

Comunicación, diseño y IoT: Perspectivas para la interfaz gráfica con diseño de voz

Everaldo Pereira⁽¹⁾, Ana Paula Scabello Mello⁽²⁾,
Flávia Janine Rosante Beo⁽³⁾, Rafael Angel Bordenabe⁽⁴⁾,
Daniel Djinishian de Briquez⁽⁵⁾ e Ivy Abreu Brandt⁽⁶⁾

Resumen: Este texto aborda la temática de la comunicación, el diseño y el Internet de las cosas (IoT), centrándose en el desarrollo de una interfaz gráfica con asistente de voz (skill) destinada al dispositivo Echo Show con sistema Alexa. Su objetivo principal es ahondar en los conceptos de Internet de las cosas como medio, basándose en investigaciones previas. Quiere comprender la experiencia de interacción y la relación entre diseño y programación en IoT. Desde una perspectiva teórica, se aplican los supuestos del pensamiento de diseño, y el Interaccionismo Simbólico. En cuanto a la metodología, se sigue el enfoque de la *Design Science Research*, orientado a abordar problemas complejos mediante la construcción y evaluación de artefactos. Esta metodología busca integrar la interpretación del fenómeno desde la perspectiva del diseño y la ciencia de datos desde el punto de vista de la IoT, con el propósito de comprender y resolver problemas complejos. Los resultados obtenidos pueden contribuir a la creación de directrices para habilidades en productos IoT desde una perspectiva de diseño, teniendo en cuenta los aspectos simbólicos, los elementos dialógicos de la IoT, la comprensión de los análisis predictivos de las organizaciones y sus consideraciones éticas, con miras a fomentar un consumo responsable.

Palabras clave: Diseño - IoT - Investigación en Ciencias del Diseño - Comunicación - Medios

[Resúmenes en castellano y en portugués en las páginas 252-253]

⁽¹⁾ **Everaldo Pereira** es Doctor en Comunicación Social, con línea de investigación en Comunicación Institucional y de Marketing. Máster en Comunicación Social, con énfasis en Procesos Comunicacionales y Comunicación de Mercado. Especialista en Planeación y Licenciado en Comunicación Social por la Universidad Metodista de São Paulo. Profesor e Coordinador del curso de Diseño en el Instituto Mauá de Tecnología. Miembro de los grupos de investigación “LabDesign” en Mauá e “IA Responsable” do Instituto de Estudos Avançados (USP). Investigador asociado a Intercom - Sociedade Brasileira de Estudos Interdisciplinares da Comunicação. Miembro de la Red de Investigadores de Diseño de la Universidad de Palermo-UP. con experiencia en las áreas de Marketing, Diseño y Comunicación. everaldo@maua.br

(2) **Ana Paula Scabello Mello** es Doctor en Ciencias, área de concentración Diseño y Arquitectura. Máster en Estructuras Ambientales Urbanas y licenciada en Arquitectura y Urbanismo por la Universidad de São Paulo USP. Profesor del curso de Diseño del Instituto Mauá de Tecnología. Ergonomista certificada por la Asociación Brasileña de Ergonomía ABERGO (2010). Miembro del grupo de investigación LabDesign: procesos creativos, experiencia e innovación en el Instituto Mauá de Tecnología CEUN-IMT. Miembro de la Red de Investigadores de Diseño de la Universidad de Palermo UP. Mentora del Programa GCSP - Grand Challenges Scholars Program. Desarrolla investigaciones relacionadas con Ergonomía, Diseño Universal, UX/UI y Comunicación, Diseño y *Internet of Things* (IoT). anamello@maua.br

(3) **Flávia Janine Rosante Beo** es Ingeniera de software e investigador al que le encanta aprender sobre nuevas tecnologías. Profesor del curso de Ciencia de la Computación del Instituto Mauá de Tecnología. Interesado en ingeniería de software, aprendizaje automático, diseño UI/UX y campos de estudio de informática y computación cuántica. flavia.janine@maua.br

(4) **Rafael Angel Bordenabe** es Diseñador Gráfico en las áreas de UX/UI e identidad visual. Estudiante de Grado en Diseño en el Instituto Mauá de Tecnología (IMT). Miembro del Programa GCSP - Grand Challenges Scholars Program (IMT). Participante del Programa de Iniciación Científica (IMT). 22.01064-5@maua.br

(5) **Daniel Djinishian de Briquez** es Estudiante de Grado en Ciencia de la Computación en el Instituto Mauá de Tecnología (IMT). Miembro del Programa GCSP - Grand Challenges Scholars Program (IMT). 22.00251-0@maua.br

(6) **Ivy Abreu Brandt** es Estudiante de Grado en Diseño en el Instituto Mauá de Tecnología (IMT). 22.01722-4@maua.br

Introducción

Los asistentes de voz están evolucionando fuertemente hacia una convergencia con las interfaces gráficas. Los asistentes de voz son aplicaciones basadas en inteligencia artificial y programación de lenguaje natural que se encuentran en varios productos y servicios en IoT, como el dispositivo Echo Show 15 de Amazon.

Los sensores presentes en el Internet de las cosas (IoT) tendrán un papel importante en la definición del perfil de lo que hace el usuario y el contexto específico de las actividades que constituyen elementos clave para una comunicación eficaz y ética. En este sentido, hay varios desafíos tecnológicos y científicos, como comprender si los dispositivos están predis-

puestos a recibir medios programáticos, cuáles son los protocolos de creación, producción y medios para el diseño gráfico con diseño de voz en el IoT, además de los productos con recursos visuales y sin recursos visuales.

