

**DEPARTAMENTO DE MEDICINA PREVENTIVA Y SALUD
PÚBLICA, CIENCIAS DE LA ALIMENTACIÓN, TOXICOLOGÍA Y
MEDICINA LEGAL.**

Programa de Doctorado 3139, Medicina.

Línea de investigación: Medicina Preventiva y Salud Pública, Higiene
y Sanidad.



**Valoración de las Vacunas desde una Plataforma Social
con Aplicaciones de Uso en Promoción de la Salud.**

TESIS DOCTORAL.

Presentada por:

Hilary Piedrahita Valdés.

Dirigida por:

Dra. Carmen Saiz Sánchez.

Dra. Patricia Guillem Saiz.

Valencia, junio de 2021.



M^a del Carmen Saiz Sánchez, Doctora en Medicina y Profesora Titular del Departamento de Medicina Preventiva y Salud Pública, Ciencias de la Alimentación, Toxicología y Medicina Legal de la Universitat de València.

Patricia Guillem Saiz, Doctora en Medicina y Catedrática de Epidemiología, Medicina Preventiva y Salud Pública. Departamento de Odontología. Facultad de Ciencias de la Salud. Universidad Europea de Valencia.

CERTIFICAN:

Que **Hilary Piedrahita Valdés**, graduada en Medicina por la Universitat de València, ha realizado bajo nuestra dirección el presente trabajo de investigación, titulado: "**Valoración de las Vacunas desde una Plataforma Social con Aplicaciones de Uso en Promoción de la Salud**", el cual posee suficiente calidad científica y reúne los requisitos exigidos, para ser presentado y defendido para la obtención del Grado de Doctora por la Universitat de València.

Y para que así conste, firmamos el presente certificado en Valencia, mayo de 2021.



Fdo. Carmen Saiz Sánchez.



Fdo. Patricia Guillem Saiz.

DEDICATORIA

A las personas que en estos momentos están leyendo esta tesis doctoral, por su interés en el trabajo realizado.

A todos los usuarios de las redes sociales, por proporcionar los datos utilizados en esta investigación.

A los desarrolladores de *software*, por su esfuerzo y aporte en la creación de nuevas herramientas de análisis de opinión.

A todos los que contribuyen en el desarrollo de las tecnologías relacionadas con el *big data*.

A todos lo que dedican sus esfuerzos a la lucha contra la desinformación sobre las vacunas.

LICENCIA

Esta tesis doctoral, titulada “Valoración de las Vacunas desde una Plataforma Social con Aplicaciones de Uso en Promoción de la Salud”, realizada por Doña Hilary Piedrahita Valdés, está sujeta a la **Licencia Internacional Pública de Atribución/Reconocimiento-NoComercial-SinDerivados 4.0 de Creative Commons** ("Licencia Pública").



Por tanto, esta obra se puede compartir y comunicar públicamente, siempre que se reconozca su autoría y no se persigan fines comerciales. Así mismo, no se permite la distribución de obras derivadas.

RESUMEN

Introducción

La exposición a información negativa o errónea sobre las vacunas es una de las causas del aumento de la duda vacunal en la última década. De hecho, según la Organización Mundial de la Salud, la reticencia vacunal fue una de las diez principales amenazas para la salud mundial en 2019. Por ello, es importante incluir la monitorización y análisis del contenido del discurso sobre vacunas en las redes sociales en el abordaje de la reticencia vacunal. Las técnicas automáticas posibilitan el acceso, el procesamiento y el análisis de la gran cantidad de datos extraídos de estas plataformas.

Material y métodos

Estudio retrospectivo de la opinión sobre las vacunas en Twitter entre el 1 de junio de 2011 y el 30 de abril de 2019. Mediante diversas librerías de Python, determinamos la geolocalización de los usuarios y realizamos un análisis semántico de los tuits. Clasificamos manualmente a los usuarios en “institucionales” o “no institucionales” y en “bots” o “no bots” con la herramienta Botometer.

Entrenamos un algoritmo híbrido basado en léxico y máquinas de soporte vectorial para la clasificación de la polaridad de sentimiento de los tuits hacia las vacunas (negativo, positivo o neutro). Comparamos el rendimiento de nuestro algoritmo con el algoritmo VADER, no específico para la opinión sobre vacunas. Construimos un clasificador automático basado en diccionarios para la categorización de los tuits en temas no excluyentes. Realizamos el análisis estadístico mediante el paquete estadístico *Statistical Package for Social Sciences* (SPSS) para Windows versión 23.0.

Resultados

Analizamos 1 499 227 tuits de 278 858 usuarios. Geolocalizamos al 55,87% de los usuarios. El análisis semántico permitió obtener una visión preliminar de los temas más frecuentes en los tuits. Nuestro clasificador de la polaridad presentó un rendimiento alto, mayor que VADER. El 69,36% de los tuits eran neutros, el 21,78% positivos y el 8,86% negativos. Hubo una polarización del discurso a lo largo del estudio, con picos de tuits positivos cada abril. A mitad de semana, aumentó la proporción de tuits positivos y disminuyó la proporción de tuits negativos. Identificamos las polaridades por localizaciones y comparamos los parámetros de interacción social por polaridades.

Los usuarios no institucionales publicaron más contenido negativo. La mayoría de los usuarios con varios tuits tuvo un discurso totalmente positivo, negativo o neutro. Describimos los temas más frecuentes según la polaridad de los tuits, el tipo de vacuna y el grupo de población diana de la vacunación. Evidenciamos correlaciones significativas en varios países entre: a) los tuits promotores de la vacuna VPH y las coberturas vacunales contra VPH, b) los tuits reticentes sobre la vacuna SRP y el número de casos de sarampión.

Conclusión

Es posible extraer, estructurar y analizar de forma rápida, económica y precisa, grandes cantidades de datos de las conversaciones en Twitter aplicando técnicas de inteligencia artificial. La información derivada de este tipo de análisis puede ser útil para mejorar las estrategias de salud pública destinadas a disminuir la duda vacunal.

Palabras clave: reticencia vacunal, duda vacunal, minería de opinión, análisis de sentimientos, modelado de tópicos, clasificación automática, Twitter.

RESUM

Introducció

L'exposició a informació negativa o errònia sobre les vacunes és una de les causes de l'augment del dubte vacunal en l'última dècada. De fet, segons l'Organització Mundial de la Salut, la reticència vacunal fou una de les deu principals amenaces per a la salut mundial en 2019. Per aquest motiu, és important incloure en l'abordatge de la reticència vacunal la vigilància i l'anàlisi del contingut del discurs sobre vacunes en les xarxes socials. Les tècniques automàtiques possibiliten l'accés, el processament i l'anàlisi de la gran quantitat de dades extretes d'aquestes plataformes.

Material y mètodes

Estudi retrospectiu de l'opinió sobre les vacunes en Twitter entre l'1 de juny de 2011 i el 30 d'abril de 2019. Mitjançant diverses llibreries de Python, determinarem la geolocalització dels usuaris i realitzarem una anàlisi semàntica dels tuits. Classificarem manualment als usuaris en "institucionals" o "no institucionals" i en "bots" o "no bots" amb la ferramenta Botometer.

Entrenarem un algoritme híbrid basat en lèxic i màquines de suport vectorial per a la classificació de la polaritat de sentiment dels tuits cap a les vacunes (negatiu, positiu o neutre). Compararem el rendiment del nostre algoritme amb l'algoritme VADER, no específic per a l'opinió sobre vacunes. Construïrem un classificador automàtic basat en diccionaris per a la categorització dels tuits en temes no excloents. Realitzarem l'anàlisi estadística mitjançant el paquet estadístic *Statistical Package for Social Sciences* (SPSS) per a Windows versió 23.0.

Resultats

Analitzàrem 1 499 227 tuits de 278 858 usuaris. Geolocalitzàrem al 55,87% dels usuaris. L'anàlisi semàntica ens permeté obtindre una visió preliminar dels temes més freqüents al tuits. El nostre classificador de la polaritat presentà un rendiment alt, major que VADER. El 69,36% dels tuits eren neutres, el 21,78% positius i el 8,86% negatius. Hagué una polarització del discurs al llarg de l'estudi, amb pics de tuits positius cada abril. A meitat de setmana, va augmentar la proporció de tuits positius i va disminuir la proporció de tuits negatius. Identificàrem les polaritats per localitzacions i comparàrem els paràmetres d'interacció social per polaritats.

Els usuaris no institucionals publicaren més contingut negatiu. La majoria dels usuaris amb diversos tuits tingueren un discurs totalment positiu, negatiu o neutre. Describírem els temes més freqüents segons la polaritat dels tuits, el tipus de vacuna i el grup de població diana de la vacunació. Evidenciàrem correlacions significatives en diversos països entre: a) els tuits promotors de la vacuna VPH i les cobertures vacunals contra VPH, b) els tuits reticents sobre la vacuna SRP i el nombre de casos de pallola.

Conclusió

És possible extraure, estructurar i analitzar de manera ràpida, econòmica i precisa, grans quantitats de dades de les conversacions en Twitter aplicant tècniques d'intel·ligència artificial. La informació derivada d'aquesta mena d'anàlisi pot ser útil per a millorar les estratègies de salut pública destinades a disminuir el dubte vacunal.

Paraules clau: reticència vacunal, dubte vacunal, mineria d'opinió, anàlisi de sentiments, modelat de tòpics, classificació automàtica, Twitter.

ABSTRACT

Introduction

Exposure to misinformation and disinformation about vaccines is one of the causes of the increase in vaccine-hesitant beliefs during the last decade. Vaccine hesitancy was one of the top ten global health threats in 2019, according to the World Health Organization. Hence, it is important to include social media content monitoring and analysis in the interventions addressing vaccine hesitancy. Automatic techniques make it possible to access and analyse the large amount of data extracted from these platforms.

Material and methods

We performed a retrospective study of the opinion about vaccines on Twitter between June 1st, 2011, and April 30th, 2019. We geolocated the users and executed a semantic analysis of the tweets using Python libraries. We manually classified the users into “institutional users” and “non-institutional users”. We used the tool Botometer to distinguish between bots and humans.

We trained a hybrid algorithm based on lexicon and support vector machines to classify the sentiment polarity of the tweets towards vaccines (negative, positive, or neutral). We compared its performance with the VADER algorithm, not specific for vaccine opinion mining. We built an automatic classifier based on dictionaries to categorise the tweets in non-exclusive topics. We conducted the statistical analysis using the Statistical Package for Social Sciences (SPSS) for Windows version 23.0.

Results

We analysed 1 49 227 tweets from 278 858 users. We geolocated 55,87% of the users. The semantic analysis gave us an approximation to the topic modelling results. Our polarity classifier had a high performance, better than VADER. 69,36% of the tweets were neutral, 21,78% positive and 8,86% negative. There was a polarization of the discourse throughout the study, with peaks of positive tweets every April. In the middle of the week, the proportion of positive tweets increased, while the proportion of negative tweets decreased. We identified the predominant polarities by location and compared the parameters of social interaction among polarities.

Non-institutional users posted more negative content than institutional users. Most of the users that had written several tweets had a totally positive, negative, or neutral speech. We described the most frequent topics according to the polarity of the tweets, the type of vaccine and the target population for vaccination. There were significant correlations in several countries between: a) tweets promoting HPV vaccine and HPV vaccine coverage, b) negative tweets about the MMR vaccine and the number of measles cases.

Conclusion

Large amounts of data from Twitter conversations can be extracted, structured, and analysed quickly, low-cost, and accurately, using artificial intelligence techniques. The information obtained from this kind of analysis can help us to improve public health strategies against vaccine hesitancy.

Keywords: vaccine hesitancy, opinion mining, sentiment analysis, topic modelling, automatic classification, Twitter.

