



Universidad
Zaragoza

Trabajo Fin de Grado

Interacción conversacional y social persona-robot
Human-robot conversational and social interaction

Autor

Adrián Arribas Mateo

Directoras

Sandra Silvia Baldassarri

Eva Mónica Cerezo Bagdasari

ESCUELA DE INGENIERÍA Y ARQUITECTURA
2024



DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD

(Este documento debe acompañar al Trabajo Fin de Grado (TFG)/Trabajo Fin de Máster (TFM) cuando sea depositado para su evaluación).

TRABAJOS DE FIN DE GRADO / FIN DE MÁSTER

D./D^a. ADRIÁN ARRIBAS MATEO,

con nº de DNI 73159709K en aplicación de lo dispuesto en el art.

14 (Derechos de autor) del Acuerdo de 11 de septiembre de 2014, del Consejo de Gobierno, por el que se aprueba el Reglamento de los TFG y TFM de la Universidad de Zaragoza,

Declaro que el presente Trabajo de Fin de (Grado/Máster)
Grado en Ingeniería Informática, (Título del Trabajo)
Interacción conversacional y social persona-robot

es de mi autoría y es original, no habiéndose utilizado fuente sin ser citada debidamente.

Zaragoza, 5 de septiembre de 2024

Fdo: _____

AGRADECIMIENTOS

A mis tutoras, Eva y Sandra. Gracias por haberme dado esta oportunidad y por toda vuestra ayuda durante la elaboración de este trabajo.

A Juan. La persona que ha ocupado el asiento justo a mi lado antes, durante y espero que después de esta carrera universitaria. Ese asiento será siempre tuyo. Gracias por confiar en mí cuando ni siquiera yo lo hacía. Vas a conseguir todo lo que te propongas. Te admiro mucho.

A Claudia. Por tu ayuda durante todo este tiempo.

A Diego. Por todo el tiempo que hemos compartido.

A Sara. No sabes la felicidad que me produce tenerte a mi lado.

A Izarbe. Por todo lo que hemos pasado juntos.

A Carlos, Carlota, Irene, Jaime, Jorge, Marina y Raúl. Habéis sido la mayor sorpresa de mi carrera universitaria, me alegro mucho de que hayamos coincidido. Habéis sido una compañía esencial durante todo el tiempo dedicado a este trabajo.

A Esperanza, mi madre; Antonio, mi padre y Leyre, mi hermana. Gracias por haberme apoyado a lo largo de este camino, os quiero mucho.

A mi abuela Josefina, no sabes cuánto agradezco que puedas verme conseguir esto.

Al resto de mi familia. Mis tíos, Roberto y Tere, Mar y Luis y Yolanda y Alberto; mis primos, Alex y Sergio y sobre todo a mis abuelos Elena y Antonio, sé que allí dónde estéis, estáis muy orgullosos de mí.

A Nanouk, mi perra; que me ha acompañado durante todos esos días y noches de estudio.

Muchas gracias a todos los que han hecho posible la elaboración de este trabajo.

Interacción conversacional y social persona-robot

RESUMEN

La aparición y el uso de los robots es cada vez más común en nuestra vida cotidiana, pudiendo observarlos en sectores sanitarios, educativos o comerciales. Sin embargo, los estudios más recientes se centran en los robots sociales ya que cuentan con una apariencia y funcionalidades que permiten ser utilizados para mejorar la calidad de vida de las personas. Desde el proyecto de investigación nacional PLEISAR, coordinado por el grupo de investigación AffectiveLab, se propone utilizar agentes interactivos sociales, como los robots, para fomentar las relaciones e interacciones sociales con el objetivo de reducir la soledad en las personas y promover el bienestar social.

Dentro de ese contexto, en este trabajo de fin de grado, se ha desarrollado un módulo conversacional para el robot social *Sanbot Elf* utilizando herramientas de Inteligencia Artificial. Este módulo permite que las personas puedan interactuar con el robot para llevar a cabo una conversación de forma fluida y natural. Para ello, se han hecho uso principalmente de las funciones de reconocimiento de voz y funciones de síntesis de habla con las que cuenta el robot optimizadas por medio de la utilización de herramientas de Inteligencia Artificial como *OpenAI*. Adicionalmente, el robot reconoce emocionalmente el estado de ánimo del usuario llegando a empatizar y reconduciendo la conversación en función de los sentimientos interpretados. Se ha hecho especial hincapié en la accesibilidad del sistema, contando con funciones de narración de textos para personas con problemas visuales, la aparición de textos en pantalla para personas con problemas auditivos o la alternativa de utilizar el módulo conversacional mediante entrada de texto, para aquellas personas con problemas del habla. Además, el módulo cuenta con una interfaz de usuario sencilla tratando de evitar deslizables o gran cantidad de botones, para aquellas personas que no estén acostumbradas al uso de dispositivos electrónicos.

Para comprobar el correcto funcionamiento del módulo, se ha llevado a cabo una evaluación con usuarios para valorar la interacción con el robot en términos de usabilidad, aceptación y agrado. Los resultados han sido muy positivos en cuanto a la facilidad de uso, la valoración de la conversación con el robot y la utilidad del mismo.

Human-robot conversational and social interaction

ABSTRACT

The appearance and use of robots is increasingly more usual in our daily life and can be seen in sectors like sanitary, educationally or commercial. However, the latest investigations are focused on social robots due to its appearance and functions allow them to be used to improve people's quality of life. National project PLEISAR, coordinated by AffectiveLab research group, seeks to use social interactive agents, like robots, to promote social relationships and interactions with the goal of reducing people's loneliness and to promote social comfort.

Within this context, in this final degree project, a conversational module has been developed for the social robot *Sanbot Elf* using Artificial Intelligence tools. This module allows people to interact with the robot in order to have a conversation in a fluent and natural way. For this, robot's voice recognition and talk synthesis functions have been mainly used and optimized applying Artificial Intelligence tools as *OpenAI*. Additionally, the robot will recognize emotionally the user's mood achieving empathy and redirecting the conversation based on the feelings interpreted. System's accesibility has been taken into consideration, having text narration for people with visual disabilities, text display on screen for people with hearing disabilities or the alternative of using conversational module by text entrance for people with talk disabilities. Besides, the module has a simple user interface, avoiding sliding options or a lot of buttons for those people who are not used to manipulate electronic devices.

To verify the correct operation of the module, a user evaluation has been made to value the interaction on terms of acceptance, effectivity and likeability. The results have been really positive on ease of use, robot's conversation assessment and robot's utility.

Índice

1. Introducción y objetivo	1
1.1. Contexto	1
1.2. Objetivo	3
1.3. Metodología	3
1.4. Estructura de la memoria	4
2. Análisis	7
2.1. Descripción del problema	7
2.2. Robot Sanbot Elf	8
2.2.1. Funcionalidades del robot	8
2.2.2. Demostración preliminar	10
2.3. Personas	11
2.4. Historias de usuario	12
3. Diseño	15
3.1. Decisiones de diseño	15
3.1.1. Kit de Desarrollo Software (SDK)	15
3.1.2. Análisis de herramientas de Inteligencia Artificial	16

3.1.3.	Modelos de Inteligencia Artificial OpenAI	18
3.1.4.	Almacenamiento de los datos	20
3.2.	Arquitectura general del Sistema	20
4.	Implementación	23
4.1.	Iteración 1: Uso del endpoint Chat Completions	23
4.1.1.	Historias de usuario cumplidas	24
4.1.2.	Ventajas y desventajas	25
4.1.3.	Conclusión	25
4.2.	Iteración 2: Uso del endpoint Audio Speech	26
4.2.1.	Historias de usuario cumplidas	27
4.2.2.	Ventajas y desventajas	27
4.2.3.	Conclusión	27
4.3.	Iteración 3: Incorporación del aspecto emocional	28
4.3.1.	Historias de usuario cumplidas	29
4.3.2.	Ventajas y desventajas	30
4.3.3.	Conclusión	31
4.4.	Iteración 4: Personalización del módulo	31
4.4.1.	Historias de usuario cumplidas	32
4.4.2.	Ventajas y desventajas	32
4.4.3.	Conclusión	33
4.5.	Iteración 5: Configuración del módulo	33
4.5.1.	Historias de usuario cumplidas	35

4.5.2.	Ventajas y desventajas	35
4.5.3.	Conclusión	35
4.6.	Iteración 6: Otras funcionalidades	36
4.6.1.	Historias de usuario cumplidas	37
4.6.2.	Ventajas y desventajas	38
4.6.3.	Conclusión	38
4.7.	Conclusiones generales	38
5.	Evaluación final con usuarios	39
5.1.	Técnicas de evaluación	39
5.2.	Usuarios y sesiones	40
5.3.	Resultados	42
5.3.1.	Preferencia de voces	42
5.3.2.	Usabilidad (Preguntas SUS)	43
5.3.3.	Usabilidad (Preguntas interacción voz)	45
5.3.4.	Likeability (Agrado)	47
5.3.5.	Aceptación	49
5.4.	Conclusiones	50
6.	Gestión del proyecto	53
6.1.	Desarrollo por Sprints	53
6.2.	Dedicación	55
7.	Conclusiones y trabajo futuro	57
7.1.	Conclusiones	57

7.2. Trabajo futuro	58
7.3. Valoración personal	59
Bibliografía	61
Lista de Figuras	63
Lista de Tablas	67
Anexos	68
A. Análisis de las funcionalidades del robot Sanbot Elf	71
A.1. Familiarización del robot	71
A.2. Primera prueba módulo conversacional	73
B. Módulos QihanOpenSDK	77
C. Explicación de los módulos del Módulo Conversacional	81
D. Endpoints de OpenAI	93
D.1. Chat Completions	93
D.2. Audio Speech	95
E. Rueda de emociones de Plutchik	97
F. Documentos utilizados en las pruebas con usuarios	101

Capítulo 1

Introducción y objetivo

En este primer capítulo se explica el contexto relacionado con los robots y las dificultades que existen actualmente al interactuar con ellos, el objetivo principal del trabajo, la metodología utilizada y la estructura de la memoria.

1.1. Contexto

Los robots originalmente tenían el simple propósito de completar tareas, principalmente de carácter industrial, pero con los avances tecnológicos se consiguió que se introdujeran en otros campos como pueden ser médicos, domésticos, militares o educativos.

Actualmente, el foco de la investigación se encuentra en los robots sociales, donde la comunicación tanto verbal como no verbal y el uso de nuevos recursos como las herramientas de Inteligencia Artificial permiten explorar nuevos aspectos como puede ser la naturalidad, la confianza o la empatía en la interacción humano-robot.

El manejo de los robots ha ido evolucionando a lo largo de los años, desde el control a través de botones o *joysticks* hasta el uso de comandos por voz para enviar tareas a las máquinas. Este último método se ha ido desarrollando hasta poder interactuar verbalmente mediante el lenguaje natural, siendo ahora una de las técnicas más utilizadas ya que proporciona naturalidad y eficiencia, además de permitir que la máquina sea controlada por humanos no expertos en la materia.

Sin embargo, la comunicación entre ambos presenta retos como pueden ser que el robot no sea capaz de responder acerca de temas que el usuario quiera hablar, la dificultad en interpretar emociones y empatizar con el usuario o la posibilidad de personalizar y adaptar

la interacción en función del perfil y preferencias de la persona, siendo las herramientas de Inteligencia Artificial una posible solución que facilite la resolución de algunos de los problemas propuestos.

Otros obstáculos a mencionar son el coloquialismo, las ambigüedades, las conversaciones dependientes del contexto o recursos como preguntas retóricas o sarcasmo que el robot es incapaz de identificar o gestionar. Pudiendo incluir también la dificultad de capturar el sonido en entornos ruidosos o cuando el emisor se encuentra a largas distancias, de identificar el usuario con el que está manteniendo la conversación o de gestionar las interrupciones o pausas durante el diálogo [1], dichos retos han quedado fuera del ámbito del trabajo.

Para el desarrollo de este trabajo se hace uso de *Sanbot*, un robot de servicio inteligente que fue desarrollado por la empresa *Qihan Technology* y presentado en el año 2016. La compañía ha desarrollado múltiples versiones del robot, entre ellas: *Sanbot Nano*, *Sanbot Max* o *Sanbot Elf*, siendo este último el prototipo utilizado como caso de estudio [2].

Sanbot Elf (Figura 1.1) fue diseñado con el propósito principal de tomar un papel importante en sectores como el comercio, la educación y el sanitario, aunque cuenta con una *Application Programming Interface* (API) de código abierto permitiendo que los desarrolladores puedan crear todo tipo de aplicaciones haciendo uso de las funciones del robot. Este robot está siendo utilizado dentro del grupo de investigación *AffectiveLab* como parte del proyecto nacional de investigación “Experiencias Lúdicas con Agentes Sociales Interactivos y Robots” (**PLEISAR** (PID2022-136779OB)) donde también participan grupos de las universidades de las Islas Baleares, de Granada y de La Laguna. Este proyecto propone utilizar agentes interactivos sociales (SIA), como los robots, con el propósito de fomentar las relaciones e interacciones sociales para reducir la soledad en las personas y promover el bienestar social [3]. Es en este proyecto en el que se enmarca este trabajo fin de grado.



Figura 1.1: Robot modelo *Sanbot Elf* diseñado por *Qihan Technology*

1.2. Objetivo

El objetivo principal de este trabajo es desarrollar un módulo conversacional sobre la plataforma del robot social *Sanbot Elf* a través de su kit de desarrollo software (SDK) que permita explorar la interacción conversacional y social entre persona y robot.

Para conseguir este propósito, se seguirán los siguientes pasos:

1. Estudiar las funcionalidades básicas del robot, focalizándose en su capacidad conversacional, expresividad facial y las características relacionadas con la interpretación y transcripción de la voz humana.
2. Investigar las dificultades actuales que existen en la interacción humano-robot, para tener en cuenta los límites y las posibles soluciones a la hora de desarrollar el módulo.
3. Estudiar la posible integración de herramientas de Inteligencia Artificial en el módulo conversacional para aumentar la naturalidad.
4. Evaluar los resultados conseguidos a través de pruebas con usuarios con el fin de valorar la interacción en términos de aceptación, efectividad y agrado (*likeability*).

1.3. Metodología

Para el desarrollo de este trabajo se han hecho uso de metodologías ágiles, considerándolas una mejor opción por el peso que toman los usuarios, por las pequeñas iteraciones que permiten recibir *feedback* por parte de las tutoras o de los usuarios mediante correcciones y evaluaciones intermedias, y por la capacidad de reconocer los errores a lo largo del trabajo y no en fases finales de desarrollo como se produce en metodologías tradicionales.

Más concretamente, las técnicas utilizadas durante el trabajo que forman parte de las metodologías ágiles son las siguientes:

1. **Método Personas:** Consiste en describir y modelar los perfiles de los usuarios que podrían acabar usando el sistema desarrollado, considerando sus necesidades y los obstáculos que podrían encontrar al utilizarlo.
2. **Historias de usuario:** Se trata de explicaciones generales e informales de una característica del *software* escrita desde la perspectiva del usuario final. Su propósito es articular cómo una característica del *software* proporciona valor al usuario.

3. **Desarrollo mediante Sprints:** Al ser un sistema totalmente inexplorado del que se desconocían sus límites, se consideró el utilizar un desarrollo basado en *Sprints* dónde, una vez acabadas un conjunto de tareas, se realizaba una reunión con las tutoras para evaluar el trabajo realizado.

1.4. Estructura de la memoria

La memoria se ha organizado mediante 7 capítulos y 6 anexos, los cuales se describen a continuación:

1. **Capítulo 1, Introducción:** Apartado en el que se detalla el contexto del trabajo, el objetivo que se plantea conseguir, la metodología usada y la estructura de la memoria.
2. **Capítulo 2, Análisis:** Apartado en el que se analiza el problema a abordar en este trabajo, se estudian las características y funcionalidades del robot *Sanbot Elf*, se definen los perfiles de los usuarios finales utilizando el método *Personas* y se concretan las historias de usuario como funcionalidades que tendrá el sistema a desarrollar.
3. **Capítulo 3, Diseño:** Apartado en el que se detallan las decisiones de diseño que se tomaron a lo largo del trabajo y se presenta la arquitectura general del sistema.
4. **Capítulo 4, Implementación:** Apartado en el que se explica las diferentes iteraciones que se fueron desarrollando de manera incremental, su funcionamiento, las ventajas y desventajas que tenían y las historias de usuario que cumplían.
5. **Capítulo 5, Evaluación final con usuarios:** Apartado en el que se detalla la metodología y los resultados obtenidos en las pruebas que se realizaron con usuarios.
6. **Capítulo 6, Gestión del proyecto:** Apartado en el que se detalla el desarrollo del trabajo y las horas de dedicación que se han dedicado a cada fase.
7. **Capítulo 7, Conclusiones y trabajo futuro:** Apartado en el que se detalla las conclusiones del trabajo y las posibles mejoras que se podrían añadir en el futuro.
8. **Anexo A, Análisis de las funcionalidades del robot Sanbot Elf:** Anexo en el que se habla acerca de las funcionalidades del robot y se dan detalles sobre una demostración desarrollada para descubrir sus límites.

9. **Anexo B, Explicación de los módulos del QihanOpenSDK:** Anexo en el que se detallan los diferentes módulos y funciones del SDK que han sido utilizados en el trabajo.
10. **Anexo C, Explicación de los módulos del Módulo Conversacional:** Anexo en el que se detallan los diferentes módulos y funciones que componen el módulo conversacional y que han sido utilizados en el trabajo.
11. **Anexo D, Endpoints de OpenAI:** Anexo en el que se detalla el funcionamiento de los *endpoints Chat Completions* y *Audio Speech* de *OpenAI*.
12. **Anexo E, Rueda de emociones de Plutchik:** Anexo en el que se presenta el funcionamiento del modelo de rueda de emociones de Plutchik.
13. **Anexo F, Documentos utilizados en las pruebas con usuarios:** Anexo en el que se incluyen el consentimiento informado, el cuestionario previo, el test POST VAVA Q, el Reysen Likeability Test y el cuestionario de usabilidad.

Capítulo 2

Análisis

En este capítulo se analizan los principales aspectos a tener en cuenta en la interacción entre humanos y robots, se presenta al robot *Sanbot Elf* utilizado como caso de estudio junto a sus funcionalidades y la demo desarrollada, se definen los perfiles de usuario aplicando el método *Personas* y se concretan las historias de usuario con las características con las que deberá contar el módulo a desarrollar.

2.1. Descripción del problema

En una conversación entre personas, ambos participantes hablan acerca de un mismo tema y el diálogo entre ellos debe estar equilibrado para evitar un monólogo.

De hecho, la capacidad de interactuar en ambos sentidos es una de las características que las personas consideran más relevantes cuando se habla de la interacción entre humanos y robots sociales, generando decepción en los casos en los que el robot no proporciona siempre una contestación al usuario [4]. Aspectos como una apariencia antropomórfica del robot, su presencia física y un diálogo coherente tienen impacto en la comodidad y confianza experimentada por el usuario al interactuar con el robot [5].

Además, las propias características de la voz del robot pueden tener una poderosa influencia en las percepciones y preferencias de la persona que interactúa con él. Por ejemplo, diversas pruebas indican que las personas se muestran más participativas conversando con voces cuyo género coincide con el suyo [6].

En una conversación, el oyente suele participar de manera no verbal en la charla con señales como el contacto visual, gestos o similares para indicar que realmente está prestando atención al emisor. De esta forma, la interacción tanto verbal como no verbal proporciona naturalidad en la comunicación humano-humano [7].

Por ello se deben incorporar técnicas multimodales (a través de varios canales de comunicación): el robot no sólo tiene que poder mantener una conversación sino que también ha de utilizar comunicación no verbal, como la proximidad, el contacto visual o la expresión facial [8].

Por otra parte, la empatía es un aspecto muy importante en la interacción humano-robot que ofrece comodidad entre ambos. El estudio e incorporación del apartado emocional es fundamental para conseguir ese objetivo ya que los robots con capacidad de representar emociones a través de expresión facial o mediante gestos son recibidos de una manera más familiar y cercana [9]. La capacidad de expresividad del robot es otro factor que facilita a la persona el entendimiento del comportamiento, del estado de ánimo, de la motivación y la racionalización del robot y además incrementa la percepción de este como alguien fiel y comprensivo [10].

Debido al clima competitivo actual de las empresas, los modelos de negocio tratan de innovar y mejorar las experiencias del cliente, dando especial importancia a la personalización a través de Inteligencia Artificial, robots conversacionales y asistentes virtuales, siendo métodos que afectan a la experiencia de usuario, buscando que esta sea más memorable e individualizada [11].

2.2. Robot Sanbot Elf

En este apartado se describen, en primer lugar, las funcionalidades con las que cuenta el robot *Sanbot*, para luego presentar una demostración con el propósito de comprobar su funcionamiento.

2.2.1. Funcionalidades del robot

Ya se ha comentado en apartados anteriores que el robot *Sanbot Elf* es el modelo a utilizar como caso de estudio, por lo que, se deben analizar sus capacidades y sus posibles limitaciones para definir las funcionalidades con las que contará el módulo conversacional.

Entre sus características, se puede destacar que cuenta con 90.2cm de altura, 42.1cm de anchura y 33.1cm de profundidad, además de tener un peso de 19kg. Su cabeza y brazos son móviles, además de contar con una base motorizada y un sistema de ruedas omnidireccionales que permiten el movimiento del robot de manera inteligente ya que puede evitar el choque con obstáculos gracias a sensores de presencia. La mayor parte de la interacción con el usuario se realiza a través de una pantalla de 10.1 pulgadas situada en su pecho (ver Figura 1.1), pero también existen otras alternativas como puede ser los múltiples sensores de contacto.

Sanbot Elf cuenta con diversos módulos que permiten al robot llevar a cabo acciones, y que se han analizado para su uso en el módulo conversacional:

1. **Control del habla:** Permite realizar funciones *Text to Speech* (TTS), es decir, reproduce un texto en forma de audio con la posibilidad de modificar parámetros como la velocidad o la entonación y funciones *Speech to Text* (STT), donde el robot reconoce el habla del usuario y la transforma en texto, permitiendo de esta manera interpretar el lenguaje natural.
 - **Utilidad para el módulo conversacional:** Estas características son interesantes en este trabajo para interpretar lo que dice el usuario y utilizar las funciones de habla del robot.
2. **Control del hardware:** Permite encender, apagar y cambiar de color los LED del robot localizados en la cabeza, brazos y base, permite gestionar los sensores táctiles que se encuentran alrededor del robot y permite utilizar una función con la que el robot reconoce la procedencia del sonido.
 - **Utilidad para el módulo conversacional:** Estas características son interesantes en este trabajo para reconocer la procedencia del sonido o para encender y apagar los LED en las distintas partes del cuerpo del robot, aunque únicamente ofrece los siguientes colores: verde, azul, rosa, morado, rojo, blanco y amarillo.
3. **Control del movimiento:** Permite controlar el movimiento de la cabeza, de los brazos y de las ruedas del robot.
 - **Utilidad para el módulo conversacional:** Estas características son interesantes en este trabajo para la comunicación no verbal del robot a través del movimiento de cabeza y brazos, permitiendo simular mediante gestos la expresión de diversas emociones que interprete durante la conversación.

4. **Control del sistema:** Permite controlar funciones relacionadas con información del robot como la obtención de su número de identificación o datos sobre la batería y también permite gestionar la expresión facial del robot formando un conjunto de 18 emociones representadas por medio de los LED que tiene como ojos (Figura 2.1).

– **Utilidad para el módulo conversacional:** Estas características son interesantes en este trabajo para reflejar sentimientos a través del cambio de la expresión facial.



Figura 2.1: Expresiones faciales del robot *Sanbot Elf*

5. **Control del audio:** Permite conocer y modificar múltiples parámetros respecto al audio del robot, como el volumen al que están los altavoces o saber si está el micrófono encendido.

– **Utilidad para el módulo conversacional:** Estas características son interesantes en este trabajo para configurar el volumen de salida de los altavoces.

6. **Control del proyector:** Permite realizar acciones con el dispositivo proyector con el que cuenta el robot, entre ellas encender, apagar, modificar los ajustes de imagen o incluso cambiar entre la proyección de pared o de techo.

7. **Gestor multimedia:** Permite trabajar con los dispositivos como la cámara de alta definición o el control de flujos de vídeo y audio. También cuenta con funciones relacionadas con el reconocimiento facial.

2.2.2. Demostración preliminar

Tras el estudio de la documentación de *Sanbot Elf*, lectura del manual de usuario y búsqueda a través de foros, se desarrolló un pequeño programa a modo de demo partiendo

del repositorio de *GitHub*¹ dónde se testeaban las funcionalidades del robot. Este programa fue presentado a un grupo de usuarios durante las primeras semanas de trabajo con el objetivo de comprobar el funcionamiento de los módulos del robot y con el propósito de encontrar posibles limitaciones de cara al desarrollo del módulo conversacional, todo el proceso ha sido documentado en el Anexo A.

Poniendo el foco en las funcionalidades del habla con las que cuenta el robot, la prueba permitió observar que la limitación más importante a destacar fue la siguiente: las conversaciones deben ser **predefinidas por completo** mediante la asignación de una respuesta tras reconocer un determinado patrón en las palabras del usuario.

Esto claramente era un gran obstáculo a la hora de desarrollar un módulo conversacional, por lo que la incorporación de herramientas de **Inteligencia Artificial** era de gran importancia para conseguir una interacción natural entre el humano y el robot.

2.3. Personas

Una parte esencial de la fase de análisis del sistema es el estudio de sus potenciales usuarios. En este trabajo, estos usuarios son los considerados en el proyecto de investigación en el que se enmarca, orientado a romper el aislamiento social de personas mayores, adolescentes y grupos intergeneracionales. Por ello se han definido las tres personas que se presentan a continuación:



Antonio: Es un hombre viudo de 75 años, que actualmente vive sólo en su propia casa en el centro de la ciudad. Ocasionalmente recibe visitas de sus dos hijos y sus dos nietos, pero la mayor parte del tiempo lo pasa sólo, exceptuando algunas tardes que va con un par de amigos a tomar un café y jugar a las cartas.

Necesidades: Antonio estaría encantado de contar con *Sanbot* y su módulo conversacional para sentirse más acompañado en las horas de soledad.

Obstáculos: Hay que tener en cuenta la avanzada edad de Antonio y su poca familiaridad con las nuevas tecnologías, seguramente el robot le genere rechazo a la hora de interactuar con él.

¹<https://github.com/igor-lirussi/CapaBot-SanBot-Robot/tree/master>



Patricia: Es una mujer desempleada de 55 años sin hijos y con una pareja que pasa gran parte del tiempo en su trabajo. Patricia disfruta mucho su día a día realizando distintas actividades en solitario como salir a pasear o ir al cine. Le llama la atención todo lo conseguido gracias a la tecnología pero le genera algo de rechazo el uso de dispositivos como puede ser el móvil o el ordenador. Aunque mantiene alguna que otra conversación por aplicaciones de mensajería como *WhatsApp* o *Gmail*, evita utilizarlas siempre que puede. Por lo tanto, su experiencia tecnológica es muy reducida.

Necesidades: Patricia podría utilizar el robot *Sanbot* para romper un poco su rutina y mantener conversaciones con alguien hasta que su pareja vuelva a casa.

Obstáculos: Se debe tener en cuenta la poca familiarización o rechazo que le puede generar un dispositivo como *Sanbot*.



