

Asistentes Virtuales

Merino Vidal Mateo Alejandro

Departamento de Informática

Universidad Mayor de San Simón

Docente: Lic. García Pérez Carmen Rosa

Materia: Inteligencia Artificial I

Cochabamba, Bolivia

202301308@est.umss.edu

Abstract—El presente artículo presenta una demostración de diversos asistentes virtuales, dando a conocer sus características, funcionalidades, plataformas en las que se ejecutan y como funcionan. Para este análisis, se tomaron en cuenta 5 asistentes virtuales, los cuales son: Siri, Alexa, Cortana, Google Assistant y Bixby.

I. INTRODUCCIÓN

En la actualidad, gracias a los avances de la inteligencia artificial, se ha podido desarrollar diversos sistemas, siendo uno de estos los denominados asistentes virtuales, los cuales han transformado la interacción entre los usuarios y la tecnología.

Gracias a esto, las tareas cotidianas son mas sencillas, permitiendo la automatización de ciertas tareas, desde la gestión laboral hasta las asistencias en entornos laborales, mediante el uso de la inteligencia artificial y el procesamiento de lenguaje natural.

Estos sistemas permiten realizar diversas acciones como las búsquedas, gestión de dispositivos, programar recordatorios y ejecutar múltiples acciones mediante comandos de voz.

Este informe analiza las características, funcionalidades y plataformas en las que operan cinco de los asistentes virtuales más representativos del mercado: Siri, Alexa, Cortana, Google Assistant y Bixby.

Se examinará su funcionamiento, las tecnologías que los respaldan y sus aplicaciones en diferentes ámbitos, con el objetivo de comprender su impacto en la vida cotidiana y el desarrollo de nuevas interacciones.

II. DESARROLLO DE CONTENIDOS

A. Procesamiento de Lenguaje Natural (PNL)

Es un proceso que le permite a las máquinas entender el lenguaje natural del ser humano. Este procedimiento empieza con la obtención y preprocesamiento del texto, en donde se

recolecta información de diversas fuentes, para después llegar a normalizar el texto, es decir, que elimina tildes, mayúsculas y signos de puntuación.

Posteriormente, se realizan los siguientes procesos:

- 1) Se realiza la **tokenización** del texto, en donde se lo descompone en palabras, frases u oraciones.
- 2) Se eliminan las *stop words*, las cuales son palabras comunes de un lenguaje que no suelen aportar algún significado clave, estando entre estas artículos, preposiciones y pronombres, como *el, la, de, en, por, y, o*, etc., en español. Posteriormente, se da la aplicación de **stemming** para reducir las palabras a su raíz o forma base, eliminando los sufijos y prefijos.
- 3) Una vez que ya se tiene el texto, se procede a transformarlo en una representación numérica mediante técnicas como **TF-IDF** y **Bag of Words (BoW)**, permitiendo capturar relaciones semánticas.
- 4) Se analiza el texto mediante **etiquetado de palabras (POS Tagging)** para reconocer categorías gramaticales, permitiendo realizar un análisis de las dependencias para comprender la estructura de las frases y reconocer entidades nombradas, con la finalidad de detectar nombres de personas, organizaciones y fechas.
- 5) Se realiza el **análisis de sentimiento**, el cual es una forma de análisis semántico, permitiendo calcular la polaridad del texto (positivo, negativo, neutro), basándose en el significado de las palabras y su contexto.

B. Siri

1) Definición Conceptual

Es un asistente virtual, incluido dentro de los sistemas operativos de los dispositivos Apple, siendo considerado como un logro de la inteligencia artificial llevado a los dispositivos digitales, que permiten reconocer la voz del

usuario y contestar las ordenes más variadas.

Este asistente es capaz de resolver requerimientos del usuario, contestar preguntas de la vida cotidiana, hacer llamadas, resolver operaciones matemáticas, hacer pronóstico del tiempo y otras cosas.

Siri se introdujo por primera vez con el iPhone 4S en octubre de 2011, recibiendo su nombre de la versión abreviada del nombre noruego Sigrid por parte del cocreador, Dag Kittlaus.

