

Specialized and updated training on supporting advance technologies for early childhood education and care professionals and graduates



Co-funded by  
the European Union



**Specialized and updated training on supporting advance  
technologies for early childhood education and care  
professionals and graduates**

## **MÓDULO VII.2**

**Intervención temprana y aplicación de recursos inteligentes.  
Intelligent personal assistants**

Docente:

Dr. D. Raúl Marticorena Sánchez  
Área de Lenguajes y Sistemas Informáticos  
Departamento de Ingeniería Informática  
UNIVERSIDAD DE BURGOS

e-EarlyCare-T



“Specialized and updated training on supporting advance technologies for early childhood education and care professionals and graduates”, eEarlyCare-T, reference 2021-1-ES01-KA220-SCH-000032661, is co-financed by the European Union's Erasmus+ programme, line KA220 Strategic Partnerships Scholar associations. The content of the publication is the sole responsibility of the authors. Neither the European Commission nor the Spanish Service for the Internationalization of Education (SEPIE) is responsible for the use that may be made of the information disseminated herein”



Specialized and updated training on supporting advance technologies for early childhood education and care professionals and graduates

## Índice de Contenidos

<b>I. INTRODUCCIÓN</b>	<b>6</b>
<b>II. OBJETIVOS</b>	<b>6</b>
<b>III. CONTENIDOS ESPECÍFICOS DEL TEMA</b>	<b>6</b>
<b>3.1. Bots o asistentes personales inteligentes</b>	<b>6</b>
3.1.1. Fundamentos y precedentes históricos	6
<b>3.2. Definiciones</b>	<b>7</b>
3.2.1. Casos de uso más habituales	8
<b>3.3. Tipología</b>	<b>9</b>
3.3.1. Criterios en la evaluación y selección de plataformas	10
<b>3.4. Aspectos genéricos</b>	<b>10</b>
3.4.1. Gestión de la conversación: incorporación (onboarding)	13
3.4.2. Secuencia de comandos funcionales ( <i>functional scripting</i> )	14
3.4.3. Extracción de entidades	15
3.4.4. Contexto y memoria	15
3.4.5. Gestión de errores	16
<b>3.5. Asistentes basados en voz</b>	<b>17</b>
<b>3.6. Soluciones tecnológicas para asistentes personales</b>	<b>18</b>
<b>3.7. Aplicaciones prácticas en salud</b>	<b>20</b>
<b>RESUMEN</b>	<b>21</b>
<b>GLOSARIO</b>	<b>22</b>
<b>BIBLIOGRAFÍA</b>	<b>22</b>
<b>Bibliografía básica</b>	<b>22</b>



## **Bibliografía complementaria**

**22**



## I. Introducción

El presente módulo introduce el concepto y uso de *bots* o asistentes personales inteligentes. Dados los avances tecnológicos de la inteligencia artificial, las redes y el computo en la nube, surge el uso de asistentes conversacionales que simulan a las personas, ayudándonos en las tareas más o menos cotidianas como gestionar la agenda, realizar las compras, etc.

En concreto, se revisarán los conceptos fundamentales de los *bots*, así como se revisarán sus aplicaciones más particulares en salud, apuntando finalmente sus aplicaciones futuras en la atención temprana.

## II. Objetivos

Los objetivos básicos fijados en el siguiente submódulo son:

- Introducir su historia y el concepto de *bot* o asistente personal inteligente.
- Explicar las diferentes características que tienen los *bots*.
- Detallar los aspectos o componentes de un *bot* y soluciones actuales.
- Revisar sus posibles aplicaciones prácticas en salud.

## III. Contenidos específicos del tema

### 3.1. Bots o asistentes personales inteligentes

En la actualidad, el uso de **bots o asistentes personales inteligentes** (*Intelligent Personal Assistant* o **IPA**), empieza a estar **generalizado** en **todos** los ámbitos, proporcionando múltiples tipos de servicios, desde resolución de dudas, búsquedas, servicios de recomendación, gestión de agendas, reserva y compra de billetes, etc.

#### 3.1.1. Fundamentos y precedentes históricos

Su fundamento teórico se basa en el **Test de Turing**, desarrollado por Alan Turing en los 50. De manera simple, dicha prueba propone que el computador debe mostrar un comportamiento "inteligente", de tal forma que pudiera engañar a otro interlocutor humano en una conversación haciéndose pasar por otro ser humano. Conseguido esto, se superaría el Test de Turing.

En 1966, Joseph Weizenbaum desarrolla ELIZA para el IBM 7094. Se trataba de un *bot* para psicoterapia que trata con los pacientes sobre sus problemas, generando fuertes reacciones emocionales, aun siendo conscientes de que trataban con un *bot*. El programa analizaba las palabras claves y fundamentaba la respuesta sobre estas, siendo probablemente el primer programa en superar el Test de Turing.

Un paso posterior en esta rama fue la aparición de PARRY, un agente conversacional que simulaba un paciente paranoide con esquizofrenia (Colby, 1975). Estos son los primeros ejemplos de aplicación válida a cuestiones de salud, superando dicho Test de Turing.

La proliferación y desarrollo de agentes conversacionales recibe un empuje que se refleja en la creación de **competiciones**. El premio Loebner se establece como plataforma anual de competición de *bots* conversacionales. ALICE (Artificial Linguistic Internet Computer Entity) gana gran atención en 1995, ganando dicho premio en tres ocasiones (2000, 2001 y 2004).

A partir de 2010, Siri emerge como la solución **comercial** de Apple integrada en sus dispositivos móviles, y posteriormente se suma Google Now en el 2012, Alexa (Amazon) en el 2014 y Cortana (Microsoft) en el 2014. En 2016 se produce una eclosión final con la integración de *bots* en redes sociales como Facebook con su propia plataforma de mensajería, a la que se irían sumando desarrollos de API.ai, LinkedIn, Viber, etc.

En épocas más recientes, Amazon creó su propia competición –“Alexa Prize” en el 2017 – con un planteamiento similar al Test de Turing. El desafío consiste en crear un “*bot* social” que converse coherentemente y que enganche a un humano, conversando sobre un tema 20 minutos. El referente actual es Mitsuku ganador de cuatro premios Loebner (Prize L, 2019).

En los últimos tiempos, la aplicación de estos *bots* conversacionales en el campo de la salud ha cobrado mucho peso con situaciones como la pandemia por la COVID-19. Como ejemplo, en el 2020 WhatsApp acordó con la Organización Mundial de la Salud (OMS) completar un servicio de *chatbot* que respondiera a cuestiones relativas a la COVID-19. Aunque no exentas este tipo de soluciones de riesgos y usos maliciosos, que han obligado a organizaciones como UNICEF a definir guías de buen uso e implementación (UNICEF, 2022a; UNICEF 2022b).

### 3.2. Definiciones

Un *bot* expone servicios software a través de una interfaz conversacional. A estos *bots* se les puede denominar como *chatbots*, agentes conversacionales, interfaces conversacionales, asistentes personales inteligentes, y de otras muchas formas, dependiendo también de la interfaz utilizada con el usuario (Shevat, 2017).

El concepto de “**conversación**” es vital. Muchas veces en estos *bots* se consideran conversaciones de una sola interacción (pregunta/respuesta) (e.g. fijar una alarma, consultar la temperatura o previsión del tiempo, realizar una llamada, etc.) cuando en el mundo real, los humanos entendemos por conversación a un intercambio más largo de preguntas/respuestas que están relacionadas entre sí. Aunque los *bots* también pueden mantener conversaciones largas sobre un tema, con varios intercambios de preguntas/respuestas.

En la siguiente Ilustración 1, se representa un ejemplo de conversación simple con un *chatbot* basado en texto, habitualmente denominados *chatbots*, como ejemplo del funcionamiento esperado. A la izquierda se muestran las interacciones del software, generadas automáticamente. Mientras que a la derecha se tienen las preguntas lanzadas por una persona. Sin información precisa, debería ser indistinguible si los mensajes del *bot* son generados por un programa o por otra persona.