Un desafío actual es crear un sistema visual y sonoro capaz de adaptarse y convertir la fragmentación actual del mundo del IoT en un lenguaje común para permitir la comunicación amplia en el IoT con enfoque en el usuario (Aksu *et al.*, 2018). Los productores de contenido para el IoT deben concebir y diseñar para destinos heterogéneos. Desde el punto de vista del diseño, esto puede resolverse desde la perspectiva de un diseño gráfico responsivo, que se adapta al formato de pantalla de destino. Desde el punto de vista del diseño de voz, es necesario integrar mensajes responsivos de acuerdo con la capacidad del dispositivo. Sin embargo, aún queda adaptarse a la experiencia del usuario también en productos sin recursos visuales.

Este texto se basa en investigaciones anteriores en las que analizamos una interfaz gráfica de anuncios interactivos en el ámbito de la Internet de las Cosas (IoT). El objetivo general es estudiar la Internet de las Cosas como medio desde una perspectiva de diseño. En este sentido, se vislumbran perspectivas interdisciplinarias centradas en la interacción entre personas y dispositivos de la Internet de las Cosas, entre diseño y programación. En nuestra investigación anterior (Pereira *et al.*, 2022), realizamos una revisión bibliográfica, principalmente sobre medios de comunicación digital, basada en enfoques de una teoría del medio, sobre la experiencia del usuario y IoT. Adquirimos el dispositivo Echo Show 8 (Ver Figura 1), de la empresa Amazon, que cuenta con el sistema de asistente de voz Alexa, como foco de experiencias en IoT, y realizamos observaciones de usos e interacciones. Realizamos investigación sobre anuncios interactivos ya existentes. Realizamos pruebas de interacción mediante software de prototipado gráfico digital Adobe XD. Llevamos a cabo entrevistas en profundidad con usuarios de dispositivos Echo, con sistema de asistente de voz Alexa. Desarrollamos un prototipo de *skill* para el 1961 Design Studio, del Instituto Mauá de Tecnología. Skill es el nombre por el cual se conocen las aplicaciones con sistema Alexa. Desarrollamos la metodología de pruebas de interfaz de usuario (UI).

En este sentido, nuestro objetivo fue comprender IoT como medio desde una interfaz de usuario de voz. Estas interfaces se caracterizan por ser naturales, conversacionales y centradas en el usuario. Al crear una experiencia de voz, debemos percibir las diversas formas en que las personas expresan significado e intención durante la interacción, haciendo que las interfaces de voz sean más flexibles que las interfaces gráficas. Sin embargo, las plataformas de IoT por comando de voz están convergiendo rápidamente hacia interfaces gráficas con asistentes de voz. Por lo tanto, el diseñador no crearía para la voz de la misma manera que crea interfaces gráficas para la web o dispositivos móviles.

Entendemos que hay una simbiosis entre las relaciones sociales y los medios de comunicación, uno actuando sobre el otro. Así como los medios fueron naturalmente creados por las relaciones sociales, como principios económicos de almacenar y transmitir información, al mismo tiempo, alteran las relaciones sociales al permitir diferentes tipos de diálogo y discursos en diferentes escalas públicas.

Comprendemos que actualmente las escalas en IoT son más rápidas, con más nodos de red, desterritorializadas y, en cierto sentido, atemporales, posibilitadas por las tecnologías

digitales en red. Las perspectivas de la IoT son que estas escalas en el futuro sean gradualmente mayores.

El entorno de IoT como medio está evolucionando rápidamente hacia un conjunto de dispositivos IoT que cooperan colectivamente para implementar y exponer al usuario múltiples funcionalidades, así como para proporcionar anuncios interactivos.

Un desafío actual es crear un sistema de interfaces capaz de adaptarse y convertir la fragmentación actual de los dispositivos IoT en un lenguaje común para permitir una red de publicidad IoT.

La IoT ha cambiado la experiencia de los usuarios y, como resultado, los métodos utilizados anteriormente para las evaluaciones de UX (experiencia del usuario) pueden no ser aplicables al contexto de la experiencia en IoT. Los métodos tradicionales relacionados con pruebas de usabilidad resultaron adecuados para evaluar sistemas IoT, pero pueden no serlo para analizar aspectos de UX y las particularidades de las aplicaciones IoT, con sus múltiples dispositivos y posibilidades de escenarios.

Estudiar la interacción de IoT como medio y las personas desde un punto de vista del diseño, con una metodología interpretativa y proposicional que resulte en un artefacto, es decir, en prescripciones, recomendaciones y orientaciones, puede profundizar el estado del arte en este campo. Los resultados pueden ayudar a organizaciones, gobiernos y personas a interpretar la Internet de las Cosas como medio de manera integral. Entendemos que el enfoque de la propuesta es innovador e interdisciplinario, ya que une, a través de la Investigación en Ciencias del Diseño, la práctica proyectual del diseño y la ciencia de la computación, construyendo una investigación no paradigmática.

Una metodología proposicional

Se llevaron a cabo discusiones y propuestas de acciones, así como el desarrollo colaborativo de textos científicos, utilizando como tema el desarrollo de una habilidad (skill) para donaciones al Instituto Uno, socio institucional de la investigación. El Instituto Uno es una organización brasileña sin fines de lucro que se enfoca en promover la educación de calidad y el desarrollo de habilidades socioemocionales en niños y jóvenes que tienden a tener rezagos educativos.

En primer lugar, profundizamos en la bibliografía, especialmente en medios de comunicación digital, inteligencia artificial aplicada al diseño de voz, diseño gráfico responsivo, principalmente aplicado en interfaces gráficas y de voz. También profundizamos en la investigación bibliográfica sobre la experiencia del usuario y la Internet de las cosas (IoT), comprendiendo cómo la red de servicios y datos recopilados a través de sensores aplicados a objetos y dispositivos ha impactado significativamente en la experiencia de los usuarios, especialmente en teléfonos inteligentes, relojes inteligentes, televisores interactivos, electrodomésticos inteligentes y sistemas de automatización residencial.

