

# Prototipo de control por comandos de voz para personas con discapacidad motriz

## *Trabajo Terminal No. 2020–B094*

*Alumnos: Domínguez Castro Israel Axel*

*Directores: Jorge Cortés Galicia, Miriam Pescador Rojas.*

*e-mail: [idinguezc1200@alumno.ipn.mx](mailto:idinguezc1200@alumno.ipn.mx)*

**Resumen** - El presente trabajo terminal tiene como objetivo principal desarrollar un sistema que permita controlar una computadora por medio de la voz, permitiendo a los usuarios controlar algunas de las funciones de la computadora con pronunciar comandos de voz. Esta idea está enfocada principalmente a personas con alguna discapacidad motriz para que así puedan realizar sus actividades con mayor facilidad. Algunas personas con dichas discapacidades no pueden usar una computadora de manera cotidiana lo que les impide obtener empleo. El propósito de esta herramienta es apoyar a este tipo de personas en el uso de computadoras.

**Palabras clave** - Comandos de voz personalizables, Inteligencia artificial, Procesamiento de lenguaje natural, Control de sistema operativo.

## 1. Introducción

Conforme pasa el tiempo, el uso de las computadoras se vuelve cada vez más indispensable y las personas hacen uso de estas para realizar ciertas tareas con mayor facilidad. En el sector público y privado se emplean sistemas que están contenidos en computadoras con las que realizan el servicio o producto que ofrecen al público, siendo estas operadas en su mayoría por personas.

Existen casos en el que las personas con alguna discapacidad motriz se ven limitados al uso de dispositivos electrónicos debido a que un alto porcentaje de estos son operados a través de las manos. Es por eso que surge la necesidad de atacar este problema y encontrar una solución para que la mayoría de los usuarios puedan controlar una PC sin tantas complicaciones y de esta manera no sufrir algún tipo de discriminación.

En la actualidad, existen algunas herramientas que ofrecen un servicio similar, pero estas tienen límites que no permiten aprovechar del todo el uso del sistema operativo (SO) que lo contiene. La aplicación que se propone desarrollar en este trabajo terminal busca abarcar la mayor cantidad de uso que se le puede dar a una computadora a través de comandos de voz.

**Inteligencia artificial.** Podemos definir IA como «la habilidad de los ordenadores para hacer actividades que normalmente requieren inteligencia humana». Pero, para brindar una definición más detallada, podríamos decir que la IA es la capacidad de las máquinas para usar algoritmos, aprender de los datos y utilizar lo aprendido en la toma de decisiones tal y como lo haría un ser humano.

Para poder empezar a reflexionar acerca del gran impacto que la IA tendrá en nuestra vida, conviene saber que las tecnologías de IA han comenzado a desarrollar habilidades de un ser humano, tales como ver (visión artificial), oír (reconocimiento de voz) y entender (procesamiento del lenguaje natural) [7].

**Procesamiento del lenguaje natural.** Cómo todo sistema de análisis lingüístico se compone de una arquitectura de niveles, la construcción del lenguaje natural se produce de forma secuencial según los niveles que lo constituyen. Partimos de una idea o concepto ajustado a la situación en la que nos encontramos y el procesamiento cerebral la transforma en aquellos sonidos que la expresan. A la hora de realizar un análisis, el sentido cambia. Partimos de la comprensión de las unidades más básicas, de esos sonidos, para darles forma y terminar por contextualizarlos adquiriendo entonces el significado justo [8].

**Asistente virtual de voz.** Los asistentes virtuales de voz son una tecnología con base en la inteligencia artificial que funciona a través de una interfaz de voz, de manera que permite que los usuarios interactúen con Internet únicamente a través del habla [9].

Con todos estos datos surge la necesidad de desarrollar un asistente virtual por voz que sea totalmente personalizable y que se pueda utilizar solo con la voz, esto ayudará a las personas con discapacidades motrices a utilizarlo sin ningún inconveniente ya que solo necesitarán el timbre de su voz para poder activarlo y enseguida comenzar a mencionar cada uno de los diversos comandos a los que podrán tener acceso y así mismo a los que podrán agregar ellos mismos dependiendo de sus necesidades actuales, ya que cada persona que utilice el asistente tendrá la accesibilidad de configurarlo totalmente a su gusto.

Los comandos de voz podrán ser personalizados por el usuario y así asignar una tarea específica, y como ya se ha mencionado, existen algunas aplicaciones que ofrecen un control similar, pero uno de los inconvenientes más relevantes es que se necesita activar y desactivar el uso cada que se va a decir un comando. El sistema propuesto tiene como objetivo permanecer a la escucha sin que el ruido ambiente pueda interferir en el funcionamiento o este tome como comando el audio que esté presente.

