

# PROTOTIPO DIGITAL PARA FOMENTAR LA CIVILIDAD EN REDES SOCIALES

## *DIGITAL PROTOTYPE TO FOSTER CIVILITY IN SOCIAL NETWORKS*

***Erick López Ornelas***

Universidad Autónoma Metropolitana, Unidad Cuajimalpa, México  
*elopez@cua.uam.mx*

***Rocío Abascal Mena***

Universidad Autónoma Metropolitana, Unidad Cuajimalpa, México  
*mabascal@cua.uam.mx*

**Recepción:** 30/octubre/2020

**Aceptación:** 27/noviembre/2020

### **Resumen**

En este artículo se describe la propuesta de un prototipo basado en una aplicación móvil diseñada para fomentar la civilidad digital dentro de la red social Facebook, llamada Facebook Civility, con la cual se propone moderar las conversaciones derivadas de comentarios generados dentro de la red social. El objetivo del prototipo es que las personas reflexionen sobre los tipos y niveles de violencia que ejercen a partir de sus comentarios en redes sociales. A través de varios ejercicios, los usuarios podrán experimentar las distintas maneras de expresar su opinión sin tener que recurrir a lenguaje ofensivo. En este artículo se describen las diferentes etapas del diseño y la evaluación realizada sobre el prototipo.

**Palabras Clave:** Civilidad digital, polarización de opinión, redes sociales.

### **Abstract**

*This article describes the proposal of a prototype based on a mobile application designed to promote digital civility within the Facebook social network, called Facebook Civility, with which it is proposed to moderate the conversations derived from comments generated within the social network. The objective of the prototype is for people to reflect on the types and levels of violence they exercise based on*

*their comments on social networks. Through various exercises, users will be able to experience different ways of expressing their opinion without resorting to offensive language. This article describes the different stages of the design and the evaluation carried out on the prototype.*

**Keywords:** *Digital civility, mixed opinions, social network.*

## **1. Introducción**

La interacción de los usuarios a través de redes sociales está plagada de violencia y frecuentemente esta violencia se traspasa del plano virtual al físico y viceversa.

En México, a partir de las elecciones presidenciales del año 2018, en redes sociales se evidenció el problema de la polarización social y política, por la diversidad de opiniones en relación con la permanencia del Partido Revolucionario Institucional (PRI) y el triunfo del populismo, representado por el candidato del partido Movimiento de Regeneración Nacional, Andrés Manuel López Obrador [Domínguez, 2018].

Si bien el objetivo de las redes sociales está ligado a la libertad de expresión también existe la corresponsabilidad que la red social Facebook estaría en posibilidades de asumir, para fomentar espacios libres de violencia en sus distintas manifestaciones. La propuesta que se plantea como solución en el presente artículo, Facebook Civility, se enfoca en dar solución al problema de violencia causado por polarización política, sin suprimir los derechos del usuario [Belaunzarán, 2019], [Tecnósfera, 2019]. El proyecto plantea que la funcionalidad de la interfaz utilice el procesamiento del lenguaje natural (Natural Language Processing [NLP]). Este método permite comprender y extraer información significativa de los mensajes y comentarios que los usuarios envían, denominados entidades; posteriormente, se utiliza la intención de estas entidades para automatizar algunas de sus respuestas, dirigir la conversación a una charla humana y recopilar datos de la audiencia [Chowdhary, 2020]. Por diversos hechos de violencia generados entre usuarios de Facebook alrededor del mundo, Mark Zuckerberg, CEO de esta empresa, tuvo que tomar acciones para mediar y buscar combatir la violencia en esta red social. Para ello, en

2018 anunció que se emplearía a 20 mil personas para vigilar los discursos de odio y mensajes reportados como violentos. Sin embargo, el mismo Zuckerberg ha aceptado que esto no es suficiente y se encuentran ante un reto difícil ya que Facebook está en más de 200 países, con su propia cultura y lenguas [Wu, 2019]. Otra red social que busca moderar los mensajes agresivos es Twitter [Kursuncu, 2019]. El 27 de junio de 2019 la compañía anunció que restará visibilidad a los mensajes que no cumplan con las normas de la red social, específicamente de las figuras públicas, como el presidente de los Estados Unidos, Donald Trump. A los usuarios les aparecerá una alerta que les avisará que están violando las reglas y le restará visibilidad para que lleguen a un menor número de usuarios en Twitter. La empresa justificó estos cambios en su algoritmo para equilibrar entre la responsabilidad de cada usuario y la libertad de expresión.