ÍNDICE

CAPÍTULO I. MARCO TEÓRICO

INTRODUCCIÓN	3
I. VACUNAS Y VACUNACIÓN	6
<i>A. Definición de Inmunización y Vacunación</i>	6
<i>B. Vacunas</i>	6
1. Concepto de vacuna	6
2. Clasificación de las vacunas	7
3. Composición de las vacunas	14
4. Características generales de las vacunas	15
<i>C. Cobertura Vacunal</i>	18
II. RETICENCIA VACUNAL (VACCINE HESITANCY)	20
<i>A. Concepto de Reticencia Vacunal</i>	20
<i>B. Modelo de Reticencia Vacunal</i>	21
<i>C. Determinantes de la Reticencia Vacunal</i>	23
<i>D. Impacto de la Reticencia Vacunal</i>	25
<i>E. Reticencia y Demanda Vacunal</i>	27
<i>F. Reticencia Vacunal a lo Largo de la Historia</i>	28
1. Reticencia vacunal: siglo XIX	28
2. Reticencia vacunal: siglo XX	30
3. Reticencia vacunal: siglo XXI	32
III. INFORMACIÓN Y DESINFORMACIÓN SANITARIA EN LA ERA DIGITAL	36
<i>A. Internet y la Red Informática Mundial</i>	36
<i>B. Redes Sociales</i>	39

IV. MÉTODOS COMPUTACIONALES PARA EL ESTUDIO DE LA DESINFORMACIÓN	41
<i>A. Minería de Textos</i>	43
<i>B. Procesamiento del Lenguaje Natural</i>	43
<i>C. Minería de Opiniones / Análisis de Sentimientos</i>	45
1. Técnicas basadas en el léxico	46
2. Técnicas basadas en el aprendizaje automático	47
3. Técnicas híbridas	57
<i>D. Modelado de Tópicos</i>	58
<i>E. Evaluación de los Métodos de Clasificación Automática de Textos</i>	59
V. INFORMACIÓN Y DESINFORMACIÓN SOBRE VACUNAS EN TWITTER	61
<i>A. Twitter</i>	61
1. Generalidades sobre Twitter	61
2. Twitter como herramienta de investigación en salud	63
3. Aplicaciones de Twitter en Salud Pública	65
<i>B. Estudios sobre la Opinión acerca de las Vacunas en Twitter</i>	67
<i>C. Justificación del Estudio</i>	74
CAPÍTULO II. HIPÓTESIS Y OBJETIVOS	
I. HIPÓTESIS	79
II. OBJETIVOS	79
<i>A. Objetivo General</i>	79
<i>B. Objetivos Específicos</i>	79

CAPÍTULO III. MATERIAL Y MÉTODOS

I. DISEÑO DEL ESTUDIO	83
II. FUENTE DE INFORMACIÓN	83
III. MARCO TEMPORAL. CRITERIOS DE INCLUSIÓN Y EXCLUSIÓN	84
IV. ENTORNO DE TRABAJO	84
<i>A. Servidor para la Base de Datos</i>	84
<i>B. Estación de Trabajo</i>	85
V. TECNOLOGÍAS Y HERRAMIENTAS UTILIZADAS	85
<i>A. Aplicación para la Extracción de los Datos de Twitter</i>	86
<i>B. Python</i>	87
<i>C. Visual Studio Code</i>	89
<i>D. Visual Studio (C#)</i>	89
<i>E. Botometer</i>	90
VI. PROCEDIMIENTO DE OBTENCIÓN DE DATOS	91
VII. ALMACENAMIENTO DE DATOS	92
VIII. DEPURACIÓN DE LOS DATOS	92
IX. ANONIMIZACIÓN DE LOS TUI TS	93
X. GEOLOCALIZACIÓN	93
XI. PREPROCESAMIENTO DE LOS DATOS	94
XII. ANÁLISIS SEMÁNTICO	95
XIII. CLASIFICACIÓN DE TUI TS SEGÚN LA POLARIDAD DE SENTIMIENTOS	95

XIV. CLASIFICACIÓN DE TUI TS POR TEMAS	97
XV. CLASIFICACIÓN DE LOS USUARIOS	101
XVI. VARIABLES DEL ESTUDIO	102
XVII. ANÁLISIS ESTADÍSTICO	105
XVIII. CONSIDERACIONES ÉTICAS	107

CAPÍTULO IV. RESULTADOS

I. ANÁLISIS GLOBAL DE LOS TUI TS	111
<i>A. Características Generales de los Tuits</i>	111
<i>B. Evolución en el Número de Tuits (junio 2011 - abril 2019)</i>	112
II. ANÁLISIS SEMÁNTICO	114
<i>A. Unigramas a Nivel de Palabra</i>	114
<i>B. Bigramas a Nivel de Palabra</i>	115
III. ANÁLISIS DE LA POLARIDAD DE SENTIMIENTOS	116
<i>A. Métricas y Comparación con el AS de Vader</i>	116
<i>B. Polaridad de Sentimientos por Idioma</i>	118
<i>C. Evolución Temporal de la Polaridad de Sentimientos</i>	119
<i>D. Polaridad de Sentimientos por Mes</i>	122
1. Frecuencia absoluta mensual de tuits por polaridad	122
2. Frecuencia relativa mensual de tuits por polaridad	123
<i>E. Polaridad de Sentimientos por Día de la Semana</i>	124
1. Frecuencia absoluta de tuits por polaridad y día de la semana	124
2. Frecuencia relativa de tuits por polaridad y día de la semana	124
<i>F. Polaridad de Sentimientos por Países</i>	125
<i>G. Interacción Social Generada según la Polaridad de los Tuits</i>	129

IV. ESTUDIO DE LOS USUARIOS	131
<i>A. Idioma de los Usuarios</i>	131
<i>B. Clasificación de los Usuarios</i>	131
<i>C. Número de Tuits por Usuario</i>	131
<i>D. Polaridad de Sentimiento según el Usuario</i>	131
<i>E. Usuarios con más Tuits</i>	132
V. ANÁLISIS DE TEMAS	135
<i>A. Campañas de Promoción de las Vacunas</i>	135
<i>B. Determinantes de Reticencia Vacunal</i>	137
<i>C. Determinantes de Reticencia Vacunal por Localización</i>	142
<i>D. Tipo de Vacuna</i>	145
<i>E. Población Diana</i>	150
<i>F. Tipo de Instituciones Mencionadas</i>	153
<i>G. Otros Conceptos Relacionados con la Reticencia Vacunal</i>	154
1. Autismo	154
2. Cáncer	156
3. Composición de las vacunas	157
4. Homeopatía	158
VI. RELACIÓN CON DATOS DE LA OMS	159
<i>A. Coberturas Vacunales frente al VPH y Campañas en Twitter</i>	159
<i>B. Casos de Sarampión y Reticencia Vacunal en Twitter</i>	160
CAPÍTULO V. DISCUSIÓN Y CONCLUSIONES	
I. DISCUSIÓN	169
<i>A. Resultados del Análisis Global de los Tuits</i>	169
1. Características generales de los tuits	169
2. Evolución en el número de tuits (junio 2011 - abril 2019)	170
<i>B. Resultados del Análisis Semántico</i>	170

<i>C. Resultados del Análisis de la Polaridad de Sentimientos</i>	172
1. Métricas y comparación con el AS de Vader	172
2. Evolución temporal de la polaridad de sentimientos y resultados por meses del año	174
3. Polaridad de sentimientos por día de la semana	176
4. Polaridad de sentimientos por países	177
5. Interacción social según la polaridad de sentimiento	178
<i>D. Resultados del Análisis de los Usuarios</i>	180
1. Resultados del estudio de todos los usuarios	180
2. Resultados del estudio de los usuarios con más tuits	181
<i>E. Resultados del Análisis de Temas</i>	183
1. Tuits según la polaridad y la variable “campaña”	183
2. Determinantes de reticencia vacunal	186
3. Tipo de vacuna	192
4. Población diana	194
5. Instituciones mencionadas	197
6. Otros conceptos relacionados con la reticencia vacunal	198
<i>F. Resultados del Análisis de la Relación con datos de la OMS</i>	201
1. Coberturas vacunales frente al VPH y campañas en Twitter	201
2. Casos de sarampión y reticencia vacunal en Twitter	203
<i>G. Limitaciones del Estudio</i>	205
<i>H. Fortalezas del Estudio y Futuras Líneas de Investigación.</i>	209
II. CONCLUSIONES	211
CAPÍTULO VI. BIBLIOGRAFÍA	
BIBLIOGRAFÍA	217
CAPÍTULO VII. ANEXO	
INFORME DEL COMITÉ DE ÉTICA DE INVESTIGACIÓN	235

ABREVIATURAS Y SIGLAS

A

AEMPS: Agencia Española del Medicamento y Productos Sanitarios.

AFR: Región de África de la OMS.

ALD: Asignación Latente de Dirichlet.

AMR: Región de las Américas de la OMS.

API: *Application Programming Interface*, Interfaz de Programación de Aplicaciones.

AS: Análisis de Sentimientos.

C

CAP: *Complete Automation Probability*, Probabilidad Completa de Automatización.

CDC: Centros para el Control y la Prevención de Enfermedades.

CNN: *Convolutional Neural Network*, Redes Neuronales Convolucionales.

CPU: *Central Processing Unit*, Unidad Central de Procesamiento.

E

EE. UU.: Estados Unidos.

EMA: Agencia Europea del Medicamento.

EMR: Región del Mediterráneo Oriental de la OMS.

EUR: Región de Europa de la OMS.

EV: Efectividad Vacunal.

J

JSON: *JavaScript Object Notation*, Notación de Objeto de JavaScript.

M

ME: Máxima Entropía.

MSV: Máquinas de Soporte Vectorial.

O

OSoMe: *Observatory on Social Media*, Observatorio en Redes Sociales.

OMS: Organización Mundial de la Salud.

ONU: Organización de las Naciones Unidas.

OR: *Odds Ratio*.

N

NLTK: *Natural Language Toolkit*.

P

PLN: Procesamiento del Lenguaje Natural.

POST: *Part-Of-Speech Tagging*, etiquetado grammatical.

R

RAE: Real Academia Española.

RAM: *Random-Access Memory*, Memoria De Acceso Aleatorio.

RI: Rango Intercuartílico.

S

SAGE: *Strategic Advisory Group of Experts on Immunization*.

SEAR: Región de Asia Sudoriental de la OMS.

SPSS: *Statistical Package for Social Sciences*.

SRP: Vacuna triple vírica (Sarampión, Rubéola, Parotiditis).

U

UNICEF: Fondo de las Naciones Unidas para la Infancia.

URL/URLs: *Uniform Resource Locator*, Localizador de Recursos Uniforme.

V

VADER: Valence Aware Dictionary and sEntiment Reasoner.

VIH: Virus de la Inmunodeficiencia Humana.

VPH: Virus del Papiloma Humano.

W

WPR: Región del Pacífico Occidental de la OMS.