Adquirimos el dispositivo Echo Dot, Echo Show 8 y Echo Show 15 (*Ver Figura 1*) de la empresa Amazon, que cuenta con el sistema de asistente de voz Alexa, como enfoque de

experiencias en IoT. Realizamos observaciones de usos e interacciones para evaluar la fragmentación de dispositivos y su impacto en el diseño responsivo.



Figura 1.
Imagem ilustrativa del dispositivo Echo Dot, Echo Show 8 e Echo Show 15 (Fuente: Divulgación de Amazon).

Realizamos una investigación para comprender cómo se lleva a cabo la interacción por voz con el asistente personal Alexa para llevar a cabo una tarea específica, hacer una donación, con el propósito de obtener información para el desarrollo de una habilidad (skill) de donaciones para Alexa. Se llevó a cabo una investigación cualitativa a través de la observación, donde voluntarios intentaron realizar una donación en dos modelos de dispositivos Echo, con y sin pantalla. De los cuatro grupos que utilizaron el dispositivo con pantalla, tres lograron completar la tarea, mientras que ninguno de los tres grupos que utilizó el dispositivo sin pantalla tuvo éxito. Se observó que predominó el uso de los comandos “hacer o realizar donación” y que esta intención no fue comprendida por la inteligencia artificial (IA). Los grupos que utilizaron Google para buscar sitios de instituciones para recibir la donación lograron completar la tarea, finalizando la acción a través de la interfaz visual. Concluimos que una habilidad de donación para Alexa debe tener en cuenta la dificultad de búsqueda por parte del usuario y explorar la interacción a través de la pantalla del teléfono celular en el caso de dispositivos sin pantalla.

Creamos y analizamos un flujo de conversación a través de la aplicación Voiceflow (Ver Figura 2) con el objetivo de desarrollar una habilidad de donación para el Instituto Uno para su uso en el dispositivo IoT Echo Show 15 con sistema Alexa. La aplicación Voiceflow es una plataforma en línea que permite la creación y diseño de aplicaciones de voz sin necesidad de codificación, lo que resulta prometedor para la IoT con comandos de voz e imagen como plataforma de medios. Una de las principales dificultades hoy en día son los diferentes tipos de dispositivos IoT existentes (Aksu et al., 2018). Observamos que Voiceflow permite a diseñadores y comunicadores crear y gestionar conversaciones de voz interactivas, como habilidades para dispositivos Echo con sistema Alexa. Una de las ventajas de esta plataforma es la capacidad de diseñar flujos de conversación complejos, con varias posibilidades de ida y vuelta, es decir, de una conversación menos unidireccional, como

es el caso de la tarea de “donación” que pretendemos desarrollar para Alexa. También evaluamos cómo crear respuestas personalizadas y agregar lógica de programación para crear una experiencia de conversación intuitiva. Desde el punto de vista de la producción comunicativa, esta aplicación ofrece una interfaz visual fácil de usar, donde los diseñadores y comunicadores pueden arrastrar y soltar elementos para crear un flujo de conversación específico. Notamos que Voiceflow también ofrece integración con otras plataformas y servicios, como APIs de otras empresas, permitiendo que diseñadores y comunicadores conecten sus aplicaciones en desarrollo a otros sistemas y servicios, como PayPal.

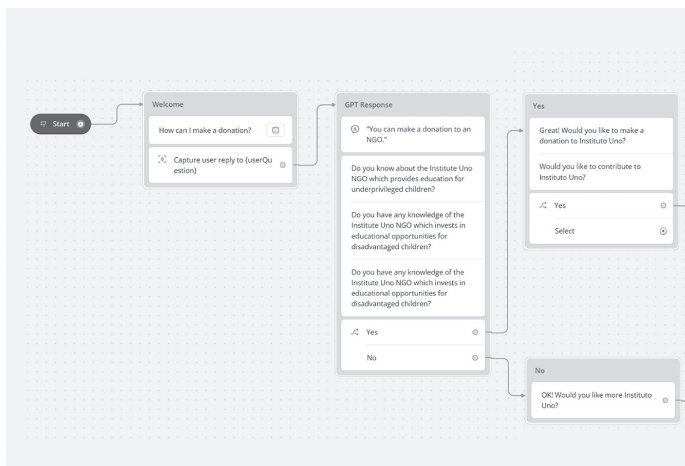


Figura 2. Modelo preliminar de diálogo para donación con el uso de Alexa (Fuente: los autores, 2023).

Para garantizar consistencia, utilizamos como base modelos preexistentes de Amazon, como el Alexa Skills Kit (ASK), que define un conjunto de palabras que los usuarios dicen para invocar una habilidad (Ver Figura 3). Sin embargo, vimos la necesidad de una evaluación crítica de los modelos existentes desde el Interaccionismo Simbólico.

El Interaccionismo Simbólico se originó en el trabajo de George Herbert Mead (1934), quien demostró que los egos de las personas son productos sociales, aunque también tienen propósitos y son creativos. Otros precursores en el área fueron Herbert Blumer y Charles Cooley. Blumer (1969), un estudioso de Mead y creador del término “Interaccionismo Simbólico”, destacó las principales perspectivas de este enfoque: las personas actúan en relación con las cosas basándose en el significado que esas cosas tienen para ellas; y esos significados son el resultado de su interacción social y se modifican por su interpretación. Blumer siguió a Mead y presentó la proposición de que las personas interactúan entre sí a través de la interpretación mutua de las acciones y la definición mutua, en lugar de

simplemente reaccionar a las acciones del otro. Sus respuestas no se dan directamente a las acciones del otro, sino que se basan en el significado que atribuyen a esas acciones. Así, la interacción humana está mediada por el uso de símbolos y significados, a través de la interpretación o determinación del significado de las acciones del otro. Blumer comparó este proceso, al que denominó “interacción simbólica”, con las explicaciones conductistas del comportamiento humano que no consideran la interpretación entre estímulo y respuesta, como en Schram (1949), entre otros. Cabe destacar que actualmente muchos objetos utilizan la interacción mediante inteligencia artificial y, por lo tanto, podrían llevar a cabo conversaciones con significados específicos de antemano. Naturalmente, estas conversaciones son programadas por otras personas que, incluso dentro de sus corporaciones, pueden generar campos de significado.

