



En bot we trust: una nueva metodología de medidas de rendimiento de chatbot



Aleksandra Przegalinska ^{a,*}, Leon Ciechanowski ^{a,B,^},
Anna Stroz ^c, Peter Gloor ^d, Grzegorz Mazurek ^a

^a Universidad Kozminski, Jagiellonska 57/59, Varsovia, Polonia

^B Universidad de Ciencias Sociales y Humanidades, Chodakowska 19/31, Varsovia, Polonia

^c Universidad de Varsovia, Krakowskie Przedmiescie 26/28, Varsovia, Polonia

^d MIT Center for Collective Intelligence, 245 First Street, E94-1509 Cambridge, MA,

Estados Unidos

PALABRAS CLAVE

Artificial
inteligencia;
Chatbots;
Chatbot
rendimiento;
Humano-
computadora
Interacción;
Rendimiento
metas;
Confianza del cliente;
Cliente
experiencia

Abstracto Los chatbots se utilizan con frecuencia en los negocios para facilitar varios procesos, en particular los relacionados con el servicio al cliente y la personalización. En este artículo, proponemos métodos novedosos para rastrear las interacciones humano-chatbot y medir el rendimiento del chatbot que tienen en cuenta las preocupaciones éticas, en particular la confianza. Nuestra metodología propuesta vincula métodos neurocientíficos, minería de textos y aprendizaje automático. Argumentamos que la confianza es el punto focal de la interacción exitosa de los chatbots humanos y evaluamos cómo la confianza como categoría relevante se está redefiniendo con el advenimiento de los chatbots compatibles con el aprendizaje profundo. Proponemos un método novedoso para analizar el contenido de los mensajes producidos en las interacciones humano-chatbot, utilizando el sistema de búsqueda de la **Tribu Condor** que desarrollamos para la minería de texto y que se basa en un motor de clasificación de aprendizaje automático.

^a 2019 Kelley School of Business, Universidad de Indiana. Publicado por Elsevier Inc. Todos los derechos reservados.

1. Interacción hombre-máquina en la actualidad

La inteligencia artificial es una disciplina académica establecida e internamente diversa que abarca muchos

décadas. Abarca subcampos como el procesamiento del lenguaje natural, el aprendizaje automático, la robótica y muchos otros. Con los principales avances en el aprendizaje automático y la llegada de big data, se ha convertido en una disciplina de suma importancia para la economía global y la sociedad en su conjunto, así como para las empresas y organizaciones que cambian rápidamente. Las capacidades actuales de aprendizaje automático incluyen competir al más alto nivel en estrategias

* Autor correspondiente

Dirección de correo electrónico: aprzegalinska@kozminski.edu.pl (A.

Przegalinska)

[^]A. Przegalinska y L. Ciechanowski contribuyeron igualmente al artículo.

<https://doi.org/10.1016/j.bushor.2019.08.005>

0007-6813 /^a 2019 Kelley School of Business, Universidad de Indiana.

Publicado por Elsevier Inc. Todos los derechos reservados.

juegos, vehículos autónomos, intelli-
 enrutamiento suave en redes de distribución,
 contenido y lo más importante el contexto de esta
 en el artículo comprensión y generación natural
 habla. Mientras miramos como inteligencia artificial lejana
 gence está llegando, haga esta pregunta: ¿Hasta dónde
 podemos y debemos llegar con más desarrollos en la
 capacidad de respuesta de los sistemas de chatbot?

Varios casos del mundo real han demostrado el
 tremendo poder que representa la IA para las empresas.
 Facebook usa inteligencia artificial para el reconocimiento
 facial para hacer que sus aplicaciones sean mejores que los
 humanos para determinar si dos imágenes son de la misma
 persona. El fabricante de motocicletas Harley-Davidson
 aumentó las oportunidades de ventas en Nueva York casi 30
 veces, de menos de cinco a alrededor de 40 por día, a través
 de comunicaciones personalizadas e hiperdireccionadas
 habilitadas por IA. Estos ejemplos muestran que la
 inteligencia artificial, además de ser un tema candente y
 una palabra de moda, se ha abierto camino en una variedad
 de organizaciones y se ha convertido en una prioridad para
 muchas corporaciones e instituciones. Sin embargo, aún
 quedan muchos problemas sin resolver. Cuanto más
 ingresa la IA al mundo de los negocios, más cuestiones
 tecnológicas, sociales y éticas surgen.

En este artículo, nos centramos en la importante
 subdisciplina del chatbot dentro de la IA. Los chatbots
 son agentes virtuales interactivos que se involucran en
 interacciones verbales con humanos. Esta tecnología es
 un caso interesante en la interacción hombre-máquina,
 ya que están diseñadas para interactuar con los usuarios
 mediante el uso de lenguaje natural basado en modelos
 formales. Esta tecnología de 60 años se diseñó
 inicialmente para determinar si los sistemas de chatbot
 podían engañar a los usuarios haciéndoles creer que
 eran humanos reales. Muy pronto, se confirmó que,
 cuando se trata de simular interacciones, era posible
 engañar a los humanos (Khan y Das, 2018; Sloman, 2014
). Sin embargo, recientemente, el objetivo principal de
 construir sistemas de chatbot no es solo imitar la
 conversación humana y engañar a los usuarios en varias
 iteraciones de la prueba de Turing (Aron, 2015; Francés,
 2012; Hernández-Orallo, 2000; Moor, 2003) pero
 también para aplicar una interacción chatbots ampliamente para
 bien formada con los usuarios en la educación de negocios,
 recuperación de información o. Durante bastante tiempo
 tiempo, aplicar chatbots era engorroso y arriesgado (Klüwer, 2011; Shawar y Atwell, 2007a). Sin embargo, en la
 actualidad, los chatbots pueden facilitar varios procesos
 comerciales, en particular los relacionados con el servicio al
 cliente y la personalización debido a su accesibilidad, costo
 bastante bajo y facilidad de uso para los consumidores
 finales. Nuestro principal objetivo en este artículo es
 rastrear la evolución de los chatbots desde sistemas
 sintácticos simples hasta entidades robustas de
 procesamiento de lenguaje natural que utilizan deep

aprendizaje (es decir, Google Duplex, Microsoft Cortana,
 GPT-2) y exponga cómo este cambio afecta a las
 empresas y organizaciones.

Proponemos una metodología novedosa para
 rastrear interacciones humano-chatbot y medir el
 rendimiento del chatbot que toma en consideración
 preocupaciones éticas. Den particular, la confianza en los
 chatbots. Nuestra investigación vincula los métodos
 neurocientíficos, la minería de textos y el aprendizaje
 profundo con la cuestión de la confianza asignada a los
 chatbots y su transparencia general. En un nivel más
 general, abordamos una amplia gama de cuestiones de
 importancia tanto teórica como práctica, incluidos los
 efectos generales de la IA en las empresas e
 instituciones, cómo se debe utilizar la IA para crear
 valor, los posibles efectos negativos de la IA y cómo
 superarlos, y las cuestiones éticas que surgieron con la
 llegada de la inteligencia artificial.

