

Editorial Dossier

Dossier Métodos Computacionales y Big Data en la Investigación en Comunicación

Carlos Arcila Calderón, Universidad de Salamanca, Salamanca, España (carcila@usal.es)

Wouter Van Atteveldt, Universidad libre de Amsterdam, Ámsterdam, Países Bajos (wouter.van.attedveldt@vu.nl)

Damian Trilling, Universidad de Amsterdam, Ámsterdam, Países Bajos (dc.trilling@uva.nl)

Los métodos computacionales se han convertido en un área interesante y prometedora en el campo de la investigación en comunicación, no solo en el Norte Global sino también en regiones como Iberoamérica. En todo el mundo, la disponibilidad de datos en redes sociales, de datos de rastreo digital y de otro tipo ofrecen nuevas oportunidades para estudiar patrones de comunicación y de comportamiento social. Debido a que estos conjuntos de datos son más grandes que los tamaños de muestra típicos en las ciencias sociales, se ha despertado el interés de los académicos en aplicar metodologías novedosas para recopilar, visualizar y analizar dichos datos. A medida que el discurso público está cada vez más moldeado por lo que sucede en línea, se necesita con urgencia un cambio hacia nuevas fuentes de datos y nuevas metodologías.

Este número especial de *Cuadernos.info* refleja cómo el enfoque computacional está siendo incluido en la agenda de investigación en países como México, Chile, Brasil o España. Los cinco estudios incluidos en este monográfico muestran la gran cantidad de contenidos y mensajes –mayoritariamente en español y portugués, dos de los idiomas más hablados del mundo– que se pueden analizar desde una perspectiva computacional. El giro del análisis basado en el inglés y el Norte Global a nuevos contextos geográficos y lingüísticos en el mundo representa

un desafío significativo para el análisis computacional de la comunicación. En primer lugar, las técnicas como el procesamiento del lenguaje natural, el análisis de sentimientos, el modelado de temas o la clasificación de texto funcionan de manera diferente en cada idioma. Específicamente, pueden ser sensibles a las características propias del lenguaje y a las decisiones relacionadas que toma el investigador (por ejemplo, con respecto al pre-procesado). Sin embargo, como la mayor parte de la literatura y de los ejemplos todavía se producen en inglés, se carece de material de referencia (y de recursos, en general) para el análisis computacional de la comunicación en idiomas como el español y el portugués. En segundo lugar, las realidades culturales en diferentes partes del mundo, como en los países de América Latina, pueden estimular la formulación de nuevas preguntas de investigación o hipótesis que podrían responderse mediante un enfoque computacional. Esto puede a su vez inspirar nuevas líneas de investigación en otras partes del mundo.

Como editores invitados de este número, tuvimos el agrado de recibir muchas propuestas para ser incluidas en el dossier dedicado a *Métodos Computacionales y Big Data en la Investigación en Comunicación*. De este conjunto inicial de manuscritos, se seleccionaron finalmente cinco artículos después de un riguroso proceso de revisión por pares con doble ciego, coordinado por la editora general el equipo editorial de la revista. El resultado es un grupo de artículos selecto y bien equilibrado, que evidencia el impulso que los estudiosos de la comunicación latinoamericanos y españoles le están dando al tema. Debemos decir que a pesar de que la convocatoria de artículos fue más amplia (es decir, al incluir modelos basados en agentes o periodismo de datos), la mayoría de los manuscritos recibidos trataba principalmente sobre diferentes tipos de análisis de contenido con métodos computacionales en redes sociales, como Twitter e Instagram. Este hecho refleja la fuerte tradición de los investigadores de la comunicación en el análisis de contenido y su creciente interés en conocer cómo las redes sociales impactan a la esfera política y cultural de la región. Junto a la recopilación automática de tuits o imágenes de Instagram, encontramos la aplicación del análisis de sentimiento o la clasificación supervisada de texto como técnicas para abordar cuestiones sociales relevantes. Asimismo, el uso del análisis de redes sociales (SNA) en algunos de los artículos demuestra cómo esta técnica computacional bien establecida es fundamental para modelar, visualizar y comprender la dinámica compleja de los contenidos generados por los usuarios.