Estos profesionales pueden definir sus habilidades para aceptar solicitudes predefinidas. Se observó que el ASK (Ver Figura 1) ofrece varios tipos diferentes de habilidades preexistentes para elegir, o permite la construcción de una habilidad personalizada utilizando predefiniciones de declaraciones, intenciones y *slots* dentro de un panel de desarrollador de Alexa. Para el desarrollo de este recorte, se analizó la creación de una habilidad para donar a una institución denominada Instituto Uno.

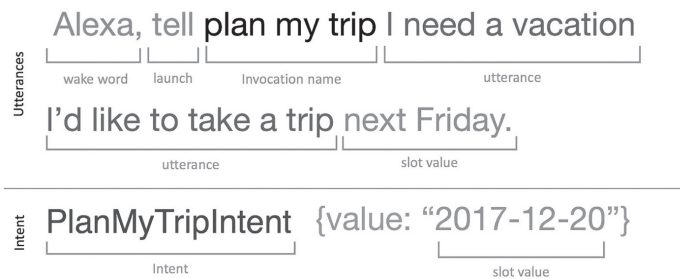


Figura 3. Conceptos del modelo ASK (Fuente: <https://developer.amazon.com/en-US/docs/alexa/custom-skills/create-the-interaction-model-for-your-skill.html>).

Los modelos ASK resaltan la importancia de comprender variables en la construcción de interfaces de voz para Alexa. Se indican los conceptos de palabras de activación, lanzamiento e invocación, cada uno desempeñando un papel específico en la interacción del usuario. El modelo ASK permite la inclusión de mensajes para dirigir interacciones y solicitar información. Personalizamos el modelo para solicitar intenciones, acciones que respondieran a las solicitudes de los usuarios, y espacios que son valores de entradas variables para ayudar a Alexa a comprender la intención del usuario. Esto nos permitió explorar diversas formas de formular solicitudes por voz.

Para el desarrollo de la habilidad de donación, probamos el entrenamiento de grandes modelos de lenguaje (Ver Figura 4), denominados *Foundation Models* (FMs), con un enfoque en el reconocimiento de voz, haciendo hincapié en el entrenamiento auto-supervisado y supervisado. Exploramos técnicas de modificación de datos para mejorar la precisión del reconocimiento del habla en asistentes de voz. En este sentido, proponemos la relación del Interaccionismo Simbólico (IS) con los FMs para generar conjuntos de datos estructurados, con el objetivo de mejorar la assertividad de los modelos de clasificación. Destacamos la necesidad de explorar la técnica de aumento de conjuntos de datos de intenciones para el entrenamiento de modelos en asistentes de voz, buscando posibles mejoras en la interacción con usuarios de dispositivos IoT.

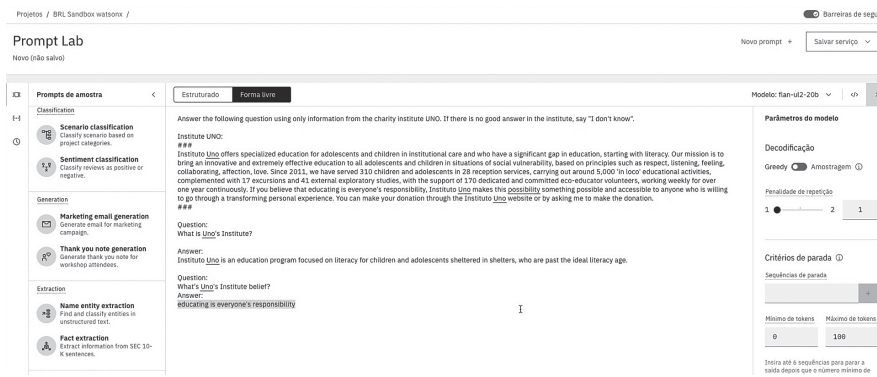
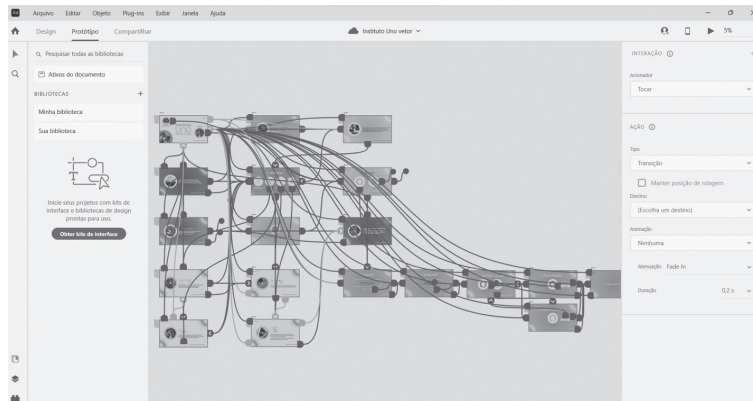


Figura 4. Entrenamiento de IA en Modelos de Fundación para una tarea de donación (Fuente: los autores, 2023).

