

## El desarrollo sociotécnico de un chatbot o ¿Cómo se construye una caja negra?

### RESUMEN

Este trabajo tiene el objetivo de analizar el proceso sociotécnico de desarrollo de *Katri*, un chatbot diseñado para apoyar la ejecución de un curso semipresencial de pedagogía en una institución brasileira de educación superior. Para tal fin, se siguió una metodología etnográfica que buscó analizar las principales actividades y elecciones sociotécnicas que influenciaron las características y la trayectoria de desarrollo del mismo, o dicho de otra manera, se siguió un abordaje etnográfico para “abrir la caja negra” de un chatbot. Se argumenta que el diseño e implementación del chatbot fue orientado por un proceso de personificación, sintetizado en 6 momentos claves que van pasando por su creación, crecimiento y madurez. Se concluye además, que en el proceso de desarrollo las capacidades de los actores para influir en las características sociotécnicas que incorpora el chatbot son desiguales.

**PALABRAS CLAVE:** chatbots, caja negra, análisis sociotécnico, algoritmos.

**Darío Reyes Reina**  
Hop Research Group, Belo Horizonte, Minas Gerais

**Leonardo Drummond Vilça**  
Instituto Ânima e CEFET-MG, Belo Horizonte, Minas Gerais

**Samira Spolidorio**  
UniBH - Centro Universitário de Belo Horizonte, Belo Horizonte, Minas Gerais e Departamento de Linguística Aplicada, Unicamp, Campinas, São Paulo

**Matheus Ícaro Martins**  
Hop Research Group, Belo Horizonte, Minas Gerais

## INTRODUCCIÓN

En un contexto, acaso cada vez más frecuente, en el cual nuevos sistemas sociotécnicos, llámense algoritmos, aplicaciones o asistentes virtuales hacen parte de nuestra día a día, es importante detenerse a reflexionar sobre su diseño, funcionamiento y articulación con la vida cotidiana.

En esta línea de ideas, el presente trabajo tiene el objetivo de analizar el proceso sociotécnico de desarrollo de Katri, un chatbot diseñado para apoyar la ejecución de un curso semipresencial de pedagogía en una institución brasilera de educación superior.

Para tal objetivo el artículo seguirá el siguiente orden. Comenzaremos presentando el referencial teórico sobre el concepto de chatbot y de caja negra. Luego será descrita la metodología, para continuar con los resultados del estudio que fueron sintetizados en 6 momentos claves del desarrollo del chatbot. Posteriormente, discutiremos sobre el proceso de desarrollo del chatbot y las posibilidades de abrir la “caja negra de los algoritmos”. Por último, se presentan algunas consideraciones finales y posibles futuras investigaciones.

## REFERENCIAL TEÓRICO

Aunque los primeros chatbots pueden remontarse a la década de los 60 del siglo pasado con el desarrollo de ELIZA (WEIZENBAUM, 1966), no es sino hasta hace pocos años que volvieron estar en el foco de atención (DALE, 2016). Aunque en los medios generalmente solo sobresalen noticias de chatbots de grandes empresas como Alexa de Amazon, Siri de Apple o Cortana de Microsoft, lo cierto es que su uso se ha extendido a múltiples contextos. Así, para 2017 se calculaba que en la plataforma Facebook Messenger podrían haber 100.000 chatbots funcionando, una cifra que en el 2018 ya se había multiplicado a 300.000 (JOHNSON, 2018).

Los chatbots han sido definidos como softwares que interactúan con las personas usando lenguaje natural (DALE, 2016). Se caracterizan por 4 elementos centrales: 1) buscan simular las conversaciones que tenemos entre humanos (CIECHANOWSKI; PRZEGALINSKA; WEGNER, 2018; MORRISSEY; KIRAKOWSKI, 2013) ; 2) tradicionalmente han proporcionado una interacción por medio de mensajes escritos, de ahí el prefijo “chat” (IO; LEE, 2018), sin negar, que recientemente se han dado avances que facilitan la interacción hablada; 3) contrario a los robots o similares son entes que “no tienen cuerpo” (*disembodied agents*) (ARAUJO, 2018); 4) se diferencian de los *avatar* pues no son representaciones de un humano en un mundo virtual (KLEVJER, 2006; PÜTTEN et al., 2010).

El segundo concepto que guía la investigación es el de caja negra. Coloquialmente, el concepto sirve para referirse a cualquier dispositivo tecnológico cuyo funcionamiento es opaco, de tal manera que solo se conocen las entradas que recibe (*inputs*) y los resultados que genera (*outputs*), más no el proceso interno o las operaciones que realiza (BUCHER, 2016; VON HILGERS, 2011).

Bucher e Von Hilger (BUCHER, 2016; VON HILGERS, 2011) señalan que el origen de este concepto se remonta a la segunda guerra mundial, cuando los ingenieros desarrollaban equipos militares, especialmente radares, cuyo funcionamiento debía ser inescrutable en caso de caer manos enemigas.

Posteriormente, dentro de los estudios sociales de la ciencia y tecnología, particularmente en los trabajos de Latour (LATOURE, 2001, 2005, 2011) el concepto ha venido jugando un papel fundamental no solo para entender a las tecnologías, sino también los hechos científicos y las corporaciones.

Igualmente, en los debates contemporáneos sobre los algoritmos este concepto ha tenido un papel protagónico. Tradicionalmente, desde la ciencia de la computación un algoritmo se ha entendido como un conjunto de instrucciones o de procesos codificados para transformar datos de entrada en algunas salidas basándose en cálculos específicos (DOURISH, 2016; GILLESPIE, 2013). Sin embargo, discusiones recientes caracterizan a los algoritmos más allá del simple código, como sistemas sociotécnicos que hacen parte de nuestra vida cotidiana produciendo y siendo producto de una serie de factores sociales, culturales, económicos y tecnológicos (GILLESPIE, 2013; SEEVER, 2017).