Alberto: Es un niño de 14 años que tiene dificultades para socializar con el resto de niños. Debido a malas experiencias en su antiguo instituto, sus padres decidieron cambiarlo de centro pero tienen miedo de que ocurra algo parecido.

Necesidades: Los padres de Alberto quieren que el niño aprenda a socializar con un agente social como es *Sanbot*. Un robot llamará mucho su atención y le animará a mantener conversaciones sobre múltiples posibles temas.

Obstáculos: Alberto, debido a su edad está muy familiarizado con la tecnología, aunque espera que el robot sea consciente de su edad y utilice vocabulario apropiado y no demasiado técnico.

2.4. Historias de usuario

A partir del análisis del problema y las personas definidas en el apartado anterior se presentan a continuación las funcionalidades del sistema en la forma de historias de usuario (Sección 1.3).

En la tabla 2.1 para cada uno de los usuarios se indica quien o quienes ven útil esa funcionalidad, qué necesidad tienen y el motivo.

Nº	Historia de usuario
1	<p>Como Antonio, Patricia y Alberto</p> <p>Quiero poder hablar al robot en lenguaje natural</p> <p>Para que pueda darme una respuesta</p>
2	<p>Como Antonio, Patricia y Alberto</p> <p>Quiero que el robot responda mi petición de forma adecuada</p> <p>Para considerar que el robot ha entendido mi petición y así poder mantener una conversación</p>
3	<p>Como Antonio, Patricia y Alberto</p> <p>Quiero comunicarme con el robot por medio de la voz</p> <p>Para que la interacción sea más cómoda y natural</p>
4	<p>Como Antonio, Patricia y Alberto</p> <p>Quiero que el robot me responda por medio de la voz</p> <p>Para que la interacción sea más cómoda y natural</p>
5	<p>Como Antonio, Patricia y Alberto</p> <p>Quiero poder hablar con el robot de cualquier tema</p> <p>Para mantener conversaciones que no sean limitadas o cerradas</p>
6	<p>Como Antonio, Patricia y Alberto</p> <p>Quiero que el robot guarde el hilo de la conversación</p> <p>Para no tener que indicar de forma específica acerca de lo que estamos hablando</p>
7	<p>Como Antonio, Patricia y Alberto</p> <p>Quiero visualizar las palabras que el robot me ha entendido</p> <p>Para poder corregirlo o repetirlo en caso de que no me haya entendido bien</p>
8	<p>Como Antonio</p> <p>Quiero observar tanto mis respuestas como las respuestas del robot en la pantalla</p> <p>Para observar la conversación en caso de tener problemas de audición</p>
9	<p>Como Antonio, Patricia y Alberto</p> <p>Quiero que el robot no dé respuestas muy extensas</p> <p>Para evitar un monólogo y que la interacción pierda naturalidad</p>
10	<p>Como Antonio, Patricia y Alberto</p> <p>Quiero que las respuestas sean reproducidas por voces no robóticas o artificiales</p> <p>Para que la conversación no me genere rechazo</p>
11	<p>Como Antonio, Patricia y Alberto</p> <p>Quiero que el robot reconozca mi estado de ánimo al conversar con él</p> <p>Para que empatice conmigo y adapte la conversación en base a él</p>
12	<p>Como Antonio, Patricia y Alberto</p> <p>Quiero que el robot refleje en su expresión facial su estado de ánimo</p> <p>Para poder interpretar mejor lo que el robot trata de decir</p>
13	<p>Como Antonio, Patricia y Alberto</p> <p>Quiero que el robot realice gestos al expresar ciertas emociones</p> <p>Para tener una conversación más agradable con el robot</p>
14	<p>Como Antonio, Patricia y Alberto</p> <p>Quiero que el robot conozca datos personales como mi nombre y edad</p> <p>Para que adapte la conversación en base a estos datos</p>

15	<p>Como Antonio, Patricia y Alberto</p> <p>Quiero poder seleccionar la voz que quiero que reproduzca las respuestas</p> <p>Para elegir aquella con la que me sienta más cómodo</p>
16	<p>Como Antonio, Patricia y Alberto</p> <p>Quiero poder personalizar características del robot como su edad o género</p> <p>Para mantener conversaciones en las que me sienta más cómodo</p>
17	<p>Como Patricia</p> <p>Quiero poder simular situaciones ficticias con el robot</p> <p>Para pasar un rato divertido sin hablar siempre de lo mismo</p>
18	<p>Como Antonio, Patricia y Alberto</p> <p>Quiero poder modificar el volumen de los altavoces del robot</p> <p>Para que se adapte al entorno en el que estoy (ruidoso, silencioso, etc)</p>
19	<p>Como Antonio, Patricia y Alberto</p> <p>Quiero poder configurar la velocidad y entonación de las voces</p> <p>Para elegir aquella que entienda mejor</p>
20	<p>Como Antonio, Patricia y Alberto</p> <p>Quiero enviar las consultas por teclado en vez de vía voz</p> <p>Para poder conversar estando en un entorno en el que no se pueda hablar</p>
21	<p>Como Antonio, Patricia y Alberto</p> <p>Quiero que la conversación no se interrumpa</p> <p>Para que la conversación sea más natural</p>
22	<p>Como Antonio y Patricia</p> <p>Quiero que el módulo cuente con un tutorial</p> <p>Para conocer sus funcionalidades y cómo se utiliza y poder verlo cuando lo necesite</p>
23	<p>Como Antonio</p> <p>Quiero repetir la última respuesta del robot</p> <p>Para escuchar de nuevo su respuesta en caso de no haberla escuchado</p>
24	<p>Como Antonio, Patricia y Alberto</p> <p>Quiero poder detener la respuesta del robot</p> <p>Para que deje de hablar en caso de estar dando una respuesta muy extensa</p>
25	<p>Como Antonio</p> <p>Quiero que la conversación aparezca en pantalla</p> <p>Para poder recordar de qué estábamos hablando si lo olvido</p>
26	<p>Como Antonio, Patricia y Alberto</p> <p>Quiero poder cambiar de conversación</p> <p>Para hablar de un tema distinto</p>
27	<p>Como Antonio, Patricia y Alberto</p> <p>Quiero poder realizar consultas de actualidad, como el tiempo meteorológico o noticias</p> <p>Para estar informado</p>

Tabla 2.1: Requisitos en forma de historias de usuario para la elaboración del módulo conversacional

Capítulo 3

Diseño

En este capítulo se comentan las decisiones de diseño que fueron surgiendo a lo largo del desarrollo del trabajo junto a la explicación de la arquitectura general del sistema.

3.1. Decisiones de diseño

Para este trabajo se han tomado decisiones relacionadas con el kit de desarrollo software utilizado, las posibles herramientas de Inteligencia Artificial y el almacenamiento de datos en el sistema.

3.1.1. Kit de Desarrollo Software (SDK)

El **Kit de Desarrollo Software** permite utilizar las funciones básicas del robot dentro del entorno de desarrollo. El SDK encontrado en la página oficial era el *SanbotOpenSDK*, al parecer, el más reciente y actualizado y el cual se utilizó para la familiarización con el robot *Sanbot Elf* durante las primeras semanas del trabajo. Sin embargo, con la intención de hacer uso de herramientas de Inteligencia Artificial, se quiso utilizar la API de *OkHttp*¹, la cual permite realizar peticiones HTTP de manera sencilla. No obstante, el programa no consiguió completar la compilación debido a un error de clases duplicadas (Figura 3.1). Tras múltiples intentos por corregirlo y consultando información por internet, se observó que la mayoría de los repositorios que trabajaban con el robot *Sanbot*, utilizaban un SDK diferente, el *QihanOpenSDK*, el cual parecía más antiguo y desactualizado. Este kit de desarrollo sí que permitió completar la compilación a diferencia del SDK anterior, por lo que se optó por sustituir el *SanbotOpenSDK* (con el que se había trabajado hasta ahora) por el *QihanOpenSDK*.

¹<https://square.github.io/okhttp/>

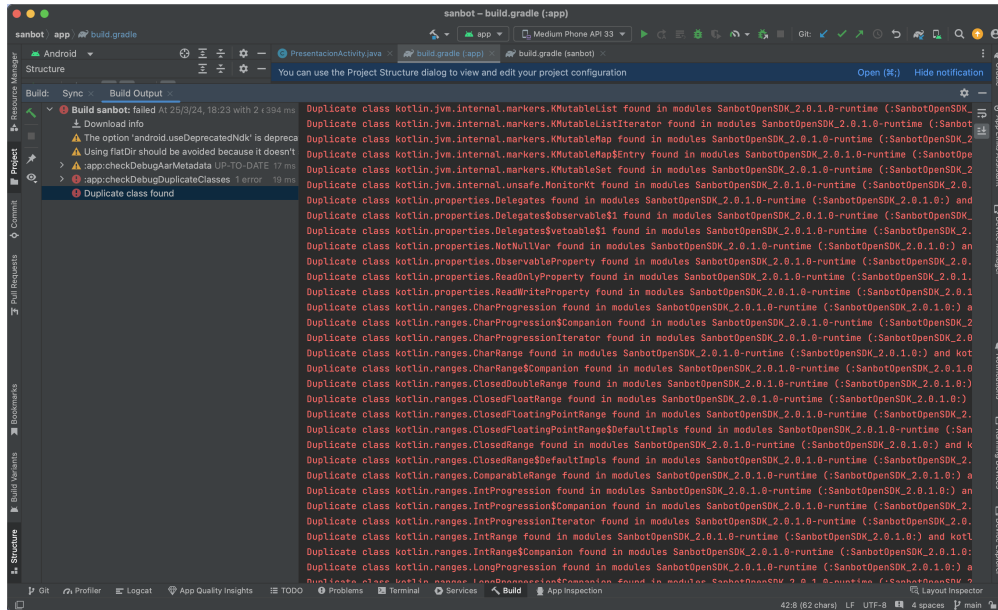


Figura 3.1: Captura de pantalla de la terminal del programa tras experimentar los problemas con el error de clases duplicadas utilizando el SDK *SanbotOpenSDK*

3.1.2. Análisis de herramientas de Inteligencia Artificial

Uno de los objetivos marcados para este proyecto era integrar herramientas de Inteligencia Artificial (IA) para conseguir una mayor naturalidad en el módulo conversacional con el propósito de ofrecer respuestas adecuadas y coherentes respecto al contexto de la conversación y reproducir respuestas con voces artificiales más similares a las humanas, evitando asociar la voz de *Sanbot* con una voz robótica que genere rechazo a los usuarios.

Para ello, se investigaron las diversas API que permitiesen cumplir esas dos funcionalidades, pudiendo observar en la tabla 3.1 las ventajas y desventajas que ofrecen cada una de ellas. A continuación se detallan los aspectos tenidos en cuenta en el análisis:

- **Herramienta:** Nombre de las herramientas evaluadas en este trabajo.
- **LLM:** Columna que indica si la herramienta es un LLM (Gran Modelo de Lenguaje), es decir, un tipo de IA capaz de realizar tareas como traducir, leer o resumir textos, pudiendo crear frases y predecir palabras. Los LLM han sido entrenados con gran cantidad de datos, por lo que son capaces de reconocer patrones textuales o aprender acerca de las características naturales y contextuales del lenguaje [12].
- **TTS:** Columna que indica si la herramienta cuenta con alguna función que permita reproducir el texto introducido mediante síntesis de voz usando *Text to Speech*.
- **Límite:** Columna que indica si la herramienta tiene algún tipo de cuota de uso limitada.

- **Precio mensual:** Columna que indica el coste mensual en dólares de utilizar la herramienta.
- **Observaciones:** Columna en el que se detalla otro tipo de información relevante acerca de la herramienta.

Herramienta	LLM	TTS	Límite	Precio mensual	Observaciones
Speechify ²	No	Sí	No conocido	11'58\$	-
Murf AI ³	No	Sí	24M de caracteres	250\$	-
Wondercraft ⁴	No	Sí	60 mins. de audio al mes	29\$	API muy centrada en temas de doblaje, podcasts, etc.
ElevenLabs ⁵	No	Sí	100,000 caracteres al mes	22\$	-
Lovo ⁶	No	Sí	300 mins. de audio al mes	24\$	-
LLaMA ⁷	Sí	Sí	-	-	Necesidad de contar con un buen equipo, un servidor adecuado, etc.
Google Cloud ⁸	Sí	Sí	-	-	Variación de los precios dependiendo de los servicios utilizados. Gran cantidad de bibliotecas API que hacen la plataforma muy confusa.
OpenAI ⁹	Sí	Sí	-	-	Variación de los precios en función del modelo LLM a utilizar.

Tabla 3.1: Comparación de herramientas de Inteligencia Artificial

²<https://speechify.com/>

³<https://murf.ai/>

⁴<https://www.wondercraft.ai/>

⁵<https://elevenlabs.io/>

⁶<https://lovo.ai/>

⁷<https://llama.meta.com/>

⁸<https://cloud.google.com/>

⁹<https://openai.com/>

A pesar de que se barajó la opción de utilizar dos herramientas de IA separadas en caso de que no contaran con funciones de tipo *LLM* y de *Text to Speech* simultáneamente, se dio preferencia a aquellas que contaban con ambas funcionalidades para facilitar su incorporación en el trabajo. Entre las restantes (**LLaMA**, **Google Cloud** y **OpenAI**), se descartaron **LLaMA** por su complejidad de instalación y por no contar con el equipo necesario y también **Google Drive** debido a su confusa biblioteca con numerosas API, quedando la herramienta de **OpenAI** como la definitiva que se utilizaría en este trabajo.

3.1.3. Modelos de Inteligencia Artificial OpenAI

Una vez seleccionada la API de *OpenAI* como herramienta a utilizar para este trabajo, se debía seleccionar el modelo de LLM que se quería usar, para ello, se investigaron las ventajas y desventajas con las que contaba cada uno, representadas en la tabla 3.2 cuyo contenido se detalla a continuación:

- **LLM**: Son los principales modelos de lenguaje de gran tamaño que ofrece la API de *OpenAI*.
- **Ventana de contexto**: Es la cantidad de información que el LLM retiene durante la interacción. Por ejemplo, en el caso de GPT-4, el modelo retendrá los últimos 8.192 *tokens* de la conversación, es decir, tiene un almacenamiento limitado de la interacción.
- **Multimodal**: Algunos de los modelos que ofrece *OpenAI* permiten datos de entrada multimodales, pudiendo realizar consultas a partir de **textos** o **imágenes**. Los modelos que no cuentan con apartado multimodal únicamente pueden realizar consultas textuales.
- **Precio input**: Es el precio que cuesta enviar **un millón** de *tokens* como entrada.
- **Precio output**: Es el precio que cuesta recibir **un millón** de *tokens* como salida.
- **Datos de entrenamiento**: La información utilizada para el entrenamiento de los modelos es previa a la fecha que aparece en esta columna.

LLM	Ventana de contexto	Multimodal	Precio input	Precio output	Limitación de los datos
GPT-3.5 Turbo	16.385 <i>tokens</i>	No	0'50\$	1'50\$	Hasta septiembre 2021
GPT-4	8.192 <i>tokens</i>	Sí	30'00\$	60'00\$	Hasta septiembre 2021
GPT-4 Turbo	128.000 <i>tokens</i>	Sí	10'00\$	30'00\$	Hasta diciembre 2023
GPT-4o mini	128.000 <i>tokens</i>	Sí	0'15\$	0'60\$	Hasta octubre 2023
GPT-4o	128.000 <i>tokens</i>	Sí	5'00\$	15'00\$	Hasta octubre 2023

Tabla 3.2: Comparación de modelos ofrecidos por la API de *OpenAI*

Durante gran parte del proceso se trabajó con **GPT-3.5 Turbo** y el resto de modelos se descartaron debido a su alto precio, limitación de ventana de contexto y porque no se pretendía utilizar la función multimodal. Sin embargo, a fecha del **18 de julio de 2024**, se incluyó en la API de *OpenAI* **GPT-4o mini** a la lista de modelos, ofreciendo un precio todavía más económico, con datos de entrenamiento más recientes y ventana de contexto mucho más amplia, por lo que se sustituyó por el modelo **GPT-3.5 Turbo** que estaba siendo utilizado en el trabajo.

Como ya se ha mencionado, *OpenAI* cuenta con funciones *Text to Speech*, ofreciendo al cliente dos posibles modelos: **TTS** y **TTS-HD**. Ya que se buscaba la menor latencia posible y la diferencia de calidad no era apenas apreciable, se seleccionó el modelo **TTS** estándar para la realización del proyecto, en la tabla 3.3 se observan las principales diferencias.

Modelo	Latencia	Calidad	Precio
TTS	Baja	Baja	15'00\$
TTS-HD	Alta	Alta	30'00\$

Tabla 3.3: Comparación de modelos *Text to Speech* ofrecidos por la API de *OpenAI*

3.1.4. Almacenamiento de los datos

Ya que el módulo conversacional es un programa a utilizar en momentos puntuales, no se vio necesario almacenar datos sobre la conversación en una base de datos.

Sin embargo, sí que se utiliza el almacenamiento local del programa mediante el paquete *SharedPreferences*¹⁰, donde distintas actividades pueden guardar u obtener valores clave utilizados durante la ejecución del programa.

3.2. Arquitectura general del Sistema

A continuación se detalla el funcionamiento del sistema utilizando dos diagramas que muestran dos niveles de profundidad del módulo conversacional. La explicación del funcionamiento de los módulos del *QihanOpenSDK* y del resto de paquetes utilizados en el módulo conversacional se describe en el Anexo B y Anexo C respectivamente.

En este diagrama (Figura 3.2) se puede observar la interacción general de las tres principales partes del sistema: la **interfaz de usuario**, el **módulo conversacional** y el **módulo de Inteligencia Artificial**.

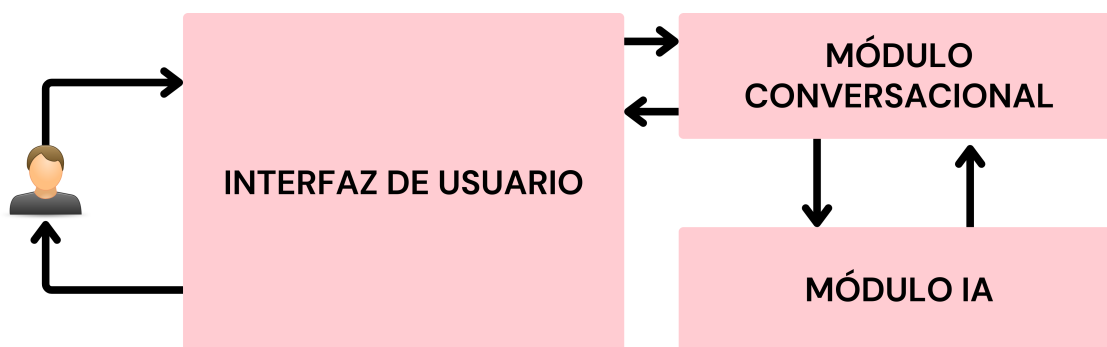


Figura 3.2: Diagrama explicativo del funcionamiento del sistema en profundidad 1

El usuario utiliza la interfaz de usuario para introducir datos ya sea por medio de la voz o a través del uso de la pantalla táctil.

La interfaz de usuario se comunica con el módulo conversacional para llevar a cabo las acciones correspondientes que el usuario desea realizar.

Por último, el módulo conversacional se comunica con el módulo de Inteligencia Artificial cuando sea necesario a través de peticiones HTTP.

¹⁰<https://developer.android.com/reference/android/content/SharedPreferences>

A continuación se explica con mayor profundidad la interacción entre los distintos componentes del sistema (Figura 3.3).

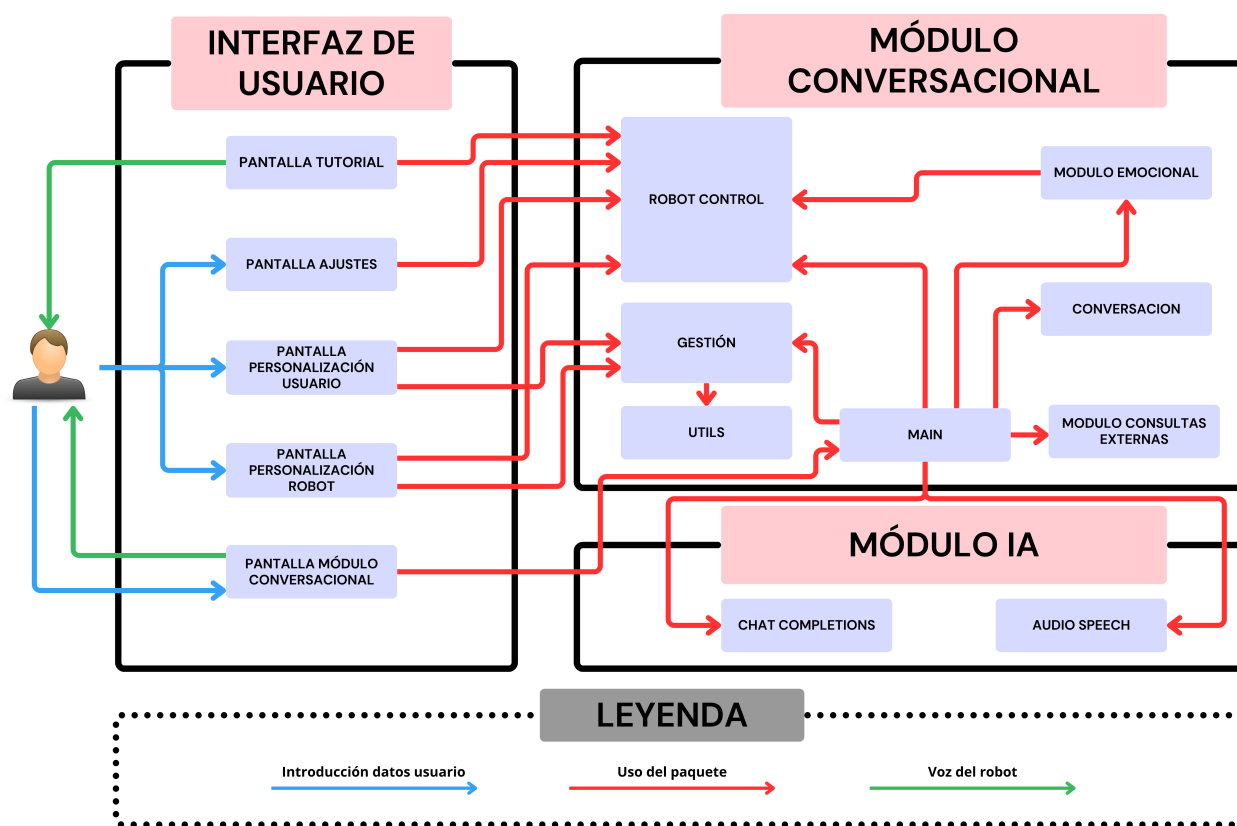


Figura 3.3: Diagrama explicativo del funcionamiento del sistema en profundidad 2

La **interfaz de usuario** se compone de las siguientes pantallas:

- **Pantalla personalización de usuario:** Permite que el usuario introduzca información personal como su nombre o su edad que será utilizada posteriormente durante la conversación.
- **Pantalla personalización del robot:** Permite que el usuario modifique ciertos parámetros del robot que se aplicarán posteriormente a la conversación, estos pueden ser el tipo de voz, el género que simulará tener, el grupo de edad al que simulará pertenecer o similares.
- **Pantalla ajustes:** Permite que el usuario modifique aspectos relacionados con la velocidad, entonación o volumen de la voz del robot. También permite gestionar otros aspectos del módulo conversacional relacionados con la comunicación por medio de voz o teclado.

- **Pantalla tutorial:** Muestra una serie de pasos que explican cómo utilizar el módulo conversacional acompañados de capturas de pantalla.
- **Pantalla módulo conversacional:** Pantalla principal del programa donde el usuario puede interactuar con el robot y comenzar la conversación. Aquí puede realizar diversas acciones como repetir lo que ha dicho el robot, detener su habla.

El **módulo conversacional** se compone de los siguientes módulos:

- **Main:** Sección principal que realiza las acciones deseadas por el usuario mediante el uso del resto de módulos.
- **Robot control:** Módulo que utiliza las funciones del SDK para realizar acciones propias del robot, entre ellas: movimiento de extremidades, gestión del habla, cambio de expresión facial y similares.
- **Modulo emocional:** Módulo encargado de gestionar el apartado emocional de la conversación.
- **Gestión:** Módulo encargado de la reproducción de audio y del almacenamiento local de datos en el programa.
- **Conversación:** Módulo encargado de construir los mensajes que se van intercambiando entre el usuario y el robot a modo de chat.
- **Módulo consultas externas:** Módulo encargado de gestionar consultas no relacionadas con *OpenAI*.
- **Utils:** Módulo encargado de permitir al módulo de gestión reproducir ristas de bytes en forma de audio.

Y por último, el **módulo IA:**

- **Chat Completions:** *Endpoint* encargado de tratar las consultas realizadas por el usuario y ofrecer una respuesta adecuada haciendo uso de la API de *OpenAI*.
- **Audio Speech:** *Endpoint* encargado de utilizar las funciones *Text to Speech* ofrecidas por la API de *OpenAI*.

Capítulo 4

Implementación

En este capítulo se comenta el proceso incremental de implementación llevado a cabo durante el trabajo.

Se parte inicialmente de la versión desarrollada durante la demo con el robot en las primeras semanas de trabajo. En esta demostración preliminar se cumplen las primeras cuatro historias de usuario planteadas en la sección 2.4 del capítulo de Análisis. Una vez terminada esta prueba, se analizaron y contrastaron las diferentes alternativas entre las posibles herramientas IA, comentadas en la sección 3.1.2, de las cuales se eligió *OpenAI* como la opción a utilizar en el trabajo.

En total se desarrollaron 6 iteraciones que se explican a continuación detallando las distintas ventajas e inconvenientes que se encontraron y las historias de usuario que se cumplen tras su implementación.

4.1. Iteración 1: Uso del endpoint Chat Completions

La primera incorporación al trabajo fue la utilización del *endpoint* de *Chat Completions* de *OpenAI*, que permite realizar consultas al LLM en formato texto (o en forma de imagen en caso de que se utilicen aquellos modelos que soporten esta funcionalidad) y recibir una respuesta coherente con la consulta formulada. El uso de este *endpoint* se explica con más detalle en el Anexo D.

El módulo *ModuloConversacionalConChatCompletions* se presenta en la figura 4.1 y su funcionamiento es el siguiente:

El usuario pulsa el botón de grabar consulta (**Paso 1**), que llama a la función *registrarConsulta* (**Paso 2**), encargada de que el robot se ponga en modo escucha (**Paso 3**)

y encargada también de llamar a la función *reconocerConsulta* (**Paso 4**), que obtiene lo que ha dicho el usuario en forma de texto mediante la función *onRecognizeResult* (**Paso 5**) y además lo muestra por pantalla (**Paso 6**).

En caso de que la consulta del usuario no se haya registrado correctamente y quiera corregirla podrá volver a pulsar el botón de grabar consulta para reiniciar el proceso. En caso contrario, el usuario enviará la consulta pulsando el botón (**Paso 7**) y es entonces cuando se llamará a la función *APIChatGPT* (**Paso 8**) que se encarga de enviar la petición al *endpoint* de *Chat Completions* de *OpenAI*, recoger su resultado (**Paso 9**) y sacarlo por pantalla (**Paso 10**). Finalmente, la función se encargará de que la respuesta sea pronunciada (**Paso 11**) a través de los altavoces del robot (**Paso 12**) y llegue al usuario (**Paso 13**).

