

## DETECCIÓN DE IDEACIÓN SUICIDA EN REDES SOCIALES

GERARDO SIERRA

El suicidio representa un grave problema de salud pública que requiere atención urgente. Conforme a los datos proporcionados por el INEGI, la tasa de suicidio en México para 2020 fue de 6.2 por cada 100 000 habitantes, siendo la segunda causa de muerte para jóvenes entre 18 a 29 años. Estos números muestran la urgencia de investigaciones dedicadas a la identificación de personas en riesgo de morir por suicidio, con el fin de crear estrategias efectivas de prevención.

Las estrategias internacionales para la prevención de suicidio parten de la premisa que no es un acto irracional o instantáneo, sino que conlleva un proceso de deliberación, incluso, se muestran signos de ideación e intento suicida, los cuales pueden ser atendidos si se detectan oportunamente. Uno de los medios en los que llegan a expresar sus sentimientos donde se manifiestan estos signos, con mayor frecuencia entre los jóvenes, es a través de las redes sociales, ya sea en Facebook, Twitter o Instagram.

El propósito de este estudio se centra en analizar el uso del lenguaje en textos de las redes sociales, con temas de depresión y suicidio, después contrastarlo con un grupo de publicaciones aleatorias con temas no específicos, para identificar categorías de palabras en español relacionadas con el pensamiento suicida. Asimismo, elaborar un conjunto de frases vinculadas a los factores de riesgo de suicidio para verificar si estos contenidos son similares a los textos con temas de depresión o suicidio, diferentes a los textos que contienen temas aleatorios no específicos.

### El lenguaje en análisis psicológicos de la personalidad

La interacción entre los rasgos de personalidad toma un papel importante en el comportamiento suicida. El *software* de análisis psicológico de textos llamado Linguistic Inquiry and Word Count (LIWC), que consiste en un contador de palabras y un gran diccionario interno con varias palabras clasificadas en categorías psicológicamente relevantes, ha sido ampliamente utilizado en estudios psicométricos previos. El programa identifica las palabras en un archivo de texto que se corresponde con el diccionario y las asigna a distintas categorías, como palabras funcionales, expresiones sociales/informales, emociones, mecanismos cognitivos, etc. Como resultado, el programa calcula la frecuencia de cada categoría de palabras en relación

con el total de palabras del texto cargado. Por tanto, las puntuaciones de cada categoría de palabra se muestran en la salida como porcentajes: la proporción de palabras del diccionario de la herramienta identificadas en el total de textos analizados. Dichos datos permiten establecer relaciones con numerosos temas a través del uso de palabras.

LIWC tiene diccionarios en diferentes idiomas, como el español. Sin embargo, una limitante del programa es que sigue un enfoque cerrado para el análisis del lenguaje, esto es, su poder analítico depende en gran medida de cuán extenso sea su diccionario. Por esta razón, el mismo programa permite ampliar el vocabulario y las categorías.

La relación entre texto y psicología ha estado presente durante mucho tiempo, pero hasta hace poco ha ganado una nueva dirección que poco a poco va creando más importancia para detectar rasgos de personalidad. Esta nueva dirección lleva el análisis psicológico de los textos al dominio de Internet, las redes sociales y sus usos particulares del lenguaje, a saber, *netspeak*. La inclusión de expresiones y términos de *netspeak* está motivada por el significado que exhibe cada expresión individual de *netspeak* y cómo se puede agregar a un análisis psicológico del texto.

Si bien la versión 2015 del diccionario LIWC cuenta a *netspeak* como una dimensión lingüística; también, categoriza expresiones de este tipo en sus correspondientes dominios psicológicos. Fue necesario actualizar el diccionario interno del español, incluyendo la categoría *netspeak* y las expresiones relevantes que expresen las dimensiones psicológicas de los usuarios con ideación o pensamiento suicida. El contenido inicial de *netspeak* se tomó de un diccionario de léxico específico de Internet obtenido para el español de México (Osuna *et al.*, 2017), que consta de 247 palabras utilizadas en la red social. Para ampliar este léxico, se realizó un análisis del corpus de *tweets* y una sesión de lluvia de ideas de palabras destinadas a complementar este conjunto. Como resultado se obtuvo un lexicón *netspeak* de 1225 palabras, un total de 644 lemas. El siguiente paso fue clasificar estas palabras en sus respectivas categorías LIWC para que fueran incluidas en el diccionario. Para llevar a cabo el proceso de categorización, se realizó un procedimiento entre anotadores con tres jueces.

También, existe un fuerte vínculo entre los rasgos de personalidad con la ideación suicida. Una teoría que goza de un buen grado de aceptación en el mundo académico actual es la teoría del Big Five o de los cinco grandes rasgos de la personalidad. Propone que las diferencias individuales o personalidad más sobresalientes y socialmente relevantes (apertura a la experiencia, escrupulosidad, extroversión, neuroticismo y amabilidad) se encuentran ya codificadas en el lenguaje natural.