Este debate nos atañe, pues lleva a cuestionarnos hasta qué punto los algoritmos que hacen parte de nuestro día a día como el buscador de Google, el sistema de recomendación de Netflix o cualquier chatbot con el que interactuamos, funcionan bajo la lógica de una caja negra y cuáles son las implicaciones de esto. En esa línea de ideas, Burrell argumenta que hay tres tipos de opacidades en los algoritmos:

(1) La opacidad como secreto corporativo o gubernamental intencional, (2) la opacidad como analfabetismo técnico, y (3) una opacidad que surge de las características de los algoritmos de aprendizaje automático y la escala requerida para aplicarlos de manera útil (BURRELL, 2016, p. 1, Traducción propia).

Otras reflexiones profundizan sobre cómo la opacidad de los algoritmos, su relativa autonomía, y la delegación de decisiones en éstos puede dificultar el establecimiento de responsabilidades (*accountability*), la transparencia y gobernanza (DIAKOPOULOS, 2016; GILLESPIE, 2013), generando inequidades y replicando sesgos (O'NEIL, 2016). Claro, también ha habido críticas sobre el concepto "caja negra" y sus capacidades para entender a los algoritmos, especialmente porque reproducen una idea de "opacidad", aun cuando hay diferentes alternativas para visibilizarlos y entenderlos. Además, diferente a su uso convencional para referirse a tecnologías estables ya "empaquetadas o cerradas", los algoritmos son fluidos, maleables y cambiantes (BUCHER, 2016; GILLESPIE, 2013; SEEVER, 2013).

En el caso de los chatbots, aunque ha habido una creciente literatura en el área, la mayoría de investigaciones solo analizan los factores que influyen la interacción humano-chatbot y las características de la misma desde abordajes psicológicos y experimentales (ARAUJO, 2018; HILL; FORD; FARRERAS, 2015; HO; HANCOCK; MINER, 2018; PÜTTEN et al., 2010; TAYLOR et al., 2014; XUETAO; BOUCHET; SANSONNET, 2009).

Este enfoque es valioso porque ayuda a entender las percepciones y actitudes de los humanos frente los chatbots. No obstante, tiene la limitación de invisibilizar las circunstancias que influenciaron el desarrollo sociotécnico de los chatbots y su lógica interna. En la literatura consultada, solo se encontró un caso que directamente quería “abrir la caja negra” de un chatbot, Alexa de Amazon (CRAWFORD; VLADAN, 2018), lo que representó una serie de desafíos metodológicos que describiremos a continuación.

## METODOLOGÍA

Para responder el objetivo de analizar el proceso sociotécnico de desarrollo de un chatbot, se siguió un abordaje etnográfico como parte del equipo encargado de la creación e implementación de Katri, un chatbot diseñado para apoyar un curso semipresencial de pedagogía en una institución brasilera de educación superior.

El registro de información se realizó de agosto a noviembre de 2018, teniendo de foco el análisis de las elecciones socio-técnicas en el desarrollo del chatbot. Específicamente, entendemos elecciones socio-técnicas como el proceso mediante el cual un artefacto incorpora características que son producto de una serie de circunstancias sociales, culturales y tecnológicas, pero que también son elecciones que terminan impulsando determinados arreglos socio-técnicos. Además, estas elecciones suelen comportarse como un *trade-off*, al seguir una opción se están rechazando muchas otras.

Los abordajes etnográficos han sido destacados como fructíferos para el análisis sociotécnico de algoritmos (KAVANAGH; MCGARRAGHY; KELLY, 2015; KITCHIN, 2017; SEEVER, 2017). Caracterizándose, porque el investigador desde una posición “externa” (*outsider*) ejecuta un proceso análogo a la ingeniería reversa para entender las elecciones sociotécnicas detrás de los algoritmos (BUCHER, 2016; SEEVER, 2017, 2018), examinando “la lógica que guía las manos, escogiendo ciertos algoritmos en lugar de otros, eligiendo representaciones particulares de datos y traduciendo ideas en código” (SEEVER, 2013, p. 10, Traducción propia)

No obstante, en este estudio el abordaje es diferente, pues se intentó visibilizar y reflexionar “desde adentro” sobre el proceso de desarrollo del chatbot. Este posicionamiento, que algunos han llamado de autoetnografía del desarrollo de algoritmos (KITCHIN, 2017), si bien permite obtener *insights* sobre sus procesos de creación, tiene limitaciones por la dificultad de los investigadores en asumir un posicionamiento crítico.

Teniendo en cuenta estos cuestionamientos, intentamos reducir los posibles sesgos de análisis triangulando fuentes de información; realizando un registro detallado de las actividades ejecutadas por medio de diarios de campo, vídeos, fotografías; la recopilación y análisis de la documentación producida, como textos, archivos Excel y líneas de código; además de realizar ejercicios de contrastación sobre las diferentes interpretaciones que podrían darse a los datos.

## RESULTADOS

En el desafío de retratar un fenómeno complejo decidimos presentar el desarrollo del chatbot Katri dividiéndolo en 6 momentos que nos ayudan a sintetizar el flujo de los acontecimientos en unidades de análisis que agrupan un conjunto de actividades y elecciones socio-técnicas (Cuadro 1).