2. Chatbots en organizaciones

Uno de los primeros chatbots, ELIZA, fue creado en 1964. mi
 1966 en el Laboratorio de Inteligencia Artificial del MIT y
 fue uno de los primeros intentos de simular una
 conversación humana. ELIZA (Weizenbaum, 1966;
 Weizenbaum y McCarthy, 1976) simulaban conversaciones
 utilizando un enfoque de sustitución y una coincidencia de
 patrones que cautivaron a los usuarios con una ilusión de
 comprensión y reciprocidad. Las directivas de interacción
 ocultas en los scripts permitieron a ELIZA procesar los
 registros de los usuarios y participar en la comunicación. El
 guión más conocido de ELIZA fue DOCTOR, que simulaba a
 un psicoterapeuta rogeriano que representaba un enfoque
 que empodera al paciente en el proceso terapéutico. El
 creador de ELIZA, Joseph Weizenbaum, consideró el
 programa como un método para exponer la superficialidad
 de la comunicación entre humanos y máquinas, pero se
 sorprendió por la cantidad de pacientes que atribuían
 sentimientos similares a los humanos a ELIZA y querían
 continuar su terapia con él. El éxito de ELIZA fue tan
 sorprendente porque el chatbot era incapaz de entender
 nada. Si bien ELIZA podría participar en un discurso
 percibido por el interlocutor como fructífero, no pudo
 conversar en un nivel semántico más profundo, lo que llevó
 a una nueva idea para desarrollar chatbots. Es decir,
 muchos desarrolladores de chatbot se sintieron animados a
 utilizar diversos tipos de trucos para engañar a los usuarios
 de chatbot. En lugar de centrarse en un trabajo engorroso
 con el procesamiento del lenguaje natural, se centraron en
 las formas más fáciles de engañar a los usuarios. Tales
 intentos a menudo tuvieron éxito como en los casos de
 PARRY (Curry y O'Shea, 2012; Phrasee, 2016) y

Eugenio Gootsman (Aamoth, 2014; Cha, Warwick, Vallverdú y Wu, 2016). Mientras que PARRY simuló efectivamente a una persona paranoica repitiendo amenazas y sospechas en cada conversación, Eugene Gootsman simuló a un niño de 14 años de Ucrania con un conocimiento limitado del mundo y una menor competencia para hablar inglés.

Pensamos en estos scripts y respuestas predeterminadas como el paradigma de simulación. Evoca el argumento de la habitación china. Sin embargo, con el aumento de las capacidades de procesamiento profundo del lenguaje natural, esto puede cambiar pronto a medida que se creen más sistemas 'conscientes'. Incluso antes de la llegada de las interfaces gráficas de usuario, las décadas de 1970 y 1980 vieron un rápido crecimiento en la investigación de interfaces de texto y lenguaje natural. Desde entonces, apareció una gama de nuevas arquitecturas de chatbot, incluidas MegaHAL, Converse y

ALICIA (Shawar y Atwell, 2005, 2007b), que almacena el conocimiento sobre los patrones de conversación en inglés en archivos de lenguaje de marcado de inteligencia artificial (AIML).

La elección de chatbots es mucho más amplia e incluye una gran cantidad de tecnologías de consumo compatibles con el aprendizaje automático como Siri y Cortana (Dunn, 2016; Hachman, 2014; López, Quesada y Guerrero, 2017). Estos pueden asumir la forma de agentes virtuales u objetos físicos que también podríamos investigar desde la perspectiva de las relaciones proxémicas que mantienen con los usuarios, así como la comunicación gestual (Ciechanowski, Przegalinska, Magnuski y Gloor, 2019). Amazon lanzó Alexa (Hayes, 2017)

proyecto como una plataforma para una mejor máquina de discos, pero la compañía lo desarrolló en un sistema de inteligencia artificial construido sobre y aprendiendo continuamente de datos humanos. El cilindro Echo impulsado por Alexa y el Dot más pequeño son ayudantes domésticos omnipresentes que pueden apagar las luces, contar chistes o permitir que los usuarios consuman las noticias con las manos libres. También recopilan una gran cantidad de datos sobre sus usuarios, que Amazon utiliza para mejorar Alexa y desarrollar sus posibilidades (Hayes, 2017). Alexa mantiene 15.000 habilidades, pero 14.997 de ellas permanecen sin explotar ya que la mayoría de los usuarios no las conocen (Przegalinska, 2018). Aunque los robots sociales como Alexa han progresado más allá de ELIZA en cuanto a usos y capacidades en entornos ruidosos de la vida real, todavía están integrados en el paradigma de la simulación. Por esta razón, tanto Google como OpenAI con sus respectivos bots compatibles con aprendizaje profundo, Duplex y GPT-2, ofrecen a Amazon una fuerte competencia a pesar de que solo cuentan con 378 habilidades para el Asistente de Google (Leviatán y Matias, 2018), considerando que ha estado entrenando su IA con petabytes de datos.

3. Medidas de rendimiento actuales del chatbot

Cuando se trata de mediciones de desempeño, existen ciertos estándares a pesar de que el uso de métricas difiere según la industria (Botanalítica, 2018). Encuestas de empresas que implementan chatbots (Asher, 2017; Castro et al., 2018) mostró cómo las empresas de todos los sectores han estado utilizando el análisis de chatbot y qué parámetros consideran cruciales para medir el rendimiento. Por ejemplo, cuando se trata de los sectores bancario y de fintech, los chatbots se utilizan principalmente para ayudar a los usuarios a realizar tareas más rápidamente, mientras se reducen los volúmenes de llamadas y se reducen los costos del servicio (Okuda y Shoda, 2018). Una de las métricas de rendimiento más importantes del chatbot es la duración y la estructura de la conversación. La mayoría de las tendencias de la industria sugieren apuntar a conversaciones más cortas con una estructura simple porque los chatbots bancarios tienen la intención de brindar soluciones rápidas (por ejemplo, enviar / recibir dinero, verificar un saldo).

La tasa de retención es otra métrica importante para medir el éxito del chatbot. Si el chatbot de una empresa tiene como objetivo reemplazar otros canales de comunicación (por ejemplo, reducir el volumen de llamadas), el objetivo es obtener una retención significativamente mayor, que es un indicador de la satisfacción del consumidor. Sin embargo, hay muchas otras opciones automatizadas que permiten a los usuarios administrar cuentas fácilmente sin hablar con un humano. Por lo tanto, si el chatbot se centra más en la atención al cliente, una alta tasa de retención no necesariamente tiene que ser la medida del éxito.

Otra medida del rendimiento del chatbot es la capacidad de proporcionar una comunicación personalizada. Un chatbot financiero debe retener y analizar la información del usuario y, posteriormente, proporcionar a los usuarios ofertas relevantes. Por ejemplo, Erica, un chatbot creado para Bank of America, ofrece a los usuarios consejos sobre cómo pueden ahorrar mejor dinero en función de sus hábitos de gasto.¹

Los chatbots minoristas brindan formas gamificadas de comprar. A través de una interfaz conversacional, las marcas se transforman en asistentes personales de compras. A diferencia de los chatbots bancarios, los creadores de chatbots minoristas generalmente buscan una gran cantidad de pasos de conversación en las métricas de participación del chatbot. Si bien los usuarios pueden buscar una solución rápida, se supone que los chatbots minoristas capturan y mantienen la atención de los usuarios alentándolos a navegar, brindando respuestas sobre productos o vendiendo una compra de manera cruzada / ascendente. Un ejemplo interesante aquí es el ShopBot de eBay, que

¹ <https://promo.bankofamerica.com/erica/>

ofrece varias opciones para fomentar la navegación y la compra.

4. Confianza en el contexto organizacional

La confianza se define generalmente como una creencia firme en la confiabilidad, la verdad o la capacidad de alguien o algo, o "un acuerdo mediante el cual una persona (un fideicomisario) tiene la propiedad como su propietario nominal para el bien de uno o más beneficiarios" (Abogados de HOMS, 2019). La confianza entre los seres humanos puede percibirse en términos de credibilidad y confianza en el juicio de los demás, así como en la previsibilidad del comportamiento de uno (Cassell y Bickmore, 2000).

En el contexto organizacional, la confianza puede definirse de múltiples formas que poseen un fundamento común de creencias (Dunn y Schweitzer, 2005). Puede entenderse como:

Competencia, definida como "la creencia de que una organización tiene la capacidad de hacer lo que dice que hará. [Incluyendo] la medida en que consideramos que una organización es eficaz; que puede competir y sobrevivir en el mercado" (Instituto de Relaciones Públicas, 2003);

Integridad, definida como la creencia de que una organización es justa y equitativa; o

Fiabilidad, definida como la creencia de que una organización "hará lo que dice que hará, que actuará de manera consistente y confiable" (Grunig, 2002).

Si bien los expertos no están completamente de acuerdo, la confianza entre una organización y su público se describe generalmente como tener varios características claramente cuantificables que incluyen, pero no limitado a:

Aspecto multinivel La confianza es el resultado de las interacciones dentro y entre organizaciones;

Aspecto basado en la comunicación y el discurso La confianza surge de interacciones y comportamientos comunicativos;

Aspecto impulsado por la cultura La confianza está correlacionada con valores, normas y creencias; y

Aspecto temporal La confianza cambia constantemente a medida que pasa por las fases de establecimiento, ruptura, interrupción, etc. (Hurley, 2011; Shockley-Zalabak, Morreale y Hackman, 2010).