### Ilustración 1. Ejemplo de interacción con un *bot* (Imagen: elaboración propia)

Se debe distinguir el *bot* del servicio que proporciona. El *bot* es solo un medio o interfaz al servicio. Aunque dicha interfaz puede ser más o menos compleja.

Ventajas de su uso:

- Mayor compromiso (fidelización) del usuario al facilitarle sus tareas.
- Facilidad de uso frente a interfaces web y apps de móviles.

Desventajas en su uso:

- No son la solución a todo tipo de problema planteado (actualmente).
- Miedo a cierto riesgo de pérdida de privacidad.
- Con interfaces basadas en voz, se produce una cierta vergüenza social

#### 3.2.1. Casos de uso más habituales

En la práctica los casos de uso de éxito más habituales son:

- Productividad y entrenamiento (coaching): centrados en recordar tareas a realizar, y gestión de tareas personales o en grupo a completar. También para ayudar a seguir dietas, gestionar gastos, realizar actividades deportivas, etc.
- Alertas y notificaciones: reemplazando al uso de correos y *apps* de notificaciones. Pueden trabajar con grupos de una manera más productiva.
- Enrutador a humanos: redirige finalmente a un interlocutor humano, pero asignando la persona más indicada para la resolución, mediante una conversación guiada.
- Servicios a clientes y respuestas a preguntas frecuentes: como soporte a las dudas más habituales que se repiten.
- Integración de terceros: para integrar servicios de otros en el producto actual.
- Juegos y entretenimiento: con el objetivo básico de entretener y divertir.

Partiendo de estos casos, vamos a identificar diferentes tipologías.



### 3.3. Tipología

Por destinatario:

- **Bot personal (Personal bot / private bot):** sirve como un asistente personal, en una conversación uno a uno (e.g. fijar en mi calendario personal una cita)
- **Bot de equipo (Team bot):** asiste a un grupo de personas para alcanzar un objetivo (e.g. fijar una fecha y hora de reunión para un grupo).

En la práctica, está más difundido el modelo personal por simplicidad, e incluso en *bots* en el hogar, con dispositivos como Amazon Echo o Google Home, se trata a todos los usuarios del mismo hogar sin distinción, como si fueran la misma persona.

Por ámbito:

- **Específico de dominio (Domain-specific):** expone un único servicio (producto, marca u objetivo).
- **Generalista (Super bot):** expone múltiples servicios al mismo tiempo.

En esta última categoría nos encontramos con soluciones como Google Assistant, Amazon Alexa, Apple Siri, etc. que encapsulan varios servicios. Básicamente agrupan diferentes funcionalidades de manera modular, de forma que se puedan incrementar de forma transparente al usuario. Estas funcionalidades modulares se suelen denominar *skills*.

Por objetivo:

- **De negocio (Business):** facilitan una tarea o proceso del negocio. El objetivo es resolver un objetivo. Orientados a tareas y flujos de trabajo.
- **De consumo (Consumer):** entretienen a la vez que facilitan una interacción comercial. Orientados a una mejor y entretenida experiencia de usuario.

Por acceso:

- **Texto:** la conversación se basa en texto introducido por teclado y visualización de la respuesta en pantalla. Se suelen denominar *chatbots* (e.g. con plataformas web que los incluyen como Slack, Facebook Telegram, WhatsApp, WeChat, etc.)
- **Voz:** la conversación se basa en el uso de audio para pregunta y devolver la respuesta sin necesidad de interactuar físicamente con los dispositivos (e.g. Amazon Alexa, Microsoft Cortana, Apple Siri, Google Assistat como el estándar de facto).
- **Multimodal:** combinan ambos elementos discrecionalmente – texto o voz – y adicionalmente pueden requerir la interacción táctil en pantallas o el uso combinado de otros dispositivos o artefactos (e.g. cámaras, relojes, dispositivos, *wearables*, etc).

En este caso no son exclusivos los modos de acceso, pudiendo tener un *bot* cuyo acceso se realice de todas las formas anteriormente indicadas.

Por integración:

- **En sistemas legados:** dan servicio sobre sistemas software ya existentes ofreciendo nuevas formas de interactuar con servicios ya preexistentes.
- **Nuevos bots:** interfaces a nuevos servicios o productos creados desde cero.



### 3.3.1. Criterios en la evaluación y selección de plataformas

A la hora de elegir una solución tecnológica aplicada, hay que tener en cuenta ciertas características:

- Audiencia a la que se dirige.
- Negocio vs. consumo.
- Forma de interactuar (texto vs. voz vs. multimodal).
- Dispositivos requeridos para interactuar.
- Costes asociados de hosting de software y compra de hardware.

En función de los criterios previos, con cuestiones más sociales o económicas en mente, no todas las soluciones de *bots* son válidas.

### 3.4. Aspectos genéricos

Se debe ser consciente que detrás de un *bot* existe básicamente un **software** ejecutado sobre plataformas hardware que hospedan dicho software.

Aunque en cada plataforma tecnológica el vocabulario utilizado puede variar, se presenta a continuación un resumen en la Tabla 1, del **vocabulario más habitual** con los conceptos básicos en un *bot* (Nota: se toma el vocabulario de DialogFlow y Alexa como referencias).

**Tabla 1. Vocabulario básico en la construcción de un *bot***

Elemento	Definición	Uso
Intent	Objetivo (o intención) que tiene un cliente al formular una pregunta.	Se definen <i>intents</i> como aglutinador de las distintas acciones. Cuando se realiza una pregunta el sistema <b>NLU</b> <sup>1</sup> busca el <i>intent</i> al que más se aproxima.
Utterance	Frase literal introducida por el usuario.	Se definen un conjunto de frases alternativas que son equivalentes para resolver un <i>intent</i> . El sistema <b>NLU</b> interpreta y resuelve la correspondencia con el <i>intent</i> .
Entity	Tipo de datos que puede ser extraído del mensaje o <i>utterance</i> del usuario.	Se utilizan como variables que pueden definirse y tomar distintos valores, para realizar acciones personalizadas en función de dichos valores.
Context	Similar a un contexto en una conversación real, definiendo variables que determinan la evolución o camino de la conversación.	Se utilizan para definir y establecer conversaciones más avanzadas, donde pueden existir distintos caminos en la conversación.
Fallback	<i>Intent</i> por defecto cuando la entrada no ha sido reconocida.	En el caso de que el <i>chatbot</i> no pueda reconocer la entrada del usuario, se debe establecer una acción típica de respuesta. Idealmente el número de veces que se ejecuta un <i>fallback</i> irá reduciéndose según se entrene y mejore el <i>chatbot</i> .
Event	Disparan la ejecución de un <i>intent</i> de forma automática sin	Permite la automatización de acciones, como por ejemplo lanzar una pregunta

<sup>1</sup> **NLU**: Natural Language Understanding (Comprensión del lenguaje natural)



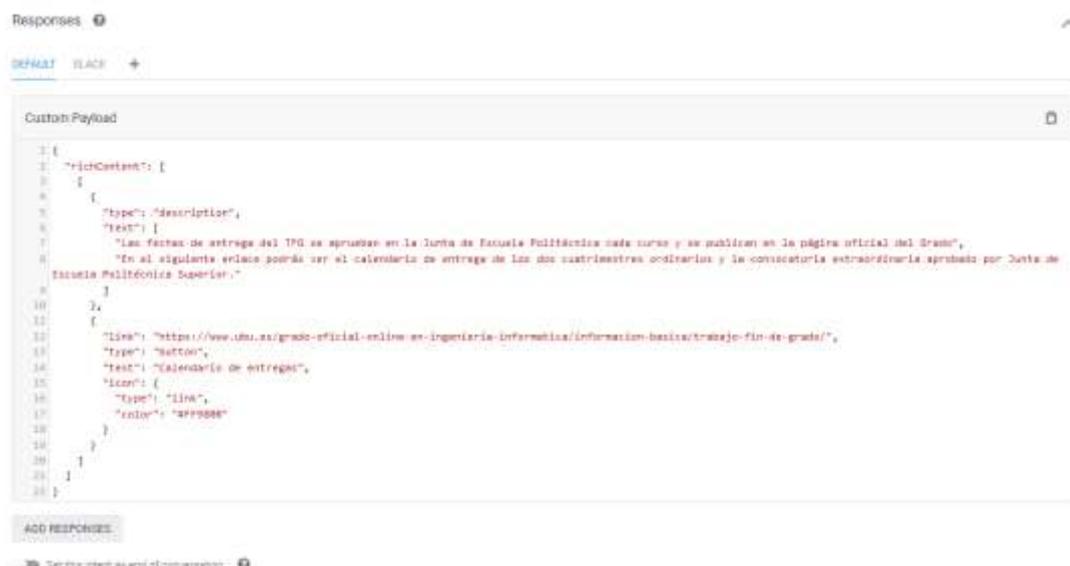
requerir una entrada del inicial cuando el usuario entra en la web usuario. que contiene el *chatbot*.