**Cuadro 1 Proceso de desarrollo del Chatbot**

Momentos	Descripción y principales elecciones sociotécnicas	Actores Involucrados
<b>Primer Momento:</b> Cocinando la creación del chatbot	Acuerdo de voluntades para desarrollar un proyecto para experimentar nuevas tecnologías en la educación.  Decisión de desarrollar un chatbot educativo	Directoría de la Universidad y de la empresa de desarrollo de software
<b>Segundo momento:</b> Creando una razón de ser y una identidad	Definición de objetivos para el chatbot: 1. Resolver dudas logísticas del curso 2. Ayudar con algunas dudas básicas de contenido  Definición de identidad de Katri (Monitora de curso, mujer con méritos académicos, nombre, imagen)	- Equipo de desarrollo: coordinador del curso, profesora, antropólogo, diseñador y programador
<b>Tercer momento:</b> Columna vertebral de Katri	Desarrollo de la columna vertebral: mensaje de recibimiento y árbol de navegación con opciones predefinidas	- Equipo de desarrollo - Un (1) alumno
<b>Cuarto Momento:</b> Alimentando y entrenando a Katri	Delimitación del alcance: definición de temas, intenciones, entidades y respuestas	- Equipo de desarrollo - Un (1) alumno
<b>Quinto momento:</b> La prueba de fuego o los primeros pasos de Katri	Test con alumnos del curso. Confirmación de identidad y alcance de Katri	- Equipo de desarrollo: - 4 alumnos: 3 mujeres, 1 hombre
<b>Sexto momento:</b> Madurez	Curaduría y Perfeccionamiento de Katri (Estabilización)	- Equipo de desarrollo - Todos los alumnos del curso (30)

Fuente: elaboración de los autores

Estos 6 momentos que describen el desarrollo del chatbot como un ser vivo que nace, crece y madura, cumplen un objetivo doble: por un lado, son un recurso narrativo para abrir la caja negra haciendo accesible a un público más amplio el proceso de desarrollo de los chatbots; por otro, sintetizan la lógica “que guía las manos” detrás de la creación de los mismos.

No debe tomarse a la ligera que durante el diseño e implementación del chatbot el equipo responsable hablara sobre el “nacimiento de Katri”, “su columna vertebral”, “sus primeros pasos” o “su manera de hablar”.

Argumentamos que el desarrollo del chatbot Katri se puede entender como un proceso guiado por una metáfora ontológica, que consiste en atribuirle a un fenómeno que no necesariamente tiene una existencia física, características de una entidad, substancia, recipiente o persona (LAKOFF; JOHNSEN, 2003). En otras palabras, entender a Katri como si fuera una persona, antropomorfizarla. A continuación los detalles de su desarrollo.

### **Primer Momento: cocinando la creación del chatbot**

No es fácil determinar porque surgió Katri, ya la presentamos como un chatbot con el objetivo de apoyar la ejecución de un curso de pedagogía, pero en el inicio del proceso esto no era claro. De hecho, el mismo nombre Katri no existía, mucho menos se sabía que solución o beneficios ofrecería, entre otras cosas, porque el problema a resolver tampoco era evidente.

La universidad que contrató el desarrollo del chatbot tenía como misión “Transformar el país por la educación, siendo referencia en prácticas innovadoras de aprendizaje y gestión”, objetivo que se reflejó en el 2018 en: 1) una alianza con una universidad finlandesa para ofrecer un curso semipresencial para profesores sobre prácticas innovadoras en educación; 2) la idea de explorar el uso de nuevas tecnologías con fines académicos, siendo los chatbots una alternativa que habían estado probando en proyectos pilotos.

Así, los encargados de la innovación académica en la universidad en cuestión pensaron que tanto el curso ofrecido por la universidad finlandesa como el objetivo de trabajar con nuevas tecnologías podrían juntarse en un proyecto de mayor envergadura que se llamó “Chatbot Proyecto Finlandia”. Una iniciativa que tuvo como objetivo experimentar el uso de nuevas tecnologías en el proceso de aprendizaje, o sea, una iniciativa que surgió sin saber exactamente cuáles serían sus características.

### **Segundo Momento: creando una razón de ser y una identidad.**

Cinco personas conformaron el equipo base de desarrollo del chatbot. Dos de la universidad, el coordinador del curso en cuestión y una profesora-traductora con experiencia en el contenido del curso; y tres de una empresa externa de desarrollo de software, un diseñador, un programador formado en ciencia de la computación y un antropólogo.

Todos juntos decidieron cual era el objetivo de desarrollar el chatbot. Dos puntos alrededor de esta discusión fueron relevantes. Primero, según el coordinador los alumnos frecuentemente tenían dudas sobre las fechas, horas, las tareas y la preparación necesaria para participar de las actividades virtuales y presenciales del curso. Estas dudas, generalmente llegaban a él o a las profesoras del curso por correo electrónico.

¿Por qué no intentar que el chatbot ayudara a resolver estas dudas? se reflexionó que tal vez hasta se conseguiría que los profesores y el coordinador del curso dedicaran más tiempo en actividades estrictamente pedagógicas y no resolviendo dudas “logísticas” del curso. Por otro parte, que los alumnos se

beneficiarían al recibir una respuesta automática del chatbot versus el uso de correo electrónico, cuya respuesta podría demorar varios días hasta que las profesoras o el coordinador tuvieran tiempo de responder. El objetivo y los usuarios del chatbot fueron entonces definidos: el chatbot sería usado por los alumnos del curso para resolver dudas logísticas.

El otro tema abordado fue sobre la relación del chatbot con el proceso de enseñanza. Una discusión que terminó influenciando directamente la identidad que el chatbot iría a tomar: un monitor de curso, teniendo en cuenta que sería alguien con: 1) información privilegiada sobre las “dudas logísticas” del curso, a diferencia por ejemplo de si fuera otro alumno; 2) capacidad de auxiliar en cosas simples el proceso de enseñanza del profesor, pero no alguien con el objetivo de reemplazarlo; 3) conocimiento básicos sobre los contenidos del curso; 4) además, de ser un rol que podría crear una relación próxima y de confianza con los alumnos.