En términos de vincular la confianza y las nuevas tecnologías con la IA en un contexto organizacional, Kaplan y Haenlein (2019, pag. 22) citó tres rasgos comunes relevantes "tanto interna como externamente: confianza, cambio y control. Las tres C de las implicaciones organizativas de la IA". Una categoría importante que analizaremos más adelante en este artículo es la confianza externa, definida aquí como la confianza en las capacidades y recomendaciones de los sistemas de IA de una organización.

5. Confía en los chatbots

La cuestión de la confianza en el rendimiento del bot es importante por varias razones. En particular, la confianza debe tenerse en cuenta durante la implementación de chatbots en entornos como la atención financiera, la atención médica y otros campos que exigen datos sensibles por los cuales los usuarios pueden estar expuestos a daños físicos, financieros o psicológicos (Bickmore y Cassell, 2001). La confianza puede ser un factor crucial porque los usuarios no quieren compartir información personal si no están seguros de la seguridad (Chung, Joung y Ko, 2017b).

Una parte crucial de la confianza está relacionada con la antropomorfización (Ciechanowski et al., 2019; Lotze, 2016; Radziwill y Benton, 2017; Stojnic, 2015). Nuestra intuición puede llevarnos a afirmar que debido a que los chatbots interactúan mediante la conversación, es apropiado antropomorfizarlos. Sin embargo, hay dos posiciones opuestas. Uno de ellos afirma que si un agente es más humano, entonces es más probable que se establezca una relación de confianza sostenible entre el agente y el usuario, lo que se conoce como perspectiva humano-humano. Dado que los humanos colocan las expectativas sociales en las computadoras, la antropomorfización aumenta la confianza de los usuarios en los agentes informáticos (es decir, cuanto más humana es la computadora, más confianza le aplican los humanos). El segundo punto de vista es que los humanos confían más en los sistemas computarizados que en otros humanos (Seeger y Heinzl, 2018). Según este punto de vista, la automatización de alta calidad conduce a una interacción más fructífera porque la máquina parece ser más objetiva y racional que un humano. Los humanos tienden a confiar en los sistemas informáticos más que en otros humanos porque se espera que los humanos sean imperfectos, mientras que lo contrario es cierto para la automatización (Dijkstra, Liebrand y Timminga, 1998; Seeger y Heinzl, 2018).

Es importante destacar, sin embargo, que el proceso de antropomorfización no se trata sólo de la atribución de características humanas superficiales, sino más importante aún, de esta esencial: una mente semejante a la humana. La gente confía antropomorfizada

tecnología más que la insensata, particularmente debido a la atribución de competencia (Waytz, Heafner y Epley, 2014) resultando en un aumento significativo de la confianza y el desempeño general de la cooperación (de Visser et al., 2012; Waytz et al., 2014). La confiabilidad de varios elementos del ecosistema de la economía digital (por ejemplo, dispositivos portátiles, sistemas y plataformas, asistentes virtuales) surge como un problema importante cuando se considera el nivel de confianza que se les asigna.

En un contexto social, la confianza tiene varias connotaciones. La confianza se caracteriza por los siguientes aspectos: una parte (fideicomitente) está dispuesta a depender de las acciones de otra parte (fiduciario) y la situación se dirige al futuro. Además, el fideicomitente (voluntaria o forzosamente) abandona el control sobre las acciones realizadas por el fiduciario. Como consecuencia, el fideicomitente no está seguro del resultado de las acciones del otro; solo pueden desarrollar y evaluar expectativas. Por lo tanto, la confianza generalmente se puede atribuir a las relaciones entre personas. Se puede demostrar que los humanos tienen una disposición natural a confiar y juzgar la confiabilidad que se puede rastrear hasta la estructura neurobiológica y la actividad de un cerebro humano. Cuando se trata de la relación entre las personas y la tecnología, la atribución de confianza es un tema de disputa (Taddeo, 2010). Esto es especialmente cierto para la tecnología de la información que altera drásticamente la causalidad en los sistemas sociales: IA, tecnología portátil, bots, asistentes virtuales y datos. Todo eso requiere nuevas definiciones de confianza.

En términos de confianza en los chatbots hoy en día, se pueden distinguir al menos otras dos dimensiones relevantes además de la antropomorfización (Nordheim, 2018):

Habilidad / experiencia: medidas de desempeño (por ejemplo, retención de clientes, duración de la conversación, generación de prospectos). La experiencia se considera un factor asociado con la credibilidad, una señal de confiabilidad. En el contexto de un sistema automatizado, se ha argumentado que la confianza se basa principalmente en las percepciones de los usuarios sobre la experiencia de los sistemas. Aquí, se supone que la experiencia del chatbot afecta la confianza de los usuarios.

Privacidad / seguridad: "Los diagnósticos de seguridad exponen vulnerabilidades y amenazas a la privacidad que existen en los asistentes virtuales inteligentes (IVA) comerciales. Los diagnósticos ofrecen la posibilidad de un IVA más seguro

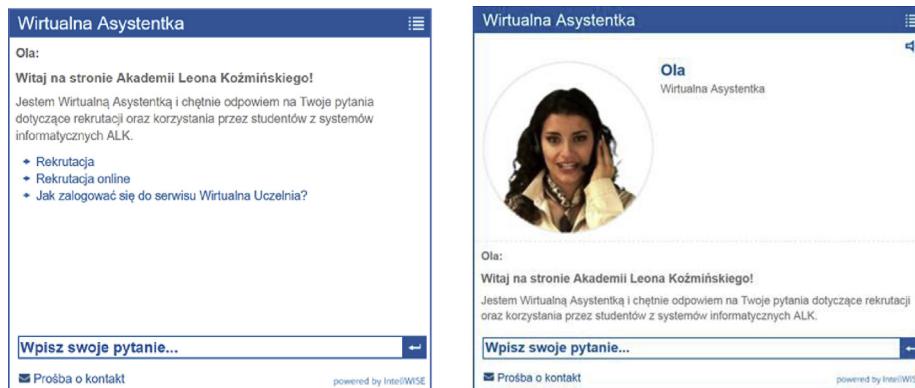
ecosistemas" (Chung, Iorga, Voas y Lee, 2017a, pag. 100).

En el pasado, se podría argumentar que la prueba de Turing ya mencionada era una forma específica de práctica relacionada con la confianza. Fue un experimento que se concibió como una forma de determinar la capacidad de la máquina para usar el lenguaje natural y demostrar indirectamente su capacidad para pensar de una manera similar a un ser humano y, como tal, tocó dos dimensiones de la confianza: antropomorfismo y pericia. En el caso de este último, la confianza se relacionó con la pericia humana para evaluar adecuadamente el estatus del interlocutor (humano / no humano). Hoy, sin embargo, incluso la transparencia se convierte en un tema problemático. Como se mencionó anteriormente, Google Duplex en temas específicos y de especialización limitada exhibe un nivel cercano al humano de comprensión y síntesis del habla humana, así como la capacidad de usar señales no verbales de que la prueba de Turing en formas anteriores ya sería inadecuada para pruebas y análisis (Harwell, 2018). Esto no quiere decir que la prueba de Turing deba suspenderse como herramienta, sino que debe reformularse. Además, DeepMind una empresa británica de inteligencia artificial adquirida por Google en 2014 está desarrollando un proyecto que puede remodelar la industria de los chatbots de una manera aún más dramática.

DeepMind explora el campo de la teoría de la mente (Premack y Woodruff, 1978) que se refiere en términos generales a la capacidad y la capacidad de los seres humanos para observar varios estados mentales como estados mentales pertenecientes a otros seres humanos. DeepMind propuso entrenar una máquina para construir tales modelos e intenta diseñar una Teoría de la Mente red neuronal (ToM-net) que utiliza el metaaprendizaje para construir modelos de los agentes que encuentra a partir de observaciones de su comportamiento únicamente (Rabinowitz et al., 2018): "A través de este proceso, adquiere un modelo previo sólido para el comportamiento de los agentes, así como la capacidad de iniciarse en predicciones más ricas sobre las características y los estados mentales de los agentes utilizando solo una pequeña cantidad de observaciones de comportamiento".