Más allá de Facebook, existen proyectos que buscan disminuir la violencia, como es el caso de ReThink [Re-Think, 2019]. Esta aplicación móvil invita a los usuarios, en su mayoría adolescentes, a pensar antes de publicar un mensaje con lenguaje ofensivo. La iniciativa de ReThink es que los jóvenes tengan una segunda oportunidad antes de emitir alguna ofensa y con esto combatir los casos más graves de cyberbullying.

Otro ejemplo de aplicación móvil es la desarrollada por la Fundación en movimiento llamada Bullysemáforo, con la cual se pretende prevenir, detectar e incluso resolver casos de acoso escolar [Evans, 2019].

A partir de esta indagación se encontró que no existen interfaces de habla hispana destinadas a combatir el problema de violencia en redes sociales, aunque esté identificado y categorizado como una alerta social en el país.

## **2. Métodos**

El método aplicado se llevó a cabo a través de un proceso conformado por el planteamiento del problema o *Initial design Testing*, la definición y creación de usuarios con la técnica *Personas*, desarrollo de casos de uso con el uso de *Storyboard*, la conceptualización del video y finalmente la realización del prototipo.

## **Initial design testing**

Como primera etapa se definió problema relacionado con la civilidad digital que se pudiera resolver a través de medios digitales e Internet.

Una vez que se tuvo la delimitación del tema a la polarización de opiniones en Facebook, se procedió a diseñar una entrevista para poder obtener información sobre cómo la gente vive la violencia en redes sociales y cuáles son sus puntos de vista.

Se realizaron un conjunto de entrevistas para obtener información cualitativa de la situación. Con los resultados se realizó un análisis para identificar las principales necesidades de los entrevistados y los perfiles generales de estos. Se generó una lista de necesidades, de la cual se eligió el crear espacios de tolerancia y respeto en redes sociales. Esta etapa fue fundamental para la detección de la problemática a abordar: la civilidad en las redes sociales.

## **Personas**

Se utilizó la técnica Personas de Allan Cooper [Haag, 2019] que consiste en crear personajes ficticios para representar al usuario de la manera más realista posible, se define su edad, algunas características, sus metas, estudios, ocupación, destrezas tecnológicas, habilidades y frustraciones.

A partir de los perfiles identificados previamente, se resumieron en dos que estuvieran implicados en el problema planteado. Este método es muy importante ya que sirve para generar empatía con los usuarios y poder determinar de manera adecuada el perfil de los usuarios con los que se trabajará.

La Persona 1 se refiere al usuario que se autocensura, no comentando su opinión o no comentando en lo absoluto, para evitar confrontaciones ya sea por temor a discutir con sus amigos o a recibir insultos (figura 1).

El segundo perfil identificado se refiere a la Persona 2, que son los que emiten su opinión, e incluso responden a otros comentarios sin importar las repercusiones que esto podría causar, tiene una postura y la defiende, los que estén en contra están mal para esta persona (figura 2).

Persona que se autocensura		Metas
	<p><b>Ramón</b></p> <p>Tiene 25 años Vive con sus papás Es de clase media baja Fue la primera vez que pudo emitir su voto, y este fue a favor de MORENA</p>	Graduarse
Estudios y ocupación	Destrezas tecnológicas	Frustraciones
<p>Estudia Sociología en la facultad de Ciencias Políticas y Sociales</p> <p>Es becado, los recursos de su familia no eran suficientes para estudiar</p>	<p>Usa redes sociales como Twitter, Facebook, Instagram, Snapchat y whatsapp</p> <p>A pesar de estar informado, a veces no revisa las fuentes de información</p>	<p>Prefiere no opinar en redes sociales puesto que en su licenciatura las opiniones son muy polarizadas y no quiere perder amigos o que lo tachen por sus opiniones</p> <p>La corrupción que se vive en el país</p>

Figura 1 Primer perfil de usuario definido por el método de *Personas*.

Persona que confronta		Metas
	<p><b>Laura</b></p> <p>Tiene 35 años Tiene 1 hijo y vive en unión libre Vive en el centro del país Se considera anti Amlo y siempre ha votado por el PRI Es clase media</p>	<p>Hacer crecer su negocio</p> <p>Que su hijo tenga una buena educación</p>
Estudios y ocupación	Destrezas tecnológicas	Frustraciones
<p>Estudió la licenciatura en administración de empresas</p> <p>Tiene su propio negocio de venta de ropa para dama</p>	<p>Usa las aplicaciones más comunes y de moda como Uber.</p> <p>Prefiere Facebook e Instagram.</p>	<p>La situación de violencia en el país</p> <p>No tener suficiente tiempo para ella</p> <p>No le gustan las noticias, pero trata de mantenerse informada.</p>

Figura 2 Segundo perfil de usuario definido por el método de *Personas*.

