**Product Description**

This section presents TalkUtopia, our innovative approach to language learning that seamlessly blends cutting-edge technology with engaging educational principles. Designed with young second language learners in mind, our solution transforms what can often be a tedious learning process into an immersive, game-like experience that captivates and motivates. To this end, TalkUtopia comprises two parts, namely a mobile application and a VR display. The app section stores language learning related data, including learning routes and user settings, and allows users to start learning directly in the mobile app. In addition, VR display devices are optional, and users can choose to learn more engrossingly in a VR device.

**TalkUtopia App**

The mobile app has several subsystem, namely user login System, Learning Objective system, Exploring system and AI NPC system, and is expected to operate on both Android and iOS devices.

**User Login System**

The user data is stored in a robust cloud-based SQL database architecture that not only safeguards user data with industry-standard security protocols but also ensures reliable, scalable information management as our user base grows.

When users register, they are required to link an email address or phone number to facilitate password recovery in case they forget their password. Additionally, during login, users can choose to log in using a username and password or with a verification code.

**Learning Objective Selection**

The TalkUtopia app breaks free from rigid curriculum structures by offering two distinct yet equally valuable approaches to structuring your language learning journey:

Our carefully crafted predefined learning paths cover the essential scenarios of everyday life—from navigating a doctor's visit and ordering at restaurants to finding specific items while grocery shopping. What brings these scenarios to life are our AI agents, who don't merely recite scripted responses but authentically embody their roles with distinctive personalities, regional accents that reflect real-world language variation, and profession-specific vocabulary that you'd encounter in genuine interactions. This thoughtful design creates a truly immersive environment where users can practice natural communication patterns rather than following rigid, predetermined dialogue flows that rarely mirror real conversations.

For users who want to create specific scenarios, we allow them to create their own personalised learning objectives and define NPCs, and TalkUtopia has a community of creators where users can upload their own learning scenarios or download content uploaded by others. This greatly enriches the user experience and facilitates communication between users.

**Explore Utopia**

After selecting your learning objectives—whether from our curated collection or your personal creations—users will step into richly detailed scenarios specifically designed to practice targeted language skills in context. Recognizing that learners have different comfort levels and learning styles, our platform accommodates various preferences through flexible interaction options: choose between natural speech recognition for pronunciation practice or text input for those focusing on written communication.

To illustrate this experience, let's walk through our popular café scenario.

Upon entering, users will find themselves in a convincingly detailed café environment where ambient sounds of coffee grinding and quiet conversation create an authentic atmosphere. A barista, who is played by AI, stands ready to assist behind a counter. The scene includes thoughtfully designed interactive elements—you can browse through a visually appealing menu just as you would in a real café, pointing to items or flipping pages as needed.

Feel free to initiate conversation naturally—perhaps asking about specialty coffee options, inquiring about the pastry selection, or even making small talk about the weather. The AI barista responds not with rigid, pre-programmed phrases, but with natural conversation that adapts to your specific questions. During the ordering process, expect realistic follow-up questions about your preferences—"Would you like that to go?" or "How sweet would you prefer your drink?"—mirroring the back-and-forth rhythm of authentic service interactions.

These conversations encourage exploration beyond scripted exchanges. If you venture into unexpected territory—perhaps asking for a dish not on the menu like fried rice—the AI doesn't break character or malfunction but responds appropriately: "I'm sorry, we don't serve fried rice here, though the Chinese restaurant just down the street has excellent options if you're hungry for that." This natural give-and-take makes language practice feel less like studying and more like navigating real-world communication challenges.

When inevitable challenges arise—perhaps you don't understand a particular phrase or struggle to express your thoughts—comprehensive learning support appears. This includes context-specific tutorial videos demonstrating common interactions, the ability to record and instantly translate portions of your conversation, and gentle grammar assistance that identifies patterns without disrupting the flow of communication. These resources remain unobtrusive until needed, ensuring you never feel overwhelmed while maintaining the integrity of the immersive experience.

**AI NPC System**

The AI NPC consists of two main parts, one is the basic AI model selection and the other is the Mask Prompt setup.

**AI Model Selection**

We offer users two flexible approaches to AI model selection:

Platform Native Model: Users can opt for our platform-integrated Qwen Max 2.5 model. This model achieves an excellent balance between performance and cost, demonstrating outstanding results across multiple authoritative benchmark tests while maintaining a reasonable cost structure.

Custom API Integration: We also support users who wish to use their own API keys, compatible with leading large language models such as ChatGPT 4o-mini, Claude 3.5, and Deepseek V3. Users will need to independently obtain and manage API keys and associated costs through the respective service providers.

This dual-track integration mechanism ensures system flexibility, allowing users to navigate smoothly while also reducing the usage costs for those who already have AI accounts, thereby attracting more users.

**Mask Prompt setup**

At the heart of every AI based NPC is an innovative mask prompt system \cite{wang2025coser}. The mask prompt system is the architectural foundation of our AI agent behaviour framework. Below is an example of using mask prompts to create a barista NPC for a coffee shop, as shown in Figure 3.

**Mask Prompt Framework** consists of three core components: **Basic Prompt**, **Knowledge Base**, and **Interaction Flow**, which work cohesively to define digital interactions.

1. **Basic Prompt**  
   Establishes the agent’s foundational identity, including role (e.g., café server), setting (e.g., retro-style café), and personality traits (e.g., friendly tone).
2. **Knowledge Base**  
   Provides domain-specific expertise (e.g., detailed café menus or specialized content) to ensure accurate, context-aware responses.
3. **Interaction Flow**  
   Manages structured conversation paths (e.g., order: greeting → order confirmation → payment → follow up → farewell) to maintain focus and prevent deviation.

In addition, the Mask Prompt framework will be used for NPCs in user-created scenarios. Users will be asked to fill in each section according to the guidelines. User-created NPCs will be reviewed before they are uploaded to the community to ensure the legitimacy of the output.