² DeepMind aplica ToM-net a agentes que se comportan en entornos simples de gridworld, lo que demuestra que aprende a modelar agentes de aprendizaje de refuerzo profundo, algorítmicos y aleatorios de poblaciones variadas y, lo que es más interesante, pasa tareas clásicas de ToM como "Sally-Anne". prueba (Baron, 1985; Wimmer y Perner, 1983) de reconocer que otros pueden tener creencias falsas sobre el mundo. Los autores argumentan que este sistema es un importante paso adelante para el desarrollo de sistemas de IA de agentes múltiples, para construir tecnología de intermediación para la interacción máquina-humano y para avanzar en el progreso de la IA interpretable.

Figura 1. Dos tipos de chatbots *



Nota: Cada usuario fue asignado al azar para hablar con uno de ellos. A la izquierda está el chatbot de texto simple, a la derecha está el avatar con sonido.

Que es esto implica en el contexto de nuestro investigarDy más en generalDes eso humano-La interacción de chatbot es un nuevo método de conceptualizar e investigar la confianza. Como ya comentamos, en el caso de los sistemas que no son blackbox (por ejemplo, chatbots basados en motores estocásticos), atribuir e investigar la confianza ya era difícil. Con los sistemas de aprendizaje profundo, se abren nuevas dimensiones de las relaciones hombre-máquina basadas en la confianza.

6. Metodología

Realizamos un estudio (el primero de este tipo) de interacción en línea entre humanos y chatbot (Ciechanowski et al., 2019; Ciechanowski, Przegalinska y Wegner, 2018) utilizando medidas subjetivas (cuestionarios) y objetivas (psicofisiología). El estudio consistió en interacciones con uno de los dos tipos de chatbot: un chatbot de texto simple o un avatar con sonido (Figura 1).

Durante la conversación con el chatbot, recopilamos datos psicofisiológicos de los participantes y al final del procedimiento, se les pidió que completaran un par de cuestionarios (Fiske, Cuddy, Glick y Xu, 2002; Fong, Nourbakhsh y Dautenhahn, 2002; Pochwatko et al., 2015) que tuvo en cuenta la cooperación entre humanos y chatbots, la realidad del chatbot y su semejanza humana. Los resultados importantes para el estudio actual indicaron que los usuarios estaban más satisfechos y felices durante la interacción con el chatbot de texto simple, lo que se confirmó tanto en los resultados del cuestionario como en los marcadores psicofisiológicos. Estos resultados también han probado la hipótesis del valle inquietante (Mori,

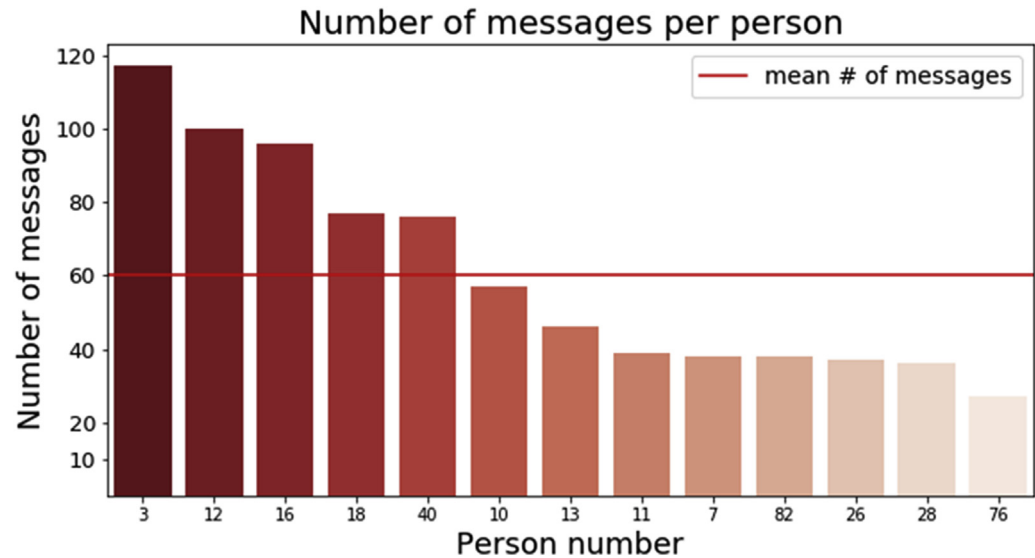
MacDorman y Kageki, 2012), que establece que los usuarios y consumidores están menos satisfechos cuando interactúan con objetos artificiales imitando ineptamente a un humano, y esto es lo que nuestro avatar chatbot estaba intentando hacer. Este resultado también se manifestó directamente en la medida del valle inquietante que aplicamos en nuestro procedimiento experimental.

Sin embargo, la pregunta de investigación que todavía nos interesaba estaba relacionada con las correlaciones lingüísticas del efecto valle inquietante. En otras palabras, quizás la imitación inepta del avatar chatbot llevó a los usuarios a comportarse verbalmente de una manera específica hacia él, y el chatbot, por lo tanto, respondió de diferentes maneras, irritándolos más que el simple chatbot de texto. Para responder a esta pregunta, analizamos los mensajes producidos por los usuarios en su interacción con el chatbot avatar.³

Para el análisis de los mensajes traducidos, utilizamos Tribefinder, un novedoso sistema basado en aprendizaje automático capaz de revelar las afiliaciones tribales de los usuarios de Twitter (De Oliveira y Gloor, 2018; Gloor, Colladon, de Oliveira y Rovelli, 2018). En este caso, una tribu es "una red de personas heterogéneas. Que están unidas por una pasión o emoción compartida" (Cova y Cova, 2002, pag. 602). En el dominio del mercado, los usuarios

³ No hay herramientas para el análisis automático de texto o la minería de texto en polaco; por lo tanto, primero necesitábamos traducir los mensajes al inglés. Con el fin de traducir archivos de registros de chatbot. Las indicaciones de los usuarios y las respuestas del chatbotDdel polaco al inglés, se utilizó la plataforma DeepL. DeepL Translator aplica la traducción automática neuronal, lo que da como resultado una precisión y calidad del texto impresionantes. Las partes de las oraciones se tradujeron manualmente. Dparticularmente debido a algunos atajos mentales y frases coloquiales introducidas por los usuarios. El número total de entradas traducidas fue de 1373 para 13 usuarios.

Figura 2. Línea media de intercambio usuario-chatbot y número de mensajes para cada usuario



o los consumidores se afilian con diferentes tribus y, por lo tanto, pueden caracterizarse por diversos comportamientos, rituales, tradiciones, valores, creencias, jerarquía y vocabulario. El análisis y el uso exitoso del tribalismo pueden ser un factor crucial en el marketing y la gestión modernos (Cova y Cova, 2002; Moutinho, Dionísio y Leal, 2007).

Inicialmente, Tribecreator se utilizó para seleccionar líderes e in fl uentes tribales, y Tribefinder se creó para establecer tribus a las que pertenecen los individuos mediante el análisis de sus tweets y la comparación de vocabulario. Técnicamente, Tribefinder está respaldado por un sistema de aprendizaje automático, que se entrenó con feeds de Twitter de 1,000 periodistas deThe New York Times, The Washington Post, The Guardian, y con miembros del Congreso de los Estados Unidos. Aplicamos este sistema, en nuestro caso, para analizar el contenido de los mensajes de los usuarios y chatbots. Como estábamos interesados en la razón por la que los usuarios estaban menos dispuestos a cooperar con el chatbot avatar, analizamos las interacciones en este grupo usando el buscador de tribus.

El sistema pudo procesar mensajes de 13 de las 15 personas del grupo avatar chatbot, ya que se basa en algoritmos de aprendizaje automático y necesita cantidades suficientes de datos para llevar a cabo una clasificación significativa. El significado número de interacciones (de fi nidas por el chatbot intercambio de usuarios de un mensaje) de seleccionado los usuarios fue 60,3 (Figura 2).

El sistema de búsqueda de tribus pudo asignar usuarios en cinco macrocategorías de tribus (ver tabla 1):

Personalidad;

Ideología;

Estilos de vida;

Realidad alternativa; y

Recreación.