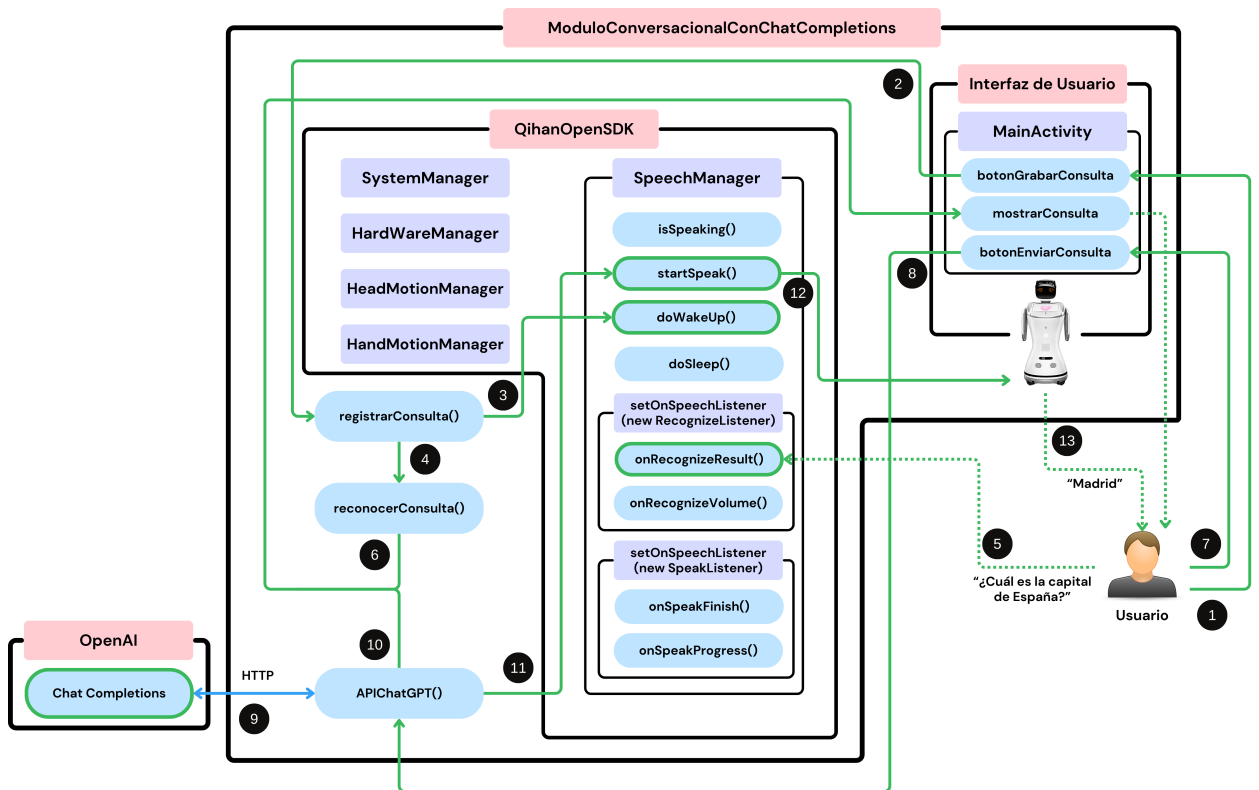


Figura 4.1: Diagrama del funcionamiento principal del módulo conversacional usando la función *Chat Completions* de la API de *OpenAI*

4.1.1. Historias de usuario cumplidas

Además de las cuatro primeras historias de usuario desarrolladas con la primera demo del módulo conversacional, ahora también se han conseguido las siguientes:

1. Quiero poder hablar con el robot de cualquier tema (Historia de usuario n^o5)

2. **Quiero que el robot guarde el hilo de la conversación** (Historia de usuario n^o6)
3. **Quiero visualizar las palabras que el robot me ha entendido** (Historia de usuario n^o7)
4. **Quiero observar tanto mis respuestas como las respuestas del robot en la pantalla** (Historia de usuario n^o8)
5. **Quiero que el robot no dé respuestas muy extensas** (Historia de usuario n^o9)

4.1.2. Ventajas y desventajas

Las ventajas y desventajas encontradas en esta iteración han sido las siguientes:

– **Ventajas:**

1. **Gran abanico de temas de conversación:** Gracias a las funcionalidades con las que cuenta *Chat Completions*, el robot siempre dará una contestación coherente a la consulta que haga el usuario, quitando la necesidad de crear una pila de preguntas y respuestas o de conocer con anterioridad el tema a tratar como sucedía en la demo.
2. **Tiempo de respuesta:** La petición a la API con el modelo **gpt-4o mini** es rápida y, exceptuando respuestas muy largas, la conversación se puede desarrollar con normalidad.

– **Desventajas:**

1. **Limitación respecto a consultas actuales:** A pesar de que el *ChatGPT* sí que consulta páginas web como fuente de información para proporcionar una respuesta a las consultas del usuario, la API de *OpenAI* no cuenta con esa funcionalidad y depende de los datos con los que ha sido entrenado. Esto impide que el usuario pida consultas como “**dime las últimas noticias**”, “**qué tiempo hará la semana que viene**” o similares.

4.1.3. Conclusión

Se ha conseguido una notable mejora respecto a la versión de demostración debido a que no se necesita predefinir la conversación ni conocer el tema a hablar con anterioridad ya que ahora al usar la API de *OpenAI* con la función *Chat Completions*, el robot ofrece una respuesta coherente con la consulta del usuario.

4.2. Iteración 2: Uso del endpoint Audio Speech

Como siguiente paso, se incorporó la utilización del *endpoint* de *Audio Speech* de *OpenAI*, el cual permite seleccionar entre seis voces similares al habla humana de forma artificial. El uso de este *endpoint* se explica con más detalle en el Anexo D.

El módulo *ModuloConversacionalConChatCompletionsYAudioSpeech* se presenta en la figura 4.2 y su funcionamiento es el siguiente:

El proceso es muy similar a la versión anterior, pero tras solicitar y obtener la respuesta de la consulta con la función de *APIChatGPT* (**Paso 9**), esta se mostrará por pantalla (**Paso 10**) y el siguiente paso será la función *APIChatGPTVoz* (**Paso 11**), dónde además de la respuesta se pasará como parámetro la voz que se quiera utilizar. En caso de que no sea la voz de *Sanbot*, se llamará al *endpoint* de *Audio Speech* por medio de una petición HTTP (**Paso 12a**). La respuesta se enviará al *MediaPlayer*, que reproducirá el audio con la función *start* (**Paso 13a**) a través de los altavoces del robot (**Paso 14**). En caso contrario, es decir, sea la voz de *Sanbot* (**Paso 12b**) se seguirá el proceso de versiones anteriores mediante la función *startSpeak* (**Paso 13b**) y los altavoces del robot (**Paso 14**).

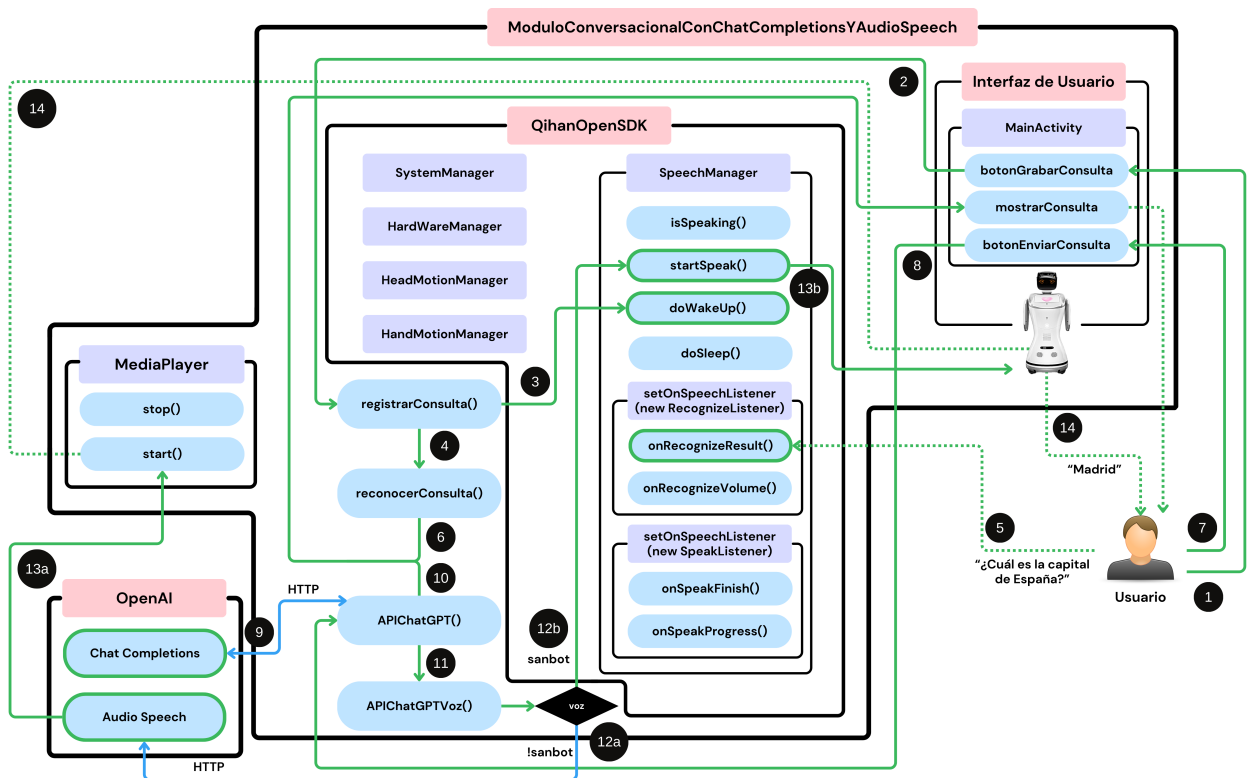


Figura 4.2: Diagrama del funcionamiento principal del módulo conversacional usando la función *Chat Completions* y *Audio Speech* de la API de *OpenAI*

4.2.1. Historias de usuario cumplidas

Además de las historias de usuario conseguidas en la iteración anterior, ahora se cumple la siguiente historia de usuario:

1. **Quiero que las respuestas sean reproducidas por voces no robóticas o artificiales** (Historia de usuario n^o10)

4.2.2. Ventajas y desventajas

Las ventajas y desventajas encontradas en esta iteración han sido las siguientes:

– **Ventajas:**

1. **Variedad de voces:** Esta función de la API de *OpenAI* permite enviar como parámetro con qué voz deseas escuchar las respuestas, por lo que el usuario del módulo conversacional puede seleccionar aquella con la que se sienta más cómodo incluyendo voces neutrales, masculinas, femeninas, más graves o más agudas.
2. **Múltiples idiomas:** La voz permite reproducir el texto a pronunciar en diversos lenguajes (más de 50), siendo muy útil en caso de añadir al módulo funciones de traducción.

– **Desventajas:**

1. **Voces con acento inglés:** A pesar de que, como se ha comentado, permite trabajar con múltiples lenguajes. Las voces han sido optimizadas en inglés, generando algo de acento al utilizar el resto de idiomas.
2. **Entonación fija:** La funcionalidad sí que permite configurar la velocidad de síntesis de voz pero no la entonación a diferencia de la voz nativa del robot *Sanbot*.
3. **Problemas de pronunciación:** En algunas ocasiones, se han experimentado problemas de pronunciación por parte de estas voces, sobre todo al pronunciar números.

4.2.3. Conclusión

Esta función perfecciona el módulo conversacional, permitiendo al usuario elegir la voz con la que se sienta más cómodo durante la conversación. Aun así, no se ha eliminado por completo la opción de utilizar la voz nativa del robot *Sanbot*, para evitar las desventajas mencionadas en el apartado anterior.

4.3. Iteración 3: Incorporación del aspecto emocional

Con motivo de mejorar el apartado social de una conversación, se ha añadido el aspecto emocional al módulo conversacional siguiendo como modelo la rueda de emociones de **Plutchik** (Figura 4.3), que cuenta con un total de **48 emociones** y cuyo funcionamiento se explica con más detalle en el Anexo E.

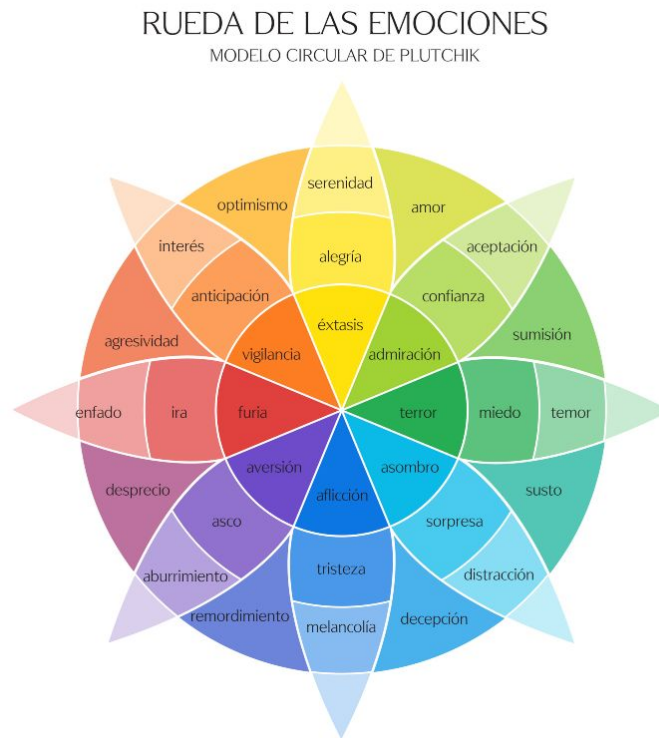


Figura 4.3: Modelo de rueda de emociones de Plutchik

Este modelo fue utilizado para mejorar la interacción social entre la persona y el robot, de manera que la conversación es adaptada en función de las emociones del usuario. Para llevar a cabo este proceso, además de mantener una conversación, se solicitó a la API de *Chat Completions* que cumpliera lo siguiente:

- Cada respuesta del usuario debía analizarla y categorizarla con una o varias emociones siguiendo el modelo de la rueda de emociones de Plutchik.
- Debía empatizar con esa respuesta lo máximo posible, dando una respuesta coherente a la conversación en función con el sentimiento interpretado y que también debía categorizar siguiendo el mismo modelo. Con esto se abrían dos posibilidades:
 1. Que el robot imitase la emoción interpretada, es decir, si el usuario estaba triste, el robot estaría triste.
 2. Que el robot tratase de animar al usuario en función de la emoción interpretada.

Siguiendo el objetivo del proyecto **PLEISAR**, se consideró que la segunda opción ayudaría a mejorar el estado de ánimo a los usuarios como podrían ser las *Personas Antonio, Patricia o Alberto*.

Tal y como se observa en la figura 4.4, la API de *OpenAI* se encarga de dar una respuesta a la consulta del usuario (**¡Hola! Qué alegría verte**) e interpretarla emocionalmente dando un valor numérico a cada emoción del modelo (**[(2)/(3)]**). Una vez conseguido, se hacen uso de las siguientes funciones que forman parte del módulo emocional: **separarRespuestaGPT**, **gestionEmocional**, **sentimientosRobot**, **sentimientosUsuario** y **expresividadGestual** para separar esa valoración emocional de la respuesta que da al usuario y, dependiendo del sentimiento reconocido, el robot realiza movimientos corporales, cambia su expresión facial o incluso enciende los LED que posee.



Figura 4.4: Ejemplo de uso del módulo emocional

4.3.1. Historias de usuario cumplidas

Además de las historias de usuario conseguidas en la iteración anterior, ahora se cumplen las siguientes historias de usuario:

1. **Quiero que el robot reconozca mi estado de ánimo al conversar con él** (Historia de usuario nº11)
2. **Quiero que el robot refleje en su expresión facial su estado de ánimo** (Historia de usuario nº12)

3. **Quiero que el robot realice gestos al expresar ciertas emociones** (Historia de usuario n^o13)

4.3.2. Ventajas y desventajas

Las ventajas y desventajas encontradas en esta iteración han sido las siguientes:

– Ventajas

1. **Conversación con empatía emocional:** Esta funcionalidad añade al robot conciencia emocional y gracias a ella reconduce la conversación que mantiene con el usuario.
2. **Mejora de la expresión no verbal:** Tal y como se comentaba en los primeros capítulos de esta memoria, la expresión no verbal toma una gran importancia en la conversación. Gracias a este análisis, el robot cambia su expresión facial y realiza movimientos o enciende luces de colores en función de las emociones interpretadas.

– Desventajas

1. **Limitación con las expresiones faciales del robot:** El robot sólo cuenta con 18 expresiones faciales, lo cual ha forzado que más de una emoción perteneciente a las 48 interpretadas por *OpenAI* muestren la misma expresión facial. Por ejemplo, en la imagen 2.1 la expresión facial “**GRIEVANCE**” se utiliza tanto para el sentimiento de melancolía como para sentimientos similares al miedo, lo cual puede ser algo confuso.
2. **Limitación en los colores LED:** A pesar de que la rueda de emociones tiene una categorización por colores, el encendido de los LED de las partes del cuerpo de *Sanbot* tiene únicamente los colores **rojo, blanco, morado, verde, rosa, azul y amarillo**. Debido a esta limitación, no todas las emociones podrían representarse con un color diferente así que las emociones básicas fueron representadas siguiendo el código de colores de los personajes de la película “**Del revés**” de **Disney Pixar** (Figura 4.5), considerándolo un modelo que muchas personas reconocerían. Aunque también se han representado de color rosa las emociones **Vergüenza** y **Amor** considerándola una representación bastante estandarizada.



Figura 4.5: Personajes de la película “Del revés” de Disney Pixar

4.3.3. Conclusión

La incorporación del aspecto emocional ha mejorado el módulo conversacional consiguiendo que el robot interprete sentimientos en las consultas del usuario y reconduzca la conversación en base a ellos, expresándolos a través de comunicación no verbal y generando así un efecto de cercanía y confianza que no estaba presente en las anteriores iteraciones.

4.4. Iteración 4: Personalización del módulo

Con la intención de permitir al usuario cambiar algunos parámetros que ofrezcan una conversación personalizada, se han añadido dos módulos más: personalización del usuario y personalización del robot.

La personalización del usuario, permite introducir el nombre y la edad en la pantalla (Figura 4.6), que serán guardados en el almacenamiento local del programa y luego serán enviados como contexto usando *Chat Completions*, de forma que durante la conversación se referirá al usuario por su nombre y adaptará la conversación en función de su edad.

Por otro lado, la personalización del robot permite al usuario cambiar a través de la pantalla (Figura 4.7) parámetros como la voz con la que quiere que el robot mantenga la conversación, el género o el grupo de edad que simulará tener o elegir un tema como contexto de la conversación.

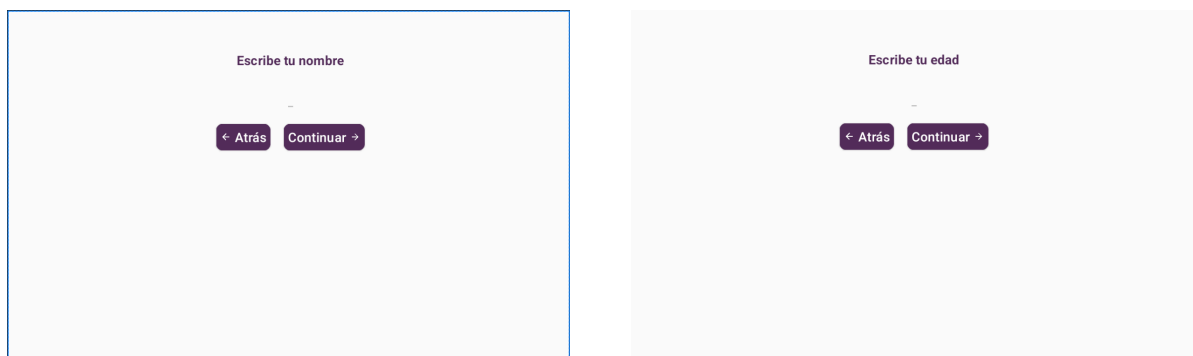


Figura 4.6: Capturas del módulo de personalización del usuario

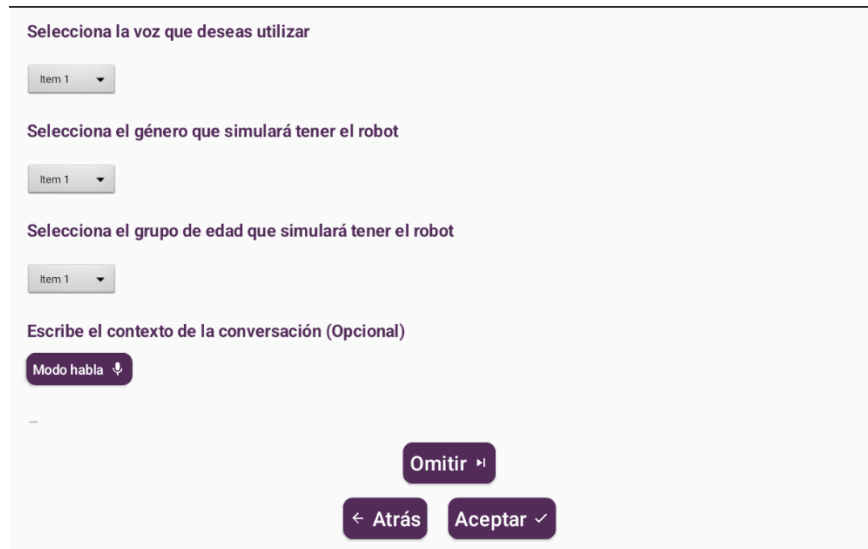


Figura 4.7: Captura del módulo de personalización del robot

4.4.1. Historias de usuario cumplidas

Además de las historias de usuario conseguidas en la iteración anterior, ahora se cumplen las siguientes historias de usuario:

1. **Quiero que el robot conozca datos personales como mi nombre y edad** (Historia de usuario n^o14)
2. **Quiero poder seleccionar la voz que quiero que reproduzca las respuestas** (Historia de usuario n^o15)
3. **Quiero poder personalizar características del robot como su edad o género** (Historia de usuario n^o16)
4. **Quiero poder simular situaciones ficticias con el robot** (Historia de usuario n^o17)

4.4.2. Ventajas y desventajas

Las ventajas y desventajas encontradas en esta iteración han sido las siguientes:

– Ventajas

1. **Adaptación del módulo:** Esta funcionalidad permite al usuario tener una conversación más natural mediante la selección de las opciones con las que se sienta más cómodo y que se aplicarán posteriormente en la conversación.

2. **Mejora de la accesibilidad:** Se puede observar en las figuras 4.6 y 4.7 que la pantalla es visible de forma completa sin necesidad de deslizar hacia abajo, evitando dificultades para personas no acostumbradas con dispositivos táctiles. Además en la figura 4.7 se puede ver el botón “**modo habla**”, mediante el cual el usuario puede rellenar la respuesta utilizando el habla, evitando así el uso del teclado.

– **Desventajas**

1. **Dificultad de uso:** La posibilidad de personalizar el módulo implica añadir pantallas, botones y selección de múltiples opciones, pudiendo complicar levemente la utilización del módulo.

4.4.3. Conclusión

La incorporación de la personalización del usuario y del robot han permitido que se pueda adaptar el módulo a las preferencias del usuario, consiguiendo un trato más individualizado durante la conversación. Además se han incluido funcionalidades que mejoran la accesibilidad del sistema, permitiendo que un mayor número de usuarios utilicen el módulo conversacional.

4.5. Iteración 5: Configuración del módulo

Con el propósito de permitir al usuario modificar distintas características del módulo conversacional, se ha añadido una pantalla (Figura 4.8) mediante la cual se pueden realizar las siguientes acciones:

1. **Seleccionar el volumen de habla del robot:** El usuario puede aumentar o disminuir el volumen de los altavoces de *Sanbot*.
2. **Seleccionar la velocidad y entonación del habla:** El usuario puede aumentar o disminuir la velocidad de síntesis de voz e incluso puede configurar la entonación del robot siempre que esté utilizando la voz propia de *Sanbot*, ya que las voces de *OpenAI* sólo permiten configurar la velocidad.
3. **Modo teclado:** El usuario puede alternar entre tener una conversación únicamente por medio oral o elegir el modo teclado, con en el que puede escribir las consultas como alternativa (Figura 4.9).

4. **Conversación automática:** El usuario puede elegir en caso de estar en modo teclado (ya que en el modo sin teclado siempre está en modo conversación automática) si quiere que el envío de consultas sea controlado. En caso contrario, siempre que el robot termine de hablar, se activará de nuevo el reconocimiento de voz, reiniciando el proceso.



Figura 4.8: Captura de la pantalla de ajustes

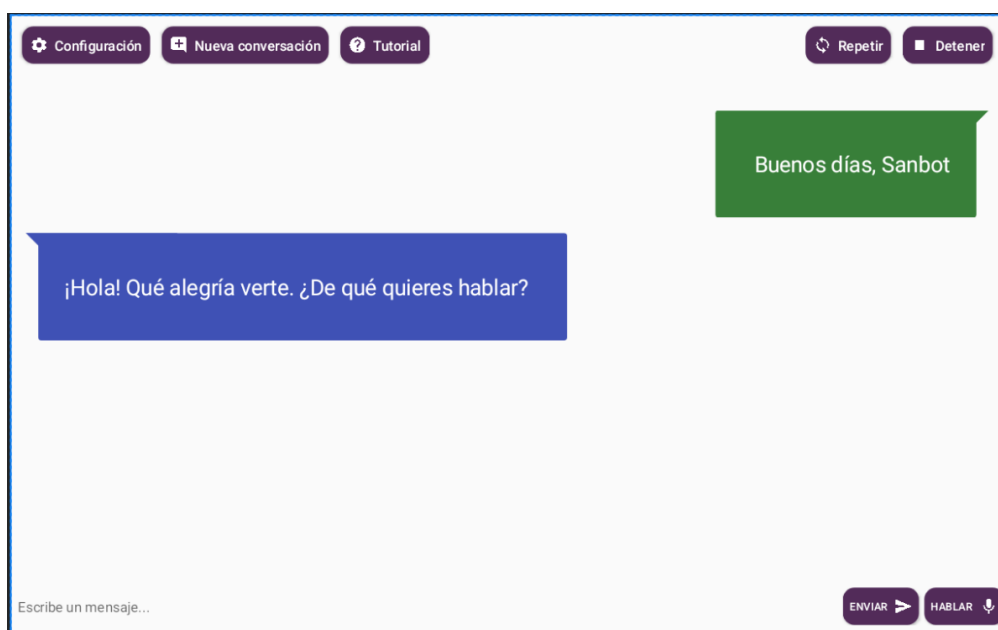


Figura 4.9: Captura del módulo conversacional con el modo teclado activado

4.5.1. Historias de usuario cumplidas

Además de las historias de usuario conseguidas en la iteración anterior, ahora se cumplen las siguientes historias de usuario:

1. **Quiero poder modificar el volumen de los altavoces del robot** (Historia de usuario n^o18).
2. **Quiero poder configurar la velocidad y entonación de las voces** (Historia de usuario n^o19).
3. **Quiero enviar las consultas por teclado en vez de vía voz** (Historia de usuario n^o20).
4. **Quiero que la conversación no se interrumpa** (Historia de usuario n^o21).