Creamos la interfaz visual (Ver Figuras 5 y 6) y profundizamos en la metodología de pruebas de interfaz de usuario (UI), aplicándola en un entorno de laboratorio a través del software Adobe XD, la plataforma Zoom y Microsoft Forms. Analizamos las pruebas de interfaz con usuarios utilizando la aplicación Adobe XD y observamos que el prototipo con el uso de la herramienta Adobe XD aún es inestable, provocando errores en la interpretación de audio y reconocimiento de voz. En algunos casos, el prototipo solo respondía a comandos en inglés. Hubo dificultades en la interpretación de diferentes tipos de voces.



5



6

Figura 5. Estudos de design visual para a habilidade de doação a la organização social (Fuente: los autores, 2023). **Figura 6.** Prototipo de habilidad para el Instituto Uno (Fuente: los autores, 2023).

Realizamos la programación de la habilidad en la Consola de Desarrollador de Alexa (Ver Figuras 7 y 8) de la empresa Amazon. Observamos que la plataforma aún requiere conocimientos especializados de programación, pero está migrando rápidamente hacia una programación de bajo código, lo que creemos que será una de las habilidades necesarias para los diseñadores que deseen desarrollar proyectos en IoT como medio.

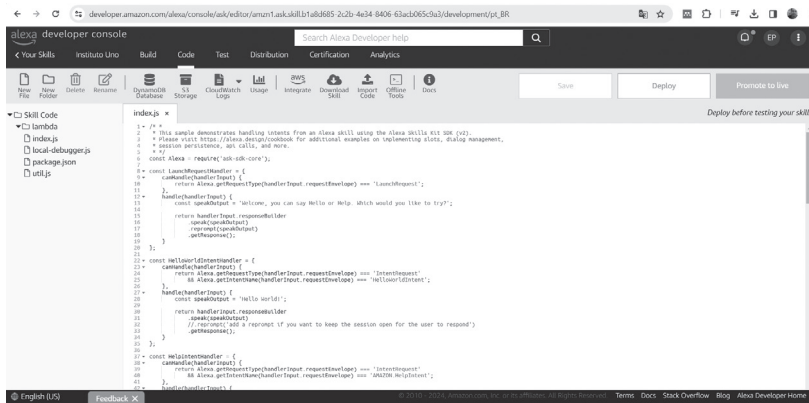


Figura 7. Programación de la habilidad en la Consola de Desarrollador de Alexa (Fuente: los autores, 2023).

Figura 8. Prototipo de la habilidad en el Echo Show 15 con Alexa (Fuente: los autores, 2023).

Análisis de usabilidad

La metodología adoptada comprendió observaciones presenciales de las interacciones del prototipo con y sin pantalla en una sala no ambientada para IoT, las cuales fueron grabadas en video mediante la plataforma Teams, junto con la aplicación de un cuestionario *online* con una duración de prueba aproximada de 30 minutos. Aunque este tiempo podría haber causado cierto malestar, no fue reportado por los participantes. Las grabaciones fueron transcritas utilizando la aplicación Google Pinpoint y analizadas tanto a través de los registros de video como de los textos extraídos. Este método permitió interpretar los gestos, tiempos de interacción, así como el contenido hablado y escuchado durante la sesión. Los resultados revelaron un perfil mayoritariamente masculino, con edades entre 20 y 24 años, cursando estudios superiores. La mayoría de los participantes se comunica en portugués e inglés, y la frecuencia de uso de dispositivos con asistentes de voz no es elevada. Aunque la interfaz fue percibida como fácil de usar, surgieron algunas preocupaciones sobre su intuición, especialmente en relación con las dificultades de reconocimiento de voz por parte del prototipo.

Las opiniones de los participantes abarcaron aspectos tanto de la interacción por voz como del diseño visual de la interfaz. A pesar del consenso en cuanto a la comprensibilidad y amigabilidad de la voz, se mencionaron algunas cuestiones relacionadas con la intuición, funcionalidad y estética de la interfaz gráfica.

Además, los análisis individuales durante las pruebas resaltaron elementos como la practicidad, velocidad del sistema y la necesidad de una palabra de activación clara (“Alexa” en este caso), proporcionando información esencial para ajustes y mejoras durante la revisión del diseño de la interfaz.

La análisis de los resultados de la investigación revela desafíos e ideas relacionadas con la interpretación de los usuarios durante la interacción con la interfaz en estudio. Algunos participantes demostraron dificultades para comprender cómo continuar con el proceso, indicando una brecha en la claridad de las instrucciones o la estructura de la interfaz. La interpretación de la información de texto como comandos ejecutables sugiere la necesidad de refinar la comunicación textual para evitar ambigüedades y garantizar una comprensión más precisa por parte de los usuarios, especialmente en información de contenido versus intenciones de los usuarios.

La observación de que la pantalla puede condicionar a los usuarios a navegar como en un sitio web, sin utilizar la interacción por voz, destaca la importancia de equilibrar las opciones de navegación para satisfacer las diversas preferencias y necesidades de los usuarios. La comparación entre el prototipo en Adobe XD y los dispositivos Echo subraya la complejidad de las intenciones, indicando que las interpretaciones de comandos pueden variar según los slots y el enfoque específico adoptado. Esto señala la necesidad de una comprensión más profunda de la dinámica de las interacciones por voz en sistemas IoT.

La constatación de que la palabra de activación (“Alexa”) no siempre es observada plantea preguntas sobre la habitualidad de los usuarios en relación con este aspecto. La importancia de recordar la palabra de activación, especialmente en entornos públicos, destaca la necesidad de estrategias claras para garantizar la eficacia de la interacción por voz en diferentes contextos, ya sean públicos o privados. El uso de botones de pantalla como medio de navegación en lugar de la interacción por voz indica la diversidad de preferencias entre los usuarios y destaca la importancia de ofrecer opciones flexibles.