Después fue debatida la representación visual del chatbot. Como previamente se había estipulado que sería un monitor, la decisión lógica fue que tuviera apariencia humana y no rasgos de animal o robot entre otras posibilidades.

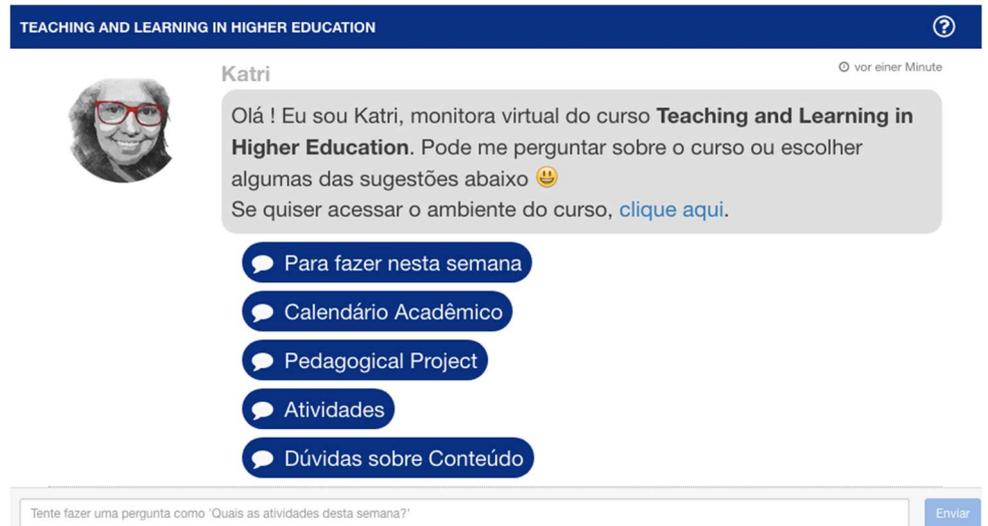
El chatbot se parecería a un humano, pero ¿cómo sería?, así se discutió si sería hombre o mujer, sobre su edad y “biografía”. Teniendo en cuenta que la mayoría de profesores que dictaban las clases presenciales del curso eran mujeres, se decidió que el chatbot también lo fuera. Como se trataba de una monitora, se supuso que tendría importantes méritos académicos y profesionales. Además, tendría alrededor de los 30 años, suponiendo de nuevo, que ya habría terminado una carrera profesional y cursaría un posgrado.

Por último, fue decidido el nombre para el Chatbot. Si los monitores de curso tienen nombre, ¿por qué este chatbot no tendría uno? igual que unos padres con deseo de tener hijos, se buscaron nombres por internet hasta llegar a uno que nos gustó, Katri, un nombre inteligible pero que generaba cierta extrañeza, como si fuera un nombre de un extranjero.

### **Tercer Momento: la columna vertebral de Katri**

Con el objetivo de garantizar una buena interacción chatbot-alumno fue diseñado un flujo patrón de conversación. Este flujo es concretizado en un mensaje de recibimiento que siempre se dispara al entrar al sistema y un árbol de navegación que guía a los usuarios por una serie de opciones predefinidas (Figuras 1 y 2).

Figura 1 Mensaje de recibimiento



Fuente: autores

Figura 2 Árbol de navegación



Fuente: autores.

Ambos, el mensaje de recibimiento y el árbol de navegación, deben ser entendidos como una columna vertebral, un eje que estructura, da consistencia, permitiendo la ejecución de un conjunto de acciones y de interacciones chatbot-humano, pero al mismo tiempo limitando la posibilidad de hacer otros.

Desde el mensaje de recibimiento presentándose como monitora, Katri orienta al usuario sobre las interacciones que puede entablar. Por ejemplo, si el usuario pregunta al chatbot “¿Qué es el proyecto pedagógico?”, teniendo en cuenta que el “Proyecto Pedagógico” hace parte del curso, Katri -como monitora- debería poder responderle. No así con otras preguntas que extralimiten el alcance del curso, por ejemplo “¿Cuál va ser el clima de hoy?”.

Por su parte, el árbol de navegación fue el resultado de una discusión entre el equipo de desarrollo y una alumna del curso. En este, fueron sintetizadas opciones predefinidas de interacción con el fin de anticipar posibles preguntas y temas de interés de los alumnos. Por ejemplo, se asumió que ellos tendrían interés en saber cuáles actividades del curso que hacer en la semana actual; para lo cual, fue predefinida la opción <<Para Hacer esta semana>>, que una vez clicada genera un mensaje con la información respectiva.

¿Qué pasaría si el usuario no escogiera una opción predefinida y escribiera otra pregunta? o si efectivamente tuviera interés en la opción predefinida <<Para hacer esta semana>> pero en vez de clicar en ella escribiera “¿qué tengo que hacer esta semana?”. Para entenderlo tenemos que pasar el cuarto momento.

#### Cuarto Momento: Alimentando y entrenando a Katri

Katri consigue interactuar con los usuarios sobre una serie de 19 temas (Cuadro 2) relacionados con los objetivos y la identidad del chatbot definidos por el equipo de desarrollo y la alumna que participó en la discusión.

Para que Katri fuera competente respondiendo las dudas de los alumnos en cada uno de los temas, es necesario, en el sentido más conductista, entrenarla para que comprenda una serie de entidades e intenciones –estímulos-, que van a generar una respuesta particular por parte de ella. Estos tres conceptos, entidades, intenciones y respuestas son parte fundamental del sistema de inteligencia artificial *IBM Watson Assistant*, usado para el desarrollo del chatbot.

