

# Design decisions of collaborative tools that reduce the impact of Fake News

Edison Caiza<sup>1</sup>, Alejandro Fernández<sup>1</sup> and Diego Torres<sup>1</sup>

<sup>1</sup> Universidad Nacional de La Plata, Calle 50 &, Av. 120, La Plata, Buenos Aires, Argentina

**Abstract.** The phenomenon of dissemination and consumption of fake news brings with it the permanent risk of harm to people, companies, or entities referenced in the content of the messages or publications. An important aspect to consider in the proliferation of disinformation is the design of collaborative tools in which fake news are disseminated; which can be improved through knowledge of the factors that facilitate or prevent its propagation. Thus, this research work proposes the construction of a reference model of the impact of fake news, which allows obtaining metrics of the damage they can cause and the factors that increase or reduce it, in order to provide a catalog of design recommendations/decisions that reduce said impact.

**Keywords.** Fake News, Collaborative tools, Impact, Reference Model.

## Introducción

Los espacios colaborativos dada su naturaleza permiten la participación de varios individuos que contribuyen a un objetivo: los cuales mantienen interacción entre sí. Bajo este escenario, si los aplicativos que permiten dicha interacción y colaboración no poseen los controles adecuados para moderar la información que se comparte, permanece latente el riesgo de difundir información errónea o falsa, o cualquier contenido asociado al término “fake news”; y, con ello, la materialización de daños a personas, empresas o entidades referenciadas en ella. Ante esta situación, resulta importante comprender la relación existente entre las características de las fake news y los espacios colaborativos en los que estas se difunden, a fin de levantar/elaborar un catálogo de decisiones de diseño que facilite la implementación de aplicativos que reduzcan el impacto de las fake news.

Las medidas adoptadas por los gigantes de las redes sociales y los sistemas de mensajería, como Facebook, Whatsapp y Twitter, han desacelerado la difusión de desinformación; sin embargo, no son suficientes, debido a que cada vez son más los interesados en difundir contenidos erróneos. Como parte de estas medidas, Facebook, por ejemplo, promueve la información creíble y elimina la información errónea de sus plataformas, además del cierre de cuentas falsas y de medios fraudulentos y la restricción de publicidad para entidades maliciosas. WhatsApp, por su parte, ha limitado la facilidad para reenvío de mensajes en su plataforma. [1] Y, Twitter

permite etiquetar un tweet como sospechoso; además de encargarse de la limpieza de bots y del control de tendencia de temas políticos.

Con base expuesto, se entiende que no existen medidas de solución definitivas y que se demanda de la implementación de medidas de solución alternativas, desde otras perspectivas: tales como el diseño de los espacios colaborativos. Es de aquí que nace la motivación de este trabajo de investigación. El primer hito a concretar, como parte de la propuesta, es la definición de un modelo referencial de los daños (impacto) que provocan la difusión y el consumo de desinformación. En lo posterior, y a partir del modelo referencial establecido, se extraerán las métricas visibles que faciliten el cálculo o la representación de la criticidad de la materialización del daño de las fake news. Y, finalmente, a partir de las representaciones de criticidad obtenidas, se trabajará en el análisis de los factores que incrementan o reducen el grado de criticidad y se construirá un catálogo de decisiones de diseño (controles) para la implementación de espacios colaborativos.