## Storyboard

El Storyboard es una técnica de diseño que se utilizó para realizar una primera propuesta de solución de la aplicación. El storyboard es muy útil para identificar la satisfacción final del usuario, además ayuda en la conceptualización de la aplicación. Se generaron 3 propuestas de casos de uso visualizadas a través de un storyboard que consistían en:

- Añadir un filtro de privacidad similar al que existe en las publicaciones, pero dirigido a los comentarios, con el fin de que éste solo pueda ser visualizado por ciertos usuarios definidos por quien está comentando.

- Detectar y bloquear las cuentas de los usuarios que emitan mensajes violentos.
- Una interfaz que suavice el lenguaje ofensivo y/o cambie el discurso de odio por uno amigable y de esta manera crear disrupción en el usuario.

Estas propuestas se evaluaron y analizaron por su conveniencia, eficiencia y originalidad, bajo estos criterios se eligió la propuesta número tres. El caso de uso consistía en un usuario que, al observar un comentario con el que no está de acuerdo, responde de manera ofensiva por lo que la interfaz cambia su discurso con palabras ofensivas por uno amable, por lo que la conversación que se pudo desarrollar de manera agresiva y ofensiva se convierte en una conversación civilizada donde ambas partes quedan satisfechas (figura 3).



Figura 3 Storyboard utilizado para implementar el escenario con mayor civilidad.

### Prototipo en papel

Para el primer boceto de la interfaz, se creó un prototipo de papel (figura 4) para conocer diferentes formas de conceptualizar la interfaz. Se seleccionaron elementos de cada idea para conformar una sola interfaz. A partir de esta propuesta conjunta se simuló el flujo de pantallas y los posibles caminos que integrarían el caso de uso.

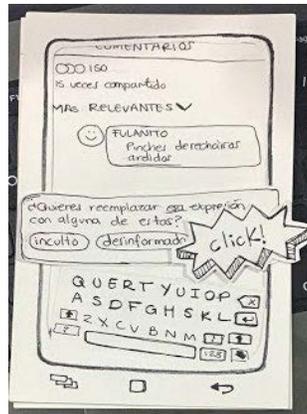


Figura 4 Prototipo en papel utilizado para la conceptualización de las ideas.

Se definieron posteriormente, dos tipos de insultos que deberían de ser detectados por la aplicación. Estos insultos tienen que ver con el uso de palabras altisonantes e insultos compuestos. La idea es que estos insultos puedan ser intercambiados por sinónimos menos agresivos y que inciten a una mejor convivencia e intercambio de mensajes de forma civilizada entre los usuarios (tabla 1).

Tabla 1 Palabras altisonantes e insultos compuestos.

Palabras altisonantes	Insulto compuesto
Pendejo	Puta madre
Puto	Pinche madre
Derechairo	Hijo de puta
Pejzombie	Pinche pendejo
Chingar	Su puta madre
Ignorante	Pinche mamón
Alv	Chingue a su madre
Imbécil	Pinche resentido
Idiota	A la chingada

### Prototipo digital

Para crear el prototipo interactivo se utilizaron dos herramientas. Para el diseño gráfico de la interfaz se utilizó el programa de ilustración vectorial Illustrator y para diagramar la funcionalidad de la interfaz la plataforma de diseño de productos digitales Marvel (<http://www.marvelapp.com>), figuras 5 y 6).

Se siguieron las métricas de estilo establecidas en Facebook, para que esta se mimetizara con la identidad gráfica de la red social. La configuración y los textos de introducción para la aplicación móvil está inspirada en la aplicación Facebook

Parejas, puesto que la intención es que Facebook Civility funcione de la misma manera forma, como un apéndice o add-on de Facebook.