Por tal razón, se probó la funcionalidad y eficiencia del lexicón obtenido de *netspeak* en español para un análisis de agrupamiento de *tuits* con base en la teoría del Big Five (Sierra *et al.*, 2020).

### Análisis en las redes sociales

Para el estudio se recogieron contenidos de Facebook y Twitter y se elaboró un conjunto de frases vinculadas a factores de riesgo suicida a partir de instrumentos validados que evalúan varias dimensiones del fenómeno suicida. El análisis de estos datos de texto se realizó utilizando LIWC, con la versión mejorada de su léxico-diccionario en español. Luego, mediante la construcción de intervalos de confianza de arranque, se comparó la diferencia en las medias de los valores resultantes por categoría de palabra por grupo para identificar diferencias y similitudes entre los grupos.

Para los efectos del estudio, se obtuvieron dos publicaciones diferentes de grupos de Facebook escritas en español mexicano cuyos contenidos específicos eran suicidio y depresión. Para determinar la muestra a analizar, se realizó una selección manual de las publicaciones que reflejaban contenidos suicidas o depresivos. Se obtuvo un total de 169 publicaciones de ambos grupos, que sumaron 2006 palabras para alimentar la herramienta para el análisis. Estas publicaciones componían el grupo etiquetado como Grupo de publicaciones temáticas (GPT), con 82 del total tomadas del grupo con temática de suicidio y 87 del grupo con temática de depresión.

Se obtuvo una lista completa de frases correspondientes a cada factor de riesgo de suicidio, evaluada por tres jueces. Con ello, se compararon sus valoraciones entre sí para determinar si las frases evaluadas representaban correctamente el factor de riesgo pretendido. Este grupo de frases fue etiquetado como conjunto de frases relacionadas con factores de riesgo de suicidio o Grupo de Factores de Riesgo (GFR).

Como punto de comparación, se obtuvieron de *Twitter* 413 publicaciones públicas sin un tema específico (quitando anuncios o hipervínculos externos), esto es, 4,639 palabras para alimentar la herramienta para el análisis. Estas publicaciones formaron el grupo etiquetado como Grupo de control aleatorio (GCA).

Las entradas de los tres grupos diferentes (GPT, GCA y GFR) se analizaron de forma independiente con LIWC y el diccionario de español mejorado, que incluye expresiones informales y *netspeak*. Este análisis se centró en las categorías de palabras informadas por la literatura asociada relevante para identificar textos relacionados con el suicidio (Tabla 1). La intención fue detectar coincidencias y discrepancias con los enunciados de la literatura.

Tabla 1. Categorías relevantes de LIWC para el estudio

CATEGORÍA LIWC	EJEMPLOS
Yo	yo, mí, me
Tú	tú, ustedes, les
Nosotros	nuestros, nosotros, nos
Verbo 1ª persona singular (V1PS)	soy, quiero, busqué
Social	amigos, personas, familia
Emoción positiva	cariño, apoyo, felicidad
Emoción negativa	solo, llorar, desesperación
Ansiedad	problema, angustiado, miserable
Enfado	inaguantable, odio, frustrado
Tristeza	deprimido, piedad, pesimista
Oír	decir, escuchar, silencio
Sexual	deseo, sexo
Casa	cocina, ventana, cama
Muerte	morir, suicidio, funeral

### Resultados

Con una prueba U-Mann Whitney se compararon por pares entre los grupos, de acuerdo con cada categoría analizada, y se encontraron las categorías LIWC que mostraron ser significativamente diferentes entre los grupos. Para el ejercicio entre el GPT y el GCA, todas las categorías resultan significativamente diferentes. Al comparar las frases del GFR con las del GPT, sólo los Verbos en primera persona del singular, la Emoción negativa y el Enfado mostraron diferencias significativas. Finalmente, al comparar el GFR y el GCA, la mayoría de las categorías resultaron significativamente diferentes, excepto la categoría Oír.

Con base en los resultados, se verificaron que el GPT contiene categorías de palabras significativamente diferentes del GCA, y el GPT exhibe categorías de palabras similares a las del GFR (Figura 1). Dado que el análisis se realizó con textos en español, la categoría de Verbos en primera persona del singular tuvo una presencia muy frecuente, situación que no se reportó en ninguno de los estudios similares.

Las categorías de palabras que informan Muerte, Hogar, Tristeza, Enfado, Ansiedad, Emociones negativas y Verbos en primera persona del singular aparecen con más frecuencia en el GPT que en el GCA. Por otro lado, las categorías Nosotros, Tú, Emoción positiva, Oír y Sexual se encontraron muchas menos veces dentro del GPT que en el GCA.

