



**Universidad de San Andrés**

**Escuela de Negocios**

**Maestría en Gestión de Servicios Tecnológicos y de Telecomunicaciones**

***Adopción de asistentes virtuales inteligentes a través de  
la comercialización de altavoces inteligentes en  
Argentina***

**Autor: Pablo Barbieri**

**DNI: 24.353.629**

**Director de Tesis: Dr. Alejandro Prince**

**Buenos Aires, 29 de Mayo de 2020**



**Universidad de San Andrés**

**Escuela de Negocios**

**Maestría en Gestión de Servicios Tecnológicos y de Telecomunicaciones**

**Adopción de asistentes virtuales inteligentes a  
través de la comercialización de altavoces  
inteligentes en Argentina.**

**Autor: Pablo Barbieri**

**DNI: 24.353.629**

**Director de Tesis: Dr. Alejandro Prince**

**Buenos Aires, 29 de Mayo de 2020**

## Agradecimientos

Como autor del presente trabajo de estudio quiero agradecer a mi familia, especialmente a mi esposa María Clara y mi hijo Pedro, por la ayuda y el aliento constante y sobretodo por la aceptación de mi ausencia en numerosos encuentros, cenas, programas y salidas de fin de semana.

A mi madre, hermanos y familia política por su incansable soporte durante el proceso de estudio.

Al director de la maestría en Gestión de Servicios Tecnológicos y Telecomunicaciones, Ing. Enrique Hofman y su equipo de trabajo, que lograron ayudarme a gestionar la finalización de la misma mientras al mismo tiempo me encontraba estableciéndome en otro país con mi familia.

Al Dr. Alejandro Prince, que con su paciencia y sapiencia como director de tesis ha logrado guiarme y aconsejarme a la distancia en la producción de conocimiento para el desarrollo del presente estudio.

A mis compañeros de maestría que han enriquecido los debates sobre las diferentes materias y al grupo núcleo de estudio que generosamente aceptó mi participación en los equipos para los trabajos prácticos tanto presenciales y como a la distancia.

Finalmente a mi padre, que ha sabido sembrar los mensajes correctos que hoy rigen el pulso de vida y al cual extraño todos los días desde hace 20 años.

Universidad de  
San Andrés

## Resumen Ejecutivo

El avance exponencial en la reducción de costos de los desarrollos tecnológicos y su correspondiente utilización ha traído aparejado la disrupción de muchas industrias y en consecuencia el ingreso y formación de nuevas compañías nativas digitales que, como denominador común llevan en su ADN el desafío de interpelar los paradigmas actuales y establecer nuevos *Statu quo*, los cuales nos obliga, cada vez con más frecuencia, como consumidores y usuarios a un cambio constante de adaptación hacia nuevas realidades.

Una pieza de esta realidad es como decidimos interactuar con nuestro ecosistema digital, el cual hace algunos años solo era factible acceder a través de una interfaz grafica mediante los dispositivos establecidos (*Smartphone, Tablet, Laptop, etc*).

En la actualidad se encuentra disponible la posibilidad de establecer la comunicación con nuestros servicios digitales en internet a través del habla y el lenguaje. Los asistentes virtuales inteligentes incorporados en altavoces se encuentran transformando nuevamente la realidad que nos rodea, los mismos no solo nos permiten acceder a servicios digitales sin la necesidad de utilizar dispositivos ya conocidos, estos asistentes virtuales tienen como desafío incorporarse a nuestras vidas de manera omnipresente y asistirnos en varias dimensiones como el monitoreo de la salud, la propuesta de entretenimiento, alertas de seguridad, asistencia en el trabajo y en la interoperabilidad del hogar inteligente.

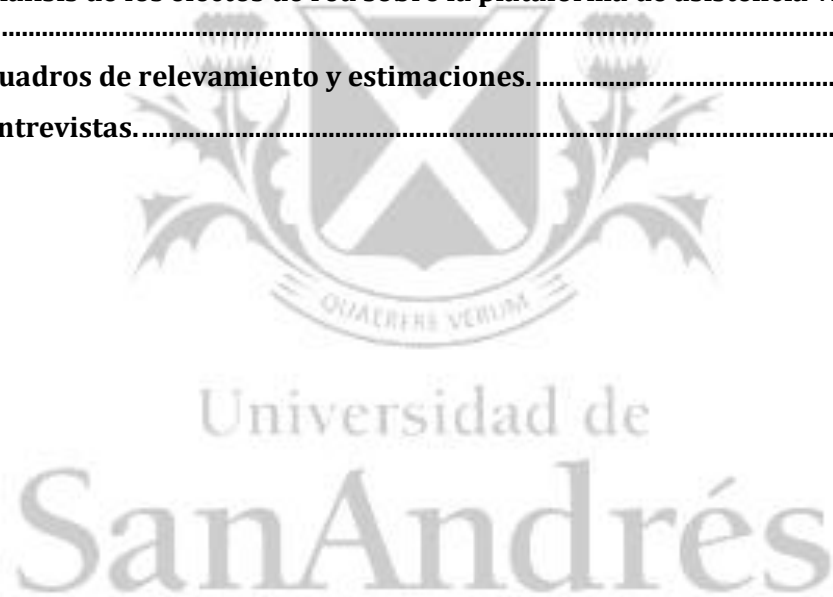
En el presente trabajo de estudio analizaremos el avance y adopción de esta tecnología en el mundo, el impacto y las implicancias derivadas de su uso, describiremos sus principales características, presentaremos las barreras que se deberán sortear para su aceptación y los cambios necesarios para que suceda la adopción en el mercado argentino.

**Palabras claves:** *Altavoz inteligente, Inteligencia artificial, Asistente virtual inteligente, Adopción de innovaciones y tecnología, Mercado argentino.*

# Índice

<b>Tabla de Figuras .....</b>	<b>5</b>
<b>1.Capítulo I – Introducción .....</b>	<b>6</b>
<b>1.2 - Objetivo Principal.....</b>	<b>9</b>
<b>1.3 - Objetivos Secundarios .....</b>	<b>9</b>
<b>1.4 - Preguntas de Investigación.....</b>	<b>10</b>
<b>2.Capítulo II – Presentación de la Metodología.....</b>	<b>11</b>
<b>2.1 – Paradigma y Tipo de Investigación .....</b>	<b>11</b>
<b>2.1.1 - Tipo de Investigación.....</b>	<b>11</b>
<b>2.2 – Instrumentos .....</b>	<b>12</b>
<b>3.Capítulo III – Marco Teórico .....</b>	<b>14</b>
<b>3.1 – Difusión y Adopción de Innovaciones .....</b>	<b>14</b>
<b>3.1.1 – Ciclo de Vida de la Adopción de Tecnología .....</b>	<b>23</b>
<b>3.1.2 – Modelo de Aceptación de Tecnología .....</b>	<b>27</b>
<b>4.Capítulo IV - Definición de Altavoz Inteligente y descripción de sus componentes. ....</b>	<b>35</b>
<b>4.1 – Definición de Altavoz Inteligente.....</b>	<b>35</b>
<b>4.2 - Plataforma de Inteligencia Artificial .....</b>	<b>41</b>
<b>4.2.1 – Evolución de Inteligencia Artificial .....</b>	<b>42</b>
<b>4.2.2 - Comunicación y Procesamiento del Lenguaje Natural .....</b>	<b>46</b>
<b>4.2.3 - Reconocimiento de la Voz o el Habla.....</b>	<b>48</b>
<b>5.Capítulo V – Historia y actualidad de los asistentes virtuales inteligentes. ....</b>	<b>52</b>
<b>5.1 – Historia y evolución de los asistentes virtuales inteligentes. ....</b>	<b>52</b>
<b>5.2 – Análisis de los principales asistentes virtuales inteligentes disponibles en el mercado. ....</b>	<b>57</b>
<b>6.Capítulo VI – Análisis y tendencia del mercado global y de Estado Unidos. ....</b>	<b>73</b>
<b>6.1 – Introducción. ....</b>	<b>73</b>
<b>6.2 – Mercado global. ....</b>	<b>74</b>
<b>6.3 – Mercado de Estados Unidos. ....</b>	<b>78</b>
<b>6.3.1 – Comportamiento y satisfacción de usuario.....</b>	<b>87</b>
<b>7.Capítulo VII – Análisis de adopción del mercado argentino y proyección de estimaciones. ....</b>	<b>92</b>
<b>7.1 – Introducción. ....</b>	<b>92</b>

<b>7.2 - Evolución y adopción de computadoras tablet en Estados Unidos y Argentina. ....</b>	<b>93</b>
<b>7.3 - Evolución y adopción de televisores inteligentes en Estados Unidos y Argentina. ....</b>	<b>97</b>
<b>7.4 - Proyección y estimaciones de adopción de altavoces inteligentes para el mercado argentino. ....</b>	<b>102</b>
<b>8. Capítulo VIII - Conclusiones. ....</b>	<b>106</b>
<b>8.1 - Introducción. ....</b>	<b>106</b>
<b>8.2 - Conclusiones. ....</b>	<b>107</b>
<b>8.3 - Preguntas de investigación. ....</b>	<b>110</b>
<b>8.4 - Reflexiones. ....</b>	<b>113</b>
<b>Bibliografía. ....</b>	<b>116</b>
<b>Anexo I - Análisis de competitividad de la industria. ....</b>	<b>126</b>
<b>Anexo II - Análisis de los efectos de red sobre la plataforma de asistencia virtual inteligente. ....</b>	<b>129</b>
<b>Anexo III - Cuadros de relevamiento y estimaciones. ....</b>	<b>131</b>
<b>Anexo IV - Entrevistas. ....</b>	<b>133</b>



## Tabla de Figuras

- Figura 1. Curva de adopción y difusión de innovaciones
- Figura 2. Curva de adopción y teoría del abismo.
- Figura 3. Principales variables constructoras identificadas para el estudio de adopción de altavoces inteligentes
- Figura 4. Diagrama arquitectura IVA
- Figura 5. Tabla de principales habilidades de un IVA
- Figura 6. Evolución de habilidades desarrolladas para Alexa
- Figura 7. Funcionalidades ejecutas en los asistentes virtuales durante 2018 por plataforma
- Figura 8. Diagrama de arquitectura referencial para el reconocimiento de voz
- Figura 9. Arquitectura del ecosistema de interacción con la plataforma Alexa
- Figura 10. Diagrama de arquitectura de integración de Google Assistant
- Figura 11. Diagrama de arquitectura de integración para aplicaciones iOS con SiriKit
- Figura 12. Cuadro comparativo de características de productos comercializados en la categoría de altavoz inteligente
- Figura 13. Resultado de asertividad de las plataformas de asistente virtual
- Figura 14. Cantidad de suscripciones de conexiones fijas a internet en el mundo
- Figura 15. Pronóstico de base instalada de altavoces inteligentes para 2019 vs base instalada en el mundo en el 2018 y abierto por los países de mayor crecimiento esperado
- Figura 16. Evolución global de adopción de altavoces inteligentes según curva propuesta por E. Rogers
- Figura 17. Número de suscripciones de banda ancha fija en los Estados Unidos de 2000 a 2018 (en millones)
- Figura 18. Porcentaje de hogares con uso de internet en los Estados Unidos de 1997 a 2018
- Figura 19. Envíos de unidades de altavoces inteligentes en los Estados Unidos de 2016 a 2020 (en millones)
- Figura 20. Porcentaje de frecuencia de utilización de altavoces inteligentes. Fuente: NPR & Edison Research
- Figura 21. Base de altavoces inteligentes en EE.UU. enero 2018 vs enero 2019
- Figura 22. Frecuencia de utilización de altavoces inteligentes en EE.UU. enero 2018 vs enero 2019
- Figura 23. Evolución Tasa de penetración de altavoces inteligentes sobre los hogares conectados a internet de EE. UU. De 2017 a 2019
- Figura 24. Proyección de tasa de penetración de los hogares con altavoces inteligentes en los Estados Unidos desde 2014 - 2025
- Figura 25. Cualidades valoradas por usuarios de altavoces inteligentes en EE.UU.
- Figura 26. Frecuencia de utilización de casos en Altavoces Inteligentes
- Figura 27. Adopción en hogares de dispositivos tecnológicos
- Figura 28. Base instalada de dispositivos sobre penetración en hogares 2019
- Figura 29. Adopción de tablets en Argentina vs EE.UU. 2010-2019
- Figura 30. Evolución cantidad de usuarios de TV conectados a internet 2015-2023
- Figura 31. Evolución de ventas SMP y TI 2015-2019
- Figura 32. Adopción de TI en Argentina vs EE.UU. 2009-2018
- Figura 33. Cuadro de estimación de hogares por clase social en Argentina 2020
- Figura 34. Estimación de escenarios para la adopción de altavoces inteligentes en Argentina vs EE.UU. 2015-2030

# 1. Capítulo I – Introducción

## 1.1 – Planteo de la Problemática

La ciencia ficción, en sus diferentes géneros, puede ser uno de los principales instrumentos donde el ser humano puede volcar y conceptualizar parte de su visión acerca del futuro y las posibilidades que la tecnología debería y podría brindar si todo siguiese un curso natural de desarrollo.

Gran parte de los escritores y cineastas de este género han soñado, escrito y diseñado guiones acerca de la posibilidad de poder entablar una comunicación de un ser humano con una máquina o computadora. Hoy esta posibilidad de poder generar y mantener una comunicación es factible gracias a la utilización de los altavoces inteligente (Smart speaker en inglés), y asistentes virtuales inteligentes.

El presente estudio se concentra sobre la investigación de la adopción de la tecnología de asistentes virtuales inteligentes (IVA por sus siglas en inglés) los cuales son posibles de utilizar gracias a los Altavoces Inteligentes o comúnmente denominados Smart Speakers. Esta relativamente nueva tecnología está tomando a nuestro entender un grado de aceptación y adopción vertiginoso y desde nuestra visión contiene ciertas aristas, las cuales pueden transformar el paradigma digital tal cual lo conocemos actualmente.

El dispositivo posibilita la ubicuidad de asistentes digitales en nuestra cotidianeidad con la facilidad de poder administrarlos, educarlos y hacer que ejecuten nuestras órdenes con el sólo uso de nuestra voz, sin necesidad de tocar o realizar ninguna otra actividad en otro dispositivo.

Entre algunas de las tareas y funciones que podemos delegar y ejecutar en los dispositivos con asistentes virtuales inteligentes se encuentran comprar un producto, reproducir una lista de canciones, enviar un mensaje de texto a nuestros contactos, activar o apagar artefactos inteligentes conectados a la red de nuestro hogar, bajar o subir la temperatura del aire acondicionado, escuchar noticias de los diarios, pedir información sobre algún producto o servicio, hacer alguna pregunta o búsqueda sobre internet y realizar compras (Hoy, 2018).

Algunos dispositivos con la funcionalidad de ir ejerciendo un aprendizaje continuo han ido desarrollado ciertas habilidades y funciones un poco más complejas como asistir personas con diversos grados de discapacidad o personas mayores que necesitan de un cuidado especial o de recordatorios específicos en sus actividades diarias, ayuda en el mantenimiento y reparación de



ciertos electrodomésticos inteligentes, contar cuentos e historias customizadas especialmente para los niños y hasta planificar actividades para poder construir una relación a largo plazo con los usuarios.

Así mismo estamos frente a la presencia de la primera generación de individuos en crecer con inteligencia artificial omnipresente en un mundo donde prácticamente cualquier pregunta puede ser respondida, la mayoría de los productos comprados en línea y muchas tareas resueltas solo con la activación de un comando a través de nuestra voz. Lo anterior plantea profundas preguntas sobre cómo los niños interactúan con la tecnología, con otras personas, y cómo esto podría moldear sus interacciones y su desarrollo (Murphy, 2018).

Se espera que el mercado global de dispositivos de altavoces inteligentes para el hogar con funciones integradas de IVA alcance una facturación cercana a los US\$ 2 mil millones de dólares para el año 2020, casi un aumento del 500 % con respecto a 2015 (Gartner, 2016). También se pronostica que el número de usuarios consumidores individuales de asistentes virtuales digitales (VDA por sus siglas en inglés) aumentará de 390 millones a 1.800 millones en todo el mundo entre los años 2015 y 2021. Mientras tanto, se prevé que el número de usuarios empresariales crecerá de 155 millones a 843 millones (Tractica, 2016).

Estos mayordomos digitales que podemos comprar e incorporar a nuestra vida y hogares están siendo desarrollados por empresas líderes de distintas plataformas e industrias como son Google, Amazon, Apple y Facebook, entre otros. Si bien el inicio o punto de partida de estos dispositivos no se encuentra muy lejano en el tiempo y data del 2014, el desarrollo y evolución de los mismos produjo algunas idas y vueltas con las estrategias de las principales empresas consideradas “Big Techs” en la industria, las cuales detallaremos más adelante. Sólo a modo de ejemplo podemos mencionar el caso de Microsoft que decidió discontinuar su tecnología Cortana para los IVA y concretó una alianza con Amazon, para incorporar el desarrollo de Amazon, denominado Alexa, a sus productos o el caso de Samsung, que está anunciando el lanzamiento inminente de Galaxy Home con su asistente virtual llamado Bixby en este nicho de la industria. Otro caso interesante es el de Facebook que había anunciado la salida al mercado de su plataforma inteligente M, pero por el momento su dispositivo llamado Portal tiene una alianza con Amazon para la utilización de Alexa como asistente virtual inteligente.

Así mismo es necesario conocer la visión de algunos de los líderes que marcan el rumbo de estas empresas. En una entrevista realizada por la revista TheEconomicTimes.com al CEO de

Alphabet, Sundar Pichai, explica la visión de la empresa en la cual se pronostica en un futuro la desaparición de los dispositivos, ya sea computadoras o smartphones, tal cual los conocemos hoy. En su lugar tendremos acceso a asistentes virtuales que nos podrán ayudar en nuestro día a día. La visión redactada también en una carta a los empleados de la empresa vislumbra que los usuarios de los productos de Google podrán moverse libremente y utilizar los servicios de Google de forma natural, el asistente de Google podrá entender el contexto y las necesidades del usuario respetando la privacidad y la protección de sus datos personales. Otro aspecto interesante de la visión de Google para el futuro es que espera que las personas tiendan a pensar menos y delegar las decisiones en asistentes virtuales cuando se trate de decidir sobre aspectos poco importantes o cotidianos de la vida diaria.

Existe un componente distintivo en la utilización de esta tecnología que es la interacción que se produce en la comunicación entre el dispositivo, su software de inteligencia artificial y el usuario. Esta nueva forma de interactuar de alguna manera genera e incorpora un componente emocional en la naturaleza de la comunicación que establecemos con el dispositivo, lo cual podría conducir a que como usuarios, de manera peligrosa y con algún grado de riesgo en la delegación de decisiones, hagamos depósito de nuestra confianza en la toma de decisiones a los algoritmos que se encuentran embebidos en estos dispositivos (Stucke and Ezrahi 2016, 2018).

En la década de 1980, los debates sobre inteligencia artificial se centraron en la cuestión de si las máquinas realmente podrían ser inteligentes. Estos debates fueron sobre los objetos mismos, sobre lo que podían o no podían hacer. Cuando se nos pide que cuidemos un objeto, cuando el objeto cuidado prospera y nos ofrece su atención y preocupación, experimentamos ese objeto como inteligente, pero algo más importante, podemos sentir un nuevo nivel de conexión emocional con el mismo. Las nuevas preguntas no son sobre si las máquinas que cuentan con la capacidades relacionales realmente tienen inteligencia o emociones, sino sobre lo que evocan en sus usuarios (Turkle, 1984).

Dado que nos encontramos en una incipiente fase de adopción de la tecnología y que también se pronostica que el rango de funciones que hoy un dispositivo con asistente virtual puede ejecutar se amplíe enormemente, se genera en consecuencia una turbulenta fase de inversión y expansión en el mercado global y un cambio dinámico en las estrategias de los principales jugadores que se encuentran entrando en la competencia en un corto plazo. Sumado a esto se espera también la internacionalización de ciertos jugadores locales de China, los cuales cuentan con gran

cantidad de recursos, no sólo económicos sino también en conocimiento, para competir en el mercado global, como Xiaomi, Baidu o Alibaba (Budzinski, Noskova, & xijie, 2019).

Otro aspecto interesante que esta tecnología intenta establecer es el rol de concentrador y moderador para acelerar la adopción de dispositivos inteligentes en hogares. Electrodomésticos, sensores, equipos de climatización, cámaras de seguridad, artefactos de iluminación son algunos de la nueva era de productos digitales inteligentes que, conectados a través del internet de las cosas (IOT por sus siglas en inglés), permiten ser operados por nuestra voz a través del IVA de nuestro parlante inteligente.

Creemos importante y fundamental describir el estado del arte de esta industria y su desarrollo de adopción, particularmente en el mercado de Estados Unidos, ya que entendemos afrontará y atraerá en el corto plazo a un alto grado de inversiones por parte de distintos jugadores que en la actualidad son líderes absolutos en sus industrias.

## 1.2 - Objetivo Principal

Analizar y describir el desarrollo tecnológico y comercial de los asistentes virtuales inteligentes en el mundo y evaluar la potencial difusión y adopción de estos productos en los hogares del mercado argentino, sus barreras de ingreso, riesgos y oportunidades.

## 1.3 - Objetivos Secundarios

- i) Describir que es un asistente virtual inteligente, su historia, surgimiento y evolución.
- ii) Comparar los diferentes productos disponibles, estrategias de producto, ventajas y desventajas de cada uno.
- iii) Examinar su evolución y adopción en los mercados globales.
- iv) Presentar y establecer pronósticos y opiniones especializadas sobre el futuro de adopción de la tecnología de asistentes virtuales inteligentes en hogares.

v) Describir las propiedades, particularidades y límites del potencial mercado argentino y presentar las características demográficas que se pueden atribuir a los posibles compradores de asistentes virtuales inteligentes.

vi) Presentar las distintas evoluciones de difusión y adopción que han seguido productos similares como los televisores inteligentes y tabletas en el mercado argentino.

vii) Identificar y reflexionar sobre las implicancias y los factores determinantes para la adopción de asistentes virtuales inteligentes.

#### 1.4 - Preguntas de Investigación

i) ¿Las principales barreras para lograr la masividad de los asistentes virtuales inteligentes en el mercado argentino están dadas por el precio en moneda extranjera del producto, la ausencia de entrenamiento de la plataforma sobre el modo y uso del lenguaje coloquial argentino y la falta de complementariedad actual que existe con otros dispositivos para interactuar en el rol de concentrador para la ejecución de tareas en los hogares inteligentes?

ii) ¿Cuáles son los factores claves y los cambios necesarios que deben ocurrir para que el mercado de hogares argentino adopte masivamente la tecnología de asistentes virtuales inteligentes?

## 2. Capítulo II – Presentación de la Metodología.

En el presente capítulo describiremos el paradigma de investigación empleado para producir el trabajo de investigación. En consecuencia, detallaremos la metodología de investigación utilizada, los instrumentos que permitieron profundizar la recolección de información a partir de distintas fuentes, el análisis propio y las inferencias de los mismos, como así también la definición de las distintas variables e indicadores sobre los cuales construimos el abordaje del presente estudio de investigación.

### 2.1 – Paradigma y Tipo de Investigación

El proceso de construcción de conocimiento encuentra diferentes puntos de interpretación y respuestas según los paradigmas de investigación a emplear. Creswell J. entiende que la principal distancia entre el paradigma cualitativo y cuantitativo se establece en el enfoque de investigación. El paradigma cualitativo resulta de utilidad para examinar y realizar conocimiento acerca de significados, tendencias y conceptos, en tanto que la investigación cuantitativa es empleada para establecer los elementos de prueba de teorías al investigar las relaciones y las causas de las mismas entre las variables que se necesitan medir (Creswell, 2009).

El presente trabajo de investigación se efectuó bajo un paradigma cualitativo, el cual encuentra sus pilares en el abordaje y análisis de trabajos existentes de investigación y difusión académica y científica, en publicaciones e informes de consultoras especializadas y reportes elaborados por agencias de investigación de mercado a nivel global.

#### 2.1.1 – Tipo de Investigación

Según lo establece Yin R. es posible definir tres tipos de investigaciones para llevar adelante el estudio que pretendemos construir, el mismo puede ser de carácter exploratorio, descriptivo o explicativo (Yin, 1994).

Una investigación descriptiva debe ser realizada cuando el área de investigación tiene una alta complejidad (Zikmund, 2000). El abordaje del trabajo de estudio se centrará en la descripción de los hechos observados para interpretarlos, comprenderlos y analizarlos dentro del contexto global en el que se producen con el fin de explicar los fenómenos asociados a los objetivos

presentados. Best W. aporta que la investigación de tipo descriptiva está relacionada a interpretar y comprender variables explicativas o combinaciones acerca de los vínculos existentes entre las mismas, formatos de prácticas que predominan en el tiempo, construcción de acuerdos y opiniones, interpretación acerca de las distintas miradas que interpelan un objeto y de los procesos de estudio que se establecen sobre la marcha de tendencias que se desarrollan (Best, 1988).

Esta tesis se encuadra dentro del paradigma de investigación cualitativo con foco en el tipo de investigación descriptivo ya que profundizará y hará hincapié en la descripción exhaustiva de las variables y hechos que han permitido la adopción y difusión de asistentes virtuales inteligentes en el mercado global a través de la comercialización de altavoces inteligentes y los desafíos, riesgos y barreras de adopción que esta tecnología puede sufrir para su incorporación y adopción en el mercado argentino.

## 2.2 – Instrumentos

Denzin N. sostiene que, con el fin de disminuir el sesgo propio de los estudios realizados bajo un paradigma cualitativo, el cual está cargado de cierto grado de subjetividad, se debe utilizar como instrumento para la recolección de datos una estrategia de investigación a través de la cual el objeto de estudio sea abordado desde diferentes miradas y perspectivas de contraste (Denzin, 1989).

La triangulación y confrontación de instrumentos y fuentes permitirá llevar a cabo las distintas interpelaciones de los datos obtenidos en el estudio de campo. Es por lo que, que este proyecto de investigación se desarrolló, con un proceso exhaustivo de compilación y análisis de información existente, utilizando como fuentes primarias y secundarias a autores académicos, consultoras expertas y publicaciones de empresas del sector. Debido a la orientación descriptiva y de alto grado de actualidad en la temática abordada, la búsqueda de información no se limitará solamente a fuentes académicas. Utilizaremos bases de datos relevantes y motores de búsqueda de consideración científica, libros, publicaciones de autores y estudiosos del tema, así como estudios realizados por consultoras de relevancia internacional.

En segundo término para efectuar una triangulación acerca de la información recolectada se realizaron entrevistas semiestructuradas con referentes del sector de empresas de retail tecnológico dentro de Argentina, expertos sobre temas relacionados con la industria de las Tecnologías de la Información y de la Comunicación (TICs) y profesionales académicos que pudieran aportar una mirada acerca de la difusión y adopción de este tipo de tecnologías en el mercado argentino.



### 3. Capítulo III – Marco Teórico.

La investigación de adopción y aceptación de tecnología proporciona modelos y teorías que intentan explicar y predecir por qué ciertos consumidores aceptan determinadas tecnologías, mientras que otros no. Debido a la naturaleza basada en la intención de estos modelos, es posible investigar los factores de aceptación relevantes y así evaluar la experiencia de los consumidores en la utilización de los productos y servicios proporcionados por las innovaciones tecnológicas.

#### 3.1 – Difusión y Adopción de Innovaciones

Rogers E. en su texto “Difusión de Innovaciones” (Diffusion of Innovations) desarrolla un modelo de difusión y adopción de innovaciones el cual contiene una teoría que aglutina distintas investigaciones sobre la temática de estudio, la mayoría con conclusiones similares y que comienzan a publicarse durante las décadas de 1940 y 1950.

Rogers define la difusión de la innovación como un proceso en el que una innovación se comunica a través de ciertos canales a lo largo del tiempo entre los miembros de un sistema social. La innovación contempla un grado de novedad sobre la idea que debe ser transmitida en el contenido del mensaje. Esta característica especial expone un nivel de incertidumbre o falta de información sobre el contenido que los sujetos receptores reciben en el mensaje (Rogers, 2003).

Las innovaciones tecnológicas están compuestas principalmente por componentes de hardware y software. La difusión de innovaciones que se encuentran compuestas mayoritariamente por software, a lo largo de los años, ha generado inconvenientes a la hora de las investigaciones acerca de su adopción ya que las mismas a diferencias de las compuestas en su mayoría por hardware presentan problemas dado su grado relativamente bajo de observabilidad y seguimiento, por lo tanto, presentan una tasa de adopción más lenta en el tiempo (Rogers, 2003).

Cuando nuevas ideas que contemplan innovaciones tecnológicas se inventan, difunden, y adoptan o rechazan, las mismas conllevan ciertas consecuencias, es aquí donde se producen los cambios sociales.

El paradigma conceptual de difusión de innovación desarrollado por Rogers permite establecer un camino para el estudio acerca de los cambios sociales que una innovación en particular puede efectuar sobre una sociedad. La investigación acerca de la difusión ofrece un



medio particularmente útil para comprender el cambio producido debido a que las innovaciones son consideradas un tipo de mensaje de comunicación cuyos efectos son relativamente fáciles de aislar y medir (Rogers, 2003).

El proceso de decisión de innovación es el proceso a través del cual un individuo (u otra unidad de toma de decisiones) pasa por varias etapas para incorporar la innovación. Desde el primer paso que consiste en obtener el conocimiento de una innovación, pasando luego a construir una actitud hacia la innovación, posteriormente tomando una decisión acerca de adoptar o rechazar la innovación, luego comienza la etapa de implementación de la nueva idea y finalmente cerrando el proceso con la confirmación de esta decisión.

El proceso definido por Rogers consta de cinco etapas: (1) conocimiento, cuando el individuo está expuesto a la existencia de la innovación y adquiere una comprensión de cómo funciona; (2) persuasión, cuando el individuo forma una actitud favorable o desfavorable hacia la innovación; (3) decisión, cuando el individuo se involucra en actividades que conducen a la opción de adoptar o rechazar la innovación; (4) implementación, cuando el individuo pone en práctica una innovación; y (5) confirmación, cuando el individuo busca refuerzo para una decisión de innovación ya tomada, pero puede revertir la decisión si se expone a mensajes contradictorios al respecto (Rogers, 2003).

Rogers define cinco atributos que debe contemplar una innovación y que a su vez tienen una correlación al momento de evaluar la tasa de adopción de una innovación en el tiempo. Los atributos enumerados por Rogers que aceleran la tasa de adopción son, en primer término, la **ventaja relativa** sobre la cual una innovación es percibida en base a su idea predecesora. En segundo lugar, se encuentra la **compatibilidad** de una innovación que consiste en la percepción que una innovación tiene en base a la existencia de valores previos, las experiencias pasadas y las necesidades actuales de los potenciales usuarios. En tercer lugar, encontramos la **complejidad** con que una innovación es percibida para ser utilizada o comprendida como un atributo cuya relación es negativa a la tasa de adopción. El cuarto atributo enumerado por Rogers es el grado de **capacidad de prueba** que tiene una innovación para ser adquirida. Finalmente el último atributo establecido en la teoría de difusión de innovación es la **observabilidad**, el mismo refiere a como los resultados de una innovación expuestos hacia otros sujetos impactan positivamente en la capacidad de adopción de la misma (Rogers, 2003).

El comportamiento innovador interpretado como una variable relativa y medida en términos del grado en que un individuo u otra unidad alternativa de adopción tiene en relación a cuanto tiempo antes adopta una innovación en comparación a sus semejantes pertenecientes al mismo sistema social, es la utilizada por Rogers para determinar las distintas categorías y perfiles de adoptantes dada una determinada innovación (Rogers, 2003).

Las cinco categorías resultantes de individuos son: innovadores, adoptantes tempranos, mayoría temprana, mayoría tardía y rezagados. A continuación, describimos ciertas características acerca de cada uno de los perfiles identificados por Rogers:

- **Innovadores:** La aventura es casi una obsesión con los innovadores. El interés de los innovadores acerca de las nuevas ideas los aleja del círculo local de redes de pares y entran en relaciones sociales más cosmopolitas. El control de recursos financieros sustanciales es útil para absorber las posibles pérdidas de una innovación no rentable. También se necesita la capacidad de comprender y aplicar conocimientos técnicos complejos. El innovador debe ser capaz de hacer frente a un alto grado de incertidumbre acerca de una innovación en el momento en que adopta. El valor más destacado del innovador es la aventura. El innovador juega un papel importante en el proceso de difusión: el de lanzar la nueva idea en el sistema importando la innovación desde fuera de los límites del sistema.
- **Adoptantes tempranos:** Los primeros en adoptar son una parte más integrada del sistema social local que los innovadores. Esta categoría de adoptantes, más que ninguna otra, tiene el más alto grado de liderazgo de opinión en la mayoría de los sistemas. El adoptante temprano es respetado por sus pares y es la encarnación del uso exitoso y discreto de nuevas ideas. El adoptante temprano disminuye la incertidumbre acerca de una nueva idea al adoptarla y luego transmitir una evaluación subjetiva de la innovación a pares cercanos a través de redes interpersonales.
- **Mayoría temprana:** La mayoría temprana adopta nuevas ideas justo antes del miembro promedio de un sistema. Interactúan frecuentemente con sus pares, pero rara vez ocupan puestos de liderazgo de opinión en un sistema social. Proporcionan interconexión en las redes interpersonales de su sistema social.

- **Mayoría Tardía:** La adopción puede ser tanto una necesidad económica para la mayoría tardía como el resultado del aumento de las presiones de los pares. Las innovaciones se abordan con un aire escéptico y cauteloso. La presión de los compañeros es necesaria para motivar la adopción. Con recursos relativamente escasos necesitan que la mayor parte de la incertidumbre sobre una nueva idea se elimine antes de incorporarla.
- **Rezagados:** Se encuentran casi aislados en las redes sociales de su sistema. El punto de referencia para el rezagado es el pasado. Los rezagados tienden a sospechar de las innovaciones y de los agentes de cambio. Su proceso de decisión de innovación es relativamente largo. Sus recursos son limitados y deben estar seguros de que una nueva idea no fallará antes de que puedan adoptar.

Al distribuir esta clasificación en una curva de frecuencia de distribución normal encontramos que el área situada a la izquierda del tiempo medio de adopción (de una innovación) menos dos desviaciones estándar incluye el primer 2.5 por ciento de los individuos en un sistema para adoptar una innovación, los innovadores. El siguiente 13.5 por ciento para adoptar la nueva idea se incluye en el área entre la media menos una desviación estándar y la media menos dos desviaciones estándar; están etiquetados como los adoptantes tempranos. El siguiente 34 por ciento de los adoptantes, llamado mayoría temprana, están incluidos en el área entre la fecha media de adopción y la media menos una desviación estándar. Entre la media y una desviación estándar a la derecha de la media se encuentra el siguiente 34 por ciento en adoptar la nueva idea, la mayoría tardía. El último 16 por ciento en adoptar se los denomina rezagados.

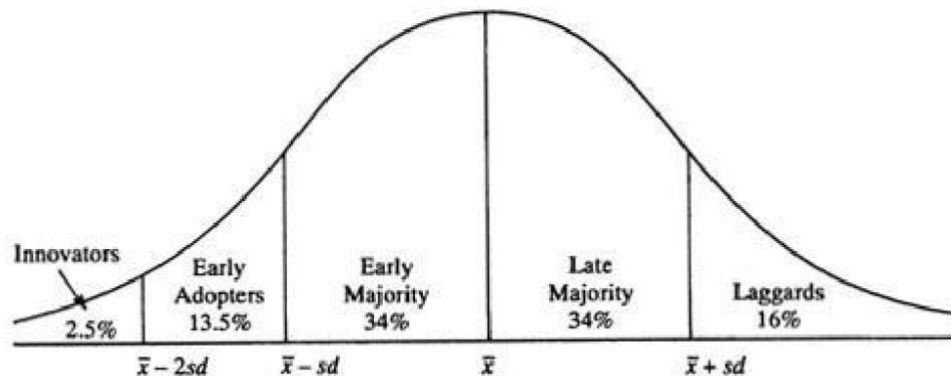


Figura 1. Curva de adopción y difusión de innovaciones. Fuente: E. Rogers

Desde una clara perspectiva acerca de los procesos que establecen la comunicación Rogers entiende que las redes de comunicación interpersonales y los programas de difusión son especialmente críticos al momento de comprender la tasa de adopción de una innovación. En el proceso de difusión es fundamental reconocer y diseñar un plan que involucre las opiniones de liderazgo de un determinado grupo de adoptantes, entender como es esa estructura de comunicación dentro de esas redes y como fluye la información dentro del sistema social.

Rogers estableció críticas a la investigación de difusión en cuatro categorías: sesgo a favor de la innovación, sesgo de la culpa del individuo, el problema del recuerdo y consecuencias de igualdad económica. El **sesgo pro-innovación** es la implicación en la investigación de difusión de que una innovación debe ser difundida y adoptada rápidamente por todos los miembros de un sistema social y que la innovación no debe reinventarse o rechazarse. El **sesgo de la culpa del individuo** es la tendencia dentro de la investigación de la difusión de volcar la promoción de la innovación sobre los agentes de cambio en lugar de los individuos que son los potenciales adoptantes. El **problema del recuerdo** está basado en uno de los pilares del estudio de la difusión, que es el tiempo. Una debilidad de la investigación de difusión es la dependencia de los datos del recuerdo auto informados de los encuestados en cuanto a su fecha de adopción de una nueva idea. La última crítica que describió Rogers sobre la teoría de difusión es la escasa atención que los investigadores prestan a las **consecuencias de igualdad económicas** de una investigación. La tendencia muestra que generalmente los beneficios de una innovación no se distribuyen

equitativamente, y tiende a ampliar la brecha socioeconómica en una población entre los grupos de mayores y menores ingresos, especialmente y con más frecuencia en naciones que se encuentran en vías de desarrollo.

Ampliando los conceptos establecidos por E. Rogers sobre de las características del producto que intenta consolidar su adopción los autores Moore y Benbasat, basados en la teoría de la difusión de innovaciones, expanden el modelo creando nuevos atributos para medir la predicción de adopción de una innovación. Su trabajo se basó en una investigación acerca de la adopción de estaciones de trabajo individuales por parte de los individuos dentro de una organización. Los autores sugieren que la percepción del usuario acerca de una innovación afecta la adopción o la no adopción. Moore y Benbasat (1991) ampliaron los indicadores que miden las variables explicativas y la escala que proponen está ampliamente difundida en estudios posteriores.

Los atributos percibidos de una innovación, como señalan los autores, incluyen: **ventaja relativa, compatibilidad, facilidad de uso, capacidad de prueba, imagen, resultados demostrables, visibilidad y uso voluntario de la innovación tecnológica.**

Según Moore y Benbasat (1991), las características percibidas de una innovación han influido constantemente en la adopción y el uso de innovaciones. Los estudios muestran diferentes medidas del poder predictivo de los atributos y algunos resultados revelan que no todos los atributos influyen en la adopción, ya que las diferencias dependen de la innovación y del período en el que se aplican. Por lo tanto, es aconsejable no excluir ningún atributo del modelo. El atributo de **ventaja relativa** refleja el grado en que una innovación se percibe como mejor que su precursor, es decir, mejor que el producto que reemplaza. Esto puede medirse en términos económicos, de prestigio social, de conveniencia y satisfacción para la innovación en cuestión. La ventaja relativa está representada por la tasa de beneficios esperados proveniente del uso de una innovación (Moore & Benbasat, 1991).

Roger (1983) se refirió al atributo de **Facilidad de Uso** como complejidad. Los miembros de un sistema social adoptarán más rápidamente nuevas ideas que son fáciles de entender y absorber, mientras que las que requieren desarrollar nuevos conocimientos y comprensiones, y en consecuencia consideradas más complejas, se adoptan más lentamente (Rogers, 2003). Moore y Benbasat (1991) intercambiaron el atributo de complejidad propuesto por Rogers (1983) por el atributo "Facilidad de Uso", con el fin de determinar el nivel en el que una innovación se percibe

como fácil de usar. Para los autores, cuanto más fácil sea usar una innovación, mayor será la probabilidad de adopción.

En los atributos de **compatibilidad, capacidad de prueba y visibilidad** (observabilidad) no se produjeron diferencias según las definiciones desarrolladas por Rogers. Además de las cinco características anteriores, Moore y Benbasat (1991) introdujeron tres nuevos atributos: **imagen, uso voluntario y demostración de resultados.**

**Imagen** refleja el grado en que el uso de una innovación se percibe como una mejora de la imagen de un individuo o el estado de un sistema social. La **voluntariedad en el uso** es definida como la percepción del grado de utilización de una innovación de manera voluntaria o libre, mide de qué manera los individuos sienten libertad en adoptar o rechazar una nueva tecnología. **Demostrabilidad de resultados**, indica el grado en que los resultados del uso de una innovación son tangibles. Venkatesh y Davis (2000) también observaron una relación directa entre los resultados demostrables y la utilidad percibida. Argumentaron que los individuos forman percepciones más positivas de un sistema si la relación entre el uso y el rendimiento se detecta fácilmente (Moore & Benbasat, 1991) (Davis & Venkatesh, 2000).

Arts, Frambach y Bijmolt concuerdan que las intenciones de adoptar una innovación son **descriptorios débiles** para predecir acerca del comportamiento de adopción o adquisición de una innovación.