LISTA DE FIGURAS

 Figura 1. Coberturas vacunales mundiales, OMS-UNICEF (2019).	18
 Figura 2. <i>Continuum</i> de la reticencia vacunal.	21
 Figura 3. Modelo de reticencia vacunal de las “tres Cs”.	22
 Figura 4. Técnicas de análisis de sentimientos.	46
 Figura 5. Técnicas de aprendizaje automático.	48
 Figura 6. Diagrama de flujo del aprendizaje supervisado.	50
 Figura 7. Enfoques de aprendizaje automático supervisado.	51
 Figura 8. Árbol de decisión.	52
 Figura 9. Diagrama de flujo de la clasificación mediante reglas.	52
 Figura 10. Estructura de las máquinas de soporte vectorial.	53
 Figura 11. Arquitectura de una red neuronal.	54
 Figura 12. Clasificación basada en la máxima entropía.	55
 Figura 13. Regiones y límites de decisión.	56
 Figura 14. Ejemplo de red bayesiana.	57
 Figura 15. Características del servidor para la base de datos.	85
 Figura 16. Características de la estación de trabajo.	85
 Figura 17. APIs de Twitter.	86
 Figura 18. Aplicación para la clasificación manual de tuits.	89
 Figura 19. Clasificación mediante Botometer.	90
 Figura 20. Geocodificación mediante Python.	94
 Figura 21. Tuits por idioma.	111
 Figura 22. Frecuencia de tuits en las regiones de la OMS.	112
 Figura 23. Número de tuits por idioma (junio 2011 – abril 2019).	113
 Figura 24. 30 palabras más frecuentes por idioma.	114
 Figura 25. 30 bigramas más frecuentes por idioma.	115
 Figura 26. Porcentaje de tuits por polaridad e idioma.	118
 Figura 27. Número mensual de tuits por polaridad de sentimiento (junio 2011 - abril de 2019).	119
 Figura 28. Frecuencia relativa de los tuits clasificados por polaridad de sentimiento frente a las vacunas (junio 2011 - abril 2019).	121
 Figura 29. Número de tuits por polaridad de sentimiento y mes.	122

Figura 30. Porcentaje de tuits por polaridad y mes.	123
Figura 31. Porcentaje de tuits por polaridad y día de la semana.	124
Figura 32. Frecuencia de tuits en las regiones de la OMS.	125
Figura 33. Porcentaje de tuits negativos (países $\geq 1\ 000$ tuits).	127
Figura 34. Porcentaje de tuits positivos (países $\geq 1\ 000$ tuits).	128
Figura 35. Porcentaje de tuits por polaridad de sentimiento según tengan o no retuits, favoritos/ me gusta y respuestas.	130
Figura 36. Porcentaje de tuits de cada polaridad según la variable “campañas”.	135
Figura 37. Evolución del número de tuits de cada polaridad según la variable “campañas”.	136
Figura 38. Número de tuits por polaridad y tema.	137
Figura 39. Número de tuits por polaridad, tema y año (2012-2018).	138
Figura 40. Porcentaje de tuits por polaridad y tema.	139
Figura 41. Porcentaje de tuits de cada tema según la polaridad y la variable “campaña”.	141
Figura 42. Tuits negativos de cada región clasificados por temas.	142
Figura 43. Tuits negativos de cada país clasificados por temas.	144
Figura 44. Número de tuits por polaridad y tipo de vacuna.	145
Figura 45. Número de tuits por polaridad, tipo de vacuna y año (2012 - 2018).	146
Figura 46. Porcentaje de tuits por polaridad y tipo de vacuna.	148
Figura 47. Número de tuits por polaridad y grupo poblacional.	150
Figura 48. Número de tuits por polaridad e institución mencionada.	153
Figura 49. Número de tuits sobre autismo por polaridad.	154
Figura 50. Correlogramas cruzados entre las series “casos mensuales de sarampión” y “tuits sobre SRP”.	161

LISTA DE TABLAS

 Tabla 1. Clasificación de las vacunas	9
 Tabla 2. Clasificación microbiológica de las vacunas	10
 Tabla 3. Vacunas víricas según su naturaleza	12
 Tabla 4. Vacunas bacterianas según su naturaleza	13
 Tabla 5. Trabajos relacionados	68
 Tabla 6. Librerías de Python utilizadas en este estudio	87
 Tabla 7. Temas	98
 Tabla 8. Variables del estudio	102
 Tabla 9. Variables del estudio de los usuarios con más tuits	104
 Tabla 10. Métricas de los algoritmos de clasificación	116
 Tabla 11. Estadísticos descriptivos de los retuits, favoritos/me gusta y respuestas según la polaridad de sentimiento de los tuits	129
 Tabla 12. Clasificación de los usuarios con más tuits en inglés	133
 Tabla 13. Clasificación de los usuarios con más tuits en castellano	134
 Tabla 14. Porcentaje de tuits negativos de cada tema según la vacuna	149
 Tabla 15. Porcentaje de tuits negativos de cada tema según la población diana	151
 Tabla 16. Porcentaje de tuits negativos de cada vacuna según la población diana	152
 Tabla 17. Correlación entre las coberturas vacunales y el número de tuits sobre VPH	159

CAPÍTULO I

MARCO TEÓRICO

INTRODUCCIÓN

Desde los albores de la humanidad, grandes hombres y mujeres de diversas ciencias han trabajado para erradicar, curar y prevenir enfermedades y epidemias que devastan a los seres humanos. En 1796, **Edward Jenner** descubrió la primera vacuna. En 1880, Louis Pasteur ya indicaba que el agua era un transmisor de microorganismos. Con ello, el hombre llegó a la potabilización del agua. Estas dos grandes intervenciones son las medidas que más han contribuido en la reducción de la mortalidad por enfermedades infecciosas.

La Organización Mundial de la Salud (OMS) (1,2) estima que la inmunización artificial evita 2,5 millones de muertes anuales entre la población infantil, lo que la convierte en una de las medidas de salud pública más costo-eficaces en la prevención de enfermedades. Además, calculan que se podrían evitar 1,5 millones de muertes adicionales si se consiguieran los objetivos de tasas vacunales mundiales para cada enfermedad. (2)

No obstante, algunas de estas vacunas están **infrautilizadas**. Por una parte, esto está justificado por el déficit de recursos económico, la inequidad en el acceso a los sistemas sanitario y la falta de infraestructuras y centros logísticos, sobre todo en países en vías de desarrollo. El uso de las vacunas también depende de la demanda activa de la vacunación por parte de las comunidades, que a su vez está influida por el grado de conocimiento de la población sobre la eficacia, seguridad e importancia de las vacunas. (1)

En el siglo XXI, las redes sociales tienen un importante papel en la difusión de mitos y mensajes científicos sobre vacunas. Un rumor

sobre una vacuna extendido por la red puede aumentar el rechazo hacia esa vacuna, lo que lleva a una disminución de la tasa de vacunación. A su vez, esto contribuye al incremento de brotes epidémicos de enfermedades inmunoprevenibles, con el aumento de la morbimortalidad que esto conlleva.

La información que circula en **Internet** y en las **redes sociales** puede utilizarse como complemento en la vigilancia activa y la monitorización a tiempo real de la duda vacunal. Asimismo, los profesionales de la salud pueden usar estos medios para difundir datos científicos que motiven a los usuarios a aceptar y a adoptar hábitos más saludables. De hecho, en el caso de la vacunación, Ahmed et al. (3) concluyen que los usuarios que consultan **Twitter** y **Facebook** como fuente de información sobre salud tienen mayor tendencia a vacunarse frente a la gripe. Esto convierte a las redes sociales en un potencial instrumento para la promoción de la salud.

Por ello, este trabajo nace con la intención de estudiar la opinión de los usuarios de Twitter respecto a la vacunación, en un periodo de ocho años, mediante técnicas de inteligencia artificial.

La presente tesis doctoral está compuesta por **siete capítulos**:



Capítulo I: marco teórico y justificación del estudio.



La primera parte del capítulo explica los conceptos básicos de la vacunología, en torno a los cuales se genera el debate y la desconfianza vacunal estudiada en capítulos posteriores.



La segunda parte del capítulo se centra en la duda vacunal: el concepto, sus causas y su evolución desde el siglo XIX hasta la actualidad.

La tercera parte del capítulo explica el papel de Internet y las redes sociales como fuente generadora de desinformación y los principales métodos de automatizados para el estudio de esta desinformación.

La cuarta parte de este capítulo incluye:

1. Una revisión bibliográfica los trabajos realizados por otros autores en el campo de la duda vacunal en las redes sociales mediante técnicas automatizadas.
2. La justificación del estudio.

 **Capítulo 2:** hipótesis y objetivos.

 **Capítulo 3:** descripción del material y métodos.

 **Capítulo 4:** resultados de la investigación original dividida en seis estudios:

- Descripción global de los tuits de la base de datos.
- Análisis semántico de los mensajes incluidos en el estudio.
- Análisis de la polaridad de sentimientos hacia las vacunas.
- Análisis de las características de los usuarios.
- Análisis de contenido de los tuits según la temática.
- Análisis de la relación entre la reticencia vacunal en Twitter y las coberturas vacunales/incidencia de casos según la OMS.

 **Capítulo 5:** discusión y conclusiones de la investigación original, con énfasis en las contribuciones en salud pública.

 **Capítulo 6:** bibliografía.

 **Capítulo 7:** anexo.

I. VACUNAS Y VACUNACIÓN

A. Definición de Inmunización y Vacunación

El “proceso de inducción o provisión de inmunidad artificial a un individuo sano susceptible mediante la administración de un producto inmunobiológico” recibe el nombre de **inmunización**. (4)

Según el mecanismo de adquisición de la inmunidad, puede distinguirse entre inmunización activa e inmunización pasiva (5):



Inmunización pasiva: Se logra mediante la transferencia de anticuerpos (inmunoglobulinas o sueros) producidos por otro animal o humano. Los efectos protectores se objetivan a corto plazo, pero suelen conferir una protección temporal, de semanas o meses de duración.



Inmunización activa (vacunación): Basada en la inducción de una respuesta inmune humoral y/o celular frente al patógeno mediante la administración de vacunas. Los anticuerpos formados como consecuencia de la inmunización activa suelen perdurar durante varios años y, en algunos casos, permanecen toda la vida.

B. Vacunas

1. Concepto de vacuna

Según la **OMS** (6), “se entiende por **vacuna** cualquier preparación destinada a generar inmunidad contra una enfermedad estimulando

la producción de anticuerpos. Puede tratarse, por ejemplo, de una suspensión de microorganismos muertos o atenuados, o de productos o derivados de microorganismo”. La **vía de administración** más habitual de las vacunas es la inyección intramuscular o subcutánea, aunque algunas de ellas pueden dispensarse por medio de vaporizadores nasales u orales.

2. Clasificación de las vacunas

En la actualidad, la **lista de enfermedades** prevenibles mediante vacunación está conformada por 25 entidades (1):

- ▶ 1. Cólera.
- ▶ 2. COVID-19.
- ▶ 3. Difteria.
- ▶ 4. Fiebre amarilla.
- ▶ 5. Fiebre tifoidea.
- ▶ 6. Encefalitis japonesa.
- ▶ 7. Enfermedad neumocócica invasiva.
- ▶ 8. Gripe.
- ▶ 9. *Haemophilus influenzae b*.
- ▶ 10. Hepatitis A.
- ▶ 11. Hepatitis B.

- ▶12. Herpes zóster.
- ▶13. Meningitis A,B,C,W e Y.
- ▶14. Parotiditis.
- ▶15. Poliomielitis.
- ▶16. Rabia.
- ▶17. Rotavirus.
- ▶18. Rubéola.
- ▶19. Sarampión.
- ▶20. Tétanos.
- ▶21. Tosferina.
- ▶22. Tuberculosis.
- ▶23. Varicela.
- ▶24. Viruela.
- ▶25. Virus del papiloma humano (VPH).

Para cada una de estas enfermedades, hay una o varias vacunas comercializadas. Existen múltiples clasificaciones de estas vacunas. En la Tabla 1 se detallan **cinco clasificaciones modificadas** a partir de la división de las vacunas realizada por la **Sociedad Española de Geriatría y Gerontología (7)**.

Tabla 1. Clasificación de las vacunas	
Clasificación según su naturaleza	☞ Atenuadas.
	☞ Inactivadas.
Clasificación microbiológica	☞ Bacterianas.
	☞ Víricas.
Clasificación según la tecnología de fabricación	☞ Clásicas.
	☞ Sintéticas.
	☞ Recombinantes.
Clasificación según su composición	☞ Monovalentes.
	☞ Polivalentes.
	☞ Combinadas.
Clasificación sanitaria	☞ Sistemáticas.
	☞ No sistemáticas.

Fuente: elaboración propia a partir de Sociedad Española de Geriátría y Gerontología (7).

a) **Clasificación según su naturaleza (5,7):** según la capacidad infectiva del microorganismo contenido en la vacuna.

 **Atenuadas:** Contienen microorganismos vivos enteros de virulencia reducida, conseguidos a través de siembras sucesivas en medios de cultivo. El microorganismo es capaz de replicarse en el interior del huésped y generar infección clínica o subclínica, que puede transmitirse a otros seres humanos.