También se podrán incluir perfiles de acuerdo a las necesidades del usuario que estén presentes en el SO, además la aplicación al ser iniciada podrá comenzar a escuchar los comandos de voz para ser ejecutados.

#### **FUNCIONES DE MOUSE**

- Clic derecho
- Clic izquierdo
- Arrastrar

#### **FUNCIONES DEL TECLADO**

- Escritura de texto y caracteres especiales
- Activar y desactivar mayúsculas
- Letra por letra

#### **FUNCIONES DEL ADMINISTRADOR DE ARCHIVOS**

- Abrir y cerrar carpetas

#### **PROGRAMAS que podrá utilizar**

- Paquetería de Office

#### **FUNCIONES DEL NAVEGADOR WEB**

- Abrir páginas web
- Realizar búsquedas
- Abrir correo electrónico

Los sistemas similares que se han desarrollado se muestran en la tabla 1.

*Tabla 1. Resumen de productos similares*

<b>Nombre del sistema</b>	<b>Proveedor</b>	<b>Descripción</b>	<b>Alcance</b>
Cortana[2]	Windows	Sistema utilizado como asistente personal para realizar ciertas actividades.	-Viene incluido en el SO -Tiene interacción con el usuario -Se actualiza constantemente
Siri[5]	Apple	Sistema de asistencia personal utilizado en dispositivos de la marca apple	-Está instalado predeterminadamente en los dispositivos -Interactúa con el usuario
Alexa[6]	Amazon	Asistente virtual de Amazon	-Plataforma de ayuda capaz de responder consultas y dudas utilizando un sistema de diálogos naturales para una persona.
Google Assistant[1]	Google	Asistente personal inteligente	-Se encuentra activo en el altavoz inalámbrico Google Home, los nuevos relojes Android Wear y las aplicaciones Allo y Duo.
Google Now[1]	Google	Asistente personal inteligente	-Está disponible dentro de la aplicación para móviles de Google Search
Prototipo propuesto	ESCOM	Prototipo de control por comandos de voz en SO	-Funciones disponibles sin conexión a internet -Personalización de comandos -Administración por perfil

## **2. Objetivo**

Desarrollar un prototipo para el control de una computadora mediante comandos de voz personalizables para usuarios con discapacidades motrices

## Objetivos particulares

- Investigar técnicas de procesamiento de lenguaje natural para asistentes virtuales.
- Determinar las funciones que podrá controlar el usuario mediante comandos de voz
- Permitir agregar y eliminar comandos de voz para su ejecución mediante el asistente
- Personalizar los comandos de voz para cubrir la necesidad de cada usuario que utilizará el asistente

## 3. Justificación

En muchos lugares se presentan casos en el que las personas no pueden usar una computadora de manera normal debido a que sufren de alguna discapacidad motriz que les impide interactuar con ellas, debido a esto es complicado que obtengan un empleo o que puedan utilizarlas por diversión, de acuerdo con la Clasificación Internacional del Funcionamiento, de la Discapacidad y de la Salud, presentada en 2001, que menciona que las personas con discapacidad “son aquellas que tienen una o más deficiencias físicas, mentales, intelectuales o sensoriales y que al interactuar con distintos ambientes del entorno social pueden impedir su participación plena y efectiva en igualdad de condiciones a las demás”. Al año 2020, una de las fechas más recientes de investigación, las personas que tienen algún tipo de discapacidad son más de 1000 millones, lo que representa 15% de la población total [4]. Para atacar ese problema se ofrece como solución el prototipo propuesto en este trabajo terminal. El enfoque principal es apoyar a las personas con alguna discapacidad, sin embargo se pretende abarcar la mayor cantidad de funcionalidades disponibles en una computadora, de esta manera se buscará generalizar el uso en la medida de lo posible.

Haciendo el uso de los comandos de voz los usuarios serán capaces de iniciar aplicaciones con solo decirlo buscando ahorrar tiempo en el proceso o simplemente para no hacer uso de las manos. Existen diversas aplicaciones que ayudan a los usuarios a optimizar esto y hacer consultas por medio de la voz, pero el inconveniente que se encuentra en la mayoría de ellos es que necesitan mantener una conexión a internet para cualquiera de sus funciones.

La aplicación que se desarrollará en este trabajo terminal tendrá funciones sin conexión a internet, también se buscará evitar los conflictos que surgen al intentar interpretar los ruidos que se presenten en alguna conversación y así no ejecutar aplicaciones incorrectas, tendrá la opción de personalizar los comandos dependiendo de las necesidades actuales y que puedan surgir a través del tiempo, para que así los usuarios puedan decidir y personalizar en todo momento la configuración del asistente virtual.