En la Ilustración 2, podemos ver cómo se definirían las frases de entrenamiento para una intención (*intent*) de comunicar las fechas de entrega de una tarea académica:



**Ilustración 2. Definición de preguntas para una intención (*intent*) en DialogFlow**

Una vez fijadas dichas frases de entrenamiento, se debe fijar también el posible conjunto de frases de respuesta. En algunos casos, se suelen dar varias respuestas equivalentes, para que aleatoriamente se elija una dotando al *bot* de una mayor variedad. O bien como en el ejemplo de la Ilustración 3, utilizar formatos más complejos (i.e. notación JSON) para generar contenidos enriquecidos con contenido multimedia.

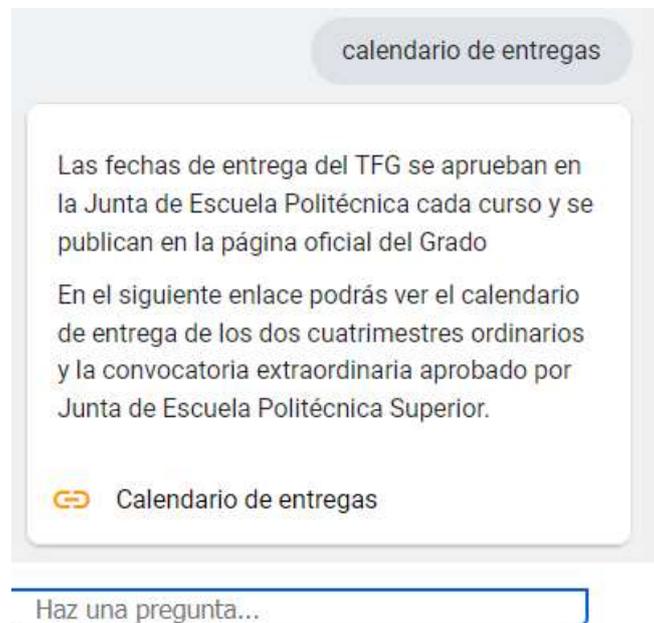


**Ilustración 3. Respuesta enriquecida para una intención (*intent*) en DialogFlow**

Como resultado de la consulta al *bot*, se obtendría en la Ilustración 4 una presentación en pantalla como la mostrada a continuación, donde además de mostrar un texto plano se



añade un elemento multimedia, con un botón como hiperenlace, que una vez pulsado abriría el contenido en un navegador web:



**Ilustración 4. Visualización del chatbot con contenido enriquecido**

Presentado el vocabulario y un ejemplo básico, se detallan otras características:

- **Estilo de interacción con el usuario:** personalidad o tono con el que se quiere comunicar. El *bot*, al igual que ocurre con las personas, debe tener una personalidad propia y reconocible, coherente a lo largo del tiempo.
- **Inteligencia artificial:** base fundamental de un *bot* que incluye varios elementos.
  - Procesado del lenguaje natural (Natural Language Processing o **NLP**): procesa la frase analizando la estructura de la misma.
  - Comprensión del texto (Natural Language Understanding o **NLU**): extrae el significado de la frase, una vez analizada previamente su estructura.
  - Conversión de voz a texto (Text To Speech o **TTS**): convierte un audio a su correspondiente texto.
  - Conversión de texto a voz (Speech To Text o **STT**): locuta un texto generando el correspondiente texto.
  - Modelos de predicción: modelos basados en técnicas de inteligencia artificial para dada una frase, predecir cual es la respuesta más adecuada. Asocia a la frase (*utterance*) del usuario la intención con la respuesta correspondiente.
  - Reconocimiento de imágenes en sistemas multimodales: el análisis de la imagen y su reconocimiento puede permitir saber qué está delante de la cámara (persona) y cuál es su estado de ánimo.
  - Gestión de la conversación: el sistema debe gestionar el flujo de preguntas respuestas de manera adecuada, recordando el contexto previo, para llegar al final de la conversación con éxito.
  - Análisis de sentimiento (*Sentiment analysis*): las frases pueden llevar implícitamente unas connotaciones que pueden hacer variar su interpretación y flujo de conversación. Para ello se aplican técnicas de análisis de sentimiento, ponderando lo positivo, negativo o neutro de los componentes de esta, afectando a la posterior respuesta.
- **Gestión de la conversación:** cómo se gestiona la conversación, desde su inicio a la consecución del objetivo fijado.



- Incorporación (*onboarding*): o cómo se prepara el acceso al *bot*, informando de sus objetivos y propósito, modos de interacción, funcionalidad aportada y cómo obtener ayuda.
- Secuencia de comandos funcionales (*functional scripting*): gestión de los flujos de conversación o historias, intentando mitigar el fracaso (pérdida del camino óptimo o "*happy path*" en la conversación).
- Retroalimentación y gestión de errores: permitiendo la mejora continua y una buena experiencia de usuario.
- Ayuda y soporte: en caso de pérdida del flujo de conversación óptimo o "*happy path*".
- **Interacciones enriquecidas:** en *bot* utilizando interfaces con texto las respuestas pueden ser enriquecidas con elementos multimedia o web que ayuden a simplificar la carga cognitiva de la conversación.
  - Ficheros: subida o descarga de ficheros (e.g. documentos de trabajo, listas de la compra, etc.)
  - Audio
  - Vídeo
  - Imágenes, mapas o gráficos
  - Botones: permite prefijar respuestas, acelerando la conversación.
  - Plantillas (dependientes de la plataforma)
  - Enlaces
  - Texto formateado con colores, estilos, etc.
  - Emojis o reacciones
  - Menús persistentes
  - Indicadores de teclado: para fingir que el *bot* está tecleando como si fueres un humano.
  - Comandos: atajos con ordenes breves para invocar acciones en forma similar a las consolas de texto.
  - Webviews: abrir una mini-vista con una porción de una página web.
- **Contexto y memoria:** recordar el contexto previo de la conversación, o conversaciones recientes, que puede cambiar completamente el flujo de la conversación actual. Los humanos realizamos esto de forma natural, pero en el caso de los *bots* es una de las cuestiones más complejas.
- **Descubrimiento e instalación:** facilitar su localización para su descarga, instalación, y facilitar la interacción inicial.
- **Métodos de fidelización (*engagement*):** métodos para conseguir que el usuario sea un cliente habitual del *bot*.
  - Notificación de nuevo contenido.
  - Guiar y ayudar a la invocación: proporcionando métodos fáciles de iniciar la conversación enseñando al usuario dicho proceso.
  - Suscripción: permitir al usuario definir sus intereses, filtrar lo que quiere, etc.
- **Monetización:** generar ingresos económicos o beneficios de otra índole con el uso del *bot* bien directa o indirectamente.

### 3.4.1. Gestión de la conversación: incorporación (*onboarding*)

Una primera cuestión es preparar bien la **primera toma de contacto** con el usuario con el *bot*, dado que una mala experiencia puede hacer que no vuelva a ser utilizado.

Las recomendaciones o heurísticas a seguir son:

- Declarar el propósito y contexto de la conversación: haciéndolo claro al usuario o grupo de usuarios. Presentar e introducir claramente qué hace el *bot*.