4.5.2. Ventajas y desventajas

Las ventajas y desventajas encontradas en esta iteración han sido las siguientes:

– Ventajas

1. **Adaptación del módulo al gusto del usuario:** Esta funcionalidad permite al usuario ajustar la voz del robot consiguiendo una conversación más cómoda.
2. **Mejora de la accesibilidad:** Gracias a esta iteración, los usuarios con problemas auditivos pueden aumentar el volumen de los altavoces del robot. Además, la activación del modo teclado permite utilizar el módulo mediante entrada de texto, consiguiendo que usuarios con dificultades del habla puedan usarlo sin problemas.

– Desventajas

1. **Dificultad de uso:** Como la iteración anterior, esta funcionalidad implica añadir pantallas, botones y selección de múltiples opciones, pudiendo complicar la utilización del módulo.

4.5.3. Conclusión

Esta iteración está muy centrada en la personalización de la voz del robot y en mejorar la accesibilidad del sistema, dando la posibilidad al usuario de utilizar el módulo conversacional a través de entrada de texto o haciendo uso de la voz como en iteraciones previas.

4.6. Iteración 6: Otras funcionalidades

Para cumplir el resto de historias de usuario, se han añadido las siguientes funcionalidades al módulo conversacional:

1. **Tutorial:** Antes de empezar la conversación, comenzará un tutorial donde se explica el funcionamiento principal del módulo y las acciones que tienen cada uno de los botones que aparecen en la pantalla (Figura 4.10).
2. **Repetición de la respuesta:** Se ha añadido un botón que permite repetir la última respuesta del robot (Figura 4.11).
3. **Detención del habla del robot:** Para evitar que se alargue la respuesta sonora del robot en caso de pronunciar una frase muy extensa, se ha añadido un botón que detiene el habla (Figura 4.11).
4. **Mejora de las respuestas en pantalla:** Anteriormente sólo se veía en pantalla la consulta del usuario y la respuesta que ofrecía el robot, pero se perfeccionó haciendo que en pantalla apareciese la conversación completa (Figura 4.11).
5. **Nueva conversación:** Se ha añadido un botón que vacía la conversación tanto de la pantalla como del contexto de las peticiones de la API de *OpenAI* (Figura 4.11).
6. **Peticiones actuales:** Ya que *OpenAI* no permite que el usuario haga consultas sobre eventos recientes, se ha añadido un módulo de peticiones externas que permite hacer consultas a otras API para conocer información sobre la actualidad. Por ejemplo, si el usuario empieza su consulta con la palabra “**Películas**”, el robot consultará las películas que se encuentran actualmente en la cartelera utilizando la API de *TMDB*¹. Aunque, por falta de tiempo no se pudo profundizar en otro tipo de peticiones como el tiempo metereológico o noticias.

¹<https://www.themoviedb.org/>

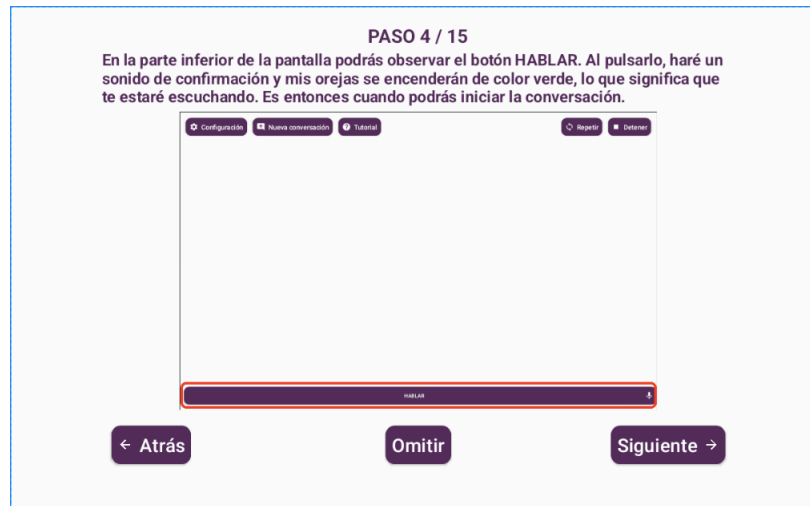


Figura 4.10: Captura de un paso intermedio del tutorial del módulo conversacional

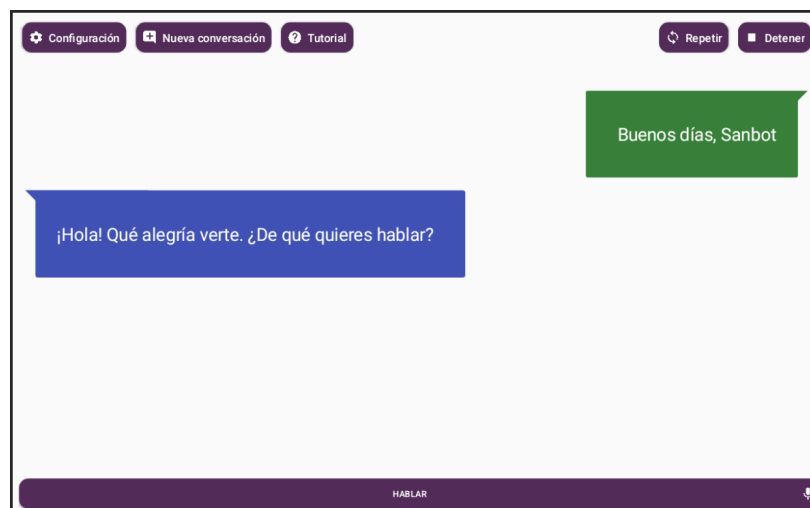


Figura 4.11: Captura del módulo conversacional con el modo teclado desactivado

4.6.1. Historias de usuario cumplidas

Además de las historias de usuario conseguidas en la iteración anterior, ahora se cumplen la siguientes historias de usuario:

1. **Quiero que el módulo cuente con un tutorial** (Historia de usuario n^o22).
2. **Quiero repetir la última respuesta del robot** (Historia de usuario n^o23).
3. **Quiero poder detener la respuesta del robot** (Historia de usuario n^o24).
4. **Quiero que la conversación aparezca en pantalla** (Historia de usuario n^o25).

5. **Quiero poder cambiar de conversación** (Historia de usuario n^o26).
6. **Quiero poder realizar consultas de actualidad, como el tiempo meteorológico o noticias** (Historia de usuario n^o27).

4.6.2. Ventajas y desventajas

Las ventajas y desventajas encontradas en esta iteración han sido las siguientes:

– **Ventajas**

1. **Mejoras del módulo:** Todas las funcionalidades han permitido mejorar el funcionamiento del módulo con aspectos relacionados con la visualización de información en pantalla, repetición y detención de la respuesta del robot o similares.
2. **Mejora de la accesibilidad:** La repetición de la respuesta del robot o la aparición completa de la conversación en pantalla permiten a personas con problemas auditivos utilizar el módulo de una forma más cómoda.

– **Desventajas**

1. **Previo aprendizaje:** A pesar de que el tutorial puede ayudar a utilizar el módulo conversacional, han sido necesarios varios pasos para explicar correctamente su funcionamiento, pudiendo llegar a ser algo tedioso para los usuarios.

4.6.3. Conclusión

Esta última iteración contiene un conjunto de funcionalidades que han mejorado ampliamente el módulo conversacional y ha permitido completar todas las historias de usuario planteadas al principio del trabajo.

4.7. Conclusiones generales

La implementación de las seis iteraciones han sido fundamentales para perfeccionar el módulo conversacional y completar las historias de usuario. Para cada una de las versiones se han evaluado las diferentes ventajas e inconvenientes, conociendo así las funcionalidades que ofrecen al módulo y las limitaciones con las que cuenta. Por último, cabe destacar que se ha hecho especial énfasis en mejorar la accesibilidad del sistema, consiguiendo que el módulo pueda ser utilizado por personas con dificultades auditivas, visuales y relacionadas con el habla, y para personas poco acostumbradas al uso de dispositivos electrónicos.

Capítulo 5

Evaluación final con usuarios

En este capítulo se presenta la evaluación final con usuarios del modulo conversacional. El objetivo fundamental ha sido comprobar su correcto funcionamiento así como valorar su usabilidad, aceptación y el impacto en la interacción de la adición de los aspectos emocionales y de personalización. Así mismo se ha consultado sobre el tipo de voz a usar y sobre el agrado suscitado en los usuarios.

En primer lugar se presentan los detalles metodológicos (técnicas, usuarios, sesiones) y después los resultados conseguidos.

5.1. Técnicas de evaluación

Las técnicas utilizadas para la evaluación con usuarios son las siguientes:

Testing A/B

La técnica **Testing A/B** consiste en dividir a los usuarios en dos grupos, cada grupo evaluará una versión del sistema desarrollado [13]. El grupo A interactuó con el sistema que tenía capacidad de realizar consultas a *OpenAI* mediante el *endpoint Chat Completions* y contaba con las voces naturales del *endpoint Audio Speech*, pudiendo personalizar los datos del usuario o del robot, además de contar con el análisis emocional del usuario y la expresión emocional del robot mediante gestos, expresión facial o luces LED. Por otro lado, el grupo B, interactuó con el sistema que únicamente tenía capacidad de realizar consultas haciendo uso del *endpoint* de *Chat Completions* por lo que escucharían en todo momento la voz original del robot *Sanbot* sin personalización, análisis emocional ni expresión emocional.

Cuestionarios

Técnica que consiste en una lista de preguntas para conocer la opinión del usuario acerca del sistema.

Para la evaluación se han utilizado los siguientes (todos ellos se pueden consultar en el anexo F):

- **Cuestionario inicial de caracterización:** Utilizado para conocer datos sobre el usuario como su edad, capacidades sensoriales al utilizar dispositivos electrónicos o frecuencia de uso de la tecnología.
- **Cuestionario POST VAVA-Q:** Cuestionario generado en el grupo de investigación para evaluar la aceptación tecnológica.
- **Cuestionario de usabilidad:** Utilizado para medir la usabilidad del sistema, las preguntas de la 2 a la 8 han sido extraídas del Test SUS (cuestionario estandarizado para medir usabilidad [14]) y el resto son preguntas específicas sobre el funcionamiento del módulo conversacional.
- **Cuestionario de likeability:** Cuestionario estándar [15] utilizado para medir el agrado que produce el sistema en el usuario.

5.2. Usuarios y sesiones

Participó un total de **18 usuarios**, 7 hombres y 11 mujeres con edades de entre 18 a los 83 años.

Las sesiones se llevaron a cabo en un laboratorio del grupo AffectiveLab de la *EINA*, durante **6 días**, con una duración por sesión de aproximadamente **30 minutos**, y en las que el autor del TFG actuó como evaluador de la prueba.

La estructura de la sesión fue la siguiente:

- Hoja de consentimiento informado, en la que se informa sobre el procedimiento de la evaluación y que el/la participante debe firmar.
- Cuestionario inicial de caracterización.
- Preferencia de voces, donde se hacía elegir a los usuarios con cuál de las seis voces se sentían más cómodos tener una conversación, tres de ellas se consideraron voces

masculinas y tres se consideraron voces femeninas. Cinco de las voces eran sacadas del *endpoint Audio Speech* de *OpenAI* y la voz restante era la propia del robot *Sanbot*.

- Conversaciones con el robot, la primera con un tema libre que elige el usuario y la segunda con un tema cerrado que elige el evaluador, con la intención de explotar la capacidad de crear conversaciones sobre situaciones ficticias (Tabla 5.1).
- Cuestionario **POST VAVA-Q**, cuestionario de **likeability** y cuestionario de **usabilidad**.

Género	Edad	Grupo	Tema libre	Tema cerrado
Hombre	56	A	Siglo de Oro	El robot actúa como esclavo del usuario durante la época romana
Hombre	22	A	Videojuegos	Simulación de una pelea en un mundo fantástico donde el robot es villano y el usuario es el héroe
Mujer	22	A	TFG	El robot y el usuario simulan estar en un safari
Mujer	23	A	Animales	Simulación de un viaje espacial
Mujer	22	B	Cantantes pop	El robot y el usuario simulan que son sirenas
Hombre	18	B	Colaboraciones entre artistas	El robot y el usuario simulan que son compañeros que van a realizar un robo
Mujer	18	A	Cantantes de reggaeton	El robot simula entrevistar al usuario que es una famosa actriz de cine
Hombre	22	B	Naturaleza	El usuario y el robot simulan ser vikingos
Mujer	19	B	Vida laboral	El usuario y el robot son supervivientes de un accidente aéreo
Hombre	19	A	Cine	El robot simula entrevistar al usuario que es ganador de un Oscar
Mujer	58	A	Día de hoy	El robot y el usuario simulan un viaje a Australia
Mujer	52	A	Vacaciones	El usuario simula ser directora de circo y el robot simula ser el público
Hombre	53	B	La radio	El usuario simula ser un <i>Sheriff</i> en el Viejo Oeste
Hombre	49	A	Extraterrestres	El usuario simula ser un famoso cantante de rap
Mujer	54	B	Perros	El usuario y el robot descubren un nuevo planeta
Mujer	74	B	Una película	La ciudad en la que vive
Mujer	77	B	Ancianos e IA	Su pueblo natal
Mujer	83	A	Su nieto	Un día en su vida

Tabla 5.1: Detalle sobre las sesiones con usuarios: género, edad, grupo y temas de conversación

5.3. Resultados

A continuación se presentan los resultados de las pruebas de evaluación con usuarios.

5.3.1. Preferencia de voces

Durante la prueba de preferencia de voces, se comprobó (ver figura 5.1) que el **72.22%** de los usuarios prefirió una voz perteneciente al *endpoint* de *Audio Speech* frente a la propia voz de *Sanbot*, siendo los **hombres** los más reacios a elegir la voz del robot (ver figura 5.2).

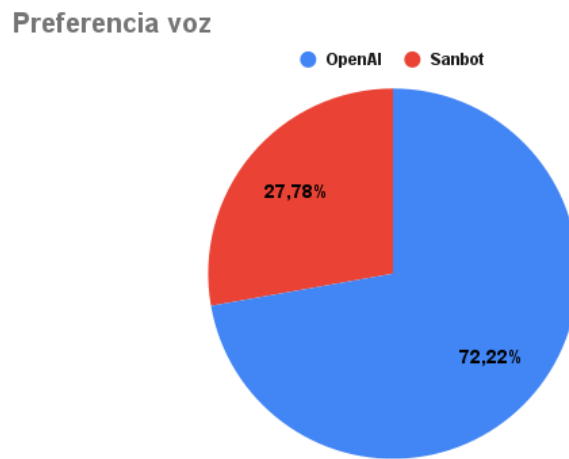


Figura 5.1: Diagrama resultados test preferencia de voz

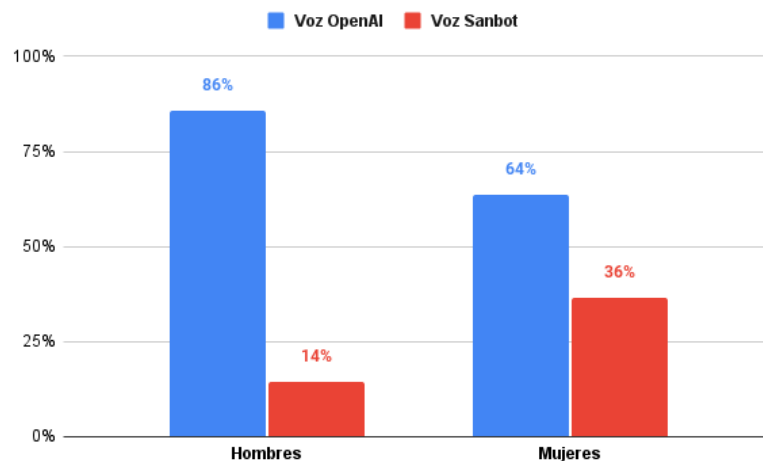


Figura 5.2: Diagrama resultados test preferencia de voz según género

También se pudo comprobar (ver figura 5.3) que el **72.22%** de las personas eligieron una voz que coincidía en género con la del usuario, aunque un mayor número de hombres y de mujeres eligieron una voz **femenina** (ver figura 5.4).



Figura 5.3: Diagrama resultados test preferencia género de voz

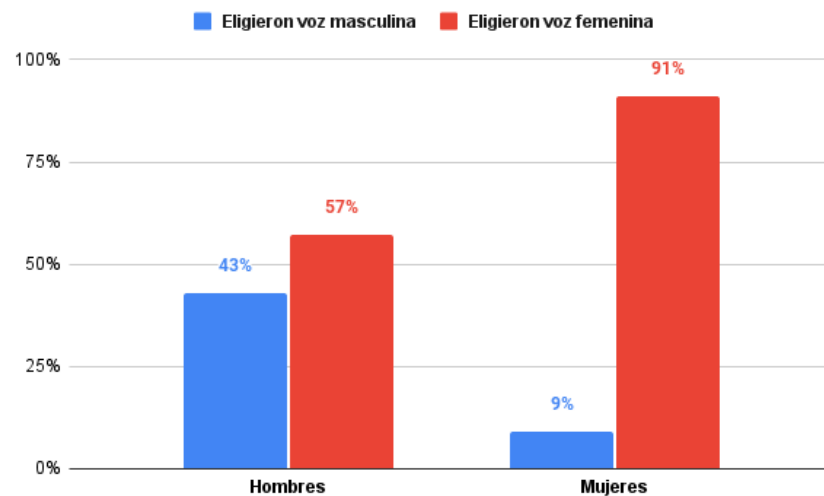


Figura 5.4: Diagrama resultados test preferencia género de voz según género

5.3.2. Usabilidad (Preguntas SUS)

Análisis general

Los resultados del test fueron muy satisfactorios, pudiendo destacar las puntuaciones a preguntas como el uso frecuente del robot o la poca complicación del módulo conversacional (ver figura 5.5).

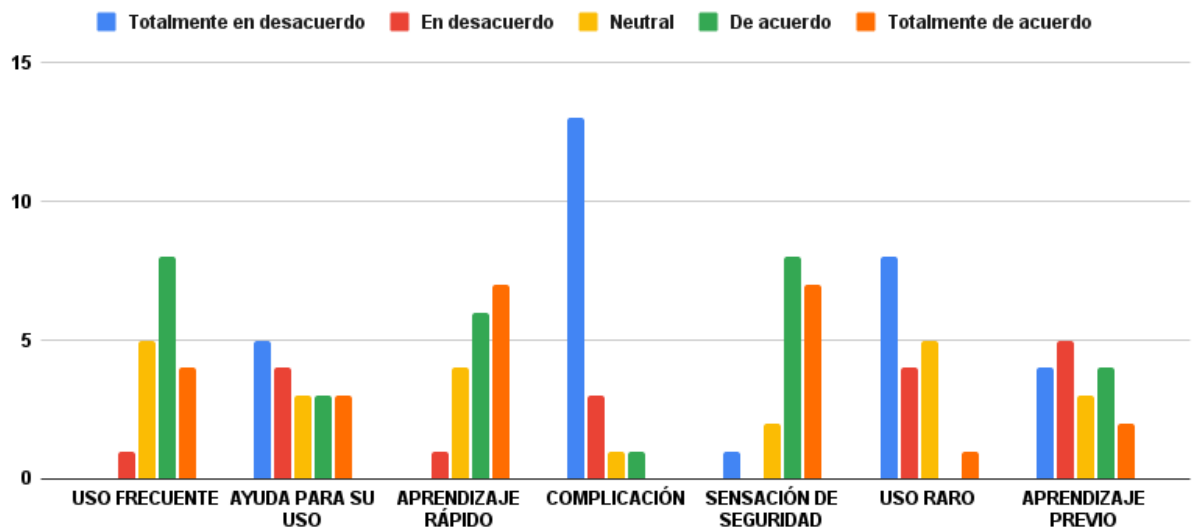


Figura 5.5: Diagrama resultados generales del test usabilidad preguntas SUS

Análisis por género

Aunque en general los resultados son similares, se puede observar en la figura 5.6 que existen algunas diferencias: las mujeres tienen más interés en usar el robot mientras que los hombres califican en mayor medida el uso del robot como raro pero valoran su aprendizaje a la hora de usarlo como más rápido.

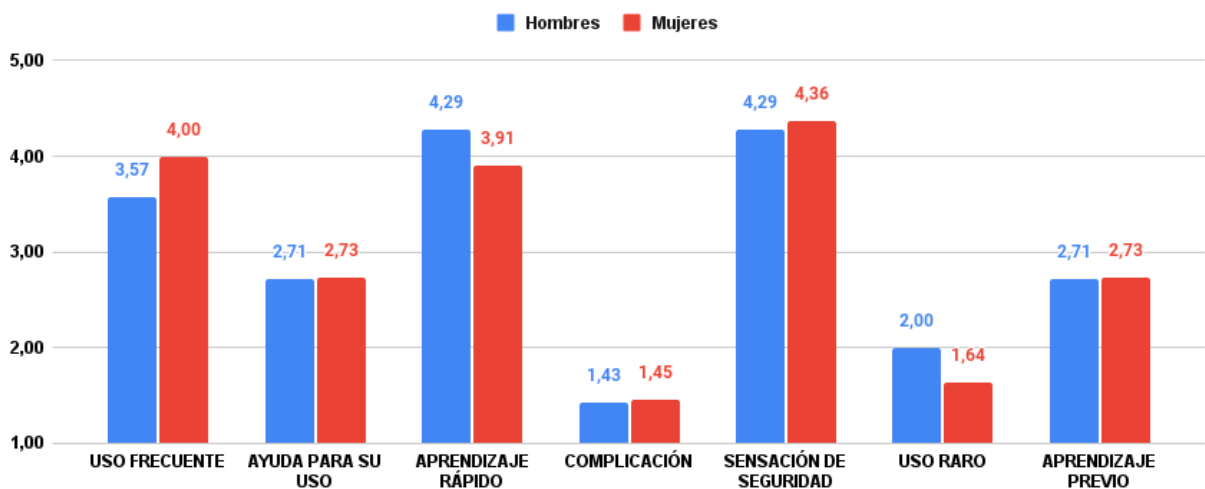


Figura 5.6: Diagrama resultados test usabilidad preguntas SUS según género

Análisis por grupo

A pesar de que no hay diferencias significativas, se puede observar en la figura 5.7 que el grupo A considera que necesitarían más ayuda para utilizar el robot, que su uso es más complicado y raro y que necesitan aprender muchas cosas antes de iniciar el módulo conversacional. Se cree que esto puede ser debido a la necesidad de pasar por las pantallas de personalización tanto del usuario como del robot antes de poder llevar a cabo la conversación con el robot. Sin embargo, consideran que aprenden más rápido a utilizarlo que el grupo B.

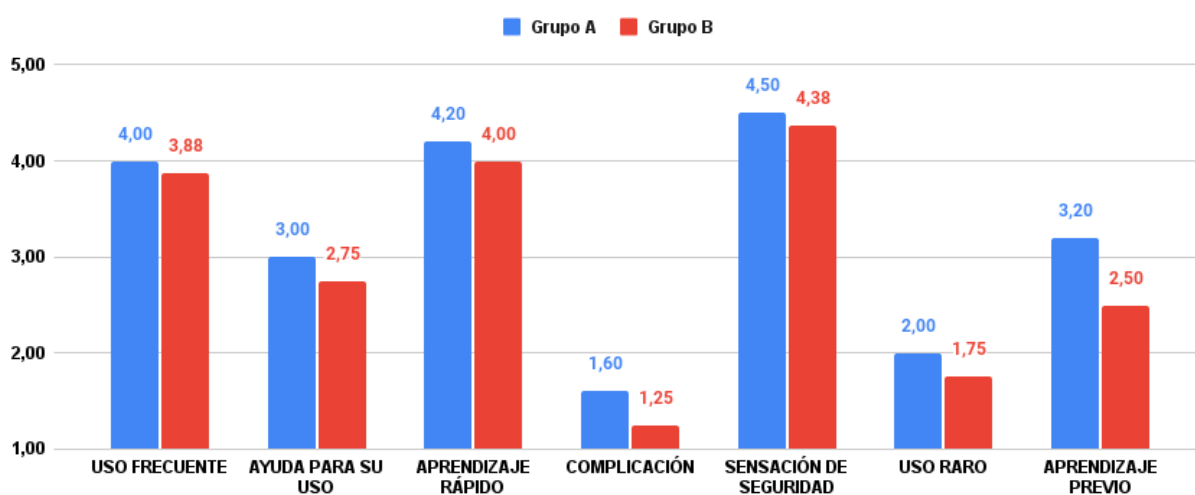


Figura 5.7: Diagrama resultados test usabilidad preguntas SUS según tipo de prueba

5.3.3. Usabilidad (Preguntas interacción voz)

Análisis general

En este test cabe destacar las altas puntuaciones por parte de los usuarios, estando muy satisfechos con la comprensión por parte del usuario y del robot (considerando que tiene una voz entendible), la coherencia de las conversaciones, la respuesta rápida del módulo o la aparición de la respuesta en pantalla. Por otro lado, independientemente de la prueba, los usuarios no han llegado a un consenso respecto a considerar la voz del robot artificial (ver figura 5.8).

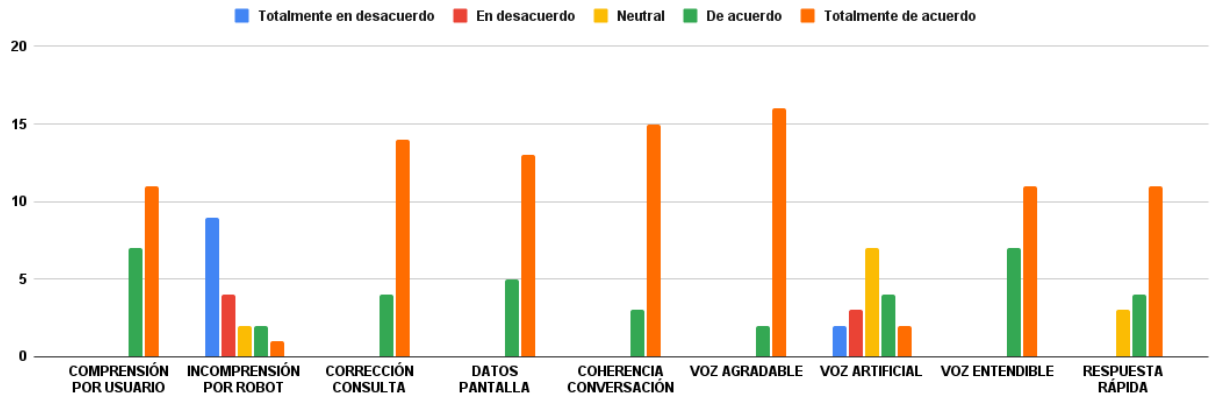


Figura 5.8: Diagrama resultados generales del test usabilidad preguntas interacción por voz

Análisis por género

Como se puede visualizar en la figura 5.9, apenas existen diferencias entre ambos grupos, pudiendo comentar que las mujeres han considerado que el robot ha entendido mejor las consultas del usuario y que sus respuestas eran más rápidas.