La ideología describe a los proponentes de la ideologías políticas, formado con 100 representantes humanos de cada una de las tres principales ideologías: socialismo, capitalismo y liberalismo. Los creadores de Tribefinder tuvieron que agregar la categoría de quejosos a esta lista, ya que algunas personas no encajaban en ninguna ideología dominante (Gloor et al., 2018). En elApéndice.

7. Resultados

Luego del análisis de los mensajes generados por los usuarios y por el chatbot en la interacción con cada usuario, observamos que todos los sujetos y chatbots (excepto en dos instancias) eran del tipo de personalidad periodística; por lo tanto, no usaron un lenguaje engañoso. De manera similar, todas las personas y los chatbots se quejaron (excepto una persona y un chatbot clasificados como liberales), por lo que fue imposible discernir su ideología.

Observamos que casi el 50% de los usuarios se puede caracterizar por mostrar características de la

Tabla 1. Definiciones de tribus asignadas a usuarios y chatbots por el sistema Tribefinder

Macro tribal categoría	Tipo de tribu	Caracterización
Personalidad	El periodista	Gente diciendo la verdad.
	Político	Personas que intentan mentir o usan lenguaje
Ideología	Liberalismo	engañoso. Personas suscritas a la ideología liberal.
	Capitalismo	Personas suscritas a la ideología capitalista. Personas
	Socialismo	suscritas a la ideología socialista.
	Denunciantes	Personas que no expresan ideas relacionadas con las 3 principales ideologías políticas.
Estilo de vida	Aptitud física	Gente aficionada a hacer deporte, hasta el punto de llegar a ser compulsivo.
	Sedentario	Lo opuesto a la categoría de aptitud física, estas personas aplican un estilo de vida inactivo.
	Vegano	Personas que evitan los productos animales en su dieta.
	YOLO	Personas que viven según el lema Solo se vive una vez. Puede mostrar un comportamiento hedonista, impulsivo o imprudente.
Alternativa realidad	Padre de familia (Nacionalista)	Conservadores, patriotas extremos, gente de familia y xenófobos.
	Nerd (tecnócratas)	Entusiastas de la tecnología, transhumanistas, fanáticos de la globalización y el networking.
	Espiritista	Gente que cree en algún tipo de poder y significado sagrado. Aficionado a la contemplación espiritual.
	Abrazador de árboles (Ambientalistas)	Personas destacando la importancia de la protección de la naturaleza. Por lo tanto, aceptan algunas tecnologías (por ejemplo, energías alternativas), pero desafían otras (por ejemplo, manipulación de genes).
Recreación	Arte	Personas interesadas en el arte, apreciando su belleza y emociones.
	Moda	Personas que siguen la moda (en el ámbito de la ropa, el peinado o el comportamiento).
	Deporte	Personas que siguen eventos deportivos en los medios. Algunos de ellos también practican deportes de forma activa.
	Viaje	Gente que ama viajar, conocer diferentes culturas y lugares.

Fuente: (Gloor et al., 2018)

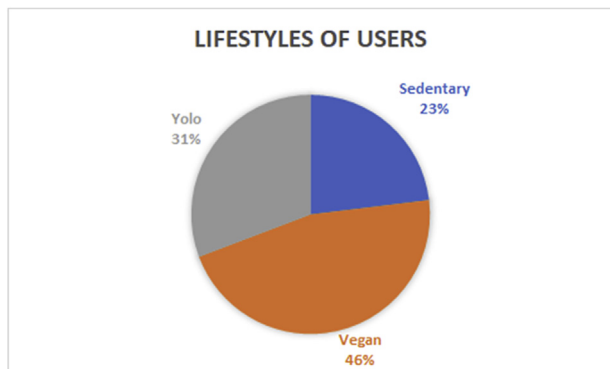
tribu vegana, 33% como YOLO y 25% como sedentario (figura 3). Descubrimos que cada instancia de un chatbot que interactúa con estos usuarios fue asignada por el buscador de tribus a la tribu sedentaria. El chatbot utilizado en nuestro experimento es estacionario y no robótico; por lo tanto, esta afiliación es correcta.

En lo que respecta a las realidades alternativas, observamos una imagen interesante. En cuanto a los usuarios, solo surgen dos categorías: espiritualismo (69%) y nerd (31%). Sin embargo, cuando observamos las tribus que los chatbots afiliaban a esta categoría, observamos a las cuatro tribus representadas (Figura 4), destacando las tribus del abrazador de árboles y el nerd. Curiosamente, ocurre una situación opuesta con la categoría de recreación (Figura 5). Estos resultados indican que los usuarios pueden describirse principalmente como usuarios de lenguaje espiritual en

conversaciones con chatbots. Esto se vuelve más comprensible cuando lo combinamos con el hecho de que los usuarios sabían que estaban teniendo una conversación con un ser artificial y estaban provocando que el chatbot dijera algo humano. Además, era probable que los usuarios que aceptaran participar en nuestro experimento fueran entusiastas de la tecnología, por lo que la otra categoría tribal de realidad alternativa que mostraron fue nerd. Las diferencias mencionadas anteriormente entre humanos y bots son más visibles cuando se presentan en un gráfico de columnas de pila (consulte Figura 6).

Figura 6 muestra las similitudes y diferencias entre los mensajes de humanos y bot. Evidentemente, el bot ajustó sus mensajes a los humanos en las categorías de personalidad e ideología, mientras que hubo diferencias significativas en las categorías de

Figura 3. Estilos de vida de los usuarios



estilo de vida, realidades alternativas y recreación. Estos resultados son comprensibles ya que el chatbot ha estado siguiendo a los humanos en las categorías más abstractas y generales, como usar un lenguaje mayoritariamente veraz (tipo periodista) y, al igual que los participantes humanos, el bot no manifestó creencias políticas (tipo quejoso). Sin embargo, esto puede tener graves consecuencias. Imaginemos un chatbot que represente a una empresa deportiva frente a sus clientes. Si el chatbot refleja las declaraciones de un usuario que odia los deportes, dirá cosas que van en contra de la misión, la visión y los valores de la empresa. Un chatbot de este tipo definitivamente no será una interfaz que haga bien su trabajo.

8. Conceptos básicos en las mediciones de chatbots para chatbots emergentes

Mirando a la [Gartner \(2018\)](#) Hype Cycle, las tecnologías que no han estado a la altura de sus expectativas terminan en la categoría obsoleta antes de la meseta. Dada la importante brecha entre el desarrollo y el uso de habilidades de Alexa, Alexa y muchos otros chatbots pueden clasificarse en esta categoría. Varios chatbots y

Los robots sociales pueden resultar útiles solo como interfaz para un motor de búsqueda. Como lo que hace preguntas de seguimiento para refinar la búsqueda del usuario y encontrar exactamente lo que el usuario está buscando.

Por otro lado, los desarrolladores creen que los dispositivos de inteligencia artificial basados en voz no son solo máquinas de discos que repiten el paradigma de simulación con mejores tecnologías. De hecho, los chatbots son el ejemplo perfecto de la implementación de inteligencia artificial de última generación orientada al consumidor que no solo simula el comportamiento humano basado en modelos formales, sino que también se adapta a él. Como tal, los chatbots siguen siendo un tema fascinante para la investigación de patrones en la interacción humana y no humana junto con cuestiones relacionadas con la asignación de roles sociales a otros, la búsqueda de patrones de interacciones (no) exitosas y el establecimiento de relaciones y vínculos sociales. En ese sentido, los chatbots se están convirtiendo gradualmente en sistemas robustos de contexto. No olvidemos que los científicos e ingenieros crearon inicialmente chatbots porque la gente quería usar el lenguaje natural para comunicarse sin problemas con los sistemas informáticos. ([Zadrozny y col. 2000](#), págs.116mi117) argumentó que la mejor manera de facilitar la interacción persona-computadora es permitir que los usuarios "expresen sus intereses, deseos o consultas de manera directa y natural, hablando, escribiendo y señalando". [Morrissey y Kirakowski \(2013\)](#) hizo un punto similar en sus criterios para el desarrollo de un chatbot más parecido al humano. Quizás los chatbots deberían volverse más sensibles a las necesidades humanas al buscar y entregar información que las personas necesitan y que de otra manera no podrían obtener ([Kacprzyk y Zadrozny, 2010](#); [Morrissey y Kirakowski, 2013](#)). Teniendo en cuenta que los científicos e investigadores están desarrollando bots y chatbots para que se conviertan en asistentes de colaboración más refinados, matizados, contextualizados y transparentes, hay pocos cambios relevantes relacionados con las dimensiones redefinidas de la confianza. Es decir, sobre la base de nuestra investigación y los principales desarrollos en