- Enseñar al usuario cómo se usa el *bot*: explicar cómo se interactúa con el *bot*, como por ejemplo indicar si hay alguna forma de activarlo (e.g. palabra de activación o "wake word" con interfaces basadas en voz), funciones principales, palabras claves o comandos, etc.
- Configuración: solicitar información adicional al usuario si se requiere para el funcionamiento del *bot* (e.g. preferencias o personalización para el usuario).
- Incitar al usuario a obtener valor del *bot*: animar o dirigir las primeras preguntas para iniciar la conversación. Incluso proporcionando algún ejemplo de conversación simulada.
- Establecer el tono y personalidad del *bot*: por coherencia el *bot* debe mantener el mismo tono durante toda la conversación. Lo ideal es que tenga una "personalidad".
- Hacer explícita la entrada del *bot* en conversaciones en grupo: de manera similar a cuando se presenta a un nuevo miembro de la plantilla, debe ser presentado a todos de manera adecuada. El *bot* debe presentarse a todos los miembros, anunciando su incorporación.

### 3.4.2. Secuencia de comandos funcionales (*functional scripting*)

El flujo de conversación puede estar dirigido básicamente a **tareas o temas**. En el primer caso, el objetivo es encontrar el conjunto óptimo de interacciones conversacionales para encontrar el objetivo (e.g. reservar un billete) o completar una **tarea** precisa.

Una forma habitual es modelar el conjunto de estados y transiciones que se dan en la conversación. También se denominan sistemas de orden-y-control (*command-and-control*).

Cuestiones a considerar en este tipo de *bots* son:

- En las respuestas se pueden cerrar el posible conjunto de respuestas para limitar la conversación y conseguir el camino óptimo o "**happy path**".
- En caso de que algún momento el usuario no de una entrada esperada se produce un **flujo divergente** y habrá que realizar una **gestión del error**.
- Para extraer información que puede venir en desorden se realiza la **extracción de entidades**. Consiste en extraer de la pregunta elementos (entidades) que tienen una semántica propia para resolver el flujo de conversación.
- **Mapeo de intenciones y control conversacional**: asociación entre las distintas entradas del usuario y las acciones/respuestas a dar. Mientras que en interfaces gráficas clásicas se da al usuario un conjunto cerrado de menús, botones, etc. en un *bot* se ofrece de otra forma. Bien implícita en la propia pregunta o bien cerrando en la respuesta un posible conjunto de comandos válidos.
  - Se denominan **intenciones (*intents*)** a las posibles acciones que puede requerir un usuario como resultado de una o conjunto de posibles entradas.
- **Taquigrafía**: si se mantiene un contexto y memoria de las acciones previas, se pueden abreviar las preguntas, utilizando el conocimiento del estado previo.
- **Historias/flujo**: permiten ramificar y agrupar conversaciones más comunes o que se repiten (e.g. "usuario quiere un producto no disponible")
- **Embudo de conversación**: a medida que un usuario va preguntando se reduce el posible camino de la conversación. Es conveniente definir estos embudos para reducir el posible conjunto de caminos.

En los flujos dirigidos por **temas**, el flujo es menos dirigido e incluso con un carácter "circular". Se conversa y discute sobre diferentes aspectos de un tema. Generalmente incluyen más interacciones que un *bot* dirigido a tarea, y su objetivo adicional es conseguir



un cierto compromiso del usuario (*engagement*) (e.g. un *bot* discutiendo sobre una película o serie).

Son bastante más complejos de implementar dadas sus características más abstractas, aplicándose los conceptos ya vistos en *bots* dirigidos a tareas (i.e. corrección de curso, extracción de entidades, etc.)

### 3.4.3. Extracción de entidades

En una frase (*utterance*) algunos de los términos utilizados pueden tener una cierta semántica especial, y que además puede condicionar la conversación posteriormente. A esto en el vocabulario de los *bots* se le denomina **entidad**, y es muy importante que el *bot* sea capaz de extraerlas y recordarlas adecuadamente.

Por ejemplo, supongamos que preguntamos:

"¿Qué temperatura hace hoy en Madrid?"

Además de responder a la pregunta, al analizar la frase se pueden detectar entidades de distinta naturaleza, un concepto a resolver como la **temperatura**, una **fecha** con valor "hoy" y una **ciudad** o destino con valor "Madrid". Si posteriormente se preguntase:

"¿Y la previsión para los siguientes días?"

En este caso el *bot*, si ha extraído las entidades previas, ya no tiene que preguntar al usuario solicitando información adicional, puesto que la pregunta ya no es ambigua, dado que el marco de referencia temporal inicial era "hoy" y la ciudad "Madrid", y estábamos preguntado por el concepto de la "temperatura". Por el contrario, si las entidades no hubiesen sido extraídas y almacenadas en el contexto, el *bot* no tendría información precisa (i.e. no tendría memoria del hilo de conversación), y no sabría de qué estamos hablando (e.g. ¿previsión de qué? ¿en qué fechas? ¿de qué lugar estamos hablando? etc.).

Otros ejemplos típicos, son recordar o mantener entidades con las cantidades solicitadas, colores favoritos, edades, nombres, etc. de tal forma que la conversación fluya, recordando las elecciones previas ya tomadas en el hilo de conversación.

### 3.4.4. Contexto y memoria

En la práctica muchos *bots* se limitan al paradigma de **pregunta/respuesta**, como por ejemplo en una resolución simple de preguntas frecuentes. Para cada nueva pregunta se parte de un nuevo contexto, no recordando la conversación previa. Sin embargo, para otro tipo de conversaciones, es necesario recordar los mensajes previamente intercambiados, que constituyen un contexto.

#### Contexto

Para aplicar el contexto se debe analizar la intención (*intent*) y el conjunto de entidades asociadas previamente en la conversación. Las variables asociadas pueden ser globales o a largo plazo (e.g. relativas al usuario y a conversaciones previas) o locales a la conversación o de corto plazo (e.g. día y hora de la reserva, color elegido, etc. en la conversación actual).

Cuando nos movemos en la conversación a otro *intent*, las variables locales pueden ser o no olvidadas, dependiendo de su utilidad, pero las globales no deberían olvidarse.



La inferencia adicional del contexto a través del uso de pronombres debe ser resuelta a través de **NLU** y queda fuera del ámbito de este módulo. Otra forma de inferir contexto es a través de contenido enriquecido que ayuda a capturar *intents*.

## Memoria

Es un concepto más general, relativo a recordar intenciones (*intents*) y entidades, o incluso conversaciones completas del pasado más lejano, y está todavía en investigación.

### 3.4.5. Gestión de errores

Los errores en una conversación son mucho más habituales de lo que uno desearía. Tanto en la vida real, como lo que particularmente se produce al conversar con un *bot*.

Si se revisan los registros de todas las conversaciones a lo largo del tiempo, se podrá comprobar la variedad de posibles entradas que no generan éxito en la conversación. Ya sean simples textos ininteligibles, preguntas o frases (*utterances*) sin sentido, frases no conectadas con la conversación previa, incluso insultos y uso de lenguaje ofensivo, fuera de lugar.

A esto se suma que muchas de las frases que deberían ser reconocidas como válidas, en las primeras etapas de configuración del *bot*, puede que no hayan sido incluidas para alcanzar una intención (*intent*) correcta. Ante los posibles errores y divergencias en las conversaciones, se pueden tomar diferentes medidas:

- **Corrección de curso:** consiste en redirigir la conversación al camino óptimo (*happy path*), contestando que no se puede resolver la pregunta previa, pero dando opciones para reconducir la conversación con opciones sí contempladas. O bien indicando que se toma nota, no pudiendo resolverse ahora, pero sí en un futuro.
- **Intervención humana:** es una solución habitual, que cuando la conversación diverge tanto de camino de éxito, se redirija a una persona real para proseguir la conversación y resolverla.
- **Reiniciar la conversación:** es la solución más fácil, pero puede generar muchas molestias, generando mala experiencia de usuario.
- **Redirigir a otro bot:** no muy utilizada por el momento, pero dada la arquitectura modular con aplicaciones modulares (*skills*) de los sistemas predominantes en el mercado, será habitual que los *bots* se especialicen y redirijan las peticiones de unos a otros (incluso planteándose ya el dilema de la conversación entre *bots*).