La intención de adopción se refiere al deseo expreso de un consumidor de comprar un nuevo producto en un futuro cercano. Se relaciona con el estado mental del consumidor antes de que ocurra el comportamiento real de compra y se basa en la información y las percepciones que tiene el consumidor en ese momento (Arts, Frambach, & Bijmolt, 2011).

El comportamiento de adopción, por otro lado, se refiere a efectuar la compra de una innovación. Los estudios sobre el comportamiento de adopción generalmente analizan las percepciones y características de los consumidores que ya han comprado la innovación en relación con aquellos que no lo han hecho. Estos últimos pueden incluir no adoptantes que tienen una alta o baja intención de adoptar o no adoptantes que incluso carecen de conocimiento de la innovación.

El número de diferentes variables utilizadas para capturar las características del adoptante es particularmente grande, ya que se ha dedicado una gran cantidad de investigaciones a encontrar rasgos de los consumidores que probablemente adopten una innovación. Las características del

adoptante capturan los rasgos personales que describen al adoptante (potencial) de una innovación, y las mismas pueden dividirse en sociodemográficas y psicográficas. (Tornatzky & Klein, 1982)

Las principales características sociodemográficas se centran particularmente en la edad de los consumidores, el nivel de educación y los ingresos. Otras variables menos relevantes que se consideran con frecuencia incluyen el tamaño del hogar, el género y el ciclo de vida familiar. Entre las características psicográficas podemos encontrar el grado de innovación de los adoptantes, el liderazgo de opinión, la propensión al consumo de los medios y el involucramiento como las variables más utilizadas para explicar la adopción. Las variables menos utilizadas incluyen, por ejemplo, conciencia de precios, familiaridad con la marca, confianza en sí mismo y dogmatismo (Gatignon & Robertson, 1985).

Arts, Frambach y Bijmolt basados en la teoría del nivel de interpretación (CLT por sus siglas en inglés) establecen que los comportamientos que están más distantes en el tiempo, como los reflejados por las intenciones de adopción de innovación, tienen más probabilidades de verse afectados por consideraciones relativamente abstractas o generales. Mientras que los comportamientos que están más cerca en el tiempo, como se refleja en el comportamiento de adopción (realización de compra), tienen más probabilidades de verse afectados por consideraciones concretas, específicas y dependientes del contexto. Esta distinción es especialmente importante para comprender cómo las características de innovación percibidas pueden afectar las etapas de intención y comportamiento del proceso de adopción (Trope & Liberman, 2003).

En el proceso de adopción de la innovación, la conveniencia y el deseo de adoptar un nuevo producto o servicio es mayor en la situación donde el adoptante potencial percibe que la innovación es ventajosa y compatible con sus necesidades. **La ventaja relativa** refleja los beneficios de la innovación sobre las ofertas alternativas, proporcionando al adoptante potencial una visión de su deseabilidad. **La compatibilidad** refleja el grado en que la innovación coincide con las necesidades y valores del adoptante potencial y, por lo tanto, es un aspecto importante de la conveniencia de la innovación para el individuo. Según la CLT, los potenciales adoptantes valorarán más los beneficios en términos de la ventaja relativa y la compatibilidad percibida de la innovación, ya que están más lejos del comportamiento de adopción real (es decir, se encuentran en la etapa de intención).

A medida que los individuos se acercan al comportamiento real de adquisición se espera que consideraciones más concretas afecten sus acciones. Con respecto a la adopción de la innovación, esto incluye la facilidad de uso del nuevo producto o servicio, es decir, su **complejidad**. Cuanto más se perciba una innovación como compleja, mayores serán los costos de aprendizaje para adoptar nuevos comportamientos (Hoeffler, 2003). Es probable que estos costos percibidos pesen especialmente en la etapa de comportamiento, ya que la necesidad de cambios de comportamiento específicos para adoptar con éxito la innovación se hace evidente. Cuanto más compleja es la innovación y, por lo tanto, mayores son sus costos percibidos, el cambio de comportamiento es menos factible, lo que inhibe a los consumidores a seguir adelante con la intención de adopción con un comportamiento real.

La **capacidad de prueba** permite al consumidor ver cómo funciona la innovación. Como tal, ayuda al adoptante potencial a evaluar el alcance del cambio de comportamiento requerido al adoptar la innovación. La capacidad de prueba mejora la preparación del consumidor, de modo que ayuda al adoptante potencial a comprender su papel y a confiar en sus habilidades utilizando la innovación. Por lo tanto, la capacidad de prueba puede afectar especialmente la viabilidad del consumidor para utilizar la innovación (Arts et al., 2011).

En la presente investigación se utilizará la teoría de difusión de innovaciones (DOI por sus siglas en inglés) descrita para comprender el marco general de adopción y en que estadio se encuentra el proceso tanto para el producto de altavoz inteligente como para su respectiva plataforma de asistencia virtual inteligente.

Si bien existen teorías posteriores a la formulada por E. Rogers que complementan y agregan capacidad de predicción en la adopción de innovaciones, en el estudio de tesis recurriremos al marco producido por E. Rogers en la teoría de difusión de innovaciones para conceptualizar la situación actual referida a la adopción de altavoces inteligentes dado la escasez de información con la cual contamos en la actualidad y la falta de estudios que analicen en profundidad los motivos y correlaciones que afectan su adopción.

Para ello utilizaremos estadísticas de comercialización validadas globalmente y producidas en los respectivos mercados sobre los cuales profundizaremos en el desarrollo del estudio.



### 3.1.1 – Ciclo de Vida de la Adopción de Tecnología

En su texto “Cruzando el Abismo” (Crossing the Chasm en inglés) G. Moore profundiza el modelo de adopción sintetizado y compilado por Rogers. El aporte que produce para esta investigación está dado por el foco de este acerca del proceso de adopción sobre productos tecnológicos disruptivos que de alguna manera modifican el comportamiento del usuario.

En los casos identificados por el autor la innovación no sólo demanda cambios significativos en el comportamiento del consumidor sino también en la infraestructura del proveedor y sus niveles de servicio, soporte y complementariedad que el oferente pudiera tener sobre un producto similar anterior (Moore, 1991).

Moore aplica una perspectiva puntualizada sobre la disciplina del Marketing como herramienta de diseño y catalizador del mensaje que debe ser transmitido entre los distintos segmentos de una comunidad o grupos con características psicológicas y sociales similares.

Los componentes del ciclo de vida no cambian, pero entre los grupos identificados por Rogers se han introducido brechas que representan de alguna manera la disociación entre los grupos con rasgos similares de personalidad, actitudes y aspiraciones. La dificultad que se describe en las brechas intercaladas entre cada uno de los grupos pasa por aceptar un nuevo producto si se presenta de la misma manera que al grupo de su izquierda inmediata.

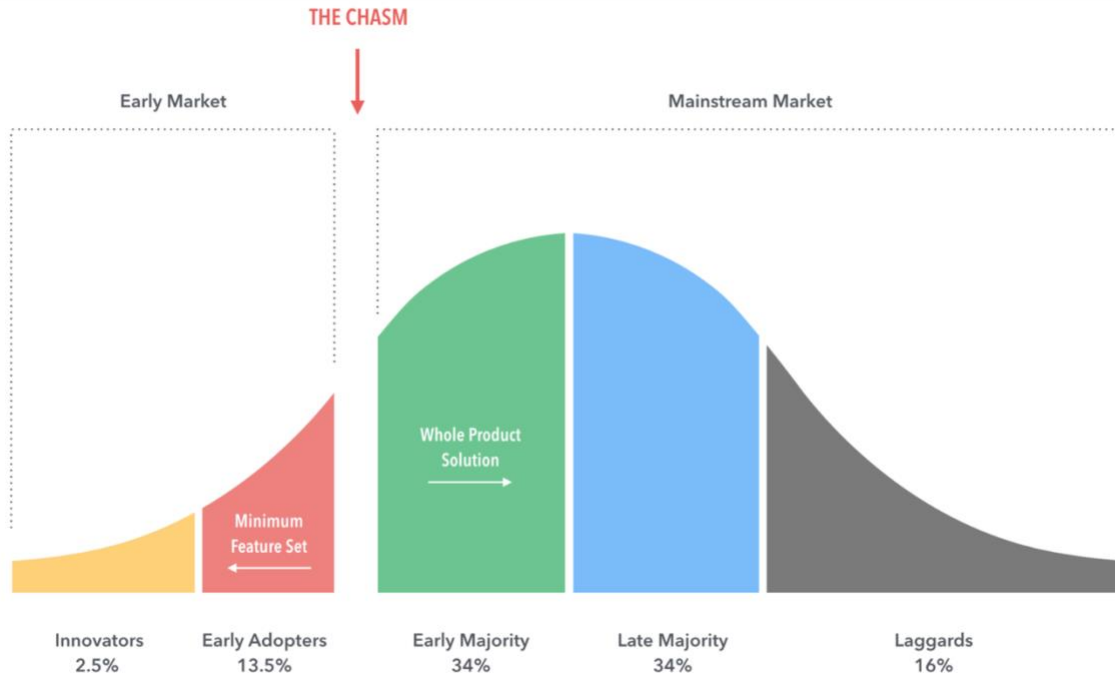


Figura 2. Curva de adopción y teoría del abismo. Fuente: <https://smithhousedesign.com/models-predicting-future-geoffrey-moores-crossing-chasm/>

Moore explica que de las fracturas presentadas entre los distintos grupos de afinidad, la más peligrosa es la que se da entre los Adoptantes Tempranos y la Mayoría Temprana porque generalmente, sobre los casos estudiados en la industria de bienes y servicios de alto impacto tecnológico, la diferencia entre los dos grupos no es reconocible.

Moore denomina abismo a la brecha que se establece entre los mercados reconocidos como Mercado Temprano y Mercado Principal. En resumen, cuando los fabricantes de productos de alta tecnología intentan hacer la transición desde una base de mercado compuesta por visionarios y primeros usuarios para penetrar en el próximo segmento de adopción, la mayoría temprana pragmática, están operando en la mayoría de los casos investigados sin una base de referencia y sin una base de apoyo dentro de un mercado altamente orientado a referencias y altamente orientado al soporte (Moore, 1991).

El fenómeno del abismo, que ocurre cada vez que se introduce una innovación discontinua, es evidenciada cuando las ventas pasan de una rápida aceleración en el desarrollo del mercado hacia de una dramática calma. Las empresas emergentes de alta tecnología llegan a un punto de

crisis donde deben dejar la relativa seguridad de su mercado inicial establecido e irse fuera en busca de un nuevo mercado, en este caso el principal (Moore, 1991).

Moore establece una analogía con el día “D” del desembarco en Normandía de la segunda guerra mundial para establecer una estrategia que permita a las empresas de alto impacto tecnológico cruzar el abismo ante el lanzamiento de un producto y/o el servicio disruptivo. La estrategia consiste en establecer un nicho de mercado específico donde la empresa entienda que puede dominar desde el principio de manera rápida y efectiva, expulsar a sus competidores de ese nicho de mercado y luego utilizar el caso de éxito logrado como una base de crecimiento para operaciones más amplias. Una característica importante que debe tener el nicho de mercado que la empresa determine es la eficiencia del proceso de comercialización. Cuanto más homogéneo sea el grupo de clientes objetivo, más fácil será crear e introducir mensajes en él y más rápido se transmitirán estos mensajes de boca en boca.

Cabeza de Playa es la denominación que Moore establece para este primer grupo de clientes objetivo en el mercado principal ya que él mismo, de ser exitoso, crearía una base de clientes referenciales para continuar la expansión hacia nuevos prospectos de mercado.

Una vez identificado el nicho de mercado principal que se decide convertir a cliente, las decisiones, tanto de canales de distribución como de precios, tienen un enorme impacto estratégico y, particularmente con los canales de distribución, ya que generalmente sólo hay una oportunidad para hacerlo exitosamente.

Moore identifica la problemática del abismo como un problema centrado exclusivamente en estrategias de marketing y describe las principales herramientas y tácticas para cruzarlo. Al mismo tiempo el autor advierte que las investigaciones efectuadas se reflejan sobre casos de innovaciones de productos y servicios de alta tecnología sobre un mercado orientado a organizaciones (B2B) y no a individuos (B2C).

Moore entiende que las empresas orientadas a la adopción de nuevos productos por parte de consumidores digitales también deben pasar a través de un proceso de adopción, pero no deben necesariamente atravesar la problemática del abismo para poder lograrlo. Para ello se identifican cuatro principales actividades que contribuyen fundamentalmente a escalar empresas que comercializan productos y servicios digitales a consumidores individuales:

- Adquirir tráfico.

- Atraer y comprometer usuarios.
- Monetizar la interacción de los usuarios.
- Alistar a los usuarios más fieles.

El orden de ejecución de las actividades es fundamental y son los pasos a seguir para llevar adelante el proceso de adopción digital. El primer paso se trata de lograr la atracción y el compromiso de los usuarios con el producto, es necesario que los usuarios demuestren cierta recurrencia de uso en este sentido. En segundo lugar, se debe comenzar con adquirir mayor tráfico sobre la innovación, en este punto resalta el autor que se debe ir iterando en ambas actividades. Esto significa ir generando crecimiento en ambos lados, tanto desde el lado de la demanda, atrayendo nuevos usuarios que demanden mayor funcionalidad o nuevos productos alternativos a la demanda de los primeros usuarios, como desde el lado de la oferta incrementando la calidad y cantidad del contenido y de las características del producto.

Si la iteración entre atraer posibles prospectos de usuarios a través del tráfico e involucrar usuarios no es suficiente para escalar masivamente la adopción de la idea, luego existe un tercer paso que es generar que los usuarios más fieles se conviertan en evangelizadores y a través de ellos lograr una mayor conversión de nuevos usuarios. La característica evangelizadora de los usuarios es clave luego para el marketing viral y la reducción de los costos en la adquisición de nuevos usuarios.

La monetización es la actividad más sensible de todas, entender en que momento del proceso se introduce es crítica para la sustentabilidad de la idea. Moore explica que al momento de introducir la monetización todas las demás actividades disminuyen su tasa de crecimiento y se vuelven más lentas (Moore, 1991).

S. Onoda, S. Kato y A. Mutoh presentan un enfoque para la difusión de innovaciones basada en el impacto de múltiples agentes sobre las estructuras de red por las cuales circula el flujo de información en la comunicación de una innovación. Como resultado indican que la velocidad de difusión de productos en el mercado ha cambiado porque la comunicación afecta los comportamientos y en consecuencia los estados mentales que influyen en las percepciones y actitudes.

Los autores entienden que el proceso de adopción se encuentra impactado por los efectos producidos por la externalidad de la red como resultante del valor de un producto y su dependencia

sobre la penetración en el mercado. Los mercados de productos que tienen externalidades de red están influenciados por la comunicación interpersonal (Onoda, Kato, & Mutoh, 2014).

La comunicación y el intercambio de información con otros individuos del mismo sistema social, y su comportamiento, afecta las acciones y psicologías de los mismos. El efecto de red (también llamado externalidad de red o economías de escala del lado de la demanda) en los mercados es un ejemplo de interacción y comunicación entre los sujetos que demandan el producto con sus semejantes que ya lo adquirieron y donde el mismo también produce un efecto importante en los individuos (Katz & Shapiro, 1985). Simplificando el concepto, el efecto de red significa que a medida que aumenta el número de usuarios en la adquisición y utilización de un elemento, ese elemento se vuelve más atractivo para el resto de los potenciales adoptantes pertenecientes al mismo entorno. En esta situación, se cree que la tasa de difusión de productos dentro de un mismo sistema social afecta la decisión de comprar o no compra del producto (Onoda et al., 2014).

En tales circunstancias, los productos de innovación con alta tecnología no necesariamente crean valor en las sociedades, sino que la mayor parte del valor se crea a través de la interacción entre consumidores y productores. El valor del producto depende de cómo se realice el consumo de los mismos y la medida en que los productos se hayan extendiendo en la sociedad (Ueda, Nishino, & Takenaka, 2009).

Así mismo cuando dos artículos similares están en una carrera para ser el más exitoso y masivo del mercado, existe un alto riesgo de que sólo se comparta un artículo exclusivamente debido a sus comentarios positivos y en consecuencia los artículos se vendan más rápido produciendo el fenómeno de “el ganador toma todo” (*Winner-Takes-All* en inglés) (Arthur, 1996).

### 3.1.2 – Modelo de Aceptación de Tecnología

Las investigaciones sobre adopción de tecnología se pueden dividir en dos categorías: en primer lugar, la adopción de nuevas tecnologías a nivel de empresa y en segundo lugar la adopción de nuevos productos y servicios tecnológicos a nivel individual.

Las investigaciones llevadas adelante sobre cuestiones y problemas de adopción de nivel organizacional se concentran respecto a cómo los empleados en la organización evalúan la satisfacción de uso y la utilidad con respecto a la adopción de una nueva tecnología en los procesos de trabajo. Por otro lado, los estudios realizados a nivel individual se refieren a cómo los usuarios

o clientes evalúan la satisfacción con respecto a las adopciones y usos de una tecnología novedosa en sus vidas desde las dimensiones de la **facilidad de uso percibida** y la **utilidad percibida**.

En general, el modelo de aceptación de tecnología siempre se ha considerado como un marco analítico útil para verificar o explorar la aceptación de los usuarios y la adopción de nuevas tecnologías avanzadas.

El Modelo de aceptación de tecnología (TAM por sus siglas en inglés) es una teoría acerca de los sistemas de información que modela cómo los usuarios aceptan y usan una tecnología (Davis, 1985). Desarrollado originalmente por Davis (1985) se ha utilizado ampliamente en los últimos años para comprender la adopción y el uso de nuevas tecnologías.

Davis entiende que el uso real del sistema es el objetivo final donde queremos evaluar el comportamiento de la adopción de la tecnología, por lo que es necesario formar la intención de comportamiento, que es el factor que lleva a las personas a usar la tecnología. La intención de comportamiento (BI por sus siglas en inglés) está influenciada por la actitud (A por su sigla en inglés), que es la impresión general hacia la tecnología.

El modelo sugiere que cuando a los usuarios se les presenta una nueva tecnología, una serie de factores que se agrupan en dos dimensiones de análisis influyen en su decisión sobre cómo y cuándo la usarán:

- **Utilidad percibida** (PU por sus siglas en inglés): Significa si alguien percibe o no que la tecnología es útil para lo que quiere hacer.
- **Facilidad de uso percibida** (PEOU por sus siglas en inglés): Davis definió esto como "el grado en que una persona cree que usar un sistema en particular estaría libre de esfuerzo". Si la tecnología es fácil de usar, entonces las barreras serán superadas. Si no es fácil de usar y la interfaz para su utilización es complicada, nadie tendrá una actitud positiva hacia ella.

Ciertas variables externas, como la influencia social, también pueden ser un factor importante para determinar la actitud hacia la aceptación de la tecnología. Davis entiende que cuando las variables descriptas anteriormente estén dadas, las personas tendrán la actitud y la

intención de usar la tecnología. Sin embargo, la percepción puede cambiar dependiendo de la edad y el género ya que cada persona es diferente (Davis, 1989).

La importancia del modelo TAM se observa en numerosas disciplinas en las que el modelo se ha utilizado para comprender la adopción de tecnologías (Rese, Baier, Geyer-Schulz, & Schreiber, 2017).

Davis (1989) subrayó que el impulso para utilizar determinada tecnología puede explicarse por las actitudes que un individuo tiene hacia la misma en base a su apreciación acerca de la utilidad percibida y la facilidad de uso percibida.

Análisis posteriores sobre TAM encontraron que la utilidad percibida y la facilidad de uso percibida explican alrededor del 40% de la variación en la intención de comportamiento de un individuo para usar una tecnología. El análisis de las investigaciones realizadas utilizando TAM muestra que los resultados no son totalmente consistentes o claros. Esto sugiere que no se incluyen factores significativos en los modelos predictores.

Las críticas realizadas a la “teoría” TAM incluyen su falta de falsabilidad, su cuestionable valor heurístico y su limitado potencial predictivo y explicatorio. TAM demuestra ser un modelo útil, pero debe integrarse en uno más amplio que incluya variables relacionadas con procesos de cambio tanto humanos como sociales, y a la adopción de un modelo de innovación (Legris, Ingham, & Collerette, 2003).

El comportamiento del consumidor deja al descubierto ciertas falencias en el modelo TAM, al revelar diferentes determinantes de adopción relacionados con la motivación, que complementan las variables generalmente consideradas por los investigadores, es decir, identificar las características personales de los consumidores. Se torna necesario diferenciar el comportamiento de adopción entre diferentes perfiles de consumidores. En consecuencia, las críticas se han dirigido a TAM debido al modelo demasiado simplificado acerca de la adopción de tecnología (San-Martín, López-Catalán, & Ramón-Jerónimo, 2013).

El modelo TAM se ha estudiado y ampliado continuamente: las mejoras principales se han logrado en el TAM 2 (Davis & Venkatesh, 2000) y en la teoría unificada de aceptación y uso de la tecnología (UTAUT por sus siglas en inglés) (Venkatesh, Morris, Davis, & Davis, 2003). También se ha propuesto un TAM 3 en el contexto del comercio electrónico con una inclusión en los efectos de la confianza y el riesgo percibido en el uso del sistema (Venkatesh & Bala, 2008).

TAM 2 incorpora tanto los procesos de influencia social (norma subjetiva, voluntariado e imagen) como los procesos instrumentales cognitivos (relevancia laboral, calidad de salida, demostrabilidad de resultados y facilidad de uso percibida), los cuales influyeron significativamente en la aceptación del usuario. Estos hallazgos evolucionan la teoría y contribuyen a la base de futuras investigaciones destinadas a mejorar la comprensión acerca del comportamiento de adopción de tecnología por parte del usuario.

El objetivo del modelo UTAUT es explicar y describir las intenciones de los usuarios al utilizar un sistema de información y su comportamiento de uso posterior.

El UTAUT fue propuesto por Venkatesh como un modelo integrador de ocho teorías o modelos de aceptación de tecnología relacionados. Esas teorías o modelos incluyen la teoría de la difusión de la innovación (DOI), la teoría de la acción razonada (TRA), la teoría del comportamiento planeado (TPB), la teoría de la motivación (MM), el modelo de aceptación tecnológica (TAM original), el modelo híbrido de TPB y TAM, el modelo de utilización de computadoras personales (MPCU) y la teoría cognitiva social (SCT).

Los resultados de la investigación realizada plantean tres determinantes directos a la intención de usar tecnología (la expectativa de rendimiento, la expectativa de esfuerzo y la influencia social) y dos determinantes directos del comportamiento de uso (la intención y las condiciones facilitadoras). Las influencias moderadoras y significativas acerca de la experiencia de utilización de tecnología como la voluntariedad, el género y la edad se confirmaron también como características integrales del modelo UTAUT. El estudio longitudinal realizado en el ensayo encontró que el modelo UTAUT puede representar hasta el 70% de la variación en la intención de uso (Venkatesh et al., 2003).

La motivación para desarrollar el modelo UTAUT fue integrar las numerosas variables superpuestas de distintos modelos de adopción y utilizarlas para explicar la adopción de tecnología y en consecuencia crear una base teórica "unificada" (Williams, Rana, Dwivedi, & Lal, 2011).

Los esfuerzos para llevar adelante una completa evaluación utilizando el modelo UTAUT y su extensión para proporcionar una base teórica unificada son los puntos débiles para poder trabajar con el modelo en las investigaciones de adopción de tecnología.

Bagozzi (2007) criticó la utilización del modelo UTAUT argumentando que un modelo con 41 variables independientes para predecir intenciones y otras con 8 variables para predecir el comportamiento alcanza la saturación.



Incluso Bagozzi (2007) argumenta que se han omitido importantes variables independientes, porque pocos de los predictores incluidos son fundamentales, genéricos o universales, y es probable que las investigaciones futuras descubran nuevos predictores no contemplados bajo las variables existentes, por lo tanto, el modelo UTAUT se vuelve de poca ayuda para informar la adopción y el uso de nuevas tecnologías (Bagozzi, 2007).

Van Raaij y Schepers (2008) critican aún más el modelo UTAUT, señalando que la varianza que explica el porcentaje de predicción sobre la variable dependiente en el modelo es alta sólo al moderar relaciones clave con cuatro variables (género, edad, experiencia y voluntariado). Esto produce que el modelo sea menos parsimonioso que TAM y TAM2. Más importante aún, se encontraron dificultades en la agrupación y el etiquetado de artículos y construcciones de variables como las condiciones facilitadoras y la influencia social (Van Raaij & Schepers, 2008).

El modelo UTAUT2 es una extensión del modelo UTAUT y se centró en las perspectivas del comportamiento individual en la adopción de tecnología. El nuevo modelo fue significativamente mejorado para explicar las variaciones en la intención de utilización de nuevas tecnologías por parte de los usuarios (Venkatesh, Thong, & Xu, 2012).

A pesar del modelo integrado UTAUT y la falta de consenso acerca de la definición de incorporar ciertas variables para la predicción de aceptación tecnológica. Venkatesh enfatiza la necesidad de incluir variables predictoras especiales que puedan usarse dentro de un contexto de uso y adopción de tecnología en individuos o consumidores (B2C). Se examinaron estudios más relacionados al comportamiento de los consumidores y se modificó la perspectiva previa establecida (de organizaciones a individuos) ajustando el modelo UTAUT para establecer un nuevo marco de predicción. Actualmente, este modelo más nuevo se ha adoptado gradualmente para investigar diversos temas, como el auto-servicio de tecnología, la adopción de dispositivos móviles inteligentes, la aceptación del software de gestión del aprendizaje y la industria del cuidado de la salud (C. Huang & Kao, 2014).

Con respecto al modelo de predicción, la variable de motivación hedónica se consideró un predictor importante y se integró al modelo UTAUT2 para una utilización más intensiva. **La motivación hedónica** se define como la diversión o el placer derivado del uso de una tecnología, y se ha demostrado que juega un papel importante en la determinación de la aceptación y el uso de la tecnología en individuos (Brown & Venkatesh, 2005).

La variable del **valor del precio** también se introdujo en el modelo UTAUT2 porque la calidad, el costo y el precio del producto influirán en las decisiones de adopción. Una diferencia importante entre un contexto de uso del consumidor y el contexto de uso de la organización, donde se desarrolló UTAUT, es que los consumidores suelen asumir el costo monetario de dicho uso, mientras que los empleados no. La estructura de costos y precios puede tener un impacto significativo en el uso de tecnología por parte de los consumidores. El valor del precio es positivo cuando se percibe que los beneficios del uso de una tecnología son mayores que el costo monetario y en consecuencia el valor del precio tiene un impacto positivo en la intención de adopción.

Venkatesh también consideró que los estudios recientes han enfatizado la importancia de los roles de la intención de comportamiento. Por lo tanto, se incorporó una nueva variable de **hábito** en el UTAUT2. La introducción del **hábito** se debió a las siguientes dos razones. Primero, el hábito se considera y se relaciona en base a un comportamiento previo existente y en segundo lugar, el hábito puede definirse como el grado en que las personas creen que el comportamiento realizado es automático. Estas nuevas construcciones agregadas se verificaron constantemente en investigaciones anteriores como los determinantes críticos para la adopción de tecnológicas por parte de consumidores individuales (Venkatesh et al., 2012).

Debido a que cada producto de innovación tecnológica tiene características específicas, ninguno de estos modelos es directamente aplicable en su totalidad para examinar la aceptación de cada tecnología y su contexto de usuario.

Por lo tanto, la mayoría de los estudios de aceptación utilizan el modelo TAM como base y lo complementan con factores relevantes para la aceptación de la tecnología específica.

Este es también el caso en estudios que examinan la aceptación de productos con características comparables con los altavoces inteligentes. Por ejemplo, los artefactos involucrados en el diseño de un hogar inteligente (Yang, Lee, & Zo, 2017), los dispositivos utilizados como accesorios o vestimenta inteligente (Kalantari, 2017) o más precisamente los relojes inteligentes (Kim & Shin, 2014) que también se encuentran conectados a una plataforma de información a través de internet y que contemplan la función de comunicar o entregar información al consumidor.

Un estudio reciente de aceptación tecnológica sobre altavoces inteligentes utilizando el marco metodológico TAM ha evidenciado que la variable facilidad de uso percibida (PEOU) de los altavoces inteligentes no tiene una correlación directa con la intención de comportamiento (BI) de los potenciales adoptantes al uso del dispositivo. El estudio concluyó que los principales

factores que influyen sobre BI para la utilización de los altavoces inteligentes son la utilidad percibida (PU), el entretenimiento percibido (ENJ) y el riesgo percibido (R).

Como los altavoces inteligentes están diseñados para ser asistentes personales en el hogar, la utilidad percibida (PU) se identifica como la medida en que un individuo cree que el uso de altavoces inteligentes aumentaría su rendimiento en la vida cotidiana en el hogar.

Los principales dos factores en tener preponderancia sobre le PU son la diversidad del sistema (SD), entendida como la percepción de las distintas funciones que un sistema puede realizar y el optimismo tecnológico (TO), el cual se encuentra relacionado con las personas que creen que la tecnológica les ofrece mayor control, flexibilidad y eficiencia en sus vidas. En tercer lugar, el estudio define a la calidad percibida del sistema (SQ), entendida como el grado de cumplimiento de las funciones de un producto en tener influencia sobre la PU.

Los altavoces inteligentes que se perciben como complejos y difíciles de usar también se considerarán que tienen problemas de rendimiento y calidad. Por el contrario, se supone que los consumidores que perciben un sistema fácil de usar también percibirán un SQ más alto.

De la misma forma que el PU, el entretenimiento percibido (ENJ) afecta positivamente el BI, mientras que el riesgo percibido tiene una correlación negativa con el BI. El riesgo percibido se encuentra compuesto por el miedo a que los datos sensibles del usuario sean robados o puestos en riesgo y el temor a estar siendo vigilados de manera constante por la escucha pasiva de los dispositivos (Kowalczyk, 2018).

San Andrés

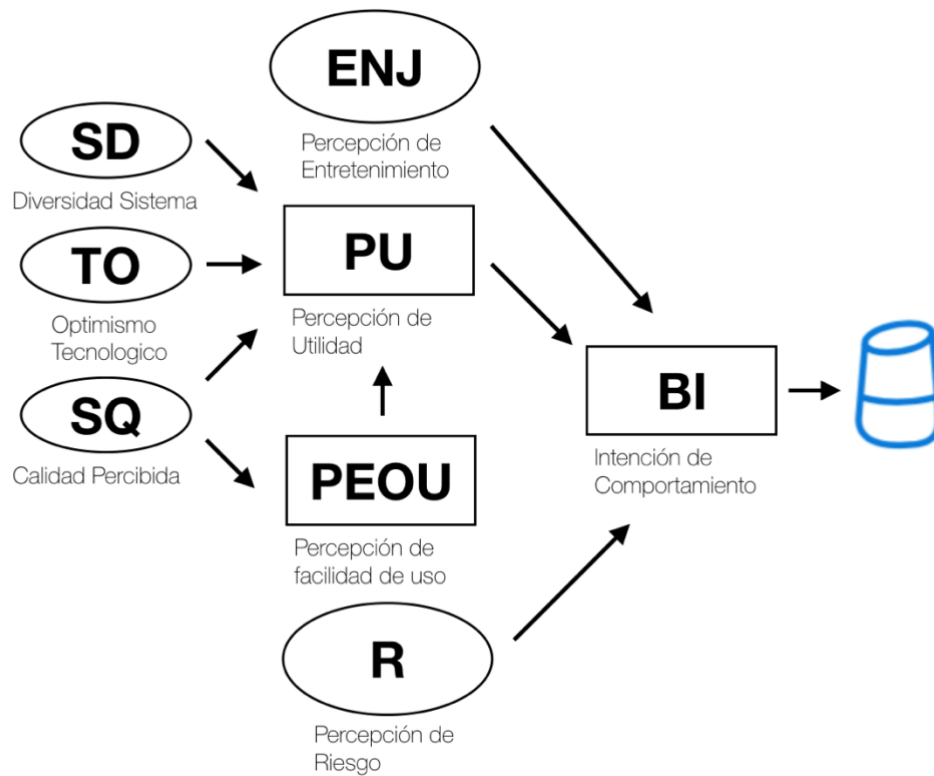


Figura 3. Principales variables constructoras identificadas para el estudio de adopción de altavoces inteligentes. Elaboración propia. Fuente: (Kowalczyk, 2018)

## 4. Capítulo IV - Definición de Altavoz Inteligente y descripción de sus componentes.

En el desarrollo del capítulo actual se aborda la definición de altavoz inteligente (Smart Speaker) y los principales productos y servicios que componen el dispositivo final.

Describiremos conceptualmente la tecnología que utilizan las plataformas donde se encuentran alojados los asistentes virtuales inteligentes y los fundamentos por los cuales las distintas características que lo componen han ido evolucionando hasta llegar a convertirse en los principios que hoy rigen sobre diferentes campos de la inteligencia artificial empleados en el dispositivo.

El objetivo de este capítulo es tener una clara comprensión acerca del producto que es objeto de estudio de esta investigación, la tecnología empleada para sus funciones y su correspondiente evolución.

### 4.1 – Definición de Altavoz Inteligente

En la actualidad los seres humanos se encuentran completamente rodeados por computadoras, la forma de interacción que se produce entre los humanos y las computadoras juega un papel muy importante. Las computadoras deben comprender las instrucciones proporcionadas por los humanos y deben ser capaces de producir resultados en consecuencia (Sinha, Shahi, & Shankar, 2010).

La Asociación de Máquinas de Computadora (ACM por sus siglas en inglés) define la interacción humano-computadora (HCI) como "una disciplina relacionada con el diseño, evaluación e implementación de sistemas informáticos interactivos para el uso humano y el estudio de los principales fenómenos que lo rodean" (ACM, 1997).

Una faceta importante de HCI es la satisfacción del usuario final, debido a que la interacción humano-computadora que estudia a un humano y una máquina en comunicación se basa en el conocimiento obtenido tanto sobre la máquina como en el ser humano. Por el lado de la máquina, podemos encontrar que las técnicas en gráficos por computadora, sistemas operativos, lenguajes de programación y los entornos de desarrollo son relevantes. En el lado humano, la teoría

de la comunicación, las disciplinas de diseño gráfico e industrial, la lingüística, las ciencias sociales, la psicología cognitiva, la psicología social y factores humanos como la satisfacción del usuario son los más relevantes.

HCI también se denomina frecuentemente como interacción humano-máquina (HMI), interacción hombre-máquina (MMI) o interacción computadora-humano (CHI). Los seres humanos interactúan con las computadoras de muchas maneras. La interfaz entre el humano y la computadora es el componente crucial para facilitar que la interacción se realice. Las aplicaciones de escritorio, los navegadores de Internet, las computadoras de mano, los sistemas ERP o CRM y los quioscos informáticos hacen uso de las interfaces gráficas de usuario (GUI). La interacción humano-computadora (sin importar el tamaño y la forma de la computadora) ha ido cambiando el desarrollo de las computadoras durante la historia (Xing Huang, 2015). Huang define que los tres dispositivos de entrada principales mayormente conocidos para realizar la interfaz HCI son el teclado, el mouse y por último los dedos que se utilizan sobre la pantalla táctil del teléfono inteligente, la tableta, terminales inteligentes, etc.

Las interfaces de usuario de voz (VUI) se utilizan para el reconocimiento de voz y los sistemas de sintetización de sonidos (Sinha et al., 2010). En la actualidad, el habla y la voz también se pueden utilizar como mecanismos de entrada para productos tecnológicos y se pueden definir como uno de los más naturales y convenientes de comunicación de información. Una gran parte de la comunicación en la vida diaria humana se produce a través del habla, esta interfaz es más rápida a comparación de otros tipos de entrada como las pulsiones de teclados individuales o los clics del mouse, al mismo tiempo la productividad del usuario puede incrementarse notablemente ya que se proporciona un canal de respuesta adicional a través del altavoz (G. L. Martin, 1989).

Si bien los asistentes de voz han sido un anhelo de las comunidades de HCI e inteligencia artificial (AI) durante décadas, esta nueva era de altavoces inteligentes representa las primeras interfaces de voz ambiental que están disponibles en entornos domésticos, para millones de personas (Bentley et al., 2018).

El altavoz inteligente, al ser un producto nuevo con una incipiente aceptación reciente, no cuenta todavía con una definición claramente establecida. En artículos, informes y papers académicos se habla de Parlante Controlado por Voz (Voice-controlled Speaker) o Parlante con Inteligencia Artificial (Artificial Intelligent Speaker) o Parlante con Asistente Digital (Digital Assistant Speaker).

Los asistentes de voz son agentes de software que pueden interpretar la voz y el habla del ser humano y responder a través de voces sintetizadas. Siri de Apple, Alexa de Amazon y el Asistente de Google son los asistentes de voz más populares y están integrados en teléfonos inteligentes y parlantes domésticos dedicados. Los usuarios pueden hacer preguntas a sus asistentes, controlar los dispositivos de automatización del hogar, la reproducción de medios a través de la voz y administrar otras tareas básicas como correo electrónico, listas de tareas y calendarios mediante comandos verbales (Hoy, 2018).

Por lo general, un altavoz inteligente significa un dispositivo de altavoz o parlante con manos libres, lo cual comprende tener incorporado un micrófono que permite realizar una conversación con un asistente de voz digital, que se encuentra embebido en una plataforma de inteligencia artificial y utiliza tecnología de interpretación y procesamiento del lenguaje natural (NLP por sus siglas en inglés) de la voz bidireccional, basada en la computación en la nube (Amazon, 2017) (E. Martin, 2017).

Por lo tanto, los usuarios no tienen que depender de teléfonos celulares o realizar acciones en otro dispositivo para interactuar con el asistente digital al momento de hacer un pedido, dar una orden o hacer una consulta a través del altavoz. Los usuarios pueden pedir Uber, consultar información, pedir comida, escuchar música o noticias o controlar cualquier dispositivo inteligente conectado al hogar.

Una publicación del Instituto Nacional de Estándares y Tecnologías de Estados Unidos (Chung, Iorga, Voas, & Lee, 2017) define a esta tecnología como Asistentes Virtuales Inteligentes (IVA por sus siglas en inglés). Existen dos tipos de asistentes virtuales inteligentes el primero denominado “built-in” donde en su tramo final de conexión utiliza distintos dispositivos con múltiples propósitos, como por ejemplo el caso de Siri de la compañía Apple y el ya discontinuado Cortana de la firma Microsoft.

La segunda categoría son los denominados “stand-alone” los cuales utilizan un dispositivo dedicado para su tramo final de conexión, en esta categoría ingresan Alexa que utiliza Echo y Echo Dot & Tab pertenecientes a Amazon y Google Assistance en el dispositivo Google Home de la compañía Alphabet.

La funcionalidad principal, podríamos decir el "cerebro" de un IVA, se aloja como un servicio en la nube que procesa datos a través del reconocimiento de la voz o el habla.

Resumiendo el proceso que utilizan los asistentes virtuales inteligentes, en el comienzo se convierte el audio de la voz capturada por el micrófono del dispositivo a un archivo de texto, luego se realiza un análisis de contexto lingüístico a través de software que interpreta y procesa el lenguaje natural (NLP), de manera posterior al proceso de NLP existe una salida o output que en este caso es la elaboración de una respuesta en un archivo de texto y finalmente el texto se vuelve a convertir a voz para su reproducción en el dispositivo o la ejecución de una instrucción como por ejemplo abrir la aplicación de Spotify y reproducir una canción determinada.

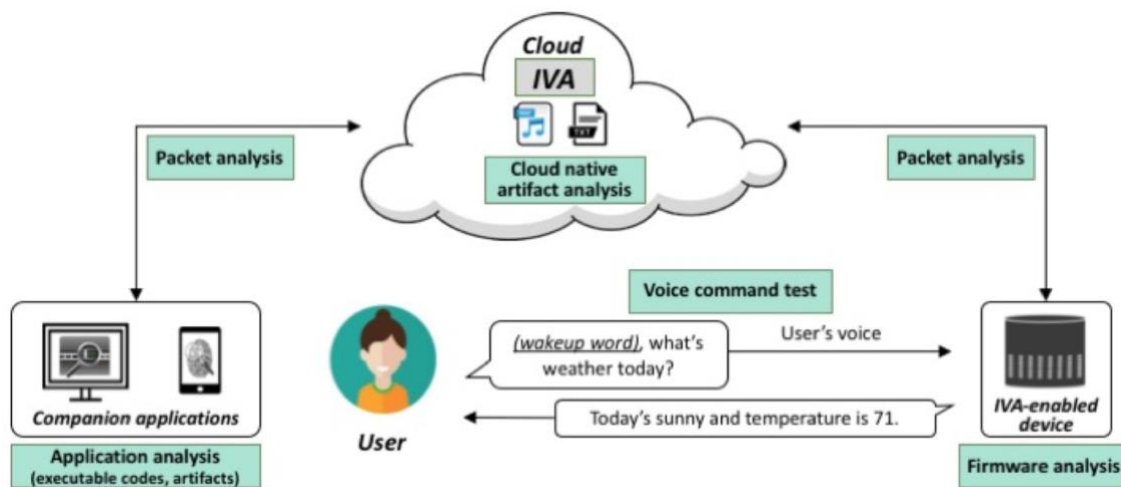


Figura 4. Diagrama arquitectura IVA. Fuente: National Institute of Standards and Technology

Los asistentes personales inteligentes buscan interactuar y poder sostener una conversación con la misma inteligencia con la que la realizaría cualquier otro ser humano. Estos dispositivos, gracias a su gran sofisticación en sus algoritmos, tienen la capacidad de aprender de manera no supervisada acerca de las rutinas, costumbres, deseos y hasta las maneras de comunicarse con los distintos integrantes de las familias en los hogares donde se encuentra instalado (Stucke & Ezrachi, 2017).