Figura 4. Realidades alternativas de usuarios y chatbots

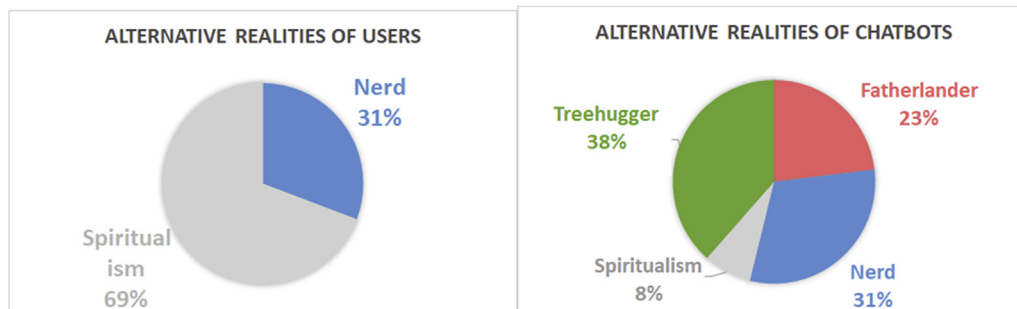
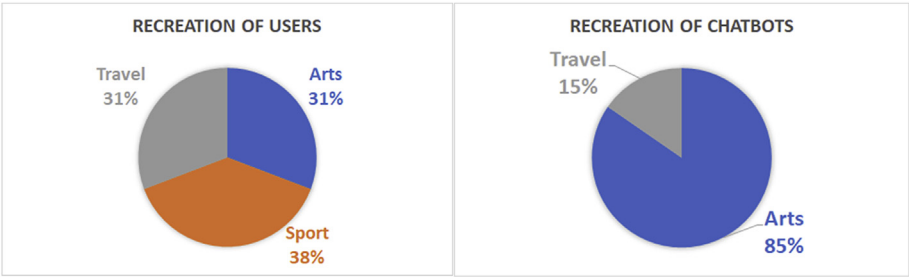


Figura 5. Recreación de usuarios y chatbots



botics, hay nuevas dimensiones de confianza que nos gustaría sugerir:

Transparencia y honestidad: el agente envía señales honestas en la comunicación, no habla como un político (según categorías de Tribefinder) y no niega su estatus.

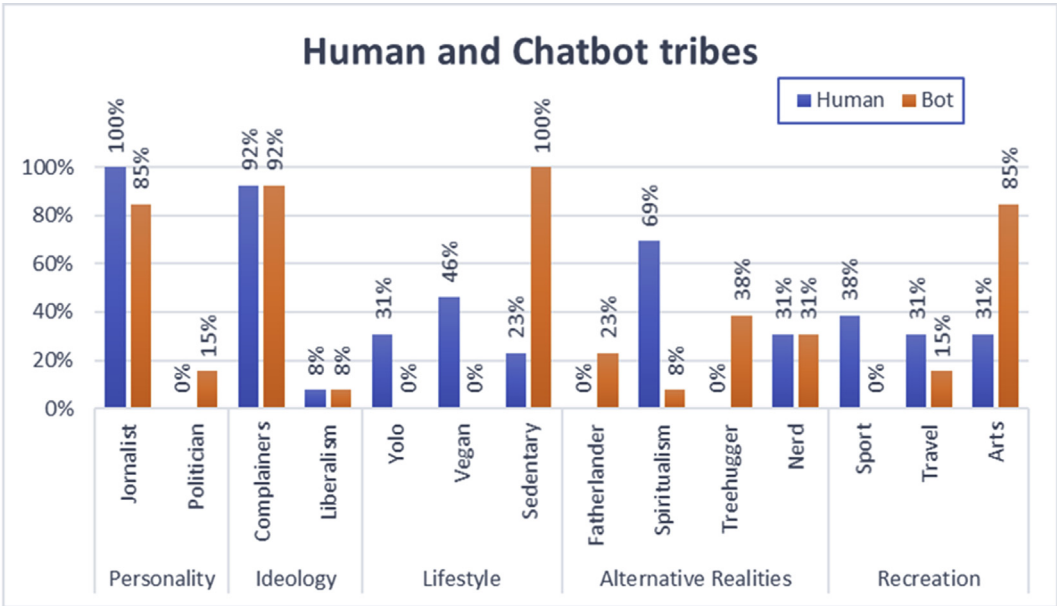
Previsibilidad, ya mencionada en publicaciones sobre confianza y chatbots (Nordheim, 2018), podría complementarse con la integridad, vista como un factor asociado a la credibilidad, y se refiere a la expectativa de los fideicomitentes de que un objeto de confianza actuará de manera coherente de acuerdo con la experiencia pasada. Si el usuario percibe la

chatbots como predecible, esto puede llevar a un sentimiento de confianza en el chatbot.

El control y la benevolencia describen el grado en que las motivaciones e intenciones del fiduciario están en consonancia con las del fideicomitente.

Este último elemento está fuertemente vinculado con el desafío de eliminar la IA de la caja negra e introducir la mayor explicabilidad del procesamiento de la información realizado por la IA (Petkovic, Kobzik y Re, 2018; Preece, 2018). Tribefinder y herramientas similares podrían investigar más la integridad y la benevolencia, así como aplicarlas en chatbots comerciales.

Figura 6. Tipos de tribus de humanos y chatbot



Nota: Presentado en macrocategorías tribales separadas (por ejemplo, personalidad). El porcentaje muestra cuántas personas / bots estaban en una macrocategoría tribal específica, por ejemplo, en la macrocategoría tribal de personalidad, el 100% de los humanos se clasificaron como periodistas, y el 0% de ellos fueron categorizados como políticos.

9. Recomendaciones para empresas

Lo queramos o no, los bots y chatbots ya se están convirtiendo en actores en el mundo en línea. Trabajar en nuevos métodos para crear agentes de conversación confiables para el futuro es una tarea de la ciencia, la investigación y los negocios, y las empresas, en última instancia, implementan los métodos. Teniendo en cuenta la tendencia emergente de cobotización (Méda, 2017), encontrar caminos para crear soluciones de chatbot seguras, útiles y sólidas es una tarea muy importante. Por lo tanto, nuestra recomendación se centra en las tres nuevas características de mejora de la confianza de los chatbots que enumeramos en la sección anterior (es decir, transparencia, integridad, explicabilidad) que permiten para un mayor control. Lo que queda por ver es si ya sea bots humanoides y robotsDa menudo empleado en ventas o atención al cliente departamentosDson realmente capaces de interactuar con los humanos y cuánta interacción real y cooperación entre un ser humano y un chatbot o robot es posible.