Otras consideraciones en el manejo de errores es la consistencia de las respuestas a lo largo de la conversación, mostrando un cierto "carácter" o "personalidad" del *bot*, que debe mantenerse de manera coherente. De igual forma que contestaría un ser humano ante la aparición de problemas en una conversación, y no generando código numéricos o mensajes que no puedan ser interpretados por una persona.

Este proceso de gestión de errores y corrección de estos no es inmediato, y se debe producir gradualmente, aprendiendo de los errores cometidos en las primeras fases. Revisando las intenciones que han fallado, se podrá deducir y mejorar lo que falla, **ajustando y afinando el conjunto de frases de entrenamiento** asociadas a una intención (*intent*). El proceso de crecimiento y mejora de un *bot* es continuo, al igual que ocurre con el aprendizaje de una persona.



### 3.5. Asistentes basados en voz

Los avances en las dos últimas décadas relativos al reconocimiento del habla (*Speech to Text* o STT) y la conversión de texto a voz (*Text to Speech* o **TTS**) han dado un empuje definitivo a la incorporación de las interfaces basadas en voz como un elemento adicional de los *bots*. En la Ilustración 5 se muestra un típico asistente con micrófono y altavoz integrado de primeras generaciones, sin pantalla.



**Ilustración 5. Ejemplo de dispositivo conversacional para el hogar [Imagen: licencia Unsplash - <https://unsplash.com/>]**

Ejemplos **comerciales** de las grandes compañías con soporte a asistentes basados en voz habituales en el día a día son Amazon Alexa, Apple Siri, Google Assistant o Microsoft Cortana. Sin embargo, debemos diferenciar entre el concepto de asistente de voz (i.e. software generalmente integrado en la nube o *cloud* en internet), como los ya citados, de sus correspondientes “altavoces inteligentes” (o con dispositivos aún más complejos y multimodales) como Amazon Echo, Google Home o Harman Kardon Invoke (para Cortana), que no dejan de ser el medio físico de acceso al agente conversacional (i.e. hardware).

Desde un punto de vista práctico, la introducción de esta interacción basada en voz, tanto para preguntar como para responder a través de esta combinación de software y hardware, añade unas ciertas cuestiones adicionales a considerar.

Inicialmente existen ciertas **ventajas** en su uso frente a la solución de introducir texto por teclado:

- Mayor rapidez en la emisión de preguntas.
- “Manos libres” dejando libertad al usuario para poder realizar otras acciones mientras usa el *bot* (y de forma más segura).
- Intuitivos: la interacción con habla es muy natural.
- Empatía: la inclusión de tono, volumen, entonación y velocidad en el habla, añaden información que ayuda a interpretar mejor la respuesta y evitar equívocos.

Adicionalmente, la reducción de pantallas en dispositivos que vestimos actualmente (e.g. teléfonos, relojes, pulseras, anillos, etc.) invitan al uso de este tipo de interfaces.

Por otro lado, también ofrecen ciertas **desventajas** a considerar en su uso:

- Uso en espacios públicos: hablar o alzar la voz en espacios públicos puede dar una extraña sensación al resto de la gente que nos rodea. El problema se agrava si varias personas hacen esto a la vez.
- Sensación de incomodidad hablando a un computador.



- Costumbre arraigada de teclear para interactuar con dispositivos.
- Privacidad: si se quieren tratar temas sensibles (e.g. salud) o escuchar cuestiones privadas (e.g. leer mensajería), no se querrá que nadie más oiga nuestra conversación.

Respecto a la gestión de errores, en estas interfaces surgen problemas adicionales a tratar como:

- Ausencia de discurso (i.e. el usuario no emite una pregunta durante un determinado tiempo). Se puede dar la conversación por acabada o preguntarle nuevamente.
- Problemas en el reconocimiento (i.e. no se reconoce la frase pese a que sí se ha oído). Se puede volver a preguntar.
- Problemas en el manejo de intenciones (i.e. se reconoce la frase, pero no hay respuesta programada adecuada o bien se da una respuesta equivocada) ya presentes en los *bots* basados en texto. Su resolución es más compleja.

A la hora de diseñar interfaces puramente basadas en voz hay que tener en cuenta algunas cuestiones adicionales. Desde un punto de inicio de la conversación, no se quiere que el dispositivo esté continuamente escuchando, por cuestiones de privacidad, y por lo tanto existe el concepto de **wake word** o **palabra de activación** que debe ser dicha por el usuario para iniciar explícitamente una conversación, activando el micrófono.

Por otro lado, las respuestas **no deberían ser excesivamente largas**, puesto que la carga cognitiva puede exceder a la capacidad del usuario. En sistemas multimodales se resuelve adicionalmente apoyándose en la presentación de información en pantalla, pero esto no siempre es posible y debe ser considerado a la hora de dar ciertas respuestas.

### 3.6. Soluciones tecnológicas para asistentes personales

A continuación, se repasan las soluciones actuales, tanto de los cuatro grandes líderes tecnológicos, en cuanto al uso de **bots conversacionales** como otras plataformas que permiten la integración de *bots*. Abordando soluciones tanto en texto, como con voz (con micrófono y altavoz integrados), y multimodal combinando imágenes, hiperenlaces, etc. siempre que el dispositivo asociado disponga adicionalmente de una pantalla de menor o mayor tamaño.

#### Alexa Amazon

Fundamentado sobre el ecosistema de Amazon en la red, es uno de los referentes en los últimos años, al tener detrás una de las mayores infraestructuras de computación en la nube (*cloud computing*). Es un aglutinador de funcionalidades denominadas *skills*. Añadiendo más o menos *skills* en Alexa, el usuario obtiene más o menos funcionalidades y tipos de conversaciones. Con una orientación hacia la domótica y control de dispositivos en el hogar, es uno de los referentes actuales.

*Dispositivos físicos asociados:* Amazon Echo, Amazon Dot o Amazon Echo Show, etc.

#### Apple Siri

Primer precedente en agentes conversacionales, integrados en los iPhone. Sin embargo, su filosofía más cerrada y de altos costes, habitual en productos Apple, dificulta la integración con terceros y su mayor difusión.



*Dispositivos físicos asociados:* iPhone.

### Google Assistant

Es la solución de Google para los agentes conversacionales, apoyándose en toda su experiencia previa en buscadores basados en texto como Google. El apoyo de dispositivos con Android, muy difundidos actualmente, dan un fuerte empuje a esta solución asociado también a menores costes. Sigue una arquitectura similar de *skills* a la de Amazon Alexa. También con una cierta orientación hacia la domótica y control de dispositivos en el hogar.

*Dispositivos físicos asociados:* Google Nest, Nest Mini, Nest Audio, Nest Hub Max, etc.

### Microsoft Cortana

Producto de Microsoft para competir directamente con Amazon Alexa y Google Assistant. Su posible abandono en próximos años hace presagiar que no tenga mayor soporte en un futuro. Su vinculación fundamental (aunque no exclusiva) a plataformas de Microsoft (e.g. Windows 10 en ordenadores) y el escaso éxito de Microsoft en móviles, han hecho que tenga una menor difusión. Se apuesta por una futura integración en productos más exitosos de Microsoft 365 como Outlook o Teams.

*Dispositivos físicos asociados:* Harman Kardon Invoke y los *Surface Headphones*.

En la Tabla 2, podemos ver un resumen de los cuatros asistente que constituyen en la actualidad el estándar *de facto*.

**Tabla 2. Asistentes conversacionales**

<b>Asistente</b>	<b>Compañía</b>	<b>Fecha de inicio</b>	<b>Dispositivo referente</b>	<b>Wake word (palabra de activación)</b>
Alexa	Amazon	Noviembre 2014	Echo	"Alexa"
Siri	Apple	Octubre 2011	iPhone	"Siri"
Assistant	Google	Mayo 2016	Nest	"Ok Google" "Hey Google"
Cortana	Microsoft	Enero 2015	PC Windows 10	"Hey Cortana"

Junto con los cuatro grandes dominadores, existen otras plataformas que permiten el desarrollo e integración de *bots*, normalmente denominados en esta rama como *chatbots*, incluso en algunos casos también vinculadas parcialmente a estas cuatro compañías. En la siguiente Tabla 3, se enumeran, pero no de forma exhaustiva, algunos de los ejemplos más conocidos.