Más allá de poder ejecutar ciertas funciones cotidianas como requerir información a través del IVA estos dispositivos se pueden convertir en la puerta de entrada a internet por el lado de los individuos y anticipar las necesidades y pedidos de sus usuarios (Stucke & Ezrachi, 2017).



El asistente virtual normalmente se encuentra en una actitud pasiva esperando interactuar con su usuario. El software se activa al escuchar una palabra clave en donde el mayordomo digital se enciende. Una vez que el IVA escucha la palabra clave, graba la voz del usuario y la envía a un servidor especializado, que procesa e interpreta el texto como un comando o una acción a realizar. Dependiendo del comando, el servidor proporciona al asistente de voz la información adecuada para que pueda ser transmitida nuevamente al usuario o, dependiendo del caso, puede reproducir los medios solicitados por el usuario o completar tareas con mediante otros servicios y dispositivos conectados. La cantidad de servicios y programas que admiten voz para la ejecución de los comandos están creciendo rápidamente y los fabricantes de dispositivos conectados a internet de las cosas también están incorporando el control por voz en sus productos.

En la siguiente tabla visualizamos algunas de las principales funcionalidades que los múltiples asistentes virtuales pueden realizar:

<b>PRINCIPALES FUNCIONES DISPONIBLES EN LOS AVI</b>
Enviar y recibir mensajes de texto del celular, realizar llamadas telefónicas, enviar y leer mensajes de email.
Responder preguntas básicas como, por ejemplo; ¿Qué hora es? ¿Cuál es el pronóstico del clima? ¿Cuál es la traducción entre distintos idiomas o unidades de medida?
Programar recordatorios de alarmas y calendarios.
Recordar listas de tareas y hacer simples cuentas matemáticas.
Controlar la reproducción y los servicios de streaming de contenido como Spotify, Netflix, Itunes, y Google Play.
Controlar dispositivos inteligentes del hogar como termostatos, electrodomésticos, luces y aires acondicionados.
Contar historias o cuentos para los niños.

*Figura 5. Tabla de principales habilidades de un IVA. Elaboración propia, fuente utilizada (Hoy, 2018).*

Además de estas tareas, los asistentes de voz pueden agregar otras funciones, a menudo llamadas "habilidades", que amplían sus funcionalidades al interactuar con otros programas

a través de comandos de voz. Alexa, la plataforma de Amazon, tiene habilidades para jugar Jeopardy (un juego televisivo muy popular en los Estado Unidos de preguntas y respuestas), ordenar la bebida habitual en un local de Starbucks o pedir un servicio de Uber utilizando datos de cuenta asociados al usuario. El asistente de Google tiene habilidades similares, pero se encuentra retrasado detrás de Amazon en la gran cantidad de habilidades disponibles, en gran parte debido a la demora en su lanzamiento y a la estrategia adoptada para la evolución de su plataforma.

### Total number of Amazon Alexa skills from January 2016 to December 2018

Growth of Amazon Alexa skills from 2016 to 2018

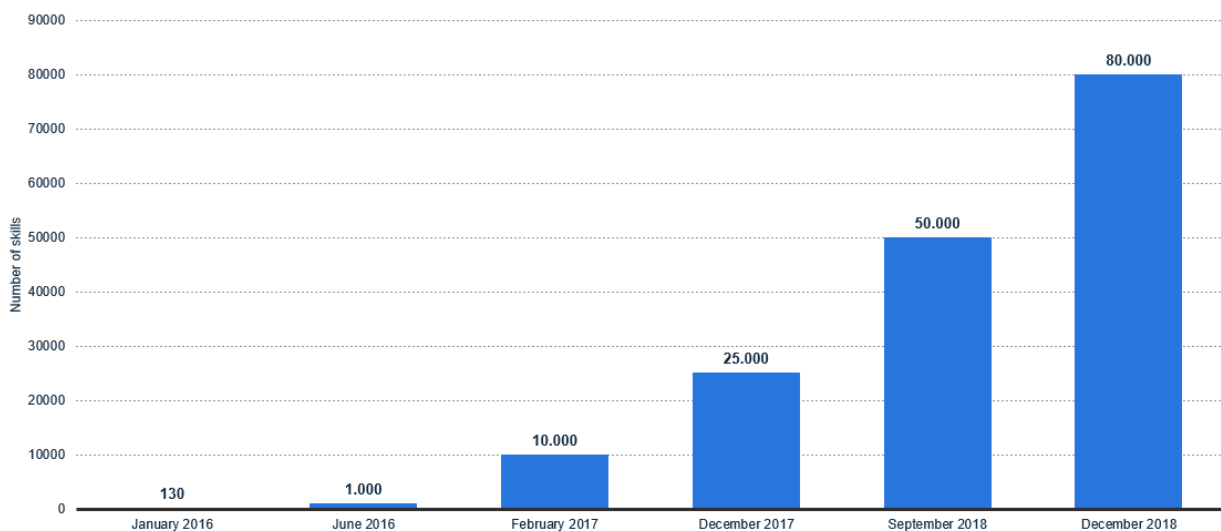


Figura 6. Evolución de habilidades desarrolladas para Alexa. Fuente: Amazon.

Si bien las tendencias de uso y funcionalidad por plataforma no demuestran todavía una utilización sofisticada de la tecnología (IHS, 2018), un informe de la organización Consumer Technology Association reveló que en el año 2018 40% de usuarios en los Estados Unidos realizaron al menos una compra a través del dispositivo con asistente virtual inteligente (CTA, 2018).

## Smart speaker with intelligent personal assistant common uses in selected countries as of 2018, by platform

Smart home voice assistants usage by platform in selected countries 2018

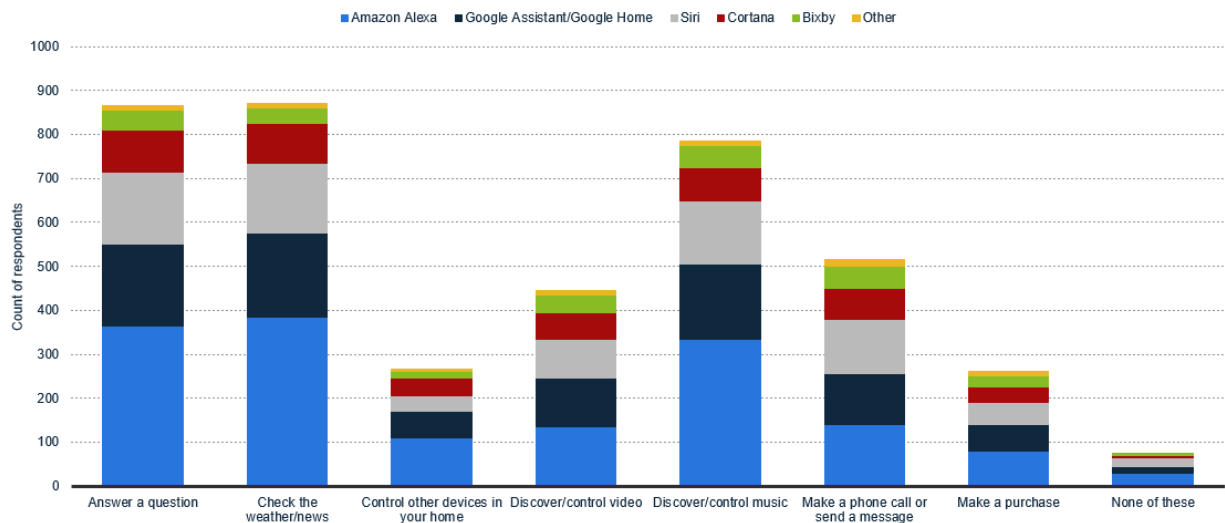


Figura 7. Funcionalidades ejecutadas en los asistentes virtuales durante 2018 por plataforma. Fuente: IHS.

### 4.2 - Plataforma de Inteligencia Artificial

Como anticipamos en el capítulo 2, un altavoz inteligente contiene dos productos asociados. El primero es el producto físico y visible que permite la interacción con un IVA a través de su micrófono, parlante, CPU y placa de red que posibilita su conexión a internet y el segundo es un servicio que trae el dispositivo que es la posibilidad de conectarse e interactuar con el IVA, el mismo se encuentra alojado en una plataforma de inteligencia artificial, mediante la cual se puede efectuar y ejecutar distintas habilidades o funciones según el fabricante del dispositivo.

La plataforma de inteligencia artificial se encuentra construida en los centros de procesamientos de información que cada competidor tiene desarrollado y utiliza el poder de cómputo de la nube para efectuar las distintas tareas que el usuario solicita.

En los capítulos siguientes entraremos en los detalles de cada una de estas plataformas de inteligencia artificial. Conceptualmente las principales funciones de inteligencia artificial utilizadas en el altavoz inteligente son el procesamiento y entendimiento del lenguaje natural y el reconocimiento de la voz o el habla.

## 4.2.1 – Evolución de Inteligencia Artificial

El primer trabajo que en la actualidad es reconocido como un trabajo de inteligencia artificial fue realizado por W. McCulloch y W. Pitts en 1943. En su ensayo propusieron un modelo de neuronas artificiales capaces de encenderse o apagarse ante un estímulo determinado proveniente de neuronas vecinas. El estudio demostró que cualquier función computable podría ser calculada por una parte de una red de neuronas conectadas entre sí.

En 1950 M. Minsky y D. Edmons en la Universidad de Harvard construyeron la primera red neuronal en computadoras capaz de simular el comportamiento de 40 neuronas humanas. En el mismo año A. Turing conceptualizó una visión sobre inteligencia artificial, que a posterior fue la que mayor impacto ha tenido hasta la actualidad. Turing propuso la idea de un “Programa o Computadora Infantil” (“Child Programme or Machine” en inglés) el cual pretendía que los esfuerzos en intentar construir un programa o sistema de computación que fuese capaz de reproducir la mente de un ser humano deberían ser concentrados en emular la mente de un niño y no la de un adulto.

En 1956, J. McCarthy junto a M. Minsky y otros investigadores realizaron el primer congreso de dos meses de duración donde se juntaron la mayoría de los investigadores relacionados a los campos de redes neuronales, teorías autómatas y estudios de la inteligencia con el objetivo de avanzar en temas relacionados a la inteligencia del aprendizaje y como las máquinas podrían ser capaces de simular ese comportamiento. El congreso no logró realizar importantes nuevos avances en la temática propuesta, pero de alguna manera aglutinó e introdujo a las principales figuras de las distintas universidades que se encontraban investigando temas relacionados a la inteligencia artificial y lo seguirían haciendo en los próximos.

J. McCarthy introdujo la iniciativa de darle al campo de la inteligencia artificial un trato propio de un campo de investigación específico de las ciencias de la computación. Es por ello que logró evitar que el estudio de la inteligencia artificial cayera en una rama de estudio de las matemáticas, ya que en el propósito de la inteligencia artificial se intenta simular facultades humanas como el lenguaje, la creatividad y la mejora del auto aprendizaje.

Los autores S. Russell y P. Norvig en su libro “Inteligencia Artificial, un enfoque moderno” establecen una mirada particular sobre la definición de inteligencia artificial. Al ser este un campo amplio de estudio podemos ordenar las diferentes definiciones realizadas sobre la materia, según los siguientes aspectos y objetivos que se encuentren en la investigación.

Por un lado, podemos abordar el concepto de inteligencia artificial con una mirada acerca del proceso por el cual se genera el resultado final o salida (output), el mismo puede contemplar tanto el modo mediante el cual se efectúa el razonamiento o pensamiento para resolver una situación o bien el comportamiento obtenido durante el proceso. Un segundo enfoque se centra sobre los resultados obtenidos, en este caso se tendrá en cuenta si los mismos fueron exitosos en términos de un ideal de performance en base al conocimiento acumulado o bien su semejanza y fiabilidad a lo que hubiese sido el resultado de un comportamiento humano (Russell & Norvig, 2016). De esta manera podemos distinguir cuatro vertientes de abordaje sobre la inteligencia artificial:

### **Actuar Humanamente**

Alan Turing en su ensayo “Computing Machinery and Intelligence” introduce el concepto y la pregunta acerca de si las máquinas pueden pensar. La ambigüedad del término “pensar” hizo que finalmente se transfiriera la incógnita a una prueba donde se expusiera la habilidad de una máquina para exhibir un comportamiento inteligente similar al de un ser humano o indistinguible de este. A. Turing propuso que una persona, en este caso un entrevistador, evaluara conversaciones y diálogos que se reproducían a través de una pantalla y un teclado entre dos entrevistados, en este caso un ser humano y una máquina diseñada para generar respuestas similares a las de un humano. La máquina pasaría la prueba de Turing si el interrogador no puede discernir si las respuestas a sus preguntas provienen de la máquina o de la persona (A. Turing, 1950).

### **Pensar Humanamente**

En esta dimensión se vuelve necesario conocer y determinar cual será el diagrama o esquema de pensamiento humano para poder establecer el estilo de comportamiento que se desea obtener para el programa o desarrollo de software que ejecutará el proceso de inteligencia artificial.

Esta orientación de la inteligencia artificial se encuentra ligada al ámbito de la ciencia cognitiva y de la psicología, donde a través de los modelos computacionales se intenta emular de que manera lo haría, pensaría y ejecutaría una mente humana.

### **Pensar Racionalmente**

Esta vertiente de la inteligencia artificial tiene sus fundamentos en la estructura del pensamiento racional. Utilizando silogismos y las premisas correctas se forman los argumentos estructurales que permiten a los programas arribar a las conclusiones acertadas. Esta mirada presenta dos limitaciones importantes, una es que no permite incorporar conocimiento informal para resolver un problema y en segundo lugar es que la teoría que se puede presentar para resolver un problema complejo dista mucho de la práctica y de la realidad. De todas formas, siempre será necesario introducir una lógica de explicación o razonamiento de la solución si se pretende construir programas o sistemas inteligentes que puedan considerar una cantidad importante de hechos y variables a ser resueltos.

### **Actuar Racionalmente**

El siguiente abordaje pone en el centro al agente o programa de la computadora que es capaz de no solo realizar la tarea para la cual fue programado sino también poder operar de manera autónoma, poder percibir información de su ambiente, adaptarse al cambio, perdurar en el tiempo y crear y lograr metas. Un agente racional es aquel que puede cumplir con las reglas anteriores y también lograr el mejor resultado y/o ante escenarios de incertidumbre poder obtener dentro de las expectativas el mejor resultado posible.

Esta última mirada de agente inteligente es la que junto al auge de internet a permitido embeber actualmente en muchas paginas webs programas de agentes específicos que resuelven con inteligencia artificial problemáticas puntuales como por ejemplo los motores de búsqueda, los sistemas de recomendación y las agregaciones que puede tener un sitio web.

A pesar de los avances puntuales en los sub campos de la inteligencia artificial como jugar y ganar partidas de ajedrez o juegos más complejos como el tradicional juego de mesa chino “Go”, manejar vehículos autónomos o reconocer conversaciones en ciertos idiomas y poder traducirlos a otros, los principales fundadores e investigadores de la inteligencia artificial entienden que la misma debería volver a poner foco en sus raíces y canalizar los esfuerzos hacia una visión mas holística donde las máquinas puedan pensar, aprender y crear (Russell & Norvig, 2016).

El término Inteligencia Artificial General (AGI por sus siglas en inglés) se consolida detrás de un grupo de investigadores, los cuales desarrollan sus trabajos detrás de una visión más cercana

a los orígenes de la práctica de la inteligencia artificial. La misma tiene como objetivo el estudio y la construcción de software y/o hardware que pueda contar con una inteligencia general comparable o en algunos casos superior a la de un ser humano (Goertzel, 2014).

Si bien no existen una definición consensuada acerca “Inteligencia Artificial General” si existe una lista de características que la define. Entre ellas las principales son:

- AGI involucra la habilidad de lograr una variedad de metas u objetivos, llevar adelante la ejecución de una variedad de tareas, en una variedad diferente de contextos y situaciones.
- AGI debe ser capaz de manejar nuevos problemas y situaciones diferentes de aquellos a los cuales fue desarrollada por su creador.
- AGI debe ser capaz de incorporar nuevo conocimiento, generalizarlo y aplicar el conocimiento transfiriéndolo de un problema o contexto a uno nuevo.

Ray Kurzweil en su libro “The Singularity Is Near: When Humans Transcend Biology” introduce el concepto de “Narrow AI”, el cual hace referencia a la idea de una versión de la inteligencia artificial angosta o reducida. El término “Narrow AI” se emplea para describir la inteligencia artificial que es utilizada para ejecutar funciones específicas donde anteriormente era necesario la utilización de la inteligencia humana, y la misma es realizada con igual o mejor resultado que un ser humano. Generalmente estas funciones son ejecutadas con mucha mayor velocidad que un ser humano y pueden considerar miles de variables de manera simultanea (Kurzweil, 2005).

La limitación que se ha encontrado en la formulación de “Narrow AI” es que en general, ante un cambio del contexto o reformulación del problema, la misma normalmente debe ser vuelta a programar o reentrenada para aceptar y resolver nuevos parámetros y variables, y de esa manera volver a conservar el mismo nivel de inteligencia (Goertzel, 2014). Es en esta característica diferencial donde radica la principal diferencia con la búsqueda de la inteligencia artificial general, donde la capacidad de poder aprender y transferir el conocimiento es la propiedad más importante de todas. (Taylor, Kuhlmann, & Stone, 2008)

#### 4.2.2 - Comunicación y Procesamiento del Lenguaje Natural

Una de las principales virtudes que separa a los seres humanos del resto de los animales es la capacidad de comunicarse confiablemente entre sí, dado un lenguaje en común que podríamos describir como un conjunto de signos que agrupándose de una determinada manera exponen un significado que podemos interpretar.

La RAE define el término de comunicación como la transmisión de señales mediante un código común entre el emisor y el receptor. Este código lo podemos interpretar como un lenguaje ya sea un idioma en particular, un lenguaje de señas y signos o mediante dibujos que envían un mensaje.

El procesamiento del lenguaje natural (NLP por sus siglas en inglés) permite a las máquinas encontrar un mecanismo para convertir el lenguaje humano en una representación formal donde se vuelve más sencillo de manipular y procesar (Collobert & Weston, 2008).

Los objetivos que persiguen los programas o los agentes de inteligencia artificial con el procesamiento del lenguaje natural es en primer lugar poder comunicarse con los seres humanos y en segundo lugar poder adquirir conocimiento del lenguaje escrito (Russell & Norvig, 2016).

Desde sus comienzos las investigaciones y avances en NLP se han focalizado en tareas como la traducción de idiomas a través de máquinas, recuperación de información, resúmenes de texto, preguntas y respuestas, extracción de información y más recientemente sobre minería de opiniones y análisis de sentimiento. Los avances se fueron logrando en primera medida en el campo de la sintaxis y gramática y luego en el análisis semántico de las sentencias.

En 1956 Noam Chomsky realizó una investigación por la cual pudo construir una teoría formal acerca de las estructuras lingüísticas. Introduce el concepto de reglas para la estructura de frases que permiten descomponer las oraciones en pequeños fragmentos y luego a través de los que Chomsky denomina reglas de transformación gramática o reglas generadoras de transformación gramática se logra reconstruir la oración en un nuevo formato gramatical. El ensayo permite reducir de manera significativa la inmensa complejidad del lenguaje actual a proporciones administrables que en el futuro serán vitales para el estudio del uso y entendimiento del lenguaje (Chomsky, 1956).



Uno de los primeros modelos exitosos en interpretar sintáctica, semántica y hasta en ciertos casos pragmáticamente expresiones de texto, fue el introducido por Barwise conocido como “First-Order Logic” (FOL por sus siglas en inglés). Las funciones de sintaxis especificaban la manera en la cual el grupo de símbolos debía ser ordenado para ser considerado apropiado para su interpretación, las funciones de semántica suponían un significado sobre los textos previamente ordenados y las funciones pragmáticas especificaban como la información contextual podía aumentar la correlación entre las distintas funciones semánticas analizadas. Basado en la estructura de la lógica el sistema deductivo implementaba una estrategia donde se utilizaban axiomas y reglas de inferencia para interpretar variables cuantitativas y predicados complejos en las oraciones (Barwise, 1977).

El modelo de utilización de reglas e inferencias basados en la lógica exponía sus debilidades al momento de incorporar nueva información al conocimiento existente y de esa manera producir cambios en el pensamientos a través de una de las capacidades de la mente humana que es el razonamiento (Cambria & White, 2014).

Un abordaje empírico hacia el NLP sugiere que las extensivas y complicadas estructuras del lenguaje se pueden aprender a través de una especificación concreta de un modelo de lenguaje, al cual se deben asignar valores sobre las distintas variables del modelo, aplicando métodos estadísticos de reconocimiento de patrones y del aprendizaje automático a grandes cantidades de ejemplos de utilidades del lenguaje (Manning, Schütze, & Weikurn, 2002).

En el campo de la inteligencia artificial, el arte de la conversación entre las máquinas y los seres humanos es uno de los aspectos más complejos de lograr. Para que las máquinas puedan entender el lenguaje o código que se efectúa en una comunicación es necesario la aplicación de distintos modelos de lenguaje, como si fueran capas de construcción que se aplican, una detrás de la otra, para procesar la conversación. Los modelos son necesarios para lograr un conocimiento profundo acerca de la conversación que se intenta realizar. Se comienzan con modelos gramaticales sobre la estructura de las frases o sentencias que componen las oraciones, luego se agregan modelos de interpretación semántica y por último modelos de traducción y de reconocimiento de voz (Russell & Norvig, 2016).

Los modelos de NLP semánticos que generan las capas de interpretación de la información harían posible escalar del procesamiento al entendimiento del lenguaje natural (NLU por sus siglas en inglés). Para poder lograr esta capacidad de procesamiento de información de una manera

precisa y sensata los modelos computacionales deberán contemplar la habilidad de proyectar los modelos semánticos y sintácticos en el tiempo, compararlos de forma simultánea y de manera dinámica de acuerdo a los distintos contextos e intenciones que se presenten (Cambria & White, 2014).

### 4.2.3 - Reconocimiento de la Voz o el Habla

El reconocimiento de la voz o el habla es la tarea por la cual se identifican una secuencia de palabras pronunciadas mediante la generación de un sonido y una señal acústica (Russell & Norvig, 2016).

Claude Shannon en 1948 introduce el estudio del “Modelo del Canal Ruidoso” (Noisy Channel Model), el modelo ayuda a decodificar una señal de audio con mínimos errores, sin importar que tan ruidoso es el canal por el cual se transmite, y siempre que el mensaje original sea codificado de varias maneras redundantes antes de su transmisión. El modelo de canal ruidoso se ha estado aplicando a diferentes agentes de inteligencia artificial más allá del reconocimiento de voz, como en la traducción automática de idiomas y las correcciones de dictado y escritura.

El modelo acústico sostiene que las ondas de sonido se propagan en el aire gracias a los diferentes cambios de presión que se aplican al emitir el mismo. Cuando las ondas llegan al micrófono la señal se digitaliza y se almacena con distintas frecuencias de muestreo. Dado que sólo se necesita conocer las palabras que fueron pronunciadas y no todos los sonidos registrados en el audio, no es necesario tenerla en cuenta cierta información para identificar la palabra pronunciada. De todas formas, mientras mayor información se almacene por segundo de sonido emitido mayores datos podremos reproducir, como por ejemplo si la voz es de una mujer o un hombre, si es una persona mayor o un niño o si la persona habla en voz baja o con voz alta.

Lingüistas han podido identificar aproximadamente 100 distintos fonemas posibles, los cuales pueden componer la mayoría de las palabras de un idioma. Un fonema es la articulación mínima de sonido posible para una consonante o una vocal de un idioma. En su gran mayoría los sistemas de reconocimiento de voz dividen las unidades de archivo de audio en varios fragmentos denominados cuadros (frames en inglés). Esos cuadros tienen una duración máxima de no más de 10 milisegundos, esto le permite aproximadamente identificar cada fonema en un cuadro distinto.

Luego los cuadros son superpuestos ordenadamente para asegurar que ningún sonido o señal se pierda (Russell & Norvig, 2016).

En 1913 el matemático A. Markov introduce un ensayo acerca de las Cadenas de Markov (MC en inglés). El modelo estadístico representa la asignación de una probabilidad de ocurrencia de un evento sobre variables que no se conocen o son aleatorias en base a variables provenientes de un evento anterior (Poritz, 1988).

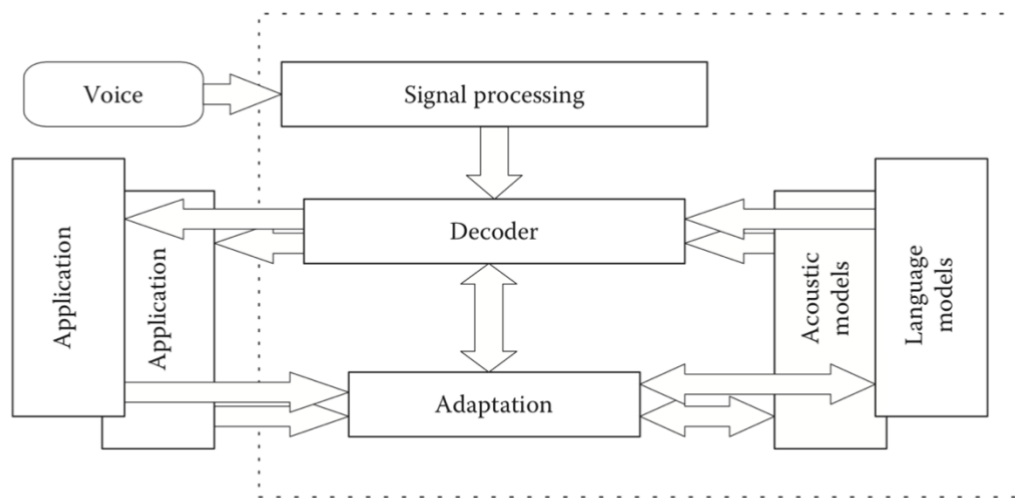
En 1970 un grupo de investigadores perfeccionan el modelo matemático de las MC introduciendo una técnica de iteración que logra efectividad al momento de maximizar el éxito de una función de probabilidad sobre el modelo estadístico de las CM (Baum, Petrie, Soules, & Weiss, 1970). E. Baum junto a varios colegas logra introducir el concepto de los Modelos Ocultos de Markov (HMM por sus siglas en inglés) como una extensión y superación al modelo estadístico de las Cadenas de Markov. El nuevo modelo logra un método eficiente que maximiza el éxito con una mejora en la estimación de los parámetros del modelo.

Los HMM son modelos estadísticos temporarios donde los parámetros para poder generar el modelo son inciertos o desconocidos. Los HMM son utilizados en los sistemas de reconocimiento de voz ya que proveen una teoría y un método que ayuda a evaluar la probabilidad de que una cierta secuencia de eventos pueda ocurrir. Los HMM empleados en el reconocimiento de la voz permiten en primer lugar codificar la probabilidad de que un patrón específico de sonidos sean encontrado para cada fonema, en segundo lugar entender como los fonemas influyen en otros y tercer lugar estimar la posibilidad de ocurrencia de aparición de otros fonemas en base al evento anterior (Kurzweil, 2005).

Los HMM han comenzado a utilizarse con éxito en el reconocimiento de voz con más frecuencia desde 1970 (Singh, Kapur, & Kaur, 2012). Los mismos contienen dos características muy relevantes. En primer lugar los HMM están basados en rigurosas teorías matemáticas, lo cual ha permitido a investigadores del reconocimiento de voz trabajar sobre la construcción de resultados matemáticos establecidos en otros campos y especialidades científicas como por ejemplo la investigación de la secuencia del genoma humano y en segundo lugar los modelos deben ser generados en base a un exhaustivo proceso de entrenamiento basados en cuerpos o conjuntos amplios y estructurados de datos reales, lo cual asegura ejecución sustentable y genera resultados que pueden ser susceptibles de mejora mediante un riguroso plan de entrenamiento. (Russell & Norvig, 2016).

Matemáticamente, el modelo para ejecutar un proceso HMM consta de un algoritmo de aprendizaje, el cual aplica un teorema de maximización de expectativas (EM por sus siglas en inglés) gracias a la posibilidad de iterar constantemente distintas combinaciones para encontrar la secuencia de parámetros que maximiza la probabilidad de reconocimiento (Xuedong Huang, Baker, & Reddy, 2014).

Una arquitectura de sistemas estándar de alto nivel de reconocimiento de voz consta de componentes básicos. En primer lugar, las aplicaciones, que son las que interactúan con el decodificador para obtener resultados del reconocimiento de la voz y además pueden utilizarse para adaptar otros componentes en la arquitectura del sistema. En segundo lugar, tenemos los modelos acústicos que incluyen la representación del conocimiento sobre acústica, fonética, variabilidad del micrófono y del entorno, diferencias de género y dialecto entre hablantes, etc. Por último los modelos de lenguaje que se refieren al conocimiento de un ecosistema de lo que constituye una palabra posible, qué palabras pueden coexistir y en que secuencia.



**FIGURE 15.2** Basic system architecture of a speech-recognition system.

*Figura 8. Diagrama de arquitectura referencial para el reconocimiento de voz. Fuente (Indurkha & Damerou, 2010)*

La semántica y las funciones relacionadas con una operación que un usuario desee realizar también pueden ser necesarias para el modelo del lenguaje. Actualmente existe mucha incertidumbre en áreas asociadas con las características del hablante, el estilo y la velocidad del

habla, el reconocimiento de segmentos básicos del habla, palabras posibles, palabras probables, palabras desconocidas, variación gramatical, interferencia de ruido, acentos no nativos y la puntuación de confianza de los resultados obtenidos (Indurkha & Damerau, 2010).



## 5. Capítulo V – Historia y actualidad de los asistentes virtuales inteligentes.

En el capítulo anterior hemos descripto los conceptos y elementos principales que integran los asistentes virtuales inteligentes. Por otro lado, se han descripto las motivaciones entorno a la ciencia y a como la investigación del campo de estudio de la inteligencia artificial ha podido diseñar soluciones que hoy, gracias al avance de la tecnología y la reducción de los costos en la producción de sus componentes, logran la posibilidad de que los seres humanos puedan interactuar con una máquina.

La tecnología de reconocimiento de voz ha recorrido un largo camino desde el lanzamiento del asistente virtual inteligente Siri perteneciente a la compañía Apple Inc. en 2011.

Los estudios realizados en el campo del aprendizaje automático, han logrado progresos considerables tanto en el reconocimiento de voz (transcripción de audio a texto) como en el procesamiento del lenguaje natural (identificación del objeto y del sentido dentro del texto). De hecho, los IVA han superado la tasa de precisión de reconocimiento de palabras por encima del 95%, lo que significa que su capacidad para comprender el lenguaje es prácticamente equivalente a la de los seres humanos.

En el desarrollo de este capítulo describiremos la evolución que han tenido los asistentes virtuales inteligentes a lo largo de la historia, analizaremos los principales actores del mercado actual y las características de sus plataformas de inteligencia artificial que en la actualidad han logrado un intenso crecimiento de sus ventas y de su utilización en los distintos mercados y hogares del planeta.

### 5.1 – Historia y evolución de los asistentes virtuales inteligentes.

Alan Turing, en 1950 con su exitoso ensayo "Maquinaria de computación e inteligencia", sugiere que en lugar de preguntar si las máquinas pueden pensar, deberíamos preguntar si las máquinas pueden pasar una prueba de inteligencia conductual, que actualmente se conoce como la prueba de Turing (Turing Test) (Turing, 1950).

La prueba establece por primera vez la consideración acerca de la construcción de un agente virtual mediante un programa informático que sea capaz de mantener una conversación (a través de mensajes escritos en línea) con un ser humano, él cual debe interrogar a la máquina

durante cinco minutos. El interrogador luego debe adivinar si la conversación realizada fue efectuada con un programa o con un ser humano. El programa pasa la prueba si engaña al interrogador el 30% del tiempo. Turing conjeturó que, para el año 2000, una computadora podría programarse lo suficientemente bien como para pasar la prueba. Estaba equivocado: los programas aún no han engañado a un juez sofisticado (Russell & Norvig, 2016).

En 1966 J. Weizenbaum crea un programa llamado ELIZA que es alojado y ejecutado en hardware perteneciente al MIT (Massachusetts Institute of Technology). El programa era capaz de ejecutar conversaciones con seres humanos utilizando funciones de conversación del lenguaje natural. ELIZA podía leer y entender las oraciones y sentencias escritas por los usuarios y responder en la misma terminal.

Simplificando conceptualmente el proceso, el mismo comienza con la entrada de un texto o una oración escrita por el usuario en la terminal, luego la oración es examinada por el programa para encontrar palabras claves en base a un esquema de ponderación y ranking establecido, a continuación la oración es transformada utilizando las reglas programadas para las palabras claves encontradas y se forma un guion que será el hilo conductor de la conversación, en el supuesto caso de no encontrar palabras claves, el programa devuelve una respuesta asociada a un contexto libre con características determinadas por la oración previamente transformada.

La particularidad que constituía el sentido de la programación de ELIZA era en imitar lo que podría ser una conversación con un psicoterapeuta, se eligió este estilo de diálogo por sus características donde mayormente el que toma la iniciativa es el paciente, en este caso interpretado por el usuario y donde el profesional tiene la libertad de desconocer el contexto del mundo real donde el paciente habita, en estas circunstancias ELIZA lograba los mejores resultados en la interacción del diálogo (Weizenbaum, 1966).

Lo interesante del experimento llevado adelante por J. Weizenbaum con ELIZA fue el involucramiento del equipo de investigación con la interacción del programa. Aún sabiendo que Eliza era una pieza de software las personas pertenecientes al equipo intentaban llevar adelante conversaciones como si en lugar de una máquina hubiese un ser humano con el cual podían intercambiar y relevar aspectos íntimos durante una conversación (Treusch, 2018).

Entre 1968 y 1970 T. Winograd desarrolla un programa en el MIT llamado SHRDLU, el cual tiene como objetivo poder ejecutar un diálogo teniendo un entendimiento sobre la materia que esta en discusión. El programa es cargado con un detallado modelo sobre una pequeña parte de

conocimiento de un ámbito específico, lo cual permite al software comprender oraciones relacionadas sobre la materia para la cual fue instruido previamente. Utilizando información semántica y del contexto puede comprender el sentido del diálogo y desambiguar sentencias para contestar preguntas, aceptar nueva información y ejecutar comandos (Winograd, 1971).

En 1975 K. Colby crea un programa llamado PARRY que a diferencia de ELIZA intenta imitar el comportamiento de una persona paranoica con respuestas de características hostiles y defensivas en los diálogos. El modelo de la simulación de comportamiento del programa consistía de dos partes, un bloque de análisis o deconstrucción de la oración que se utilizaba como entrada de información y una segunda instancia correspondiente a un modulo de acción-interpretativa compuesto por estructuras de datos y reglas de producción de respuestas para el diálogo, que evaluaban el estado actual de la entrevista y decidían que acciones (lingüísticas) realizar para cumplir las intenciones del modelo (Colby, 1981).

Tanto ELIZA como PARRY fueron experimentos diseñados en inteligencia artificial para imitar el comportamiento humano y por lo tanto con el objetivo de pasar el Test de Turing, los resultados de la investigación con estos programas demostraron que las interacciones hombre / máquina no solo eran posibles, sino que merecían más atención. Una de las vías que los investigadores comenzaron a explorar fue el desarrollo del software de entrevistas y exámenes asistidos por computadora (CAI y CAT por sus siglas en ingles) (Epstein & Klinkenberg, 2001).

Las computadoras con programas que ejecutaban diálogos en un formato de pregunta y respuesta debían lidiar con el desconocimiento del contexto en el cual se producía el intercambio. Una limitación básica de estos programas era intentar administrar el problema del significado con el entendimiento de una sola oración cuando en realidad no existe la posibilidad de que la utilización del uso humano del lenguaje se produzca en un ámbito tan artificial. Normalmente somos parte de un contexto y ese contexto nos ayuda a generar las interpretaciones y la heurística en un proceso de comunicación (Winograd, 1973).

En la misma dirección que T. Winograd y R. Shank ampliando el concepto cognitivo acerca de la construcción del conocimiento crean un programa que simula la habilidad humana de comprender historias. El mismo tiene la capacidad de responder preguntas acerca de una historia aún cuando la información suministrada en la pregunta nunca fue ingresada como dato concreto durante el relato de la historia en cuestión (Searle, 1980).



En 1972 R. Shank introduce la teoría de la dependencia conceptual (Conceptual Dependency en inglés) en la cual describe la investigación de un método para obtener la capacidad del entendimiento humano del lenguaje natural para las máquinas (R. C. Schank, 1972). Shank sostiene un abordaje cognitivo hacia la problemática de construir máquinas inteligentes y entiende que no existe una manera adecuada de desarrollar capacidades de “comprensión o entendimiento” en computadoras, sin proveer un método en el cual se incorpore información acerca del contexto y el problema con el cual la máquina debería trabajar en la elaboración de su interacción (R. Schank & Abelson, 1977).

En 1976 el proyecto de investigación avanzada para la agencia de defensa de los Estados Unidos (DARPA por sus siglas en inglés) comienza el camino de una serie de investigaciones sobre el entendimiento del habla (SUR por sus siglas en inglés). El grupo de desarrolladores pudo construir una serie de sistemas de reconocimiento del habla que permitían entre sus principales funciones controlar robots a través de la voz, reconocimiento de la voz sobre vocabularios extensos, reconocimiento del habla dependiente de los oradores y dictado de vocabularios prácticamente sin restricciones (Reddy, 1976) .

El programa de investigación DARPA ha logrado desarrollar capacidades del reconocimiento del habla en distintos proyectos que se mantienen vigente en la actualidad. Sólo por nombrar algunos de los principales logros en el avance del reconocimiento de la voz podemos citar los siguientes ejemplos.

HEARSAY-1 fue uno de los primeros sistemas capaces de lograr el reconocimiento continuo de la voz y utilizar información semántica para reducir significativamente el número de alternativas consideradas en el resultado del reconocimiento. DRAGON fue uno de los primeros desarrollos que utilizó los modelos ocultos de Markov (HMM) para el procesamiento del discurso del habla. El sistema HARPY introdujo el concepto de búsqueda de grafos, donde el lenguaje de reconocimiento de voz se representa como una red conectada derivada de representaciones léxicas de palabras, con reglas de producción sintácticas y reglas de límites de palabras, el cual luego ha sido utilizado durante décadas como método exitoso para la búsqueda eficiente en memoria y traducción simultánea en máquinas.

Otros desarrollos del programa DARPA fueron SPHINX-I, construido en 1987, fue el primer sistema en demostrar reconocimiento de la voz independiente del orador. El sistema SPHINX-II, desarrollado en 1992, se benefició en gran medida del modelo de parámetros

vinculados para balancear el entrenamiento y la eficiencia de los algoritmos de búsqueda en la utilización de los métodos combinados de Gauss y de los modelos ocultos de Markov, el mismo alcanzó el nivel más alto de precisión en el reconocimiento de la voz según los indicadores de evaluación diseñados durante el programa DARPA en 1992, los cuales hoy todavía se mantienen vigentes para evaluar la precisión de los distintos desarrollos en el reconocimiento del habla.

El proceso de decodificación para encontrar la mejor secuencia de palabras (Variable X) que coincida con el discurso de la variable de entrada (Variable Z) es más complejo que un simple problema de reconocimiento de patrones, ya que uno enfrenta un número exponencial de patrones de palabras para buscar como coincidencia.

El desafío del proceso de decodificación en la operación del reconocimiento de la voz es encontrar una secuencia de palabras cuyos modelos acústicos y de lenguaje correspondientes coincidan mejor con la secuencia de vectores consideradas en la variable de entrada. Por lo tanto, dicho proceso de decodificación son modelos acústicos y de lenguaje entrenados frecuentemente y se conocen como procesos de búsqueda. Los algoritmos de búsqueda de gráficos, que previamente se han explorado en profundidad en los campos de la inteligencia artificial, la investigación de operaciones y la teoría de juegos, sirven como la base para la resolución del problema de búsqueda en el reconocimiento de voz (Xuedong Huang et al., 2014).

Durante la década de 1990 en adelante el programa DARPA concentró el foco de la investigación en el entendimiento natural del lenguaje y sus fronteras con el reconocimiento del habla. Más tarde, el énfasis se amplió a una gama de diferentes áreas de aplicación de comprensión del habla, junto con un nuevo enfoque en la transcripción de noticias durante su transmisión y el reconocimiento y entendimiento del habla conversacional (Anusuya & Katty, 2009).

En el 2005 se crea dentro del programa DARPA el proyecto EARS ( Effective, Affordable Reusable Speech to text por sus siglas en inglés) el cual logra desarrollar una tecnología de alta efectividad en la precisión de la transcripción y el reconocimiento del habla hacia un procesador de texto. El objetivo del proyecto era hacer posible que las máquinas hicieran un trabajo más preciso al momento de detectar, extraer, resumir y traducir información importante, permitiendo así a los seres humanos comprender lo que se decía leyendo transcripciones en lugar de escuchar las señales de audio (Anusuya & Katty, 2009).