 **Inactivadas:** Con microorganismos “muertos”, enteros o en subunidades, a los que se les ha desactivado la capacidad infectiva mediante procesos físicos o químicos. Suelen ser menos reactógenas que las atenuadas, aunque suelen requerir

adyuvantes y la protección que otorga suele ser menos duradera.

b) **Clasificación microbiológica (7):** las vacunas pueden ser bacterianas o víricas según el antígeno que contengan (Tabla 2).

Tabla 2. Clasificación microbiológica de las vacunas	
Bacterianas	☞ Acelulares
	☞ Enteras.
	☞ Polisacáridos capsulares.
	☞ Proteínas de superficies.
	☞ Conjugadas polisacárido + proteínas.
	☞ Toxoides.
Víricas	☞ Enteras.
	☞ Subunidades.

Fuente: elaboración propia a partir de Sociedad Española de Geriátria y Gerontología (7).

c) **Clasificación según la tecnología de fabricación (5,7,8):** según aspectos relacionados con la síntesis de las vacunas.

 **Clásicas:** Cultivadas in vitro.

 **Sintéticas:** Formadas por cadenas de aminoácidos sintetizadas in vitro que imitan a la secuencia primaria de aminoácidos de los microorganismos.

 **Recombinantes:** Obtenidas mediante la clonación de genes del patógeno que codifican las proteínas capaces de generar respuesta inmune en el ser humano. Mediante esta tecnología se

han producido vacunas frente a la hepatitis B, la influenza o el VPH.

d) **Clasificación según su composición.** (7) En función de la cantidad de tipos antigénicos que contengan, las vacunas pueden ser:

 **Monovalentes:** Con un componente antigénico. Ejemplo: vacuna antirrábica.

 **Polivalentes:** Con varias clases de antígenos de una misma especie. Ejemplo: vacuna antineumocócica polisacárida 23-valente.

 **Combinadas:** Con antígenos de distintas especies o microorganismos. Ejemplo: vacuna triple vírica (SRP), con virus atenuados de sarampión, rubéola y parotiditis.

e) **Clasificación sanitaria** (7): según criterios epidemiológicos.

 **Sistemáticas:** Incluidas en los calendarios vacunales recomendados a toda la población por su interés en salud pública. Las vacunas incluidas en este grupo son distintas en cada país.

 **No sistemáticas:** Recomendación individualizada según los antecedentes médicos del paciente, riesgo de exposición y situación epidémica.

En la Tabla 3 y en la Tabla 4, se presentan las vacunas más empleadas en la actualidad, divididas según su clasificación microbiológica y su naturaleza.

Tabla 3. Vacunas víricas según su naturaleza		
	Atenuadas	Inactivadas
Enteras	<ul style="list-style-type: none"> • Fiebre amarilla. • Herpes zóster. • Poliomielitis oral (Sabin). • Rotavirus. • Triple vírica. • Varicela. 	<ul style="list-style-type: none"> • Encefalitis centroeuropea o por garrapatas. • Encefalitis japonesa. • Gripe. • Hepatitis A. • Poliomielitis parenteral (Salk). • Rabia.
Subunidades		<ul style="list-style-type: none"> • Gripe fraccionada o de subunidades. • Hepatitis B. • Virus del papiloma humano.

Fuente: modificado de Sociedad Española de Geriátria y Gerontología (7).

Tabla 4. Vacunas bacterianas según su naturaleza		
	Atenuadas	Inactivadas
☞ Acelulares		<ul style="list-style-type: none"> • Tosferina acelular.
☞ Enteras	<ul style="list-style-type: none"> • Cólera oral. • Fiebre tifoidea oral. • Tuberculosis o BCG (no disponible en España). 	<ul style="list-style-type: none"> • Cólera parenteral. • Fiebre tifoidea parenteral.
☞ Proteínas de superficie		<ul style="list-style-type: none"> • Antimeningocócica B.
☞ Polisacáridos capsulares		<ul style="list-style-type: none"> • Antineumocócica 23 valente. • Antimeningocócica tetravalente.
☞ Conjugadas polisacárido + proteína		<ul style="list-style-type: none"> • <i>Haemophilus influenzae</i> tipo b. • Antimeningocócica C • Antimeningocócica tetravalente. • Antineumocócica 7 valente. • Antineumocócica 10 valente. • Antineumocócica 13 valente.
☞ Toxoides		<ul style="list-style-type: none"> • Difteria. • Tétanos.

Fuente: modificado de Sociedad Española de Geriatria y Gerontología (7).

3. Composición de las vacunas

Las vacunas suelen contener antígenos, adyuvantes, conservantes, estabilizantes, surfactantes y residuos derivados del proceso de fabricación. (9)

 **Antígeno** (9,10): Es el ingrediente activo capaz de inducir inmunidad específica.

 **Adyuvantes** (9): Son un conjunto de sustancias incorporadas a las vacunas con el fin de potenciar la respuesta inmune producida por éstas. Algunas de las sustancias autorizadas para su uso como adyuvantes son las endotoxinas bacterianas, las emulsiones de aceite en agua y las sales de aluminio. La cantidad de aluminio contenida en una vacuna en la que se han utilizado sales de aluminio como adyuvantes es menor que la aportada por la dieta diaria. El uso de adyuvantes hace posible que se genere la inmunidad deseada con menor cantidad de antígeno, el cual es difícil de producir. Por su naturaleza, las vacunas atenuadas no requieren adyuvantes, mientras que la mayoría de las inactivadas sí precisan de estos componentes.

 **Antibióticos** (10): En algunas vacunas se añaden aminoglucósidos o polimixina B con el fin de impedir el crecimiento bacteriano.

 **Conservantes** (9,10): Tienen como objetivo retrasar la degradación de las vacunas. Las vacunas con conservantes basados en mercurio, como el tiomersal, no se comercializan en Europa y EE.UU. desde el año 2002, a excepción de algunas vacunas frente a la influenza.

-  **Estabilizantes** (9): Mantiene juntos los ingredientes de las vacunas y les confieren resistencia frente a condiciones químicas y ambientales adversas. La albúmina, la gelatina, la lactosa y la sacarosa son algunos ejemplos de estabilizantes.
-  **Líquido de suspensión** (10): Puede ser una solución salina, agua destilada o productos derivados de los cultivos de las vacunas.
-  **Surfactantes** (9): Tensioactivos, como el polisorbato 80, que impiden que el resto de los componentes sedimenten o aglutinen, manteniéndolos en suspensión en el líquido.
-  **Residuos** (9): Es inevitable la presencia de restos de sustancias empleadas en el proceso de síntesis de las vacunas, tales como formaldehído, antibióticos, proteína de huevo o gelatina.

4. Características generales de las vacunas

Según L. Salleras (11), las cinco **propiedades** que debe tener una vacuna son: **inmunogenicidad, eficacia protectora, eficiencia, estabilidad y seguridad.**

La **inmunogenicidad** es la capacidad de una vacuna de inducir una respuesta inmunitaria específica frente al antígeno administrado en el sujeto al que se vacuna. (11)

El “valor protector de una vacuna aplicada en condiciones ideales”, esto es, la **eficacia protectora**, depende directamente de la inmunogenicidad. (12) Una vacuna con alta eficacia e inmunogenicidad es capaz de inducir una respuesta inmune suficiente - humoral, celular o ambas - frente al agente infeccioso seleccionado, en el tejido en el que es necesaria esta protección. (11)

No obstante, la protección proporcionada por las vacunas en condiciones reales (**efectividad**) suele diferir de la observada en los ensayos clínicos aleatorizados (eficacia). La efectividad vacunal (EV) puede calcularse en un estudio de casos (vacunados) y controles (no vacunados) mediante la siguiente fórmula: $EV = (1 - OR) \times 100$, donde OR es la Odds Ratio, es decir, la probabilidad de contraer la infección en vacunados, dividido entre la probabilidad de enfermar en las no vacunados. (13) La efectividad también puede valorarse mediante estudios de cohortes. En ese caso, la OR se sustituye en la fórmula por el riesgo relativo. (12)

Una vez calculada la efectividad, pueden realizarse análisis coste-efectividad para determinar la **eficiencia** de la vacunación en grupos poblacionales con distinta incidencia de enfermedad y realizar una adecuada distribución de recursos según los resultados. (11)

Las diferencias entre la efectividad y la eficacia están justificadas por la presencia de barreras de accesibilidad a las vacunas por parte de la población diana, incumplimientos de las pautas vacunales recomendadas, la capacidad de la vacuna para generar memoria inmunológica y factores relacionados con la correcta conservación y administración de las vacunas. (11)

En cuanto a la conservación de las vacunas, hay que tener en cuenta que sus componentes suelen degradarse ante la exposición lumínica y el almacenamiento a temperaturas inadecuadas, habitualmente menores de 0°C o mayores de 8°C. (4,11) Esta degradación de componentes se denomina pérdida de **estabilidad** y se traduce en una disminución de la inmunogenicidad. Además, la estabilidad de las vacunas frente a los factores ambientales varía según su naturaleza. Las vacunas toxoides son las más estables, seguidas de las

inactivadas, siendo las atenuadas las de menor estabilidad. La capacidad antigénica de una vacuna también puede disminuir conforme transcurre el tiempo desde su fabricación, con el límite aceptable marcado por la fecha de caducidad. (11)

Todo ello se debe tener en cuenta para asegurar una correcta **seguridad** vacunal. (14) Ninguna vacuna está exenta de complicaciones y no existen unas características universales para considerar como segura a una vacuna. Se determina el grado de seguridad que se exige a una vacuna en función de la gravedad de la infección frente a la que se vacuna y la morbimortalidad secundaria a la vacuna que la población diana esté dispuesta a aceptar. (11) Los **efectos adversos** de una vacuna pueden deberse a (14):

-  Las características intrínsecas de la vacuna.
-  Defectos en su fabricación, almacenamiento o manipulación.
-  La influencia de los aspectos socioculturales y la patología de base de los pacientes en su respuesta fisiológica y psicológica.

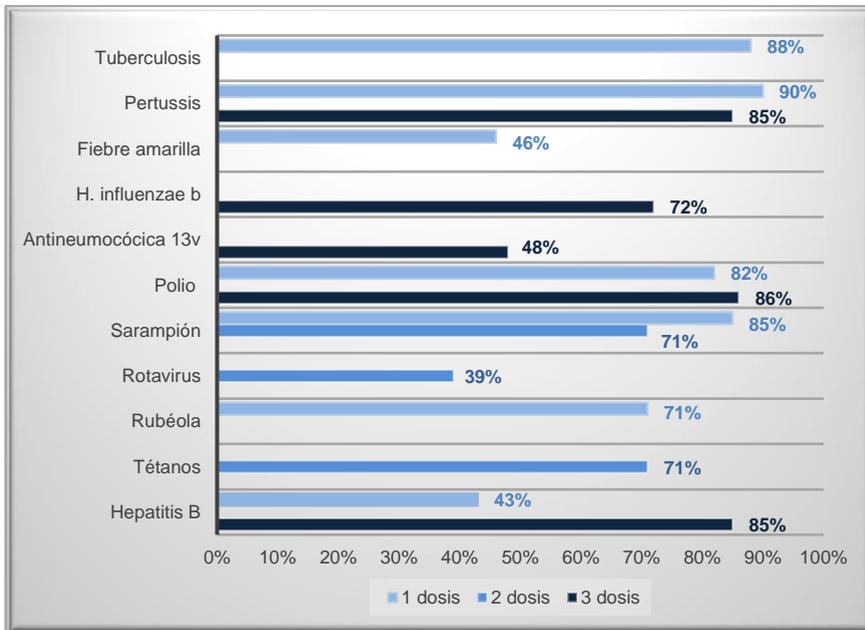
Cabe destacar que sólo hay **cuatro contraindicaciones permanentes** para la vacunación (15):

-  Antecedente de reacción alérgica grave (anafiláctica) a algún componente o a una dosis previa de la vacuna.
-  Debut de encefalopatía no atribuible a otra etiología en los primeros siete días desde la vacunación frente a tosferina.
-  Antecedentes de invaginación intestinal (para la vacuna frente al rotavirus).
-  Inmunodeficiencia combinada grave (para la vacuna frente al rotavirus).