**Cuadro 2 Temas del Chatbot**

1	Actividades Semanales	10	Video Introductorio
2	Calendario Académico	12	Teorías Pedagógicas *
3	Proyecto Pedagógico	13	Motivación e Involucramiento*
4	Actividades	14	Alineación Constructiva*
5	Webinars	15	Metodologías de Enseñanza*
6	Encuentros Presenciales	16	Otros temas*
7	Observación de Pares	17	Juego sobre Teorías de aprendizaje
8	Estudio de Caso	18	Acceso a Recursos
9	Entrevista a estudiantes	19	Sobre Katri

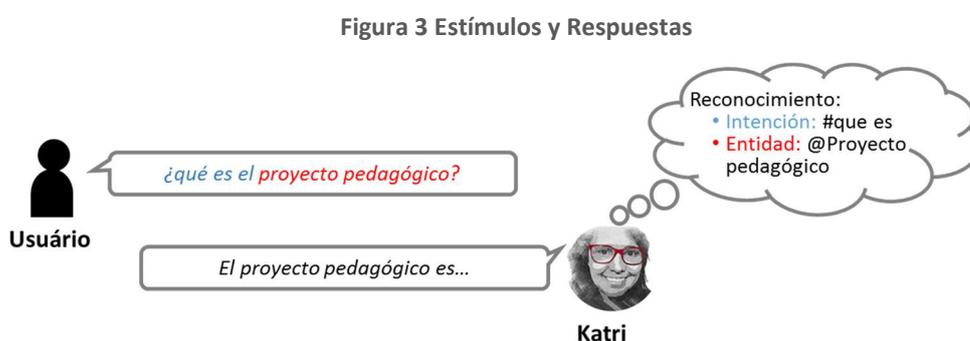
\*Temas 12 a 16 hacen parte de la opción predefinida dudas de contenido

Fuente: autores

Las entidades se refieren a palabras, generalmente sustantivos, que el chatbot tiene capacidad de reconocer, son su “vocabulario básico”. Por ejemplo, Katri reconoce las palabras “calendario académico” y sus sinónimos como “cronograma del curso”, etc.

Las intenciones se tratan de propósitos u objetivos de los usuarios, en este caso, expresados por el usuario por mensajes escritos en el chat. Así, un alumno podría tener la intención de “saber los horarios de las clases” o “conocer las fechas de los encuentro presenciales” por mencionar algunas alternativas que el chatbot aprende a diferenciar.

Un ejemplo nos ayuda a entenderlo. Katri fue entrenada para reconocer la entidad <<Proyecto Pedagógico>> y una serie de intenciones relacionadas como: ¿El proyecto pedagógico es un trabajo obligatorio? ¿Cuál es la fecha límite para entregar el proyecto pedagógico? entre otras opciones. Así, cuando un usuario le pregunta “¿Qué es el proyecto pedagógico?”, ella consigue responder al estímulo y darle la respuesta correcta (Figura 3).



Fuente: autores.

Vale la pena preguntarse cómo surgieron las intenciones y las consecuentes respuestas. Las intenciones se basaron en los correos electrónicos que recibía el coordinador del curso con las dudas de los alumnos, las preguntas que ellos mismos hacían en los foros del *Moodle* del curso, las recomendaciones de la alumna que participó de las discusiones, así como las sugerencias del equipo de desarrollo.

Por otra parte, las respuestas fueron desarrolladas por el coordinador del curso y la profesora ayudante. El foco del resto del equipo fue garantizar que las respuestas tuvieran coherencia con la identidad del chatbot, cuestionándose permanentemente: “Si Katri estuviera viva, ¿cómo respondería esta pregunta o comentario?”

### Quinto Momento: la prueba de fuego o los primeros pasos de Katri

Con cuatro estudiantes del curso, tres mujeres y un hombre, se realizó un *Think Aloud Test*, entendido como una técnica de investigación en la cual se le pide a los participantes que mientras usan un software piensen continuamente en voz alta, es decir, que intenten verbalizar sus pensamientos mientras usan el sistema (NIELSEN, 1993)

A los cuatro estudiantes se les solicitó que interactuaran con Katri, hicieran preguntas y exploraran el árbol de navegación. Igualmente, en el transcurrir de estas actividades se les realizaron preguntas sobre los objetivos y las características del chatbot. Específicamente, se tenía el objetivo de evaluar si valía la pena tener

un chatbot para el curso, si la identidad de Katri como mujer monitora de curso hacía sentido para ellos, y analizar si los temas que cubría cumplían sus expectativas.

En el momento en que los test se realizaron buena parte de las intenciones, entidades y respuestas ya habían sido mapeadas, pero todavía no estaban terminadas, ni Katri había “acabado su entrenamiento” para interiorizarlas. Justamente, se deseaba una retroalimentación rápida, de tal manera que de ser necesario se conseguiría hacer cambios de fondo antes de “cerrar la caja negra”.

Similar a las escenas en que los padres llevan de la mano a sus hijos cuando están aprendiendo a caminar, el *Think Aloud Test* fue pensado como una técnica de investigación que permitiría un ambiente “controlado” de interacción entre el chatbot y los alumnos. No se quería, partiendo de su incompletud e “inmadurez”, que el acceso al chatbot fuera totalmente libre.

Los alumnos concordaron que tener un chatbot para el curso hacía sentido y podría ayudar a resolver dudas básicas. Particularmente, el mayor valor percibido de Katri fue su capacidad para responder dudas sobre las actividades del curso. Una intención cuya respuesta encontraban fácilmente en la opción predefinida <<Para hacer esta semana>> (Cuadro 3, comentarios 1-2).