La repetición de sugerencias de intención como indicativo del recorrido conversacional refuerza la relevancia de estas sugerencias en la orientación de los usuarios durante la interacción. Sin embargo, la identificación de problemas en la ejecución del comando de voz relacionado con el módulo de donación destaca la necesidad de mejoras específicas en esta área.

Otras observaciones, como la falta de familiaridad del usuario con la palabra de activación y la identificación de intenciones no registradas, subrayan la importancia continua de la investigación y ajuste de la interfaz para satisfacer las expectativas y competencias de los usuarios. La interpretación de los resultados también señala oportunidades de refinamiento en la comunicación, usabilidad y adaptabilidad de la interfaz, buscando una interacción más eficiente y satisfactoria.

Las observaciones extraídas de los resultados de la investigación sugieren mejorar la enseñanza-aprendizaje en sistemas de conversación, especialmente en el contexto de la interacción con la interfaz en estudio. El reconocimiento de la importancia de investigar

las dinámicas de introversión y extroversión en el uso de Alexa sugiere la necesidad de considerar las características individuales de los usuarios para optimizar la experiencia de interacción y puede ser objeto de futuras investigaciones.

En cuanto al método de análisis, se observó que la transcripción automática mediante Pinpoint pierde datos, resaltando la importancia del tratamiento adecuado de los datos transcritos para garantizar la integridad y precisión de la información recopilada, lo que requiere una revisión del investigador antes del análisis de contenido. La sugerencia de realizar un entrenamiento inicial para los estudiantes de Iniciación Científica (IC) destaca la importancia de capacitar a los jóvenes investigadores para abordar efectivamente las sutilezas de la investigación.

Se reconoce la necesidad de investigar la interferencia del investigador, ya que es crucial minimizar cualquier influencia externa que pueda distorsionar los resultados de la interacción, especialmente dado que la mayoría de los dispositivos IoT se acceden en entornos sin instrucciones externas al dispositivo. Además, el análisis de la diferencia entre entornos público y privado en el uso de comandos de voz destaca la relevancia de adaptar la interfaz a las diferentes situaciones de uso y podría ser objeto de futuras investigaciones.

Las instrucciones de prueba se destacan como un punto de atención, sugiriendo la importancia de hacer el procedimiento de prueba claro y menos sujeto a la interferencia del investigador. Las fallas en el retorno de voz, percibidas en Adobe XD, evidencian desafíos técnicos que deben abordarse para garantizar una evaluación precisa de la interfaz incluso en prototipo. También se destaca la necesidad de probar en una sala aislada para la prueba y una sala ambientada.

La complejidad de hacer que la conversación sea más natural se identifica como un desafío, indicando la necesidad de refinamientos en la interfaz para lograr una comunicación más fluida e intuitiva.

Conclusiones

Destacamos la necesidad de profundizar en diversos aspectos debido a la diversidad de equipos e interacciones. Enfatizamos la importancia de las discusiones interdisciplinarias para prescribir contenidos en comunicación, diseño gráfico y experiencia del usuario en IoT. Observamos la complejidad de la experiencia del usuario en sistemas IoT y sugerimos adaptaciones en los métodos tradicionales de pruebas de usabilidad, UX y UI, especialmente en contenidos por voz e imagen. Presentamos un modelo de proceso de desarrollo de diseño gráfico y de voz, destacando la etapa de briefing, estudios de UX, elaboración de un plan de habilidad, desarrollo de un *roadmap* flexible para interacción y prototipado responsiva. También destacamos los desafíos relacionados con la falta de compartimiento fácil de interfaces entre diferentes sistemas de equipos IoT. Como prescripciones para diseñadores, se sugiere la colaboración entre diseñadores y ciencia de datos, pruebas de software de prototipado y la inclusión de cuestiones más intuitivas y orientadas hacia una conversación natural en *roadmaps* de interacción. Se observa la importancia de considerar

la memoria y contextos de uso y datos en la programación de mensajes para sistemas de conversación IoT.

Comprendemos, en la intersección entre Diseño y Ciencias de la Computación, que una interfaz de voz que utiliza un FM asociado al Interaccionismo Simbólico puede ser capaz de entender mejor comandos de voz complejos, responder a preguntas con más precisión o generar respuestas más naturales y contextualmente relevantes, importantes en el contexto de la propuesta de modelo de proceso comunicacional Persona-IoT (*Ver Figura 9*) presentado en nuestro informe anterior. Además, esta asociación podría usarse para mejorar la capacidad de las interfaces de voz para realizar tareas específicas, como hacer donaciones, realizar reservas o proporcionar información detallada sobre un tema específico. Dado que los FM se entrenan en grandes cantidades de texto que contienen información simbólica y significados culturales y aprenden a comprender y generar lenguaje, incluyendo símbolos y significados asociados a palabras y frases, es relevante observarlos desde una perspectiva teórica que se centra en cómo los seres humanos desarrollan el sentido y el significado a través de la interacción social. Al utilizar FM en interfaces de voz e imagen (IVIU), podríamos explorar cómo estos modelos pueden ayudar en la comprensión y generación de lenguaje simbólico.

Al evaluar estos modelos de fundación en tareas centradas en el ser humano y explorar más profundamente sus capacidades, se observaron aciertos significativos de los modelos. Algunos análisis manuales detallados también revelaron las limitaciones de estos grandes modelos de lenguaje en términos de comprensión, utilización del conocimiento, razonamiento y cálculo. Esto alienta a promover el desarrollo de modelos más alineados con la cognición humana, permitiendo que manejen una gama más amplia de tareas complejas y centradas en el ser humano con mayor precisión y confiabilidad.