El embarazo y la inmunosupresión son **contraindicaciones temporales** para la administración de vacunas atenuadas. (15)

C. Cobertura Vacunal

La OMS (16) define como **cobertura vacunal** mundial, la “proporción de niños de todo el mundo que reciben las vacunas recomendadas”. En la Figura 1, se presenta las coberturas vacunales mundiales en 2019, estimadas por la OMS y el Fondo de las Naciones Unidas para la Infancia (UNICEF), desglosadas por vacunas y dosis.



Notas: fiebre amarilla – en países de riesgo; polio – la cobertura de la 1ª dosis es de la vacuna inactivada, la cobertura de la 3ª dosis incluye la oral y la inactivada; rotavirus – 2 o 3 dosis, en países donde está incluido en el calendario vacunal; tétanos - dos o más dosis de vacuna toxoide. *Fuente:* elaboración propia a partir de OMS (17).

Figura 1. Coberturas vacunales mundiales, OMS-UNICEF (2019).

La cobertura vacunal depende fundamentalmente de **tres factores**: la **disponibilidad** de las vacunas, la **accesibilidad** a los servicios de inmunización y la **aceptación** de la vacunación por parte de la población. Ante una baja cobertura vacunal en una región, la primera medida que se debe aplicar es la evaluación de los problemas de acceso a las vacunas y corregirlos. A continuación, hay que determinar el grado de aceptación de las vacunas en la comunidad y, si es bajo, aplicar medidas encaminadas a aumentarlo. (18)

No obstante, es imprescindible tener en cuenta que la aceptación de la vacunación por parte de un grupo no excluye la existencia de duda vacunal en éste. Diversas publicaciones han constatado que una parte de los individuos que aceptan la vacunación lo hacen teniendo dudas sobre ella. (19) Por ello, en el siguiente apartado procedemos al abordaje de la **duda vacunal** como barrera frente a la vacunación.

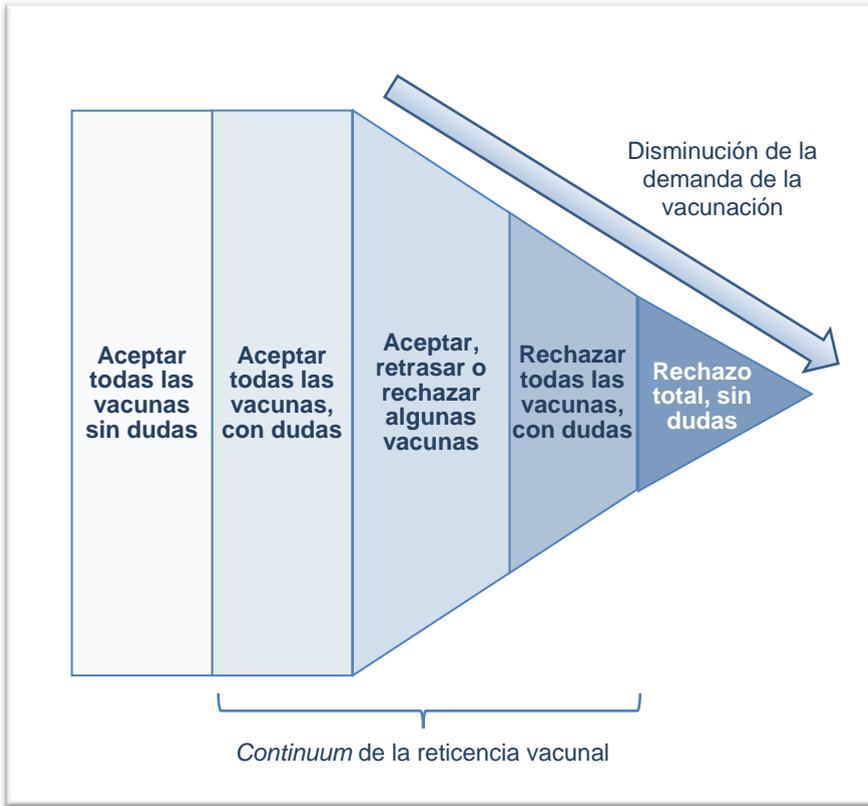
II. RETICENCIA VACUNAL (VACCINE HESITANCY)

A. Concepto de Reticencia Vacunal

El Grupo Consultivo Estratégico de Expertos sobre inmunización de la OMS (*Strategic Advisory Group of Experts on Immunization - SAGE*) (20) define la reticencia vacunal como “el retraso en la aceptación o la negativa a la vacunación a pesar de la disponibilidad de los servicios de vacunación”. Es un fenómeno complejo, multifactorial, variable a través del tiempo y lugar, distinto para cada subtipo de vacuna. Abarca un gran espectro de opiniones entre la aceptación total de las vacunas sin dudas y la negativa total a las vacunas sin dudas (***continuum*** de la reticencia vacunal) (Figura 2). Incluye, por lo tanto, la aceptación o el rechazo de todas las vacunas con dudas, la aceptación de algunas vacunas y rechazo de otras y la elección del retraso de la vacunación.

La reticencia a la vacunación, también conocida como **renuencia vacunal** o ***vaccine hesitancy*** en inglés, engloba cuatro conceptos: indecisión, incertidumbre, retraso y reluctancia (21).

Las vacunas más antiguas y empleadas en los planes de inmunización habituales suelen estar relacionadas con un menor grado de reticencia vacunal. Por el contrario, las vacunas más recientes y las vacunas empleadas en el contexto de campañas de vacunación masiva suelen generar mayor reticencia. (22)



Fuente: modificado de MacDonald et al. (20)

Figura 2. Continuum de la reticencia vacunal.

B. Modelo de Reticencia Vacunal

El grupo SAGE (20) desarrolló el **modelo de las “tres Cs”** para facilitar la comprensión del concepto de reticencia vacunal. Este modelo agrupa los factores que influyen en la reticencia vacunal en tres grupos interrelacionados (Figura 3).



Fuente: modificado a partir de MacDonald et al. (20)

Figura 3. Modelo de reticencia vacunal de las “tres Cs”.

1. La **confianza** se entiende como creencia en:

-  La seguridad y efectividad de las vacunas.
-  Los servicios de salud y los profesionales sanitarios que aconsejan y proveen las vacunas.
-  La administración sanitaria que determina las vacunas que deben administrarse a la población.

2. La **complacencia** hace referencia a la percepción de que las vacunas no son necesarias porque el riesgo de adquirir una enfermedad prevenible mediante vacunación es bajo, sin tener en cuenta que la vacunación es una de las principales razones de disminución de la prevalencia de estas enfermedades.

3. La **conveniencia** engloba:

-  La accesibilidad geográfica y económica de las vacunas.
-  La calidad real y percibida de los servicios sanitarios.
-  Lenguaje y alfabetización en salud.

Aunque una comunicación inadecuada o deficitaria puede aumentar el rechazo a las vacunas, el grupo SAGE (20) decidió no incluir la comunicación en este modelo por considerarla una herramienta para la difusión de mensajes relativos a la vacunación.

C. Determinantes de la Reticencia Vacunal

Tras establecer el modelo de las “tres Cs”, el grupo SAGE (20) diseñó una **matriz** más compleja de los determinantes de la reticencia vacunal. En ella, se agrupan los aspectos que influyen en la reticencia vacunal en tres categorías: aspectos relacionados con el contexto, aspectos individuales y de grupo y aspectos específicos de las vacunas y la vacunación. Los **determinantes** incluidos en esta matriz son los siguientes:

Aspectos relacionados con el contexto:

-  Comunicación y medios.
-  Líderes de opinión, encargados de programas de inmunización, *lobby* anti y pro-vacunación.
-  Influencias históricas.
-  Religión, cultura, género y factores socioeconómicos.
-  Políticas y regulaciones.

-  Barreras geográficas.
-  Percepción de la industria farmacéutica.



Aspectos individuales y de grupo:

-  Experiencia personal, de familiares y/o de miembros de la comunidad con vacunaciones anteriores, incluyendo el dolor.
-  Creencias y actitudes acerca de la salud.
-  Nivel de conocimiento y de conciencia que se tenga en relación con las vacunas.
-  Confianza y experiencia personal en el sistema sanitario y con los proveedores.
-  Riesgo/beneficio (percepción, heurística).
-  Inmunización como “norma social” versus “falta de necesidad y daño”.



Aspectos específicos de las vacunas y la vacunación:

-  Riesgo/beneficio (evidencia epidemiológica y científica).
-  Introducción de nuevas vacunas, formulaciones o recomendaciones para una vacuna existente.
-  Formas de administración.
-  Diseño del programa de vacunación y forma de distribución (p.ej. programa rutinario o campaña de vacunación).
-  Confiabilidad de la fuente de suministro de vacuna y/o equipo de vacunación.
-  Calendario de vacunación.

-  Costos.
-  Fuerza de la evidencia científica de la recomendación y/o base del conocimiento. Actitud de los profesionales de la salud.

D. Impacto de la Reticencia Vacunal

La reticencia vacunal supone un **grave problema de salud pública**. Prueba de ello es el hecho de que la OMS (2) la incluyera en la lista de las **diez principales amenazas para la salud global** en **2019**.

La duda vacunal constituye una de las principales oposiciones en la erradicación de las enfermedades prevenibles mediante vacunación y ha tenido un papel clave en el resurgimiento de éstas en diversos países en la última década. (2,21) Los movimientos contrarios a las vacunas tienen un gran **impacto** sociosanitario y económico, ya que contribuyen a (23):

-  La generación de **brotos** por enfermedades previamente controladas mediante vacunación, con el subsecuente aumento en la morbilidad y mortalidad por estas infecciones.
-  El aumento de los **costes directos e indirectos** por estas enfermedades, incluidos los costes derivados del aumento del ausentismo laboral o escolar de los individuos enfermos.
-  La **pérdida de productividad** y la **retirada del mercado** de fabricantes de vacunas, lo cual se deriva en la menor disponibilidad de vacunas.

Debido a la heterogeneidad entre los distintos países y la escasez de datos recogidos en países fuera de Europa y América, es difícil determinar el impacto de la reticencia vacunal en cada región. (6)

En general, la reticencia vacunal se traduce en una menor **cobertura vacunal** en los subgrupos que expresan esta reticencia. Pero cuando los subgrupos reticentes abarcan un porcentaje pequeño de una población, no hay variación significativa en las coberturas vacunales de esa población ni en la incidencia de las enfermedades prevenibles mediante vacunación, en ausencia de contexto epidémico. (22)

A diferencia de los determinantes de salud, los determinantes de la vacunación no son unidireccionales, sino que pueden tener una influencia distinta en la vacunación según el país. Por ejemplo, un **nivel educativo alto** es un determinante de salud relacionado con un mejor estado de salud global en cualquier región. En cambio, el nivel educativo alto como determinante de reticencia vacunal constituye una barrera para la aceptación de las vacunas en algunos países, como China, Líbano, Israel, Bangladesh o Estados Unidos (EE. UU.), mientras que actúa como promotor de la vacunación en otros países, como Grecia, Países Bajos, Nigeria o Pakistán.

En el otro extremo se encuentra el **nivel educativo bajo**. Aunque constituye una barrera para la vacunación por la falta de conocimientos que conlleva, no siempre está asociada con el rechazo a la vacunación, como se ha observado en India. En cambio, en Nigeria y Kirguistán, el nivel educativo bajo sí se ha relacionado con mayor tendencia a la adopción de actitudes contrarias a la vacunación. (22)

E. Reticencia y Demanda Vacunal

La **OMS** y **UNICEF** (24) definen la **demanda vacunal** como las acciones de los individuos y las comunidades para buscar, apoyar y / o abogar por las vacunas y las inmunizaciones. Esta definición implica que el término demanda hace referencia a comportamientos, más que actitudes, ya se fundamenta en tres verbos:

-  **Buscar:** relativo al comportamiento de los individuos.
-  **Apoyar:** expresar una norma social.
-  **Abogar:** defender y reclamar los derechos e influir en los responsables de la toma de decisiones.