La posibilidad de almacenar y procesar grandes cantidades de datos gracias a los desarrollos realizados en el campo de “Big Data” y a la capacidad de utilizar y poder incrementar

el poder de cómputo debido al avance de la computación en la nube han permitido desbloquear el valor de las proyecciones tecnológicas y generar avances considerables en el entrenamiento de los modelos acústicos y del lenguaje para el reconocimiento de la voz. (Xuedong Huang et al., 2014).

El avance logrado en el campo de las estructuras estadísticas del aprendizaje automático de máquinas (ML por sus siglas en inglés) en gran parte sostenido por la aparición de una infraestructura de procesamiento cada vez más poderosa y la posibilidad de contar con grandes volúmenes de datos para el entrenamiento de los algoritmos estadísticos constituyen las principales fuerzas conductoras en el avance de los desarrollos de los asistentes virtuales inteligentes (Xuedong Huang et al., 2014).

## 5.2 – Análisis de los principales asistentes virtuales inteligentes disponibles en el mercado.

Las plataformas tecnológicas Alexa de Amazon, Siri de Apple y Google Assistant de Google son los asistentes virtuales inteligentes disponibles y más utilizados en el mercado. Estas plataformas se encuadran y se sitúan académicamente según el dominio de estudio de la Interacción Humano-Computadora (HCI por sus siglas en inglés) dentro del campo de la Interface Natural del reconocimiento del habla del Usuario (NUI por sus siglas en inglés).

HCI junto con el estudio de la interacción entre humanos y agentes (HAI por sus siglas en inglés) forman parte de la incipiente área de investigación de comunicación entre humanos y computadoras (HMC por sus siglas en inglés) la cual tiene como objetivo de estudio comprender como las personas perciben y establecen su comunicación con la tecnología (Guzman, 2019).

Los NUI son sistemas de aplicaciones utilizados por los usuarios mediante acciones intuitivas relacionadas al comportamiento natural del ser humano y a través de instrucciones emitidas por medio de la voz o el habla (López, Quesada, & Guerrero, 2018).

Por primera los seres humanos pueden interactuar libremente con tecnología, la cual identifican como inteligencia artificial, y pueden establecer comunicaciones con la misma de maneras radicalmente distintas a cualquier otra interacción que hayan realizado con máquinas en el pasado (Guzman, 2019).

Una característica importante de la industria y el mercado al cual pertenecen los IVA, es que los servicios de asistentes virtuales inteligentes no tienen un precio de venta ni de suscripción

para su utilización. Todos los competidores actuales comercializan el altavoz inteligente con el servicio del IVA sin costo adicional por su uso.

A continuación, describiremos las principales características de los asistentes virtuales inteligentes que se comercializan en los distintos competidores disponibles a través de integrar las plataformas en los altavoces inteligentes.

### 5.2.1 Alexa - Amazon

El principal líder del mercado Amazon tiene el producto Echo como su dispositivo de parlante inteligente. Echo fue lanzado en el año 2015 y se comercializa en dos versiones, la más económica es Echo Dot a partir de U\$S 39.99 y Echo Plus desde U\$S 99.99 dólares estadounidenses.

Ambos productos tienen embebido el asistente virtual inteligente llamado Alexa. Cuando el software de Alexa es iniciado en el dispositivo, Amazon Echo deja de ser sólo un altavoz y comienza a operar como un asistente virtual inteligente y personal.

El asistente virtual inteligente Alexa, fue desarrollado por Amazon Lab126 en el 2014 (Carr, 2015). Es capaz de interactuar a través del reconocimiento de voz y ejecutar comandos como reproducción de música, hacer listas de tareas, configurar alarmas, transmitir podcasts, reproducir audiolibros y proporcionar información sobre el clima, el tráfico, los deportes y otra información en tiempo real, como noticias. Alexa también puede controlar varios dispositivos inteligentes utilizándose como un sistema de automatización del hogar (Amazon, 2019).

Alexa se ejecuta como un servicio en la nube, puede interactuar con varios dispositivos habilitados para Alexa, como Echo o el altavoz inteligente de Facebook Portal, también puede comunicarse con otros dispositivos inteligentes conectados a internet compatibles y aplicaciones de terceros al convertir las solicitudes de voz al protocolo de comunicación nativo de otros servicios.

Además, para personalizar estos entornos relacionados con Alexa, los usuarios pueden acceder al servicio en la nube utilizando clientes o terminales complementarios, como computadoras personales o dispositivos móviles (Android e iOS). Por lo tanto, el ecosistema creado por todos estos dispositivos interconectados, aplicaciones de terceros y clientes complementarios es complejo y heterogéneo dentro del servicio IVA de Amazon. Nos referiremos en adelante a este ecosistema como la plataforma Alexa.

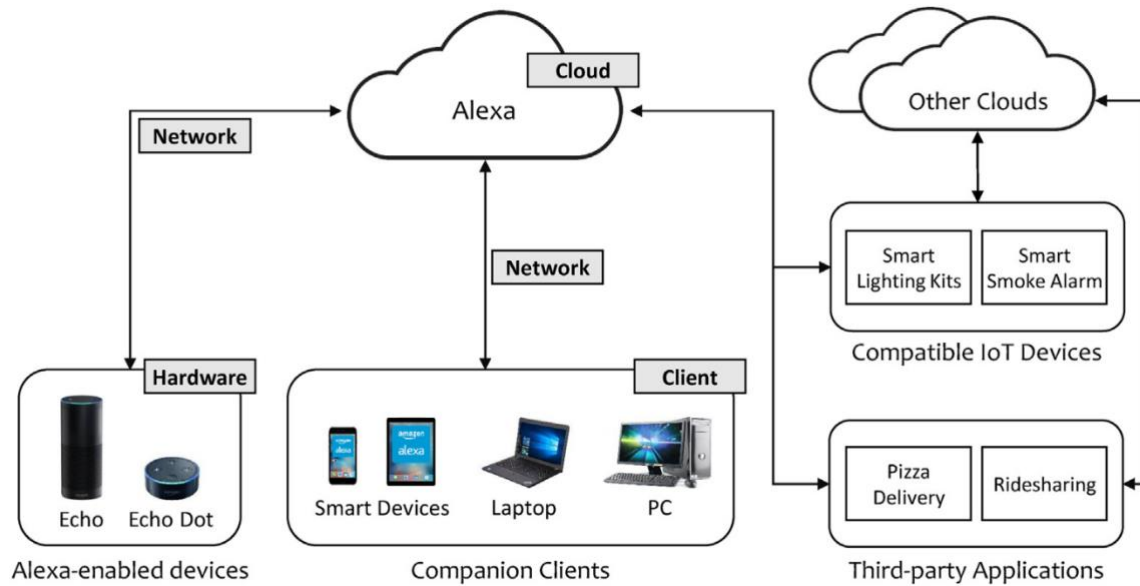


Figura 9. Arquitectura del ecosistema de interacción con la plataforma Alexa. Fuente (Chung, Park, & Lee, 2017)

La plataforma Alexa se encuentra abierta para programadores independientes o de otras empresas no pertenecientes a la firma, los cuales pueden crear y desarrollar nuevas funcionalidades o habilidades para la plataforma basándose en la programación de funciones y software de inteligencia artificial disponibles en la plataforma de Amazon Web Services (AWS por sus siglas en inglés). AWS es una vertical de la línea de negocios que Amazon comercializa como servicios de computación en la nube principalmente a otras empresas.

El intenso crecimiento de las funcionalidades de Alexa se apalanca sobre la base de la inversión efectuada por Amazon en AWS para sus productos de cómputo en la nube y de esa manera poder ofrecer sus servicios de IVA en los distintos dispositivos. La plataforma contiene funciones de programación específicas para los desarrollos de inteligencia artificial y aprendizaje automático de máquinas, los mismos también se encuentran disponibles para que terceros puedan seguir evolucionando los servicios y funciones de Alexa (Rouhiainen, 2018).

Un factor importante para mencionar acerca del crecimiento de la utilización de la plataforma son las distintas alianzas que Amazon ha llevado adelante para incorporar el software de Alexa en distintos fabricantes de parlantes de alta fidelidad como Bose, Sonos, Harman Kardon

y Ultimate Year. Otra alianza importante para Amazon fue la lograda con Facebook para incorporar el asistente virtual en los dispositivos fabricados por el líder de las redes sociales.

La plataforma de Alexa tiene ciertas características que predisponen a los usuarios a personificarlo e integrarlo en la vida social. En primer lugar, para operar el dispositivo, los usuarios deben interactuar con Alexa, el agente de conversación y reconocimiento del habla. Por lo tanto, el altavoz inteligente Echo se convierte inherentemente en un dispositivo socialmente interactivo, ya que requiere de una interacción social para funcionar. En segundo término, el servicio de Alexa se encuentra personificado, ya que el agente está dotado de un nombre, un género y una personalidad. Dichas opciones de diseño alientan a los usuarios a despertar sentimientos humanos sobre el dispositivo (Purinton, Taft, Sannon, Bazarova, & Taylor, 2017).

El antropomorfismo es un mecanismo inconsciente en el que las personas infieren que una entidad no humana tiene características similares a la del ser humano y por lo tanto la persona garantiza al objeto un tratamiento similar al ser humano (Nass & Moon, 2000). La plataforma Alexa se encuentra programada con la capacidad de interactuar de una manera lúdica, generando una comunicación más atractiva e inteligente y cuenta con el poder de alterar la dinámica del entorno que la rodea (Purinton et al., 2017).

En cuanto a las principales características y funcionalidades que podemos destacar del sitio web de la plataforma son las siguientes:

### **Inteligencia Basada en la Nube**

Alexa cuenta con la característica de poder evolucionar su inteligencia con nuevas capacidades a través del aprendizaje automático y el poder de procesamiento escalable en la nube. Las tecnologías complejas como el reconocimiento de voz automático y la comprensión del lenguaje natural se administran a través de servicios alojados en la nube de Amazon.

### **Facilidad en la Programación de Interfaz de Aplicaciones**

Alexa cuenta dentro de su plataforma con espacios de ayuda para el desarrollo de interfaces sencillas para funciones básicas de servicios como reproducción de audio, control de volumen o texto a voz. Estas interfaces tienen la posibilidad de introducirse como librerías de programación y ser fácilmente reutilizadas para el desarrollo de capacidades incorporadas dentro de Alexa y las habilidades o funciones construidas por terceros.

### **Control de Hogares Inteligentes**

Alexa en su rol de concentrador de hogares inteligentes incorpora un conjunto de comandos que permiten el control a través de la voz para dispositivos inteligentes en hogares, como luces, interruptores, termostatos y cerraduras. Estableciendo alianzas con los principales fabricas diseñadoras de estos dispositivos como Belkin Wemo, Philips Hue, SmartThings, Insteon y Wink entre otros.

### **Reproducción de Audio**

La plataforma permite a través del reconocimiento de voz vincularse para la reproducción de servicios de transmisión como Amazon Music, SiriusXM, iHeartRadio y TuneIn, como así también reproducir resultados de búsqueda en la web, resultados deportivos, informes meteorológicos, chistes y alertas informativas.

### **Localizaciones**

Los servicios de la plataforma Alexa proporcionan lenguajes y habilidades localizadas en un número creciente de regiones internacionales.

### Fortalezas

- Precios competitivos.
- Plataforma abierta a nuevos desarrollos. Más de 80.000 skills o habilidades desarrolladas y en crecimiento dentro de la plataforma Alexa a nivel global.
- Integración completa con el ecosistema de Amazon y su portal de comercio electrónico para realizar compras a través del dispositivo.
- La aplicación Alexa de Amazon puede ser descargada tanto del Apple Store como de Google Play en otros dispositivos como tabletas y celulares.

### Debilidades

- Privacidad sobre los datos que la plataforma almacena.
- Control de calidad sobre la disponibilidad de las habilidades y funciones desarrolladas para Alexa.

## 5.2.2 Google Assistant - Google

Google en el año 2019 ha logrado equiparar a Amazon en cuanto a las ventas de altavoces inteligentes, el líder mundial de búsquedas en internet comercializa tres versiones de producto, la primera se denomina Google Home Mini y se puede adquirir desde U\$S 49, la opción intermedia se trata de Google Home con precios desde U\$S 99 y por último Google Home Max, la opción más costosa que se comercializa desde U\$S 299 dólares estadounidenses.

Google es una de las compañías basada en publicidad y datos más grandes del mundo, posee datos de millones de usuarios obtenidos a través de búsquedas de Google.com, YouTube, Gmail y otros productos y servicios propiedad de Google. Tener acceso a una cantidad tan enorme de datos ha proporcionado a Google la ventaja perfecta para convertirse en uno de los líderes mundiales en inteligencia artificial (Rouhiainen, 2018).

La inteligencia artificial es una de las prioridades clave para Google y su empresa matriz, Alphabet Inc. Google es considerada una de las compañías de inteligencia artificial más avanzadas, ya que casi todos sus nuevos productos y servicios aplican tecnología de inteligencia artificial de alguna manera (Rouhiainen, 2018).



Al igual que Amazon, Google ha ido detrás de alianzas con fabricantes de altavoces, pero en este caso orientadas a empresas grandes de tecnología como Sony, Lenovo o LG y no tanto a empresas de nicho dedicadas a la producción de altavoces de alta fidelidad.

Para brindar una experiencia consistente también fuera del hogar la plataforma de Google Assistant puede ser descargada y utilizada en celulares y tabletas, en estos dispositivos el asistente digital logra una mejor experiencia de usuario si es utilizado bajo el sistema operativo propietario de Alphabet Inc. llamado Android (Lynley, 2016).

La plataforma de Google Assistant permite integrar varios servicios internos y de terceros en Google Home. Estos servicios permiten a los usuarios interactuar a través de comandos de voz para controlar la interacción con ellos. Algunos ejemplos de servicios compatibles con Google Assistant incluyen Google Play Music, Spotify e iHeartRadio para audio, Netflix, YouTube y Google Photos para videos y fotos, Google Calendar y Google Keep para tareas, y CNN, CNBC, BBC y The Wall Street Journal para actualizaciones de noticias (Westenberg, 2019).

A diferencia de su predecesor, Google Now, Assistant puede entablar conversaciones bidireccionales con los usuarios. En octubre de 2016 Google contrató a escritores de los estudios cinematográficos de películas Pixar y del periódico satírico estadounidense The Onion para desarrollar una personalidad para el IVA, con el objetivo a largo plazo de poder invocar una sensación de relación emocional hacia los usuarios (Hartmans, 2016).

Los usuarios pueden conectar y agrupar múltiples altavoces dentro del hogar para la reproducción sincronizada de música en distintas habitaciones (Google, 2019).

Google Home tiene la capacidad, a través de su micrófono, de poder reconocer hasta seis voces diferentes para un usuario, obteniendo al mismo tiempo información de sus diferentes cuentas dependiendo de quién este hablando e interactuando con el dispositivo. También permite tener integración con el calendario de cada usuario, para agregar recordatorios y establecer citas verbalmente (Lardinois, 2017).

Google Home incluye funciones de automatización del hogar, lo que permite a los propietarios usarlo como un concentrador para controlar dispositivos inteligentes dentro del hogar. Los ejemplos de dispositivos compatibles incluyen el reproductor de medios digitales Chromecast y productos de Nest, SmartThings, Philips Hue, LIFX, LightwaveRF y Logitech Harmony entre otros (Lagace, 2019).

En febrero de 2017, Google anunció que los usuarios de Google Home podrían comprar productos a través del reconocimiento de voz por medio de su servicio de compras Google Express, con productos disponibles en Whole Foods Market, Costco, Walgreens, PetSmart y Bed Bath & Beyond (Steele, 2017).

En mayo de 2018, Google lanzó Duplex, una extensión de Google Assistant que permite llevar a cabo conversaciones naturales imitando la voz humana (Google, 2018). El asistente puede completar tareas de forma autónoma, como llamar a una peluquería para reservar una cita, programar una reserva de restaurante o llamar a las empresas para verificar el horario de las tiendas de vacaciones (Nieva, 2018). Duplex fue creado para poder desarrollar una capacidad de habla con una voz y un lenguaje más naturales al incorporar las deficiencias del habla, como los sonidos que se utilizan de relleno durante una conversación, y el uso de frases comunes dentro de un lenguaje e idioma propio, junto con una entonación más humana e incorporando latencia en la respuesta del diálogo (Yaniv, 2018).

A mediados de 2019, Google Assistant relanza una plataforma de desarrollo para la construcción de habilidades y funcionalidades por partes de terceros programadores y comienza a abrir el ecosistema a nuevos programadores para que puedan diseñar e integrar aplicaciones al ecosistema de Google Home (Lardinois, 2019).

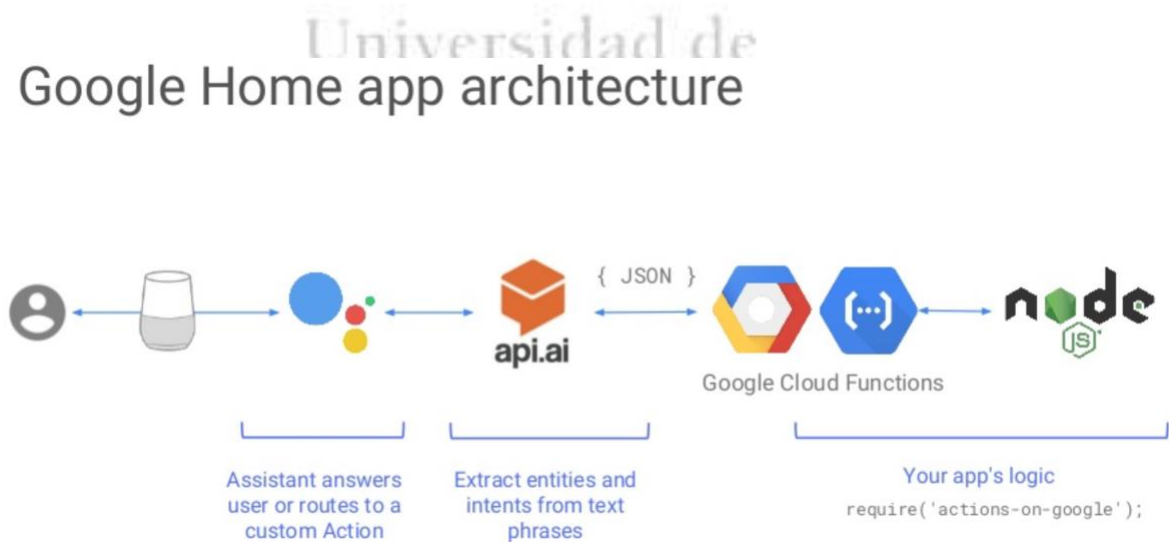


Figura 10. Diagrama de arquitectura de integración de Google Assistant. Fuente (Bret McGowen, Developer Advocate at Google).

En cuanto a las principales características y funcionalidades que podemos destacar del sitio web de la plataforma se encuentran:

### **Información y Respuestas**

La plataforma de Google Assistant permite obtener respuestas relacionadas con deportes, clima, finanzas, cálculos, traducciones y recetas de cocina entre otros.

### **Entretenimiento**

A través del comando por voz es posible reproducir música, podcasts y estaciones de radio, o transmitir videos en televisores que utilicen el dispositivo Google Chromecast.

### **Planificación de Actividades**

Google Assistant a través de la integración con Google Calendar permite recibir ayuda con el resumen del día, la información de vuelos o viajes cotidianos al trabajo. También permite acceder a las novedades del clima y el tráfico del área por la cual se encontrará el usuario.

### **Administrar Tareas**

Google Assistant puede agregar artículos a la lista de compras. También es posible establecer alarmas, administrar el calendario y hacer llamadas telefónicas.

### **Control del Hogar**

Google Assistant permite establecer la temperatura del hogar a través de Nest, o apagar las luces con la integración realizada con los dispositivos Philips Hue. También permite activar o desactivar los dispositivos del hogar con interruptores inteligentes compatibles a la plataforma.

Fortalezas

- Precios competitivos y una opción superior para un segmento destacado interesado en altavoces que reproducen audio de alta fidelidad.
- Asertividad en el proceso de reconocimiento de voz y de interpretación del lenguaje natural.
- Gran base de usuarios Android y de utilitarios como Gmail, Google Drive y Google Calendar para integrar las principales habilidades del asistente virtual.

#### Debilidades

- Plataforma en la actualidad abierta para el desarrollo de habilidades y funciones de Google Assistant, pero con bajo índice de crecimiento.
- Privacidad de la información. En este punto Google se encuentra aún un poco más comprometido dado el acceso a la identificación del usuario vía el sistema operativo Android y la plataforma de Google Chrome.

### 5.2.3 Siri - Apple

El asistente virtual de Apple llamado Siri fue lanzado al mercado en el 2011 junto con el lanzamiento del iPhone 4, luego de que Apple Inc. adquiriese la empresa que construyó el desarrollo del software llamada SRI International (Bostic, 2013).

El asistente virtual Siri nació como un desprendimiento del programa del Centro Internacional de Inteligencia Artificial SRI, y es una rama del proyecto “Asistente Cognitivo que Aprende y Organiza” (CALO por sus siglas en inglés), financiado por la Agencia de Proyectos de Investigación Avanzada de Defensa de los Estados Unidos (DARPA) (Markoff, 2008).

El proyecto combinó enfoques de inteligencia artificial tradicionalmente aislados para tratar de crear un programa de asistente personal que mejore el interactuar con los usuarios. En principio, Siri fue diseñado como un software de asistente personal, que ayudaría a los usuarios a completar tareas en lugar de solo recopilar información (Naone, 2009).

Posteriormente a ser lanzado en el 2011, el asistente inteligente Siri fue evolucionando primero con la adaptación a varios idiomas y luego al ser incluido a través del sistema operativo iOS en todos los dispositivos y plataformas disponibles de Apple en la presentación de los nuevos productos de la empresa en el año 2016 (Statt, 2016).

En febrero del 2018 Apple Inc. Lanzó Homepod, el altavoz inteligente de Apple con un foco orientado más hacia a la calidad del sonido que hacia la inteligencia artificial embebida de Siri dentro del altavoz (Gartenberg, 2017).

El precio de Homepod es de U\$S 299 dólares estadounidenses, es un precio relativamente alto comparado con sus principales competidores, los principales fundamentos para la disposición del precio a pagar por el producto está dado por la alta calidad del sonido que el dispositivo puede brindar y la inteligencia embebida que tiene para poder entender las dimensiones del lugar donde se encuentra y ajustar la intensidad del volumen del audio para lograr un sonido uniforme (Rouhiainen, 2018).

Si bien las aplicaciones desarrolladas para el ecosistema de Apple bajo el sistema operativo iOS permiten invocar los servicios de Siri a través de un conjunto de funciones disponibles en SiriKit, la plataforma de asistente inteligente virtual de Apple no se encuentra abierta para que terceros programadores puedan desarrollar nuevas habilidades o funcionalidades para Siri.

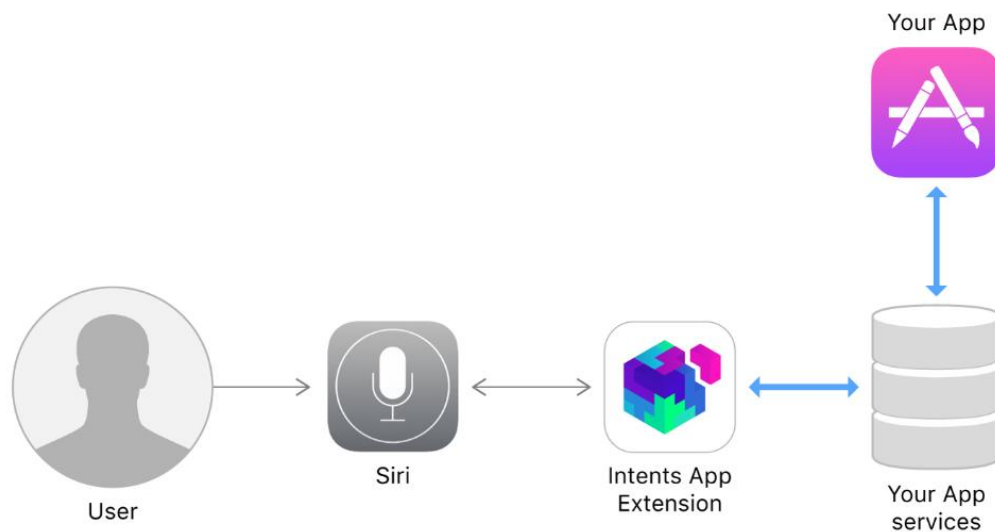


Figura 11. Diagrama de arquitectura de integración para aplicaciones iOS con SiriKit.  
Fuente ( <https://developer.apple.com/documentation/sirikit> at Apple.com).

En cuanto a las principales características y funcionalidades que podemos destacar del sitio web de la plataforma se encuentran:

### **Llamadas y Mensajes de Texto**

La plataforma de asistencia inteligente virtual Siri puede hacer llamadas o enviar mensajes de texto a través del reconocimiento de voz. También puede reproducir mensajes en otros dispositivos conectados. Cuenta con la posibilidad de ofrecer sugerencias proactivas, como enviar mensajes de texto a los contactos.

### **Tareas Cotidianas**

Siri permite establecer alarmas, temporizadores, recordatorios y obtener las direcciones de las reuniones agendadas en su calendario. Según la rutina establecida en el calendario, Siri cuenta con la posibilidad de anticipar lo que podría necesitar para el resto del día.

### **Música**

Apple Music y Siri trabajan en forma conjunta. Es posible solicitarle a Siri que encuentre una nueva canción en la biblioteca de música de Apple Music. Siri puede responder preguntas sobre las bandas favoritas y también puede sugerir su lista de reproducción cuando el usuario llega a distintas locaciones como por ejemplo el gimnasio o comience un viaje de regreso a su casa.

### **Hogar Inteligente**

Siri permite controlar los dispositivos inteligentes del hogar, verificar su estado o incluso ejecutar comandos en simultáneo, utilizando sólo el reconocimiento de voz.

### **Conocimiento**

Siri permite verificar rápidamente hechos, hacer cálculos o traducir frases a otro idioma. Siri trabaja detrás de escena como un asistente personal. La plataforma de voz utiliza tecnología neuronal avanzada de texto a voz (Text-to-Speech) para sonar natural, especialmente al emitir frases más largas, como leer las noticias o responder preguntas de conocimiento general.

### **Sonido Inteligente**

El altavoz inteligente HomePod combina tecnología de audio personalizada diseñada por Apple y software avanzado para ofrecer un sonido de precisión que llena la habitación.

#### Fortalezas

- Integración con los dispositivos y servicios de Apple (Apple Music & Podcasts) y Siri para poder ejecutar funciones y comandos con la voz y en sentido inverso poder tomar notas o establecer recordatorios en el teléfono.
- Alta fidelidad e inteligencia para distribuir el sonido.
- Resguardo y privacidad de la información.

#### Debilidades

- Actualmente Siri no es una plataforma abierta que tenga como objetivo o visión de largo plazo generar nuevos desarrollos o habilidades para el autoaprendizaje e interacción de nuevos servicios a ser consumidos a través de la voz o el diálogo.
- Habilidad o funcionalidades limitadas al ecosistema Apple. Por Ej. no es posible ejecutar comandos de voz para reproducir música fuera de la aplicación de Apple Music.
- No es posible integrar el parlante con otros Homepods dentro del hogar.
- Precio superior al de sus competidores.

### **5.2.4 Facebook & Microsoft**

La investigación realizada para el presente estudio incluye dentro de su análisis los productos de altavoz inteligente de las compañías Amazon, Google y Apple, debido principalmente a su participación de mercado y en segundo término a raíz de poseer una estrategia definida acerca del desarrollo y evolución del producto y de su respectiva plataforma IVA.

Facebook ha intentado progresar en este mercado utilizando su asistente virtual “M” pero no ha logrado tener pruebas satisfactorias. Es por ello que la decisión estratégica de la empresa fue embeber el IVA de Amazon en su dispositivo de interacción Portal y orientar el mismo hacia el uso de las aplicaciones de comunicación como Facebook Messenger, utilizando un agregado de pantalla táctil para poder realizar video conferencia.

En el caso de Microsoft, si bien las pruebas de exactitud en la interpretación del lenguaje natural realizada todos los años por la firma de fondos de inversión de riesgo, Venture Capital Loup, arrojaron resultados aceptables para su plataforma IVA denominada Cortana, Microsoft desistió de competir en la comercialización de altavoces inteligentes luego de haber realizado una alianza con el fabricante de altavoces *Harman Kardon* para distribuir el asistente virtual en los altavoces de alta fidelidad llamado “Invoke”. Al momento de realizar el trabajo de campo del presente estudio se desconocen los fundamentos de la decisión de no competir por el momento en el sector.

### 5.2.5 Resumen

En cuanto a las descripciones de los productos comparados no se encuentran amplias diferencias, salvo por la estrategia de producto diferenciada que pretende establecer Apple con su dispositivo Homepod. Apple intenta establecer una visión más amplia que la de ofrecer un altavoz inteligente con mejor sonido. La compañía está intentando generar de Siri una presencia ubicua que pueda conectar y controlar todos sus dispositivos y servicios conectados, utilizando como principal beneficio la poca fricción al momento de la interoperabilidad que hoy existe entre sus diversas líneas de productos.

En cuanto a las estrategias de desarrollo de producto de Google y Amazon ambos siguen caminos similares. Según la dimensión de utilidad que se desea evaluar es donde emergen ciertas características que pueden ser más ventajosas en la plataforma IVA debido principalmente al ecosistema de aplicaciones y servicios anteriormente desarrollados y por los cuales cada uno es protagonista en mercados donde el ganador toma todo.

El producto y servicio de Amazon logra ventajas más significativas para el usuario en lo referente a entretenimiento y comercio electrónico debido a la integración de la experiencia de usuario de la plataforma de comercio electrónico y el portal de entretenimientos Amazon Prime. Por otro lado, el producto de Google logra mejores sinergias con las funciones de búsqueda y noticias y las herramientas de escritorio dada la integración con el portal de búsquedas Google.com, Gmail, Google Calendar y los beneficios de integración con dispositivos celulares con sistema operativo Android. En el análisis de la dimensión de hogares inteligentes existe una ventaja del producto de Amazon sobre Google debido a la estrategia inicial adoptada para el



desarrollo de las aplicaciones por parte de terceros dentro de la plataforma, lo cual produjo la construcción de un gran número de aplicaciones para integrar distintos fabricantes de dispositivos inteligentes para el hogar.

	Amazon	Google	Apple
Precio	\$39.99 - \$99.99	\$49   \$99   \$299	\$299
Plataforma	<ul style="list-style-type: none"> <li>• Abierta para desarrollos.</li> <li>• + 80.000 Habilidades.</li> </ul>	<ul style="list-style-type: none"> <li>• Abierta para desarrollos.</li> <li>• + 5.000 Habilidades.</li> </ul>	<ul style="list-style-type: none"> <li>• Cerrada.</li> <li>• Funciones sobre Apps nativas de Apple</li> </ul>
Entretenimiento	<ul style="list-style-type: none"> <li>• Integración con los principales servicios de streaming (no Youtube &amp; Google Music).</li> <li>• Fire TV</li> </ul>	<ul style="list-style-type: none"> <li>• Integración con algunos servicios de streaming (no Amazon Prime o Apple Music).</li> <li>• Chromecast</li> </ul>	<ul style="list-style-type: none"> <li>• Integración solo con Apple Music.</li> <li>• Apple TV</li> </ul>
Comercio Electrónico	<ul style="list-style-type: none"> <li>• Proceso de compra completo a través de <a href="https://www.amazon.com">amazon.com</a>.</li> </ul>	<ul style="list-style-type: none"> <li>• Necesita integración con otras aplicaciones.</li> </ul>	<ul style="list-style-type: none"> <li>• Necesita integración con otras aplicaciones.</li> </ul>
Búsquedas & Noticias	<ul style="list-style-type: none"> <li>• Búsqueda de productos nativa en el portal de Amazon.com.</li> <li>• Integración con otras aplicaciones de Búsqueda &amp; Noticias</li> </ul>	<ul style="list-style-type: none"> <li>• Integración nativa con Google Search &amp; News.</li> </ul>	<ul style="list-style-type: none"> <li>• Integración con otras aplicaciones de Búsqueda &amp; Noticias</li> </ul>
Escritorio	<ul style="list-style-type: none"> <li>• Integración con aplicaciones de terceros.</li> </ul>	<ul style="list-style-type: none"> <li>• Integración nativa con las herramientas de escritorio del ecosistema Google.</li> </ul>	<ul style="list-style-type: none"> <li>• Integración nativa con las herramientas de escritorio del ecosistema Apple.</li> </ul>
Hogar Inteligente	<ul style="list-style-type: none"> <li>• Integración con aplicaciones de terceros. + de 100.000 productos soportados y 9.500 marcas.</li> </ul>	<ul style="list-style-type: none"> <li>• Integración nativa con los productos Nest y con aplicaciones de terceros. + de 30.000 productos soportados y 3.500 fabricantes.</li> </ul>	<ul style="list-style-type: none"> <li>• Integración con aplicaciones de terceros. + de 290 dispositivos soportados.</li> </ul>

Figura 12. Cuadro comparativo de características de productos comercializados en la categoría de altavoz inteligente. Fuente: Elaboración propia.

En lo que respecta al análisis del funcionamiento integral de los dispositivos vinculados a la plataforma de asistencia inteligente, el fondo de inversión de capitales de riesgo Loup VC realiza todos los años un test de inteligencia a las distintas plataformas donde evalúa el nivel de comprensión del lenguaje de las distintas plataformas y las respuestas emitidas ante cada requerimiento y pregunta.

Google Home tiene un rendimiento superior comparado al resto de los altavoces, respondiendo correctamente en el 88% de los casos y entendiendo prácticamente la totalidad de las 800 preguntas del test. Apple con su dispositivo HomePod respondió correctamente el 75% y

solo malinterpretó 3 preguntas, mientras que Echo de Amazon respondió correctamente el 73% y malinterpretó 8 preguntas. Todas las preguntas mal comprendidas normalmente involucraban un nombre propio, a menudo el nombre de una ciudad o restaurante local. Tanto el reconocimiento de voz como el procesamiento del lenguaje natural de los asistentes digitales en todos los ámbitos han mejorado hasta el punto en que, dentro de lo razonable, podrán comprender todo lo que se les diga en el corto plazo (Loup, 2018).

## Query Results

	Answered Correctly	Understood Query
Google Assistant	87.9%	100%
Siri	74.6%	99.6%
Alexa	72.5%	99.0%

Figura 13. Resultado de asertividad de las plataformas de asistente virtual. Fuente: <https://loupventures.com/annual-smart-speaker-iq-test/>



## 6. Capítulo VI – Análisis y tendencia del mercado global y de Estado Unidos.

En el capítulo anterior hemos introducido los principales productos disponibles en el mercado de altavoces inteligentes, las características de sus respectivas plataformas IVA y también hemos analizado sus ventajas y desventajas considerando las distintas dimensiones de análisis según el propósito de su función. A continuación, describiremos como es la situación de adopción y difusión de la tecnología en el mercado global y en el mercado de mayor adopción en la actualidad que es Estados Unidos. También analizaremos rasgos importantes del comportamiento del usuario, su satisfacción hacia el producto y las principales complicaciones encontradas para su adopción.

### 6.1 – Introducción.

Tomando como base las funcionalidades que hemos descrito anteriormente sobre los distintos altavoces inteligentes disponibles en el mercado, los mínimos requerimientos necesarios para su utilización son la disponibilidad de la alimentación eléctrica y una conexión a internet inalámbrica.

También consideramos necesario para maximizar las ventajas de su utilización que el altavoz se encuentre en un lugar estable, geo localizado y donde los miembros que interactúan con el producto no varíen frecuentemente. La portabilidad del producto no es algo que sea descrito por los fabricantes como una característica funcional del mismo.

La plataforma de IVA embebida en el altavoz inteligente tiene en su naturaleza de concepción la necesidad de ir comprendiendo y aprendiendo sobre los distintos comportamientos de los usuarios, sus rutinas, sus necesidades y pedidos, y de esta manera el IVA logrará ir evolucionando, siendo más efectivo y eficiente sobre los requerimientos que recibe.

Dada las condiciones mínimas de requerimientos para su uso y el contexto necesario para una mejor utilización del producto, las principales variables duras que emplearemos en la investigación para comprender el grado de adopción y difusión de la tecnología y sus correspondientes tendencias son los hogares conectados a internet.

## 6.2 – Mercado global.

La agencia Unión Internacional de Telecomunicaciones (ITU), pertenecientes al organismo de las Naciones Unidas (UN) estimó que en el año 2019 existieron aproximadamente 1.145 millones de suscripciones de conexión fija a internet, de lo cual el informe considera que cerca de 4.100 millones de personas se encuentra utilizando la conectividad de internet en el mundo. La tendencia vuelve a mostrar una recuperación en la tasa de crecimiento de cobertura, luego de haber caído de 9,7% al 6,4% de incremento del 2017 al 2018, las estimaciones indican que la cobertura de internet en hogares ha crecido del 2018 a 2019 con una tasa superior al 7% (ITU, 2019).

### Number of fixed broadband internet subscriptions worldwide from 2005 to 2019 (in millions)

Number of fixed broadband internet subscriptions worldwide 2005-2019

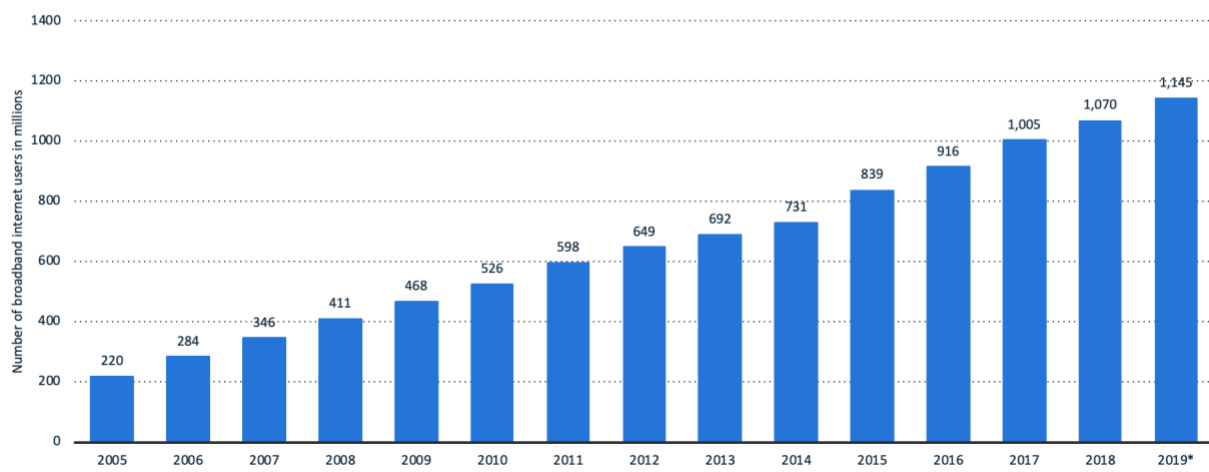


Figura 14. Cantidad de suscripciones de conexiones fijas a internet en el mundo. Fuente: ITU. [https://www.itu.int/en/ITU-D/Statistics/Documents/statistics/2019/ITU\\_Key\\_2005-2019\\_ICT\\_data\\_with%20LDCs\\_28Oct2019\\_Final.xls](https://www.itu.int/en/ITU-D/Statistics/Documents/statistics/2019/ITU_Key_2005-2019_ICT_data_with%20LDCs_28Oct2019_Final.xls)

La consultora especializada en investigación de mercados tecnológicos, Canals pronostica que la base instalada de altavoces inteligentes en todo el mundo crecerá 82.4% de 114.0 millones de unidades en 2018 a 207.9 millones en 2019. Entre los principales países que impulsan la tendencia se espera que los EE. UU. mantenga su liderazgo en términos de adopción de altavoces inteligentes, lo que representa el 42,2% de la base instalada mundial de altavoces inteligentes. Es necesario destacar que los mercados de Asia Oriental también fueron partícipes de esta tracción,

particularmente China, ya que estos países presentan grandes oportunidades de crecimiento debido a la masiva cantidad de hogares de clase media que han surgido en los últimos cinco años y a una población activa dispuesta a adoptar nuevas tecnologías.