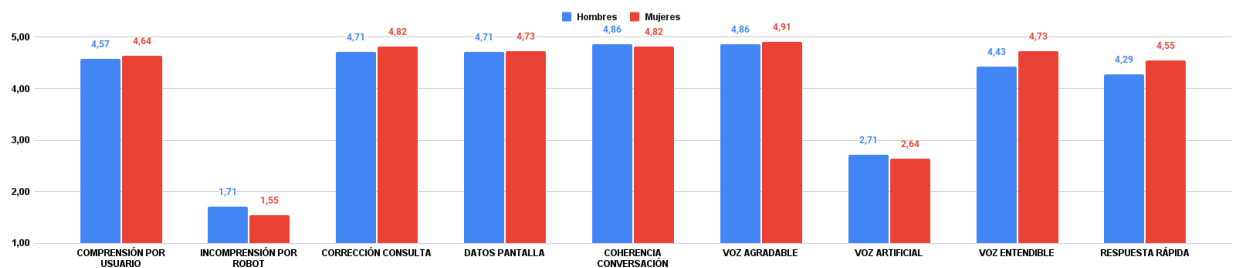


Figura 5.9: Diagrama resultados del test usabilidad preguntas interacción por voz según género

Tal y como se comenta al principio del capítulo, los temas de conversación que los usuarios tuvieron con el robot fueron muy variados, incluyendo conversaciones menos habituales con los temas cerrados que propuso el investigador (Tabla 5.1). Ninguno de los usuarios experimentó incoherencia en la conversación y se puede ver reflejado en la buena puntuación de la pregunta de **coherencia conversación**.

Análisis por grupo

En la figura 5.10 se puede visualizar que el grupo B ha comprendido mejor las respuestas del robot y consideran la voz más entendible, probablemente debido a los problemas de

pronunciación de las voces *OpenAI*, aunque opinaron que el robot entendía peor las consultas debido a que en ese grupo no se aplicaba ninguna personalización. Por último, el grupo A consideró las respuestas del robot más lentas, seguramente debido a que la petición a *OpenAI* contenía más datos de tipo emocional, personalización o similares.

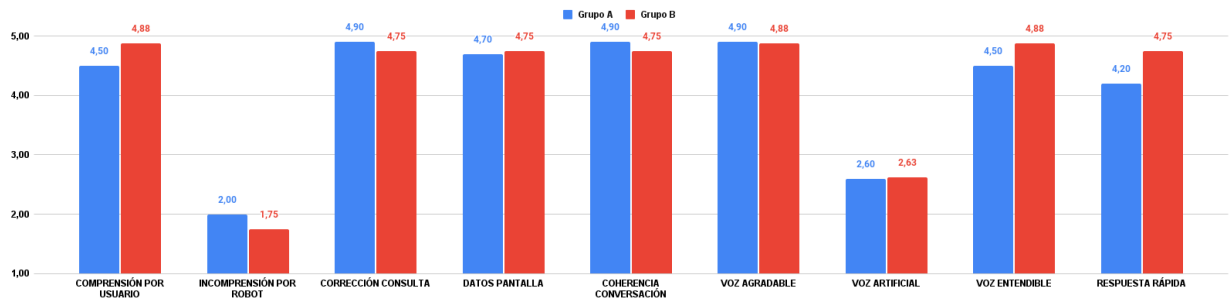


Figura 5.10: Diagrama resultados del test usabilidad preguntas interacción por voz según tipo de prueba

5.3.4. Likeability (Agrado)

Análisis general

Las puntuaciones en este test han sido más dispares al ser preguntas más personales relacionadas con la confianza y opinión que tienen los usuarios acerca del robot. Sin embargo, se valora positivamente el trato amistoso y simpático que encuentran en él (ver figura 5.11).

A pesar del buen resultado, se tienen dudas sobre si todos los usuarios entendieron de la misma forma la pregunta sobre la accesibilidad del robot.

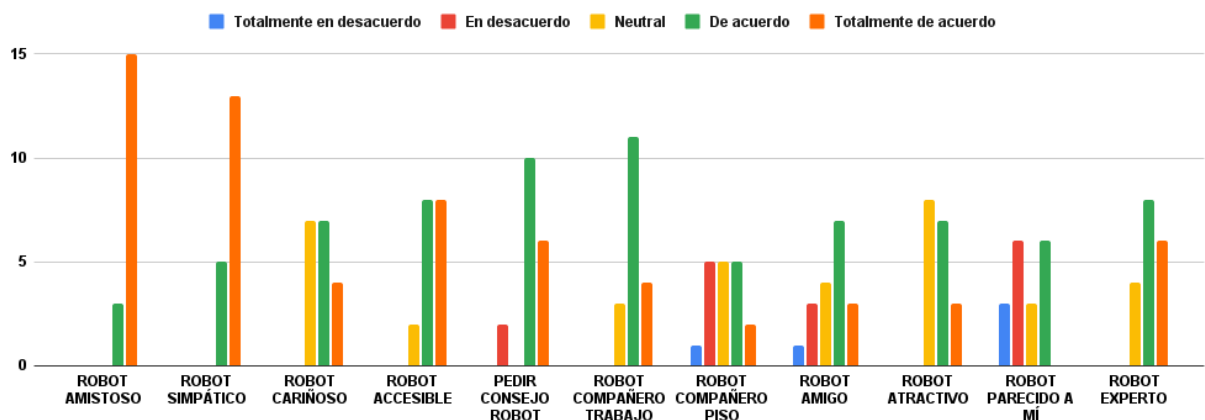


Figura 5.11: Diagrama resultados generales del test de agrado

Análisis por género

Se puede ver en la figura 5.12 que no existen grandes diferencias en la gráfica, pudiendo comentar que los hombres ven al robot más atractivo o lo preferirían como compañero de piso frente a las mujeres, que lo elegirían como compañero de trabajo y lo consideran más experto.

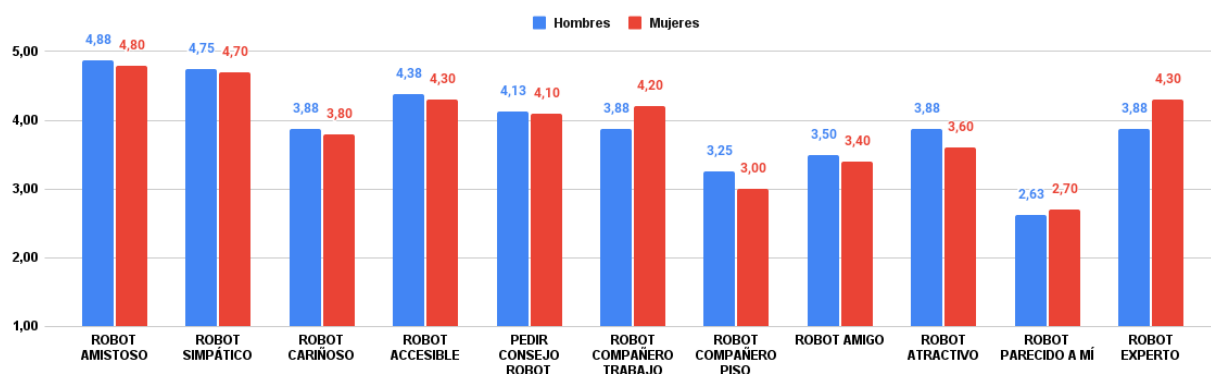


Figura 5.12: Diagrama resultados del test de agrado según género

Análisis por grupo

En la figura 5.13 se puede observar que el grupo A consideró al robot más cariñoso y más parecido a los usuarios, seguramente por la incorporación del módulo emocional. Sin embargo, el grupo B lo utilizaría para pedirle consejo, como compañero de trabajo o de piso o como amigo, causado muy probablemente por la facilidad de configuración que tiene esta prueba al no contar con personalización.

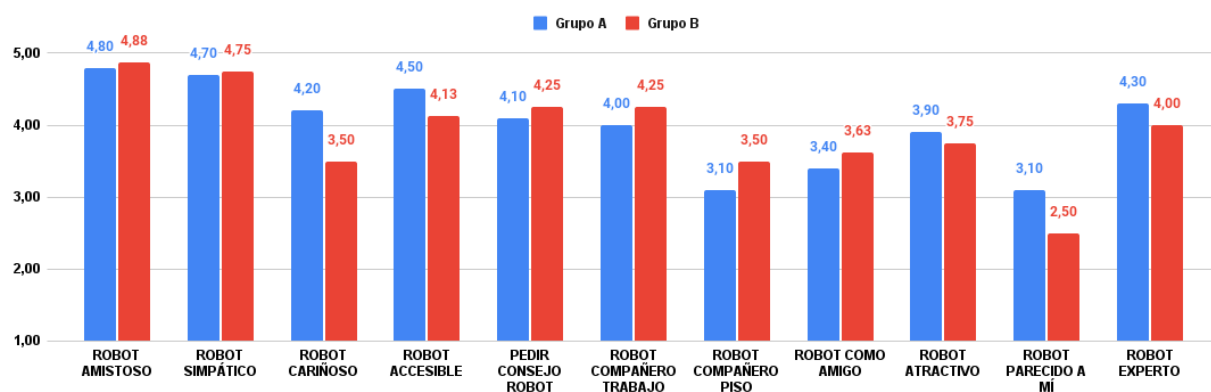


Figura 5.13: Diagrama resultados del test de agrado según tipo de prueba

5.3.5. Aceptación

Análisis general

Los resultados de este test han sido muy satisfactorios, pudiendo observar que los usuarios apenas consideraron el módulo difícil de utilizar a pesar de las diversas edades o experiencia con la tecnología, que muchos de ellos tenían intención de usarlo de nuevo, que lo veían útil en base a sus necesidades y que todos ellos consideraron la prueba muy divertida (ver figura 5.14).

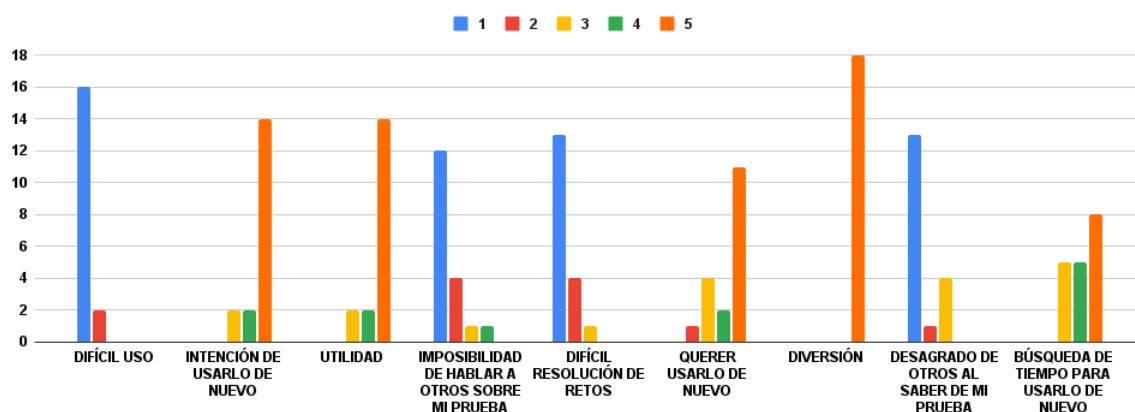


Figura 5.14: Diagrama resultados generales del test de aceptación. Más detalle de las preguntas en el Anexo F

Análisis por género

En la figura 5.15 se puede ver que la diferencia más significativa es que las mujeres encuentran más útil el uso del robot, sin observarse relación con otras variables como la edad o experiencia tecnológica. Por otro lado, en las preguntas de intención de usarlo de nuevo y querer usarlo de nuevo se encuentra cierta contradicción en la que los hombres tienen mayor puntuación en la primera y las mujeres mayor puntuación en la segunda.

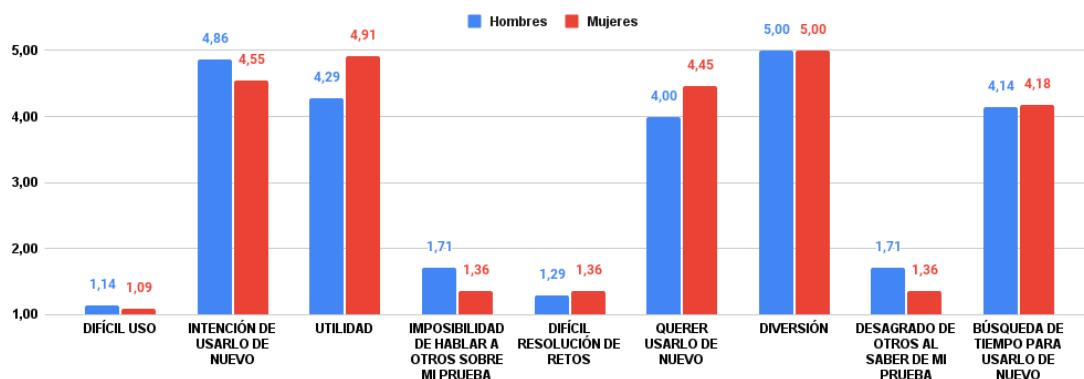


Figura 5.15: Diagrama resultados del test de aceptación según género

Análisis por grupo

En la figura 5.16 se muestra que, de manera muy similar a lo que pasaba con el análisis anterior, el grupo B encuentra el robot más útil y existe cierta contradicción entre las preguntas intención de usarlo de nuevo y querer usarlo de nuevo, lo que hace pensar en la necesidad de reformularlas.

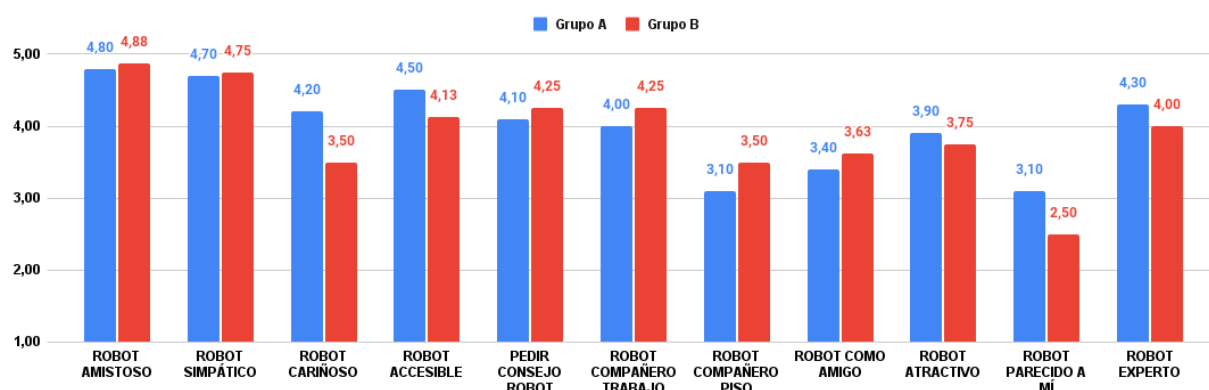


Figura 5.16: Diagrama resultados del test de aceptación según tipo de prueba

5.4. Conclusiones

Además de los resultados comentados en los apartados anteriores, se quiere recalcar que:

- El **100 %** de los usuarios consideraron el uso del robot fácil o muy fácil.
- El **100 %** de los usuarios han considerado la prueba de *Sanbot* como divertida.
- El **100 %** de los usuarios está de acuerdo o muy de acuerdo con que entendieron bien lo que decía el robot.
- El **100 %** de los usuarios está de acuerdo o muy de acuerdo con que la conversación que han tenido con el robot ha seguido un hilo coherente.
- El **100 %** de los usuarios está de acuerdo o muy de acuerdo con que el robot ha entendido correctamente sus consultas y ha sabido darles una respuesta con sentido.
- El **100 %** de los usuarios está de acuerdo o muy de acuerdo con que la aparición de la respuesta en la pantalla era de gran ayuda.
- El **88.9 %** de los usuarios utilizarían el robot para pedirle consejo.

- El **88.9%** de los usuarios considera el robot como una idea útil o muy útil en base a lo que les puede aportar.
- El **75%** de los usuarios está de acuerdo o muy de acuerdo con que la respuesta del robot es rápida.
- El **66.67%** de los usuarios está de acuerdo o muy de acuerdo con que les gustaría utilizar el robot con frecuencia
- Sólo el **16.67%** de los usuarios está de acuerdo o muy de acuerdo con que les costaba entender lo que decía el robot.

En cuanto a la preferencia de voces los participantes se decantan claramente por las voces de *OpenAI* con un género coincidente con el suyo.

En relación al impacto del módulo emocional, si bien incrementa el agrado de los usuarios y usuarias en cuanto a percibirlo como más cariñoso y parecido a ellos mismos y aumenta la sensación de que el robot entiende mejor las consultas, parece afectar a la usabilidad, ya que opinan que necesitarían más ayuda para utilizarlo.

Capítulo 6

Gestión del proyecto

En este capítulo se explica el desarrollo del trabajo clasificado por *Sprints* y el tiempo que se ha dedicado para llevar a cabo este proyecto.

6.1. Desarrollo por Sprints

A lo largo del trabajo se han aplicado diferentes metodologías ágiles, entre ellas el desarrollo por medio de *Sprints*, el cual permitió dividir el proyecto en grupos de tareas, teniendo una reunión con las tutoras tras completarlas para corregir los fallos y comentar el estado del trabajo. La explicación de cada tarea se detalla a continuación:

- **Familiarización robot:** Proceso que incluye tareas como investigar las funciones con las que cuenta el sistema del robot y la comprobación del funcionamiento del SDK.
- **Demo:** Proceso principalmente de implementación, en el que se desarrolló la primera prueba del módulo conversacional para observar las funcionalidades y limitaciones con las que cuenta el robot.
- **Iteración 1:** Proceso que incluye tareas de análisis, diseño e implementación para la utilización del *endpoint Chat Completions* de *OpenAI*.
- **Iteración 2:** Proceso que incluye tareas de análisis, diseño e implementación para la utilización del *endpoint Audio Speech* de *OpenAI*.
- **Iteración 3:** Proceso que incluye tareas de análisis, diseño e implementación para la incorporación del módulo emocional.

6.2. Dedicación

El tiempo dedicado al trabajo se ha recogido mediante la aplicación *Clockify*¹, la cual permite registrar las horas empleadas en un proyecto pudiendo categorizarlo en diferentes tareas.

El uso de esta herramienta ha permitido contabilizar la cantidad de horas que se han invertido en la realización de este trabajo, además, cuenta con múltiples diagramas y gráficos que permiten observar información relacionada con el proyecto.

En la figura 6.2 se muestran las horas totales dedicadas al trabajo y las horas empleadas a realizar cada tarea, siguiendo la misma distribución que el apartado anterior, en la que gran parte del análisis, diseño y sobre todo implementación se llevó a cabo en las iteraciones de la 1 a la 6.

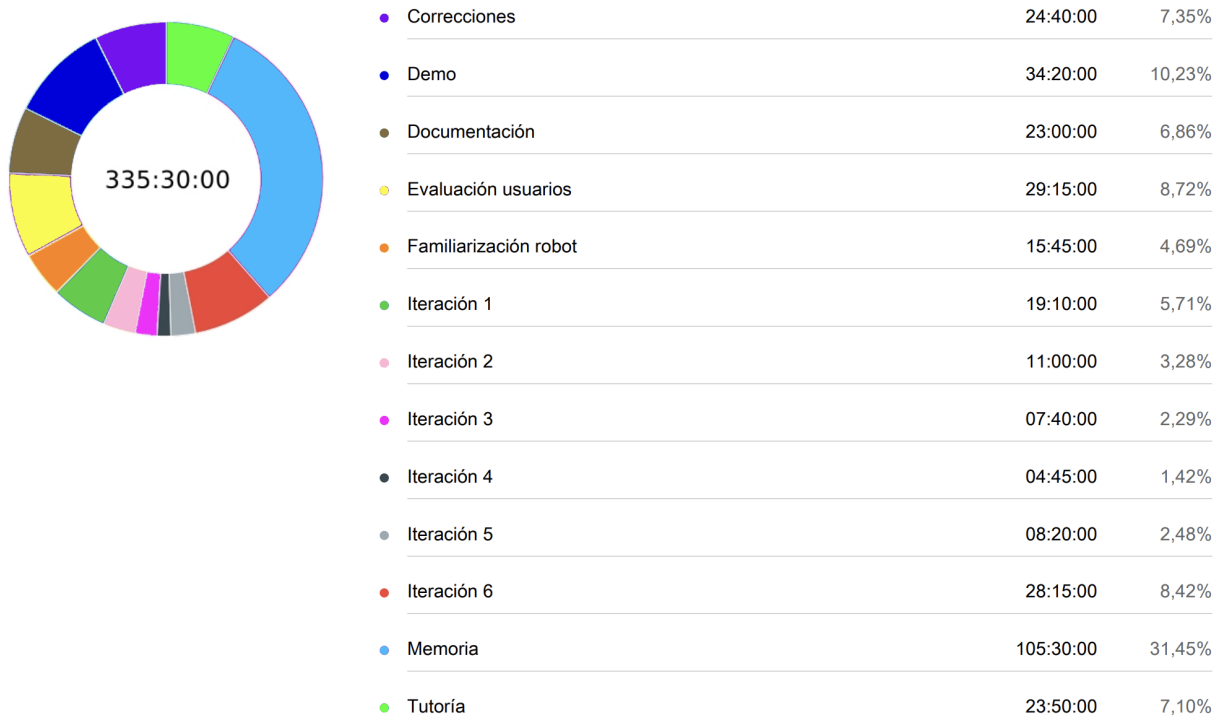


Figura 6.2: Diagrama del tiempo dedicado al proyecto clasificado por tareas

¹<https://clockify.me/es/>

Capítulo 7

Conclusiones y trabajo futuro

En este capítulo se comentan las conclusiones tras el desarrollo del trabajo, las posibles mejoras que se podrían incluir en un futuro y la valoración personal.

7.1. Conclusiones

Se han logrado cumplir los objetivos marcados al principio del trabajo ya que se ha conseguido completar el desarrollo de un módulo conversacional que permite a los usuarios interactuar con el robot *Sanbot Elf* para llevar a cabo una conversación de una forma natural y fluida, teniendo en cuenta el contexto de la conversación y muy centrado en la interpretación emocional del diálogo, además de permitir al usuario modificar y personalizar ciertas características del módulo.

Este módulo ha sido mejorado gracias a la incorporación de herramientas de **Inteligencia Artificial** cuyo uso ha permitido profundizar ampliamente el contexto de la conversación y también ha permitido la utilización de múltiples voces muy similares a la humana, consiguiendo que los usuarios se sientan cómodos al interactuar con el robot.

Además, dando importancia a la comunicación no verbal durante la conversación, se ha incorporado un módulo emocional que, tomando como referencia el modelo de emociones de **Plutchik**, permite reconducir la conversación en función de las emociones que el robot interprete y también consigue que realice acciones de movimiento, encendido de luces LED o cambio de su expresión facial según el sentimiento a representar.

Por último, se ha incluido la posibilidad de que el usuario personalice el módulo conversacional a su gusto, pudiendo introducir datos personales que influirán en la posterior

conversación o incluso la modificación de aspectos como el género o edad del robot que después se verán reflejados en la interacción.

El usuario también puede utilizar el módulo conversacional para consultar información relacionada con la actualidad como la fecha y hora actual o las películas que están actualmente en cartelera.

Para comprobar todas estas funcionalidades se han realizado pruebas de evaluación con usuarios cuyos resultados han sido muy satisfactorios, consiguiendo que las personas disfrutasen y valorasen positivamente la interacción con el robot.

Tras analizar las necesidades, obstáculos y capacidades de las *Personas* definidas en la sección 2.3, se ha puesto especial énfasis en la accesibilidad, pudiendo mencionar que aquellos usuarios no acostumbrados al uso de dispositivos tienen la posibilidad de introducir parámetros por medio de la voz como alternativa al teclado táctil, además de que se han reducido las pantallas deslizables o con excesivos botones. Las personas con problemas o discapacidades sensoriales podrán utilizar el módulo conversacional gracias a las siguientes ayudas: todos los textos aparecen en pantalla en caso de que el usuario tenga problemas de audición, incluso tienen la posibilidad de configurar el volumen de salida de los altavoces del robot; la mayoría de textos son narrados vocalmente por el robot, permitiendo así el uso a personas con problemas o discapacidades visuales y, por último, la alternativa de uso de teclado permite a personas con dificultades del habla seguir utilizando el módulo conversacional.

7.2. Trabajo futuro

Las posibles implementaciones que mejorarían el módulo conversacional son las siguientes:

1. **Simplificación del módulo:** Tal y como se ha podido observar tras realizar la evaluación con usuarios, la posibilidad de personalizar y configurar distintas características del módulo ha provocado una sensación de dificultad a los usuarios, siendo interesante la simplificación del módulo en un futuro.
2. **Identificación de los usuarios:** Sería interesante añadir un sistema de perfiles para identificar el usuario que está utilizando el módulo. De esta manera, se podrían guardar datos relevantes acerca del usuario como nombre de familiares o anécdotas que haya mencionado en conversaciones mediante la incorporación de una base de datos. Esto

podría ayudar a personas a recordar aspectos relacionados con el usuario, pudiendo combatir la mala memoria o incluso enfermedades relacionadas con el Alzheimer.

3. **Utilizar las funcionalidades multimodales del GPT-4o mini:** Tal y como se ha explicado en el apartado 3.1.3, **GPT-4o mini** cuenta con un apartado multimodal para que, además de texto como entrada, se pueden pasar imágenes que la herramienta es capaz de analizar. El módulo conversacional podría ampliarse dando la oportunidad al usuario de enviar imágenes como parte de la conversación.
4. **Ampliación del módulo de peticiones externas:** Sería interesante profundizar en el módulo de peticiones externas añadiendo API que sean capaces de consultar el tiempo meteorológico, noticias de actualidad o similares.
5. **Explotación del modelo de Plutchik:** Ya se ha explicado el funcionamiento de la rueda de emociones de Plutchik y cómo relaciona los sentimientos haciendo uso del modelo. Sin embargo, la relación entre ellos no se ha explotado lo suficiente y simplemente se ha pedido a *OpenAI* que reconozca en la consulta del usuario alguna de las 48 emociones que componen el modelo.

7.3. Valoración personal

La elaboración de este proyecto me ha dado la oportunidad de trabajar con agentes sociales, en este caso, los robots, y conocer su impacto en la sociedad y las dificultades a las que se enfrenta dentro de la comunicación e interacción con los humanos.

Desarrollar el módulo conversacional ha sido una experiencia totalmente nueva al haber trabajado con un dispositivo como *Sanbot Elf* y su kit de desarrollo software, pudiendo observar el impacto que tenía mi programa reflejado en el robot: encendido de LED, movimiento de extremidades o similares.

Sin duda, las pruebas con usuarios ha sido la parte más satisfactoria para mí, pudiendo observar como las personas disfrutaban de las conversaciones con el robot ya sea por sus respuestas o las gesticulaciones que realizaba tras evaluar las consultas del usuario. Esto para mí significa que este sistema realmente se podría utilizar para reducir la soledad en las personas y promover el bienestar social tal y como defiende el proyecto **PLEISAR**.

Todavía hay muchas mejoras que se pueden implementar para conseguir un módulo conversacional mucho más útil y que tenga un impacto mayor, pero estoy satisfecho con el trabajo realizado al haber completado los objetivos marcados al principio del proyecto.