Por ello, no es posible entender la demanda fuera del sistema proveedor de vacunas, pues por definición, la demanda requiere de la interacción entre los individuos de la sociedad y el sistema. Así pues, la demanda es un fenómeno complejo y dinámico, variable según el contexto (tiempo, lugar, vacuna, individuo, sociedad y sistema sanitario). (24)

La reticencia vacunal es uno de los factores que disminuye la demanda de las vacunas. Por lo tanto, una alta reticencia vacunal conlleva una baja demanda de la vacunación. (20) Por ello, el **Plan de Acción Global de Vacunación 2011 - 2020** de la **OMS** (25) recomienda a los países tomar medidas para contrarrestar la reticencia vacunal como una de las estrategias destinadas a aumentar la demanda de vacunas.

No obstante, una baja reticencia vacunal no asegura una alta demanda de vacunas. Esto se debe a que en la demanda de las vacunas influyen otros aspectos como: el contexto socioeconómico,

las características de cada comunidad y las estrategias de promoción de cada vacuna. (20)

Por otra parte, algunos autores (24) han propuesto que la demanda también puede influir en la reticencia vacunal. De esta manera, una alta demanda de vacunas por parte de la población podría contribuir a que la sociedad de su entorno tenga menos dudas acerca del riesgo-beneficio de las vacunas y, por tanto, exprese en menor medida pensamientos reticentes.

F. Reticencia Vacunal a lo Largo de la Historia

1. Reticencia vacunal: siglo XIX

En la era de la prevacunología (siglos XII–XVIII), el método de inmunización frente a la viruela era la inoculación o **variolización**. Consistía en la inhalación voluntaria de polvo de costras de viruela por parte de una persona sana o en la introducción de pus de un enfermo en una incisión realizada previamente a la persona sana. Las personas variolizadas desarrollaban una respuesta inmune específica frente a la viruela, pero en numerosas ocasiones, también manifestaban una forma grave de viruela, no exenta de mortalidad. Además, el riesgo de transmisión de otras enfermedades a través de la inoculación era alto. (26)

Por ello, Edward Jenner (1749–1823) propuso sustituir la inoculación de viruela humana por la inmunización con viruela vacuna (*vaccina*) en 1796. (26) Así se produjo el descubrimiento de la **primera vacuna**, aunque el término “vacunación” no lo acuñó Pasteur hasta 1880. (26,27) A pesar de que Jenner demostró la seguridad y eficacia

de la vacunación frente a la viruela, ésta encontró opositores en los distintos estratos de la sociedad desde que se inició su administración a la población a principios del siglo XIX. (28)

En vista a la gran mortalidad por las epidemias de viruela en **Reino Unido**, el gobierno promulgó la **primera ley de vacunación obligatoria** en **1853**, que exigía la vacunación de todos los niños entre 0–3 meses de edad, con multas o penas de cárcel para los padres incumplidores de la ley. (29,30)

Inmediatamente, la población de distintas regiones se manifestó en contra de la ley y un grupo de personas reticentes fundó la **Liga Anti-Vacunación** en Londres. A pesar de ello, en 1867, el gobierno extendió la obligatoriedad de vacunación hasta los catorce años y se introdujeron penas acumulativas. Esto agravó la preocupación del pueblo acerca de la restricción de las libertades civiles y derivó en numerosas publicaciones que expresaban rechazo a la vacunación a lo largo de las dos décadas siguientes. (29)

El centro de la resistencia se estableció en **Leicester**, bajo el mando de líderes religiosos y políticos locales. Allí crearon la **Liga Anti-Vacunación Obligatoria** en **1869**, que estableció un sistema de desinfección y de cuarentena de los casos sospechosos de sarampión para evitar epidemias en el pueblo, con predominio de población no vacunada. (30)

En **1885**, se convocó en este pueblo una manifestación contra la vacunación que reunió entre 80 000 y 100 000 personas. Tuvo tal repercusión, que los miembros del Parlamento debatieron sobre este evento en la Cámara de los Comunes y el gobierno británico decidió

formar una comisión para la investigación de la conveniencia de la vacunación.

Tras nueve años, en **1896**, la comisión concluyó que la vacuna frente a la viruela ofrecía protección frente a la enfermedad, pero recomendó retirar su obligatoriedad. A raíz de estas conclusiones, en **1898**, el gobierno eliminó las penas acumulativas y permitió a los padres reticentes la obtención de certificados de excepción vacunal. Esta fue la primera vez que en la ley inglesa se empleó el concepto “objeción de conciencia”. (30)

Simultáneamente a estos hechos en Reino Unido, a lo largo del siglo XIX, también surgieron grupos reticentes en el resto de Europa, especialmente en Francia y en EE. UU. (29,31) En **EE. UU.**, al igual que en Reino Unido, el conflicto estaba basado en la inconformidad por la obligatoriedad de la vacunación. (29) En cambio, en el **resto** de los **países europeos** predominaban las dudas acerca de la seguridad de la vacunación y existía una fuerte oposición liderada por el clero, que defendía que la vacunación interfería con la voluntad de Dios. (31)

2. Reticencia vacunal: siglo XX

En **1910**, el médico canadiense Sir William Osler expresó su incomprensión por la negatividad a la vacunación. Se ofreció a acompañar en la siguiente epidemia de viruela a diez personas vacunadas y a diez no vacunadas (tres del Parlamento, tres médicos y cuatro propagandistas antivacunas) y se comprometió a cuidar a los que enfermaran de este grupo y a organizar el funeral de los que murieran. (28)

Conforme fue transcurriendo el **siglo XX**, se produjo un auge en el desarrollo científico de la vacunología, que derivó en el descubrimiento de nuevas vacunas y en el perfeccionamiento de los métodos de manufactura. Por otra parte, aumentó el nivel educativo de la población y se evidenció una mejoría en la concienciación pública sobre la peligrosidad de las enfermedades prevenibles mediante vacunación. Todo ello se produjo en una población con aumento de la natalidad que deseaba proteger a los niños susceptibles frente a las infecciones. (28)

En consecuencia, entre **1940 y 1970**, se evidenció un repunte en la **aceptación vacunal**, junto con una disminución significativa de la incidencia y mortalidad por enfermedades inmunoprevenibles. Pero esto a su vez generó que la población volviera a percibir que las enfermedades prevenibles mediante vacunación no eran una amenaza para su salud.

A partir de **1970**, empezaron a aumentar progresivamente las expresiones de preocupación sobre la seguridad de las vacunas y sobre la conveniencia de añadir nuevas vacunas al calendario de inmunizaciones. (28)

En **1974**, **Kulenkampff et al.** (32) publicaron un artículo en el que sugerían una relación causal entre el componente ***pertussis*** de la vacuna frente a difteria, tétanos y tosferina y el desarrollo de **complicaciones neurológicas** tras la vacunación, tales como convulsiones, espasmos, irritabilidad, vómitos y disminución del nivel conciencia.

El debate continuó en **1982**, tras la emisión de un programa de televisión estadounidense titulado “**DPT: Vaccine Roulette**”. (28) Ambos hechos tuvieron como consecuencia una disminución de la tasa de vacunación frente a tosferina en diversos países del mundo e, incluso, produjo la eliminación temporal de la vacuna en algunos calendarios vacunales, con el consecuente aumento en el número de brotes por esta enfermedad entre 1970 y 1990. (28)

España tampoco estuvo exenta de la duda vacunal. En **1989**, en Barcelona, un grupo de padres y madres y profesionales de la salud creó la primera asociación española de personas reticentes a las vacunas, llamada “**Liga para la Libertad de Vacunación**”. (33)

3. Reticencia vacunal: siglo XXI

La renuencia vacunal continúa siendo una barrera para la vacunación en el siglo XXI. Uno de los eventos que más ha contribuido al aumento de la reticencia vacunal en las dos primeras décadas del siglo XXI es el artículo *Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children*, escrito por **Wakefield et al.** (34) y publicado en la revista *The Lancet* en **1998**.

En este trabajo, los autores describían doce casos que presuntamente habían presentado anomalías gastrointestinales y **problemas neurocognitivos**, incluyendo regresión del desarrollo y síndromes del espectro autista, después de haber recibido la **vacuna SRP**. Posteriormente, publicaron una adenda donde informaron que habían observado estas anomalías en treinta y nueve casos más, de cuarenta pacientes estudiados. Concluían que eran necesarios más

estudios para esclarecer si existía una asociación causal entre la inmunización y el desarrollo del síndrome anteriormente descrito.

En 2004, el periodista Brian Deer (35) reveló que había disparidades entre los datos publicados por Wakefield et al. (34) en *The Lancet* en 1998 y la historia clínica de los casos. También sugirió que Wakefield falsificó los datos con el fin de obtener ganancias económicas.

A partir de este descubrimiento, el Consejo Médico General de Reino Unido inició su propia investigación. En enero de 2010, dictaminó que Wakefield y dos de los coautores del artículo eran culpables de numerosos delitos, tales como realizar el estudio fuera de las condiciones aprobadas por el Comité de Ética, someter a los niños a procedimientos invasivos no indicados (36) y recibir financiación por parte de representantes legales de padres que creían que la vacuna SRP había causado daño a sus hijos (37).

Acto seguido, el artículo fue retractado por *The Lancet* en febrero de 2010 (36) y al gastroenterólogo Wakefield le retiraron su licencia médica en Reino Unido en mayo de ese mismo año.

A pesar de la retractación y de que varios estudios (38–44) han demostrado la ausencia de relación causal entre la inmunización con la vacuna SRP y el autismo, se ha evidenciado una disminución en la inmunización con la vacuna SRP en diversos países desde la publicación del artículo.

En consecuencia, se ha producido un aumento en la frecuencia y en la magnitud de los brotes epidémicos de sarampión, (45) los últimos de ellos en 2019, en países donde previamente la incidencia era baja, tales como EE.UU., Ucrania, Francia, Italia, Brasil, Venezuela, Filipinas o Japón. (46)

Otros ejemplos de reticencia vacunal en los últimos 20 años y sus consecuencias son:



Reemergencia de la **poliomielitis** en **varios países** entre **2002 y 2004** como consecuencia de las recomendaciones de no vacunar frente a esta enfermedad dadas por líderes religiosos musulmanes y políticos de la **India y Nigeria**.

En un contexto de conflictos internacionales entre los países musulmanes y EE. UU., esto líderes difundieron el mensaje de que la vacuna frente a la poliomielitis oral era un medio de ataque ideado por EE.UU. para propagar el virus de la inmunodeficiencia humana (**VIH**) y causar infertilidad en la población musulmana. (47)



Baja aceptación de la vacuna contra el **VPH** por parte de la población española tras la difusión de la noticia de que dos adolescentes de la Comunidad Valenciana habían sido diagnosticadas de estatus epiléptico después la administración de esta vacuna en 2009. (48)

Aunque la investigación por parte de la Agencia Española del Medicamento y Productos Sanitarios (AEMPS) y la Agencia Europea del Medicamento (EMA) concluyó que no había relación causal entre la vacunación contra el VPH y los síntomas, la amplia cobertura mediática que se le dio a los casos antes de conocerse el dictamen definitivo generó un alto grado de duda sobre la seguridad de la vacuna entre los miembros de la comunidad. (48)

-  Baja cobertura vacunal frente a la influenza A(**H1N1**)pdm2009 en la pandemia de 2009 debido a los mensajes de desconfianza en la seguridad de esta vacuna propagados por distintos grupos reticentes. (49)
-  Notificación de un caso de **difteria** en **España** en **2015**, después de treinta años sin casos autóctonos. Se trataba de un niño de seis años no vacunado por decisión de los padres, que falleció a consecuencia de la infección. (49)

Las razones y las consecuencias de esta reticencia continúan siendo similares a las observadas en 1800 y 1900, aunque ahora los mensajes se difunden a través de **medios de comunicación** y **redes sociales**. (28)

III. INFORMACIÓN Y DESINFORMACIÓN SANITARIA EN LA ERA DIGITAL

A. Internet y la Red Informática Mundial

La Real Academia Española (RAE) (50) define **información** como “comunicación o adquisición de conocimientos que permite ampliar o precisar los que se poseen sobre una materia determinada”. Otra acepción aceptada para este término es “conocimientos comunicados o adquiridos mediante una información”.