Además, los FM pueden entrenarse junto con datos que incluyen interacciones sociales reales para capturar matices y sutilezas del lenguaje simbólico utilizado por los seres humanos. Esto puede ayudar a que las interacciones sean más naturalmente significativas y relevantes para los usuarios. Sin embargo, es importante tener en cuenta que los modelos de fundación son técnicas de aprendizaje automático y no poseen comprensión o intención verdaderas. Aprenden a identificar patrones en los datos de entrenamiento, pero hasta el momento, no podemos afirmar que tengan una comprensión conceptual o simbólica del mundo. Por lo tanto, al utilizar estos modelos en aplicaciones basadas en el Interaccionismo Simbólico, es necesario considerar sus limitaciones y garantizar que la interpretación y generación de símbolos se realicen de manera adecuada y ética, teniendo en cuenta el contexto y la diversidad cultural.

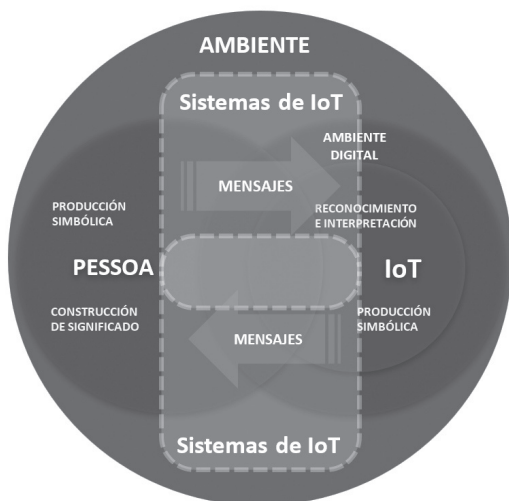


Figura 9.
Propuesta de modelo de proceso comunicacional Persona-IoT (Fuente: los autores, adaptado de Pereira, 2018).

Los análisis permitieron desarrollar procesos creativos para el diseño en IoT con enfoque en prescripciones para un proceso de desarrollo, pero aún no para un futuro sistema gestor de entrega de interfaces para dispositivos IoT y, por lo tanto, para el usuario, como un protocolo de medios programáticos. Este sistema gestor en medios IoT podría ser objeto de investigación futura en conjunto con el campo de la Administración, comprendiendo estudios de diseñadores, editores, HUBs de Medios, industrias de IoT y usuarios; se puede investigar cómo crear estrategias de contenido en IoT a partir de la conciencia del contexto del usuario.

La investigación del comportamiento del usuario, antes de la creación y evaluación de la experiencia del usuario (UX) de los prototipos de anuncios interactivos, fue importante para revelar hábitos de uso y actitudes de conversación distintos de los hábitos de uso en dispositivos exclusivamente con pantallas. El prototipo ayuda en el desarrollo de recomendaciones para mejorar la experiencia de los usuarios con las interfaces visuales y sonoras de dispositivos IoT, considerando la complejidad de las intenciones por voz. Una investigación futura en este campo podría evaluar si hay convergencias o divergencias entre dispositivos originalmente con pantallas que incorporaron asistentes de voz, como Smart TVs, y dispositivos originalmente de asistencia de voz que incorporaron pantallas, como Echos. Una hipótesis es que no haya muchas divergencias entre estos dispositivos y también en los hábitos de los usuarios de estos dispositivos.

Los contenidos interactivos por asistentes de voz aún carecen de profundización en diversos aspectos, como una mayor aproximación de profesionales de varias áreas; deben considerar la amplia diversidad de equipos y sistemas IoT en el momento; deben analizar

y comprender el comportamiento del usuario en relación con los contenidos en sistemas IoT.

Los resultados aquí presentados contribuyen a la comprensión de la comunicación a través del diseño en IoT desde una perspectiva de diseño basada en aspectos simbólicos, aspectos dialógicos de IoT, la comprensión de los análisis predictivos de las organizaciones y sus consideraciones éticas para un consumo responsable.

Referencias bibliográficas

- Aksu, H., et al. (2018). Advertising in the IoT Era: Vision and Challenges. Department of Electrical and Computer Engineering Florida International University, Miami, FL, USA. arXiv, v1, 31 Jan 2018. ISSN 1802.04102. Disponible em <https://arxiv.org/abs/1802.04102>.
- Almeida, R. L. A. (2018). CHASE: checklist para avaliação da experiência do usuário em ambientes de internet das coisas. Dissertação de Mestrado, Universidade Federal do Ceará.
- Amin-Nejad, A., Ive, J., & Velupillai, S. (2020). Exploring transformer text generation for medical dataset augmentation. In Proceedings of the Twelfth Language Resources and Evaluation Conference, pp. 4699-4708.
- Andler, S. (1979). Predicate path expressions. In Proceedings of the 6th ACM SIGACT-SIGPLAN Symposium on Principles of Programming Languages (POPL '79), 226-236. DOI: <https://doi.org/10.1145/567752.567774>.
- Blumer, H. (1969). Symbolic Interactionism: Perspective and Method, Englewood Cliffs, Prentice-Hall.
- Blumer, H. (1969). Symbolic Interactionism: Perspective and Method. Englewood Cliffs, NJ: Prentice-Hall.
- Bommasani, R., Hudson, D. A. (2021). On the opportunities and risks of foundation models. arXiv:2108.07258.
- Cho, E., & Kumar, S. (2018). A conversational neural language model for speech recognition in digital assistants. In 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE.
- Deibel, D., & Evanhoé, R. (2021). Conversations with Things: UX Design for Chat and Voice. New York: Rosenfeld Media. ISBN: 1-933820-26-8.
- Dresch, A. (2015). Design Science Research. Porto Alegre: Bookman.
- Hassenzahl, M. (2007). The hedonic/pragmatic model of user experience. Towards a UX manifesto, 10, 10-14.
- Jurafsky, D., & Martin, J. H. (2000). Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Prentice Hall PTR, Upper Saddle River, NJ, USA.
- Mahesh, B. (2020). Machine Learning algorithms-a review. International Journal of Science and Research (IJSR), 9(1), 381-386.
- McDonald, D. D. (2010). Natural language generation. Handbook of natural language processing, 2, 121-144.
- Mead, G. H. (1934). Mind, self, and society. Chicago: University of Chicago Press.