Bibliografía

- [1] Alessandro Marin Vargas, Lorenzo Cominelli, Felice Dell’Orletta, and Enzo Pasquale Scilingo. Verbal Communication in Robotics: A Study on Salient Terms, Research Fields and Trends in the Last Decades Based on a Computational Linguistic Analysis. *Frontiers in computer science*, 2, 2 2021.
- [2] Patented sanbot robot, ai robot design — sanbot robotics. <http://en.sanbot.com/product/sanbot-elf/design>.
- [3] I3A. El grupo de investigación affectivelab y el planetario de huesca estudian posibles vías de colaboración. <https://i3a.unizar.es/es/noticias/el-grupo-de-investigacion-affectivelab-y-el-planetario-de-huesca-estudian-positi>
- [4] Anna Henschel, Guy Laban, and Emily S. Cross. What Makes a Robot Social? A Review of Social Robots from Science Fiction to a Home or Hospital Near You. *Current Robotics Reports*, 2(1):9–19, 2 2021.
- [5] Michael Lewis, Katia Sycara, and Phillip Walker. *The Role of Trust in Human-Robot Interaction*. 1 2018.
- [6] Charles Crowell, Matthias Scheutz, Paul Schermerhorn, and Michael Villano. Gendered voice and robot entities: perceptions and reactions of male and female subjects. *ResearchGate*, 10 2009.
- [7] Franziska Babel, Johannes Kraus, Linda Miller, Matthias Kraus, Nicolas Wagner, Wolfgang Minker, and Martin Baumann. Small Talk with a Robot? The Impact of Dialog Content, Talk Initiative, and Gaze Behavior of a Social Robot on Trust, Acceptance, and Proximity. *International journal of social robotics*, 13(6):1485–1498, 1 2021.
- [8] Andrea Bonarini. Communication in Human-Robot Interaction. *Current Robotics Reports*, 1(4):279–285, 8 2020.

- [9] Marina Sardà Gou, Vasiliki Vouloutsi, Klaudia Grechuta, Stéphane Lallée, and Paul F. M. J. Verschure. *Empathy in humanoid robots*, page 423–426. January 2014.
- [10] Junchao Xu, Joost Broekens, Koen Hindriks, and Mark A. Neerinx. Mood contagion of robot body language in human robot interaction. *Autonomous Agents and Multi-Agent Systems*, 29(6):1216–1248, August 2015.
- [11] Mohammad Shafiquzzaman Bhuiyan. The role of ai-enhanced personalization in customer experiences. *Journal of Computer Science and Technology Studies*, 6(1):162–169, February 2024.
- [12] Pere Munar and Cyberclick. ¿qué son los grandes modelos de lenguaje o llm (large language models)? <https://www.cyberclick.es/numerical-blog/que-son-los-grandes-modelos-de-lenguaje-o-llm-large-language-models>.
- [13] Mailchimp. ¿qué es una prueba a/b? <https://mailchimp.com/es/marketing-glossary/ab-tests/>.
- [14] Cris Busquets and Cris Busquets. Medir la usabilidad con el sistema de escalas de usabilidad (sus). <https://www.uifrommars.com/como-medir-usabilidad-que-es-sus/>, December 2021.
- [15] Stephen Reysen. Construction of a new scale: The reysen likability scale. *Social Behavior and Personality: an international journal*, 33:201–208, 01 2005.
- [16] Qihan Technology. Lanzamiento de prensa del robot Sanbot, 9 2016.
- [17] OpenAI Platform. Uso del endpoint chat completions de openai. <https://platform.openai.com/docs/guides/chat-completions>.
- [18] OpenAI Platform. Uso del endpoint audio speech de openai. <https://platform.openai.com/docs/guides/text-to-speech>.
- [19] Simon Whatley. Plutchik’s wheel of emotion. <https://www.simonwhatley.co.uk/writing/plutchik-wheel-of-emotion/>, April 2013.

Lista de Figuras

1.1.	Robot modelo <i>Sanbot Elf</i> diseñado por <i>Qihan Technology</i>	2
2.1.	Expresiones faciales del robot <i>Sanbot Elf</i>	10
3.1.	Captura de pantalla de la terminal del programa tras experimentar los problemas con el error de clases duplicadas utilizando el SDK <i>SanbotOpenSDK</i>	16
3.2.	Diagrama explicativo del funcionamiento del sistema en profundidad 1	20
3.3.	Diagrama explicativo del funcionamiento del sistema en profundidad 2	21
4.1.	Diagrama del funcionamiento principal del módulo conversacional usando la función <i>Chat Completions</i> de la API de <i>OpenAI</i>	24
4.2.	Diagrama del funcionamiento principal del módulo conversacional usando la función <i>Chat Completions</i> y <i>Audio Speech</i> de la API de <i>OpenAI</i>	26
4.3.	Modelo de rueda de emociones de Plutchik	28
4.4.	Ejemplo de uso del módulo emocional	29
4.5.	Personajes de la película “Del revés” de Disney Pixar	31
4.6.	Capturas del módulo de personalización del usuario	31
4.7.	Captura del módulo de personalización del robot	32
4.8.	Captura de la pantalla de ajustes	34
4.9.	Captura del módulo conversacional con el modo teclado activado	34

4.10. Captura de un paso intermedio del tutorial del módulo conversacional	37
4.11. Captura del módulo conversacional con el modo teclado desactivado	37
5.1. Diagrama resultados test preferencia de voz	42
5.2. Diagrama resultados test preferencia de voz según género	42
5.3. Diagrama resultados test preferencia género de voz	43
5.4. Diagrama resultados test preferencia género de voz según género	43
5.5. Diagrama resultados generales del test usabilidad preguntas SUS	44
5.6. Diagrama resultados test usabilidad preguntas SUS según género	44
5.7. Diagrama resultados test usabilidad preguntas SUS según tipo de prueba	45
5.8. Diagrama resultados generales del test usabilidad preguntas interacción por voz	46
5.9. Diagrama resultados del test usabilidad preguntas interacción por voz según género	46
5.10. Diagrama resultados del test usabilidad preguntas interacción por voz según tipo de prueba	47
5.11. Diagrama resultados generales del test de agrado	47
5.12. Diagrama resultados del test de agrado según género	48
5.13. Diagrama resultados del test de agrado según tipo de prueba	48
5.14. Diagrama resultados generales del test de aceptación. Más detalle de las preguntas en el Anexo F	49
5.15. Diagrama resultados del test de aceptación según género	49
5.16. Diagrama resultados del test de aceptación según tipo de prueba	50
6.1. Diagrama de Gantt con el desarrollo del trabajo dividido en <i>Sprints</i>	54
6.2. Diagrama del tiempo dedicado al proyecto clasificado por tareas	55

A.1. Foto de grupo sacada tras las pruebas del robot <i>Sanbot</i>	72
B.1. Diagrama representativo de los módulos utilizados pertenecientes al <i>QihanOpenSDK</i>	79
C.1. Diagrama que representa los paquetes que forman el módulo <i>robotControl</i>	83
C.2. Diagrama que representa los paquetes que forman el módulo <i>activities</i>	86
C.3. Diagrama que representa los paquetes que forman el módulo <i>modulos</i>	89
C.4. Diagrama que representa los paquetes que forman el módulo <i>conversacion</i>	90
C.5. Diagrama que representa el paquete que forma el módulo <i>utils</i>	90
C.6. Diagrama que representa los paquetes que forman el módulo <i>gestion</i>	92
D.1. Ejemplo de uso del <i>endpoint Chat Completions</i> de <i>OpenAI</i>	94
D.2. Ejemplo de uso del <i>endpoint Audio Speech</i> de <i>OpenAI</i>	96
E.1. Resto de parejas de la combinación de las emociones de la rueda de emociones de Plutchik	98

Lista de Tablas

2.1. Requisitos en forma de historias de usuario para la elaboración del módulo conversacional	14
3.1. Comparación de herramientas de Inteligencia Artificial	17
3.2. Comparación de modelos ofrecidos por la API de <i>OpenAI</i>	19
3.3. Comparación de modelos <i>Text to Speech</i> ofrecidos por la API de <i>OpenAI</i> . .	19
5.1. Detalle sobre las sesiones con usuarios: género, edad, grupo y temas de conversación	41
E.1. Ejemplo de funcionamiento de la rueda de emociones de Plutchik	98

Anexos

Anexo A

Análisis de las funcionalidades del robot Sanbot Elf

En este anexo se explica el proceso de familiarización con las funcionalidades del robot durante las primeras semanas del trabajo y se entra en detalle acerca de la primera versión del módulo conversacional.

A.1. Familiarización del robot

Durante las primeras semanas de trabajo, se ha analizado la documentación del robot para conocer sus funciones básicas y buscado más información a través de foros, páginas relacionadas con los servicios que ofrece la compañía *Qihan* (la mayoría de ellos privados, con solicitud de permiso para acceder o páginas desactualizadas que no llevaban a ningún lado) y unos pocos repositorios con errores y bastante incompletos subidos por parte de usuarios de alrededor del mundo.

Se consiguió desarrollar una pequeña demo en la que poníamos a prueba algunas de las funciones del robot haciendo uso del kit de desarrollo de software encontrado entre la documentación de la compañía, siendo este el SDK *SanbotOpenSdk2.0.1.0*. Entre estas funciones se encontraba: movimiento de extremidades (cabeza, tronco y brazos), movimiento del robot, cambios de la expresión facial, reconocimiento facial del usuario, cambio de color de los LED encontrados en la cabeza o en los brazos, reproducción de textos a través del sintetizador de voz y reconocimiento y transcripción de la voz.

El objetivo de esta primera prueba era mantener lo más parecido a una conversación con el robot, para ello, se hizo un programa en el que *Sanbot* daba respuestas concretas dependiendo de las palabras que reconocía por parte del usuario.

La prueba sí que permitió conseguir una mínima interacción entre el humano y el robot pero predefinir la conversación era claramente **ineficiente**, ya que este proceso fuerza conocer palabras clave que el usuario debe pronunciar y una respuesta cerrada que el robot siempre acabará contestando, por lo tanto, el tema de la conversación debe ser conocido con anterioridad y la duración vendrá delimitada por la pila de respuestas que haya almacenada en el programa.

Esta demostración se evaluó con usuarios pertenecientes al **Planetario de Huesca**, dónde se mencionó como problemas la posible dificultad que conllevaría programar la conversación o la sensación de excesiva velocidad que encontraron en el habla del robot (ver figura A.1).



Figura A.1: Foto de grupo sacada tras las pruebas del robot *Sanbot*

A.2. Primera prueba módulo conversacional

Aquí se incluye un fragmento del código que conforma el programa de la primera prueba del módulo conversacional:

```
1 public void dialogoPlanetario() {
2     // definimos la velocidad y entonación de la conversación
3     SpeakOption speakOption = new SpeakOption();
4     speakOption.setSpeed(60);
5     speakOption.setIntonation(50);
6     // el robot dice el dialogo incial
7     speechManager.startSpeak("Hola, Estoy buscando a Drako,"
8 + "Le habéis visto ? ", speakOption);
9     try {
10        Thread.sleep(4000);
11    } catch (InterruptedException e) {
12        e.printStackTrace();
13    }
14    // empieza el reconocimiento de voz
15    speechManager.doWakeUp();
16    reconocerRespuesta(speakOption);
17 }

1 public boolean reconocerRespuesta(SpeakOption so){
2     speechManager.setOnSpeechListener(new RecognizeListener() {
3         // variable (no utilizada)
4         boolean preguntaEncontrada = false;
5         @Override
6         public boolean onRecognizeResult(Grammar grammar) {
7
8             // paso la gramática reconocida a String
9             String cadenaReconocida = grammar.getText();
10            // paso la cadena a minúsculas
11            cadenaReconocida = cadenaReconocida.toLowerCase();
12            cadenaReconocida = Normalizer.normalize(cadenaReconocida,
13
14            Normalizer.Form.NFD);
15            // elimino las tildes de la cadena reconocida para poder
16            // compararlo cómodamente con cadenas concretas
17            cadenaReconocida = cadenaReconocida.
18            replaceAll("[^\\p{ASCII}]", "");
19
20            // recorro el vector para ver si la pregunta está
21            // dentro de la conversación
22            for(int i=0; i<pregunta_respuesta.length; i++){
23                preguntaEncontrada = false;
24                // expresión regular para ver si la cadena a comparar
25                // está dentro de la cadena reconocida
26                String regex = "\\b" + pregunta_respuesta[i][0] + "\\b";
27                Pattern pattern = Pattern.compile(regex,
28                Pattern.CASE_INSENSITIVE);
29                if(pattern.matcher(cadenaReconocida).find()){
30                    preguntaEncontrada = true;
31                    // si la cadena reconocida es "fin conversación"
32                    // el robot para la conversación
33                    if(cadenaReconocida.equals("fin conversacion")){
34                        speechManager.doSleep();
```

```

35     }
36     // en caso contrario responde la respuesta
37     // correspondiente y su tiempo de espera
38     else{
39         speechManager.startSpeak(pregunta_respuesta[i][1],
40                                 so);
41         try {
42             Thread.sleep(Integer.parseInt
43                             (pregunta_respuesta[i][2])*1000);
44         } catch (InterruptedException e) {
45             throw new RuntimeException(e);
46         }
47         // se vuelve a despertar para
48         // continuar la conversación
49         speechManager.doWakeUp();
50     }
51 }
52 }
53 return true;
54 }

```

La sucesión de palabras reconocidas se comparaban con unos patrones fijos definidos en el programa y que seguían el siguiente formato:

```

1 private String [][] pregunta_respuesta =
2 {
3     {"planeta mas cercano al sol", "Sabéis cuál es ?", "3"},
4     {"mercurio", " Y cómo es Mercurio ? Es el planeta más, ?", "3"},
5     {"pequeno", "Y es el planeta más, ?", "3"},
6     {"rapido", "muy bien", "3"},
7     {"bueno sigamos buscando", "qué planeta viene después ?", "3"},
8     {"tienes razon", "Entonces, qué planeta viene luego ?", "3"},
9     {"marte", " y qué color tiene ?", "3"},
10    {"planetas gaseosos", " gaseosos ? qué quieres decir ?", "3"},
11    {"jupiter", " y qué característica tiene ?", "3"},
12    {"claro", "chicos, cuál es el siguiente ?", "3"},
13    {"frio", "No creo que esté allí, sólo falta uno, cuál es?", "4"},
14    {"pluton", "Ay, sí, es un planeta enano o planetoide", "3"},
15    {"fin conversacion"}
16 };

```

Este vector de *Strings* se compone por otros vectores con tres componentes:

- Primera componente: Cadena que el robot tiene que reconocer entre la respuesta del usuario.
- Segunda componente: Respuesta que el robot debe reproducir tras reconocer la cadena del primer componente.
- Tercera componente: Son enteros en forma de cadena, indican el tiempo que el robot tiene que esperar tras reproducir la respuesta. Se utiliza para no solapar acciones del robot.

Si la cadena reconocida formaba parte de alguno de los ejemplos, daba la respuesta y esperaba los segundos indicados. Después, el robot volvía a ponerse en modo escucha volviendo a repetir el proceso. Si el usuario decía “fin conversación”, el robot detenía la interacción y no seguía interpretando las palabras del usuario.

El ejemplo del código simula una conversación entre un robot y un niño, donde el robot realiza preguntas acerca del sistema solar. Cuando el robot reconoce que el niño ha dicho “mercurio”, responderá con “Y cómo es Mercurio? Es el planeta más?”, y esperará tres segundos antes de volver a reconocer otra respuesta del usuario. Si el niño responde “pequeño”, el robot continuará diciendo “Y es el planeta más?”, y así sucesivamente simulando una conversación.

Anexo B

Módulos QihanOpenSDK

En este anexo se explican los diferentes módulos y funciones que han sido utilizadas en el trabajo y que pertenecen al *QihanOpenSDK*, representado en la figura B.1.

1. **SystemManager**: Es el módulo utilizado para cambiar la expresión facial del robot mediante los LED que tiene como ojos.
 - **showEmotion**: Cambia la expresión facial del robot a aquella que se le pase como parámetro.
2. **HardWareManager**: Es el módulo utilizado para encender y apagar los LED que posee el robot en partes como la cabeza, brazos y base.
 - **setLED**: Permite encender o apagar los LED de determinadas partes del robot.
3. **HeadMotionManager**: Es el módulo utilizado para realizar movimientos con la cabeza del robot.
 - **doRelativeAngleMotion**: Permite indicar la dirección y ángulo a la que la cabeza va a moverse.
 - **doAbsoluteAngleMotion**: Se ha utilizado para recuperar la posición original de la cabeza.
4. **HandMotionManager**: Es el módulo utilizado para realizar movimientos con los brazos del robot.
 - **doAbsoluteAngleMotion**: Permite indicar el brazo o brazos, la dirección y ángulo en la que se van a posicionar.

5. **AudioManager**: Es el módulo utilizado para consultar y modificar el volumen de los altavoces.
- **getStreamVolume**: Permite conocer el volumen de los altavoces del robot.
 - **setStreamVolume**: Permite modificar el volumen de los altavoces del robot.
6. **SpeechManager**: Es el módulo utilizado para la síntesis de voz e interpretación de diálogo.
- **isSpeaking**: Permite saber si el robot está hablando o no.
 - **startSpeak**: Función utilizada para hacer que el robot pronuncie una cadena a una entonación y velocidad pasados como parámetro.
 - **stopSpeak**: Función utilizada para hacer que el robot detenga el habla.
 - **doWakeUp**: Permite poner al robot en modo escucha, lo cual reconocerá lo que el usuario diga hasta que deje de hablar o hasta que el robot entre en modo *Sleep*.
 - **doSleep**: Permite al robot salir del modo escucha.
 - **setOnSpeechListener(new RecognizeListener)**: Permite pasar como parámetro un *RecognizeListener*, es decir, el fragmento que el robot ha escuchado desde que se puso en modo *WakeUp* hasta que entró en modo *Sleep*.
 - **onRecognizeResult**: Función que permite reconocer lo que el usuario ha dicho durante ese fragmento y pasarlo a modo texto.
 - **onRecognizeVolume**: Función que permite conocer el volumen que alcanza el fragmento.
 - **setOnSpeechListener(new SpeakListener)**: Permite pasar como parámetro un *SpeakListener*, es decir, el fragmento en el que el robot está hablando.
 - **onSpeakFinish**: Función que se ejecuta cuando el robot ha terminado de pronunciar la frase.
 - **onSpeakProgress**: Función que se ejecuta mientras el robot continúa pronunciando la frase.
7. **SpeakOption**: Permite modificar la velocidad y entonación de la voz del robot.
- **getSpeed**: Función que permite conocer la velocidad del habla.
 - **setSpeed**: Función que permite modificar la velocidad del habla.
 - **getIntonation**: Función que permite conocer la entonación del habla.
 - **setIntonation**: Función que permite modificar la entonación del habla.

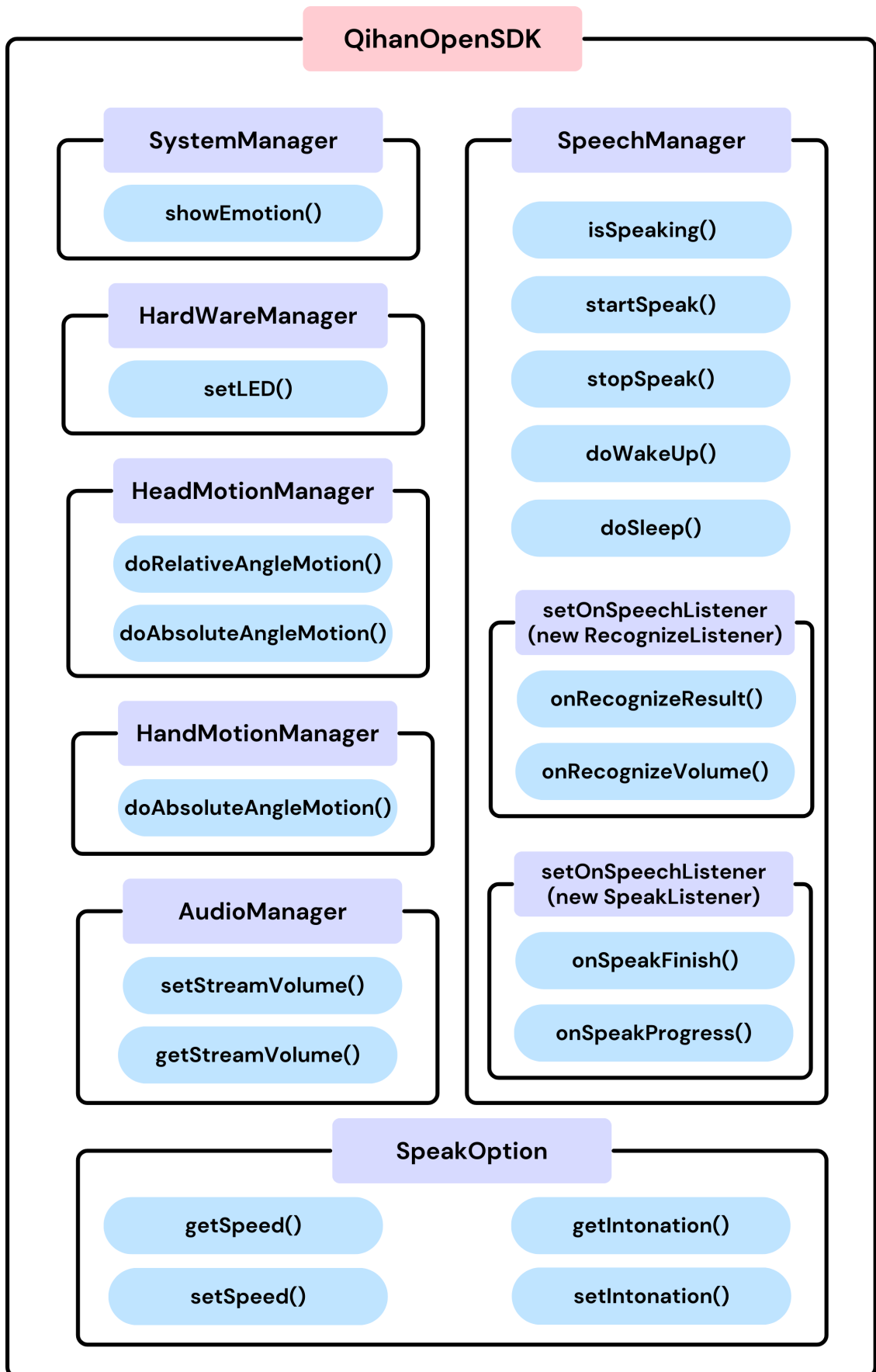


Figura B.1: Diagrama representativo de los módulos utilizados pertenecientes al *QihanOpenSDK*

Anexo C

Explicación de los módulos del Módulo Conversacional

En este anexo se explican los diferentes módulos, funciones y variables utilizadas para el desarrollo del módulo conversacional, los cuales se detallan a continuación:

1. **robotControl**: Es el módulo que incluye todas las funcionalidades pertenecientes al robot *Sanbot Elf* que se han querido utilizar en el trabajo (ver figura C.1).
 - **HandsControl**: Hace uso del *HandMotionManager* explicado en el apartado anterior.
 - **TipoBrazo**: Variable de tipo *enum* que puede tomar tres valores: BRAZO DERECHO, BRAZO IZQUIERDO o AMBOS.
 - **AccionesBrazos**: Variable de tipo *enum* que puede tomar dos valores: SUBIR o BAJAR.
 - **controlBasicoBrazos**: Función que tiene como dos parámetros las dos variables anteriores. Realiza la acción indicada sobre el brazo indicado haciendo uso de las funcionalidades de movimiento del *HandMotionManager*.
 - **reiniciar**: Función que permite volver al estado original al que estaban los brazos del robot, normalmente, bajados.
 - **HardWareControl**: Hace uso del *HardwareManager* explicado en el apartado anterior.
 - **encenderLED**: Función que permite encender los LED de la parte que se indique como primer parámetro y del color que se indique como segundo parámetro.

- **apagarLED**: Función que permite apagar los LED de la parte que se indique como primer parámetro.
- **HeadControl**: Hace uso del *HeadMotionManager* explicado en el apartado anterior.
- **AccionesCabeza**: Variable de tipo enum que puede tomar cinco valores: IZQUIERDA, DERECHA, ABAJO, ARRIBA, CENTRO.
 - **controlBasicoCabeza**: Función que tiene como parámetros la variable anteriores. Realiza la acción indicada haciendo uso de las funcionalidades de movimiento del *HeadMotionManager*.
 - **reiniciar**: Función que permite volver al estado original al que estaba la cabeza del robot, normalmente, centrada.
- **AudioControl**: Hace uso del *AudioManager* explicado en el apartado anterior.
- **getVolumen**: Permite obtener el volumen de los altavoces del robot.
 - **setVolumen**: Permite modificar el volumen de los altavoces del robot al valor pasado como parámetro.
- **SpeechControl**: Hace uso del *SpeechManager* explicado en el apartado anterior.
- **isRobotHablando**: Función que indica si el robot está hablando o no mediante sus funciones de síntesis de voz.
 - **hablar**: Función que permite pronunciar una cadena pasada como parámetro mediante las funciones de síntesis de voz.
 - **pararHabla**: Función que permite detener el habla del robot.
 - **SpeakOption**: Hace uso del *SpeakOption* explicado en el apartado anterior.
 - **getVelocidadHabla**: Permite conocer la velocidad del habla del robot.
 - **getEntonacionHabla**: Permite conocer el nivel de entonación del habla del robot.
 - **setVelocidadHabla**: Permite modificar la velocidad del habla del robot.
 - **setEntonacionHabla**: Permite modificar el nivel de entonación del habla del robot.
 - **modoEscucha**: Función que hace uso de la función *onRecognizeResult* explicado en el apartado anterior.
 - **heAcabado**: Función que hace uso de la función *onSpeakFinish* y espera a que el robot termine de hablar.
- **SystemControl**: Hace uso del *SystemManager* explicado en el apartado anterior.

- **cambiarEmocion**: Función que permite cambiar la expresión facial del robot a la indicada como parámetro.

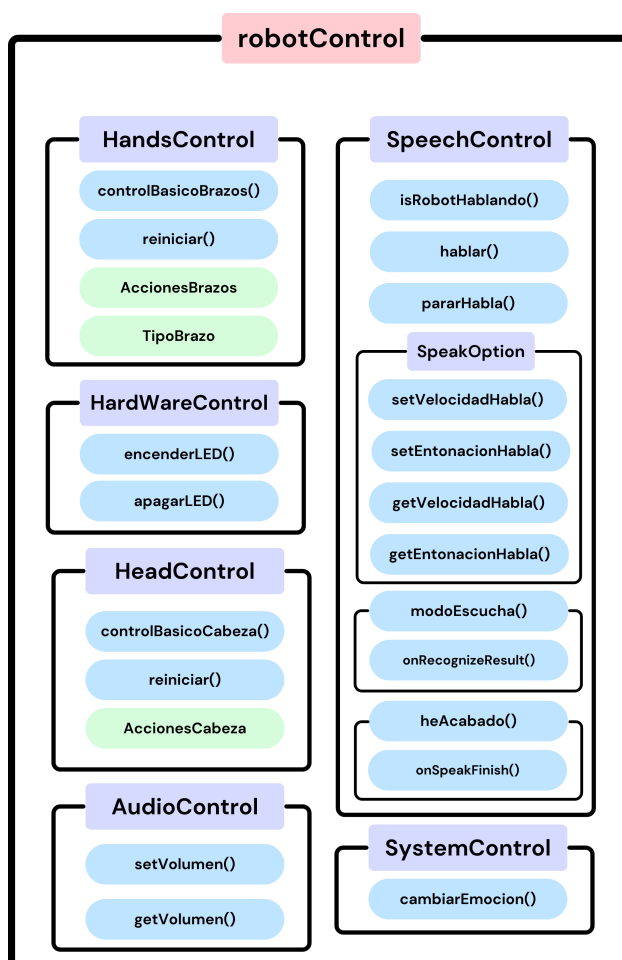


Figura C.1: Diagrama que representa los paquetes que forman el módulo *robotControl*

2. **activities**: Es el módulo que incluye todas las actividades actualmente existentes dentro del módulo conversacional (ver figura C.2).
 - **MenuConfiguracionActivity**: Pantalla que permite elegir entre pasar a la pantalla de *AjustesActivity*, *PersonalizacionRobotActivity* o *NombreUsuarioActivity*.
 - **AjustesActivity**: Permite modificar el volumen, entonación o velocidad del habla del robot. Además permite activar o desactivar el modo teclado, que incluye la opción de enviar consultas por medio de texto en el módulo conversacional.
 - **PersonalizacionRobotActivity**: Pantalla que permite modificar la voz que utilizará el robot, el género que simulará tener, el grupo de edad al que simulará pertenecer y permite añadir un contexto previo a la conversación.