**Tabla 3. Plataformas de desarrollo de chatbots**

Producto	Descripción	Características
Amazon Lex	Producto de Amazon para el desarrollo de <i>chatbots</i> .	Utiliza el mismo motor de <b>NLU</b> que Alexa, pero implica costes posteriores al segundo año de implantación, debido al uso de la plataforma de Amazon.
Chatcompose	Plataforma de <i>chatBots</i> para marketing y soporte.	Ofrece la opción de <i>live chat</i> , permitiendo la inclusión de agentes humanos en la



Chatfuel	Integración de <i>chatbots</i> .	conversión. Limitado el número de <i>chatbots</i> en su versión gratuita. Solo dispone integración para Facebook Messenger e Instagram.
DialogFlow	Producto de Google para el desarrollo de <i>chatbots</i> .	Simplicidad en su uso y gratuito. Sin límite de <i>chatbots</i> y con una interfaz gráfica intuitiva para la creación.
Microsoft Bot Framework	Denominado también Azure Bot Service. Solución de Microsoft para la creación e integración de <i>chatbots</i> .	Ofrece integración en el ecosistema Microsoft con Office y <i>Teams</i> . Con ciertas limitaciones en la versión gratuita.
Rasa	<i>Framework</i> de código abierto para el aprendizaje automático y creación de <i>chatbots</i> .	Desarrollado en Python, con alta curva de aprendizaje, sin hosting en la nube y en su versión gratuita sin interfaz gráfica.
Watson Assistant	Producto de IBM para el desarrollo de <i>chatbots</i> .	Enfocada a desarrollos más complejos, con modelos conversacionales más complejos.

---

### 3.7. Aplicaciones prácticas en salud

Tomando como base el estudio de (Car et al., 2020), se fija como objetivo de su aplicación **mejorar la accesibilidad, personalización y eficiencia en el cuidado de los pacientes a través de bots**. Su inclusión como tecnología emergente se debe realizar siempre con este objetivo, mejorar en el tratamiento, y no por el mero hecho de introducir un elemento tecnológico adicional. Dicho trabajo realiza una revisión de trabajos sobre bases de datos como MEDLINE, EMBASE, PubMed, Scopus and Cochrane Central, focalizados en el uso de términos como *chatbots*, agentes conversacionales, **IA** conversacional, etc.

Existiendo ya precedentes en la literatura del uso de mensajes de texto y SMSs como herramienta de ayuda para el tratamiento de la salud mental y física (Hall et al., 2015; Rathbone & Prescott, 2017), estos han sido reemplazados por soluciones más modernas basadas en web, integradas en las *apps* de móviles o con interfaces más avanzadas de voz. En particular los **móviles (teléfonos inteligentes o smart phones)**, como dispositivos ubicuos de uso generalizado, se han convertido masivamente en objeto de estudio para su aplicación en la intervención terapéutica.

Por otro lado, los avances en la inteligencia artificial hacen que se dé un empuje a *bots* mucho más avanzados. Aplicados a *coaching (entrenadores virtuales)*, agentes personales personificados o aplicaciones específicas para el control de cambios de comportamiento (Car et al., 2020).

En salud las líneas principales de aplicación en los últimos años son:

- Tratamiento y monitorización.
- Apoyo a servicios de salud.
- Educación de pacientes.



En este tipo de contextos, los trabajos suelen estar más orientados a **conversaciones orientadas a temas**, más que a tareas, y trabajan sobre el paciente, más que al profesional. Suelen ser más habituales las aplicaciones con interfaz de textos frente al uso de voz. Solo en el contexto particular de personas mayores, por cuestiones de problemas en la vista y movilidad, el uso de la voz está más extendido.

Por otro lado, su uso en la *web* favorece su uso y reduce abandono, frente a su uso con *apps* móviles que dificulta algo más la incorporación u *onboarding*, con la salvedad de *apps* de mensajería muy difundidas como Facebook, Messenger, Telegram, Whatsapp, etc.

En (Car et al., 2020) aunque reconociendo la amplitud de campos en los que se aplica en salud, se remarcan algunos campos particulares como:

- Salud mental (Abd-Alrazaq et al., 2020; Bérubé et al., 2021; Piette et al., 2013).
- Neurodegeneración (Li et al., 2020; Rahman et al., 2021).
- Obesidad y diabetes (Steinberg et al., 2014)
- Salud sexual (Bauermeister et al., 2017).

Aunque otras líneas de aplicación también incipientes son:

- Atención primaria (Lee et al., 2021; Fan et al., 2021; Schario et al., 2022).
- Cardiología (Nahar & Lopez-Jimenez, 2022).
- Coaching para adolescentes (Gabielli et al., 2020).
- Dermatología.
- Discapacidad (Masina et al., 2020).
- Enfermería (educación) (Shorey et al., 2019).
- Enfermedades cardiovasculares (Kowalska et al., 2020).
- Enfermedades renales (Fink et al., 2016).
- Enfermedades pulmonares (Gross et al., 2020; Kim et al., 2021).
- Geriátricos (Gudala et al., 2022; Bennion et al., 2020)
- Gestión del estrés (Mauriello et al., 2021).
- Obstetricia (Chung et al., 2021).
- Oncología (Bibault et al., 2019; Greer et al., 2019; Chaix et al., 2019; Greer et al., 2019; Hong et al., 2021).
- Ortopedia (Bian et al., 2020).
- Pediatría (Wong et al., 2021; Espinoza et al., 2020).
- Vacunación (Ferrand et al., 2020; Wijesundara et al., 2020).

Desde la panorámica de la educación temprana en la infancia ha sido abordado con el uso de *PopBots* (Crompton et al., 2018; Williams et al., 2019). Pero desde un punto de vista más cercano a la robótica industrial con los que se interactúa de manera constructivista, con elementos de Lego, sensores, motores, tabletas y *apps* móviles, no en un sentido conversacional y orientadas al paciente. Sin embargo, en la línea más concreta de la **atención temprana**, la aplicación de este tipo de tecnologías es aún escasa, más desde el punto de vista de **asistencia al terapeuta**, y por lo tanto se abre toda una línea de investigación en el futuro sobre el uso de *bots* conversacionales en dicho campo.

## Resumen

Este módulo plantea los elementos básicos para entender en qué consisten y cómo funcionan a un nivel abstracto los *bots*, entendidos como asistentes personales inteligentes. Mostrando además a su posible uso y utilidad en la rama de salud, y apuntando a sus futuras aplicaciones en la **atención temprana**.



## Glosario

**IA:** Inteligencia Artificial

**IPA:** Intelligent Personal Assistant (Asistente Personal Inteligente)

**NLP:** Natural Language Processing (Procesado del Lenguaje Natural)

**NLU:** Natural Language Understanding (Comprensión del Lenguaje Natural)

**TTS:** Text To Speech (Texto a voz)

**STT:** Speech To Text (Voz a texto)

## Bibliografía

### Bibliografía básica

Colby, K. M. (1975). *Artificial Paranoia: Computer Simulation of Paranoid Processes*. Elmsford, N.Y.; Pergamon Press.

Pearl, C. (2016). *Designing voice user interfaces: Principles of conversational experiences*. O'Reilly Media.

Shevat, A. (2017). *Designing bots: Creating conversational experiences*. O'Reilly Media.