## **Estado del Arte**

Desafortunadamente, el encontrarse expuesto a información inexacta puede conducir a ciertos problemas. Los lectores, después de consumir declaraciones inexactas, muestran claros efectos de esos contenidos en sus decisiones. La exposición a este tipo de información genera confusión acerca de lo que es verdadero, dudas sobre entendimientos precisos y la subsiguiente dependencia de falsedades. [2] Son incontables los eventos que se pueden citar como resultado de la materialización del daño provocado por la desinformación. La muerte, la escalada del conflicto, la hostilidad política y el pánico social son algunos de los efectos adversos de la proliferación de las fake news. [3] En trabajos como [4], por ejemplo, se identifican datos alarmantes del consumo de información falsa respecto del “mito popular de que el alcohol altamente concentrado podría desinfectar al cuerpo y matar al coronavirus” (cerca de 800 personas muertas, 5876 hospitalizadas y 60 ciegos luego de beber metanol). Otro ejemplo del daño a la integridad física de las personas es el asesinato o linchamiento de personas inocentes provocado por la falta de educación digital de quienes consumen información inexacta. En el ámbito social, por citar un último ejemplo, las fake news afectan de sobre manera a las decisiones democráticas de los países, influenciando de forma negativa en la opinión de los electores. [5]

Pese a que existen demasiados ejemplos de materialización de los daños provocados por las fake news, una búsqueda preliminar en revistas y artículos indexados, además de sitios especializados en desinformación, dan cuenta de la falta de métricas estandarizadas para evaluación del impacto de las fake news; lo que representa un problema y un reto para la implementación de controles adecuados, en espacios colaborativos, que reduzcan los riesgos asociados al consumo de desinformación. [6][7]

## Planteamiento del Problema y Contribuciones

La proliferación de fake news en los espacios colaborativos es motivo de estudio de un sin fin de trabajos de investigación, en distintas disciplinas de la ciencia; debido principalmente a los efectos negativos (daños) que estas producen en las personas, empresas y entidades. La literatura disponible se enfoca principalmente en el estudio de la problemática social y en la búsqueda u optimización de los métodos de detección; sin embargo, es importante abordar también la relación existente entre las características e impacto de las fake news y las decisiones de diseño de los espacios colaborativos en los que estas se difunden, a fin de establecer los controles necesarios para reducir el impacto de su consumo.

Bajo este contexto, se han planteado las siguientes preguntas de investigación:

**P1:** ¿Se puede determinar con precisión el impacto de las fake news?

**P2:** ¿Se puede determinar con precisión la relación existente entre la caracterización de las fake news y las decisiones de diseño de espacios colaborativos?

**P3:** ¿Se puede contribuir en la toma de decisiones para construcción de espacios colaborativos saludables, en los que se reduzca o impida la propagación de fake news?

Para responder a estas preguntas, el trabajo de investigación planteado propone levantar un conjunto de recomendaciones/decisiones de diseño para implementación de espacios colaborativos “seguros”: a través de la búsqueda de información relevante en literatura y la construcción y validación de un modelo de referencia del impacto de las fake news.

En vista de que no se han identificado artículos en los que se aborde la lucha contra las fake news desde la perspectiva del diseño de los espacios colaborativos, la contribución de este trabajo de investigación se centrará en la discusión de la estandarización de métricas del impacto de las fake news y de la forma en que estas se encuentran relacionadas con el diseño de las aplicaciones de interacción social. [5]

## Metodología y enfoque de investigación

La ejecución de este trabajo de investigación se dividirá en tres fases:

1. ***Construcción de un Modelo de Referencia del impacto de las fake news***, que permita la extracción de métricas que faciliten la explicación de la relación entre el impacto de las fake news y las decisiones de diseño de los

espacios colaborativos en los que estas se difunde. Para implementación del modelo se empleará la Metodología de Investigación de Diseño (DRM) [8]; que, una vez evaluada, se muestra como la más apta para este tipo de representación.

2. ***Compilación de un catálogo de decisiones de diseño***, que se obtendrá a partir del análisis de la relación existente entre las métricas de impacto de las fake news y el diseño de las aplicaciones colaborativas.
3. ***Evaluación del catálogo de decisiones de diseño***, para validación de consistencia de las recomendaciones levantadas. Esta fase se describe a detalle en el punto “Plan de evaluación” de este documento.