Apéndice

Referencias

- Aamoth, D. (2014). Entrevista con Eugene Goostman, el falso niño que pasó la prueba de Turing. Tiempo. Disponible en: <https://time.com/2847900/eugene-goostman-turing-test/>
- Aron, J. (2015). Olvídate de la prueba de TuringDhay mejores formas de juzgar la IA. Científico nuevo. Disponible en: <https://institutions.newscientist.com/article/dn28206-forget-the-turing-test-there-are-better-ways-of-judging-ai/>
- Asher, N. (2017). Una bienvenida más cálida: Aplicación de un chatbot como facilitador para la incorporación de nuevos empleados. Disponible en: <http://www.diva-portal.org/smash/record.jsf?pidZdiva2%3A1116842&dswidZ4708>
- Baron, J. (1985). Racionalidad e inteligencia. Cambridge, Reino Unido: Prensa de la Universidad de Cambridge.
- Bickmore, T. y Cassell, J. (2001). Agentes relacionales: un modelo e implementación del fomento de la confianza de los usuarios. En Actas de la conferencia SIGCHI sobre factores humanos en sistemas informáticos (págs. 396mi403). Nueva York, NY: SIGCHI.
- Botanalítica. (2018). En qué se diferencian las métricas de rendimiento del chatbot en industria. Disponible en: <https://chatbotslife.com/how-chatbot-performance-metrics-differ-by-industry-b380d4bd7f6b>
- Cassell, J. y Bickmore, T. (2000). Manifestaciones externas de confiabilidad en la interfaz. Comunicaciones de la ACM, 43 (12), 50mi56.
- Castro, F., Tedesco, P., Alves, H., Quintino, JP, Steffen, J., Oliveira, F. y col. (2018). Desarrollo de un chatbot corporativo para un programa de participación del cliente: una hoja de ruta. En Teorías y aplicación de la computación inteligente (págs. 400mi412). Cham, Suiza: Springer International.
- Chung, H., Iorga, M., Voas, J. y Lee, S. (2017a). Alexa, puedo ¿confiar en ti? Computadora, 50 (9), 100mi104.

Mensaje de usuario	Mensaje de chatbot
Qué hiciste hoy? ¿Cansado?	Creo que trabajo las veinticuatro horas del día.
¿Estás cansado?	Intente formular su pregunta de otra manera.
Eso es bueno	No estoy cansado, más bien al contrario. Estoy lleno de energía.
Excelente	¡Excelente!
¿Cuáles son tus planes para la noche?	Me alegro también.
	Me sentaré aquí. Así es como paso la mayor parte de las tardes.
Este extracto es del usuario nº 3. Tanto el usuario como esta instancia de chatbot fueron asignados por el sistema Condor al periodista. tipo de personalidad (usando lenguaje veraz).	

Mensaje de usuario	Mensaje de chatbot
Hola :)	¡Oye!
Yo soy oskar	¡Estoy muy contento con esto!
¿Cómo estás?	Genial, muchos clientes.
Buen tiempo.	No me he ido desde ayer, no me preguntes por el tiempo. Todos los días tengo un color diferente: 365 tipos para todo el año. Me alegro.
Ok, tengo algunas preguntas. Ajá, genial.	
Tengo algunas preguntas.	Estoy esperando tus preguntas.
Este extracto es del usuario nº 16. El sistema Condor asignó al usuario el tipo de personalidad del periodista, mientras que esta instancia del chatbot se asignó al político (utilizando un lenguaje engañoso).	

- Chung, M., Joung, H. y Ko, E. (2017b). El papel de las marcas de lujo Agentes conversacionales: Comparación entre cara a cara y chatbot. En Actas de la conferencia global de gestión de la moda de 2017 en Viena. Disponible en: http://gfmcproceedings.net/html/sub3_01.html?codeZ326161 Ciechanowski, L., Przegalinska, A., Magnuski, M. y Gloor, P. (2019). Las sombras del valle inquietante: un estudio experimental de humanos e interacción de chatbot. FGCS: Sistemas informáticos de generación futura, 92, 539mi548.
- Ciechanowski, L., Przegalinska, A. y Wegner, K. (2018). los necesidad de nuevos paradigmas en la medición de la interacción humano-chatbot. En M. Hoffman (Ed.), Avances en la toma de decisiones transculturales. Cham, Suiza: Springer International.
- Cova, B. y Cova, V. (2002). Marketing tribal: la tribalización de la sociedad y su impacto en la realización del marketing. Revista europea de marketing, 36 (5/6), 595mi620.
- Curry, C. y O'Shea, JD (2012). La implementación de un chatbot narrador de historias. Avances en la investigación de sistemas inteligentes, 1 (1), 45mi52.
- De Oliveira, JM y Gloor, PA (2018). GalaxyScope: Encontrar la "verdad de las tribus" en las redes sociales. En F. Grippa, J. Leitão, J. Gluesing, K. Riopelle y P. Gloor (Eds.), Redes de innovación colaborativa: creación de organizaciones adaptativas y resilientes (págs.153mi164). Cham, Suiza: Springer International.
- de Visser, EJ, Krueger, F., McKnight, P., Scheid, S., Smith, M., Chalk, S. y col. (2012). El mundo no es suficiente: Confianza en los agentes cognitivos. En Actas de la reunión anual de la sociedad de factores humanos y ergonomía (págs. 263mi267). Santa Mónica, CA: HFES.
- Dijkstra, JJ, Liebrand, WBG y Timminga, E. (1998). Capacidad de persuasión de los sistemas expertos. Tecnología de la información y el comportamiento, 17 (3), 155mi163.
- Dunn, J. (2016). Ponemos a Siri, Alexa, Google Assistant y Cortana a través de una maratón de pruebas para ver quién gana la carrera de asistente virtual mi esto es lo que encontramos. Business Insider. Disponible en: <https://www.businessinsider.com/sirivs-google-assistant-cortana-alexa-2016-11>
- Dunn, JR y Schweitzer, ME (2005). Sentir y creer: La influencia de la emoción en la confianza. Revista de personalidad y psicología social, 88 (5), 736mi748.
- Fiske, ST, Cuddy, AJC, Glick, P. y Xu, J. (2002). Un modelo de contenido estereotípico (a menudo mixto): la competencia y la calidez se derivan respectivamente del estado percibido y la competencia. Revista de personalidad y psicología social, 82 (6), 878mi902.
- Fong, T., Nourbakhsh, I. y Dautenhahn, K. (2002). Una encuesta de Robots socialmente interactivos: conceptos, diseño y aplicaciones (CMU-RI-TR02-29). Pittsburgh, PA: Universidad Carnegie Mellon.
- Francés, RM (2012). Yendo más allá de la prueba de Turing. Comunicaciones del ACM, 55 (12), 74mi77.
- Gartner. (2018). Surgen 5 tendencias en el ciclo de Gartner Hype para tecnologías emergentes, 2018. Disponible en: <https://www.gartner.com/smarterwithgartner/5-trends-emerge-in-gartnerhype-cycle-for-emerging-technologies-2018/>
- Gloor, PA, Colladon, AF, deOliveira, JM y Rovelli, P. (2018). Identificar tribus en Twitter a través de un contexto compartido. En Y. Song, F. Grippa, PA Gloor y J. Leitão (Eds.), Redes colaborativas de innovación. Cham, Suiza: Springer.
- Grunig, JE (2002). Métodos cualitativos para evaluar las relaciones: Entre organizaciones y públicos. Disponible en: https://instituteforpr.org/wp-content/uploads/2002_AssessingRelations.pdf
- Hachman, M. (2014). Batalla de los asistentes digitales: Windows teléfono Cortana vs Google ahora vs Siri. Mundo PC. Disponible en: <https://www.pcworld.com/article/2142022/the-battle-of-the-digital-assistants-windows-phone-cortana-vs-google-now-vs-siri.html>
- Harwell, D. (2018). Un programa de Google puede pasar como humano el teléfono. ¿Debería exigirse que le diga a la gente que es una máquina? The Washington Post. Disponible en: https://www.washingtonpost.com/news/the-switch/wp/2018/05/08/agoogle-program-can-pass-as-a-human-on-the-phone-should-it-be-required-to-tell-people-es-una-máquina/?utm_term=.cf13615d7898
- Hayes, A. (2017). Amazon Alexa: un principiante de inicio rápido guía. Scotts Valley, CA: Plataforma de publicación independiente CreateSpace.
- Hernández-Orallo, J. (2000). Más allá de la prueba de Turing. Diario de Lógica, lenguaje e información, 9 (4), 447mi466.
- Abogados de HOMS. (2019). Fideicomisos: Implicaciones fiscales. Disponible en: <https://www.homs.ie/publications/trusts-tax-implications/> Hurley, RF (2011). La decisión de confiar: cómo crean los líderes organizaciones de alta confianza. Hoboken, Nueva Jersey: John Wiley & Sons.
- Instituto de Relaciones Públicas. (2003). Directrices para medir confianza en las organizaciones. Disponible en: <https://instituteforpr.org/pautas-para-medir-la-confianza-en-organizaciones-2/> Kacprzyk, J. y Zadrozny, S. (2010). Computar con palabras es un paradigma implementable: consultas difusas, resúmenes de datos lingüísticos y generación de lenguaje natural. Transacciones IEEE en sistemas difusos, 18 (3), 461mi472.
- Kaplan, A. y Haenlein, M. (2019). Siri, Siri, en mi mano: ¿Quién es la más bella de la tierra? Sobre las interpretaciones, ilustraciones e implicaciones de la inteligencia artificial. Horizontes de negocios, 62 (1), 15mi25.
- Khan, R. y Das, A. (2018). Introducción a los chatbots. En R. Khan, y A. Das (Eds.), Cree mejores chatbots: una guía completa para comenzar con los chatbots (págs. 1mi11). Berkeley, CA: Apress.
- Klüwer, T. (2011). Desde chatbots hasta sistemas de diálogo. En D. Perez-Marin e I. Pascual-Nieto (Eds.), Agentes conversacionales e interacción del lenguaje natural: Técnicas y prácticas efectivas (págs. 1mi22). Hershey, PA: IGI Global.
- Leviathan, Y. y Matias, Y. (2018). Google Duplex: un sistema de inteligencia artificial para realizar tareas del mundo real por teléfono. Blog de IA de Google. Disponible en: <https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversion.html>
- López, G., Quesada, L. y Guerrero, LA (2017). Alexa contra Siri vs. Cortana vs. Asistente de Google: una comparación de interfaces de usuario naturales basadas en voz. En IL Nunes (Ed.), Avances en la interacción de sistemas y factores humanos (págs. 241mi250). Cham, Suiza: Springer.
- Lotze, N. (2016). Chatbots. Peter Lang. Disponible en: <https://www.peterlang.com/view/title/18967>
- Méda, D. (2017). El futuro del trabajo: el significado y el valor de trabajar en Europa. Archivos HAL. Disponible en: <https://hal.archives-ouvertes.fr/hal-01616579/>
- Moor, JH (2003). El estado y el futuro de la prueba de Turing. Mentes y máquinas, 11 (1), 77mi93.
- Mori, M., MacDorman, KF y Kageki, N. (2012). Lo inquietante Valle. Revista IEEE Robótica y Automatización, 19 (2), 98mi100.
- Morrissey, K. y Kirakowski, J. (2013). 'Realidad' en chatbots: Establecimiento de criterios cuantificables. En M. Kurosu (Ed.), Interacción hombre-ordenador: modalidades y técnicas de interacción (págs.87mi96). Berlín, Alemania: Springer.
- Moutinho, L., Dionísio, P. y Leal, C. (2007). Comportamiento tribal del surf: una aplicación de marketing deportivo. Inteligencia y planificación de marketing, 25 (7), 668mi690.
- Nordheim, CB (2018). Confíe en los chatbots para el servicio al cliente: Hallazgos de un estudio con cuestionario. Disponible en: https://www.duo.uio.no/bitstream/handle/10852/63498/1/CecilieBertinussenNordheim_masteroppgaveV18.pdf