**Cuadro 3 Comentarios de los alumnos - Think Aloud Test**

1	“Me va ayudar a organizar en términos de tiempo y espacio. En un segundo momento, me ayuda a aclarar algunas nociones básicas, como observación en pares, alineación constructiva, nociones importantes para los proyectos y las metodologías que el programa propone” (alumna 4, Traducción Propia)
2	“(La opción) <<Para hacer esta semana>> es importante, me motivaría a venir aquí todas las semanas a conversar con ella, descubrí que estoy encima del tiempo, atrasado... no estaba recordando esa actividad (señalando la respuesta de Katri sobre la actividad que debería hacer durante esa semana). Me gustó mucho esa funcionalidad, yo la usaría bastante” (alumno 2, Traducción Propia)
3	“Si fuera un coordinador sonaría como una cobranza. Monitor suena mejor, como una persona que está ahí para ayudar” (Alumna , Traducción Propia)
4	“Me gusta la figura, da credibilidad (señalando la imagen de mujer), que sea monitora está bien, porque es más próximo al alumno, si fuera un profesor o un coordinador, ¿Qué puede pasar por la cabeza de nosotros? que también estamos siendo evaluados por la interacción con el chatbot” (Alumna 1, Traducción Propia).
5	“Si colocan (al chatbot) como un profesor o profesora se asume una gran responsabilidad, sobre todo hablando de las dudas de contenido, si usted pregunta al profesor y él no sabe exactamente alguna cosa, depreciaría al profesor” (alumno 2, Traducción Propia).

Fuente: autores

Sobre la identidad, los alumnos estuvieron de acuerdo con que Katri fuera monitora de curso, sin embargo, agregaron nuevos elementos a la discusión, alegando que si el chatbot se presentara como un profesor o coordinador sería incómodo al sentirse evaluados o insatisfechos por su incapacidad hipotética de responder dudas (Cuadro 3, comentarios 3-5). En cuanto Katri como mujer, los alumnos argumentaron que representaba una figura de credibilidad y con un saber

valioso como fue la intención del equipo de desarrollo. No obstante, en las expresiones de los alumnos aparecieron implícitamente algunos elementos como la amabilidad, proximidad y capacidad de ayuda de Katri, que sugieren posibles reproducciones de los estereotipos de la mujer cuidadora y atenta, un fenómeno que según Fioretti, probablemente explique porque la mayoría de los asistentes virtuales tienen “voz de mujer” (FIORETTI, 2018)

### Sexto Momento: madurez

Después de los test con los cuatro alumnos, se sintetizaron los aprendizajes obtenidos, se realizaron algunos ajustes en Katri, que finalmente fue liberada para interactuar con todos los alumnos del curso.

El equipo de desarrollo continuó trabajando en el chatbot, pero ahora con algunas particularidades. Mientras que los momentos anteriores estuvieron marcado por elecciones socio-técnicas, ahora el trabajo era de curaduría perfeccionando las características ya definidas y aumentando algunas entidades, intenciones y respuestas basado en el análisis de la interacción de los alumnos con Katri.

## DISCUSIÓN

Hemos presentado el proceso de desarrollo de Katri por medio de 6 momentos que van desde su creación, pasando por su crecimiento, hasta su madurez. Esta tendencia a antropomorfizar las tecnologías ha sido bien documentado en el campo de la interacción humano-máquina (DUFFY, 2003; EPLEY; WAYTZ; CACIOPPO, 2007) y humano-chatbot (ARAUJO, 2018), lo interesante de nuestro caso, es evidenciar que este fenómeno no solo se da una vez la tecnología está lista, sino también durante el mismo proceso de desarrollo. Pensar el chatbot como una persona ayuda a orientar su diseño e implementación.

Ahora bien, esta “biografía” de Katri fue marcada por una serie de elecciones socio-técnicas que le dieron la consistencia que obtuvo al final del proceso. Al inicio el chatbot estaba lleno de interrogantes ¿por qué crearlo? y ¿Cómo debería ser?, ya en el sexto momento la finalidad, identidad y alcance de Katri eran evidentes. Es claro entonces, que las características y las implicaciones de los 6 momentos no son iguales.

En los primeros dos momentos fueron realizadas la mayoría de las elecciones sociotécnicas, proponiéndose que la solución debería ser un chatbot y que atendería unos objetivos particulares. En el tercer y cuarto momento las implicaciones de estas decisiones fueron profundizadas. En el quinto las decisiones fueron puestas a prueba en los test, dando como resultado, por lo menos en este caso, la reafirmación de las elecciones anteriores. Finalmente, en el sexto momento las elecciones fueron mínimas y el trabajo básicamente pasó a ser de refinamiento de elementos ya definidos.

Retomando las discusiones sobre las cajas negras, no es sino hasta el sexto momento que el arreglo sociotécnico se estabilizó. Aunque el chatbot continúe siendo modificado para agregársele algunas entidades, intenciones y respuestas,

su estructura base se mantiene. Así, consideramos que las posturas que caracterizan a los algoritmos como fluidos, maleables y dinámicos deben ser matizadas (BUCHER, 2016; GILLESPIE, 2013; SEAVER, 2013). En este caso, aunque el algoritmo detrás de Katri no para de transformarse, a partir del sexto momento lo hace de una manera paulatina y superficial.

Adicionalmente, parte importante de las discusiones sobre los algoritmos han girado sobre su opacidad, capacidad de agencia y la delegación de decisiones que los humanos dejamos en éstos (DIAKOPOULOS, 2016; GILLESPIE, 2013; NEYLAND; MÖLLERS, 2016). En esta línea de ideas, los chatbots, por su antropomorfismo, podrían dar la sensación de ser “autónomos” en sus interacciones con las personas, como si tuvieran una propia personalidad y escogieran qué y cómo responder, pero, ¿son realmente autónomos?