2) *Funcionalidades*

- Realizar búsquedas en Internet y responder preguntas básicas, proporcionando información general del tema, como también es capaz de analizar artículos relevantes y mostrar enlaces útiles.
- Realizar operaciones matemáticas como la suma, resta, división, multiplicación. También permite la conversión de unidades, como la transformación de la temperatura de grados Kelvin a Celsius.
- Realizar llamadas telefónicas y FaceTime, permitiendo comunicarse con contactos mediante comandos de voz y realizar video-llamadas.
- Redactar y enviar mensajes a través de iMessage, WhatsApp y otras aplicaciones compatibles, mediante dictado por voz. Asimismo, también es capaz de leer mensajes entrantes y responder de forma automática.
- Agendar eventos de calendario o recordatorios, como también programar reuniones, estableciendo recordatorios por ubicación y organizando listas de tareas pendientes.
- Traducir idiomas en tiempo real, permitiendo la conversión de frases y palabras a varios idiomas, mediante la pronunciación de voz.

3) *Características*

- Esta integrada en el ecosistema Apple, funcionando de manera sincronizada en iPhone, iPad, Mac, Apple Watch, HomePod y Apple TV mediante iCloud.
- Se activa por voz con "Hey Siri" o manteniendo presionado un botón físico, ofreciendo un acceso rápido sin necesidad de tocar la pantalla.
- Es compatible con aplicaciones de terceros como WhatsApp, Spotify y Google Maps, permitiendo el

control de funciones dentro de esas apps mediante comandos de voz.

- Funciona en modo manos libres y con opciones de accesibilidad, incluyendo control por voz, lectura de pantalla y dictado, facilitando su uso a personas con discapacidad.
- Garantiza la seguridad y privacidad, ya que procesa las solicitudes sin la necesidad de guardar datos personales, como también de compartir la información personal del usuario sin su consentimiento.

4) *Plataformas de Ejecucion*

- **iOS (iPhone):** Disponible en todos los modelos compatibles con Siri, permitiendo el control por voz para realizar llamadas, enviar mensajes, reproducir música y más.
- **iPadOS (iPad):** Funciona de manera similar a iOS, ofreciendo soporte para multitarea y gestión de aplicaciones mediante comandos de voz.
- **macOS (Mac):** Integrado en macOS, permite ejecutar comandos para abrir aplicaciones, buscar archivos, hacer cálculos y ajustar configuraciones del sistema.
- **watchOS (Apple Watch):** Disponible en Apple Watch, permitiendo el uso de comandos de voz para responder mensajes, iniciar llamadas y gestionar la actividad física.
- **CarPlay:** Integrado en CarPlay, permite la navegación GPS, llamadas, mensajes y control de música sin necesidad de tocar la pantalla del vehículo.

C. *Alexa*

1) *Definición Conceptual*

Es un asistente virtual, basado en el enfoque de inteligencia artificial, desarrollado por Amazon. Se puede acceder a ella a través de dispositivos habilitados como el Echo, que actúan como altavoces inteligentes y se conectan a la red Wi-Fi del hogar.

Es capaz de realizar múltiples funciones como responder preguntas, reproducir música y controlar dispositivos domésticos inteligentes.

Este asistente funciona a través de una combinación de tecnologías avanzadas, incluyendo el procesamiento de lenguaje natural (NLP), el aprendizaje automático y la inteligencia artificial. Las cuales le permiten a Alexa

comprender y responder de forma precisa y eficiente a los comandos de voz del usuario.

2) Funcionalidades

- Controlar dispositivos domésticos inteligentes, permitiendo encender y apagar luces, ajustar la temperatura del termostato, cerrar puertas con cerraduras inteligentes y gestionar electrodomésticos como cafeteras y aspiradoras robot mediante comandos de voz.
- Proporcionar información y noticias actualizadas, conectándose a una amplia variedad de fuentes en línea para responder preguntas de manera rápida y precisa.
- Gestionar tareas y listas, permitiendo la creación de recordatorios, alarmas y listas de pendientes para una mejor organización.
- Ejecutar habilidades y aplicaciones desarrolladas por terceros, ampliando sus funcionalidades y permitiendo realizar tareas como reservar en restaurantes o solicitar un servicio de transporte como Uber.