- Mead, G. H. (1934). *Mind, self, and society*. Chicago: University of Chicago Press.
- Mello, A. P. S., Pereira, E., Mellito, M. A. (2023). “Design, Comunicação e IoT: considerações a respeito da experiência de usuários do assistente pessoal Alexa em dispositivos Echo”. In: *Anais do ERGODESIGN & USIHC 2023 & JOP’Design 2023*. São Paulo: Blucher. ISSN 2318-6968, DOI 10.5151/ergodesign2023-74.
- Narayanan, A., et al. (2018). Toward Domain-Invariant Speech Recognition via Large Scale Training. in *Proc. SLT*.
- Pereira, E. (2018). *Comunicação de mercado e ressignificação de marca*. São Paulo: Edições Acadêmicas.
- Pereira, E., & Mello, A. P. S. (2023). Internet das coisas como mídia: perspectivas e desafios para comunicação e design de anúncios. *Signos do Consumo*, 15(1), e210974. <https://doi.org/10.11606/issn.1984-5057.v15i1e210974>
- Pereira, E., Mello, A. P. S. (2023). Internet das coisas como mídia: perspectivas e desafios para comunicação e design de anúncios. *Signos do Consumo*, 15(1), e210974. DOI: 10.11606/issn.1984-5057.v15i1e210974.
- Pereira, E., Mello, A. P. S., BEO, F. J. R., Bordenabe, R. A., De Briquez, D. D., & Sawada, E. T. 2023. *Design de voz: conversando com as coisas*. En *XVIII Encuentro Latinoamericano de Diseño - Semana Internacional de Diseño - Universidad Palermo*, Buenos Aires, Argentina. Recuperado de <https://www.youtube.com/watch?v=7JW1chjBrr0>
- Pereira, E., Mello, A. P. S., Correa, J. A., Mellito, M. A. (2022). Comunicação, design e tecnologia: perspectivas e desafios da IoT como mídia para anúncios interativos. *Anais do 45º Congresso Brasileiro de Ciências da Comunicação*, ISSN 2175-4683.
- Pinker, S. (2002). *O instinto da linguagem. como a mente cria a linguagem*. São Paulo: Martins Fontes.
- Raffel, C., et al. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. *J. Mach. Learn. Res.*, 21(140), 1–67.
- Schramm, W. (Ed.). (1949). *Mass Communications*. Urbana, IL: University of Illinois Press.
- Schramm, W. (Org.). (1949). *Mass Communications*. Urbana, IL: University of Illinois Press.
- Trindade, E., Perez, C., & Teixeira Filho, C. (2019). Tendências das pesquisas em publicidade e consumos nos periódicos nacionais e internacionais de comunicação: um panorama sobre o estudo do algoritmo. In *Anais...* Porto Alegre: Escola de Comunicações e Artes, Universidade de São Paulo. Recuperado de <https://www.eca.usp.br/acervo/producao-academica/002961807.pdf>

Abstract: This text addresses the topic of communication, design and the Internet of Things (IoT), focusing on the development of a graphical interface with voice assistant (skill) for the Echo Show device with Alexa system. Its main objective is to delve deeper into the concepts of the Internet of Things as a medium, building on previous research. You want to understand the interaction experience and the relationship between design and programming in IoT. From a theoretical perspective, the assumptions of design thinking, and Symbolic Interactionism are applied. In terms of methodology, it follows the ap-

proach of Design Science Research, oriented to address complex problems through the construction and evaluation of artefacts. This methodology seeks to integrate the interpretation of the phenomenon from the design perspective and data science from the IoT point of view, with the purpose of understanding and solving complex problems. The results obtained can contribute to the creation of guidelines for IoT product skills from a design perspective, taking into account the symbolic aspects, the dialogical elements of IoT, the understanding of organisations' predictive analytics and their ethical considerations, with a view to fostering responsible consumption.

Keywords: Design - IoT - Design Science Research - Communication - Media

Resumo: Este texto aborda o tema da comunicação, do design e da Internet das Coisas (IoT), com foco no desenvolvimento de uma interface gráfica com assistente de voz (skill) para o dispositivo Echo Show com o sistema Alexa. Seu principal objetivo é aprofundar os conceitos da Internet das Coisas como um meio, com base em pesquisas anteriores. Você quer entender a experiência de interação e a relação entre design e programação na IoT. Em uma perspectiva teórica, são aplicadas as premissas do design thinking e do Interaçionismo Simbólico. Em termos de metodologia, segue-se a abordagem da Design Science Research, orientada para tratar de problemas complexos por meio da construção e avaliação de artefatos. Essa metodologia busca integrar a interpretação do fenômeno a partir da perspectiva do design e da ciência de dados do ponto de vista da IoT, com o objetivo de entender e resolver problemas complexos. Os resultados obtidos podem contribuir para a criação de diretrizes para as habilidades de produtos da IoT sob a perspectiva do design, levando em conta os aspectos simbólicos, os elementos dialógicos da IoT, a compreensão da análise preditiva das organizações e suas considerações éticas, com o objetivo de promover o consumo responsável.

Palavras-chave: Design - IoT - Pesquisa em ciência do design - Comunicação - Mídia