- Okuda, T. y Shoda, S. (2018). Servicio de chatbot basado en IA para industria financiera. *Revista científica y técnica de Fujitsu*, 54 (2), 4mi8.
- Petkovic, D., Kobzik, L. y Re, C. (2018). Aprendizaje automático y análisis profundo para la biocomputación: exige una mejor explicabilidad. *Simposio Pacífico de Biocomputación*, 23, 623mi627.
- Phrasee. (2016). PARRY: El chatbot de IA de 1972. Disponible en: <https://phrasee.co/parry-the-ai-chatterbot-from-1972/> Pochwatko, G., Giger, J.-C., Różańska-Walczyk, METRO.
- Swidrak, J., Kukiela, K., Mozdyn, J. y col. (2015). Versión polaca de la actitud negativa hacia los robots a escala (NARSPL). *Journal of Automation Mobile Robotics and Intelligent Systems*, 9 (3), 65mi72.
- Preece, A. (2018). Preguntar 'por qué' en IA: explicabilidad de la inteligencia sistemas gent mi perspectivas y desafíos. *Sistemas inteligentes en contabilidad, finanzas y gestión*, 25 (2), 63mi72.
- Premack, D. y Woodruff, G. (1978). ¿Tiene el chimpancé una teoría de la mente? *Ciencias del comportamiento y del cerebro*, 1 (4), 515mi526.
- Przegalinska, A. (2018). Tecnologías portátiles en organizaciones ciones: Privacidad, e fi ciencia y autonomía en el trabajo. Nueva York, NY: Springer.
- Rabinowitz, Carolina del Norte, Perbet, F., Song, HF, Zhang, C., Ali Eslami, SM y Botvinick, M. (2018). Teoría mecánica de la mente. Universidad de Cornell. Disponible en: <https://arxiv.org/abs/1802.07740.pdf>
- Radziwill, NM y Benton, MC (2017). Evaluar la calidad de chatbots y agentes conversacionales inteligentes. Universidad de Cornell. Disponible en: <http://arxiv.org/abs/1704.04579.pdf>
- Seeger, A.-M. y Heinzl, A. (2018). Humano versus máquina: Factores de contingencia del antropomorfismo como estrategia de diseño inductora de confianza para agentes conversacionales. En F. Davis, R. Riedl, J. vom Brocke, P.-M. Léger y A. Randolph (Eds.), *Sistemas de información y neurociencia* (págs. 129mi139). Cham: Springer Internacional.
- Shah, H., Warwick, K., Vallverdú, J. y Wu, D. (2016). Poder las máquinas hablan? Comparación de Eliza con los sistemas de diálogo modernos. *Computadoras en el comportamiento humano*, 58, 278mi295.
- Shawar, BA y Atwell, ES (2005). Usando corpora en máquina-aprendizaje de sistemas de chatbot. *Revista Internacional de Lingüística de Corpus*, 10 (4), 489mi516.
- Shawar, BA y Atwell, E. (2007a). Chatbots: ¿son realmente útiles? *Foro LDV*, 22 (1), 29mi49.
- Shawar, BA y Atwell, E. (2007b). Diferentes medidas métricas para evaluar un sistema de chatbot. En *Actas del taller sobre cómo cerrar la brecha: investigación académica e industrial en tecnologías de diálogo* (págs.89mi96). Stroudsburg, PA: Asociación de Lingüística Computacional.
- Shockley-Zalabak, PS, Morreale, S. y Hackman, M. (2010). Creación de una organización de alta confianza: estrategias para respaldar cinco dimensiones clave de la confianza. Hoboken, Nueva Jersey: John Wiley & Sons.
- Sloman, A. (2014). Juzgando chatbots en la prueba de Turing 2014. Disponible en: <https://www.cs.bham.ac.uk/research/proyectos/cogaff/misc/turing-test-2014.pdf>
- Stojnic, A. (2015). Antropomorfismo digital: Intérpretes, avatares y chat-bots. *Investigación de rendimiento*, 20 (2), 70mi77.
- Taddeo, M. (2010). Confianza en la tecnología: un distintivo y un relación problemática. *Conocimiento, tecnología y políticas*, 23 (3), 283mi286.
- Waytz, A., Heafner, J. y Epley, N. (2014). La mente en el máquina: el antropomorfismo aumenta la confianza en un vehículo autónomo. *Revista de psicología social experimental*, 52, 113mi117.
- Weizenbaum, J. (1966). ELIZADun programa de computadora para el estudio de la comunicación en lenguaje natural entre el hombre y la máquina. *Comunicaciones del ACM*, 9 (1), 36mi45.
- Weizenbaum, J. y McCarthy, J. (1976). Poder de la computadora y Razón humana: Del juicio al cálculo. Nueva York, NY: WH Freeman y compañía.
- Wimmer, H. y Perner, J. (1983). Creencias sobre creencias: Representación Representación y función restrictiva de creencias erróneas en la comprensión del engaño por parte de los niños pequeños. *Cognición*, 13 (1), 103mi128.
- Zadrozny, W., Budzikowska, M., Chai, J., Kambhatla, N., Levesque, S. y Nicolov, N. (2000). Diálogo en lenguaje natural para una interacción personalizada. *Comunicaciones de la ACM*, 43 (8), 116mi120.