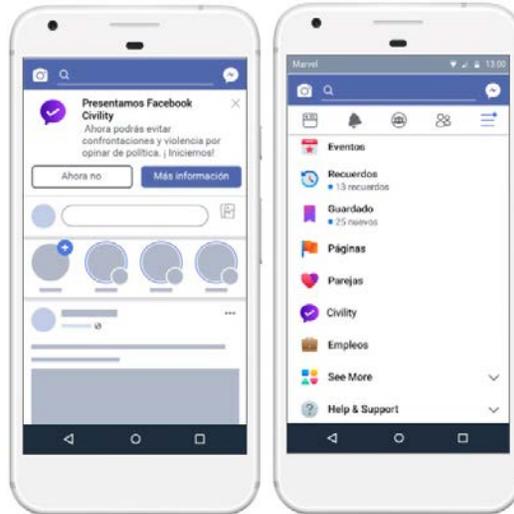


Figura 5 Prototipo digital utilizado para la conceptualización de la aplicación.

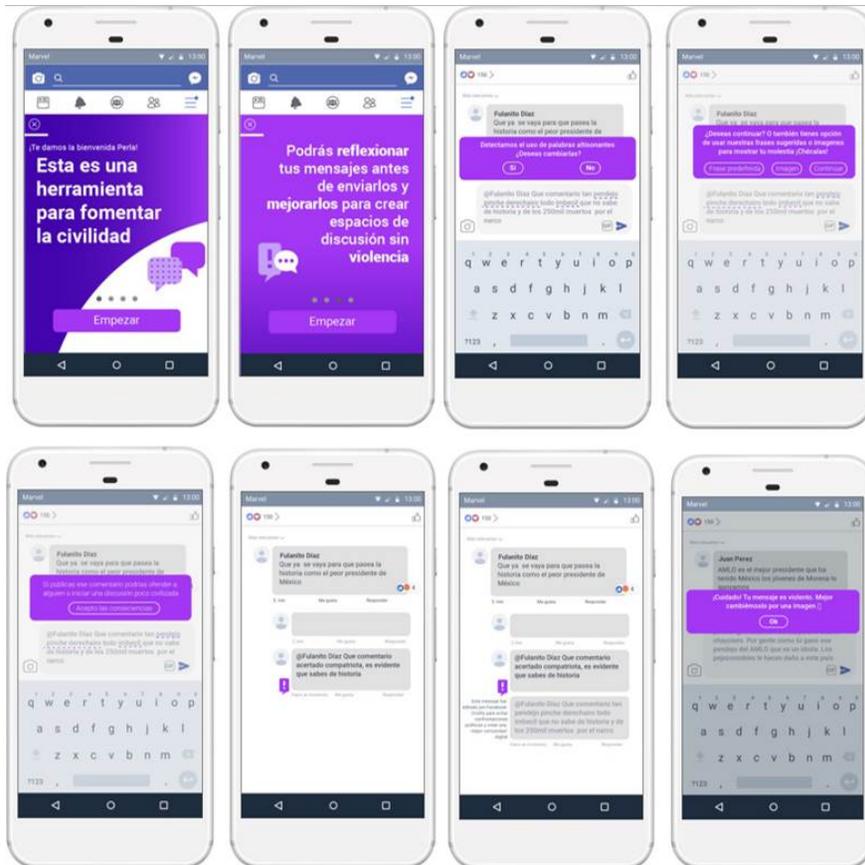


Figura 6 Prototipo digital donde se ejemplifican las funcionalidades implementadas.

El funcionamiento del diagrama inicia con la instalación de la aplicación, para esta acción existen dos caminos, el primero se realiza a través de la configuración general de Facebook y el otro a través de una sugerencia para activar la aplicación (figura 5). El siguiente paso es, que al encontrar una noticia y dirigirse a la sección de comentarios, selecciona uno al cual responder por qué no coincide con sus ideales y quiere demostrar su enojo o disgusto, por lo que la interfaz ofrece una primera opción cambiar las palabras identificadas como agresivas.

Si el usuario acepta cambiar las palabras, se cambian las altisonantes por sinónimos más amables y con connotación menos ofensiva, mientras que las palabras compuestas se eliminan, dejando un mensaje con el mismo sentido del discurso, pero menos violento. Si el usuario no acepta se despliegan tres opciones más, donde intenta hacer recapacitar al usuario ofreciendo otras opciones: usar una imagen o GIF de la base de imágenes creada, con las cuales demostrar principalmente enojo o desacuerdo, sin tener connotaciones políticas u ofensivas ya que se eligieron cuidadosamente memes populares en la cultura mexicana que fueran neutros, no se prestaran a dobles sentidos o sarcasmo y su significado fuera obvio; usar una frase predefinida por el sistema para expresar el disgusto del usuario de manera educada y como última opción, continuar para publicar su comentario. Si el usuario rechaza todas las oportunidades de cambiar su comentario y decide publicarlo con las palabras ofensivas, se cambian las palabras detectadas como insultos con antónimos, cambiando de esta manera el sentido del discurso del usuario, aunque también ofrece la opción de ver el comentario original para evitar caer en la censura. Si cualquier comentario tiene errores gramaticales o más de 10 palabras identificadas como insulto, la interfaz automáticamente omite las opciones de cambio de palabra y ofrece al usuario la opción de publicar una imagen o GIF que exprese su opinión.