Concordamos con Seaver en su argumento de que siempre hay un actor humano cuando hablamos de algoritmos (SEEVER, 2018), aunque no siempre sean fáciles de visualizar. Los procesos de desarrollo de algoritmos se caracterizan por funcionar en ciclos (*loops*) (SEEVER, 2018): primero uno o dos programadores crean un algoritmo que ellos mismos prueban, luego lo muestran a unas cuantas personas cercanas, después se involucra un equipo, luego un pequeño grupo de usuarios experimenta con este, hasta que similar a una bola de nieve el fenómeno va creciendo a tal punto que los usuarios difícilmente conocen quienes fueran las personas y equipos involucrados en los primeros ciclos. Una lógica que se hizo evidente en el propio desarrollo de Katri.

¿Cuál es la relevancia de este argumento? el señalar que en cada uno de estos ciclos, aunque opacos para la mayoría del público, hubo humanos que realizaron una serie de elecciones sociotécnicas que, por acción u omisión, esbozaron la trayectoria y las características que iba a tener la tecnología.

De hecho, del análisis del desarrollo de Katri se desprende la idea que las elecciones sociotécnicas y los actores que participaron de ellas tienen una influencia desigual en el desarrollo de los algoritmos. Si los actores involucrados en el primer momento hubieran decidido crear un aplicativo educacional en vez de un chatbot, el resultado sería totalmente diferente. O si el equipo de desarrollo hubiera decidido que el chatbot no sería un monitor, sino que representaría a otro alumno del curso, la lógica de Katri sería otra. Así, los actores encargados de las primeras decisiones sociotécnicas tuvieron una mayor relevancia en las características que terminó incorporando el chatbot, versus aquellos actores que solo participaron en los últimos ciclos.

## CONSIDERACIONES FINALES

A lo largo del artículo esperamos haber hecho visible, evidente y concreto el proceso de desarrollo de un algoritmo, en este caso de un chatbot. Por medio de un abordaje etnográfico buscamos “abrir una caja negra”, cuestionándose sobre el diseño del chatbot, las características sociotécnicas que incorporó y el papel fundamental que jugaron múltiples actores humanos en su desarrollo.

Como parte cada vez más frecuente de la cotidianidad, es de suma importancia seguir cuestionándonos sobre el funcionamiento y desarrollo de las

nuevas tecnologías. Así como los chatbots son un tipo particular de algoritmo, entre los chatbots también hay una gran variedad de tipologías, de este modo, futuras investigaciones podrían analizar si los procesos de desarrollo de chatbots en otros sectores (salud, servicios, etc.) tienen algunas particularidades. Adicionalmente, sugerimos indagar las maneras como los creadores de chatbots deciden que intenciones, entidades y respuestas usar para alimentarlos, al igual que los factores que contemplan para elegir y moldear sus “identidades”.

# The sociotechnical development of a chatbot or how is a black box constructed?

## ABSTRACT

This work aims to analyze the socio-technical development process of Katri, a chatbot designed to support the execution of a blended course of pedagogy in a Brazilian institution of higher education. For this purpose, an ethnographic methodology was followed that sought to analyze the main sociotechnical activities and choices that influenced the characteristics and development trajectory of the same, or in other words, an ethnographic approach was followed to "open the black box" of a chatbot. It is argued that the design and implementation of the chatbot was guided by a personification process, which was synthesized in 6 key moments that go through its creation, growth and maturity. It is also concluded that in the development process the capacities of the actors to influence the sociotechnical characteristics incorporated in the chatbot are unequal.

**KEYWORDS:** chatbots, black box, socio-technical analysis, algorithms

## REFERENCIAS

ARAUJO, T. Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. **Computers in Human Behavior**, v. 85, p. 183–189, 2018.

BUCHER, T. Neither Black Nor Box: Ways of Knowing Algorithms. In: KUBITSCHKO, S.; KAUN, A. (Eds.). . **Innovative Methods in Media and Communication Research**. London: Palgrave, 2016. p. 81–98.

BURRELL, J. How the Machine “Thinks:” Understanding Opacity in Machine Learning Algorithms. **Big Data & Society**, n. June, p. 1–12, 2016.

CIECHANOWSKI, L.; PRZEGALINSKA, A.; WEGNER, K. The Necessity of New Paradigms in Measuring Human-Chatbot Interaction. In: HOFFMAN, M. (Ed.). . **Advances in Cross-Cultural Decision Making**. Cham: Springer, 2018. p. 205–214.

CRAWFORD, K.; VLADAN, J. **Anatomy of an AI System: The Amazon Echo As An Anatomical Map of Human Labor, Data and Planetary Resources**. Disponível em: <<https://anatomyof.ai>>. Acesso em: 9 maio. 2019.

DALE, R. The return of the chatbots. **Natural Language Engineering**, v. 22, n. 5, p. 811–817, 2016.

DIAKOPOULOS, N. Accountability in algorithmic decision making. **Communications of the ACM**, v. 59, n. 2, p. 56–62, 2016.

DOURISH, P. Algorithms and their others: Algorithmic culture in context. **Big Data & Society**, v. 3, n. 2, p. 205395171666512, 2016.

DUFFY, B. R. Anthropomorphism and the social robot. **Robotics and Autonomous Systems**, v. 42, p. 177–190, 2003.

EPLEY, N.; WAYTZ, A.; CACIOPPO, J. T. On Seeing Human : A Three-Factor Theory of Anthropomorphism. **Psychological Review**, v. 114, n. 4, p. 864–886, 2007.