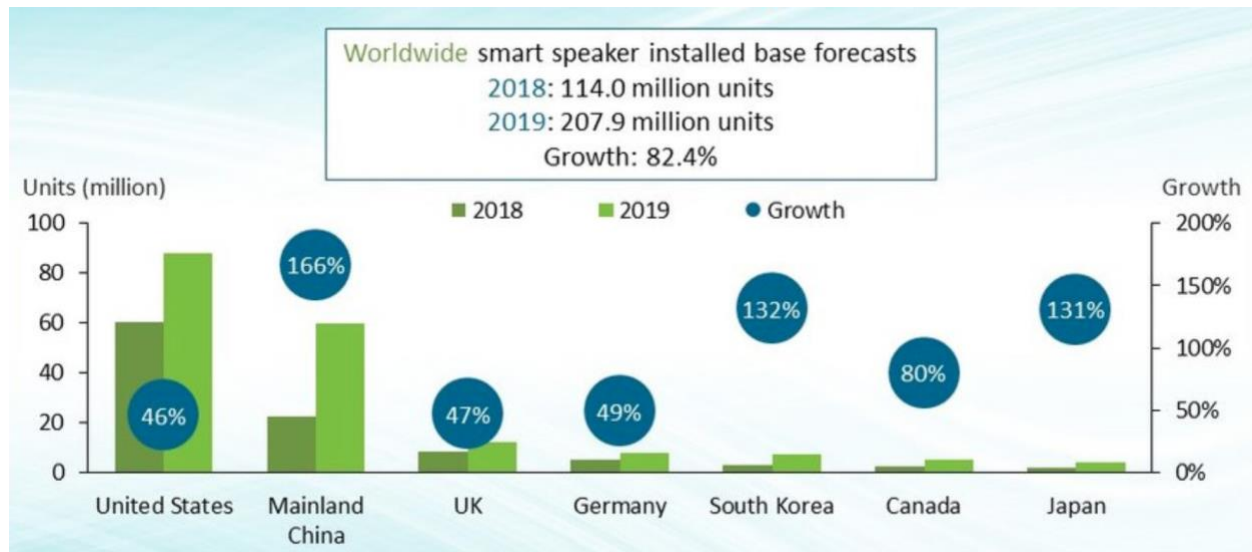


Figura 15. Pronóstico de base instalada de altavoces inteligentes para 2019 vs base instalada en el mundo en el 2018 y abierto por los países de mayor crecimiento esperado. Fuente: Canalys. <https://www.canalys.com/newsroom/canalys-global-smart-speaker-installed-base-to-top-200-million-by-end-of-2019>

En el mismo informe se espera que las ventas de altavoces inteligentes lleguen a 500 millones de dispositivos en 2023, superando de esta manera la base instalada de tabletas y dispositivos wearables pronosticada para el año 2021 (Canalys, 2019).

Un estudio realizado por la consultora Deloitte considera que los altavoces inteligentes tienen, literalmente, un mundo de oportunidades para el crecimiento. Gran parte de esa oportunidad proviene de la expansión a países cuyo idioma principal no es el inglés. A finales de 2017, las ventas de altavoces inteligentes se limitaron en gran medida a los mercados de habla inglesa, con más del 95 por ciento de las ventas en los Estados Unidos y el Reino Unido. Sin embargo, en cuanto estos dispositivos puedan lograr una madurez de funcionalidad en otros idiomas, las ventas deberían comenzar a crecer en países cuyas principales lenguas sean chino (mandarín o cantonés), francés, español, italiano o japonés. En la mayoría de estas geografías, es

altamente probable que el producto de altavoz inteligente evidencie un crecimiento más veloz en adquisición y envíos en relación a la comercialización de otros dispositivos inteligentes.

Según la investigación de Deloitte, el altavoz inteligente fue el dispositivo con mayor crecimiento de ventas año tras año durante el 2018, en seis de los siete mercados en los cuales se encontraba disponible con las múltiples marcas principales (China, Estados Unidos, Japón, Reino Unido, Canadá y Australia). A mediados del año 2018, la penetración de altavoces inteligentes lograba su mayor índice en las ciudades urbanas de China, con el 22 por ciento de los adultos con acceso a un altavoz inteligente, seguido por Estados Unidos, con el 19 por ciento de los adultos con acceso a una unidad. En estos mercados, el altavoz inteligente también fue el de más rápido crecimiento en comparación con todos los dispositivos conectados a internet que han emergido en los últimos años.

Deloitte Global estima que la demanda potencial de altavoces inteligentes podría estar en los miles de millones de unidades, posiblemente más que los teléfonos inteligentes. El informe estima que se podría instalar un altavoz en cada habitación de una casa u hotel, en cada oficina de un edificio, en cada aula de una escuela o en cada cama de un hospital (Lee, Paul; Stewart, Duncan; Loucks, Jeff; Arkenberg, 2019).

Durante la investigación de este estudio se hizo pública la cifra de ventas globales de altavoces inteligentes del año 2019. Según la consultora de investigación Strategy Analytics, el mercado global de altavoces inteligentes alcanzó un nuevo máximo en el año 2019 con ventas de 146,9 millones de unidades, un aumento del 70% en comparación al año 2018 (Strategy Analytics, 2019).

Amazon confirmó que sigue siendo la marca líder del mercado con una participación del 26,2%, por debajo del 33,7% de mercado obtenido en el año 2018. Google mantuvo el segundo lugar con una participación del 20,3%, también por debajo del 25,9% del año anterior. Los principales fabricantes chinos Baidu, Alibaba y Xiaomi aumentaron sus participaciones, mientras que Apple se mantuvo en el sexto lugar con un 4,7%.

Las ventas en el cuarto trimestre también alcanzaron un nuevo récord de 55,7 millones de unidades, la cifra de ventas trimestral más alta de la historia, según Strategy Analytics. Las mismas fueron impulsadas por las fuertes ventas navideñas en los EE. UU. y Europa, como así también por una recuperación en el negocio de altavoces inteligentes de Google luego de la presentación de nuevos productos, un mejor suministro de componentes y una fuerte actividad promocional.

Esta última cifra de ventas del año 2019, sumado a la base instalada del año 2018 de 114 millones de unidades, incrementa la base instalada mas de 250 millones de dispositivos comercializados, lo cual ubica globalmente a la adopción de la tecnología en hogares por encima del 23%. Esto implica que la adopción se encuentra en la etapa de mayoría temprana siguiendo la curva de adopción de E. Rogers (Rogers, 2003) y comienza a traspasar las puertas del mercado principal tras haber logrado exitosamente cruzar el abismo según propone G. Moore (Moore, 1991).

La localización del lenguaje puede imponer algunas limitaciones a la expansión global de los altavoces inteligentes. Es probable que la creación de soporte para nuevos idiomas requiera mucho tiempo y capital debido a la complejidad de los idiomas. En China, hay aproximadamente 130 dialectos hablados. En India, mientras que la mayoría de las personas hablan hindi, hay aproximadamente 10 diferentes variaciones de ese idioma, y la cantidad de contenido hindi disponible para aprendizaje automático actualmente es limitada. Según el análisis de Deloitte, el 90 por ciento de todos los asistentes de voz digital en India solo admiten inglés. Pero estos problemas no son insuperables y el tamaño de estos mercados ofrece un amplio incentivo para que los fabricantes de altavoces inteligentes reasignen inversiones para hacerles frente.



*Figura 16. Evolución global de adopción de altavoces inteligentes según curva propuesta por E. Rogers. Elaboración propia. Fuente: Datos recolectados de diferentes investigaciones de mercado.*

Algunas consideraciones sobre la construcción del indicador de adopción global de altavoces inteligentes. Si bien el dispositivo (hardware) no sufre variaciones dependiendo del país donde se comercializa, la plataforma de asistencia virtual inteligente sí evidencia distintos grados de maduración según el idioma y los dialectos para su utilización.

Es por ello que los criterios empleados en la teoría de aceptación tecnológica, expuestos por Davis (Davis, 1985), acerca de la percepción de utilidad y la facilidad de uso, se encuentran impactados en distintas proporciones según el grado de avance y maduración logrado por el IVA en la interpretación del lenguaje natural (NLP) del idioma y/o dialecto utilizado para su interacción.

Al mismo tiempo las cifras de dispositivos comercializados se ven distorsionadas en primer lugar, por que no se encuentra información acerca del porcentaje de dispositivos comercializados con pantalla táctil o si se trata de un altavoz inteligente adicional al ya existente para el hogar. En segundo lugar, por el auge y penetración de los dispositivos en el mercado chino, el cual se encuentra fuertemente regulado y donde las condiciones de competencia son en el mejor de los casos desiguales para los competidores extranjeros.

### 6.3 – Mercado de Estados Unidos.

Siguiendo las fuentes y los lineamientos de construcción para los indicadores de adopción de altavoces inteligentes encontramos que la cantidad de hogares con conexión fija a internet en Estados Unidos es superior a 110 millones de hogares conectados a internet para el año 2018 según la agencia ITU de Naciones Unidas (ITU, 2019).



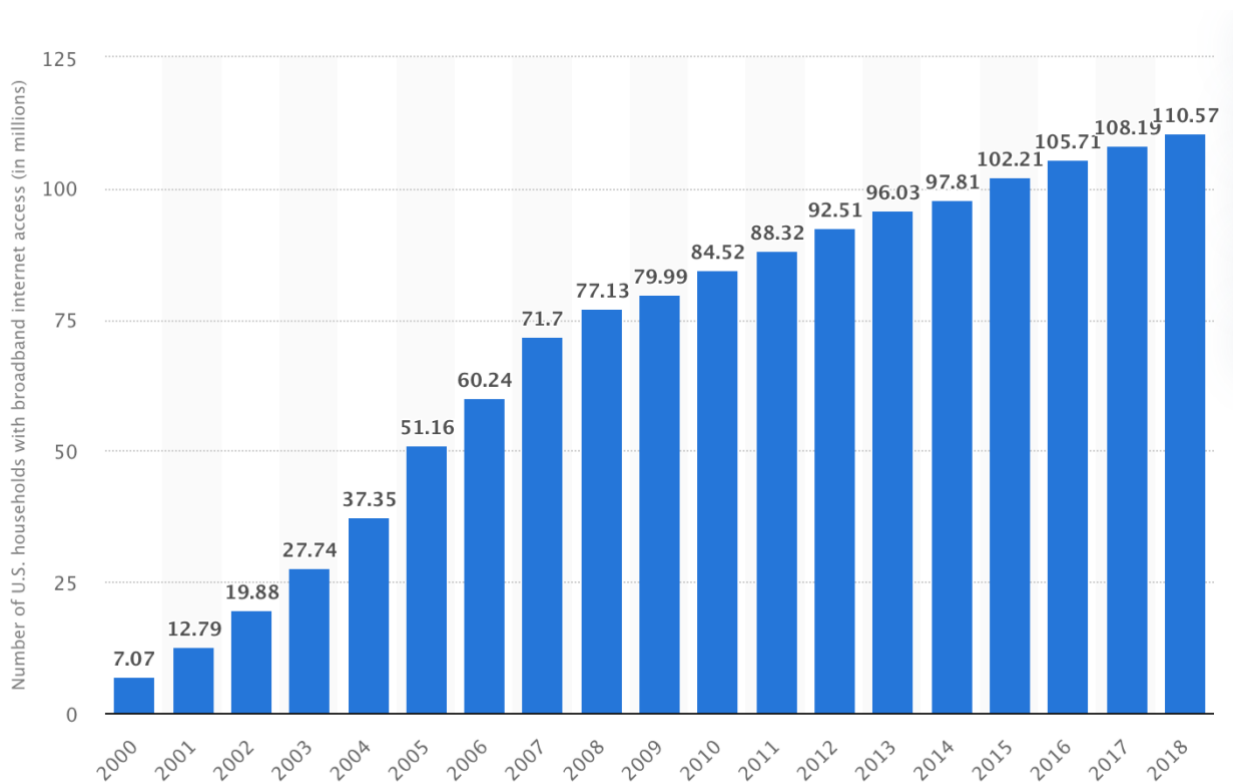


Figura 17. Número de suscripciones de banda ancha fija en los Estados Unidos de 2000 a 2018 (en millones). Fuente: ITU. [https://www.itu.int/en/ITU-D/Statistics/Documents/statistics/2019/Fixed\\_broadband\\_2000-2018\\_Dec2019.xls](https://www.itu.int/en/ITU-D/Statistics/Documents/statistics/2019/Fixed_broadband_2000-2018_Dec2019.xls)

En la plataforma [Data.census.gov](https://data.census.gov) para acceder a datos y contenido digital de la Oficina del Censo de EE. UU., para el año 2018 se publicó que el 85.3 % del total de hogares de los EE. UU. posee alguna suscripción a Internet y el formato más popular de suscripciones de Internet entre los hogares en los Estados Unidos fue la banda ancha, con casi el 76 % de los hogares suscritos a este servicio.

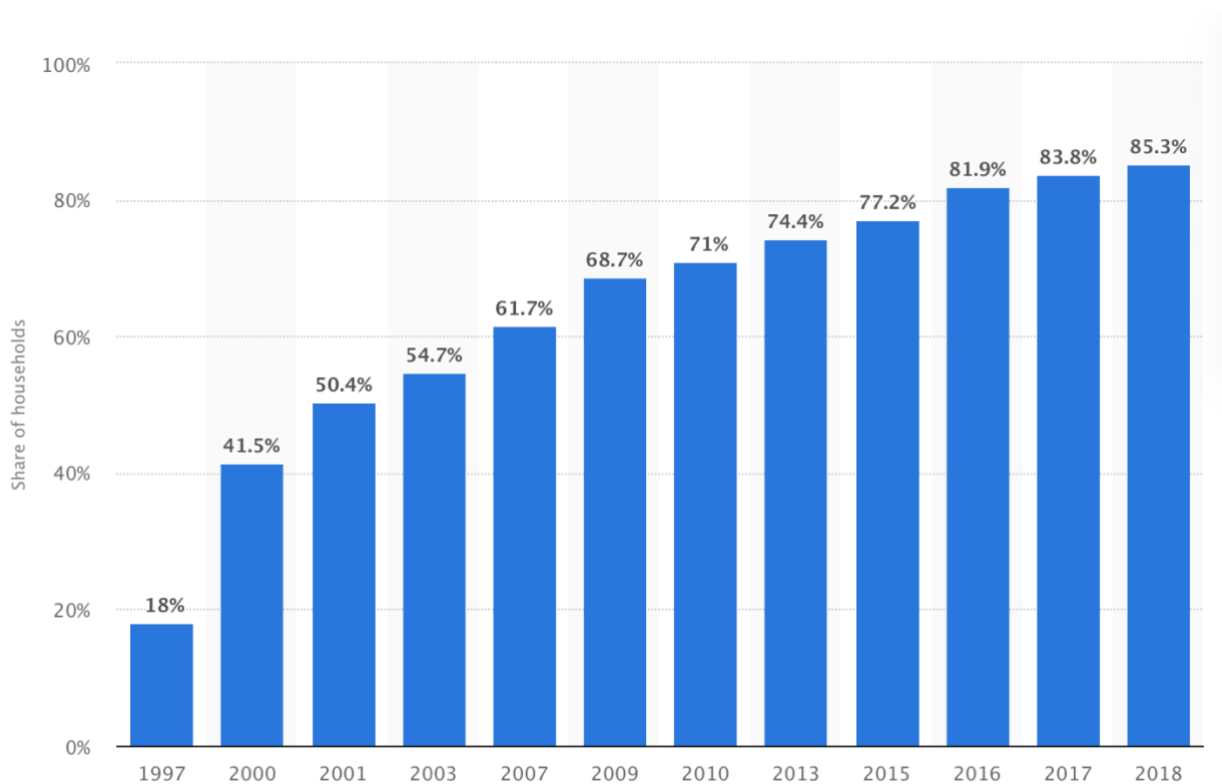


Figura 18. Porcentaje de hogares con uso de internet en los Estados Unidos de 1997 a 2018. Fuente: Oficina del Censo Gob. EE.UU. <https://data.census.gov/cedsci/table?q=home%20broadband&lastDisplayedRow=12&table=B28002&tid=ACSDTIY2018.B28002&hidePreview=true&tp=false>

La Asociación de Consumo Tecnológico de EE.UU. (CTA por sus siglas en inglés) estima que gran parte de los consumidores se encuentra actualizando sus altavoces inteligentes a dispositivos habilitados para voz con una pantalla de visualización. Las ventas de altavoces inteligentes para el año 2019 se ubicaron en 37.1 millones de unidades y se proyecta que las ventas de unidades alcancen los 39 millones de unidades vendidas (un 5% más que el 2019) y U\$S 4,2 mil millones de dólares estadounidenses en ingresos (14% mayor al 2019) en 2020 (CTA, 2019).

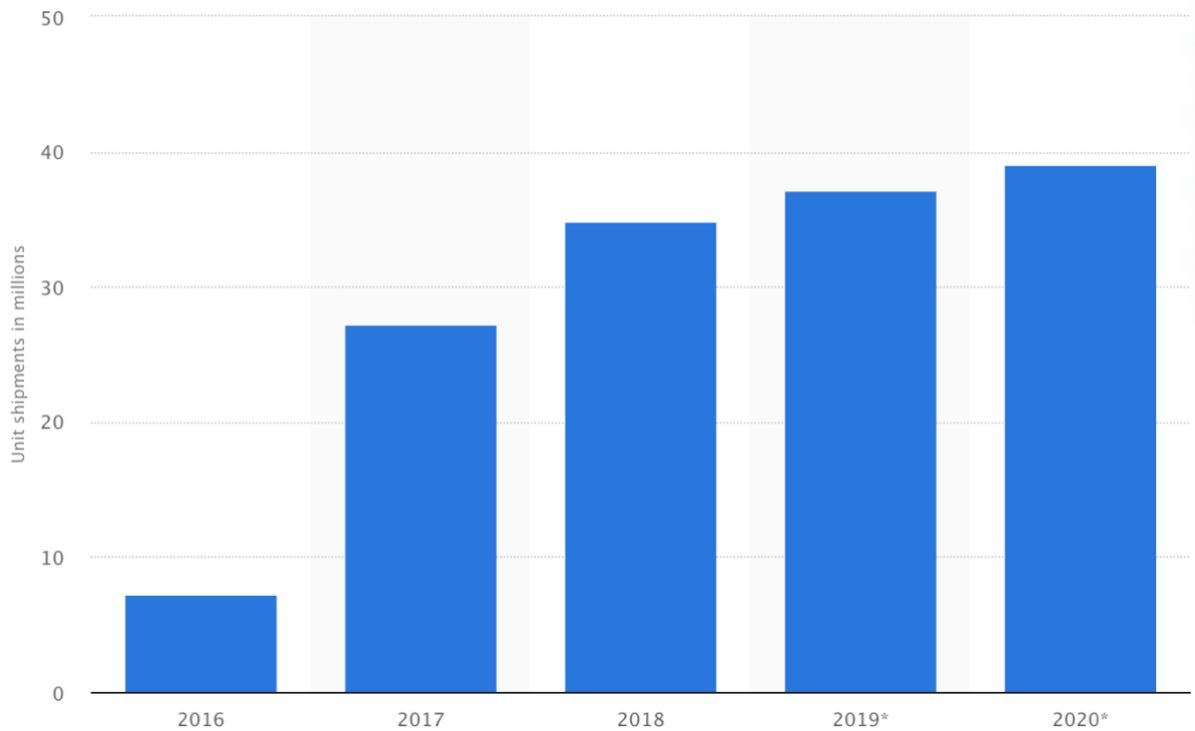


Figura 19. Envíos de unidades de altavoces inteligentes en los Estados Unidos de 2016 a 2020 (en millones). Fuente: CTA. [https://www.cta.tech/Resources/Newsroom/Media-Releases/2020/January/Consumer-Tech-U-S-Sales-to-Reach-Record-\\$422-B-\(1\)](https://www.cta.tech/Resources/Newsroom/Media-Releases/2020/January/Consumer-Tech-U-S-Sales-to-Reach-Record-$422-B-(1))

Continuando con los lineamientos básicos seguidos para la construcción de indicador global de adopción podemos calcular el porcentaje de adopción para el mercado de EE.UU. La base instalada al año 2019 es de aprox. 106.8 millones de dispositivos (69.7 millones en 2018), con un 53% de crecimiento interanual en la base instalada.

En el año 2018 los hogares conectados a internet con suscripción fija fueron 110.6 millones lo que resulta en un índice de adopción del 63% en el año 2018. Si proyectamos la misma tasa de crecimiento de los hogares conectados a internet del período 2017/18 (2,2%) para el año 2019, podemos estimar que los hogares conectados a internet se encuentran cercanos a los 113 millones de suscripciones. Lo que resulta en una tasa de adopción de altavoces inteligentes estimada de 94,5%.

En la construcción del análisis de adopción para el mercado de EE. UU. Nos encontramos con informes e indicadores que permiten construir un índice de adopción enriquecido considerando

variables más complejas como la relación de unidades de dispositivos por hogar, las características socio-demográficas de los compradores y resultados de las encuestas realizadas acerca de la utilización y satisfacción del producto.

Un estudio efectuado por las consultoras NPR y Edison Research concluye que la adopción de altavoces inteligentes creció hasta el 24% de la población adulta de EE. UU. En el 2019, o lo que sería equivalente a alrededor de 60 millones de personas, un crecimiento de 14% con respecto al año anterior, según la encuesta realizada en diciembre de 2019. El estudio también encontró que los hogares con altavoces inteligentes contienen un promedio de 2.6 dispositivos por hogar, en comparación con 2.3 del año anterior y 1.7 en diciembre de 2017 (NPR, 2019).

La encuesta nacional telefónica realizada a más de 1,000 personas mayores de 18 años también muestra que el 54% de la población de EE. UU. ha usado algún tipo de tecnología de comando de voz, como asistentes de voz en teléfonos inteligentes, altavoces inteligentes y otros dispositivos. De los que usan asistentes de voz, el 60% dice que es un usuario frecuente, utilizando su dispositivo al menos una vez al día.

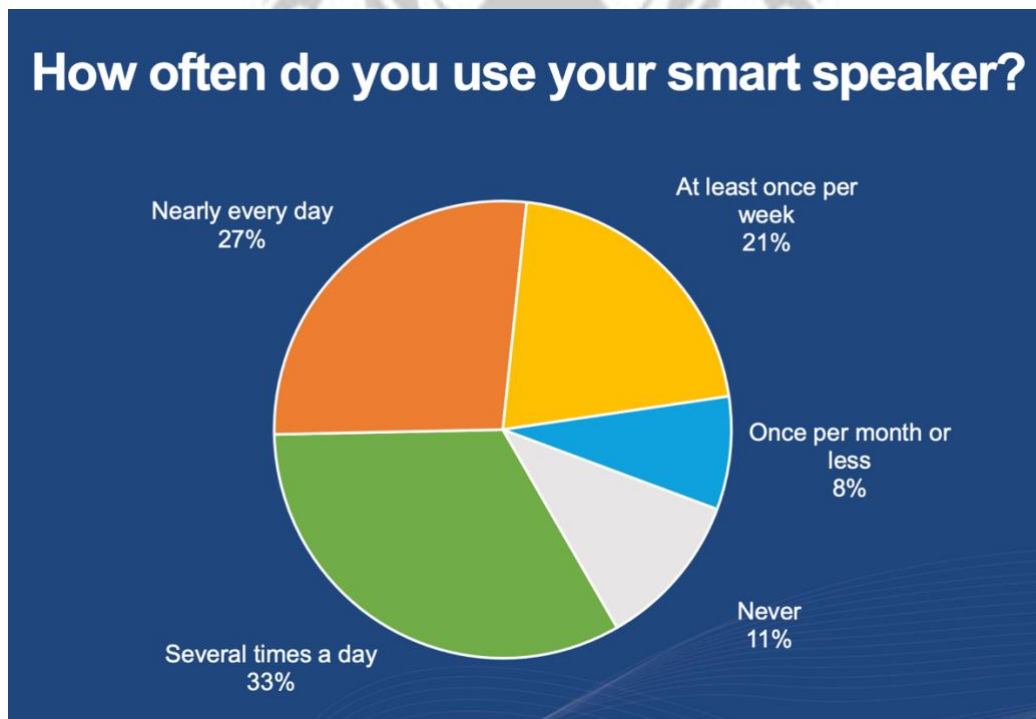
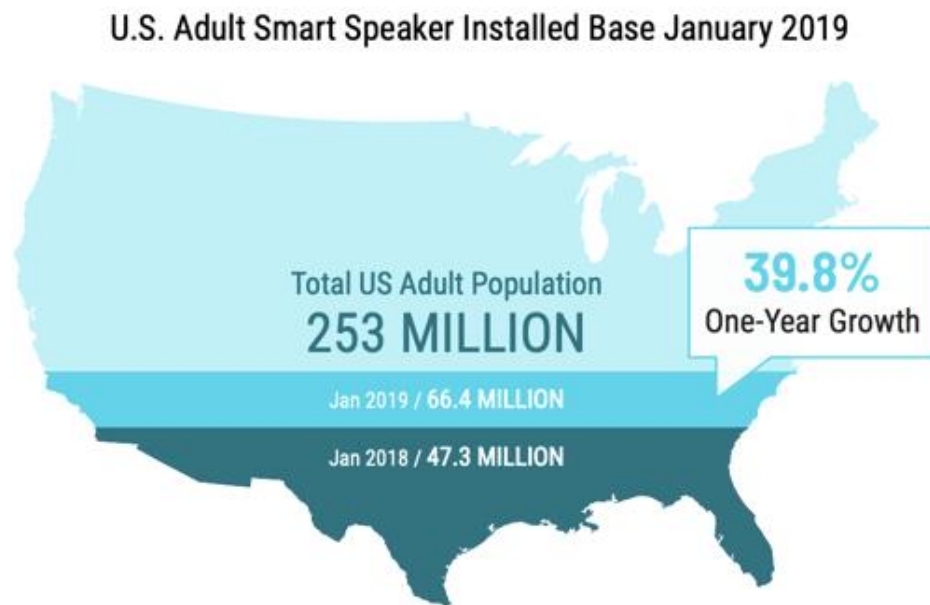


Figura 20. Porcentaje de frecuencia de utilización de altavoces inteligentes. Fuente: NPR & Edison Research. <https://www.nationalpublicmedia.com/uploads/2020/01/The-Smart-Audio-Report-Winter-2019.pdf>

La publicación anual del sitio Voicebot.ai en 2019 estima que el ratio de parlantes inteligentes por usuario aumentó más del 10% de 1.8 en 2018 a 2.0 en 2019. Sin embargo, el informe establece que la expansión y la popularidad de los altavoces inteligentes también ha atraído a nuevos usuarios ocasionales por lo que en cuanto a la frecuencia de uso se identificó que mientras más del 60% de los propietarios de altavoces inteligentes en enero de 2018 se identificaron como usuarios diarios, menos del 50% lo hizo al año posterior y el número de propietarios de dispositivos que afirman usar sus altavoces inteligentes nunca o rara vez se duplicó al 26% (Voicebot.ai, 2019).

El porcentaje de adultos estadounidenses que poseen altavoces inteligentes aumentó un 40,3% en 2019, pasando de 47,3 millones a 66,4 millones durante el año. Este aumento significa que más de uno sobre cada cuatro adultos de EE. UU. ahora tiene acceso a un asistente de voz basado en altavoces inteligentes. Los altavoces inteligentes tardaron menos de cuatro años en lograr un 25% de adopción desde la introducción inicial restringida a los miembros de Amazon Prime.



*Figura 21. Base de altavoces inteligentes en EE.UU. Enero 2018 vs Enero 2019. Fuente: Voicebot Consumer Adoption Report January 2019. [https://voicebot.ai/wp-content/uploads/2019/03/smart\\_speaker\\_consumer\\_adoption\\_report\\_2019.pdf](https://voicebot.ai/wp-content/uploads/2019/03/smart_speaker_consumer_adoption_report_2019.pdf)*

Amazon Echo Dot es el altavoz inteligente más adoptado por un margen significativo. El dispositivo de precio de lista de menos de U\$S 50 dólares está normalmente disponible a través de ofertas en diferentes canales de distribución por precios inferiores y los modelos restaurados se pueden adquirir por menos de U\$S 20 dólares. Este dispositivo ha demostrado ser más popular que las ofertas de mayor precio de Amazon, como Echo, Echo Plus, Echo Spot y Echo Show.

En el portafolio de Google, los productos Home y Home Mini parecen ser igualmente populares con un 11.2% de participación cada uno. El informe indica que es probable que haya más productos Home Mini en uso en términos de dispositivos totales ya que el análisis de la consultora refleja el número de usuarios con acceso a un dispositivo y no la cantidad de dispositivos por hogar. Esta situación puede ser común ya que el 87% de los propietarios de altavoces inteligentes de Google informan tener ambos dispositivos (Voicebot.ai, 2019).

Probablemente el mayor cambio en la composición de los patrones de uso de los propietarios de altavoces inteligentes es el comportamiento acerca de la recurrencia de usuarios más ocasionales sobre los dispositivos. Casi el 64% de los propietarios de dispositivos en enero de 2018 informaron ser usuarios diarios. En enero de 2019 ese número cayó a solo alrededor del 47%. Los usuarios mensuales fueron bastante similares, con la diferencia compensatoria de los usuarios poco frecuentes que aumentaron del 13% a más del 26%. Un comportamiento esperado sería que los primeros en adoptar la tecnología tengan más probabilidades de incorporar los altavoces rápidamente en sus hábitos diarios que los consumidores que tienden a adoptarla más tarde. Tres de cada cuatro propietarios de parlantes inteligentes aún informan que son usuarios activos mensualmente. El informe indica que mientras veamos ese tipo de uso constante junto con un crecimiento continuo en la venta de los dispositivos, los altavoces inteligentes seguirán creciendo en importancia como un canal de asistente de voz para la participación del consumidor (Voicebot.ai, 2019).

## U.S. Smart Speaker Frequency of Use 2018

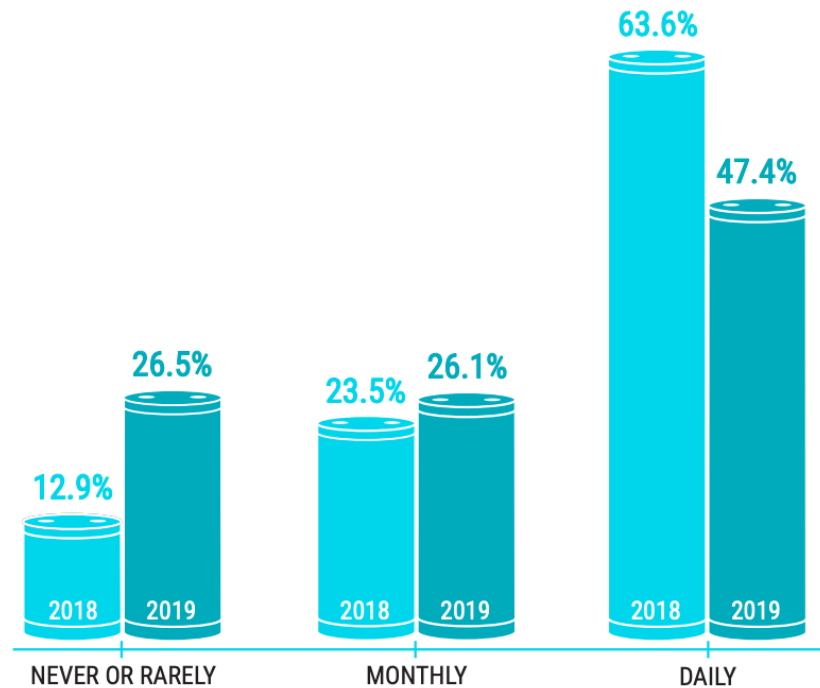


Figura 22. Frecuencia de utilización de altavoces inteligentes en EE.UU. Enero 2018 vs enero 2019. Fuente: Voicebot Consumer Adoption Report January 2019. [https://voicebot.ai/wp-content/uploads/2019/03/smart\\_speaker\\_consumer\\_adoption\\_report\\_2019.pdf](https://voicebot.ai/wp-content/uploads/2019/03/smart_speaker_consumer_adoption_report_2019.pdf)

Otra interpretación posible es que los resultados evidenciados en el informe podrían indicar una tendencia en la disminución del uso por insatisfacción del usuario en cuanto a beneficios y prestaciones concretos del producto o podría demostrar una evolución natural en la curva de adopción pasando de los usuarios con características de adoptantes tempranos a nuevos usuarios con rasgos de mayoría temprana. Un rasgo a considerar sobre los diferentes conjuntos de adoptantes es que, por lo general, el conjunto de la mayoría temprana otorga una mayor percepción de valor a la amplitud de aplicaciones y funcionalidades y también a las integraciones posibles con otros dispositivos (Rogers, 2003).

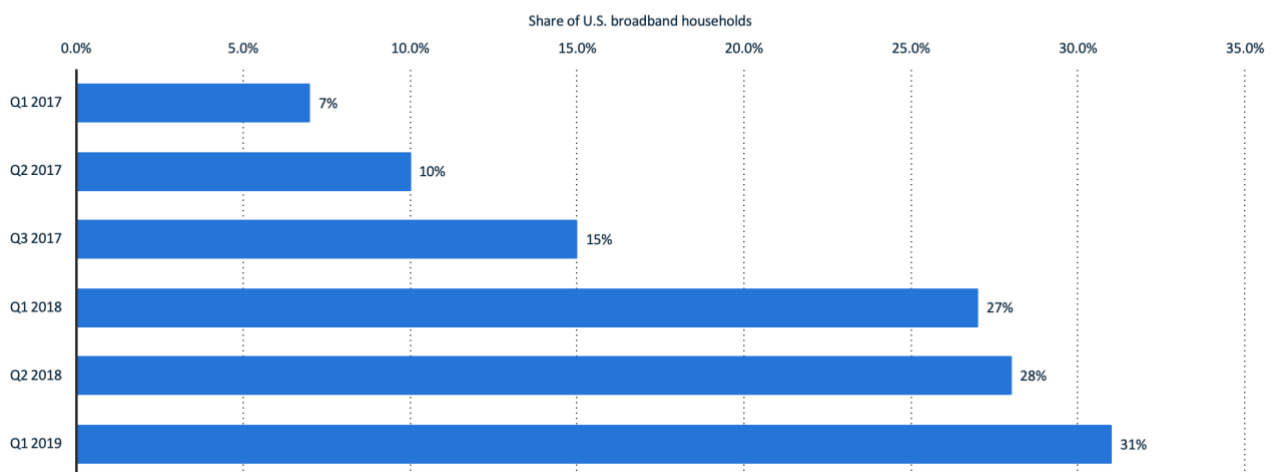
El informe pronostica que se debe esperar que los fabricantes de altavoces inteligentes direccionen los esfuerzos de desarrollos hacia la construcción de nuevas aplicaciones, mejorar la calidad de los accesos y las integraciones de terceros en lo que respecta a la evolución del producto.

Sonia Agnese, Analista Sr. De Omdia, establece una alarma al respecto, dado que las últimas encuestas realizadas a usuarios en America Latina, casi el 40% de los usuarios no pueden encontrar una utilidad al dispositivo y se encuentran disconformes porque el IVA no ha logrado comprender el mensaje y/o la instrucción (S. Agnese, comunicación por email, 16 de Abril de 2020)

Parks Associate, consultora especializada en investigación y análisis de mercado para productos y servicios relacionados con internet de las cosas (IoT), hogares inteligentes y tecnologías de entretenimiento y conexión, publicó en abril del año 2019 la medición de su índice de penetración de altavoces inteligentes en hogares con conexión a internet y el mismo se ubicó en 31% para el primer trimestre del año 2019 (Park Associates, 2019).

**Smart speaker with intelligent personal assistant ownership rate among U.S. broadband households from 2017 to 2019\***

Smart home voice assistants U.S. household ownership 2017-2019



*Figura 23. Evolución Tasa de penetración de altavoces inteligentes sobre los hogares conectados a internet de EE. UU. De 2017 a 2019. Fuente: Parks Associates. <http://www.parksassociates.com/blog/article/cus-2019-pr4>*

El fondo de capital de riesgo Loup Ventures, que realiza todos los años las pruebas de inteligencia de las plataformas de IVA, publicó en junio del 2019 una proyección actualizada de sus modelos de adopción en base a los nuevos aumentos de las cifras de ventas de altavoces



inteligentes del 2019. El estudio indica que más de un tercio de todos los hogares de EE. UU. tendrán un altavoz inteligente, un dispositivo que no existía hace solo cinco años, a fines de 2019 (Loup, 2019).

### Smart speaker household penetration rate in the United States from 2014 to 2025\*

Smart speaker household penetration rate U.S. 2014-2025

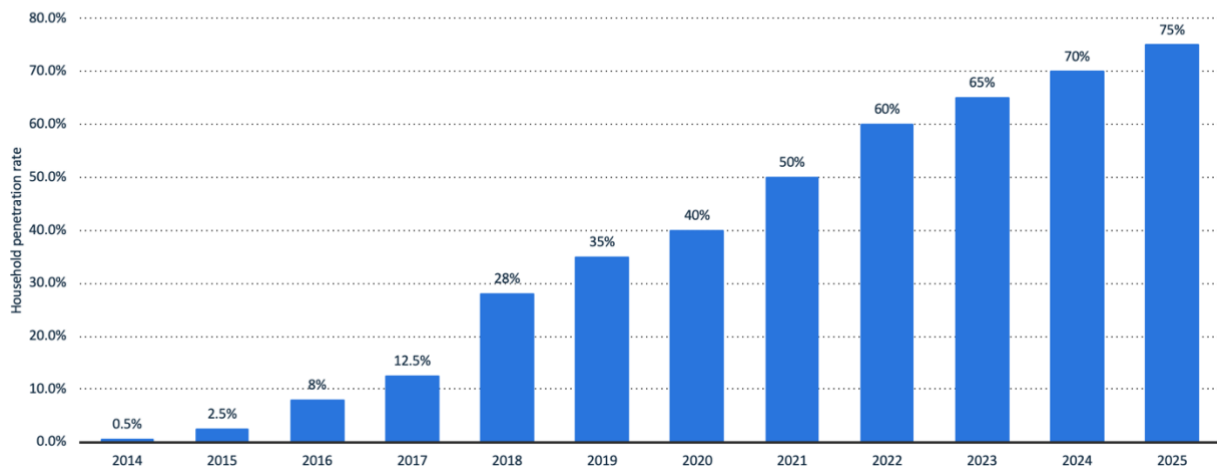


Figura 24. Proyección de tasa de penetración de los hogares con altavoces inteligentes en los Estados Unidos desde 2014 - 2025. Fuente: Loup VC. <https://loupventures.com/smart-speaker-market-share-update/>

#### 6.3.1 – Comportamiento y satisfacción de usuario.

A continuación, analizaremos las recientes investigaciones y estudios de satisfacción de cliente para obtener un panorama más claro acerca de si la adopción de los altavoces inteligentes en EE.UU. encuentra fundamentos para su crecimiento en los beneficios que brinda sobre la plataforma IVA y los futuros cambios de comportamiento que ayudarían a mejorar los avances de una próxima adopción en otros países.

En una investigación de mercado realizada por Microsoft a comienzos del 2019 se destaca que el 80% de los usuarios entrevistados que poseen un altavoz inteligente se encuentran satisfechos con las prestaciones de su asistente virtual, sólo el 6% está disconforme y un 14% se mantiene neutral en su opinión.

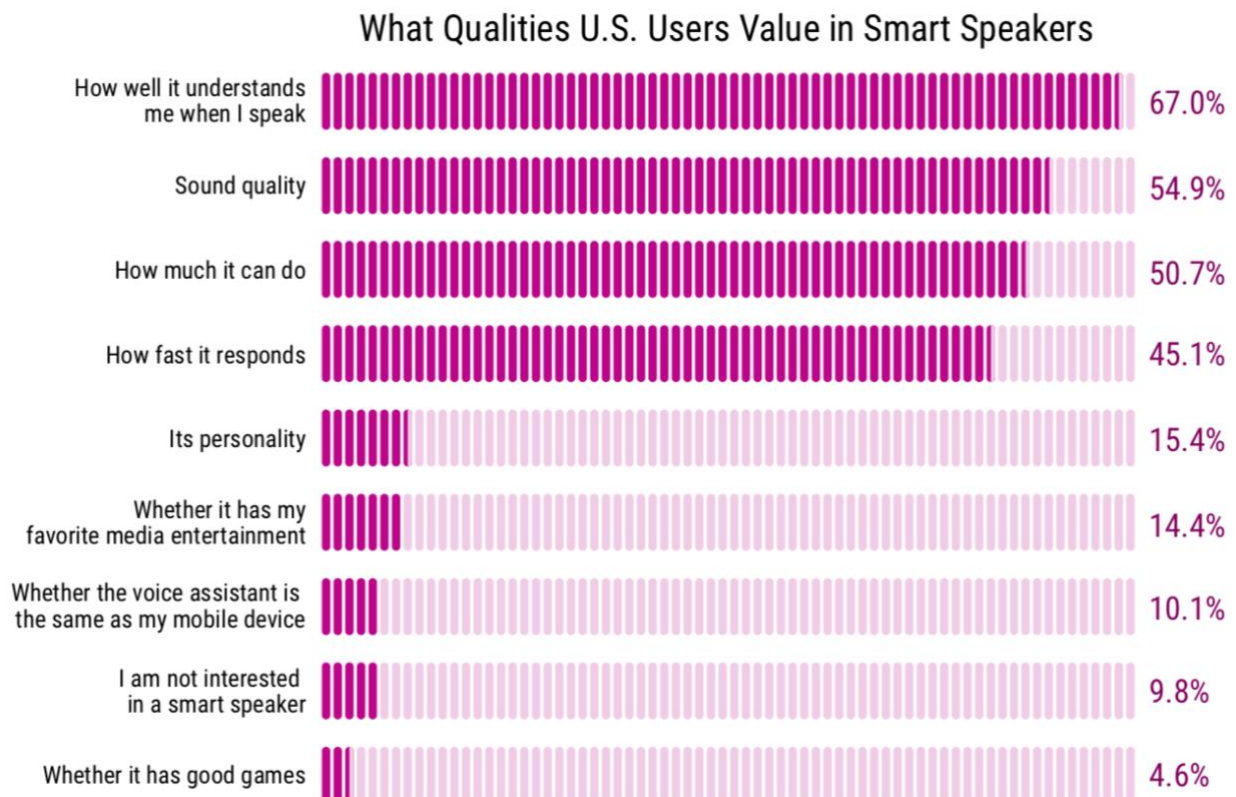
Una de las principales barreras identificadas que podría estar frenando la utilización del IVA en situaciones cotidianas y con un mayor grado de complejidad detectadas en el estudio, es la preocupación acerca de la privacidad de la información. El 41% de los encuestados respondió tener ciertos reparos al momento de la utilización del dispositivo debido al temor de estar siendo grabados durante la escucha pasiva y un 51% expresó preocupación sobre la seguridad concerniente a los datos almacenados.