3) Características

- Comprende el lenguaje humano de forma natural gracias a su tecnología de procesamiento de lenguaje natural, lo que permite un diálogo fluido, preciso e intuitivo con los usuarios.
- Cuenta con un avanzado sistema de seguridad y privacidad, permitiendo a los usuarios revisar y eliminar grabaciones de voz, así como desactivar el micrófono cuando lo deseen.
- Su diseño y compatibilidad la hacen adaptable a distintos entornos, desde hogares hasta oficinas, integrándose estéticamente con diferentes estilos y necesidades.
- Ofrece opciones de personalización al aprender de los hábitos del usuario, permitiendo respuestas adaptadas y la creación de skills personalizadas mediante la plataforma blueprints.
- Requiere de activación por voz mediante la palabra de activación predeterminada “Alexa”, seguida de la instrucción deseada, como por ejemplo: “Alexa, ¿Que tiempo hará hoy?”.

4) Plataformas de Ejecucion

- **Dispositivos Echo (Echo, Echo Dot, Echo Show, Echo Studio, Echo Flex):** Altavoces inteligentes de Amazon que permiten el uso de Alexa para gestionar dispositivos del hogar inteligente y responder consultas.
- **Fire TV (Fire TV Stick y Fire TV Cube):** Dispositivos de streaming con Alexa integrada, permitiendo el control por voz para la navegación, búsqueda y reproducción de contenido.
- **Fire Tablets (Amazon Fire HD y Fire Kids Edition):** Tabletas de Amazon con acceso a funciones de Alexa sin necesidad de dispositivos Echo.
- **Automóviles con Alexa Built-in:** Vehículos de marcas como Ford y BMW con Alexa integrada, permitiendo navegación, llamadas y entretenimiento por voz sin depender de un teléfono.
- **Electrodomésticos y dispositivos inteligentes:** Alexa se encuentra disponible en Smart TVs, refrigeradores y otros dispositivos del hogar, permitiendo su control mediante comandos de voz sin necesidad de dispositivos adicionales.

D. Cortana

1) Definición Conceptual

Es un asistente virtual creado por Microsoft para Windows 10, Windows 10 Mobile, Windows Phone 8.1, altavoz inteligente Invoke, Microsoft Band, etc. Tuvo su primera presentación oficial el 2 de abril de 2014 durante la conferencia Build 2014 de Microsoft. Sin embargo, debido a su baja popularidad y el surgimiento de otros asistentes más populares como Google Assistant, Alexa y Siri, terminó siendo discontinuado.

Este asistente podía establecer recordatorios, reconocer voz natural sin la necesidad de usar el teclado y responder preguntas utilizando información del motor de búsqueda de Bing.

2) Funcionalidades

- Crear recordatorios sin necesidad de mucha intervención. Por ejemplo, se puede decir: “Hola Cortana, recuérdame que asista a la conferencia telefónica de Tom a las 6 p. m.”.
- Controlar dispositivos domésticos inteligentes, como las luces Philips Hue, con solo dar comandos de voz.

- Buscar palabras o frases mientras se navega por la web. Esta función está integrada en el navegador Microsoft Edge y facilita la búsqueda en línea.
- Obtener previsiones meteorológicas mediante comandos de voz, como “Hola Cortana, ¿cuál es el nivel de humedad actual en Los Ángeles?”.
- Brindar indicaciones y ayudar en la navegación, permitiendo la búsqueda de rutas y proporcionando instrucciones para llegar a un destino.

3) Características

- Permite adaptar la interacción al aprender preferencias, como nombres, ubicaciones frecuentes, intereses y rutinas diarias. De este modo, brinda información relevante y sugerencias sin necesidad de realizar búsquedas manuales.
- Se integra con aplicaciones como Microsoft Office, Spotify y LinkedIn, permitiendo realizar diversas tareas sin salir de la aplicación. Por ejemplo, es posible enviar correos electrónicos o consultar mensajes de LinkedIn directamente desde Cortana.
- Ofrece recomendaciones basadas en el historial de búsqueda y preferencias. Por ejemplo, al buscar un restaurante, proporciona información como ubicación, horario de atención y opiniones de clientes, además de sugerir opciones similares según búsquedas previas.
- Permite la interacción sin necesidad de tocar el teléfono o la computadora, lo que resulta especialmente útil al conducir o realizar otras tareas que requieren el uso de las manos.
- Utiliza diversas tecnologías para mejorar su precisión y capacidad de interacción. Entre ellas se incluyen la conversión de voz a texto (STT) y texto a voz (TTS), control y supresión de ruido, procesamiento del lenguaje natural (PLN), aprendizaje profundo, visión artificial (CV) e inteligencia emocional (IE).