Por el contrario, la **desinformación** es la comunicación de información errónea por desconocimiento (*misinformation*) o con intención de manipular, causar daño u obtener beneficios personales, políticos o económicos (*disinformation*). (51)

A lo largo de la **historia**, la información se ha transmitido por distintos canales. Inicialmente, predominaba la comunicación **oral**. Con la invención de la escritura, se inició la difusión de datos a través de los **manuscritos** y, más tarde, a través de los **impresos mecanografiados** y las **publicaciones digitales**. (52)

En las publicaciones digitales, la información está presentada en forma de **hipertextos**, esto es, un “conjunto estructurado de textos, gráficos, etc., unidos entre sí por enlaces y conexiones lógicas”. (50,53) El acceso a estos hipertextos se realiza por medio de Internet. (53) Jošanov et al. (54) definen **Internet** como “un sistema global de redes de ordenadores interconectados, donde la *web* es el servicio más importante”.

La **red informática mundial** (en inglés, *World Wide Web*), abreviado como red (*web*) es un sistema de documentos en hipertexto que

pueden ser consultados vía Internet a través de aplicaciones que reciben el nombre de navegadores. El científico inglés Sir Tim Berners-Lee escribió las dos primeras propuestas de sistema *web* entre 1989 y 1990. Finalizó el diseño de la primera *web* en diciembre de 1990 y lo publicó en agosto de 1991. (53)

Al principio de la **era digital**, los usuarios sólo podían leer la información disponible en la red, pero no modificarla. La evolución tecnológica ha permitido que la red evolucione hacia un entorno más dinámico, bautizado como **Web 2.0**, en el que los usuarios pueden compartir su propio contenido e interactuar con otros usuarios. (52) La participación de los usuarios en el proceso de comunicación les empodera, ya que les permite tomar decisiones informadas, a la vez que adquieren mayor confianza en su propio criterio y disminuye su ansiedad. (55)

La comunicación de datos relativos al **ámbito de la salud** a través de la *Web 2.0* ha generado un modelo en el que cualquier persona con acceso a Internet puede producir, difundir y consumir información sobre salud y enfermedad, sin intermediarios que verifiquen previamente su veracidad científica. (52)

Respecto a la época predigital, Internet facilita y amplifica la difusión de multitud de datos de calidad y fiabilidad **heterogénea**, desde rumores y mitos hasta afirmaciones avaladas por la ciencia. (56) De esta manera, Internet funciona tanto como medio de **promoción de la salud** como vía de **desinformación sanitaria**. (57)

Pese a que los seres humanos continuamos obteniendo información sobre salud a partir de nuestro entorno social, los centros sanitarios, la radio, la televisión y los documentos en papel, cada vez hay más

personas que consultan la red como **primera fuente** de información sobre salud, especialmente la población más joven, con mayores ingresos y mayor nivel educativo. (58)

Según un estudio realizado por el grupo de investigación **TNS** (59) para la **Comisión Europea** en 2014, el 54% de los encuestados había consultado información sobre una enfermedad en Internet en el último año y el 55% había buscado en la red información general sobre cuidados en salud y estilo de vida. De ellos, el 70% afirmaba que la información que había encontrado era **útil**, el 60% opinaba que la información le había resultado fácil de **comprender** y fácil de **encontrar** y el 50% identificaba la fuente consultada como **fiable**. En **España**, las cifras eran ligeramente mayores que la media: 87% para la utilidad de los datos de la red, 91% para la comprensión y la facilidad de encontrar la información y 74% para la fiabilidad de la fuente.

En el caso de la **vacunación**, el uso de Internet para resolver dudas sobre la misma está ampliamente extendido. El porcentaje de padres que consulta Internet como fuente principal de información sobre vacunas es variable según el país. Por ejemplo, la Asociación Española de Vacunología (60) indica que el 58% de los padres encuestados recurre en primer lugar a Internet para informarse sobre vacunas, mientras que en Malasia (61) la cifra asciende al 80%.

Estas cifras son relevantes debido a que la información disponible en la *web* puede influir en la decisión de aceptar, posponer o rechazar la vacunación. (52,62) Sin embargo, es importante tener en cuenta que los padres confían más en la información sobre vacunas obtenida a partir de los pediatras de sus hijos que en los datos divulgados en la red. (63,64)

El **impacto** sobre la salud de un individuo que tiene la información expuesta en la red individuo es variable y depende de (65,66):

- 🖱 Las barreras económicas, educativas, geográficas y demográficas que condicionan el **acceso** a Internet.
- 🖱 Factores socioculturales, el nivel educativo y el nivel cognitivo que delimitan la capacidad individual de distinguir, seleccionar y confiar en las **fuentes** fiables.
- 🖱 La **inteligibilidad** de la información por parte de los usuarios no profesionales. Los usuarios suelen seleccionar la información que les resulta más fácil de comprender, aunque identifiquen que la fuente tiene una fiabilidad baja.

B. Redes Sociales

La *Web 2.0* permite la creación de **plataformas** de comunicación social con contenido producido por los consumidores como los blogs, los foros o las redes sociales. (67) Estas plataformas sociales son una importante fuente de información en salud. En **Europa**, el 47-48% de los usuarios que utilizan Internet como fuente de información sobre salud consulta sitios *webs* específicos, incluyendo blogs o foros, mientras que el 16 – 23 % recurren a las redes sociales. (59)

Boyd y Ellison (68) definen los **sitios web de redes sociales** (en adelante, redes sociales) como “servicios basados en la red que permiten a las personas”:

- 🖱 “Construir un **perfil** público o semipúblico dentro de un sistema acotado”.

- 🖱️ “Articular una **lista** de otros usuarios con quienes comparten una conexión”.
- 🖱️ “Ver y recorrer su lista de **conexiones** y las realizadas por otros dentro del sistema”.

El uso de las redes sociales no presenta diferencias significativas según el sexo, pero sí según los **grupos de edad**. El 80% de los usuarios de Internet con edades comprendidas entre 14 y 34 años hacen uso de las redes sociales, mientras que esta cifra desciende al 20% entre los usuarios mayores de 65 años. (67)

Los mensajes intercambiados en las redes sociales tienen niveles de **privacidad** variables. Según la plataforma y la configuración seleccionada por el usuario, las personas pueden compartir los mensajes sólo con uno o varios contactos previamente seleccionados, con todos los usuarios incluidos en la lista de conexiones, con cualquier usuario inscrito en la plataforma o con cualquier persona que consulte el perfil sin necesidad de inscripción previa. (67)

Las redes sociales actúan como plataformas de información y desinformación sobre la salud, incluidas las vacunas. La definición de **desinformación** aplicada a las vacunas hace referencia a la difusión de afirmaciones “que han sido investigadas y rechazadas con suficiente evidencia en la literatura revisada por pares”. (57)

IV. MÉTODOS COMPUTACIONALES PARA EL ESTUDIO DE LA DESINFORMACIÓN

Dada la gran cantidad de datos disponibles en Internet y las redes sociales, los profesionales pueden recurrir a **algoritmos automatizados** para el estudio y abordaje de los tres **componentes** de la **desinformación**: fuente, consumidores y contenido. (69)

Fiabilidad de la fuente de información:

Los responsables de la difusión de la desinformación no son sólo personas reticentes, sino que también participan **bots** (usuarios automatizados). (62) Al igual que las personas, los **bots** pueden seguir a otros usuarios, marcar mensajes como favoritos, participar en conversaciones y compartir contenido tanto a favor como en contra de las vacunas. (69,70)

Aunque en algunas ocasiones el uso supervisado de los **bots** puede ser **útil** para disminuir el trabajo manual necesario para transmitir noticias y mensajes positivos, en numerosas ocasiones los **bots amplifican** la difusión de datos no verificados y rumores creados con la intención de manipular el discurso público y cambiar las tendencias. Por ello, el estudio de la fiabilidad de la fuente de la información mediante métodos automatizados se ha centrado en la diferenciación de los **bots** de los usuarios humanos. (69,70)

Los **métodos automatizados** utilizados para la caracterización del **tipo de usuario** que participa en una red social combinan el análisis de varios elementos, como el contenido de los mensajes, el lenguaje, la ratio del número de tuits por unidad de tiempo, el patrón de comportamiento del usuario o los datos de su perfil. (69)

Confianza de los consumidores en el mensaje:

El impacto de la desinformación depende en gran medida de la **confianza** que los usuarios depositen en el mensaje.

Algunos autores (71–74) han desarrollado herramientas que ayudan a los consumidores de datos a **detectar rumores** en Twitter según el análisis del contenido de los mensajes, las respuestas que generan o las características de los autores. Adicionalmente, algunos de estos sistemas permiten a los usuarios realizar **seguimiento** a tiempo real de la propagación de los rumores. (73,74)

No obstante, es necesario esclarecer la **utilidad** de estas herramientas. Lazer et al. (75) sugieren un posible efecto contrario al que se pretende, ya que estos recursos digitales incrementan la exposición de los individuos a la desinformación y éstos tienden a recordar mejor la información al familiarizarse con ella, independientemente del contexto.

Comunicación entre la fuente y los consumidores:

Diversas técnicas de análisis de conexiones sociales permiten determinar quiénes participan en la difusión de la información y cómo se relacionan los usuarios entre sí. (76,77)

Análisis del contenido:

Permite **identificar** cuáles son las dudas, los rumores y los mitos circulantes. El análisis del contenido de los mensajes de las redes sociales es posible gracias a la **minería de textos**.

A. Minería de Textos

La **minería de textos** engloba todas las técnicas utilizadas para analizar textos y descubrir nueva información a partir de ellos (78):

- 👉 **Limpieza de datos:** métodos básicos de preprocesamiento de texto.
- 👉 **Análisis de nivel de palabras:** averiguar la frecuencia de cada palabra en un texto.
- 👉 **Análisis de asociación de datos:** determinación de las relaciones entre las palabras.
- 👉 **Minería de datos avanzada:** clasificación de textos, agrupación de textos según sus similitudes, modelado de temas, análisis de opiniones y extracción de información.

Gran parte de estas técnicas están basadas en el **procesamiento del lenguaje natural (PLN)**.

B. Procesamiento del Lenguaje Natural

El **lenguaje natural** es el sistema de comunicación oral o escrito que empleamos los humanos de forma habitual entre nosotros. (50,79) El **PLN** es la rama de conocimiento de la inteligencia artificial encargada de la investigación y desarrollo de mecanismos computacionales capaces de comprender y realizar una comunicación efectiva por medio de lenguajes naturales.

Dale et al. (80) desglosan de manera pedagógica el PLN en cinco **fases** consecutivas:

- 📌 **Segmentación del texto (*Tokenization*):** delimitación de palabras del texto original. (80)
- 📌 **Análisis léxico o morfológico:** extracción de morfemas (unidades mínimas de significado). (50,79)
- 📌 **Análisis sintáctico:** definir la estructura gramatical de las frases. (80)
- 📌 **Análisis semántico:** estudio del significado literal, independiente de contexto. (79,80)
- 📌 **Análisis pragmático:** determinación del significado según el contenido del texto, es decir, en función de las frases precedentes. (79,80)

Las principales **aplicaciones** del PLN son:

- 📌 Desarrollo de asistentes virtuales y *chatbots* capaces de mantener conversaciones orales o escritas con personas en lenguaje natural. (81)
- 📌 Reconocimiento y síntesis de voz. (82)
- 📌 Traducción automática de voz o texto. (82,83)
- 📌 Revisión automática de la ortografía y gramática de los textos. (84)
- 📌 Extracción de información de textos. (85)
- 📌 Resumen de textos. (86)
- 📌 Análisis de textos: modelado de tópicos (87) y análisis de la opinión (88).