El informe también pronostica el crecimiento del comercio de voz (*vCommerce*), donde el 40% de los encuestados han intentado hacer una compra utilizando su voz a través de su asistente digital o un altavoz doméstico inteligente. De los encuestados que hicieron una compra a través del IVA, el 74% informó que pudo completar su transacción con bastante facilidad, el 21% informó que tenía dificultades para completar la transacción y el 5% informó que no pudo completar su transacción (Microsoft, 2019).

El informe de adopción del portal *Voicebot.ai*, especializado en tecnologías relacionadas con la voz y la inteligencia artificial, indica que dos tercios de los consumidores entienden como la cualidad más importante del altavoz inteligente la exactitud con la cual los comprende el asistente de voz. A esto le sigue la calidad de sonido (54.9%), la cantidad de funcionalidades que se pueden realizar (50.7%) y la velocidad de respuesta (45.1%). Cualquier otra cualidad está muy por detrás de estas cuatro características principales (Voicebot.ai, 2019).

Amazon, Apple y Google han comunicado muchas veces sobre su enfoque en agregar personalidad a los asistentes de voz a pesar de que solo el 15,4% de los propietarios de altavoces inteligentes lo consideran importante. Esta baja calificación puede estar influenciada por el hecho de que la personalidad es una característica ofrecida por todos los proveedores líderes de asistentes de voz, pero claramente no es algo que tenga impacto hoy en día en la percepción del usuario.

Un indicio importante que se desprende del estudio de mercado es que la propiedad de teléfonos inteligentes solo influyó en la selección de altavoces inteligentes en aproximadamente uno de cada diez consumidores.



*Figura 25. Cualidades valoradas por usuarios de altavoces inteligentes en EE.UU.*  
 Fuente: Voicebot Consumer Adoption Report January 2019. [https://voicebot.ai/wp-content/uploads/2019/03/smart\\_speaker\\_consumer\\_adoption\\_report\\_2019.pdf](https://voicebot.ai/wp-content/uploads/2019/03/smart_speaker_consumer_adoption_report_2019.pdf)

En términos de funcionalidad el informe revela que, por segundo año consecutivo, hacer preguntas generales es el caso de uso más utilizado por los propietarios de altavoces inteligentes. Sin embargo, no es el principal caso de uso empleado en forma mensual o diaria. Esa distinción se refiere a escuchar los servicios de transmisión de música como sucedió también en 2018. El tercer lugar en ambos años fue consultar sobre el clima seguido por la utilización de despertadores y alarmas en las posiciones cuarta y quinta.

Cuatro de los cinco casos de uso principales son los que se consideran servicios de primera parte. Eso significa que son proporcionados por el asistente de voz de manera nativa. Dos de los seis casos de uso principales involucran música que son servicios de entretenimiento de terceros. Las posiciones 7, 8 y 9 son los servicios de terceros más tradicionales, muchos de los cuales fueron creados por desarrolladores independientes de habilidades y acciones de Alexa y Google

respectivamente. Por lo tanto, estableciendo un orden de frecuencia de uso a nivel de categoría las primeras son las utilidades provistas de primera parte por el dispositivo, en segundo lugar las aplicaciones de entretenimiento de terceros y en tercer término las aplicaciones y servicios de terceros.

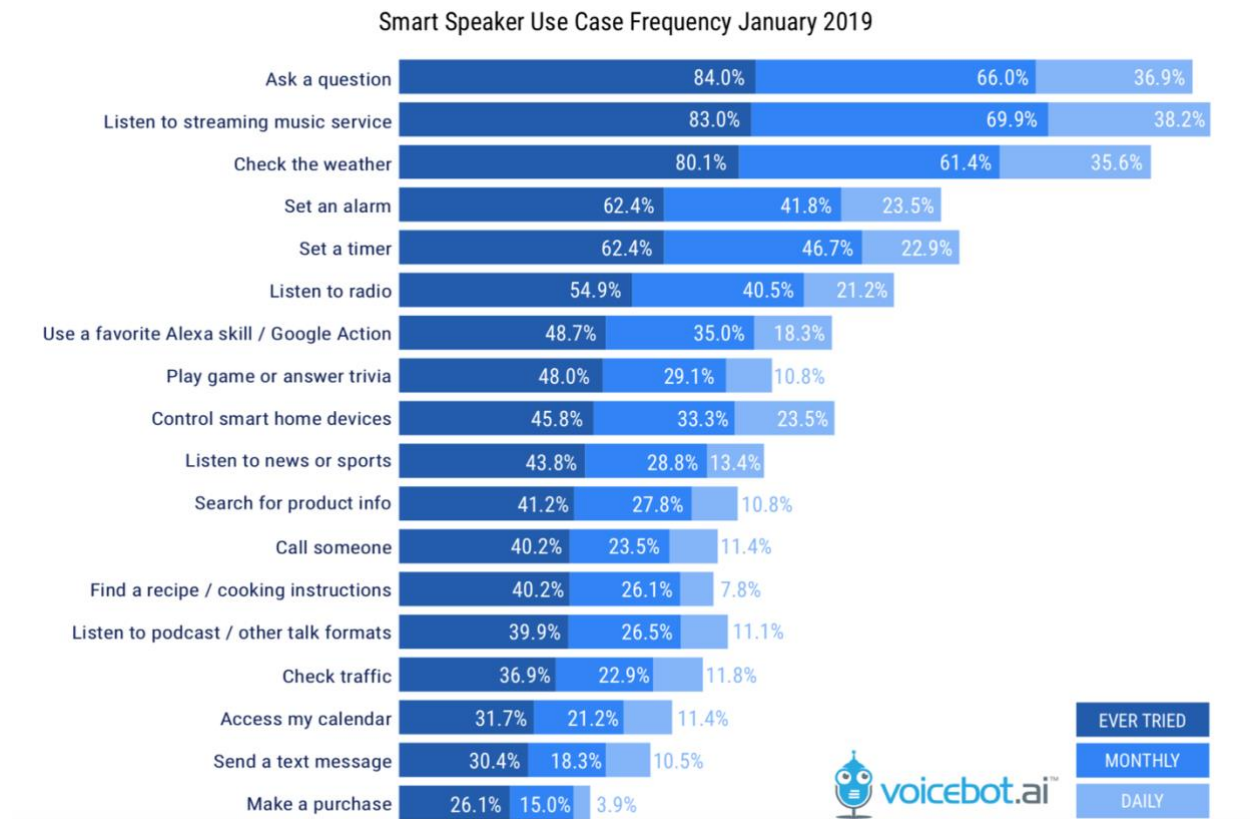


Figura 26. Frecuencia de utilización de casos en Altavoces Inteligentes. Fuente: Voicebot Consumer Adoption Report January 2019. [https://voicebot.ai/wp-content/uploads/2019/03/smart\\_speaker\\_consumer\\_adoption\\_report\\_2019.pdf](https://voicebot.ai/wp-content/uploads/2019/03/smart_speaker_consumer_adoption_report_2019.pdf)

La mayor variación entre la frecuencia de utilización se produce en la funcionalidad de control inteligente del hogar, que se ubica noveno en términos de "probado alguna vez" y cuarto para "uso activo diario". Esto se debe al efecto de complementariedad donde es necesario tener un dispositivo doméstico inteligente para usar esta característica. Sin embargo, controlar las luces o los termostatos ya son funciones diarias y si tiene dispositivos domésticos inteligentes para estas funciones, entonces cambiar sus hábitos de control de aplicaciones de teléfonos inteligentes a interacción de voz es un cambio relativamente fácil. Esto no solo conduce a una mayor percepción

de valor por parte de los consumidores, sino que también conduce a la fidelización del mismo, lo que significa para los fabricantes tener una menor probabilidad de que los consumidores los retiren o los cambien por un producto de la competencia.

Otra particularidad que tuvieron los casos de uso relacionados con el hogar inteligente fue el crecimiento interanual de utilización de esta característica. En 2018 fue el decimocuarto caso de uso más común y subió al noveno en 2019. Más del 45% de los propietarios de altavoces inteligentes han utilizado los altavoces inteligentes para controlar dispositivos domésticos inteligentes, cuando solo se trató del 38% en 2018 (Voicebot.ai, 2019).

Juan Gnius, Director de Telracom, re afirma la perspectiva sobre el crecimiento en la utilización de IVA como administrador del hogar inteligente con su visión acerca de la principal utilidad que puede ayudar a acelerar su adopción en hogares (J. Gnius, comunicación por email, 3 de Abril de 2020).

El informe de Voicebot.ai al igual que el estudio de Microsoft, hacen hincapié en la dimensión de seguridad como una de las barreras más importantes a sobrepasar para lograr una tracción más frecuente de usuarios y funciones. Dos tercios de los consumidores expresan al menos algo de preocupación relacionada a la privacidad, mientras que el 26% de los encuestados admite estar muy preocupados.

Un dato complementario es que las preocupaciones de privacidad alcanzan a todos los consumidores, aquellos que no poseen altavoces inteligentes tienen prácticamente preocupaciones idénticas. Por ejemplo, el 27.7% de los consumidores sin altavoces inteligentes dijeron también estar “muy preocupado” por los problemas de privacidad en comparación con el 21,9% de los propietarios de dispositivos. Ignacio Perrone, Director de Frost & Sullivan advierte que solo el rumor del hecho de la posible filtración de información de datos privados del usuario en ciertos casos alcanza para disuadir la adquisición de este tipo de dispositivos (I. Perrone, comunicación por email, 7 de Abril de 2020)

## 7. Capítulo VII – Análisis de adopción del mercado argentino y proyección de estimaciones.

En el capítulo anterior describimos el estado de adopción de los altavoces inteligentes en el mundo y en el mercado de EE.UU., también describimos los diferentes análisis realizados de penetración a nivel usuario como a nivel hogares y construimos los indicadores de adopción para el mercado de hogares con conexión a internet global y de EE.UU. En el desarrollo del presente apartado tenemos como objetivo estructurar las estimaciones de adopción que se ajustarían en un futuro cercano al mercado argentino, para ello evaluaremos las curvas de adopción sufridas por productos con características similares y sus respectivas analogías sobre la evolución de adopción del mercado estadounidense.

### 7.1 – Introducción.

En el desarrollo de este capítulo describiremos como fue la evolución de la situación de adopción de las computadoras tablet y de los televisores inteligente (TI) en el mercado estadounidense y su comparación con el mercado argentino. De esta manera poder estimar proyecciones y trazar algún tipo de paralelismo al posible marco de adopción de altavoces inteligentes en el mercado argentino.

Durante la investigación identificamos a las tablets y los TI como los productos con mayor similitud y contemporaneidad a los altavoces inteligentes dadas las estadísticas que existen en términos de su medición de adopción a nivel de hogar y no a nivel de persona o individuo.

La justificación de la elección se debe ha que hemos considerado a la tablet y al TI de alguna manera bienes complementarios dentro del ecosistema de productos electrónicos del hogar junto al altavoz inteligente ya que creemos que ambas demandas se encuentran fuertemente relacionadas como un posible hilo de adquisición en el transcurso del tiempo. Nuestra premisa de estimación parte de la base que el conjunto de hogares que poseen ambos dispositivos son los principales candidatos a adoptar tempranamente un altavoz inteligente e introducir en sus vidas la utilización del IVA.

En ambos casos los productos, al igual que los altavoces inteligentes, se utilizan para establecer una interfaz de comunicación, dentro del hogar, hacia una plataforma a través de una

conexión a internet con diversas aplicaciones y que brindan distintos tipos de servicios con disímiles utilidades finales.

## 7.2 – Evolución y adopción de computadoras tablet en Estados Unidos y Argentina.

Una computadora tablet, o comúnmente denominada tablet, es una computadora portátil que utiliza una pantalla táctil como dispositivo de interfaz principal. La mayoría de las tablets son más pequeñas y pesan menos que la computadora portátil promedio.

Si bien se conocen varios desarrollos de tablets anteriores al lanzamiento del Ipad realizado por Apple Inc. en el año 2010. El Ipad fue el primer diseño de producto en lograr generar adopción masiva en este formato de computadora sin teclado en EE.UU.

La Tablet, en términos de precio, es un producto que a diferencia de los altavoces inteligentes se lanzó al mercado con precios relativamente altos y luego de su éxito en ventas el mismo fue disminuyendo gracias a la competencia y a la entrada de productos competitivos provenientes de otros fabricantes de tecnología.

Por otro lado, la tablet amplió el uso y funcionamiento habitual de los teléfonos inteligentes (*smartphones*) sin la necesidad de instaurar un cambio de hábito o de comportamiento con respecto al usuario final.

Otra característica distintiva de la tablet que es importante tener en cuenta durante el análisis de adopción es que a diferencia de lo que sucede con los altavoces, las tablets no requieren la utilización de una plataforma de inteligencia artificial para su funcionamiento, ya que la principal utilidad percibida de la tablet esta relacionada al ecosistema de aplicaciones disponibles que permiten sus diferentes usos, en este caso la variable de crecimiento de la tablet conserva una relación directa con la misma variable a la cual se encontraron expuestos los *smartphones*.

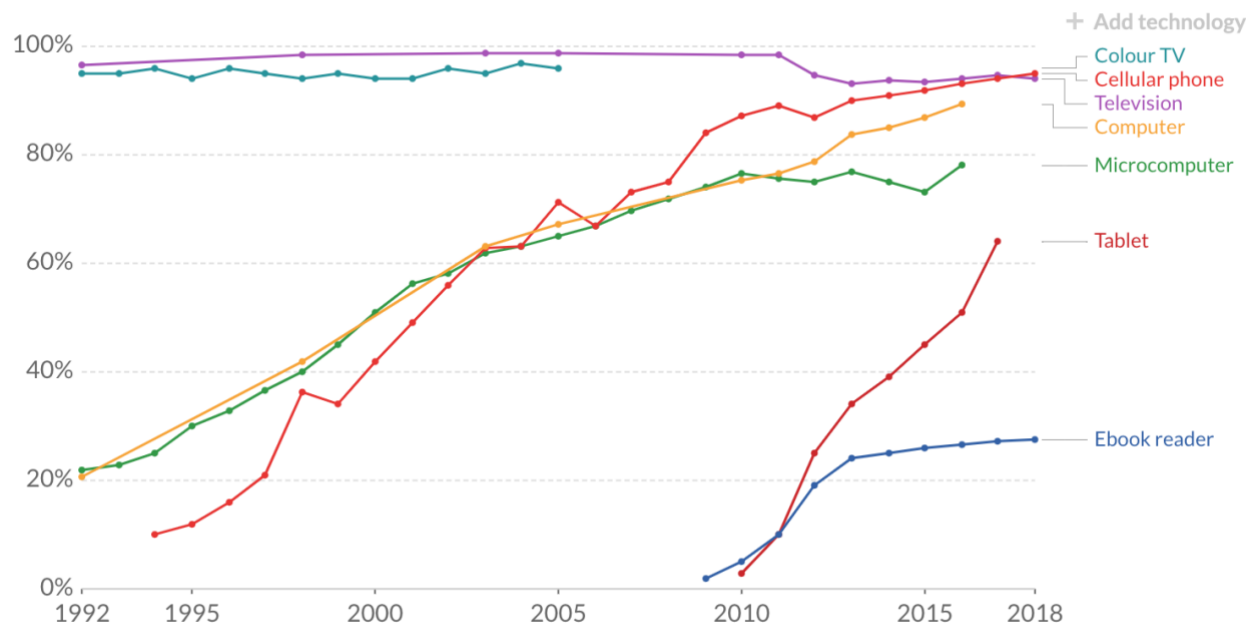
Por lo tanto en virtud de lo anterior, podemos encontrar distintos tipos de comportamiento sobre la curva de adopción dada las distintas intensidades que pueden existir en la tasa de penetración.

Un informe elaborado por la Asociación de Consumidores de Tecnología (CTA por sus siglas en inglés) y la consultora centris.com en términos de penetración en hogares del mercado estadounidense las tablets han logrado en el primer trimestre del año 2019 una penetración cercana

al 65%, logrando así la tasa de adopción en hogares más elevada y la de mayor velocidad de adopción para un dispositivo tecnológico hasta la aparición del altavoz inteligente en EE.UU. (CTA, 2019).

## Technology adoption in US households

Technology adoption rates, measured as the percentage of households in the United States using a particular technology.



Source: Comin and Hobijn (2004) and others

Note: See the sources tab for definitions of household adoption, or adoption rates, by technology type.

CC BY

Figura 27. Adopción en hogares de dispositivos tecnológicos. Fuente: Our World Data. Recuperado en: <https://ourworldindata.org/technology-adoption>

El mismo informe elaborado un año atrás describe la penetración de tablets en hogares con un ratio del 64%, por lo que el crecimiento interanual en el año 2019 ha sido el más bajo observado por el producto hasta el momento. En términos de adopción, pareciera estar atravesando una meseta final de usuarios correspondientes a la mayoría tardía en el mercado estadounidense. La CTA estima en base a la encuesta realizada, por primera vez de manera digital, que la penetración en hogares es de 78 millones, por lo que se considera que cada hogar posee aproximadamente en promedio 1,6 dispositivos tablet.



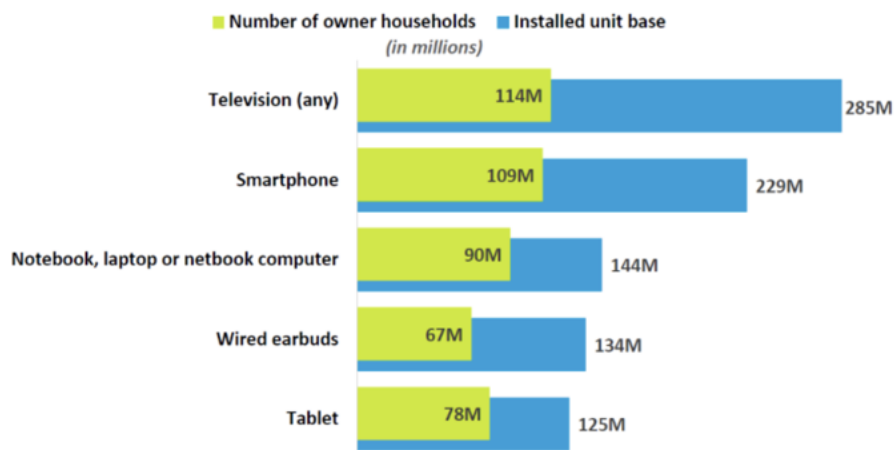


Figura 28. Base instalada de dispositivos sobre penetración en hogares 2019. Fuente: CTA. <https://www.cta.tech/News/Press-Releases/2019/May/Americans-Adopt-AI-Smart-Speakers-See-Largest-Gain.aspx>

Un estudio elaborado a finales del año 2018 por el Centro de Investigación Pew Research en EE.UU. estima que la penetración de tablets en ese país, sobre la población adulta, se ubica entorno al 53% (Pew Research, 2018).

Una investigación llevada a cabo por la consultora Daniel Research Group, utilizando como fuente de información las unidades de tablets comercializadas, indica que la base instalada de tablets en EE.UU. se encuentra cercana a los 194 millones de unidades (70 millones más que las estimaciones realizadas por CTA). Si bien la base de cálculo para realizar la estimación difiere sobre el estudio realizado de penetración de tablets ya que analiza población en lugar de hogares, se puede observar que la curva de adopción al igual que en hogares comienza a aplanarse dando lugar a una tasa de adopción menor luego de haber alcanzado el grupo de usuarios correspondientes a la mayoría tardía (DRG, 2019).

En cuanto a las tendencias y cortes demográficos que facilitaron la adopción de tablets en EE.UU. se encuentran los grupos familiares con niños pequeños, hogares con ingresos superiores a 75.000 U\$S dólares al año, graduados del colegio secundario y grupos etarios entre los 35 y 44 años de edad.

Para efectuar el análisis de la situación de adopción de tablets en el mercado argentino se consultaron diversas fuentes. En principio se obtuvieron estadísticas acerca de la venta del

dispositivo proveniente de la Cámara Argentina de Multimedia, Ofimática Comunicaciones y Afines (CAMOCA). Por otro se obtuvieron datos de hogares conectados a internet pertenecientes a ENACOM e INDEC para el mismo período.

Para la construcción del análisis comparativo se tomó el mismo ratio de dispositivos por hogares utilizado para la adopción en EE.UU. que es de 1,6 dispositivos tablet por hogar.

En ambos casos se observa una curva de adopción con forma de “S”, con una velocidad de adopción mayor en EE.UU., mientras que la etapa de adoptantes tempranos tomo sólo un año en el mercado de estadounidense, en el argentino se extendió por casi 3 años más. En la etapa de adopción del conjunto considerado Mayoría Temprana ambas curvas de adopción tuvieron una performance similar de duración, con aproximadamente 3 años de diferimiento en el tiempo entre un mercado y otro.

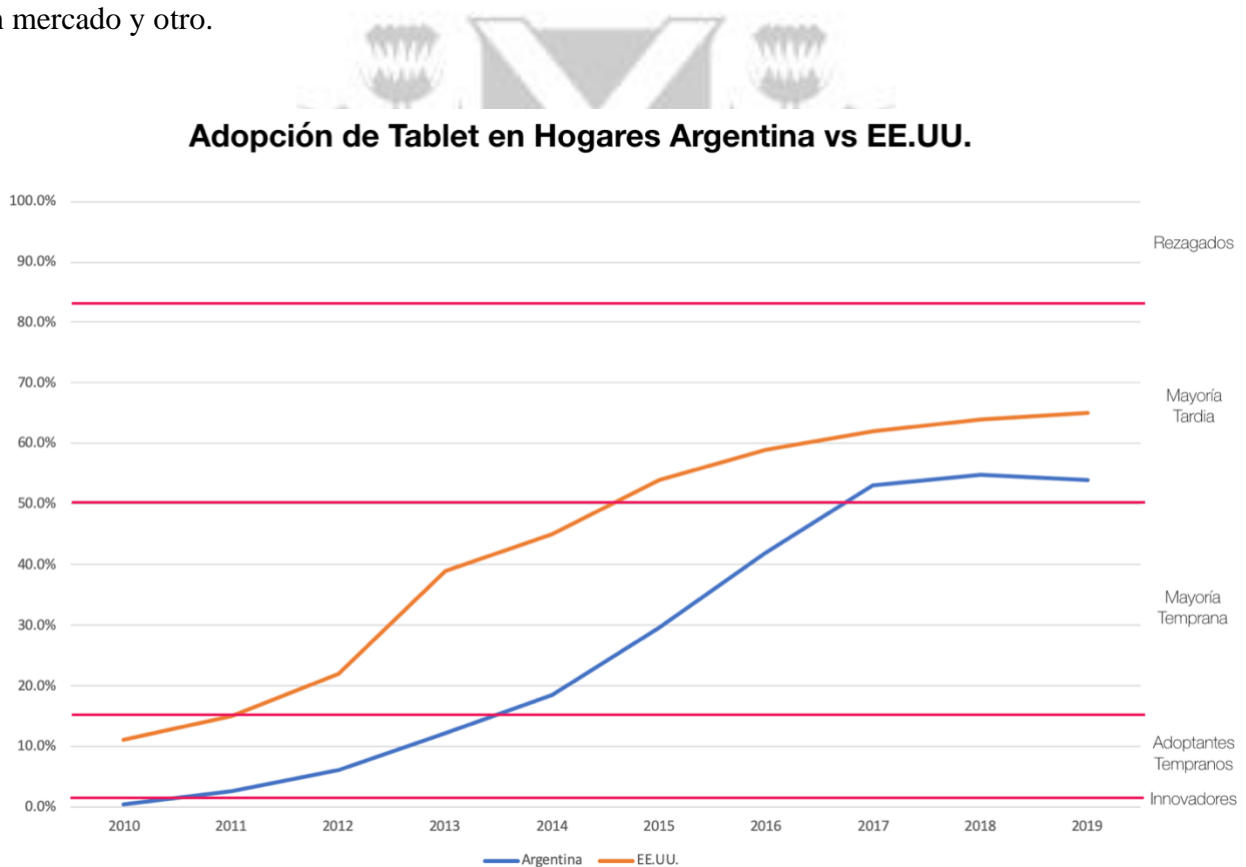


Figura 29. Adopción de tablets en Argentina vs EE.UU. 2010-2019. Fuente: CTA, DRG, CAMOCA, INDEC y ENACOM. Elaboración propia.

En esta fase de adopción también se puede observar que la pendiente de la curva correspondiente al mercado argentino tiene una derivada más agresiva, no tanto en forma de “S”, lo cual denota una tasa de crecimiento mayor.

Un análisis efectuado por la consultora Enrique Carrier y Asociados indica que las tablets han podido encontrar un lugar de crecimiento en el mercado argentino, tanto en los hogares como en las empresas.

Por un lado, adultos (generalmente mayores de 40 años) que las eligen por su mayor tamaño respecto de un smartphone (sobre todo los de gama media y baja), lo que las hace más cómodas para el consumo de contenidos y para operar aplicaciones que requieren múltiples funcionalidades. Por el otro lado, los más chicos, niños y preadolescentes que aún no tienen un smartphone medianamente moderno. En estos casos suele ser sustituto de la PC, siendo un dispositivo más económico, pero con amplias capacidades para convertirse en consola de videojuegos y también como en un centro para el consumo de contenidos audiovisuales (Por ejemplo: Youtube y Netflix).

También por su aptitud para usos laborales y profesionales dada su mayor capacidad para la creación de contenidos, produjo que las tablets fueran un complemento para los ámbitos corporativos (Carrier, 2015).

Siguiendo las tendencias y los lineamientos de adopción destacados en EE.UU. dentro del mercado argentino, las tablets también tuvieron una penetración mayor a medida que sube el nivel socioeconómico de los grupos familiares, al mismo tiempo no sorprende que su mayor adopción se encuentre también en los hogares con hijos menores, ya que aquí se encuentran los segmentos etarios más propensos a tener una tablet, los adultos y los niños y preadolescentes. Un indicador interesante relativo a la tendencia de adopción es que se trata de un dispositivo que gana adhesión en la medida en que ya hay otros equipos conectados en el hogar, llegando a un 50% de penetración en hogares donde ya existen 3 o más dispositivos conectados (Carrier, 2014).

### 7.3 – Evolución y adopción de televisores inteligentes en Estados Unidos y Argentina.

Un televisor inteligente (TI) es un televisor digital, es esencialmente una computadora con conexión a internet y con capacidad de almacenamiento especializada para la recepción de transmisión y reproducción de contenido audiovisual. Los televisores inteligentes están

disponibles como productos independientes, pero algunos modelos de televisores previos al TI también se pueden convertir a "inteligentes" a través de decodificadores que permiten funciones avanzadas. Además de las funciones tradicionales de los televisores, los decodificadores pueden proporcionar TV por Internet, medios interactivos en línea, contenido over-the-top (OTT), medios de transmisión bajo demanda y acceso a redes inalámbricas de internet.

Los TI son productos de actualización tecnológica, los mismos son el reemplazo natural de distintos tipos de dispositivos destinados al entretenimiento. El televisor es un producto que ya ha atravesado su curva de adopción y se encuentra en un proceso de recambio tecnológico por la adquisición de TI y de adaptadores (Apple TV, Amazon Fire, Roku, ChromeCast, etc.) que se utilizan como suplementos para la transmisión y reproducción de contenido multimedia (SMP por sus siglas en inglés) y mediante los cuales permite convertir antiguos modelos de TV, como LED o LCD, en televisores que pueden acceder a aplicaciones a través de su conexión a internet .

La consultora Activate Consulting estima que en los EE.UU. existen 198 millones de usuarios con televisores conectados a internet, con un crecimiento interanual promedio de 5% (Activate, 2020).

### Number of connected TV users in the United States from 2016 to 2023 (in millions)

Number of connected TV users in the U.S. 2017-2023

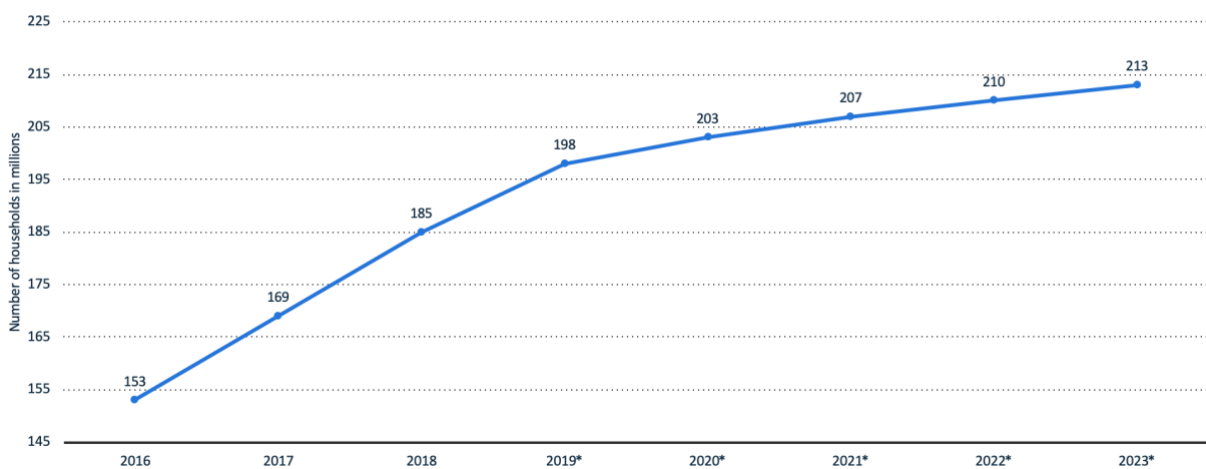


Figura 30. Evolución cantidad de usuarios de TV conectados a internet 2015-2023.

Fuente: Activate Consulting. <https://www.slideshare.net/ActivateInc/activate-technology-media-outlook-2020-185417815>

Un informe de CTA predice que los TI constituirán el 74% de todos los televisores nuevos vendidos en 2020, un 4% más que el total estimado para 2019 (CTA, 2019).

El informe destaca que se comercializan en la actualidad más televisores inteligentes por año que SMP, a pesar de que este último es una compra mucho más económica para el consumidor. CTA sugiere que los consumidores prefieren soluciones integradas, como lo demuestran las estrategias de Roku y Amazon Fire, los cuales se encuentran haciendo acuerdos para integrar sus dispositivos SMP en TI de ciertos fabricantes. CTA estima que en el año 2019 se vendieron 19,9 millones de SMP en los EE. UU., un 9% más que en 2018. Se espera que las ventas volverán a aumentar en el 2020, para llegar a 20,7 millones (un 4% más). Además, CTA proyecta que poco más de la mitad de los hogares de EE. UU. Tendrá un SMP para finales de este año (CTA, 2019).

Un análisis efectuado por la consulta Nscreenmedia.com, ampliando el concepto acerca de los factores que influyen en la compra de un TI, indica que la funcionalidad integrada del SMP dentro del TI es simplemente uno de los factores que inciden en la adquisición. Las personas deciden comprar por otros motivos como por ejemplo el precio, la calidad de la imagen, la marca y el tamaño de la pantalla, los cuales se consideran más influyentes en los procesos de compra. Eso no quiere decir que la gente no quiera la función inteligente. Simplemente no es un factor decisivo en el momento de la compra.

El estudio advierte también que la calidad de la imagen de alta definición (Ultra HD) se está convirtiendo en un importante impulsor de la adopción del TI. UHD es ahora un requisito crítico para la calidad de imagen. CTA espera que los televisores UHD representen casi la mitad de las ventas de televisores en 2020. Prácticamente todos los televisores UHD también incluyen la funcionalidad de TI. El informe agrega que, los fabricantes de TI son conocidos por poner fin al soporte de sus modelos de televisores inteligentes más antiguos. Tan pronto como un usuario no pueda encontrar el servicio que desea o una nueva versión de una aplicación en su TI por problemas de actualización de software, se estima que el usuario en primera instancia comprará un SMP para poder hacer uso de la aplicación (Nsreenmedia, 2019).

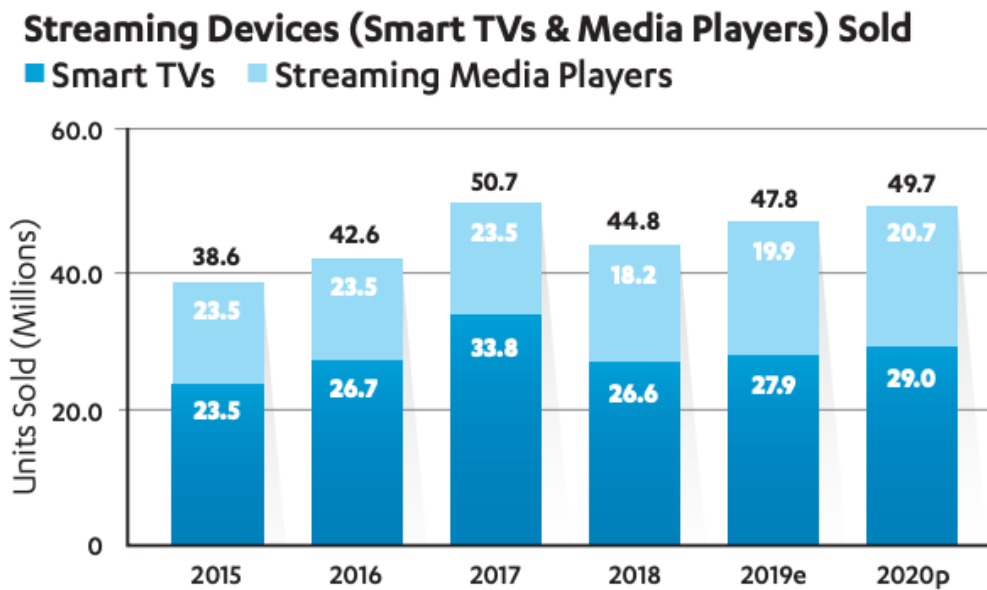


Figura 31. Evolución de ventas SMP y TI 2015-2019. Fuente: CTA. [https://cdn.ces.tech/ces/media/pdfs/white-papers/2020-cta-variety\\_8-tech-trends-entertainment.pdf](https://cdn.ces.tech/ces/media/pdfs/white-papers/2020-cta-variety_8-tech-trends-entertainment.pdf)

La consultora de investigación de mercado Enrique Carrier y Asociados subraya, en concordancia con CTA, que en Argentina la relevancia de contar con aplicaciones nativas integradas en los TI, en especial Netflix y Youtube, simplificaron el acceso a los contenidos dando lugar a la expansión de la base de usuarios potenciales.

La investigación indica que la popularización de los TI (y sus aplicaciones) ayudaron a que los servicios OTT terminaran de masificarse, ofreciendo la conveniencia del consumo online combinado con la calidad de una gran pantalla y una experiencia de uso similar al de la TV tradicional. En la actualidad, el TI es el dispositivo más utilizado para acceder a contenido audiovisual vía internet, siendo elegido por el 71% de quienes consumen servicios OTT en Argentina.

Con respecto a la curva de adopción de los TI en el mercado argentino podemos apreciar que su despliegue comenzó con 3 años de demora en comparación al mercado estadounidense, pero con una velocidad de crecimiento más agresiva que permitió alcanzar la mayoría tardía prácticamente al mismo tiempo a finales del año 2016.

A diferencia del análisis efectuado para el mercado de las tablets se aprecia una curva más plana, con una progresión de tendencia más lineal a diferencia de lo que podemos observar en forma de “S” para las tablets. Como ya hemos mencionado, en el caso de los TI, nos encontramos evaluando un producto con una tecnología que ya fue adoptada prácticamente por la totalidad de los hogares y que se encuentra impactada durante el período de análisis por el recambio tecnológico de dispositivos con funcionalidades ampliadas, el aumento de la velocidad de los servicios que proveen internet, el crecimiento de las plataformas que transmiten contenido sobre la red en formato de suscripción (OTT) y la posibilidad de adquirir productos complementarios y más económicos, como el SMP, que permite emular las principales características de los TI.

### Adopción de Smart Tv en Hogares Argentina vs EE.UU.

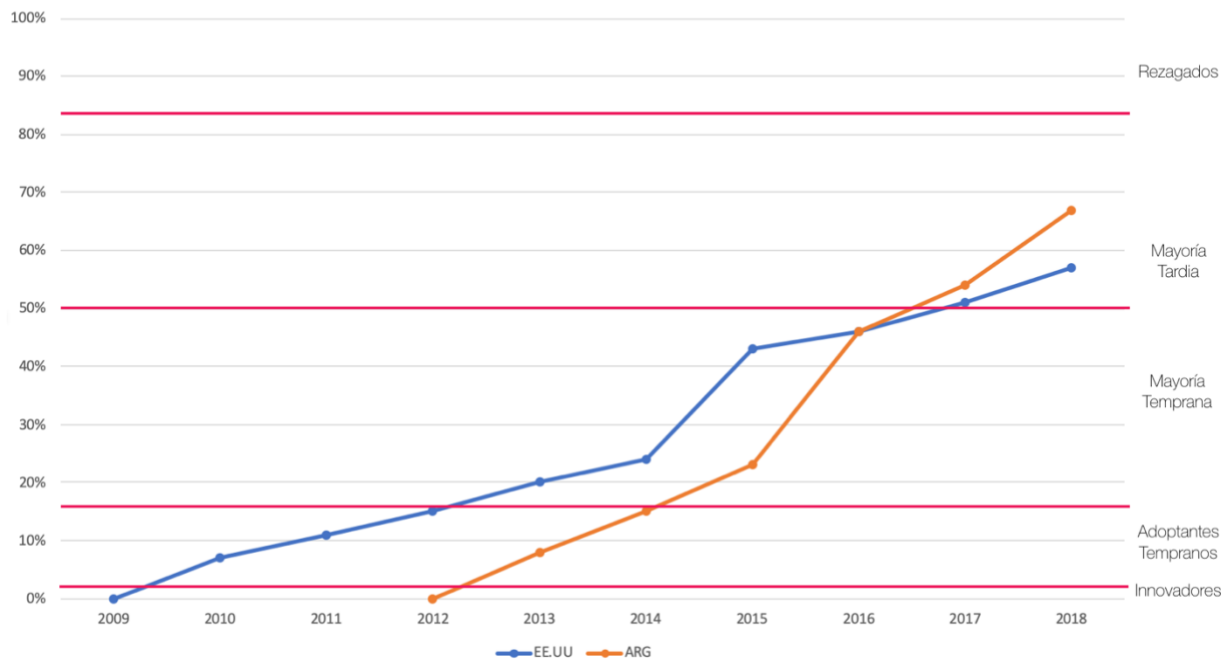


Figura 32. Adopción de TI en Argentina vs EE.UU. 2009-2018. Fuente: CTA, IHS, Enrique Carrier y Asociados, Despachante de Aduana. Elaboración propia.

#### 7.4 – Proyección y estimaciones de adopción de altavoces inteligentes para el mercado argentino.

Las proyecciones del Instituto Nacional de Estadísticas y Censos (INDEC), utilizando como base el último censo realizado en el año 2010, estima para el año 2019 una población en el orden de los 45 millones de habitantes, los cuales se agrupan aproximadamente en 10 millones de hogares en el total del país (INDEC, 2010).

El Ente Nacional de Comunicaciones (ENACOM) y el INDEC reflejan que en la actualidad existen 66% de hogares conectados a internet a través de un servicio de banda ancha o abono de línea digital asimétrica (ADSL por sus siglas en inglés).

En relación con los ingresos promedio de la población, el informe de ingresos del INDEC entrega datos por ingreso individual, total familiar y per cápita familiar, para la población total, la población ocupada y los hogares. En el último informe del INDEC que midió el ingreso por hogar se precisó que el 60% de los hogares logra reunir hasta 738 U\$S dólares estadounidenses (\$ 46.500 pesos argentinos) por mes, durante el tercer trimestre del año 2019. Con una canasta básica situada en 544 U\$S dólares estadounidenses (\$ 34.270 pesos argentinos) por mes, aproximadamente el 50% de los hogares no llega a cubrir la canasta básica establecida por el INDEC (INDEC 2018).

Para la Comisión Económica para América Latina y el Caribe (CEPAL) los estratos de ingresos medios son definidos como aquellos cuyos ingresos per cápita varían entre 1,8 veces y 10 veces la línea de pobreza. En el caso de la Argentina, donde la línea de pobreza está determinada en 200 U\$S dólares estadounidenses, el segmento medio estaría conformado por personas con un ingreso per cápita entre 360 U\$S dólares estadounidenses y 2000 U\$S dólares estadounidenses. Dentro de este espectro se puede distinguir tres niveles: un estrato medio-bajo, con ingresos per cápita de 360 U\$S a 600 U\$S; un estrato medio-intermedio, de 601 U\$S a 1200 U\$S; y medio-alto, de 1201 U\$S a 2001 U\$S. Y quienes perciben ingresos mayores a esa cifra serían considerados según el organismo como parte de la clase alta (CEPAL, 2019).