4) Plataformas de Ejecucion

- **Windows 10:** Integrado como asistente digital en el sistema operativo, permitiendo interacción por voz y texto para ejecutar comandos, realizar búsquedas y gestionar tareas.

- **Windows 10 Mobile:** Disponible en dispositivos móviles con Windows 10, facilitando el acceso a funciones del sistema y aplicaciones mediante comandos de voz.
- **Windows Phone 8.1:** Introducido como el primer sistema operativo móvil de Microsoft con Cortana, ofreciendo asistencia personalizada y gestión de recordatorios.
- **Microsoft Band:** Integrado en la pulsera inteligente de Microsoft para permitir el control por voz de notificaciones, recordatorios y actividades de salud.
- **Xbox One:** Incorporado en la consola de videojuegos para ejecutar comandos de voz, como iniciar juegos, ajustar configuraciones y controlar contenido multimedia.

E. Google Assistant

1) Definición Conceptual

Es un asistente virtual desarrollado por Google que utiliza inteligencia artificial para procesar comandos de voz y texto, permitiéndole responder preguntas, automatizar tareas y controlar dispositivos inteligentes.

Fue presentado por primera vez el 18 de mayo de 2016 durante la conferencia Google I/O. Este asistente se integra con una variedad de plataformas, como smartphones, altavoces inteligentes, televisores, etc.

2) Funcionalidades

- Permite establecer alarmas, recordatorios y planificar agendas. Los usuarios pueden consultar su calendario con comandos de voz como “¿qué hay en mi calendario mañana?” o “¿cuál es mi agenda de hoy?”.
- Facilita la gestión de la cuenta de Google mediante comandos de voz, permitiendo acceder a ajustes, configuraciones de privacidad e historial de búsqueda.
- Permite gestionar diversas funciones del dispositivo, como activar/desactivar Wi-Fi, Bluetooth, linterna y modo avión. También posibilita ajustar el brillo, el volumen, hacer fotos y abrir aplicaciones con comandos de voz.
- Proporciona acceso inmediato a datos y consultas sin necesidad de búsquedas manuales. Los usuarios pueden obtener información sobre celebridades, estrenos de películas, recetas y otros temas simplemente preguntando.

3) Características

- Permite realizar varias preguntas seguidas sin necesidad de repetir "Hey Google" en cada interacción.
- Tecnología avanzada que le permite realizar llamadas telefónicas en nombre del usuario para reservar citas en restaurantes o peluquerías de manera natural.
- Tiene acceso directo a la base de datos de Google, proporcionando respuestas más precisas y actualizadas en comparación con otros asistentes.
- Permite configurar una secuencia de comandos con un solo comando de voz, como encender luces, leer noticias y ajustar el termostato con una sola orden.
- Puede traducir conversaciones en tiempo real en varios idiomas, funcionando como un traductor instantáneo de voz.

4) Plataformas de Ejecucion

- **Google Nest (Nest Hub, Nest Audio, Nest Mini):** Dispositivos inteligentes de Google diseñados para el hogar, integrando Google Assistant para control por voz de luces, termostatos, cámaras y otros dispositivos conectados.
- **Google Nest Secure:** Sistema de seguridad para el hogar con Google Assistant integrado, permitiendo el control por voz de alarmas, sensores y cámaras de seguridad.
- **Wear OS (Relojes inteligentes):** Sistema operativo de Google para relojes inteligentes, permitiendo el acceso a Google Assistant para consultar información, gestionar eventos y controlar dispositivos conectados.
- **Google Clips:** Cámara inteligente de Google que utilizaba inteligencia artificial y Google Assistant para capturar automáticamente momentos sin intervención del usuario.
- **Chromebooks:** Laptops con Chrome OS que integran Google Assistant para realizar búsquedas, gestionar tareas y controlar dispositivos conectados mediante comandos de voz.