Para algunos lectores, el alcance de este número especial podría parecer muy técnico. Sin embargo, la mayoría de los manuscritos recibidos abordaban algunas cuestiones teóricas muy relevantes y tenían fuertes implicancias prácticas y sociales: desde la polarización o el discurso de odio, hasta la reacción del asesinato

de un líder indígena o las conversaciones durante la pandemia de COVID19. Esto revela el enorme potencial que tiene el enfoque computacional para abordar tanto las preocupaciones tradicionales como las actuales en la investigación de la comunicación. De hecho, no es casualidad que las principales asociaciones de investigadores de la región ya hayan creado grupos permanentes para discutir la aplicación de estos métodos. Es el caso del Grupo Comunicación digital, redes sociales y procesos, de la Asociación Latinoamericana de Investigadores de la Comunicación (ALAIC), o el Grupo emergente Métodos Computacionales y Análisis de Big Data en Comunicación, de la Asociación Española de Investigadores de la Comunicación (AE-IC), ambas iniciativas en línea con la división de Métodos Computacionales de la *International Communication Association* (ICA), a la que pertenecen los tres editores de este número especial.

El primer manuscrito de este dossier, escrito por los académicos mexicanos Manuel Cebral Loureda y Gabriela Elisa Sued Palmeiro, está dedicado a un análisis a gran escala de la conversación pública en mensajes de Twitter en español durante el inicio de la pandemia de COVID19. Es una excelente muestra del uso de técnicas computacionales para el análisis de texto (como el modelado de temas y el análisis de sentimientos) para comprender una realidad social en curso tan compleja como la pandemia. El segundo trabajo –escrito por los académicos chilenos Marcelo Santos, Oscar Jaramillo Castro y Daniel Aguirre Azócar–, combina SNA, analítica de medios sociales y análisis de contenido para cubrir la reacción en Twitter del asesinato de un miembro del pueblo originario Mapuche. Nos muestra una relevante discusión teórica de los procesos de radicalización dentro de las llamadas cámaras de eco y filtro burbuja. El tercer artículo, de los colegas brasileños Rita de Cássia Romeiro Paulino y Mariane Pires Ventura, mantiene su interés en Twitter y utiliza SNA, análisis textual y de emojis, para ver las discusiones mantenidas durante una protesta en línea (#Somos70porcento) en un contexto de crisis sanitaria. Este documento demuestra la importancia de las redes sociales como un nuevo medio para la discusión y el debate públicos. El cuarto manuscrito, de los académicos españoles Miguel Varela y Miguel Vicente, se centra en Instagram, al explicar un método para hacer *scraping* de imágenes a escala. Este informe mantiene la atención en un proceso de comunicación sobre la salud (una campaña en línea contra el cáncer), pero también tiene una relevancia más general, ya que sus conocimientos sobre el análisis de imágenes de una red social muy popular pueden ser útiles para una audiencia más amplia. Finalmente, el trabajo de los académicos emergentes españoles Javier Amores, David Blanco, Patricia Sánchez y Maximiliano Frías explica cómo utilizar la clasificación de texto supervisado (con algoritmos de aprendizaje superficial y profundo)

para detectar el discurso de odio motivado por razones políticas en español. Esta investigación muestra cómo la combinación de un etiquetado manual de buena calidad y algoritmos de última generación (es decir, *embeddings* y redes neuronales profundas) puede producir excelentes predicciones cuando es aplicado a contenidos en español.

Creemos que el surgimiento de una ciencia computacional de la comunicación está brindando más oportunidades a los investigadores para abordar preguntas de investigación nuevas y antiguas, y esto, a su vez, contribuirá a un cambio en la calidad de las investigaciones en medios y comunicación. Esperamos que disfrute y encuentre útiles estos documentos.

CARLOS ARCILA CALDERÓN, WOUTER VAN ATTEVELDT Y DAMIAN TRILLING

Editores invitados y autores del próximo libro

Computational Analysis of Communication. A practical introduction to the analysis of texts, networks, and images with code examples in Python and R.