## Conclusiones y trabajo futuro

Los presentes hallazgos son un primer paso hacia el análisis del lenguaje en los grupos de redes sociales, donde se puede obtener información valiosa sobre los pensamientos e ideas de las personas. Sin embargo, no se puede afirmar que las personas que utilizan palabras relacionadas con el suicidio tengan mayor probabilidad de morir por suicidio. Como se mencionó anteriormente, la muerte por suicidio involucra varias fases y sólo los pensamientos no son suficientes para explicar por qué las personas mueren por suicidio, por lo que sería necesario seguir investigando sobre el uso del lenguaje. Futuros estudios basados en el lenguaje deberían tener en cuenta las teorías psicológicas para considerar la transición del pensamiento a la tentativa y a la muerte por suicidio. Dicho trabajo implica desarrollar diccionarios que incluyan frases representativas de la capacidad adquirida, la pertenencia frustrada y la carga percibida.

También se ha presentado el *software* existente para esta tarea y algunos métodos de análisis que se están utilizando en una investigación en curso. La incorporación de estos elementos

del lenguaje en el análisis psicológico de las palabras a través de programas estilo LIWC representa tanto un desafío como una oportunidad. Además, queda por mejorar la implementación de algoritmos de agrupamiento para la investigación en curso. Ciertamente, hay muchos otros métodos y enfoques con los que se puede experimentar para desarrollar aún más este campo emergente. A medida que más investigaciones sigan adoptando esta nueva dirección del análisis psicológico del lenguaje en Internet, quedará más claro qué tipo de aspectos lingüísticos son clave para comprender las dinámicas individuales y sociales reflejadas en esta era mediada por Internet.

## Agradecimientos

Se agradece el apoyo del equipo de trabajo, integrado por Patricia Andrade Palos de la Facultad de Psicología y responsable del proyecto; Gemma Bel Enguix del Instituto de Ingeniería; Adriana Cabrera Mora y Alejandro Osornio, de la Facultad de Psicología; Luis García Nieto, de la Facultad de Ciencias; y Tamara Sierra, del Instituto de Terapia Cognitivo Conductual. Se realizó bajo el patrocinio de PAPIIT IG400119.

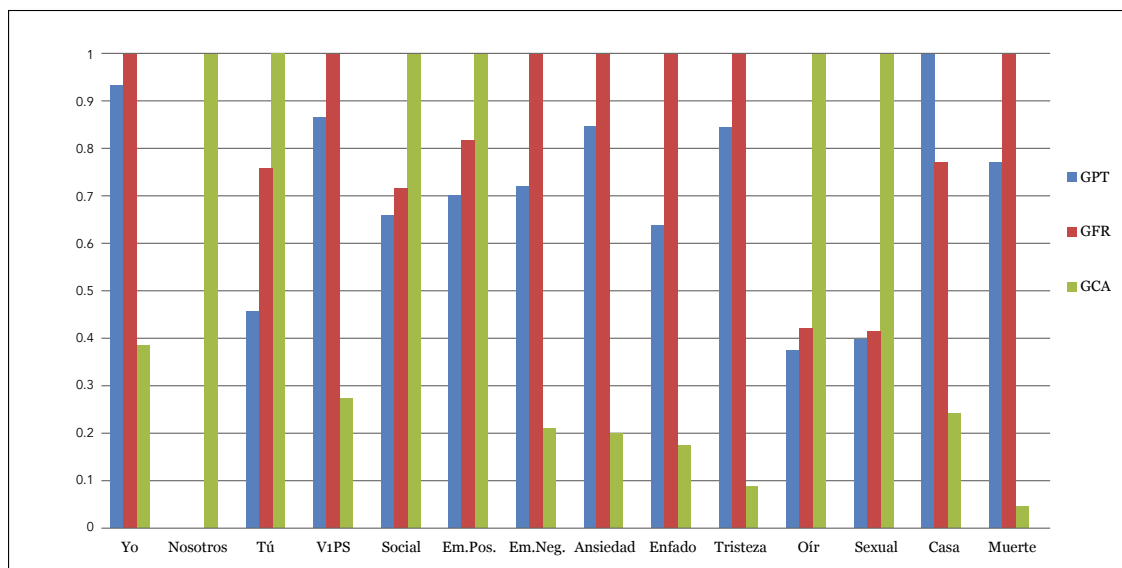


Figura 1. Frecuencias normalizadas de las categorías de LIWC para los tres grupos

## Referencias

- Osuna, F.; Minjares Ávila, K. J. y Sidorov, G. (2017). Compilation of Spanish Specific Lexicon of Social Networks. Centro de Investigación en Computación, IPN.
- Sierra, G.; Bel-Enguix, G.; Osornio-Arteaga, A.; Cabrera-Mora, A.; García-Nieto, L.; Bustos, A.; Romo-Anaya, A. M. y Silva-Cuevas, V. (2020). An exploration of personality traits detection in a Spanish Twitter corpus. Proceedings of the LREC 2020 workshop on: Resources and processing of linguistic data from people with various forms of cognitive/ psychiatric/developmental impairments (RaPID-3), 132–141.