## 4. Productos o Resultados esperados

- 1.- Prototipo del sistema.
- 2.- Documentación Técnica del sistema.
- 3.- Manual del usuario.

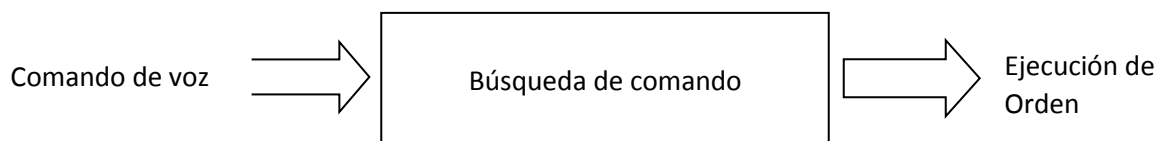


Figura 1. Sistema de petición de comandos.

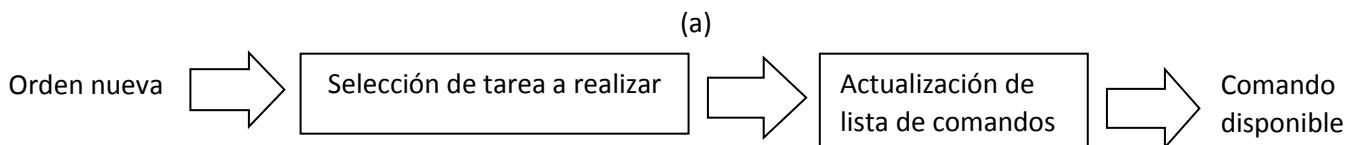


Figura 2. Módulo de personalización de comandos

(b)

Figura 1. Arquitectura del sistema

La figura 1 se explica de la siguiente manera:

- a. El comando de voz es escuchado por el asistente virtual que se encarga de hacer la búsqueda de dicho comando en la base de datos para así entonces proceder a ejecutarlo.

b. En cuanto a la personalización de los comandos, se introduce una orden nueva se selecciona la tarea que podrá realizar de acuerdo a las necesidades específicas, se introduce en la base de datos y se actualiza la lista para tener así el nuevo comando disponible para su uso en el momento requerido.

## 5. Metodología

El Proceso Unificado de Rational es un proceso de desarrollo de software que en conjunto con el Lenguaje Unificado de Modelado (UML), constituye la metodología estándar para el análisis de diseño e implementación de sistemas orientados a objetos lo cual se adapta al desarrollo del sistema en los módulos para actualización y petición de comandos ya que permite organizar las tareas en fases e iteraciones.

Se usarán cuatro fases.

- Inicio: Se analizarán los requerimientos del proyecto tomando en cuenta los alcances y su objetivo
- Elaboración: Especificación de casos de uso, elaboración de casos de uso, asignaciones de tareas dentro del equipo
- Construcción: Realización de diagramas de clase, diagramas de secuencia, implementación y realización de pruebas. En esta etapa se realizarán las iteraciones necesarias hasta obtener la nueva implementación del producto.
- Transición: Garantizar que el sistema esté listo para realizar tareas con los comandos de voz que los usuarios digan. Presentación del trabajo terminal

Se utilizará el modelo en espiral para el desarrollo del sistema, ya que el modelo en espiral está orientado a cambios y/o modificaciones, que nos brindará la facilidad de ir avanzando paulatinamente con el sistema. [3]

### Prototipo A

1. Contempla la definición de los comandos (el alcance del proyecto y sus requerimientos)

Módulos:

Comandos internos del sistema operativo  
Comandos del sistema de archivos  
Comandos de control de programas  
Comandos de control de navegador web

### Prototipo B

2. Contempla la personalización y configuración de nuevos comandos, para adaptarlos a las necesidades concretas del usuario añadiendo comandos personalizados en cada uno de sus módulos.

Configuración: comando, acción y respuesta

Personalización de: programas, carpetas, páginas web, tareas específicas

- Insertar comando
- Eliminar comando

La elección de esta metodología en comparación con otras revisadas y existentes, nos brinda la ventaja de que es una combinación del modelo lineal y el modelo iterativo, de forma importante suele utilizarse en proyectos que empiezan siendo pequeños e investigando los riesgos que se pueden tolerar para así hacerse grandes poco a poco, esto se acopla perfectamente al desarrollo del asistente virtual por voz ya que la metodología propuesta empezará con un pequeño asistente funcional que a medida que se programen los comandos y funcionalidades irá creciendo hasta completar el objetivo.