## **Plan de evaluación**

El foco de evaluación de este trabajo de investigación será el catálogo de decisiones de diseño, para lo cual se combinarán las siguientes estrategias:

1. Las recomendaciones/decisiones de diseño se encontrarán soportadas principalmente en la literatura disponible y a través de la recolección de datos empíricos.
2. Para las recomendaciones/decisiones de diseño que sean factibles, se validará su asertividad a través de experimentación o por medio del juicio de expertos.
3. Las recomendaciones/decisiones de diseño que resulten demasiado complejas de experimentación y no se cuente con el juicio del experto, se validarán a través de la percepción de los usuarios (mediante opinión o encuestas).

## **Resultados preliminares o intermedios**

Como parte del trabajo realizado, se pueden citar los siguientes resultados preliminares:

1. Se ha determinado el Estado del Arte de las Fake News y de los medios en los que se difunden; a través de lo cual se ha conseguido la publicación de un artículo, a modo de revisión sistemática de literatura.
2. Se ha conseguido una primera aproximación al modelo de referencia de impacto de las fake news, a través de la síntesis de literatura y la representación gráfica con DRM.

## Conclusiones y Lecciones Aprendidas

El desarrollo y resultados de este trabajo de investigación contribuirán a la lucha en contra de la difusión de contenidos erróneos y de desinformación, desde la perspectiva de evaluación del daño que pueden ocasionar y el enfoque en las recomendaciones para diseño de espacios colaborativos un tanto más seguros para compartir información. El modelo de impacto de fake news que se obtenga como resultado de este trabajo, puede servir como base para la discusión en trabajos futuros en los que se analicen métricas y decisiones de diseño de espacios colaborativos.

## Etapa del Doctorado

Doctoral ES

## Referencias

1. Nsoesie, O. Elaine, Oladeji, Olubusola (2020). Identifying Patterns to Prevent the Spread of Misinformation during Epidemics, The Harvard Kennedy School (HKS) Misinformation Review, Volume 1, Special Issue on COVID-19 and Misinformation.
2. Rapp, David & Salovich, Nikita. (2018). Can't We Just Disregard Fake News? The Consequences of Exposure to Inaccurate Information. *Policy Insights from the Behavioral and Brain Sciences*. 5. 232-239. 10.1177/2372732218785193.
3. Apuke, Oberiri & Omar, Bahiyah. (2020). FAKE NEWS PROLIFERATION IN NIGERIA: CONSEQUENCES, MOTIVATIONS, AND PREVENTION THROUGH AWARENESS STRATEGIES. *Humanities & Social Sciences Reviews*. 8. 318-327. 10.18510/hssr.2020.8236.
4. Islam, M. S., Sarkar, T., Khan, S. H., Mostofa Kamal, A., Hasan, S. M. M., Kabir, A., Yeasmin, D., Islam, M. A., Amin Chowdhury, K. I., Anwar, K. S., Chughtai, A. A., & Seale, H. (2020). COVID-19-Related Infodemic and Its Impact on Public Health: A Global Social Media Analysis, *The American Journal of Tropical Medicine and Hygiene*, 103(4), 1621-1629.
5. Machado, Caio & Kira, Beatriz & Narayanan, Vidya & Kollanyi, Bence & Howard, Philip. (2019). A Study of Misinformation in WhatsApp groups with a focus on the Brazilian Presidential Elections. *WWW '19: Companion Proceedings of the 2019 World Wide Web Conference*. 1013-1019. 10.1145/3308560.3316738.
6. Xinyi Zhou and Reza Zafarani. 2018. Fake News: A Survey of Research, Detection Methods, and Opportunities. *ACM Comput. Surv.* 1, 1 (December 2018).
7. S. B. Parikh, V. Patil, R. Makawana and P. K. Atrey, "Towards Impact Scoring of Fake News," 2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), 2019, pp. 529-533, doi: 10.1109/MIPR.2019.00107.
8. Lucienne T.M. Blessing, Amaresh Chakrabarti, DRM, a Design Research Methodology, Springer-Verlag London 2009, XVII, 397 páginas, <https://doi.org/10.1007/978-1-84882587-1>.