### Bibliografía complementaria

Abd-Alrazaq, A. A., Rababeh, A., Alajlani, M., Bewick, B. M., & Househ, M. (2020). Effectiveness and safety of using chatbots to improve mental health: Systematic review and meta-analysis. In *Journal of Medical Internet Research* (Vol. 22, Issue 7, p. e16021). JMIR Publications Inc. <https://doi.org/10.2196/16021>

Bauermeister, J., Giguere, R., Leu, C. S., Febo, I., Cranston, R., Mayer, K., & Carballo-Diéguez, A. (2017). Interactive voice response system: Data considerations and lessons learned during a rectal microbicide placebo adherence trial for young men who have sex with men. *Journal of Medical Internet Research*, 19 (6), e7682. <https://doi.org/10.2196/jmir.7682>

Bérubé, C., Schachner, T., Keller, R., Fleisch, E., Wangenheim, F. v., Barata, F., & Kowatsch, T. (2021). Voice-based conversational agents for the prevention and management of chronic and mental health conditions: Systematic literature review. In *Journal of Medical Internet Research* (Vol. 23, Issue 3, p. e25933). JMIR Publications Inc. <https://doi.org/10.2196/25933>



- Bian, Y., Xiang, Y., Tong, B., Feng, B., & Weng, X. (2020). Artificial intelligence-assisted system in postoperative follow-up of orthopedic patients: Exploratory quantitative and qualitative study. *Journal of Medical Internet Research*, 22 (5), e16896. <https://doi.org/10.2196/16896>
- Bennion, M. R., Hardy, G. E., Moore, R. K., Kellett, S., & Millings, A. (2020). Usability, acceptability, and effectiveness of web-based conversational agents to facilitate problem solving in older adults: Controlled study. *Journal of Medical Internet Research*, 22 (5), e16794. <https://doi.org/10.2196/16794>
- Bérubé, C., Schachner, T., Keller, R., Fleisch, E., Wangenheim, F. v., Barata, F., & Kowatsch, T. (2021). Voice-based conversational agents for the prevention and management of chronic and mental health conditions: Systematic literature review. In *Journal of Medical Internet Research* (Vol. 23, Issue 3, p. e25933). JMIR Publications Inc. <https://doi.org/10.2196/25933>
- Bibault, J. E., Chaix, B., Guillemassé, A., Cousin, S., Escande, A., Perrin, M., Pienkowski, A., Delamon, G., Nectoux, P., & Brouard, B. (2019). A chatbot versus physicians to provide information for patients with breast cancer: Blind, randomized controlled noninferiority trial. *Journal of Medical Internet Research*, 21 (11), e15787. <https://doi.org/10.2196/15787>
- Car, L. T., Dhinakaran, D. A., Kyaw, B. M., Kowatsch, T., Joty, S., Theng, Y. L., & Atun, R. (2020). Conversational agents in health care: Scoping review and conceptual analysis. In *Journal of Medical Internet Research* (Vol. 22, Issue 8, p. e17158). JMIR Publications Inc. <https://doi.org/10.2196/17158>
- Chaix, B., Bibault, J. E., Pienkowski, A., Delamon, G., Guillemassé, A., Nectoux, P., & Brouard, B. (2019). When chatbots meet patients: One-year prospective study of conversations between patients with breast cancer and a chatbot. *JMIR Cancer*, 5(1), e12856. <https://doi.org/10.2196/12856>
- Chung, K., Cho, H. Y., & Park, J. Y. (2021). A chatbot for perinatal women's and partners' obstetric and mental health care: development and usability evaluation study. *JMIR Medical Informatics*, 9 (3), e18607. <https://doi.org/10.2196/18607>
- Crompton, H., Gregory, K., & Burke, D. (2018). Humanoid robots supporting children's learning in an early childhood setting. *British Journal of Educational Technology*, 49 (5), 911–927. <https://doi.org/10.1111/bjet.12654>
- Espinoza, J., Crown, K., & Kulkarni, O. (2020). A guide to chatbots for COVID-19 screening at pediatric health care facilities. *JMIR Public Health and Surveillance*, 6 (2), e18808. <https://doi.org/10.2196/18808>
- Fan, X., Chao, D., Zhang, Z., Wang, D., Li, X., y Tian, F. (2021). *Utilization of Self-Diagnosis Health Chatbots in Real-World Settings: Case Study*. *J Med Internet Res*, 23(1), e19928. <https://doi.org/10.2196/19928>



- Ferrand, J., Hockensmith, R., Houghton, R. F., & Walsh-Buhi, E. R. (2020). Evaluating smart assistant responses for accuracy and misinformation regarding human papillomavirus vaccination: Content analysis study. *Journal of Medical Internet Research*, 22 (8), e19018. <https://doi.org/10.2196/19018>
- Fink, J. C., Doerfler, R. M., Yoffe, M. R., Diamantidis, C. J., Blumenthal, J. B., Siddiqui, T., Gardner, J. F., Snitker, S., & Zhan, M. (2016). Patient-Reported Safety Events in Chronic Kidney Disease Recorded With an Interactive Voice-Inquiry Dial-Response System: Monthly Report Analysis. *Journal of Medical Internet Research*, 18 (5), e5203. <https://doi.org/10.2196/jmir.5203>
- Greer, S., Ramo, D., Chang, Y. J., Fu, M., Moskowitz, J., & Haritatos, J. (2019). Use of the chatbot “vivibot” to deliver positive psychology skills and promote well-being among young people after cancer treatment: Randomized controlled feasibility trial. *JMIR MHealth and UHealth*, 7(10), e15018. <https://doi.org/10.2196/15018>
- Gabrielli, S., Rizzi, S., Carbone, S., & Donisi, V. (2020). A chatbot-based coaching intervention for adolescents to promote life skills: Pilot study. *JMIR Human Factors*, 7 (1), e16762. <https://doi.org/10.2196/16762>
- Gross, C., Kohlbrenner, D., Clarenbach, C. F., Ivankay, A., Brunschwiler, T., Nordmann, Y., & v Wangenheim, F. (2020). A Telemonitoring and Hybrid Virtual Coaching Solution “CAir” for Patients with Chronic Obstructive Pulmonary Disease: Protocol for a Randomized Controlled Trial. *JMIR Research Protocols*, 9 (10), e20412. <https://doi.org/10.2196/20412>
- Gudala, M., Ross, M. E. T., Mogalla, S., Lyons, M., Ramaswamy, P., & Roberts, K. (2022). Benefits of, Barriers to, and Needs for an Artificial Intelligence-Powered Medication Information Voice Chatbot for Older Adults: Interview Study With Geriatrics Experts. *JMIR Aging*, 5 (2), e32169. <https://doi.org/10.2196/32169>
- Gupta, J., Singh, V., y Kumar, I. (2021). "Florence- A Health Care Chatbot," 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS) (p. 504-508). <https://doi.org/10.1109/ICACCS51430.2021.9442006>
- Hall, A. K., Cole-Lewis, H., & Bernhardt, J. M. (2015). Mobile Text Messaging for Health: A Systematic Review of Reviews. *Annual Review of Public Health*, 36(1), 393–415. <https://doi.org/10.1146/annurev-publhealth-031914-122855>
- Hong, G., Folcarelli, A., Less, J., Wang, C., Erbası, N., & Lin, S. (2021). Voice assistants and cancer screening: A comparison of alexa, siri, google assistant, and cortana. *Annals of Family Medicine*, 19 (5), 447–449. <https://doi.org/10.1370/AFM.2713>
- Kim, A. J., Yang, J., Jang, Y., & Baek, J. S. (2021). Acceptance of an informational antituberculosis chatbot among korean adults: Mixed methods research. *JMIR MHealth and UHealth*, 9 (11), e26424. <https://doi.org/10.2196/26424>