En cuanto a la ejecución el modelo en espiral tiene ciclos crecientes de 4 fases cada uno, la parte más importante de la elección de este modelo es que es un modelo evolutivo que mientras más avanzan los ciclos, también lo hace el tiempo de ejecución y así mismo el código fuente desarrollado, permitiendo así aumentar la complejidad del prototipo y agregar funciones gradualmente.

En la sección de anexos se incluye el cronograma de actividades para implementar cada uno de los prototipos propuestos.

## 6. Cronograma

Revisar el anexo A

## 7. Referencias

- [1] Propiedad de Google, (2020). (7 de Noviembre del 2020). Recuperado de: <https://assistant.google.com/>
- [2] Microsoft inc. (2020). ¿Qué es Cortana?. (7 de Noviembre del 2020). Recuperado de: <https://support.microsoft.com/es-mx/help/17214/windows-10-what-is>
- [3] Pressman, Roger S. (2010). Ingeniería de software un enfoque práctico. México. McGRAW-HILL
- [4] INEGI. (2020). Análisis de población. (7 de Noviembre del 2020). Recuperado de: <http://cuentame.inegi.org.mx/poblacion/discapacidad.aspx?tema=P>
- [5] Apple inc. (2020). (7 de Noviembre del 2020). Descripción de Siri. Recuperado de: <http://www.apple.com/mx/ios/siri/>
- [6] Alexa Internet, Inc. (1996-2020). (7 de Noviembre del 2020). Recuperado de: <http://www.alexa.com/>
- [7] Rouhiainen, L., 2018. *Inteligencia iva Artificial*. 1st ed. Barcelona: Alienta, pp. 17,23.
- [8] 2020. *Procesamiento Del Lenguaje Natural*. [Práctica] <http://www.it.uc3m.es/jvillena/irc/practicas/13-14/02.pdf>, Madrid.
- [9] Liege, J., Lostalé E. (2018). La era de la voz: Asistentes virtuales y voice marketing. Kanli(6). Recuperado de: [https://www.kanli.com/wp-content/uploads/2018/09/LA\\_ERA\\_DE\\_LA\\_VOZ\\_ASISTENTES\\_VIRTUALES\\_Y\\_VOICE\\_MARKETING.pdf](https://www.kanli.com/wp-content/uploads/2018/09/LA_ERA_DE_LA_VOZ_ASISTENTES_VIRTUALES_Y_VOICE_MARKETING.pdf)

## 8. Alumnos y Directores

*Domínguez Castro Israel Axel.-* Alumno de la carrera de Ing. en Sistemas Computacionales en ESCOM, Especialidad Sistemas, Boleta: 2013630027 Tel: 5573396523 Email: [idominguezc1200@alumno.ipn.mx](mailto:idominguezc1200@alumno.ipn.mx)

Firma: \_\_\_\_\_

*Dr. Jorge Cortés Galicia.-* Doctorado en Ciencias de la Computación CIC 2014, Maestría en Ciencias de la Computación CIC 2001, Ingeniería en Electrónica UAM 1998. Profesor del Depto. de Ingeniería en Sistemas Computacionales desde 2002. Áreas de interés: Sistemas distribuidos, sistemas operativos, lenguajes de modelado, teoría de juegos, Tel: 57296000 Ext. 52039, Email: [jcortesg@ipn.mx](mailto:jcortesg@ipn.mx)

Firma: \_\_\_\_\_

*Pescador Rojas Miriam.-* Dra. en Ciencias en Computación por el CINVESTAV-IPN en 2019, M. en C. en Computación por el CINVESTAV-IPN en 2010, Ing. en Sistemas Computacionales por la ESCOM-IPN en 2008, Profesora de carrera en ESCOM-IPN desde 2010 a la fecha, áreas de interés Inteligencia Artificial, Cómputo Evolutivo, Aprendizaje Máquina, Ext. 52022, email: [mpescadorr@ipn.mx](mailto:mpescadorr@ipn.mx)

Firma: \_\_\_\_\_

CARÁCTER: Confidencial  
FUNDAMENTO LEGAL: Artículo 11 Fracc. V y Artículos 108, 113 y 117 de la Ley Federal de Transparencia y Acceso a la Información Pública. PARTES CONFIDENCIALES: Número de boleta y teléfono.

## ANEXO A

CRONOGRAMA Nombre del alumno(a): Domínguez Castro Israel Axel

TT No.:

**Título del TT: Prototipo de control por comandos de voz para una computadora personal**

[illegible]