FIORETTI, L. **Por que a tecnologia tem voz de mulher (e por que isso importa)?** Disponível em: <<http://ofuturodascoisas.com/porque-a-tecnologia-tem-voz-de-mulher-e-porque-isso-importa/>>. Acesso em: 20 ago. 2019.

GILLESPIE, T. The Relevance of Algorithms. In: BOCZKOWSKI, P. J.; FOOT, K. A.; TARLETON, G. (Eds.). . **Media Technologies: Essays on Communication, Materiality, and Society**. Cambridge,: MIT Press, 2013. p. 167–194.

HILL, J.; FORD, W. R.; FARRERAS, I. G. Real conversations with artificial intelligence : A comparison between human – human online conversations and human – chatbot conversations. **Computers in Human Behavior**, v. 49, p. 245–250, 2015.

HO, A.; HANCOCK, J.; MINER, A. S. Psychological , Relational , and Emotional Effects of Self-Disclosure After Conversations With a Chatbot. **Journal of Communication**, v. 00, n. June, p. 1–22, 2018.

IO, H. N.; LEE, C. B. Chatbots and conversational agents: A bibliometric analysis. **IEEE International Conference on Industrial Engineering and Engineering Management**, v. 2017- Decem, p. 215–219, 2018.

JOHNSON, K. **Facebook Messenger passes 300,000 bot**. Disponível em: <<https://venturebeat.com/2018/05/01/facebook-messenger-passes-300000-bots/>>. Acesso em: 9 maio. 2019.

KAVANAGH, D.; MCGARRAGHY, S.; KELLY, S. **Ethnography in and around an Algorithm**. 30th EGOS Colloquium: Sub-theme 15: (SWG) Creativity, Reflexivity and Responsibility in Organizational Ethnography. **Anais...Atenas: (EGOS) European Group for Organizational Studies**, 2015

KITCHIN, R. Thinking critically about and researching algorithms. **Information, communication & society**, v. 20, n. 1, p. 14–29, 2017.

KLEVJER, R. **What is the Avatar? Fiction and Embodiment in Avatar-Based Singleplayer Computer Games**. [s.l.] University of Bergen, 2006.

LAKOFF, G.; JOHNSON, M. **Metaphors we live by**. London: University of Chicago Press., 2003.

LATOUR, B. **La esperanza de Pandora: ensayos sobre la realidad de los estudios de la ciencia**. Barcelona: Gedisa, 2001.

LATOUR, B. **Reassembling the social: an introduction to Actor-Network Theory**. Oxford: Oxford University Press, 2005.

LATOUR, B. **Ciência em ação : como seguir cientistas e engenheiros sociedade afora**. Segunda Ed ed. São Paulo: Editora UNESP, 2011.

MORRISSEY, K.; KIRAKOWSKI, J. “Realness” in chatbots: Establishing quantifiable criteria. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, v. 8007 LNCS, n. PART 4, p. 87–96, 2013.

NEYLAND, D.; MÖLLERS, N. Algorithmic IF ... THEN rules and the conditions and consequences of power. **Information, communication & society**, v. 4462, n. March, 2016.

NIELSEN, J. **Usability Engineering**. First Edit ed. San Francisco, United States: Morgan Kaufmann, 1993.

O’NEIL, C. **Weapons of math destruction: how big data increases inequality and threatens democracy**. New York: Crown/Archetype, 2016.

PÜTTEN, A. M. VON DER et al. “ It doesn ’ t matter what you are !” Explaining social effects of agents and avatars. **Computers in Human Behavior**, v. 26, p. 1641–1650, 2010.

SEEVER, N. **Knowing algorithms**. Media in Transition 8: public media, private media. **Anais...Cambridge, MA: MIT Massachusetts Institute of Technology**, 2013.

Disponível em: <<http://nickseaver.net/papers/seaverMIT8.pdf>>

SEEVER, N. Algorithms as culture: Some tactics for the ethnography of algorithmic systems. **Big Data & Society**, v. 4, n. 2, p. 205395171773810, 2017.

SEEVER, N. What should an anthropology of algorithms do? **Cultural Anthropology**, v. 33, n. 3, p. 375–385, 2018.

TAYLOR, P. et al. Toward an Objective Linguistic-Based Measure of Perceived Embodied Conversational Agent Power and Likeability. **International Journal of Human-Computer Interaction**, v. 30, n. 6, p. 495–516, 2014.

VON HILGERS, P. The History of the Black Box: The Clash of a Thing and its Concept. **Cultural Politics**, v. 7, n. 1, p. 41–58, 2011.

WEIZENBAUM, J. ELIZA--A computer program for the study of natural language communication between man and machine. **Communications of the ACM**, v. 9, n. 1, p. 36–45, 1966.

XUETAO, M.; BOUCHET, F.; SANSONNET, J. **Impact of agent ' s answers variability on its believability and human-likeness and consequent chatbot improvements.** Conference: Proc. of the Symposium Killer Robots vs Friendly Fridges -- The Social Understanding of Artificial Intelligence (AISB 2009). **Anais...2009**

**Recebido:** 08 jun 2019.

**Aprovado:** 19 out 2019.

**DOI:** 10.3895/rts.v16n39.10213

**Como citar:** REINA, D.R.; *et.al.*, El desarrollo sociotécnico de un chatbot o ¿Cómo se construye una caja negra? **R. Technol. Soc.**, Curitiba, v. 16, n. 39, p.23-40, jan/mar. 2020.

Disponível em: <https://periodicos.utfpr.edu.br/rts/article/view/10213>. Acesso em: XXX.

**Correspondência:**

**Direito autoral:** Este artigo está licenciado sob os termos da Licença Creative Commons-Atribuição 4.0 Internacional.