### **3. Resultados**

Cuando se tuvo un primer prototipo digital se realizó la primera evaluación, la cual fue una Evaluación Heurística (EH). Una EH es un método de inspección de la usabilidad sin usuarios (Nielsen, 1994). Esta, consiste en examinar la calidad de

uso de una interfaz por parte de varios evaluadores expertos, a partir del cumplimiento de unos principios reconocidos de usabilidad: los heurísticos. Esta evaluación nos ayudó a verificar los principios de usabilidad y qué heurísticas no se estaban cumpliendo, por lo que algunos de los cambios que se realizaron fueron:

- Control y libertad: Se tenían solo dos opciones de mensajes, por lo que se replanteó el funcionamiento y se añadieron más opciones de cambio de palabra como frases predefinidas e imágenes.
- Prevención de errores: Se añadió una introducción de la aplicación, para explicar al usuario que es Facebook Civility y cómo funciona, así mismo se cambió su funcionamiento de aplicación externa a un complemento add-on dentro de Facebook.

Como segunda etapa de evaluación se realizó una prueba de usabilidad con un usuario real, la cual consistió en mostrarle el prototipo y pedirle que, siguiendo las métricas del protocolo Think-Aloud (McDonald, 2013), describiera su interacción con la interfaz, al igual que dudas o ideas que surgieran durante el proceso. Al mismo tiempo, se registró presencialmente la interacción, mientras que un grupo de expertos, lo observaba de manera remota a través de dos cámaras colocadas en la sala [Rubin, 2008]. Una de las cámaras fue posicionada directamente sobre el Smartphone para conocer el desempeño de la interfaz; y otra permaneció dirigida hacia el usuario, para observar sus gestos y actitud hacia la interfaz.

Los hallazgos detectados en esta fase de evaluación fueron:

- La falta de algún botón para regresar a la pantalla inicial. Como el prototipo muestra varias posibles opciones y rutas, se volvía tedioso regresar al inicio sólo con el botón de ir hacia atrás.
- Evidenciar las rutas activas del prototipo. Por el mimetismo del prototipo con Facebook, el usuario quería interactuar con él como lo hace con la aplicación real y recurría a dar clic en todos los elementos hasta encontrar el que si tenía funcionamiento programado.
- La barra de botones colocada para simular el teléfono móvil confundía al usuario, ya que, por inercia, intentaba interactuar con esta.

En general, se obtuvieron resultados positivos y las funcionalidades implementadas resultaron de agrado y benéficas para los usuarios, los cuales comentaron que si les será de gran ayuda, sobre todo cuando existe un intercambio de mensajes entre usuarios.

#### **4. Discusión**

El hablar de civilidad en las redes sociales, es un elemento complejo debido a que las redes sociales brindan un cierto anonimato y una forma de ser desinhibido en la interacción que se realiza. Es por eso que al interactuar se debe buscar siempre tener cierta empatía y amabilidad en cada interacción, y tratar a todos con los que se conecta en línea con dignidad y respeto.

De igual forma hay que tomar en cuenta que todos somos diferentes por lo que se deben de considerar las diversas perspectivas y cuando surjan desacuerdos, participar cuidadosamente y evitar insultos y ataques personales.

Es importante hacer una pausa y reflexionar antes de responder a cosas con las que no está de acuerdo, y no publicar ni enviar nada que pueda dañar a otra persona, afectar una reputación o amenazar la seguridad de alguien. De igual manera apoyar a los demás y denunciar actividades amenazantes y preservando evidencia de comportamiento inapropiado o inseguro. Estos son algunos de los elementos que, en este artículo, se intenta reflexionar. Al ser un prototipo no todas las funcionalidades han sido implementadas, pero si deja ver el alcance y la forma en que éste ayudará al usuario y fomentará la civilidad en las redes sociales.