- Kowalska, M., Gładys, A., Kalańska-Lukasik, B., Gruz-Kwapisz, M., Wojakowski, W., & Jadczyk, T. (2020). Readiness for voice technology in patients with cardiovascular diseases: Cross-sectional study. *Journal of Medical Internet Research*, 22 (12), e20456. <https://doi.org/10.2196/20456>
- Lee, H., Kang, J., & Yeo, J. (2021). Medical specialty recommendations by an artificial intelligence chatbot on a smartphone: Development and deployment. *Journal of Medical Internet Research*, 23 (5), e27460. <https://doi.org/10.2196/27460>
- Li, J., Maharjan, B., Xie, B., & Tao, C. (2020). A personalized voice-based diet assistant for caregivers of alzheimer disease and related dementias: System development and validation. *Journal of Medical Internet Research*, 22 (9), e19897. <https://doi.org/10.2196/19897>
- Mauriello, M. L., Tantivasadakarn, N., Mora-Mendoza, M. A., Lincoln, E. T., Hon, G., Nowruzi, P., Simon, D., Hansen, L., Goenawan, N. H., Kim, J., Gowda, N., Jurafsky, D., & Paredes, P. E. (2021). A suite of mobile conversational agents for daily stress management (popbots): Mixed methods exploratory study. *JMIR Formative Research*, 5 (9), e25294. <https://doi.org/10.2196/25294>
- Masina, F., Orso, V., Pluchino, P., Dainese, G., Volpato, S., Nelini, C., Mapelli, D., Spagnolli, A., & Gamberini, L. (2020). Investigating the accessibility of voice assistants with impaired users: Mixed methods study. *Journal of Medical Internet Research*, 22 (9), e18431. <https://doi.org/10.2196/18431>
- Nahar, J. K., & Lopez-Jimenez, F. (2022). Utilizing Conversational Artificial Intelligence, Voice, and Phonocardiography Analytics in Heart Failure Care. In *Heart Failure Clinics* (Vol. 18, Issue 2, pp. 311–323). Elsevier Inc. <https://doi.org/10.1016/j.hfc.2021.11.006>
- Piette, J. D., Sussman, J. B., Pfeiffer, P. N., Silveira, M. J., Singh, S., & Lavieri, M. S. (2013). Maximizing the value of mobile health monitoring by avoiding redundant patient reports: Prediction of depression-related symptoms and adherence problems in automated health assessment services. *Journal of Medical Internet Research*, 15 (7), e2582. <https://doi.org/10.2196/jmir.2582>
- Prize L. Mitsuku Wins 2019 Loebner Prize and Best Overall Chatbot at AISB X. AISB – The Society for the Study of Artificial Intelligence and Simulation of Behaviour. 2019. URL: <https://aisb.org.uk/mitsuku-wins-2019-loebner-prize-and-best-overall-chatbot-at-aisb-x/>
- Pumplun, L., Fecho, M., Wahl, N., Peters, F., y Buxmann, P. (2021). Adoption of Machine Learning Systems for Medical Diagnostics in Clinics: Qualitative Interview Study. *J Med Internet Res*, 23(10), e29301. <https://doi.org/10.2196/29301>
- Rathbone, A. L., & Prescott, J. (2017). The use of mobile apps and SMS messaging as physical and mental health interventions: Systematic review. In *Journal of Medical Internet Research* (Vol. 19, Issue 8, p. e7740). JMIR Publications Inc. <https://doi.org/10.2196/jmir.7740>



- Rahman, W., Lee, S., Islam, M. S., Antony, V. N., Ratnu, H., Ali, M. R., Mamun, A. al, Wagner, E., Jensen-Roberts, S., Waddell, E., Myers, T., Pawlik, M., Soto, J., Coffey, M., Sarkar, A., Schneider, R., Tarolli, C., Lizarraga, K., Adams, J., ... Hoque, E. (2021). Detecting parkinson disease using a web-based speech task: Observational study. *Journal of Medical Internet Research*, 23 (10), e26305. <https://doi.org/10.2196/26305>
- Schario, M. E., Bahner, C. A., Widenhofer, T. v., Rajaballey, J. I., & Thatcher, E. J. (2022). Chatbot-Assisted care management. *Professional Case Management*, 27 (1), 19–25. <https://doi.org/10.1097/NCM.0000000000000504>
- Shan, Y., Ji, M., Xie, W., Qian, X., Li, R., Zhang, X., y Hao, T. (2022). Language Use in Conversational Agent–Based Health Communication: Systematic Review. *J Med Internet Res*, 24(7), e37403. <https://doi.org/10.2196/37403>
- Shorey, S., Ang, E., Yap, J., Ng, E. D., Lau, S. T., & Chui, C. K. (2019). A virtual counseling application using artificial intelligence for communication skills training in nursing education: Development study. *Journal of Medical Internet Research*, 21 (10), e14658. <https://doi.org/10.2196/14658>
- Steinberg, D. M., Levine, E. L., Lane, I., Askew, S., Foley, P. B., Puleo, E., & Bennett, G. G. (2014). Adherence to self-monitoring via interactive voice response technology in an eHealth intervention targeting weight gain prevention among black women: Randomized controlled trial. *Journal of Medical Internet Research*, 16 (4), e2996. <https://doi.org/10.2196/jmir.2996>
- Wijesundara, J. G., Fukunaga, M. I., Ogarek, J., Barton, B., Fisher, L., Preusse, P., Sundaresan, D., Garber, L., Mazor, K. M., & Cutrona, S. L. (2020). Electronic health record portal messages and interactive voice response calls to improve rates of early season influenza vaccination: Randomized controlled trial. *Journal of Medical Internet Research*, 22 (9), e16373. <https://doi.org/10.2196/16373>
- Williams, R., Park, H. W., Oh, L., & Breazeal, C. (2019). Popbots: Designing an artificial intelligence curriculum for early childhood education. *33rd AAAI Conference on Artificial Intelligence, AAAI 2019, 31st Innovative Applications of Artificial Intelligence Conference, IAAI 2019 and the 9th AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019*, 33 (01), 9729–9736. <https://doi.org/10.1609/aaai.v33i01.33019729>
- Wong, J., Foussat, A. C., Ting, S., Acerbi, E., van Elburg, R. M., & Chien, C. M. (2021). A chatbot to engage parents of preterm and term infants on parental stress, parental sleep, and infant feeding: Usability and feasibility study. *JMIR Pediatrics and Parenting*, 4 (4), e30169. <https://doi.org/10.2196/30169>

## Web

- Amazon (2022a) Amazon Alexa official site: What is Alexa? (n.d.). Amazon (Alexa). <https://developer.amazon.com/es-ES/alexa> Última consulta 17 de julio del 2022.



Amazon (2022b) Chatbot | Deep learning | Amazon Lex. (n.d.). Amazon Web Services, Inc. <https://aws.amazon.com/es/lex/>. Última consulta 17 de julio del 2022.

Apple (2022) Siri. (n.d.). Apple (España). <https://www.apple.com/es/siri/> Última consulta 17 de julio del 2022.

Chatcompose (2022) Plataforma de chatbots para marketing Y soporte. (n.d.). ChatCompose - Chatbot Platform for Sales and Support Automation. <https://www.chatcompose.com/es.html>. Última consulta 17 de julio del 2022.

Chatfuel (2022, 11). Chatfuel | Customer support and sales automation. <https://chatfuel.com/>. Última consulta 17 de julio del 2022.

IBM (2022) IBM Watson assistant - Virtual agent. (n.d.). IBM - United States. <https://www.ibm.com/products/watson-assistant>. Última consulta 17 de julio del 2022.

Google (2022). Dialogflow. <https://dialogflow.cloud.google.com/>. Última consulta 17 de julio del 2022.

Meta (2022) Wit.ai. <https://wit.ai/> Última consulta 17 de julio del 2022.

Microsoft (2022a) ¿Que es Cortana? (n.d.). Microsoft Support. <https://support.microsoft.com/es-es/topic/-qu%C3%A9-es-cortana-953e648d-5668-e017-1341-7f26f7d0f825>. Última consulta 17 de julio del 2022.

Microsoft (2022b) Microsoft bot framework. (n.d.). <https://dev.botframework.com/> Última consulta 17 de julio del 2022.

Open source conversational AI. (2020, December 1). Rasa. <https://rasa.com/> Última consulta 17 de julio del 2022.

UNICEF (2022a) Safer Chatbots | UNICEF. (n.d.). Retrieved July 14, 2022, from <https://www.unicef.org/documents/safer-chatbots>

UNICEF (2022b) Safer Chatbots Implementation Guide | UNICEF. (n.d.). Retrieved July 14, 2022, from <https://www.unicef.org/documents/safer-chatbots-implementation-guide>