F. Bixby

1) Definición Conceptual

Es un asistente virtual desarrollado por Samsung, diseñado

para ofrecer una experiencia de usuario más intuitiva e integrada en los dispositivos de la marca.

Tuvo su lanzamiento en 2017, a partir del cual ha ido evolucionado y mejorando la interacción con los dispositivos Samsung, permitiendo realizar una amplia gama de tareas a través de comandos de voz, texto y toques.

A diferencia de otros asistentes virtuales, Bixby está profundamente integrado en el ecosistema Samsung, lo que le permite controlar una amplia variedad de funciones y aplicaciones del dispositivo.

2) Funcionalidades

- Lo que se puede hacer con el tacto, se puede hacer con la voz. Permite abrir aplicaciones, modificar configuraciones y realizar acciones avanzadas sin tocar el dispositivo.
- Permite crear rutinas personalizadas basadas en patrones de uso, activando funciones automáticamente según la hora, ubicación o estado del dispositivo.
- Ofrece integración total con SmartThings, permitiendo controlar electrodomésticos, televisores y dispositivos inteligentes desde un solo asistente.
- Permite asignar frases cortas a tareas complejas, facilitando la ejecución de múltiples acciones con un solo comando de voz.
- Distingue entre diferentes usuarios en un mismo dispositivo, proporcionando respuestas y configuraciones personalizadas según la voz de cada persona.

3) Características

- Diseñado específicamente para el ecosistema de Samsung, lo que permite un control avanzado de dispositivos Galaxy, electrodomésticos y Smart TVs de la marca.
- Ofrece una pantalla de inicio personalizada con información relevante basada en el uso del usuario, integrando notificaciones, recordatorios y sugerencias en una interfaz dinámica.
- Permite el reconocimiento de objetos, texto y códigos QR a través de la cámara del teléfono, facilitando la traducción, compras y búsqueda de información visual.
- Permite interactuar con comandos de voz sin necesidad

de repetir la activación "Hi Bixby", manteniendo un flujo de conversación natural.

- Aprende del comportamiento del usuario para proporcionar respuestas más personalizadas y adaptadas a las preferencias individuales.

- La privacidad y seguridad siguen siendo aspectos clave en el uso de asistentes virtuales, ya que estos están constantemente interactuando y recopilando información personal del usuario, por lo que cada empresa ha implementado medidas para la protección de los datos.

4) Plataformas de Ejecucion

- **Samsung Galaxy Smartphones:** Integrado en dispositivos Galaxy como S y Note, permitiendo el control avanzado del sistema mediante comandos de voz y Bixby Routines.
- **Samsung Galaxy Watch:** Disponible en relojes inteligentes de Samsung con Wear OS y Tizen, permitiendo la gestión de notificaciones, llamadas y funciones del dispositivo con comandos de voz.
- **Samsung Smart TVs:** Incorporado en televisores inteligentes de Samsung, facilitando el control de canales, volumen y aplicaciones mediante comandos de voz sin necesidad de control remoto.
- **Samsung Family Hub Refrigerators:** Integrado en refrigeradores inteligentes de Samsung, permitiendo la gestión de listas de compras, control del calendario familiar y visualización de contenido multimedia por voz.
- **Samsung Galaxy Buds:** Auriculares inalámbricos con compatibilidad con Bixby, permitiendo la ejecución de comandos sin necesidad de interactuar con el teléfono.

III. CONCLUSIONES

- Los asistentes virtuales han ido evolucionando a lo largo del tiempo, gracias a los avances de la inteligencia artificial y el procesamiento de lenguaje natural, permitiendo establecer una interacción mas intuitiva entre el ser humano y la tecnología.
- Cada asistente virtual posee características y funcionalidades únicas, que lo diferencia de otros, como se puede observar:
 - Siri: Se integra profundamente con el ecosistema Apple.
 - Alexa: Destaca en la automatización del hogar, a traves de electrodomésticos inteligentes.
 - Google Assistant: Destaca en la búsqueda inteligente.
 - Cortana: Destaca en la integración con Microsoft.
 - Bixby: Se integra profundamente en el ecosistema Samsung.