La consultora privada W estima que la clase media Argentina agrupa el 45 % de la población y que este índice no ha fluctuado demasiado desde el 2007. La construcción de las clases sociales establecida por la consultora tiene una lógica europea, la misma se encuentra basada en la capacidad de comprensión cultural. Es por esto que los dos valores que más se ponderan en la medición y construcción de las clases sociales son el nivel educativo y el nivel de ocupación o



empleo. En el mismo sentido que la CEPAL, la consultora establece que dentro de la clase media se incluyen varios estratos. Una clase media alta que tiene un ingreso promedio por familia de 1270 U\$S netos mensuales y que se extiende hasta los 2063 U\$S; y una clase media baja que tiene ingresos de hasta 666 U\$S mensuales (W, 2019).

De acuerdo a la consultora, la clase alta comienza en una línea muy similar a la que postula el INDEC en el decil más rico de la población ocupada, cuyo intervalo contempla ingresos mensuales a partir de 1507 U\$S dólares hasta 31.746 U\$S por mes. En esa línea, la consultora estima que sólo un 5% de hogares se considera de clase alta, mientras que un 20% de hogares se encuentran por debajo la línea de la pobreza y un 30% se considera de clase baja.

Estos datos son suficientes para estimar el mercado potencial de adopción de altavoces inteligentes en Argentina. En el mismo estableceremos la premisa que solo hogares de clase media intermedia, de media alta y clase alta son potenciales adoptantes de esta tecnología. Consideramos que este grupo de hogares ya se encuentran con una conexión fija a internet y además posee otros dispositivos complementarios como tablet y/o televisor inteligente.

En la última encuesta de hogares realizada por el INDEC se proyecta que el 50% de los hogares tienen un ingreso superior a 600 U\$S, aproximadamente 5 millones de hogares tiene la capacidad de lograr este ingreso. Estratificando la población según el ingreso neto por hogares estimado por el INDEC y utilizando la clasificación de poder adquisitivo definida por la CEPAL podemos establecer la siguiente relación Ingreso - Clase Social - Cantidad de Hogares:

<b>Clase Social (CEPAL)</b>	<b>Hogares</b>	<b>% Total Hogares</b>
Clase Media Intermedia – 601 / 1200 U\$S	3.200.000	32%
Clase Media Alta – 1200 / 1999 U\$S	1.300.000	13%
Clase Alta > 2000 U\$S	500.000	5 %

*Figura 33. Cuadro de estimación de hogares por clase social en Argentina 2020. Fuente: CEPAL, INDEC, Consultora W. Elaboración propia.*

En el informe realizado por el INDEC en el 2018 acerca del consumo realizado por los hogares se puede determinar que la estructura de los hogares tiene la siguiente relación porcentual,

9.8% corresponde a viviendas unipersonales menor de 65 años, el 7.1% a unipersonal mayor a 65 años, 14.4% a parejas sin hijos, 49.9% a parejas con hijos y 18.8% a hogares extendidos, lo cual implica en este último caso que en el hogar se encuentran habitantes que no guardan una relación conyugal o familiar entre sí (INDEC, 2018).

Sobre este escenario consideramos que el grupo habitacional correspondiente a viviendas unipersonales menores a 65 años, los hogares de parejas sin y con niños tienen mayor probabilidad de adoptar tempranamente un dispositivo según la tendencia evidenciada en la adopción de tablets en Argentina, mientras que hogares configurados como unipersonales mayores a 65 años tiene una probabilidad de adopción sobre el producto menor o tendiente a cero, correspondiendo su mayoría al grupo identificado por E. Rogers como rezagados.

Así mismo el conjunto de hogares conformados por mayores a 65 años, demuestra en la investigación índices de consumo en relación al gasto del hogar menores a los grupos restantes sobre ítems relacionados a *Equipos Audiovisuales, Fotográficos y de Procesamiento de Información y Efectos Personales* de los rubros *Recreación y Cultura y Bienes y Servicios Varios* respectivamente (INDEC, 2018).

A continuación hemos proyectado tres escenarios posibles contemplando una situación optimista de adopción, un escenario de condiciones normales y por último un ejercicio con un contexto desfavorable y una lenta evolución en la tasa de adopción.

En los tres escenarios se estipuló un comienzo de adopción posterior al ocurrido en el mercado estadounidense, si bien hay una base instalada de altavoces inteligentes en Argentina, la misma no representa cifras significativas ya que en la actualidad no existen datos que puedan representar el porcentaje de penetración en hogares. En este caso estipulamos la premisa de partir de un escenario de penetración del 3% de los hogares.

En la figura 34 se describe el escenario de adopción de altavoces inteligentes en EE.UU. y su esperada proyección para los años siguiente. El escenario optimista de Argentina simula una tasa de adopción similar a la que atraviesa hasta el momento el mercado estadounidense, y al mismo tiempo una tasa de crecimiento que emula la pendiente de la curva de adopción. El escenario intenta representar una evolución de adopción similar a la que las proyecciones que indican atravesará el producto para el conjunto de la *Mayoría Temprana*.

El segundo escenario de estimación intenta emular la misma tasa de adopción atravesada por las tablet con un CAGR de 67% y un período de aproximadamente 9 años para alcanzar a conquistar el conjunto de la *Mayoría Temprana*.

Por último, nuestro escenario más pesimista refleja un período 11 años para la adopción de la *Mayoría Temprana* con la menor tasa de adopción y un CAGR de 53%.

Cabe destacar que la mayoría de los expertos entrevistados coinciden en un escenario de adopción masivo en el mediano plazo de 3 a 5 años, lo cual ubicaría el pronóstico establecido entre las estimaciones del escenario #1 (optimista) y el #2 (normal).

### Adopción de Altavoces Inteligentes en Hogares Argentina vs EE.UU.

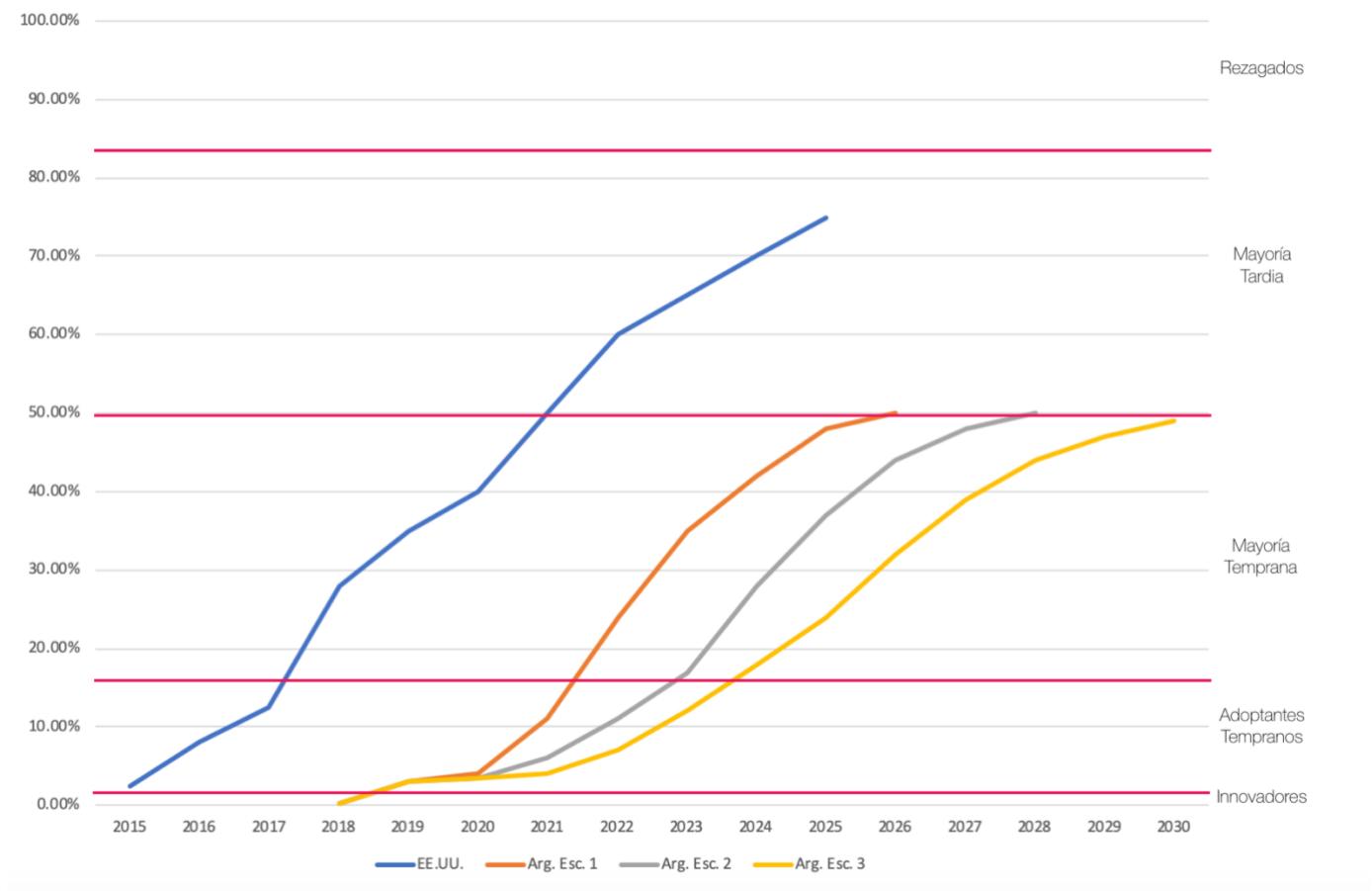


Figura 34. Estimación de escenarios para la adopción de altavoces inteligentes en Argentina vs EE.UU. 2015-2030. Fuente: Loup Ventures para EE.UU. y estimaciones propias para Argentina. Elaboración propia.

## 8. Capítulo VIII – Conclusiones.

### 8.1 – Introducción.

En el presente estudio de investigación hemos desarrollado un análisis profundo acerca de la situación actual en la que se encuentra la adopción de asistentes virtuales inteligentes a través de la comercialización de altavoces inteligentes. Durante el mismo fuimos conceptualizando y desarrollando los siguientes capítulos.

En el capítulo I hemos planteado los objetivos de la investigación y descripto las justificaciones, actuales problemáticas e implicancias que la adopción de los IVA ha comenzado a despertar en el mundo y principalmente en aquellos adoptantes tempranos que comienzan a incrementar día a día la funcionalidad de estos dispositivos.

En el capítulo II y III establecimos los lineamientos de nuestro trabajo de investigación, su alcance y su correspondiente marco teórico donde estipulamos las variables y teorías que creímos convenientes establecer y de las cuales nos basamos para poder conducir y llevar adelante un análisis estructural del trabajo de campo realizado.

Durante el capítulo IV detallamos en esencia el producto que estamos analizando, como logra su funcionamiento, cuales son sus elementos claves y servicios principales y también realizamos una reseña acerca de como sus principales características han ido evolucionando en el tiempo.

En el desarrollo del capítulo V describimos los principales productos que en la actualidad se están comercializando, sus estrategias, cuales son sus ventajas y desventajas y realizamos un análisis comparativo identificando las principales dimensiones de análisis para su evaluación.

En el capítulo VI analizamos la situación del crecimiento comercial ocurrido en el mercado global y en el mercado de EE.UU. en cuanto a la comercialización de altavoces inteligentes y los principales servicios que las estadísticas reflejan acerca de su utilización por los usuarios.

Finalmente, en el capítulo VII investigamos como fue la situación y evolución de adopción de productos tecnológicos similares, como la computadora tablet y el televisor inteligente, en el mercado argentino. Establecimos las variables demográficas que permiten identificar el conjunto de la población como mercado potencial a abastecer y realizamos las estimaciones de adopción considerando tres tipos de escenarios a largo plazo.

## 8.2 – Conclusiones.

El desarrollo de plataformas de inteligencia artificial ha traído aparejada la construcción de nuevos tipos de servicios e interfaces disponibles para los seres humanos, los cuales de alguna manera intentan simplificar las interacciones y los procesos utilizados por los dispositivos comúnmente conocidos hasta ahora como los teléfonos inteligentes, computadoras personales y tablets.

Parte de la visión que estas importantes empresas tecnológica del mundo intentan plasmar detrás de los desarrollos de las plataformas de inteligencia artificial es poder lograr la ubicuidad de los asistentes virtuales inteligentes en la cotidianeidad de nuestra actividades. Además que los mismos reflejen, a través del aprendizaje, un conocimiento y una personalización de su servicio con la capacidad, entre otras cosas, de tomar decisiones autónomas sobre las tareas rutinarias de nuestras vidas de un modo mucho más eficiente y conveniente de la que normalmente realizaríamos de manera propia. Para lograr este objetivo será necesario que los IVA no solo estén presentes en nuestro hogar a través de un altavoz con conexión a internet, también será importante incorporarlos en otros espacios de nuestra vida como el auto, el teléfono móvil y el entorno laboral.

En este sentido los IVA embebidos en altavoces han logrado simplificar ciertas actividades que antes requerían de varias interacciones manuales y táctiles en los dispositivos a sólo poder ejecutarlas con el habla y el reconocimiento de la voz. Este avance tecnológico requiere por sobre todas las cosas, un cambio de hábito y comportamiento por parte del usuario. Es importante considerar que el cambio en sí mismo no es menor y no es meramente evolutivo, como lo fue el paso del teléfono celular al teléfono inteligente, sino que en este caso se requiere la adaptación, preparación y utilización de otros sentidos del ser humano, como el habla y la escucha, que anteriormente no eran involucrados en la interacción con la tecnología para la realización y ejecución de funciones.

Así mismo el avance de nuevos productos denominados “inteligentes” que se desarrollan tanto para los hogares como para el día a día de nuestras actividades y, que en conjunto con el progreso de internet de las cosas (IOT), han ido generando la necesidad de contar con un único dispositivo que pueda lograr la concentración e integración de la ejecución de comandos de las distintas funcionalidades de una manera sencilla y confortable para el usuario.

Claramente podemos escindir el altavoz inteligente en un producto físico por el cual se canaliza la transmisión de la comunicación, el cual no varía sus características y prestaciones, de la plataforma de inteligencia artificial a la cual se conecta. La misma sí se encuentra constantemente evolucionando, ya sea aprendiendo el comportamiento y la manera de comunicarse con cada usuario, como así también incorporando nuevas funcionalidades o habilidades que permiten ir direccionando una mayor cantidad de transacciones en su utilización.

Como anticipamos en el desarrollo del marco teórico del cap. III efectuar el análisis de la situación de adopción utilizando fundamentalmente la teoría de difusión de innovación (DOI) que nos aporta el marco conceptual acerca del contexto sobre el cual está inmersa la tecnología en análisis (Rogers, 2003). De esta manera fue posible identificar que tanto en el maduro mercado estadounidense como en el relevamiento global, el grupo considerado *Mayoría Temprana* se encuentra en pleno desarrollo y aceptación de la tecnología.

Así mismo, siguiendo la línea de investigación planteada en el marco teórico, hemos podido establecer que los productos que en la actualidad lideran la participación en la comercialización del mercado han logrado atravesar la teoría del “abismo” elaborado por R. Moore como una extensión válida de la teoría de difusión de innovaciones (Moore, 1991). Por el cual se identifica que los productos que tienen un alto impacto de innovación en lo que respecta al comportamiento y hábito del usuario necesitan cruzar el abismo que se produce entre el conjunto de consumidores identificados por E. Rogers como *Adoptantes Tempranos* y el grupo de la *Mayoría Temprana*. Este hito no sólo significa que grupos de clientes con diferentes características demográficas y psicográficas comienzan a aceptar y adquirir el dispositivo, si no también que las políticas y estrategias de marketing utilizadas para educar y conquistar mercados masivos están surtiendo efecto positivo.

Desde su lanzamiento en el año 2014 la adopción tecnológica de los altavoces inteligentes ha sido intensa, sobretodo en los últimos dos años (2018 / 2019), y las proyecciones de distintas consultoras especializadas estiman que posiblemente su comercialización se incremente en el futuro cercano. Especialmente en el mercado estadounidense es donde el producto ha logrado índices de penetración en hogares y en individuos que marcan incrementos en las tasas de adopción de la curva E. Rogers y que ha logrado sobrevivir al abismo ilustrado por G. Moore como el salto de productos tecnológicos de un mercado inicial (*Early market*) al de un mercado principal (*Mainstream market*) el cual puede llevar a su adopción masiva.

Otro aspecto que introdujimos en el marco teórico fue la posibilidad de comprender que características del producto y formas de utilización pudieran ser predictores para una futura adopción. Sobre este camino decidimos trabajar con la teoría de aceptación tecnológica (TAM por sus siglas en inglés) (Davis, 1985). La *Facilidad de Uso Percibida* y la *Utilidad Percibida* por el usuario son las dimensiones que tienen mayor peso e injerencia sobre la actitud y la intención de comportamiento hacia la utilización de una tecnología en particular. En este aspecto analizamos las últimas investigaciones de mercado realizadas por grandes consultoras y consultoras especializadas en la temática de usuarios de altavoces inteligentes para comprender si las dimensiones de impacto en el uso de la tecnología se encuentran alineadas con los descriptores de adopción.

Un aspecto interesante que se destacan en las encuestas es la reducción de las interacciones diarias con el dispositivo en el mercado con la más alta penetración como es el estadounidense. Como advertimos anteriormente, una causa probable puede ser atribuida el hecho de haber llegado al conjunto de usuarios considerados *Mayoría Temprana*, los cuales requieren una curva de aprendizaje mayor en la utilización del producto y donde probablemente suceda que la cantidad de dispositivos inteligentes a integrar dentro del hogar también sea menor. De todas formas, sin dudas es un indicador que, sin dudas, debemos seguir en el corto plazo de cerca, ya que su constante decremento futuro podría comenzar a develar un déficit en la dimensión de la *Utilidad Percibida*.

Otro concepto para resaltar dentro de los estudios analizados es lo que los usuarios entienden como características más valoradas de los altavoces inteligentes y cuales son las principales funcionalidades utilizadas por los mismos. Podemos apreciar que todavía, entre el gran abanico de opciones disponibles para ejecutar, la mayoría de los usuarios opta por la utilización de habilidades simples como por ejemplo solicitar información, reproducir música, consultar el tiempo o el estado de tránsito de una ruta. Los usuarios valoran, sobre otras características cuan preciso puede ser el altavoz al momento de comprender una instrucción, su calidad de sonido y la cantidad de habilidades que el IVA posee. Podemos interpretar de estos resultados que todavía existe un extenso terreno de funciones donde el IVA puede comenzar a entregar valor, aumentando la confianza con el usuario y en consecuencia generando impacto positivo sobre la *Utilidad Percibida* y la *Intención de Comportamiento*.

### 8.3 – Preguntas de investigación.

Al comienzo del trabajo de investigación nos planteamos las siguientes hipótesis que intentaremos responder a continuación:

*¿Las principales barreras para lograr la masividad de los asistentes virtuales inteligentes en el mercado argentino están dadas por el precio en moneda extranjera del producto, la ausencia de entrenamiento de la plataforma sobre el modo y uso del lenguaje coloquial argentino y la falta de complementariedad actual que existe con otros dispositivos para interactuar en el rol de concentrador para la ejecución de tareas en los hogares inteligentes?*

Las principales barreras que han surgido del análisis efectuado y que podemos identificar para la adopción de los altavoces inteligentes en el mercado argentino son las siguientes:

**Idioma:** Los dispositivos que en la actualidad tienen mayor preponderancia en el mercado son los producidos por empresas estadounidenses y en consecuencia es el idioma inglés el que mayor tiempo y cantidad de iteraciones tiene para que los IVA puedan mejorar su aprendizaje en relación al mensaje que intenta transmitir el usuario. El aprendizaje del asistente virtual sobre la lengua castellana y su correspondiente coloquial argentino será crítico para que los usuarios puedan tener una experiencia acorde a las expectativas sobre la percepción de utilidad del producto.

**Complementariedad:** Una característica importante que tiene el altavoz inteligente es la posibilidad de ejecutar el rol de concentrador de funciones sobre otros dispositivos considerados “inteligentes”. El término “inteligente” se atribuye a aquellos artefactos que incluyen una serie de sensores, los cuales envían información a una plataforma digital para que el usuario pueda interactuar y ejecutar distintas operaciones sobre el producto. Si el avance de la adopción de estos productos conectados a internet que dan lugar a los denominados “hogares inteligentes” no se realiza o se difiere mucho en el tiempo el altavoz inteligente perderá una de sus funciones clave para lo cual puede ser adquirido.



**Privacidad:** Más del 40% de los propietarios de altavoces inteligentes han demostrado tener algún tipo de preocupación sobre la escucha pasiva del dispositivo y el almacenamiento de la información en la plataforma del IVA. Los fabricantes deben poder mejorar sus políticas de privacidad de acuerdo a normas internacionales de protección de datos personales y establecer términos de servicio transparentes para la utilización de los datos cuya finalidad sea establecer un servicio de asistencia virtual más personalizado y confeccionado a medida de las preferencias de cada usuario.

**Distribución:** Si bien el comercio electrónico crece año a año en la Argentina, será imprescindible para los fabricantes de los dispositivos estar presente en los distintos puntos de ventas del país. Establecer las alianzas necesarias para estructurar los canales de comercialización minoristas en tiendas de venta de electrodomésticos y equipos electrónicos, como así también en supermercados, será importante para que los potenciales compradores tengan su primer contacto y puedan conocer e interactuar con el dispositivo.

**Pantalla y Cámara:** Durante la producción del presente estudio de investigación los fabricantes introdujeron al mercado una nueva versión del producto de altavoz inteligente, el cual incluye una cámara y una pantalla táctil donde se pueden visualizar los pedidos de información o noticias, como así también las distintas peticiones que se realizan al IVA. Esta nueva característica aportará un nuevo abanico de habilidades y en consecuencia otorgará mayor versatilidad al producto y a la utilización de la plataforma de inteligencia artificial. Sin embargo, esta nueva característica traerá aparejado nuevos desafíos para la plataforma de inteligencia artificial, ya que se dispondrá de toda una nueva dimensión de información en imágenes que podría dañar el contrato de privacidad de información que se encuentra estipulado.

**Escenario Económico:** Si bien los precios iniciales de comercialización de los altavoces inteligentes son relativamente bajos en comparación a otros dispositivos como los smartphones o las tablets, la inestabilidad económica y la incertidumbre del contexto país a mediano y largo plazo produce que la decisión de compra de determinado tipo de bienes se postergue en el tiempo afectado los tiempos de adopción de esta tecnología.

**Escenario Político:** Será imprescindible contar con políticas claras en término de regulación de importaciones de productos que no pueden ser fabricados localmente en el país. Cualquier nueva regulación que limite la importación de productos electrónicos relacionado a altavoces inteligentes irá en detrimento de la posible adopción de la tecnología.

*ii) ¿Cuáles son los factores claves y los cambios necesarios que deben ocurrir para que el mercado de hogares argentino adopte masivamente la tecnología de asistentes virtuales inteligentes?*

En primer lugar, desde nuestro análisis se desprende que es clave poder implementar y estructurar los canales de distribución minoristas y lograr la presencia de los productos en tiendas especializadas de ventas de electrónica, cadenas de retail y electrodomésticos y grandes supermercados.

Si bien las ventas electrónicas siguen creciendo año tras año en el país, el año 2019 tuvo un incremento del 76% en facturación, 22% en productos o ítems y un 18% en órdenes de compra en relación al 2018 (CACE, 2020) la tendencia también muestra que cada vez más los compradores se mueven entre el mundo online y offline para concretar la compra, frecuentemente se realizan consultas en la web antes de ir a la tienda a comprar el producto. En la actualidad el comercio electrónico no supera el 3% de la totalidad de las ventas de comercios minoristas en Argentina, mientras que el porcentaje del comercio electrónico en EE.UU. es del 12%.

Otro aspecto que toma relevancia para la adopción de los IVA es la adecuación de los canales digitales de venta y atención de las principales plataformas digitales del país. Contar con la posibilidad de interactuar con las plataformas de comercio electrónico, atención al cliente de servicios públicos, transporte, turismo, servicios financieros, entretenimiento, logística y salud serán indispensables para incrementar la transaccionalidad sobre la utilización del asistente y en consecuencia poder hacer crecer el ecosistema que pueda acercar una real experiencia personalizada al usuario. En ese sentido entendemos que no basta con sólo poder establecer un recordatorio acerca del pago de una cuota del colegio, de la factura de un servicio de internet o de los productos faltantes en el hogar para la próxima compra, sino que la experiencia será mucho más enriquecedora cuando el IVA pueda conectarse a nuestro banco, billetera electrónica o portal de supermercado y realizar la transacción sin la intervención del usuario.

Uno de los denominadores comunes entre las implementaciones exitosas de la tecnología de voz es la capacidad de reducir la fricción para el cliente. La fricción representa el tiempo y el esfuerzo que un usuario dedica a realizar una tarea. Entonces, si se presenta la oportunidad de hacer que una tarea específica sea más eficiente a través de la tecnología de voz en comparación con el método ya establecido, las empresas deberían aprovecharla. Para ello es necesario que las empresas dediquen esfuerzos en el desarrollo de habilidades específicas para su organización en las distintas plataformas de inteligencia artificial disponibles, lo que deja abierta la pregunta a investigaciones futuras sobre el impacto y la preponderancia que estos canales de atención tendrán sobre la experiencia del cliente y como las compañías que logren desarrollar estos canales podrán usufructuar y monetizar su relación con el cliente de una manera más cercana.

#### 8.4 – Reflexiones.

Por el momento no está claro el horizonte de capacidades que las plataformas de inteligencia artificial pueden realizar a través de un formato lo más parecido a un mayordomo o asistente personal, pero sin embargo entendemos que no parecería descabellado plantear que parte de la economía digital que hoy es factible realizar predominantemente a través de un *smartphone* estará migrando en un corto plazo a un formato de plataforma de inteligencia artificial, donde pequeños actos rutinarios de decisión como elegir la mejor ruta para ir al trabajo o identificar las compras semanales necesarias para el hogar recaerán en un asistente virtual que en la actualidad ya tiene las capacidades para encargarse de las mismas.

Por otro lado, la sofisticación del producto sobre el cual estamos analizando la industria permite distintos grados de utilización del mismo y por lo tanto requiere determinada educación por parte de los usuarios, la cual debe ir actualizándose periódicamente a medida que nuevas funcionalidades se van incorporando a la plataforma. Es por esto que consideramos que existe un efecto tardío en el avance de la utilización de funcionalidades más complejas del asistente virtual en relación al avance de la penetración de la tecnología en los hogares.

A medida que los usuarios comiencen a elevar el tiempo de uso y conviertan a estos dispositivos como sus principales opciones de búsqueda de información, consulta de sitios web y precios de producto o lo utilicen como principal hub concentrador dentro de su hogar para ejecutar

acciones como escuchar música, bajar la temperatura del aire acondicionado u ordenar una pizza, hará que posiblemente ciertas compañías dueñas de plataformas que en la actualidad se encuentran liderando mercados digitales se vean amenazadas y en consecuencia deban repensar sus estrategias. Algunas de las implicancias sobre las posibles industrias que pueden ser afectadas:

- Publicidad: ¿La producción de publicidad deberá incorporar un nuevo formato y adaptarse a la publicidad con un estilo más acorde a la radio o al guion de cine?
- Entretenimiento: ¿Podrá Spotify o Netflix mantener el dominio de esta industria digital o habrá lugar para nuevos jugadores?
- Comercio Electrónico: Nuevos consumidores utilizando un asistente virtual para las compras de productos y servicios. ¿Nace el “Conversational o Voice Commerce”?
- CRM Software: ¿Se encuentran las empresas preparadas dentro de sus departamentos de Inteligencia Comercial para interpretar necesidades y acciones a través de la voz del cliente?
- Smartphone: ¿Seguirá utilizándose el dispositivo dentro del hogar?

En el anexo I se realiza una evaluación de la competitividad de la industria que estamos investigando. Del mismo se desprenden algunas consideraciones importantes que debemos tener en cuenta al momento de establecer posibles pronósticos acerca de la evolución de la industria:

- Nos encontramos ante una industria con alta rivalidad competitiva por parte de los actuales incumbentes, los cuales en la búsqueda de economías de escala y ventas cruzadas según sus distintos ecosistemas de negocio generan diferentes estrategias agresivas de comercialización.
- Se requieren de grandes inversiones y economías de escala para participar en el mercado.
- Las barreras de entrada son altas debido a la sofisticación del producto y al alto costo de cambio al cual se expone el usuario.
- Debemos considerar que el alcance de la industria que estamos evaluando no trata únicamente acerca de la producción y consumo de bienes electrónicos. Es una industria que tiene un mercado que se encuentra evolucionando y madurando en el entendimiento de los beneficios que este tipo de producto puede traer. Por lo tanto se torna complejo establecer si los usuarios van

a obtener las prestaciones acordes al precio al cual se encuentran dispuestos a pagar por el producto.

- Si bien la parte más importante del producto, que es la utilización de la plataforma de asistente virtual inteligente, se encuentra completamente subsidiada, todavía no está clara la propuesta de valor final que esta tendría en el futuro y de qué manera la misma podría generar ingresos para los distintos competidores.

- Dadas estas circunstancias entendemos que la ventaja competitiva en la industria de altavoces inteligentes está dada por el capital humano e intelectual y las habilidades que los competidores pueden derivar de las mismas y a través de las cuales pueden desarrollar distintas prestaciones y funcionalidades para los dispositivos.

C. Shapiro y H. Variant, en su ensayo acerca de la guerra de estándares, definen estas son cruciales para el éxito de las empresas y sus negocios cuando se detectan fuertes efectos de red que producen altos costos de compatibilidad entre los productos para los usuarios. Siguiendo los finales posibles de una guerra de estándares, los mismos pueden finalizar o bien en una tregua entre los competidores, en un mercado duopólico o en una lucha a muerte “Fight to death”.

En su estudio se pone énfasis en que las principales batallas “Fight to death” son cruciales cuando existen poderosas retroalimentaciones positivas derivadas de la generación de fuertes efectos de red (Shapiro & Variant, 1999). Si los actuales competidores comienzan a generar ingresos y conocimientos derivados de la intensificación proveniente de los efectos de red (Anexo II) por una mayor utilización de la plataforma de asistente virtual inteligente, nos encontraremos posiblemente en una situación de mercado donde suceda un escenario de guerra de plataformas en el formato “*Fight to death*” entre los principales competidores.

Los principales competidores en la industria, Amazon, Google y Apple tienen por detrás ecosistemas de negocio diferentes en los cuales en muchos casos se encuentran compitiendo en diferentes verticales de industria y hacia diferentes segmentos de clientes, pero como similitud existe la particularidad que tanto Amazon como Google son exitosos en mercados donde “*El ganador toma todo*”, por lo tanto, ambos competidores cuentan con experiencia en este tipo de enfrentamientos.

## Bibliografía.

- ACM. (1997). En: SIGCHI.org. Recuperado en:  
[https://web.archive.org/web/20140817165957/http://old.sigchi.org/cdg/cdg2.html#2\\_1](https://web.archive.org/web/20140817165957/http://old.sigchi.org/cdg/cdg2.html#2_1)
- Activate (2020). Recuperado en: <https://www.slideshare.net/ActivateInc/activate-technology-media-outlook-2020-185417815>
- Amazon Inc. (2019). En: Developer/Amazon.com . Recuperado en:  
<https://developer.amazon.com/en-US/docs/alexa/alexa-voice-service/api-overview.html>
- Amazon. (2017). Amazon Echo-Black, <https://www.amazon.com/Amazon-Echo-Bluetooth-Speaker-with-WiFi-Alexa/dp/B00X4WHP5E>
- Anusuya, M., Katty, S. (2009). Speech Recognition by Machine: A Review. (IJCSIS) International Journal of Computer Science and Information Security, Vol. 6, No. 3, 2009.
- Arthur, B. (1996). Increasing Returns and the Two Worlds of Business. Harvard Business Review, (July-August, 1996).
- Arts, J. W. C., Frambach, R. T., & Bijmolt, T. H. A. (2011). Generalizations on consumer innovation adoption: A meta-analysis on drivers of intention and behavior. *International Journal of Research in Marketing*, 28(2), 134–144.  
<https://doi.org/10.1016/j.ijresmar.2010.11.002>
- Bagozzi, R. P. (2007). The legacy of the technology acceptance model and a proposal for a paradigm shift. *Journal of the Association for Information Systems*, 8(4), 244–254.  
<https://doi.org/10.17705/1jais.00122>
- Barwise, J. (1977). An Introduction to First-Order Logic. *Studies in Logic and the Foundations of Mathematics*. [https://doi.org/10.1016/S0049-237X\(08\)71097-8](https://doi.org/10.1016/S0049-237X(08)71097-8)
- Baum, L. E., Petrie, T., Soules, G., & Weiss, N. (1970). A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains. *The Annals of Mathematical Statistics*.
- Bentley, F., Luvogt, C., Silverman, M., Wirasinghe, R., White, B., & Lottrjidge, D. (2018). Understanding the Long-Term Use of Smart Speaker Assistants. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(3), 1–24.  
<https://doi.org/10.1145/3264901>
- Best, W. (1988). *Cómo investigar en educación?*. Madrid: Morata, S.L.
- Bostic, K. (2013). “Nuance confirms its voice technology is behind Apple's Siri”. En *AppleInsider*”. Recuperado en: <https://appleinsider.com/articles/13/05/30/nuance-confirms-its-technology-is-behind-apples-siri>

- Brown, S., & Venkatesh, V. (2005). Model of a Adoption of Technology in Households : a Baseline Model Test and Extension Incorporating. *MIS Quarterly*, 29(3), 399–426. Retrieved from <http://www.vvenkatesh.com/wp-content/uploads/2015/11/Brown&VenkateshMISQ2005.pdf>
- Budzinski, O., Noskova, V., & xijie, zhang. (2019). The Brave New World of Digital Personal Assistants: Benefits and Challenges from an Economic Perspective. *SSRN Electronic Journal*, 24(118). <https://doi.org/10.2139/ssrn.3306169>
- Canalys (2019). Recuperado en: <https://www.canalys.com/newsroom/canalys-global-smart-speaker-installed-base-to-top-200-million-by-end-of-2019>
- Cambria, E., & White, B. (2014). Jumping NLP curves: A review of natural language processing research. *IEEE Computational Intelligence Magazine*, 9(2), 48–57. <https://doi.org/10.1109/MCI.2014.2307227>
- Carr, A (2015). “Following Fire Phone Flop, Big Changes At Amazon’s Lab126”. En FastCompany.com. Recuperado en: <https://www.fastcompany.com/3040383/following-fire-phone-flop-big-changes-at-amazons-lab126>
- Carrier (2014). “Tablets al podio”. En Enrique Carrier y Asociados. Recuperado en: <https://comentarios.info/index.php/2014/09/12/tablets-al-podio/>
- Carrier (2015). “Tablets en su lugar”. En Enrique Carrier y Asociados. Recuperado en: <https://comentarios.info/index.php/2015/08/21/tablets-en-su-lugar/>
- Carrier (2019). “Smart OTT”. En Enrique Carrier y Asociados. Recuperado en: <https://comentarios.info/index.php/2019/08/30/smart-ott/>
- CEPAL (2019). “Estudio Socioeconómico”. Recuperado en: [http://www.momarandu.com/notix/noticia/07241\\_argentina-el-dcil-ms-rico-tiene-un-ingreso-medio-de-87108-pesos-por-mes.htm](http://www.momarandu.com/notix/noticia/07241_argentina-el-dcil-ms-rico-tiene-un-ingreso-medio-de-87108-pesos-por-mes.htm)
- Chomsky, N. (1956). *Three models for the description of language*.
- Chung, H., Iorga, M., Voas, J., & Lee, S. (2017). Alexa, Can i Trust You? *Computer*, 50(9), 100–104. <https://doi.org/10.1109/MC.2017.3571053>
- Chung, H., Park, J., & Lee, S. (2017). Digital forensic approaches for Amazon Alexa ecosystem. *DFRWS 2017 USA - Proceedings of the 17th Annual DFRWS USA*, 22, S15–S25. <https://doi.org/10.1016/j.diin.2017.06.010>
- Colby, K. M. (1981). Modeling a paranoid mind. *Behavioral and Brain Sciences*, 4(4), 515–534. <https://doi.org/10.1017/S0140525X00000030>
- Collobert, R., & Weston, J. (2008). *A unified architecture for natural language processing*. 160–

167. <https://doi.org/10.1145/1390156.1390177>

Creswell, J. (1998). *Qualitative Inquiry and Research Design: Choosing among five traditions*. London and Thousand Oaks: Sage Publications.

Creswell, J. W. (2009). *Research design: Qualitative and mixed methods approaches*. London and Thousand Oaks: Sage Publications.

CTA (2018). “Voice Commerce over Smart Speakers”. Recuperado en: <https://www.cta.tech/News/i3/Articles/2018/September-October/Smart-Speakers-Shopping-Via-Voice.aspx>

CTA (2019). “8 Key Consumer Trend for Entertainment”. En CES Report. Recuperado en: [https://www.cta.tech/Resources/Newsroom/Media-Releases/2020/January/Consumer-Tech-U-S-Sales-to-Reach-Record-\\$422-B-\(1\)](https://www.cta.tech/Resources/Newsroom/Media-Releases/2020/January/Consumer-Tech-U-S-Sales-to-Reach-Record-$422-B-(1))

CTA (2019). “Consumer Tech U.S. Sales to Reach Record \$422 Billion in 2020”. En CES Report. Recuperado en: [https://cdn.ces.tech/ces/media/pdfs/white-papers/2020-cta-variety\\_8-tech-trends-entertainment.pdf](https://cdn.ces.tech/ces/media/pdfs/white-papers/2020-cta-variety_8-tech-trends-entertainment.pdf)

Davis, F. D. (1985). A technology acceptance model for empirically testing new end-user information systems: Theory and results [Dissertation]. *PhD Thesis - Massachusetts Institute of Technology*. [https://doi.org/10.1016/S0378-7206\(01\)00143-4](https://doi.org/10.1016/S0378-7206(01)00143-4)

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly: Management Information Systems*, 13(3), 319–339. <https://doi.org/10.2307/249008>

Davis, F. D., & Venkatesh, V. (2000). *Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies*. 46(April), 186–204. <https://doi.org/10.1287/mnsc.46.2.186.11926>

Denzin, N. K. (1989). *Strategies of Multiple Triangulation. The Research Act: A Theoretical Introduction to Sociological Methods*. Chicago: Aldine Publishing Company.

DGR (2019). “United States Personal Device History and Forecast, 1975-2023 November 2019 Update”. En Danielresearchgroup.com. Recuperado en: <http://www.danielresearchgroup.com/Portals/0/Downloads/DRG%20United%20States%20Personal%20Devices%201975-2023%20History%20and%20Forecast%20November%202019.pdf>

Epstein, J., & Klinkenberg, W. D. (2001). From Eliza to Internet: A brief history of computerized assessment. *Computers in Human Behavior*, 17(3), 295–314. [https://doi.org/10.1016/S0747-5632\(01\)00004-8](https://doi.org/10.1016/S0747-5632(01)00004-8)

Gatignon, H., & Robertson, T. S. (1985). *A Propositional Inventory for New Diffusion Research*.



*Journal of Consumer Research*, 11(3), 849–867.

Gartenberg, C. (2017). “Apple announces HomePod speaker to take on Sonos”. En The Verge. Recuperado en: <https://www.theverge.com/2017/6/5/15732144/apple-homepod-speaker-announced-siri-price-release-date-wwdc-2017>

Gartner (2016). “Amazon Echo, other smart speakers to rake in \$2B by 2020”. En CNET.com. Recuperado en: <https://www.cnet.com/news/amazon-echo-alexa-google-assistant-apple-smart-speaker-2-billion-2020-gartner/>

Goertzel, B. (2014). Artificial General Intelligence: Concept, State of the Art, and Future Prospects. *Journal of Artificial General Intelligence*, 5(1), 1–48. <https://doi.org/10.2478/jagi-2014-0001>

Google LLC. (2018). Google Duplex: An AI System for Accomplishing Real-World Tasks Over the Phone. En Google AI Blog. Recuperado de: <https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html>

Google LLC. (2019). Google.com. En Google/Product/Features. Recuperado de: [https://store.google.com/product/google\\_home](https://store.google.com/product/google_home)

Guzman, A. L. (2019). Voices in and of the machine: Source orientation toward mobile virtual assistants. *Computers in Human Behavior*, 90, 343–350. <https://doi.org/10.1016/j.chb.2018.08.009>

Hartmans, A. (2016). “Google hired writers from Pixar and The Onion to make Assistant more personable”. En Business Insider. Recuperado en: <https://www.businessinsider.com/google-assistant-pixar-the-onion-2016-10>

Hoeffler, S. (2003). *Hoeffler - 2003 - Measuring Preferences for Really New Products(2).pdf*. *XL(November)*, 406–420.

Hoy, M. B. (2018). Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants. *Medical Reference Services Quarterly*, 37(1), 81–88. <https://doi.org/10.1080/02763869.2018.1404391>

Huang, C., & Kao, Y. (2014). UTAUT2 Based Predictions of Factors Influencing the Technology Acceptance of Phablets by DNP Chi-Yo. *Automotive Engineer (London)*, 24(5), 34–36.

Huang, Xing. (2015). *Xing Huang VIRTUAL REALITY / AUGMENTED REALITY TECHNOLOGY : THE NEXT CHAPTER OF HUMAN-COMPUTER INTERACTION Xing Huang VIRTUAL REALITY / AUGMENTED REALITY TECHNOLOGY : THE NEXT CHAPTER OF HUMAN-*.