- **NombreUsuarioActivity**: Pantalla que permite al usuario introducir el nombre con el que se presentará ante el robot.
- **EdadUsuarioActivity**: Pantalla que permite usuario introducir la edad con la que se presentará ante el robot.
- **TutorialModuloConversacionalActivity**: Pantalla que muestra una sucesión de pasos dónde el robot explicará mediante la síntesis de voz y capturas de pantalla el funcionamiento principal del módulo.
 - **update**: Función que permite mostrar el siguiente o anterior paso del tutorial. Lo cual mostrará dependiendo del paso en el que se encuentre el usuario una determinada explicación o captura de pantalla.
- **ModuloConversacionalActivity**: Pantalla principal del módulo conversacional.
 - **registrarConsulta**: Función que permite reconocer la consulta que el usuario ha dicho mediante el habla.
 - **reconocerConsulta**: Función que permite reconocer si la consulta del usuario va a ser derivada a la API de *OpenAI* o al contrario, a otro tipo de API.
 - **enviarConsulta**: Función que muestra la consulta del usuario en la pantalla.
 - **clasificarConsulta**: Función que, si la consulta no es de tipo *OpenAI*, utilizará las funciones definidas en el módulo *ModuloPeticonesExternas* para satisfacer la consulta del usuario.
 - **consultaChatCompletions**: Función que obtiene la respuesta de la consulta tras utilizar la API de *OpenAI*, la muestra por pantalla y pide al robot que la pronuncie utilizando la síntesis de voz.
 - **accionReproduccionVoz**: Variable de tipo enum que puede tomar los valores HABLAR, REPETIR O DETENER.
 - **gestionVoz**: Función que como primer parámetro toma la voz con la que el usuario desea realizar la acción y como segundo parámetro toma una variable de tipo *accionReproduccionVoz*. Si la acción es hablar y la voz es la propia del robot Sanbot, reproducirá la consulta a través de los altavoces y, cuando termine, volverá a poner el robot en modo escucha, en caso contrario, utilizará la API de *OpenAI* para dar la respuesta con una de las voces sintetizadas que ofrece la plataforma. Si la acción es repetir utilizará de nuevo las funciones de habla del robot (si la voz seleccionada es *Sanbot*) o reproducirá de nuevo el reproductor de media (si la voz seleccionada es *OpenAI*). Si la acción es

detener, utilizará las funciones de habla del robot para parar el habla (si la voz seleccionada es *Sanbot*) o detendrá la reproducción del reproductor de media (si la voz seleccionada es *OpenAI*).

- **capitalizeCadena:** Función utilizada para capitalizar la primera letra de los mensajes que aparecen en pantalla.
- **gestionarFinHablaSanbot:** Función utilizada para que, después de que *Sanbot* deje de pronunciar una cadena, vuelva a ponerse en modo escucha.
- **gestionarFinReproduccionMediaPlayer:** Función muy similar a la anterior pero para voces de tipo *OpenAI*, ya que su reproducción se gestiona por medio del *MediaPlayer*. Así que, cuando terminen de pronunciar una cadena, el robot volverá a ponerse en modo escucha.
- **recuperarConversacion:** En caso de que el usuario cambie de pantalla y acabe regresando a la pantalla del módulo, esta función gestionará la recuperación de la conversación para volver a mostrarla en pantalla.
- **actualizarVistaConversacion:** Función que permite situarse en los últimos items del *ChatArrayAdapter*, es decir, mostrando las respuestas de las últimas consultas realizadas por el usuario.
- **gestionarPantallaModoTeclado:** Función que, si el modo teclado está activado, mostrara una barra para introducir texto y un botón para enviar la consulta escrita cuando el usuario lo desee. Si el modo teclado está desactivado estos dos componentes se ocultarán, mostrando únicamente el botón hablar para grabar las consultas mediante la voz.
- **gestionModuloConversacional:** Función que permite conocer qué datos ha introducido el usuario en las pantallas de personalización y adaptar la conversación en función de estos.
- **clasificarRoleSystem:** Función que indica el contexto de la conversación en función del resultado de la función anterior, es decir, si el usuario no ha rellenado el nombre y la edad, no podrá indicar a la API de *OpenAI* estos datos, pero sí que lo añadirá como premisa en caso de que el usuario los haya rellenado.

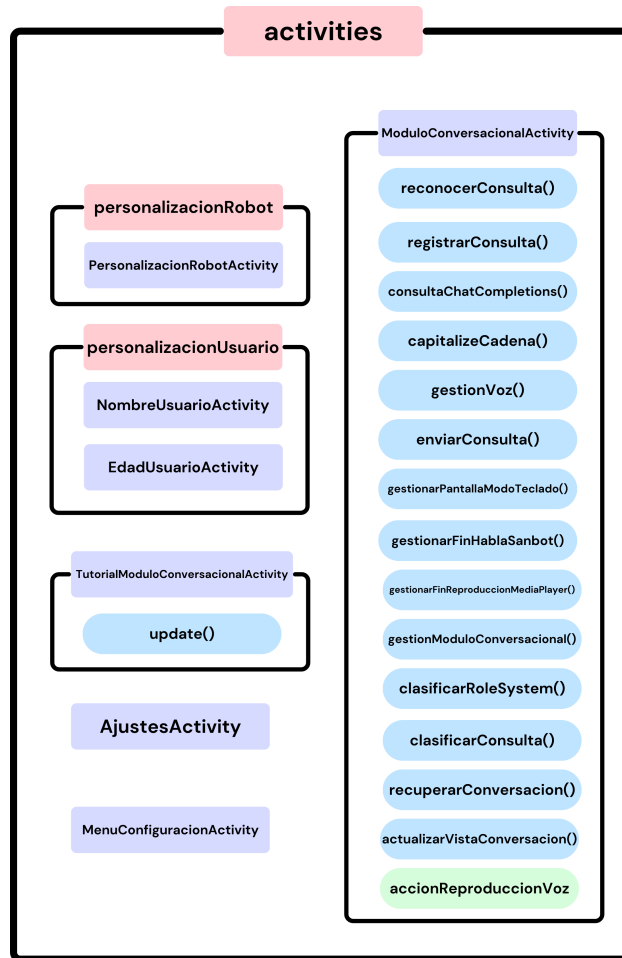


Figura C.2: Diagrama que representa los paquetes que forman el módulo *activities*

3. **modulos**: Conjunto de módulos de mayor tamaño o complejidad que están presentes en el trabajo (ver figura C.3).

– **moduloOpenAI**: Módulo que contiene los dos módulos pertenecientes a los dos *endpoints* de la API de *OpenAI* utilizados en este proyecto.

- **ModuloOpenAIChatCompletions**: Módulo que utiliza las funcionalidades de *Chat Completions* que ofrece la API de *OpenAI*.

- **anadirRoleSystem**: Permite añadir los datos pertenecientes al *Role System* del *endpoint* de *Chat Completions*. El *Role System* permite añadir un contexto que siempre estará presente durante las consultas que se hagan a la API, como por ejemplo, “quiero tener una conversación contigo y simular que somos aventureros”.

- **clearRoleSystem**: Permite eliminar los datos que contiene el *Role System*.

- **anadirRoleUser**: Función que permite añadir datos pertenecientes al *Role User* del *endpoint* de *Chat Completions*. El *Role User* es

- normalmente, el papel que tomará el usuario, aquí aparecerán las consultas que el usuario vaya preguntando a la API, como por ejemplo, “Mira, encontré una joya dentro del cofre”.
- **anadirRoleAssistant**: Función que permite añadir datos pertenecientes al *Role Assistant* del *endpoint* de *Chat Completions*. El *Role Assistant* es normalmente, el papel que tomará el LLM, aquí aparecerán las respuestas a las consultas que haga el usuario, como por ejemplo, “¡Increíble, seguro que está valorada en mucho dinero!”
 - **consultaOpenAI**: Función que realiza la consulta a la API de *OpenAI* utilizando el *endpoint* de *Chat Completions*, por medio de una petición HTTP. Esta petición devolverá un apartado *Role Assistant* con la respuesta del LLM. En la petición se añadirá el modelo de LLM que se desea utilizar, la lista de mensajes intercambiados entre el LLM y el usuario (nunca se llegará a superar la ventana de contexto mencionada en 3.1.3)
 - **construirConversacion**: Función que permite ir añadiendo las consultas que realiza el usuario y las respuestas que da el LLM, para conseguir que se vaya almacenando el hilo de la conversación.
 - **clearConversacion**: Función que permite borrar el hilo de la conversación.
 - **getRespuestaGPT**: Función que permite conocer la respuesta a la última consulta que el usuario ha realizado.
 - **ModuloOpenAIAudioSpeech**: Módulo que utiliza las funcionalidades de *AudioSpeech* que ofrece la API de *OpenAI*.
 - **peticionVozOpenAI**: Función que realiza la consulta a la API de *OpenAI* utilizando el *endpoint* de *Audio Speech*, por medio de una petición HTTP. Esta petición devolverá una ristra de bytes con la respuesta sonora al pronunciar la cadena pasada como primer parámetro utilizando el tipo de voz indicada como segundo parámetro.
 - **getGPTVoz**: Función que devuelve la ristra de bytes de la última consulta realizada a la API.
 - **ModuloEmocional**: Módulo que permite gestionar las emociones interpretadas por el LLM utilizado.
 - **emociones**: Vector que contiene el conjunto de 48 emociones presentes en el modelo de Plutchik.

- **separarRespuestaGPT**: Función que separa la respuesta a la consulta que ha hecho el usuario de la cadena extra que contiene la codificación de las emociones interpretadas por el LLM.
 - **separarEmociones**: Función que, de esa codificación ya separada de la respuesta de la consulta del usuario. Vuelve a separar en las emociones interpretadas en la consulta del usuario de las emociones que considera que tiene que transmitir el robot con su siguiente respuesta.
 - **gestionEmocional**: Función que permite mapear los números codificados en la respuesta del LLM dentro del vector de emociones ya definido.
 - **sentimientosRobot**: Función que añade los números que representan las emociones que va a transmitir el robot según el modelo de Plutchik a una lista.
 - **sentimientosUsuario**: Función que añade los números que representan las emociones que ha interpretado en la consulta del usuario según el modelo de Plutchik a otra lista.
 - **expresividadGestual**: Función que, según el sentimiento que el robot desea transmitir, el robot cambiará su expresión facial, moverá la cabeza, moverá los brazos o encenderá los LED de su cuerpo.
- **ModuloPeticonesExternas**: Módulo utilizado para gestionar peticiones que no van dirigidas a la API de *OpenAI*.
- **funcionesRobot**: Función que clasifica peticiones que el usuario ha comenzado pronunciando como primera palabra “Robot”.
 - **diaYHora**: Función que obtiene la fecha y hora actual.
 - **peliculasAPI**: Función que realiza una petición a la API de *TMDB* para obtener las películas que actualmente se encuentran en cartelera.

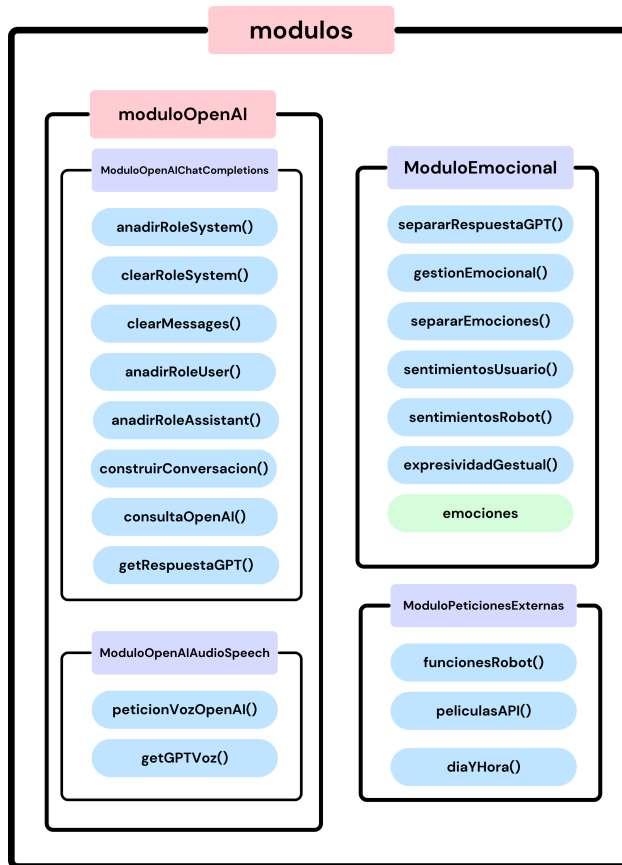


Figura C.3: Diagrama que representa los paquetes que forman el módulo *modulos*

4. **conversacion**: Módulo que permite gestionar el chat entre el LLM y el usuario que aparecerá en el centro de la pantalla del módulo conversacional (ver figura C.4).

- **MensajeChat**: Clase que representa un mensaje individual.
 - **left**: Atributo booleano que, si es true, indica que el mensaje es entrante (enviado por el LLM) y si es false, indica que el mensaje es saliente (enviado por el usuario).
 - **message**: Atributo de tipo String que indica el texto que aparecerá en el mensaje.
- **ChatArrayAdapter**: Componente de tipo *ArrayAdapter* utilizado para crear chats usando la clase *MensajeChat* definida anteriormente.
 - **add**: Función que permite añadir un mensaje de tipo *MensajeChat* al Adapter.
 - **getCount**: Función que indica el número de mensajes que existen en el chat.
 - **getItem**: Función que devuelve el mensaje situado en el índice pasado como parámetro.

- **getView**: Función que permite mostrar personalizar la vista dónde se encuentra el *Adapter*.
- **clear**: Función que vacía los mensajes del chat.

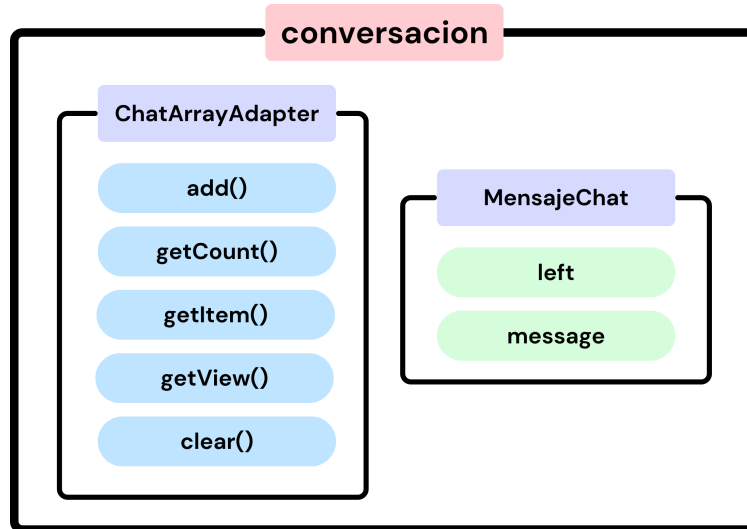


Figura C.4: Diagrama que representa los paquetes que forman el módulo *conversacion*

5. **utils**: Módulo que incluye módulos extra que son necesarios para determinadas acciones del módulo conversacional (ver figura C.5).
 - **ByteArrayMediaDataSource**: El componente *MediaPlayer* utilizado en el proyecto para reproducir las voces de la API de *OpenAI*, permite trabajar con archivos de tipo audio como pueden ser *mp3*, *wav*, etc. Sin embargo, la API de *OpenAI* nos devuelve como respuesta una ristra de bytes. Este módulo permite trabajar con la ristra de bytes de una manera similar a los archivos de audio mencionados, pudiendo utilizar el *MediaPlayer* sin problemas.
 - **readAt**: Función que permite leer el byte en el índice pasado como parámetro para poder realizar acciones con el *MediaPlayer*.
 - **getSize**: Función que permite conocer el tamaño de la ristra de bytes.

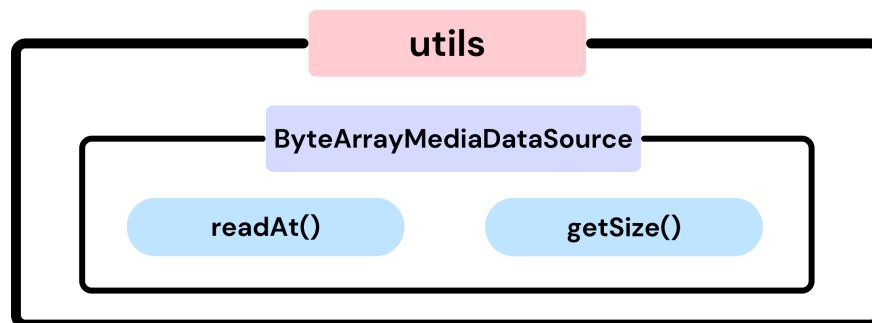


Figura C.5: Diagrama que representa el paquete que forma el módulo *utils*

6. **gestion**: Módulo que permite gestionar relevantes componentes utilizados en el trabajo (ver figura C.6).

- **GestionMediaPlayer**: Módulo que permite realizar acciones con el componente *MediaPlayer* de *Android*.
 - **isMediaPlayerReproduciendose**: Función que indica si hay un *MediaPlayer* reproduciéndose.
 - **reproducirMediaPlayer**: Función que permite reproducir la ristra de bytes que se le pasa como parámetro.
 - **pararMediaPlayer**: Función que permite detener la reproducción de un *MediaPlayer*.
 - **heAcabado**: Función que espera a que el *MediaPlayer* termine.
- **GestionSharedPreferences**: Módulo que permite realizar acciones con el componente *SharedPreferences* de *Android*, lo cual permite obtener o guardar valores en el almacenamiento local del programa.
 - **getStringSharedPreferences**: Permite obtener el valor de tipo *String* que está almacenado en la *SharedPreferences* cuyo nombre se pasa como primer parámetro. En caso de que no exista, se pondrá como valor por defecto aquello que se pase como segundo parámetro.
 - **getIntSharedPreferences**: Permite obtener el valor de tipo *int* que está almacenado en la *SharedPreferences* cuyo nombre se pasa como primer parámetro. En caso de que no exista, se pondrá como valor por defecto aquello que se pase como segundo parámetro.
 - **getBooleanSharedPreferences**: Permite obtener el valor de tipo *Boolean* que está almacenado en la *SharedPreferences* cuyo nombre se pasa como primer parámetro. En caso de que no exista, se pondrá como valor por defecto aquello que se pase como segundo parámetro.
 - **putStringSharedPreferences**: Permite almacenar el valor de tipo *String* en la *SharedPreferences* cuyo nombre se indica como parámetro.
 - **putIntSharedPreferences**: Permite almacenar el valor de tipo *int* en la *SharedPreferences* cuyo nombre se indica como parámetro.
 - **putBooleanSharedPreferences**: Permite almacenar el valor de tipo *Boolean* en la *SharedPreferences* cuyo nombre se indica como parámetro.
 - **clearSharedPreferences**: Permite vaciar los valores existentes en la *SharedPreferences* cuyo nombre se indica como parámetro.

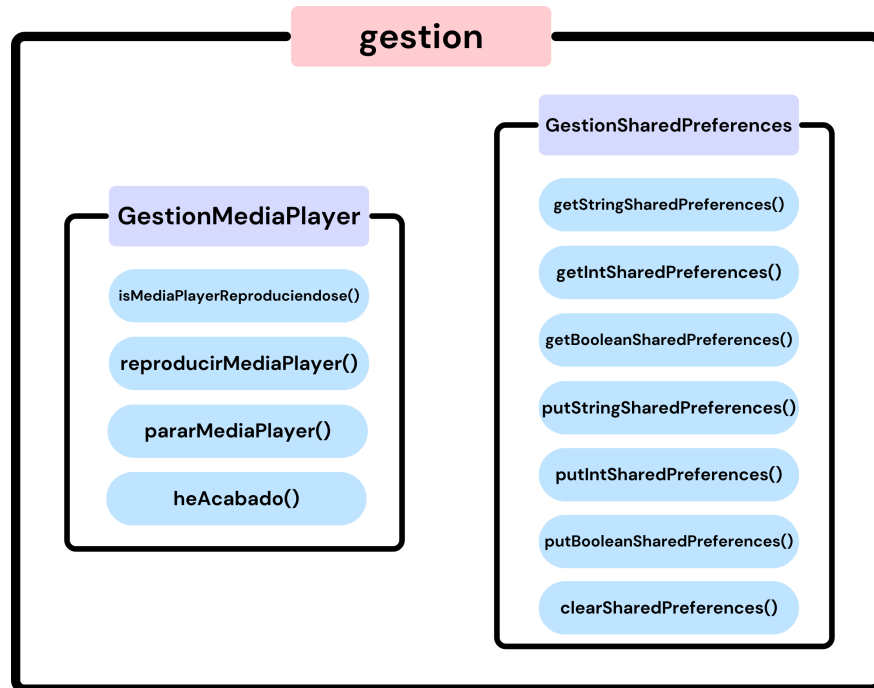


Figura C.6: Diagrama que representa los paquetes que forman el módulo *gestion*

Anexo D

Endpoints de OpenAI

En este anexo se explica en detalle el funcionamiento de los dos *endpoints* de *OpenAI* utilizados en el trabajo.

D.1. Chat Completions

El *endpoint* de *Chat Completions* [17] se utiliza para recibir respuestas a consultas utilizando funciones de grandes modelos de lenguaje.

Esta API soporta la entrada tanto de texto como de imagen en ciertos modelos que ofrece *OpenAI* y como salida devuelve un objeto *JSON* con el contenido que el usuario puede necesitar, entre ellos, la respuesta a la consulta.

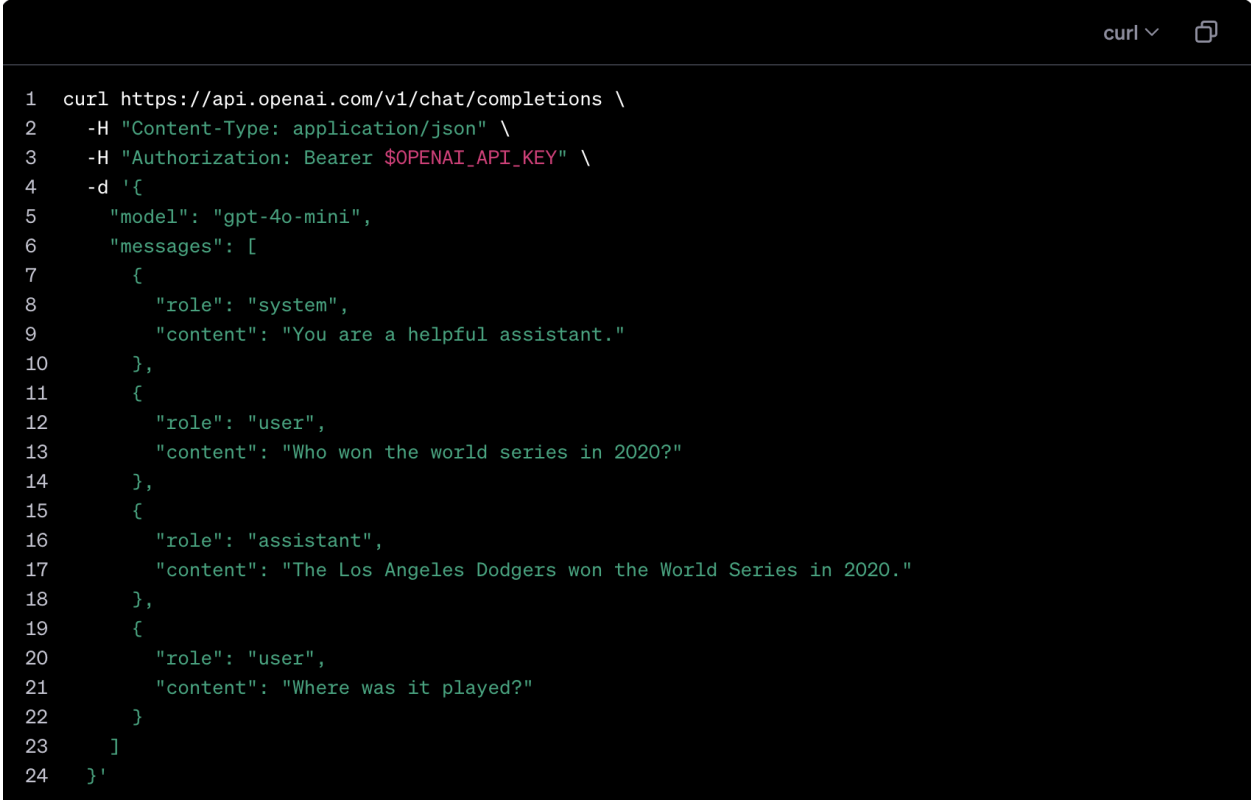
Este *endpoint* funciona mediante mensajes que siguen un sistema de roles los cuales pueden ser:

- **Rol de sistema (Role System)**: Es opcional y se utiliza para determinar el comportamiento del asistente a utilizar.
- **Rol de usuario (Role User)**: Son las consultas o comentarios que el usuario puede enviar al asistente con la intención de que sean respondidas.
- **Rol de asistente (Role Assistant)**: Se utiliza para guardar respuestas previas del propio asistente pero también puede ser modificada por el programador para definir un comportamiento esperado.

Estos mensajes funcionan exactamente igual que funcionaría el *ChatGPT*, si quieres preguntar acerca de un tema concreto las consultas del usuario se almacenarán en los mensajes bajo un rol de usuario y aquellas respuestas por parte del asistente se almacenarán bajo un rol de asistente.

Sin embargo, sólo se guardarán los últimos mensajes en caso de superar la ventana de contexto mencionada en la sección 3.1.3 y que depende del tipo de modelo utilizado. Si no se supera esa ventana de contexto, se guardarán el conjunto de mensajes por completo, pero si este límite se supera, se irán desechando los primeros mensajes para ir almacenando los mensajes entrantes más recientes.

En la figura D.1 se puede observar un ejemplo sencillo de uso de la API:



```
1 curl https://api.openai.com/v1/chat/completions \  
2 -H "Content-Type: application/json" \  
3 -H "Authorization: Bearer $OPENAI_API_KEY" \  
4 -d '{  
5   "model": "gpt-4o-mini",  
6   "messages": [  
7     {  
8       "role": "system",  
9       "content": "You are a helpful assistant."  
10    },  
11    {  
12      "role": "user",  
13      "content": "Who won the world series in 2020?"  
14    },  
15    {  
16      "role": "assistant",  
17      "content": "The Los Angeles Dodgers won the World Series in 2020."  
18    },  
19    {  
20      "role": "user",  
21      "content": "Where was it played?"  
22    }  
23  ]  
24 }'
```

Figura D.1: Ejemplo de uso del *endpoint Chat Completions* de *OpenAI*

Uso del endpoint en el módulo conversacional

El uso de esta funcionalidad en el sistema es muy simple, todas las consultas o comentarios que el usuario transmitía al robot se añadían al conjunto de mensajes bajo el rol de usuario y aquellas respuestas por parte del robot se añadían al conjunto de mensajes bajo el rol de asistente, de esta forma se construía una conversación.

