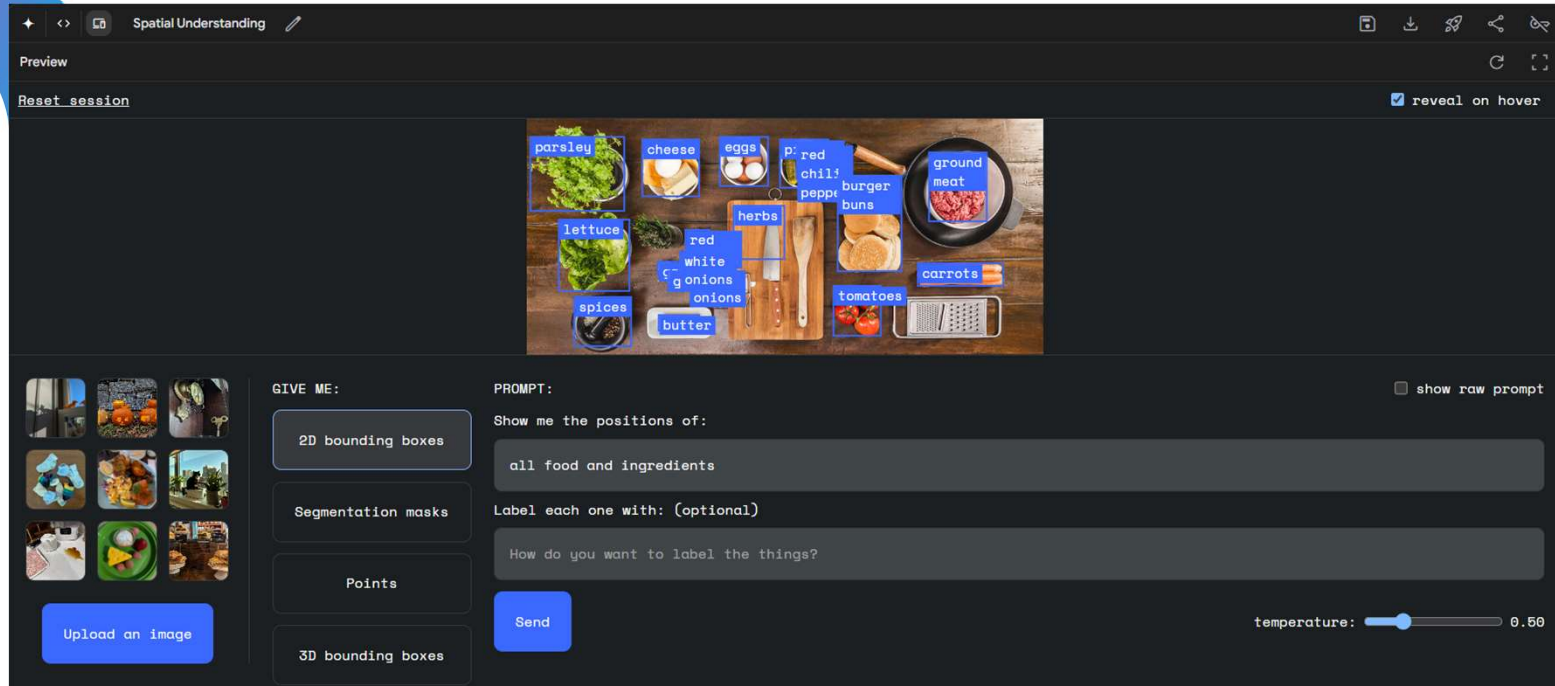
```
string filePath = Path.Combine(Application.persistentDataPath, jsonNameFile);
```

- Parte di codice che genera il path in cui salvare i file JSON e la foto

```
string json = File.ReadAllText( path: Application.streamingAssetsPath + "/config.json");
```

- Lettura JSON di config

- › Gemini è una LLM sviluppata da Google
- › Riconoscimento di **immagini** e documenti
- › Generazione di **testo** in base a determinati **prompt**
- › Funzionalità di **Spatial Understanding**
- › **API** gratuita con **rate limits** molto alti per modelli di base (Gemini 2.0 Flash-Lite)



› Schermata di Gemini Studio

Utilizzo di Gemini nell'applicazione

- Utilizzo dello **Structured output** per facilitare la lettura dei dati elaborati dalla LLM
- Due richieste POST: una per **generare gli ingredienti**, altra per le **ricette**
- La foto viene convertita in Base64 per essere inviata in POST

```
byte[] imageBytes = System.IO.File.ReadAllBytes(imagePath);  
string base64Image = System.Convert.ToBase64String(imageBytes);
```

- Conversione in Base64

```
var www = new UnityWebRequest(url, method: "POST");  
byte[] bodyRaw = Encoding.UTF8.GetBytes(body);  
www.uploadHandler = (UploadHandler)new UploadHandlerRaw(bodyRaw);  
www.downloadHandler = (DownloadHandler)new DownloadHandlerBuffer();  
www.SetRequestHeader(name: "Content-Type", "application/json");  
  
// Send the request  
yield return www.SendWebRequest();
```

- Richiesta POST

- Il menu è composto da tre bottoni
- Due dei quali aprono le diverse **finestre**, chiamate Plates,
- Il terzo, separato del resto, che permette di **fissare il menu nello spazio**
- La barra bianca in basso serve a **spostare il menu** nello spazio