Huang, Xuedong, Baker, J., & Reddy, R. (2014). A historical perspective of speech recognition.

- Communications of the ACM*, 57(1), 94–103. <https://doi.org/10.1145/2500887>
- IHS (2018). “Smart Homes”. En IHS.com. Recuperado en :  
[https://news.ihsmarket.com/INFO/press\\_releases](https://news.ihsmarket.com/INFO/press_releases)
- INDEC (2018). “Consumo en Hogares”. En Indec.gov.ar. Recuperado en:  
[https://www.indec.gov.ar/ftp/cuadros/sociedad/engho\\_2017\\_2018\\_informe\\_gastos.pdf](https://www.indec.gov.ar/ftp/cuadros/sociedad/engho_2017_2018_informe_gastos.pdf)
- INDEC (2019). “Ingresos 3er trimestre”. En Indec.gov.ar. Recuperado en:  
[https://www.indec.gov.ar/uploads/informesdeprensa/ingresos\\_3trim190357631751.pdf](https://www.indec.gov.ar/uploads/informesdeprensa/ingresos_3trim190357631751.pdf)
- Indurkha, N., & Damerau, F. J. (2010). Handbook of natural language processing, second edition. In *Handbook of Natural Language Processing, Second Edition*.
- ITU (2019) Cantidad de suscripciones de conexiones fijas a internet en el mundo. Recuperado en: [https://www.itu.int/en/ITU-D/Statistics/Documents/statistics/2019/ITU\\_Key\\_2005-2019\\_ICT\\_data\\_with%20LDCs\\_28Oct2019\\_Final.xls](https://www.itu.int/en/ITU-D/Statistics/Documents/statistics/2019/ITU_Key_2005-2019_ICT_data_with%20LDCs_28Oct2019_Final.xls)
- ITU (2019). Número de suscripciones de banda ancha fija en los Estados Unidos de 2000 a 2018 (en millones). Recuperado en: [https://www.itu.int/en/ITU-D/Statistics/Documents/statistics/2019/Fixed\\_broadband\\_2000-2018\\_Dec2019.xls](https://www.itu.int/en/ITU-D/Statistics/Documents/statistics/2019/Fixed_broadband_2000-2018_Dec2019.xls)
- Kalantari, M. (2017). Consumers adoption of wearable technologies: literature review, synthesis, and future research agenda. *International Journal of Technology Marketing*, 12(1), 1. <https://doi.org/10.1504/ijtmkt.2017.10008634>
- Katz, M. L., & Shapiro, C. (1985). Network externalities, competition, and compatibility. *American Economic Review*, 75(3), 424–440. <https://doi.org/10.2307/1814809>
- Kim, K., & Shin, D. (2014). An acceptance model for smart watches. *Department of Interaction Science, Sungkyunkwan University, Seoul, Korea Abstract*. <https://doi.org/10.12740/APP/66485>
- Kowalczyk, P. (2018). Consumer acceptance of smart speakers: a mixed methods approach. *Journal of Research in Interactive Marketing*, 12(4), 418–431. <https://doi.org/10.1108/JRIM-01-2018-0022>
- Kurzweil, R. (2005). *The Singularity Is Near*.
- Lagace, M. (2019). “These Products and Services Work with Google Home“. En Android Central. Recuperado en: <https://www.androidcentral.com/these-products-and-services-work-google-home>
- Lardinois, F (2019). “Google launches new Assistant developer tools”. En TechCrunch. Recuperado en: <https://techcrunch.com/2019/05/07/google-launches-new-developer-tools-for-the-google-assistant/>

- Lardinois, F (2017). “Google Home can now recognize up to six voices and give personalized responses”. En TechCrunch. Recuperado en: <https://techcrunch.com/2017/04/20/google-home-can-now-recognize-up-to-six-voices-and-give-personalized-responses/>
- Lee, Paul; Stewart, Duncan; Loucks, Jeff; Arkenberg, C. (2019). Telecommunications Predictions 2019. *Deloitte Insights*.
- Legris, P., Ingham, J., & Collette, P. (2003). Why do people use information technology? A critical review of the technology acceptance model. *Information and Management*, 40(3), 191–204. [https://doi.org/10.1016/S0378-7206\(01\)00143-4](https://doi.org/10.1016/S0378-7206(01)00143-4)
- López, G., Quesada, L., & Guerrero, L. A. (2018). Alexa vs. Siri vs. Cortana vs. Google Assistant: A Comparison of Speech-Based Natural User Interfaces. *Advances in Intelligent Systems and Computing*, 592, 241–250. [https://doi.org/10.1007/978-3-319-60366-7\\_23](https://doi.org/10.1007/978-3-319-60366-7_23)
- Loup (2019) “Update Market Share” Fuente: Loup VC. Recuperado en: <https://loupventures.com/smart-speaker-market-share-update/>
- Loup VC (2018). “Annual Smart Speaker IQ Test”. En Loup VC Research. Recuperado en : <https://loupventures.com/annual-smart-speaker-iq-test/>
- Lynley, M (2016). “Google unveils Google Assistant, a virtual assistant that’s a big upgrade to Google Now”. En TechCrunch. Recuperado en: <https://techcrunch.com/2016/05/18/google-unveils-google-assistant-a-big-upgrade-to-google-now/>
- Manning, C. D., Schütze, H., & Weikurn, G. (2002). *Foundations of Statistical Natural Language Processing*.
- Markoff, J. (2008). “A Software Secretary That Takes Charge”. En The New York Times. Recuperado en: [https://www.nytimes.com/2008/12/14/business/14stream.html?\\_r=1](https://www.nytimes.com/2008/12/14/business/14stream.html?_r=1)
- Martin, E. (2017). How echo, Google home, and other voice assistants can change the game for content creators. *EContent*, 40(1).
- Martin, G. L. (1989). The utility of speech input in user-computer interfaces. *International Journal of Man-Machine Studies*, 30(4), 355–375. [https://doi.org/10.1016/S0020-7373\(89\)80023-9](https://doi.org/10.1016/S0020-7373(89)80023-9)
- Microsoft (2019). Voice Report. Recuperado en: [https://advertiseonbing-blob.azureedge.net/blob/bingads/media/insight/whitepapers/2019/04%20apr/voice-report/bingads\\_2019\\_voicereport.pdf](https://advertiseonbing-blob.azureedge.net/blob/bingads/media/insight/whitepapers/2019/04%20apr/voice-report/bingads_2019_voicereport.pdf)
- Moore, & Benbasat. (1991). Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation. *Information Systems Research*, Vol. 2, pp. 192–222.

- Moore, G. (1991). *Crossing the Chasm*.
- Murphy, S. (2018). “Growing up with Alexa: A child's relationship with Amazon's voice assistant”. En CNN Business. Recuperado en: <https://www.cnn.com/2018/10/16/tech/alexa-child-development/index.html>
- Naone, E. (2009). Intelligent software assistant. *Technology Review*, 112(2), 48.
- Nass, C., & Moon, Y. (2000). Mindfulness Theory and Social Issues - Machines and Mindlessness - Social Responses to Computers. *Journal of Social Issues : A Journal of the Society for the Psychological Studies of Social Issues*, 56(1), 81–103. Retrieved from <http://www.coli.uni-saarland.de/courses/agentinteraction/contents/papers/Nass00.pdf>
- Nieva, R. (2018). Exclusive: Google's Duplex could make Assistant the most lifelike AI yet. En CNET. Recuperado de: <https://www.cnet.com/news/google-assistant-duplex-at-io-could-become-the-most-lifelike-ai-voice-assistant-yet/>
- NPR (2019) “Porcentaje de frecuencia de utilización de altavoces inteligentes”. Recuperado en: <https://www.nationalpublicmedia.com/uploads/2020/01/The-Smart-Audio-Report-Winter-2019.pdf>
- Nscreenmedia (2019). “Smart TVs and SMPs jointly drive connected TV market”. En [nscreenmedia.com](https://nscreenmedia.com). Recuperado en: <https://nscreenmedia.com/smart-tvs-and-smps-jointly-drive-connected-tv-growth/>
- Onoda, S., Kato, S., & Mutoh, A. (2014). *The Effect of the Network Structure Differences on the Diffusion of Items*. 809–814. <https://doi.org/10.7551/978-0-262-32621-6-ch130>
- Park Associates (2019) Fuente: Parks Associates. Recuperado en: <http://www.parksassociates.com/blog/article/cus-2019-pr4>
- Pew Research (2018). “Internet, social media use and device ownership in U.S. have plateaued after years of growth”. Pew Research Center. Recuperado en: <https://www.pewresearch.org/fact-tank/2018/09/28/internet-social-media-use-and-device-ownership-in-u-s-have-plateaued-after-years-of-growth/>
- Poritz, A. B. (1988). *Hidden Markov Models: A Guided Tour*. <https://doi.org/10.1109/icassp.1988.196495>
- Porter, M. E. (1991). Towards a dynamic theory of strategy. *Strategic Management Journal*, 12(2 S), 95–117. <https://doi.org/10.1002/smj.4250121008>
- Purinton, A., Taft, J. G., Sannon, S., Bazarova, N. N., & Taylor, S. H. (2017). “Alexa is my new BFF”: Social roles, user satisfaction, and personification of the Amazon Echo. *Conference on Human Factors in Computing Systems - Proceedings, Part F1276*, 2853–2859. <https://doi.org/10.1145/3027063.3053246>

- Reddy, R. (1976). Speech Recognition by Machine: A Review. *IEEE Computational Intelligence Magazine Colloquium (Digest)*, 64(499), 129–134. <https://doi.org/10.2139/ssrn.3367658>
- Rese, A., Baier, D., Geyer-Schulz, A., & Schreiber, S. (2017). How augmented reality apps are accepted by consumers: A comparative analysis using scales and opinions. *Technological Forecasting and Social Change*, 124, 306–319. <https://doi.org/10.1016/j.techfore.2016.10.010>
- Rogers, E. (2003). *Diffusion of Innovations*.
- Rouhiainen, L. (2018). *Artificial Intelligence 101 Things You Must Know Today About Our Future*.
- Russell, S., & Norvig, P. (2016). Artificial Intelligence a Modern Approach. In *Book* (Vol. 53).
- San-Martín, S., López-Catalán, B., & Ramón-Jerónimo, M. A. (2013). Mobile Shoppers: Types, Drivers, and Impediments. *Journal of Organizational Computing and Electronic Commerce*, 23(4), 350–371. <https://doi.org/10.1080/10919392.2013.837793>
- Schank, R., & Abelson, R. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*.
- Schank, R. C. (1972). Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology*, 3(4), 552–631. [https://doi.org/10.1016/0010-0285\(72\)90022-9](https://doi.org/10.1016/0010-0285(72)90022-9)
- Searle, J. R. (1980). Minds and brains without programs. *Mindwaves*, 3, 1–19.
- Shapiro, C., & Variant, H. (1999). *Standards Wars*.
- Singh, B., Kapur, N., & Kaur, P. (2012). *Speech Recognition with Hidden Markov Model : A Review*. 2(3).
- Sinha, G., Shahi, R., & Shankar, M. (2010). Human Computer Interaction. *Proceedings - 3rd International Conference on Emerging Trends in Engineering and Technology, ICETET 2010*, 1–4. <https://doi.org/10.1109/ICETET.2010.85>
- Statt, N. (2016). “Apple to release macOS Sierra on September 20th”. En The Verge. Recuperado en: <https://www.theverge.com/2016/9/7/12758218/apple-mac-os-sierra-update-release-date-announced>
- Strategy Analytics (2019). Recuperado en: <https://voicebot.ai/2020/02/17/amazon-again-topped-q4-global-smart-speaker-sales-followed-by-google-and-baidu-according-to-strategy-analytics-smart-speaker-shipments-set-new-record/>
- Steele, B. (2017). “Google Assistant now helps with your shopping on Google Home (updated)”. En Engadget. Recuperado en: <https://www.engadget.com/2017/02/16/google-home-shopping-with-google-assistant/>

- Stucke, Maurice E., A. E. (2004). *How Digital Assistants Can Harm Our Economy, Privacy and Democracy*.
- Stucke, M. E., & Ezrachi, A. (2017). How Digital Assistants Can Harm our Economy, Privacy, and Democracy. In *Ssrn*. <https://doi.org/10.2139/ssrn.2957960>
- Taylor, M. E., Kuhlmann, G., & Stone, P. (2008). Transfer learning and intelligence: An argument and approach. *Frontiers in Artificial Intelligence and Applications*.
- Tornatzky, L. G., & Klein, K. J. (1982). Tornatzky\_Klein\_1982 Innovation Characteristics and Innovation Adoption-Implementation A Meta-Analysis of Findings. *Ieee Transactions on Engineering Management*, Vol. 29, pp. 28–43.
- Tractica (2016). “Virtual Digital Assistants”. En Tractica.com. Recuperado en: <https://tractica.omdia.com/wp-content/uploads/2016/08/VDA-16-Brochure.pdf>
- Treusch, P. (2018). *Re-reading ELIZA : Human – machine Interaction as Cognitive*. 32(94), 411–426.
- Trope, Y., & Liberman, N. (2003). Temporal Construal. *Psychological Review*, 110(3), 403–421. <https://doi.org/10.1037/0033-295X.110.3.403>
- Turing, A. M. (1950). Computing machinery and intelligence. In *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer*. [https://doi.org/10.1007/978-1-4020-6710-5\\_3](https://doi.org/10.1007/978-1-4020-6710-5_3)
- Turkle, S. (1984). The Second Self. In *Book*.
- Ueda, K., Nishino, N., & Takenaka, T. (2009). Producer decision-making in markets with network externalities. *CIRP Annals - Manufacturing Technology*, 58(1), 413–416. <https://doi.org/10.1016/j.cirp.2009.03.108>
- Van Raaij, E. M., & Schepers, J. J. L. (2008). The acceptance and use of a virtual learning environment in China. *Computers and Education*, 50(3), 838–852. <https://doi.org/10.1016/j.compedu.2006.09.001>
- Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences*, 39(2), 273–315. <https://doi.org/10.1111/j.1540-5915.2008.00192.x>
- Venkatesh, V., Morris, M., Davis, G., & Davis, F. (2003). *USER ACCEPTANCE OF INFORMATION TECHNOLOGY: TOWARD A UNIFIED VIEW*. 27(3).
- Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer Acceptance and Use of Information Technology. *MIS Quarterly*, 36(1), 157–178. Retrieved from [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2002388](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2002388)

- Voicebot.ai (2019) “Consumer Adoption Report January 2019” Recuperado en:  
[https://voicebot.ai/wp-content/uploads/2019/03/smart\\_speaker\\_consumer\\_adoption\\_report\\_2019.pdf](https://voicebot.ai/wp-content/uploads/2019/03/smart_speaker_consumer_adoption_report_2019.pdf)
- W (2019). “Estudio Socioeconómico” Recuperado en:  
[http://www.momarandu.com/notix/noticia/07241\\_argentina-el-dcil-ms-rico-tiene-un-ingreso-medio-de-87108-pesos-por-mes.htm](http://www.momarandu.com/notix/noticia/07241_argentina-el-dcil-ms-rico-tiene-un-ingreso-medio-de-87108-pesos-por-mes.htm)
- Weizenbaum, J. (1966). ELIZA—A Computer Program For the Study of Natural Language Communication Between Man And Machine. *Communications of the ACM*, 26(1), 23–28.  
<https://doi.org/10.1145/357980.357991>
- Westenberg, J. (2019). “The best Google Home services you should know about”. En Adroid Authority. Recuperado en: <https://www.androidauthority.com/google-home-services-749968/>
- Williams, M., Rana, N., Dwivedi, Y., & Lal, B. (2011). S UTAUT REALLY USED OR JUST CITED FOR THE SAKE OF IT? A SYSTEMATIC REVIEW OF CITATIONS OF UTAUT’ s. *European Conference on Information Systems (ECIS)*, 231. Retrieved from <http://aisel.aisnet.org/ecis2011/231>
- Winograd, T. (1971). *MIT AI Technical Report 235*.
- Winograd, T. (1973). *A procedural model of language understanding*.
- Yaniv, L.; Yossi, M. (2018). "Google Duplex: An AI System for Accomplishing Real-World Tasks Over the Phone". En Google AI Blog. Recuperado de:  
<https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html>
- Yang, H., Lee, H., & Zo, H. (2017). User acceptance of smart home services: An extension of the theory of planned behavior. *Industrial Management and Data Systems*, 117(1), 68–89.  
<https://doi.org/10.1108/IMDS-01-2016-0017>
- Yin, R. (1994) Case study research: design and methods. Segunda Edición. Thousand Oaks. SAGE. California. Estados Unidos.
- Zikmund, W. G., & Carr, G. (2000). Business Research Methods. 7th. USA, Dryden.

## Anexo I – Análisis de competitividad de la industria.

La publicación de M. Porter acerca de la teoría dinámica de la estrategia sostiene que el éxito de una firma deviene de una o una serie de ventajas competitivas las cuales finalmente derivan en el éxito de los resultados financiero sustentables de la misma. En consecuencia es una función que se da entre el atractivo que surge del análisis de la estructura de una industria y la posición relativa de competitividad que la empresa tenga en relación a esa industria (Porter, 1991).

Para comprender acerca de la estructura de la industria y su atractivo se realizará un análisis acerca de la situación competitiva en la industria de altavoces inteligentes utilizando el modelo de Porter y el análisis de las 5 fuerzas que interactúan en la búsqueda de una estrategia que pueda desarrollar una ventaja competitiva.

Referencia utilizada para describir las intensidades de las fuerzas.

Muy débil  Muy Fuerte

Poder de negociación de los compradores 

- Alta atomización de los compradores.
- Costos de cambio relativamente bajos para la adquisición de un nuevo dispositivo, existe un alto costo de entrenamiento del IVA. A medida que se utilizan más habilidades o skills, los costos de cambio y entrenamiento del nuevo dispositivo aumentan.

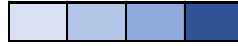
Poder de negociación de los proveedores 

- Alta concentración de los proveedores que suministran los insumos tecnológicos para la construcción de la plataforma del IVA, de todas formas, existe una capa importante de valor agregado producido por los incumbentes hacia el producto final en el desarrollo del software y entrenamiento de las habilidades de inteligencia artificial.
- Los proveedores que suministran los insumos para el dispositivo físico del altavoz se encuentran en un entorno más atomizado.



- Baja o escasa probabilidad de realizar una integración hacia adelante por parte de los proveedores de tecnología.

#### Amenaza de nuevos participantes



- Industria de alto capital intensivo para poder competir en lo referente a la tecnología para las plataformas de IVA.
- Nuevas amenazas posibles en el corto y mediano plazo se darán con el ingreso de jugadores chinos como Baidu y Alibaba a nivel global y la salida al mercado del smart speaker de Samsung con Bixby como su asistente virtual.
- Marcas de altavoces de alta fidelidad necesitan establecer alianzas con los principales incumbentes para incluir la plataforma de IVA en sus productos y de esa manera poder participar en las utilidades de la industria, pero únicamente en lo que respecta a ingresos por venta de dispositivos.
- Bajos costos de cambio para los dispositivos si los usuarios se mantienen dentro de la plataforma del IVA que se encontraban utilizando.
- Alta probabilidad de adquisición lateral por parte de los incumbentes hacia empresas que producen dispositivos de altavoces de alta fidelidad.

#### Amenaza de productos o servicios sustitutos



- El producto de Smart Speaker es muy difícil de sustituir si pensamos en la tecnología aplicada en el dispositivo. Esta tecnología requiere de una alta inversión de capital para obtener poder de cómputo para el procesamiento e interpretación del lenguaje natural y atraer talento que en la actualidad es muy escaso para desarrollar las soluciones que componen la plataforma de inteligencia artificial.
- Interesados más por la utilización de la plataforma que por la venta de los dispositivos, los dos principales competidores de la industria ya han lanzado sus versiones económicas con precios reducidos.

## Rivalidad entre competidores



- Alta rivalidad y competencia entre los principales competidores existentes. La misma se pone de manifiesto en varios escenarios;
  - En primer lugar, el escenario más visible es el que existe por las utilidades que genera la venta de dispositivos, a esto se suman fricciones entorno a las alianzas desarrolladas con los productores de altavoces de alta fidelidad, por Ej. Sonos ya se encuentra desarrollando su versión de producto con Google Home, y en el desarrollo de varias opciones de producto en cuanto a la calidad y funcionalidades para poder competir en todos los segmentos posibles.
  - En segundo lugar, pero más importante que el anterior, se encuentra el escenario que se da por la pelea en la atracción de usuarios a las distintas plataformas, es una pelea más silenciosa que se da puertas adentro en el desarrollo de las distintas habilidades que se pueden poner a disposición para ejecutar a través del IVA.

## Anexo II – Análisis de los efectos de red sobre la plataforma de asistencia virtual inteligente

Como hemos mencionado anteriormente y considerando a la plataforma que provee los servicios de asistente virtual inteligente como el verdadero core del producto y sobre la cual se generan las distintas estrategias de utilización y consumo, describiremos los posibles efectos de red indirectos que la plataforma puede generar (Stucke, Maurice E., 2004).

### I. Atracción de empresas de manufacturas y desarrolladores

Empresas que producen manufacturas y en especial aquellas relacionadas con la casa inteligente e internet de las cosas desarrollarán software y aplicaciones en sus productos para las plataformas donde se encuentren la mayor cantidad de usuarios. Lo mismo ocurre para los desarrolladores y programadores, los cuales incrementarán su número si la plataforma permanece abierta para que terceras puedan desarrollar software que pueda ser utilizado por más usuarios.

### II. Mejora en la calidad del producto

A medida que la plataforma gana tracción de uso entre los usuarios, al igual que ocurrió en su momento con los buscadores de internet y el algoritmo de búsqueda, el IVA se retroalimenta con mejoras constantes. La mejora en el procesamiento e interpretación del lenguaje natural sólo puede producirse con entrenamiento. Tomando como ejemplo el reconocimiento de la voz, mientras más usuarios utilicen el asistente, el mismo será capaz de aprender diferentes pronunciaciones, estructuras gramaticales y diferentes maneras de pedir la ejecución de una orden.

### III. Mayores datos y personalización de ofertas y búsquedas

El incremento en el uso de los asistentes virtuales trae consigo la generación de mayores sets de datos y por consiguiente mejoras en términos de conocimiento del usuario, preferencias de usos, entornos y contextos, hábitos de consumo y predicción de necesidades cuyo principal output está

relacionado a la personalización de ofertas que pueden llegar al usuario por la plataforma o por otro medio de comunicación.

#### IV. Ofertas de productos y publicidad

Por último, el efecto de atracción que se produce ante una masiva utilización de los asistentes virtuales es que muchos productos estén disponibles para la compra en un formato de Voice o Conversational Commerce, lo cual a su vez también atrae nuevos formatos de publicidad que desean incorporarse a la plataforma.



## Anexo III – Cuadros de relevamiento y estimaciones.

Adopción de tablets en Argentina vs EE.UU. 2010-2019. Fuente: CTA, DRG, CAMOCA, INDEC y ENACOM. Elaboración propia.

AÑO	Unidades CAMOCA	Acumulado Un.	Ratio Tablets por Hogar 1.6	Enacom HH Int	Argentina	EE.UU.
2010	28695	28695	17934	4725000	0.38%	11.0%
2011	175152	203847	127404	5197500	2.45%	15.0%
2012	354848	558695	349184	5717250	6.11%	22.0%
2013	635390	1194085	746303	6210000	12.02%	39.0%
2014	754397	1948482	1217801	6598000	18.46%	45.0%
2015	1351680	3300162	2062601	6984000	29.53%	54.0%
2016	1569768	4869930	3043706	7252000	41.97%	59.0%
2017	1815611	6685541	4178463	7870000	53.09%	62.0%
2018	759548	7445089	4653181	8474000	54.91%	64.0%
2019	455729	7900818	4938011	9165000	53.88%	65.0%

Adopción de Televisores Inteligentes en Argentina vs EE.UU. 2009-2018. Fuente: CTA, IHS, Enrique Carrier y Asociados, Despachante de Aduana.

Año	EE.UU	Argentina
2009	0%	-
2010	7%	-
2011	11%	-
2012	15%	-
2013	20%	8%
2014	24%	15%
2015	43%	23%
2016	46%	46%
2017	51%	54%
2018	57%	67%

Estimación de escenarios para la adopción de altavoces inteligentes en Argentina vs EE.UU. 2015-2030. Fuente: Loup Ventures para EE.UU. y estimaciones propias para Argentina.

Año	EE.UU.	Arg. Esc. 1	Arg. Esc. 2	Arg. Esc. 3
2015	2.50%			
2016	8.00%			
2017	12.50%			
2018	28.00%	0.30%	0.30%	0.30%
2019	35.00%	3.00%	3.00%	3.00%
2020	40.00%	4.00%	3.50%	3.50%
2021	50.00%	11.00%	6.00%	4.00%
2022	60.00%	24.00%	11.00%	7.00%
2023	65.00%	35.00%	17.00%	12.00%
2024	70.00%	42.00%	28.00%	18.00%
2025	75.00%	48.00%	37.00%	24.00%
2026		50.00%	44.00%	32.00%
2027			48.00%	39.00%
2028			50.00%	44.00%
2029				47.00%
2030				49.00%
<b>CAGR</b>	<b>41%</b>	<b>89.6%</b>	<b>66.8%</b>	<b>52.9%</b>

## Anexo IV – Entrevistas.

Para realizar el presente estudio de investigación se llevo adelante una serie de entrevistas estructuradas con referentes profesionales y académicos del mercado de las TIC. El cuestionario fue enviado por correo electrónico a los participantes y los mismos optaron por el tipo de respuesta, ya sea escrita o en formato de entrevista por conferencia. A continuación, presentamos el cuestionario, el perfil de los entrevistados y sus respuestas:

### Preguntas:

- i) ¿Cuáles considera pueden ser las principales barreras o impedimentos en el mercado argentino para lograr la adopción de asistentes virtuales inteligentes como consecuencia de la comercialización de altavoces inteligentes?
- ii) ¿Qué cambios de índole regulatorio, tecnológicos, funcionales de producto u otros a su entender considera necesariamente deben ocurrir para que la adopción de los altavoces inteligentes logre masividad en el mercado argentino?
- iii) Según su perspectiva, ¿Qué horizonte de tiempo estima razonable para la adopción masiva de altavoces inteligentes en el mercado argentino?
  1. Corto plazo: 1 o 3 años
  2. Mediano plazo: 3 a 5 años
  3. Largo plazo: + de 5 años

### Respuestas:

#### **Juan Gnius:**

*Director de Telracom. Especialista en la industria de telecomunicaciones para América Latina y el Caribe.*

- i) En principio encuentro dos impedimentos principales para la comercialización de altavoces digitales y la consecuente adopción de asistentes digitales. En primer lugar, el conocimiento que

puede tener el consumidor sobre estos dispositivos, su utilidad y funcionamiento, entre otros puntos. En segundo lugar, la existencia de dispositivos hogareños que puedan conectarse a los altavoces inteligentes para ser comandados por estos. Si bien los consumidores que tienen smartphones pueden estar acostumbrados al uso de la interfaz de voz para buscar o pedirle cosas al teléfono inteligente, considero que la capacidad principal de los altavoces inteligentes es comandar otros dispositivos. Y en este punto puede haber falencias en los hogares.

En otras palabras, si usamos un altavoz inteligente para escuchar música y hacerle preguntas que se le pueden hacer al teléfono es una subutilización del dispositivo.

El precio, en principio, no me parece una barrera para el público al que va dirigido.

ii) Deberíamos definir masividad para este tipo de dispositivos. En principio no veo impedimentos regulatorios para su uso, ya que en todo caso le corresponderían las normas que tratan sobre privacidad de los datos, etc. De existir regulación específica, iría en desmedro del uso de altavoces inteligentes.

Refiero a mi respuesta anterior para una mayor adopción de estos dispositivos en la existencia de un ecosistema de otros gadgets que puedan ser comandados por los altavoces. En teoría pueden hacer de todo: encender luces, poner música, controlar la temperatura del hogar, levantar persianas y mil cosas más que no tiene caso mencionarlas aquí. Ahora bien, estas mil cosas más deben poder conectarse a los altavoces.

Por otro lado, algo que debería suceder para una mayor adopción de altavoces inteligentes son los canales de venta y/o distribución de estos productos. Esta situación la padecen gadgets como relojes inteligentes o pulseras del tipo Fitbit. Los operadores móviles no parecen tener interés en la promoción de estos dispositivos y ellos cuentan con una gran capilaridad de puntos de venta; los retail (Garbarino, Frávega, etc.) tampoco.



iii) Nuevamente hay que definir masividad. Te cambio el concepto por popularidad, en el sentido que la gente sepa de qué se tratan estas cosas. Masivo y popular hoy son los smartphones, no creo que los altavoces inteligentes alcancen esta masividad y popularidad.

En este punto sí no puedo dejar de lado la realidad económica de Argentina que ya afecta, por ejemplo, el tiempo de recambio de smartphones y notebooks.

En este escenario, un horizonte razonable puede ir de 3 a 5 años.

**Ignacio Perrone:**

*Director de Investigación en Tecnologías de la Información y las Comunicaciones en Frost & Sullivan.*

i) Hay varias barreras:

- 1) Idioma: algunos asistentes (como el de Amazon) están sólo en inglés
- 2) Precio: todavía en la Argentina son caros, en especial en relación con las funcionalidades que prestan y si se quieren obtener algunas funcionalidades (como apagar luces, etc), hay que invertir todavía más dinero (ver siguiente)
- 3) Funcionalidades limitadas: en muchos casos el sistema no puede interactuar con versiones locales de otros sistemas (como diarios o radios locales, o portales de compra electrónica), por lo que no todas las funcionalidades se pueden aprovechar.
- 4) Desconfianza del consumidor: dado el relativo atraso del comercio electrónico, no todo el mundo confía en hacer compras con estos asistentes y con tan poca intervención humana, la percepción del riesgo todavía es alta.
- 5) Privacidad: aunque no es una gran barrera, algunas personas que podrían ser clientes de este producto están al tanto de posibles filtraciones de información (aunque no esté confirmado, el rumor alcanza para disuadir muchas veces).

ii) Lo principal sería atender los puntos mencionados antes (del 1 al 5). No creo que se necesite una regulación especial, los proveedores deberían poder resolverlo.

iii) Depende mucho de qué se entienda por masivo. No lo tengo confirmado, pero creo que en USA la penetración está en torno al 40% de los hogares. Yo diría que eso es muy difícil de lograr en Argentina, por diversas razones. De todos modos, podría llegarse a un 20% pero en el largo plazo (más de 5 años).

**Sebastián Novoa:**

*Analista Sr. De Mercado de Dispositivos de Consumo Comercial.*

i) Creo que la única barrera hoy en día tal vez sea la interpretación de ciertos modismos y el entendimiento sobre cosas nuestras.

ii) A nivel regulatorio, no le estoy encontrando un freno, ahora igualmente con el efecto COVID la comercialización de estos productos está siendo impactada.

iii) Para mi es de 1 a 3, más para 2 a 3 y 5 años sería demasiado gap de tiempo.

**Enrique Hofman:**

*Director de la Maestría en Tecnología y Negocios de la Universidad de San Andrés. Profesor de Gestión de las Operaciones y Gestión de Servicios dentro de la Escuela de Administración y Negocios.*

i) Yo no veo mayor diferencia en el mercado argentino que cualquier otro mercado salvo que tengas una crisis macroeconómica en algún momento, una problemática de compras o de importaciones, no veo porque el mercado argentino deberá ser muy distinto al brasilero que norteamericano no lo veo si hay alguna diferencia estimó que va hacer una situación macroeconómica pasajera.

No veo algún impedimento especial en la Argentina.

Con respecto de los parlantes, tienen que ver con muchas cosas, cuan han afinado está el reconocimiento de voz porque eso varía en los distintos idiomas, vocabulario, las frases, etc.

Creo que los asistentes van a tener más éxito en la medida que estén ligados a internet de las cosas y que vigilen tu latido cardíaco, el jabón del lavarropas, lo veo más relacionado con eso, que con

las preguntas que uno le puede hacer o que sea un asistente, un conserje que te contesta cualquier cosa y todo lo sabe o poner música en Spotify.

Creo que cuando recibas tu lavarropas Samsung que tiene conexión a internet, va a estar ligado a Alexa y eso va a dar un impulso y desde el punto de vista en lo personal médico puede ser para tu compañía de seguros o para tu prepaga donde informe los latidos cardíacos, cuántos pasos caminaste hoy, todo eso lo veo más conectado en el ecosistema de la casa con personas y aparatos. Alexa es un caballo de Troya que quiere dominar el hogar más que interactuar con los humanos desde mi punto de vista.

ii) Desde mi punto de vista yo no incluiría ninguna regulación, nada más que vos tengas que aceptar las condiciones como cuando tenes una tarjeta Visa, leer la letra chica y los términos de servicio. Internet se manejó muy bien en forma anárquica estos 15 años y creo que todos los pretextos para regular son para proteger ciertas industrias. Hay una guerra entre Europa y Estados Unidos donde Europa es el ambiente más regulador y Estados Unidos es el menos regulador.

Creo que discutir hoy en día un Android o Alexa o Siri no tiene sentido. En el único lugar del mundo que las cosas son un poco distintas es en China porque hay asistentes locales.

No tiene que haber marco regulatorio, para eso hay un contrato de adhesión en donde vos sabés los pros y los contras y si quisieras te sumas o no. Respecto del hackeo y de la confidencialidad y de la seguridad, si existe un delito penal alguien deberá ir a la cárcel.

iii) Entiendo que cuando se instale será explosivo como los celulares, creo que de 3 a 5 años va a tener entre un 10% y un 15% de adopción, y luego será cuestión de poco tiempo para que se dispare masivamente.

El crecimiento estará más ligado al control de los aparatos y al control de la salud humana. Para mí el altavoz inteligente tiene dos funciones, una de internet de las cosas y la otra es la función de conserje, la función de conserje. Considero será indispensable en la vida de acá a 5 años.

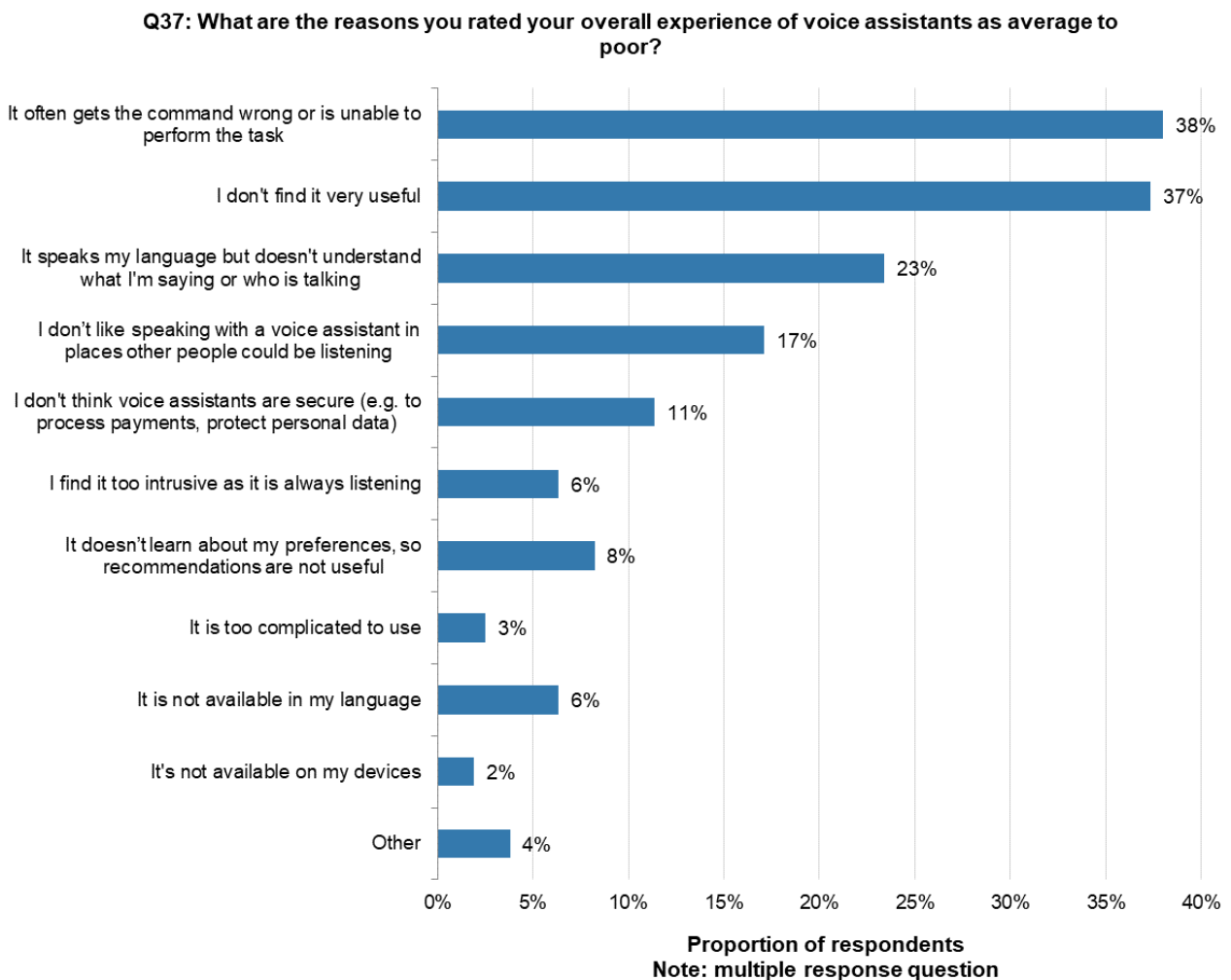
**Sonia Agnese:**

*Analista Senior para Latino América de Omdia (Omdia incluye Ovum, IHS Markit technology research, Heavy Reading y Tractica)*

i) Adjunto este estudio de relevancia

Estudio Omdia sobre problemas detectados de consumidores digitales en Argentina con los asistentes de voz (noviembre 2019)

Como verás de las respuestas, los usuarios mencionan como principales impedimentos la falta de utilidad o que la experiencia que han tenido no ha sido muy buena (no hace la tarea o no entiende el lenguaje)



Source: Digital Consumer Insights Omdia 2019 Latin America

ii) Se requieren mejoras en temas de reconocimiento de voz como primer punto. Aún falta que entiendan correctamente lo que se le dice.

Entiendo que el principal tema es la evolución tecnológica del servicio.

iii) Adopción masiva a mediano plazo. De 3 a 5 años

**Juan de la Cruz Ramirez Silva:**

*Gerente Comercial de Frávega (Celulares -IT-Audio-Gaming-Accesorios-Servicios (Garantías, Asistencias, Seguros, Instalaciones y otros)*

i) La principal barrera es la económica. La valuación del dólar actualmente es elevada e incrementa el precio del producto. Esto genera que no haya interés por parte del cliente y en consecuencia que no se produzca la oferta y los distribuidores no los están ingresando al país.

ii) No hay desarrolladas en el país interfaces con otros dispositivos inteligente que pueden generar desarrollo de hogares inteligentes. Las funcionalidades quedan reducidas y son prácticamente las mismas que se pueden realizar de un smartphone o altavoz estándar sin la característica de inteligencia.

iii) Me imagino mediano plazo, 3 a 5 años y dependerá en fuerte medida del precio y las funcionalidades asociadas que sin dudas en otros países se encuentran más avanzadas.

**Alan Lerner:**

*Head of HR & Organizational Transformation | Director - Advisory, Consulting @ KPMG Argentina. Profesor a tiempo parcial con especial interés en Gestión de Operaciones, Estrategia, Gestión de Procesos, Sistemas de Información de Gestión, Controles Internos y Estrategia en la Universidad de San Andrés (UdeSA), Universidad del Centro de Estudios Macroeconómicos de Argentina (UCEMA), Universidad Argentina de la Empresa (UADE) y FCE-UBA.*

i) En mi opinión tienen varios ejes: 1) Adopción del comportamiento humano y organizacional, la mayoría de las personas están acostumbradas a comunicarse por los canales establecidos (Chat, Dialogo, Teléfono, Reuniones), hay dudas respecto al potencial que puede tener el entrenamiento a los IVA. La clave pasa por que los humanos dediquen tiempo al entrenamiento. 2) Tecnológico, falta de inversión en la infraestructura de comunicación y TI. Las TELCOS que deberían potenciar estos productos a través de sus servicios de valor agregado, tal vez con una app o algún otro medio

no están haciendo esfuerzo en Argentina. También hay un factor económico, de caso de negocio, ya que las empresas que los producen por distintas circunstancias no se encuentran en el país.

ii) Creo que a nivel tecnológico y funcional los procesos deben redefinirse para adecuarse a los cambios continuos y constantes dentro de las empresas y generar visibilidad de los procesos To Be para incorporar la inteligencia artificial dentro de las empresas. El uso doméstico esta sólo vinculado actualmente a las personas que pueden viajar y traer un dispositivo. Al no haber casos de usos visibles y tangibles de asistentes virtuales corporativos, como el de Watson de IBM o Amilia de Apisoft se hace difícil incorporarlos en las empresas que deben coordinar su interoperabilidad y su infraestructura tecnológica.

iii) Creo que puede traccionar en el mediano plazo, 3/4 años para poder contar con un asistente virtual en cada hogar.