Antes de empezar una conversación se añadía el contexto de esta en función de los datos que habían sido rellenados por el usuario mediante un mensaje bajo el rol de sistema, a continuación se presentan algunos ejemplos:

- Si el usuario no había rellenado sus datos personales ni datos respectivos a la personalización del robot, el mensaje con rol de sistema que se añadía era el siguiente: “quiero que mantengamos una conversación”.
- Si el usuario sí había rellenado sus datos personales, al mensaje anterior también se añadía: “También quiero que a veces me llames por mi nombre que es [**nombreUsuario**] y que adaptes la conversación teniendo en cuenta que mi edad es de [**edadUsuario**] años”.
- Si el usuario además había rellenado los datos de personalización del robot, al mensaje anterior también se añadía: “Además quiero que actúes como que tu genero es [**generoRobot**], que perteneces al grupo de edad [**grupoEdadRobot**] y además [**contextoConversacion**]”.

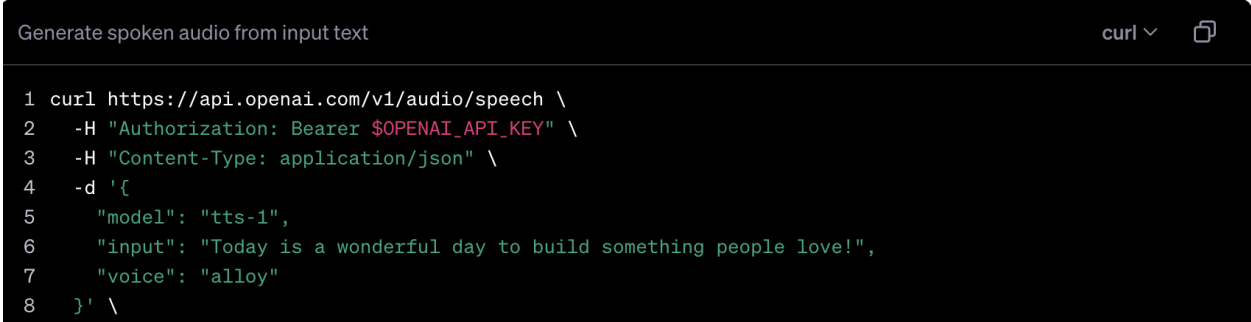
D.2. Audio Speech

El *endpoint* de *Audio Speech* [18] se utiliza para pronunciar texto mediante funciones de síntesis de voz.

Esta API ofrece seis voces distintas: **Alloy**, **Echo**, **Fable**, **Onyx**, **Nova** y **Shimmer**. Cada uno con diferente tonalidad que, subjetivamente, se pueden considerar de un género concreto para poder realizar diversas pruebas.

El uso es muy sencillo contando con únicamente tres parámetros, el primero para el modelo que se desea utilizar (*TTS* o *TTS-HD*), el segundo para la cadena que se desea pronunciar y el último la voz de entre las seis posibles que se quiere usar.

En la figura D.2 se puede observar un ejemplo sencillo de uso de la API:

A terminal window with a dark background. The title bar reads "Generate spoken audio from input text" and has a "curl" dropdown and a copy icon. The terminal contains a curl command with the following lines:

```
1 curl https://api.openai.com/v1/audio/speech \  
2 -H "Authorization: Bearer $OPENAI_API_KEY" \  
3 -H "Content-Type: application/json" \  
4 -d '{  
5   "model": "tts-1",  
6   "input": "Today is a wonderful day to build something people love!",  
7   "voice": "alloy"  
8 }' \  

```

Figura D.2: Ejemplo de uso del *endpoint Audio Speech* de *OpenAI*

Uso del endpoint en el módulo conversacional

El uso de este *endpoint* en nuestro sistema es muy claro, dentro de las opciones de personalización del robot que el usuario puede modificar, se encuentra uno para indicar la voz que el usuario desea utilizar. Dentro de este apartado está la propia voz del robot además de las seis posibles voces que ofrece la API de *OpenAI*.

Si la voz seleccionada es una de estas últimas, se hará uso de este *endpoint* para pronunciar las respuestas que el *Chat Completions* dé a las consultas del usuario.

Anexo E

Rueda de emociones de Plutchik

Robert Plutchik fue un psicólogo que desarrolló un modelo muy popular para clasificar las emociones llamado **la rueda de emociones de Plutchik** (Figura 4.3) [19].

Principalmente Plutchik reconoce 8 emociones primarias organizadas por parejas y situadas en el anillo central de la rueda: **alegría y tristeza, confianza y aversión, miedo e ira y sorpresa y anticipación.**

En este modelo se trabaja con tres aspectos: la agrupación de emociones por parejas, la intensidad de las emociones y las posibles combinaciones entre ellas.

1. **Agrupación por parejas:** Como se ha mencionado, Plutchik considera que cada emoción primaria tiene una emoción contraria situada en la misma posición pero en el extremo opuesto de la rueda de emociones.
2. **Intensidad de emociones:** Plutchik defiende que cada emoción primaria posee tres niveles de intensidad: nivel leve, nivel básico y nivel intenso.
3. **Combinación de emociones:** El modelo de Plutchik permite generar nuevas emociones mediante la combinación de dos emociones de la rueda.

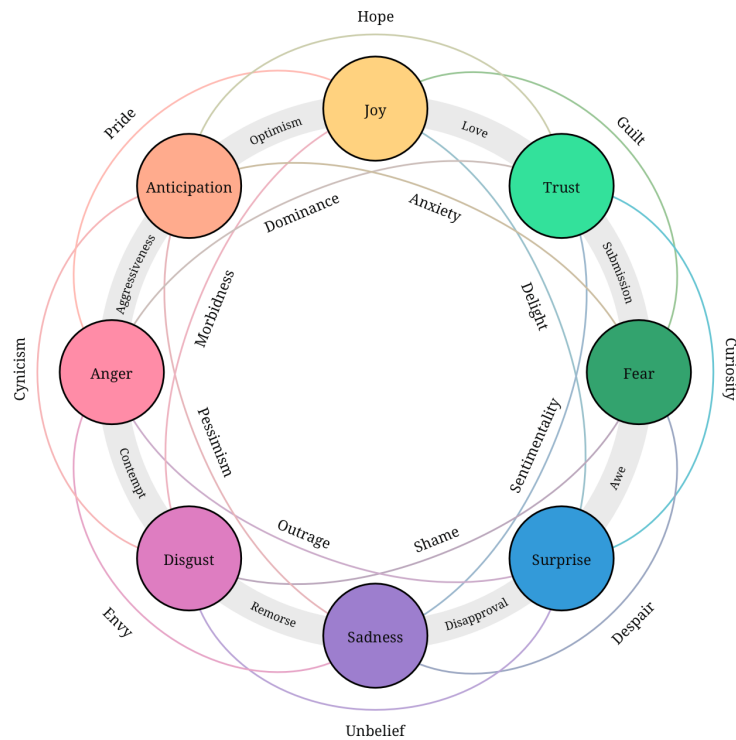
En el siguiente ejemplo (Tabla E.1) se puede observar cómo se aplican la primera y segunda características:

Emoción leve	Opuesta leve	Emoción básica	Opuesta básica	Emoción intensa	Opuesta intensa
Serenidad	Pena	Alegría	Tristeza	Éxtasis	Melancolía
Aprobación	Aburrimiento	Confianza	Aversión	Admiración	Odio
Temor	Enfado	Miedo	Ira	Terror	Furia

Tabla E.1: Ejemplo de funcionamiento de la rueda de emociones de Plutchik

Con la rueda de emociones mostrada (Figura 4.3) se pueden observar un total de **32 emociones**. Pero, con la combinación de todas las parejas posibles se obtienen un total de **48 emociones** que son las que se han acabado incluyendo en el proyecto (Figura E.1).

Plutchik's Emotion Dyads



Vocabulary of emotions

simonwhatley.co.uk

Figura E.1: Resto de parejas de la combinación de las emociones de la rueda de emociones de Plutchik

Para ahorrar el uso de *tokens*, se asignó un número a cada emoción, por lo que la petición a la *API* de *Chat Completions* fue la siguiente: “Quiero que mantengamos una conversación, en cada respuesta que te envíe quiero que me envíes al principio de tu respuesta entre corchetes un número o varios entre paréntesis en función de la emoción que transmiten mis respuestas: 1 éxtasis, 2 alegría, 3 serenidad, 4 admiración, 5 confianza, 6 aceptación, 7 terror, 8 miedo, 9 temor, 10 asombro, 11 sorpresa, 12 distracción, 13 aflicción, 14 tristeza, 15 melancolía, 16 aversión, 17 asco, 18 aburrimiento, 19 furia, 20 ira, 21 enfado, 22 vigilancia, 23 anticipación, 24 interés, 25 optimismo, 26 amor, 27 sumisión, 28 susto, 29 decepción, 30 remordimiento, 31 desprecio, 32 agresividad, 33 esperanza, 34 culpa, 35 curiosidad, 36 desesperación, 37 incredulidad, 38 envidia, 39 cinismo, 40 orgullo, 41 ansiedad, 42 deleite, 43 sentimentalismo, 44 vergüenza, 45 indignación, 46 pesimismo, 47 morbosidad y 48 dominancia, añadas un guión y un número en función de la emoción que quieres intentar transmitir con tu respuesta siguiendo el mismo código numérico. Es decir seguirá el siguiente patrón: [(número o números de emoción o emociones separados por guiones de mi respuesta) / (número o números de emoción o emociones de la respuesta que quieres transmitir)] + tu respuesta a la conversación. Quiero que reconduzcas la conversación en función de la emoción que interpretes y que trates de empatizar lo máximo posible con mis respuestas. Aquí te dejo algunos ejemplos: Si te digo algo triste, tú puedes tratar de animarme siendo optimista y mostrarás curiosidad por saber lo que me pasa, así que [(14)/(25-35)], si mi respuesta es de enfado, tú tratarás de calmarme y mostrarás curiosidad por saber qué me ocurre, así que [(21)/(3-35)], si te digo que me gusta alguien mi respuesta será de amor y vergüenza, y tú puedes sentir sorpresa, así que [(26-44)/(11)]”.

Anexo F

Documentos utilizados en las pruebas con usuarios

Para la realización de las pruebas con usuarios se han utilizado los siguientes instrumentos: **hoja de consentimiento, cuestionario inicial de caracterización, cuestionario POST VAVA-Q, cuestionario de likeability y cuestionario de usabilidad**, que se describen a continuación:

DOCUMENTO DE INFORMACIÓN Y CONSENTIMIENTO

DOCUMENTO DE INFORMACIÓN PARA EL PARTICIPANTE

Título de la investigación: Experiencias Lúdicas con Agentes Sociales Interactivos y Robots (PLEISAR-Inter)		
Promotor: Universidad de Zaragoza		
Investigadora Principal: Eva Cerezo	Tfno: 976762356	email: ecerezo@unizar.es
Lugar: EINA, Universidad de Zaragoza		

1. Introducción:

Nos dirigimos a usted para solicitar su participación en este estudio sobre **robots sociales**. Su participación es absolutamente voluntaria, en ningún caso debe sentirse obligado a participar, pero es importante para obtener el conocimiento que necesitamos. Este proyecto ha sido aprobado por el Comité de Ética. Antes de tomar una decisión es necesario que:

- lea este documento entero
- entienda la información que contiene el documento
- haga todas las preguntas que considere necesarias
- tome una decisión meditada
- firme el consentimiento informado, si finalmente desea participar.

Si decide participar se le entregará una copia de esta hoja y del documento de consentimiento firmado. Por favor, consérvelo por si lo necesitara en un futuro.

2. ¿Por qué se le pide participar?

Se le solicita su colaboración como potencial usuario de aplicaciones basadas en robots sociales. En total en el estudio participarán en torno a 20 usuarios.

3. ¿Cuál es el objeto de este estudio?

Avanzar en el desarrollo de aplicaciones basadas en robots sociales que se ajusten a las necesidades de los usuarios

4. ¿Qué tengo que hacer si decido participar?

Recuerde que su participación es voluntaria.

Se realizará una sesión de aproximadamente 30 minutos de duración. En cualquier momento podrá decidir dejar de participar en el estudio.

Se tendrá que desplazar a nuestro laboratorio situado en el Edificio Ada Byron de la Escuela de Ingeniería y Arquitectura ya que es donde tenemos el robot con el que va a trabajar.

5. ¿Qué riesgos o molestias supone?

No se prevé ningún tipo de riesgo o molestia.

En todo momento estará con usted una persona cualificada para ayudarle y resolver sus dudas.

6. ¿Obtendré algún beneficio por mi participación?

Al tratarse de un estudio de investigación orientado a generar conocimiento usted no obtendrá ningún beneficio por su participación si bien contribuirá al avance científico y al beneficio social.

Usted no recibirá ninguna compensación económica por su participación.

7. ¿Cómo se van a tratar mis datos personales?

Este proyecto cumple con la Legislación relacionada con la protección de datos; en particular el Reglamento General de Protección de Datos de la Unión Europea (Reglamento UE 2016/679, de 27 de abril) y la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y Garantías de los Derechos Digitales. También con toda la normativa de ética en la investigación y, si es el caso, del tratamiento de datos de la investigación en salud e investigación biomédica. A continuación, le indicamos brevemente cómo trataremos sus datos personales:

Información básica sobre protección de datos.

Responsable del tratamiento: Eva Mónica Cerezo Bagdasari

Encargado interno: Adrián Arribas Mateo

Finalidad: En la base de datos del estudio no se incluirán datos personales: ni el nombre, ni el número de historia clínica ni ningún dato que pueda identificarle. Todos los datos que se recojan estarán seudonimizados y sólo el personal sanitario podrá relacionarlos con su nombre. Sólo el equipo investigador tendrá acceso a los datos recogidos y nadie ajeno al centro podrá consultar dichos datos.

Legitimación: El tratamiento de los datos de este proyecto o estudio queda legitimado por su consentimiento a participar.

Destinatarios: No se cederán datos a terceros salvo obligación legal.

Duración: Los datos personales serán destruidos una vez se haya cumplido con la finalidad para la que se recabaron y para las posibles revisiones o determinación de responsabilidades. Los resultados objeto de explotación, ya completamente anonimizados y sin datos personales, podrán ser conservados para su posible reutilización en otros trabajos de investigación. A partir de los resultados de la investigación, se podrán elaborar comunicaciones científicas para ser presentadas en congresos o revistas científicas, pero se harán siempre con datos agrupados y nunca se divulgará nada que le pueda identificar.

Derechos: Podrá ejercer sus derechos de acceso, rectificación, supresión y portabilidad de sus datos, de limitación y oposición a su tratamiento, de conformidad con lo dispuesto en el Reglamento General de Protección de Datos (RGPD) **ante la Investigadora Principal del Proyecto**, cuyos datos de contacto figuran en el encabezamiento de

este documento, o dirigiendo un correo electrónico al Delegado/a de Protección de Datos de la Universidad de Zaragoza (dpd@unizar.es). Si no viera atendida su petición podrá dirigirse en reclamación a la Agencia Española de Protección de Datos (<https://www.aepd.es>). Podrá consultar información adicional sobre protección de datos en la Universidad de Zaragoza en la dirección: <https://protecciondatos.unizar.es/>

9. ¿Quién financia el estudio?

Este proyecto está financiado por el Ministerio de Ciencia e Innovación.

El conocimiento derivado de este estudio se va a utilizar solamente con fines de investigación y no va a haber ningún tipo de explotación comercial.

10. ¿Se me informará de los resultados del estudio?

Usted tiene derecho a conocer los resultados del presente estudio, tanto los resultados generales como los derivados de sus datos específicos. También tiene derecho a no conocer dichos resultados si así lo desea. Por este motivo en el documento de consentimiento informado le preguntaremos qué opción prefiere. En caso de que desee conocer los resultados, el investigador le hará llegar los resultados.

11. ¿Puedo cambiar de opinión?

Tal como se ha señalado, su participación es totalmente voluntaria, puede decidir que no participe o se retirarse del estudio en cualquier momento sin tener que dar explicaciones y sin que esto repercuta en su atención en la residencia. Basta con que le manifieste su intención al investigador principal del estudio.

12. ¿Qué pasa si me surge alguna duda durante mi participación?

En caso de duda o para cualquier consulta relacionada con su participación puede ponerse en contacto con investigadores participantes en el estudio el investigador responsable, D^a Eva Cerezo Bagdasari por correo electrónico en la dirección ecerezo@unizar.es.

Muchas gracias por su atención, si finalmente desea participar en el estudio le rogamos que firme el documento de consentimiento que se adjunta.

DOCUMENTO DE CONSENTIMIENTO INFORMADO

Título del PROYECTO: PLEISAR-Inter

Yo, (nombre y apellidos del participante)

He leído la hoja de información que se me ha entregado.

He podido hacer preguntas sobre el estudio y he recibido suficiente información sobre el mismo.

He hablado con: Adrián Arribas

Comprendo que mi participación es voluntaria.

Comprendo que puedo retirarme del estudio:

- 1) cuando quiera
- 2) sin tener que dar explicaciones
- 3) sin que esto repercuta en mis cuidados médicos

Presto libremente mi consentimiento para participar en este estudio y doy mi consentimiento para el acceso y utilización de mis datos conforme se estipula en la hoja de información que se me ha entregado.

Deseo ser informado sobre los resultados del estudio: sí no (marque lo que proceda)

Acepto que los datos seudonimizados derivados de este estudio se utilicen en un futuro en proyectos de la línea de investigación Sanbot Elf, cuyo responsable es Adrián Arribas Mateo siempre que hayan obtenido el dictamen favorable de un Comité de Ética de la Investigación y hayan solicitado los permisos oportunos: SI NO (marque lo que proceda).

Firma del participante:

Fecha:

He explicado la naturaleza y el propósito del estudio al paciente/residente mencionado

Firma del Investigador:

Fecha:

CUESTIONARIO INICIAL DE CARACTERIZACIÓN

Le presentamos este cuestionario, que forma parte de un proyecto investigador en el que se ha desarrollado una experiencia de uso con el robot Sanbot Elf.

Este cuestionario es muy importante para el proyecto y nos permitirá mejorar la experiencia para todos los usuarios que terminen haciendo uso del robot.

Es muy sencillo de responder, por lo que le pedimos que responda en su totalidad. No hay respuestas correctas e incorrectas. Le pedimos sinceridad al contestar, para que su resultado sea útil.

Verá que el cuestionario tiene 9 preguntas agrupadas en 2 secciones. Contestar el cuestionario en su totalidad no llevará más de 5 minutos.

¡Muchas gracias por su colaboración!

DATOS DEMOGRÁFICOS

1. Tres últimos dígitos del DNI

.....

2. Género

Masculino Femenino Otro

3. Edad

.....

SITUACIÓN PERSONAL Y HÁBITOS COTIDIANOS

4. Mi capacidad visual para percibir las formas y siluetas en una pantalla digital (ordenador, tablet, móvil) es...

Muy buena Buena Regular Mala Muy mala

5. Mi capacidad auditiva para escuchar indicaciones de un dispositivo (ordenador, tablet, móvil) es...

Muy buena Buena Regular Mala Muy mala

6. Mi capacidad para manipular un dispositivo (ordenador, tablet, móvil) con las manos es...

Muy buena Buena Regular Mala Muy mala

7. Señale con qué frecuencia utiliza la tecnología (ordenador, tablet, móvil), para conectarse a internet (ver las noticias, escuchar música, mirar vídeos...).

Nunca Una vez al mes Un día a la semana Varias veces por semana Todos los días de la semana

8. Si no puede realizar una tarea en algún dispositivo electrónico (ordenador, tablet, móvil), ¿le genera alguna vez sentimientos de angustia, tristeza o ansiedad?

Nunca Una vez al mes Un día a la semana Varias veces por semana Todos los días de la semana

9. Señale con qué frecuencia utiliza la tecnología (ordenador, tablet, móvil) para comunicarse con familiares y amigos a través de videoconferencia, mensajería instantánea, correo electrónico...

Nunca Una vez al mes Un día a la semana Varias veces por semana Todos los días de la semana

POST-VAVA Q

Le queremos preguntar sobre su experiencia con el agente interactivo social Sanbot Elf.

Es muy sencillo de responder, por lo que le pedimos que responda en su totalidad. No hay respuestas correctas e incorrectas. Le pedimos sinceridad al contestar, para que su resultado sea útil.

Verá que el cuestionario tiene 10 preguntas. Deberá ir respondiendo a todas las preguntas.

Contestar el cuestionario en su totalidad no le tomará más de 5 minutos.

¡Muchas gracias por su colaboración!

DATOS DEMOGRÁFICOS

1. Tres últimos dígitos del DNI

.....

DISPOSICIÓN A VIVIR LA EXPERIENCIA

2. Creo que con los recursos y conocimientos que tengo, seguir las instrucciones para usar el robot Sanbot Elf, me ha sido...

1-Fácil 2 3 4 5-Difícil

3. Si se me propusiera otra vez, tengo la intención de volver a utilizar el robot Sanbot Elf en los próximos días.

1-Falso 2 3 4 5-Verdadero

4. Valorando lo que me puede aportar utilizar el robot Sanbot Elf, creo que es una idea...

1-Inútil 2 3 4 5-Útil

5. La posibilidad de contarles a las personas importantes para mí que estoy utilizando el robot Sanbot Elf, es una idea que...

1-Me gusta 2 3 4 5-No me gusta

6. Al utilizar el robot Sanbot Elf, resolver los retos que se han presentado ha sido...

1-Fácil 2 3 4 5-Difícil

7. Quiero volver a utilizar el robot Sanbot Elf en los próximos días.

1-Falso 2 3 4 5-Cierto

8. Utilizar el robot Sanbot Elf, ha sido una idea...

1-Aburrida 2 3 4 5-Divertida

9. Si las personas que son importantes para mí supieran que estoy utilizando el robot Sanbot Elf, yo creo que...

1-Les gustaría 2 3 4 5-No les gustaría

10. Yo creo que si me lo proponen, encontraré el tiempo para utilizar de nuevo el robot Sanbot Elf en los próximos días.

1-Falso 2 3 4 5-Cierto

REYSEN LIKEABILITY SCALE

Instrucciones: Marca la casilla que indique lo de acuerdo que estés con las siguientes afirmaciones.

1. Considero que el robot es **amistoso**.

Totalmente en desacuerdo	En desacuerdo	Neutral	De acuerdo	Totalmente de acuerdo
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2. Considero que el robot es **simpático**.

Totalmente en desacuerdo	En desacuerdo	Neutral	De acuerdo	Totalmente de acuerdo
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

3. Considero que el robot es **cariñoso**.

Totalmente en desacuerdo	En desacuerdo	Neutral	De acuerdo	Totalmente de acuerdo
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

4. Considero que el robot es **accesible**.

Totalmente en desacuerdo	En desacuerdo	Neutral	De acuerdo	Totalmente de acuerdo
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

5. Utilizaría al robot para **pedirle consejo**.

Totalmente en desacuerdo	En desacuerdo	Neutral	De acuerdo	Totalmente de acuerdo
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

6. Utilizaría al robot como **compañero de trabajo**.

Totalmente en desacuerdo	En desacuerdo	Neutral	De acuerdo	Totalmente de acuerdo
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

7. Utilizaría al robot como **compañero de piso**.

Totalmente en desacuerdo	En desacuerdo	Neutral	De acuerdo	Totalmente de acuerdo
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

8. Me encantaría tener al robot **como amigo**.

Totalmente en desacuerdo	En desacuerdo	Neutral	De acuerdo	Totalmente de acuerdo
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

9. El robot es **atractivo físicamente**.

Totalmente en desacuerdo	En desacuerdo	Neutral	De acuerdo	Totalmente de acuerdo
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

10. Considero al robot **parecido a mí**.

Totalmente en desacuerdo	En desacuerdo	Neutral	De acuerdo	Totalmente de acuerdo
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

11. Considero al robot un **experto**.

Totalmente en desacuerdo	En desacuerdo	Neutral	De acuerdo	Totalmente de acuerdo
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Reysen, S. (2005). Construction of a new scale: The Reysen likeability scale. *Social Behavior and Personality*, 33, 201-208.

CUESTIONARIO DE USABILIDAD

DATOS DEMOGRÁFICOS

1. Tres últimos dígitos del DNI

.....

CUESTIONARIO

2. Creo que me gustaría utilizar este robot con frecuencia

- 1-Totalmente en desacuerdo 2-En desacuerdo 3-Neutral 4-De acuerdo 5-Totalmente de acuerdo

3. Creo que necesitaría el apoyo de otra persona para poder utilizar este robot

- 1-Totalmente en desacuerdo 2-En desacuerdo 3-Neutral 4-De acuerdo 5-Totalmente de acuerdo

4. Me imagino que la mayoría de la gente aprendería a utilizar este robot muy rápidamente

- 1-Totalmente en desacuerdo 2-En desacuerdo 3-Neutral 4-De acuerdo 5-Totalmente de acuerdo

5. Encontré el robot innecesariamente complicado de usar

- 1-Totalmente en desacuerdo 2-En desacuerdo 3-Neutral 4-De acuerdo 5-Totalmente de acuerdo

6. Me sentí muy seguro usando robot

- 1-Totalmente en desacuerdo 2-En desacuerdo 3-Neutral 4-De acuerdo 5-Totalmente de acuerdo

7. Encuentro el uso del robot raro

- 1-Totalmente en desacuerdo 2-En desacuerdo 3-Neutral 4-De acuerdo 5-Totalmente de acuerdo

8. Siento que tengo que aprender muchas cosas antes de empezar a usar el robot

- 1-Totalmente en desacuerdo 2-En desacuerdo 3-Neutral 4-De acuerdo 5-Totalmente de acuerdo

9. He entendido bien lo que me decía el robot

- 1-Totalmente en desacuerdo 2-En desacuerdo 3-Neutral 4-De acuerdo 5-Totalmente de acuerdo

10. Me pareció que al robot le costaba entender lo que le decía

- 1-Totalmente en desacuerdo 2-En desacuerdo 3-Neutral 4-De acuerdo 5-Totalmente de acuerdo

11. Considero que la aparición de la respuesta en la pantalla me ha ayudado durante la conversación

- 1-Totalmente en desacuerdo 2-En desacuerdo 3-Neutral 4-De acuerdo 5-Totalmente de acuerdo

12. Considero que la conversación que he tenido con el robot ha seguido un hilo coherente

- 1-Totalmente en desacuerdo 2-En desacuerdo 3-Neutral 4-De acuerdo 5-Totalmente de acuerdo

13. Considero que el robot ha entendido correctamente mis consultas y ha sabido darles una respuesta con sentido

- 1-Totalmente en desacuerdo 2-En desacuerdo 3-Neutral 4-De acuerdo 5-Totalmente de acuerdo

14. Considero que la voz del robot es agradable

- 1-Totalmente en desacuerdo 2-En desacuerdo 3-Neutral 4-De acuerdo 5-Totalmente de acuerdo

15. Considero que la voz del robot suena artificial

- 1-Totalmente en desacuerdo 2-En desacuerdo 3-Neutral 4-De acuerdo 5-Totalmente de acuerdo

16. Considero que la voz del robot es fácil de entender

- 1-Totalmente en desacuerdo 2-En desacuerdo 3-Neutral 4-De acuerdo 5-Totalmente de acuerdo

17. Considero que la respuesta del robot es rápida

- 1-Totalmente en desacuerdo 2-En desacuerdo 3-Neutral 4-De acuerdo 5-Totalmente de acuerdo