#### **5. Conclusiones**

En esta sección resulta importante resaltar el proceso de diseño que se siguió para la generación de la propuesta, desde la búsqueda de necesidades y poder identificar la problemática, hasta la aplicación de un conjunto de técnicas como el storyboard y la utilización de prototipos para la generación de la propuesta. Se aplicaron de igual forma dos tipos de evaluaciones, la verificación de heurísticas y la pruebas con usuarios reales, se recibió una crítica positiva respecto a la interfaz y la premisa en la que está basada.

El objetivo de crear disrupción en el usuario se cumple y genera en él una reflexión para no ser tan agresivo, al mismo tiempo que brinda opciones de cómo mejorar sus comentarios para evitar ofender y de esta manera ser ofendido en redes sociales.

Facebook civility, es una propuesta que busca fomentar la civilidad en las interacciones realizadas en Facebook y donde los usuarios tengan la capacidad de discernir entre un buen comentario amable y un comentario agresivo. Esto por supuesto fomenta una mejor interacción entre los usuarios, el intercambio de ideas y aumenta la civilidad que debería siempre existir en el mundo virtual, donde las redes sociales nos han invadido en la actualidad.

Es importante mencionar que, en el presente artículo, se busca encontrar una solución viable al problema de la civilidad en las redes sociales. En el artículo no detalla la manera de identificación automática de mensajes ni las técnicas de tratamiento de lenguaje natural utilizadas.

Como trabajo a futuro, es interesante el poder realizar la implementación real de la aplicación, donde forzosamente se deberá de contar con una base de datos además de la implementación de algoritmos de procesamiento de lenguaje natural para poder realizar la verificación y el intercambio de conceptos que se viertan por parte de los usuarios de la aplicación.

## **6. Bibliografía y Referencias**

- [1] Belaunzarán F., (2019). México polarizado. Excelsior: <https://www.excelsior.com.mx/opinion/opinion-del-experto-nacional/mexico-polarizado/1303797>.
- [2] Chowdhary K. R., (2020). Natural Language Processing. In: *Fundamentals of Artificial Intelligence*. Springer, New Delhi.
- [3] Domínguez González G., (2018). Elección, polarización y hartazgo social en México. Nueva sociedad: <https://nuso.org/articulo/eleccion-polarizacion-y-hartazgo-social-en-mexico/>.
- [4] Evans C. B. R., Smokowski, P. R., Rose R. A. et al., (2019). Cumulative Bullying Experiences, Adolescent Behavioral and Mental Health, and Academic Achievement: An Integrative Model of Perpetration, Victimization,

- and Bystander Behavior *J. Child Fam Stud* 28, 2415–2428. <https://doi.org/10.1007/s10826-018-1078-4>.
- [5] Haag M., Marsden N., (2019). Exploring personas as a method to foster empathy in student IT design teams. *Int J Technol Des Educ* 29, pp. 565–582. <https://doi.org/10.1007/s10798-018-9452-5>.
- [6] Kursuncu U., Gaur M., Lokala U., Thirunarayan K., Sheth A., Arpinar I. B., (2019). Predictive Analysis on Twitter: Techniques and Applications. In: Agarwal N., Dokoochaki N., Tokdemir S., (eds) *Emerging Research Challenges and Opportunities in Computational Social Network Analysis and Mining*. Lecture Notes in Social Networks. Springer, Cham.
- [7] McDonald S., McGarry K., Willis L.M., (2013). Thinking-aloud about web navigation: the relationship between think-aloud instructions, task difficulty and performance. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 57, pp. 2037–2041. SAGE Publications, Los Angeles.
- [8] Nielsen J., (1994). Heuristic evaluation. In Nielsen J., and Mack R.L. (Eds.), *Usability Inspection Methods*, John Wiley & Sons, New York, NY.
- [9] Re-Think Illinois, (2019). Estados Unidos: <http://www.rethinkwords.com>
- [10] Rubin J., (2008). *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests*. Wiley Publishing Inc.
- [11] Tecnósfera, (2019). Cambio en Twitter podría reducir visibilidad de Trump. El Tiempo: <https://www.eltiempo.com/tecnosfera/novedades-tecnologia/el-cambio-de-politicas-de-twitter-que-puede-quitar-visibilidad-a-trump-381554>.
- [12] Wu, C. et al., (2019). Machine Learning at Facebook: Understanding Inference at the Edge, 2019 IEEE International Symposium on High Performance Computer Architecture (HPCA), Washington, DC, USA, pp. 331-344, doi: 10.1109/HPCA.2019.00048.