Interfaccia grafica

Plate Fotocamera e Ingredienti

10

- In alto troviamo due bottoni che ci permettono di visualizzare la schermata per la fotocamera e per gli ingredienti
- Nella **schermata per la fotocamera** possiamo scattare una foto, eliminarla e visualizzarla e generare gli ingredienti con Gemini
- Nella **schermata degli ingredienti** possiamo visualizzarli, modificarli ed aggiungerli, inoltre si possono generare le tre ricette

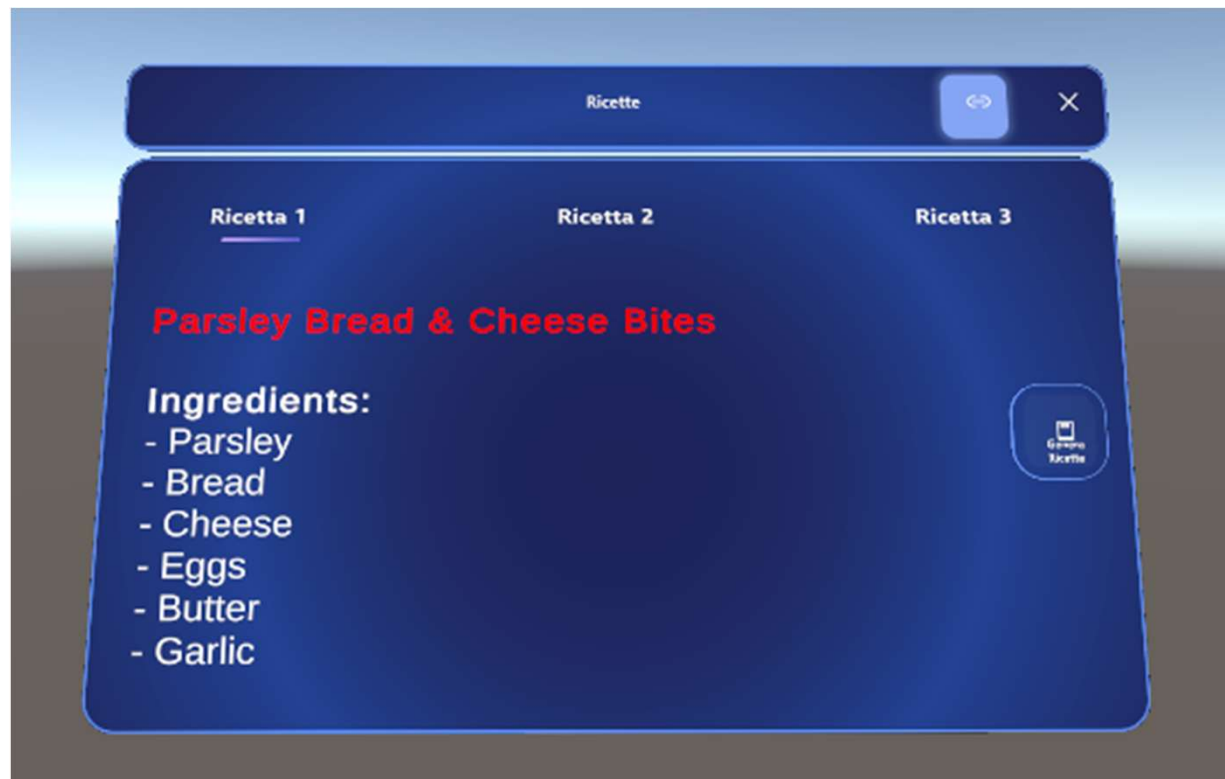
➤ Schermata per la fotocamera



➤ Schermata per gli ingredienti

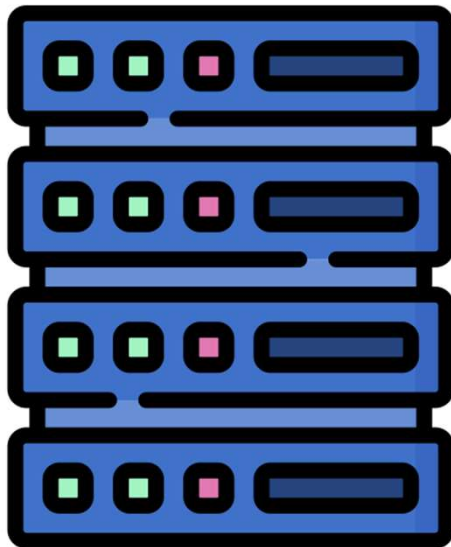


- Nella Plate delle ricette possiamo selezionare **tre possibili ricette**
- A destra è presente un bottone per **rigenerare le ricette** se quelle generate non sono di gradimento per l'utente





- **Tracking** degli ingredienti utilizzando le coordinate fornite da Gemini
- **Maggiori informazioni** visualizzabili sopra ogni singolo ingrediente
- Sistema **server** e **login** per l'utente per visualizzare e gestire da altri dispositivi ricette e ingredienti
- Sviluppo per **altri visori** che supportano la realtà mista e l'accesso alle fotocamere



- Esperienza con lo sviluppo di applicazioni XR in realtà mista
- Utilizzo di software per sviluppo di applicazioni 3D
- Studio e utilizzo del linguaggio C#

Problematiche riscontrate

- Scarsa documentazione
- Supporto inesistente
- Problematiche sull'SDK
- Processo di sviluppo lento

Corso di Laurea in Informatica

Grazie per l'attenzione