C. Minería de Opiniones / Análisis de Sentimientos

Los mensajes en lenguaje natural pueden incluir valoraciones subjetivas que una persona realiza con respecto a algo o alguien (opiniones) y pueden reflejar estados afectivos del ánimo (sentimientos). (50)

El uso de técnicas de **PLN** para extraer y determinar las opiniones o sentimientos expresados en un texto se denomina **análisis de sentimientos (AS)**. (89) Por tanto, el AS engloba la **minería de opinión**, por lo que a menudo se utilizan los dos términos como sinónimos. (89)

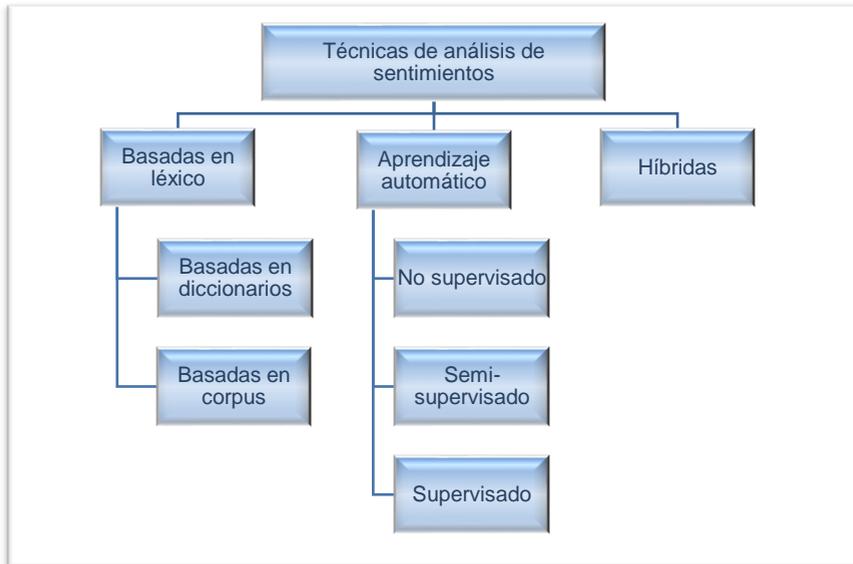
Mediante el AS, podemos catalogar la opinión de los mensajes según diversos criterios como la **polaridad** de la opinión (positiva, negativa o neutra) (90), el tipo de sentimiento (felicidad, tristeza, miedo, aburrimiento...) (91) o el grado de acuerdo/satisfacción según una escala de valoración numérica (92).

El AS tiene **aplicaciones** en diversas ciencias, tales como la economía, la ingeniería de software o la medicina. (93)

Distinguimos tres **niveles de AS** según su objetivo (94):

- 🖱 **AS a nivel del documento:** Orientado a averiguar la opinión global de un documento. Sólo aporta resultados válidos en documentos que tratan sobre un único tema.
- 🖱 **AS a nivel de la frase:** Indica el tipo de opinión de una frase.
- 🖱 **AS basado en aspectos o en características:** Clasifica la opinión e indica frente a qué entidad se ha emitido esta opinión.

Las **técnicas de AS** se clasifican en tres grupos (Figura 4): técnicas basadas en el léxico, técnicas basadas en el aprendizaje automático (*machine learning*) y técnicas híbridas. (95)



Fuente: modificado de Chauhan et al. (95)

Figura 4. Técnicas de análisis de sentimientos.

1. Técnicas basadas en el léxico

Son técnicas muy utilizadas para el AS por su sencillez y eficiencia. (96) En primer lugar, los investigadores elaboran una lista de palabras relevantes para la clasificación, determinan manualmente la polaridad de cada palabra y les asignan una puntuación indicativa de la polaridad. (96,97) Para la selección de palabras de interés, los investigadores pueden utilizar como referencia conjuntos de texto sobre el tema a estudio (aproximación basada en **corpus**).

Como alternativa, pueden completar el inventario de palabras mediante la búsqueda de sinónimos y antónimos en diccionarios (aproximación basada en **diccionarios**). (98)

Posteriormente, los algoritmos asignan una puntuación a cada unidad del texto a estudio según el catálogo elaborado previamente. Por último, calculan la polaridad predominante en el texto a partir de las puntuaciones de sus componentes. (96,97)

2. Técnicas basadas en el aprendizaje automático

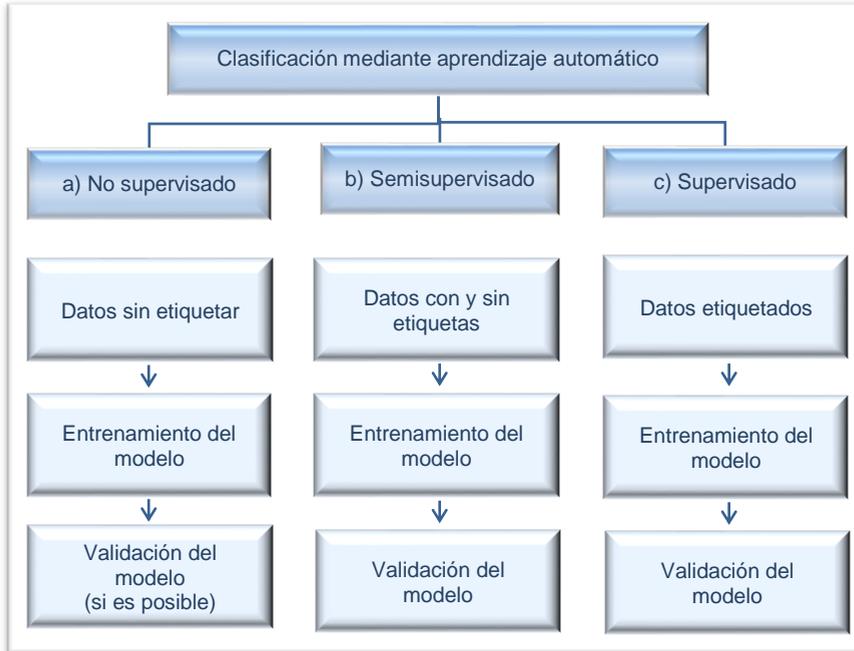
El aprendizaje automatizado es la disciplina de las ciencias de la computación encargada de generar sistemas capaces de enlazar y modificar datos, reconocer patrones, tomar decisiones y ajustar su comportamiento al exponerse a nueva información. (95,99)

Los algoritmos de *machine learning* denominados **clasificadores** entrenan modelos para la identificación automatizada de patrones entre los datos y la posterior clasificación de información según estos patrones.

El aprendizaje automático puede realizarse mediante técnicas no supervisadas, supervisadas o semisupervisadas (Figura 5).

a) Técnicas de aprendizaje no supervisado

Los modelos de aprendizaje no supervisado agrupan automáticamente los datos según sus similitudes, sin instrucciones de los investigadores sobre cómo deben etiquetar ni cuáles son las categorías que deben encontrar en el texto. (100) La función principal de los modelos no supervisados es la identificación de nuevos patrones. (99)



Fuente: modificado de Zhao et al. (101).

Figura 5. Técnicas de aprendizaje automático.

b) Técnicas de aprendizaje semisupervisado

Utilizan datos etiquetados y no etiquetados para diseñar el modelo de clasificación más apropiado. (102)

c) Técnicas de aprendizaje supervisado

Los métodos de clasificación basados en aprendizaje supervisado dividen los datos tomando como referencia conjuntos de datos a los que el investigador o la investigadora ha asignado una categoría previamente. El aprendizaje supervisado es útil para predecir resultados sobre un grupo de categorías preestablecidas. (99)

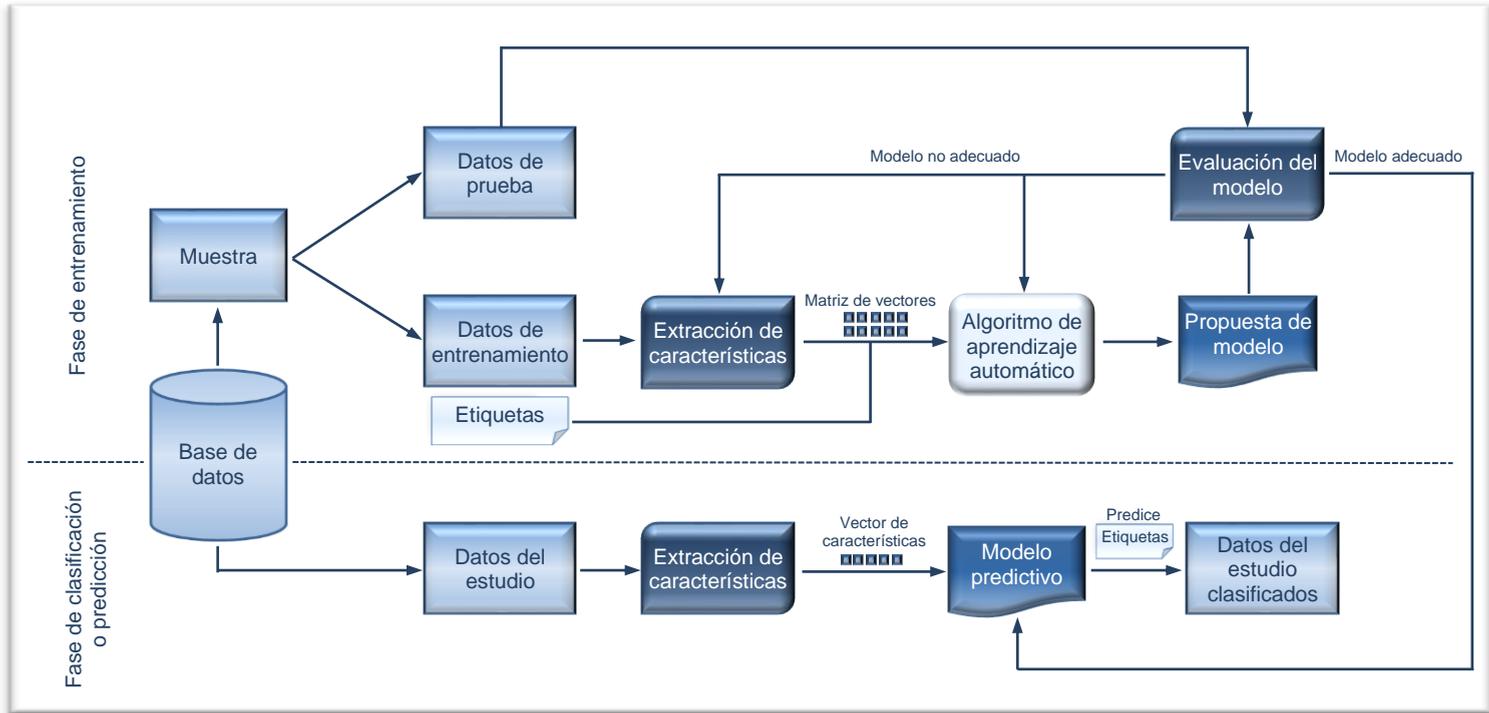
La Figura 6 representa el **diagrama de flujo** simplificado de la clasificación mediante aprendizaje supervisado. En primer lugar, el investigador o la investigadora obtiene una muestra representativa de la base de datos del estudio. El tamaño muestral depende de: el tamaño de la base de datos, el número de categorías en las que se quiere dividir los datos y la cantidad de atributos que caracterizan a cada categoría. (99)

A continuación, divide la muestra en dos grupos de iguales características (datos de prueba y datos de entrenamiento), usualmente con una distribución 80% - 20% respectivamente (99) y etiqueta manual o automáticamente los datos de entrenamiento (103).

El algoritmo de aprendizaje automático utiliza los datos de entrenamiento etiquetados para detectar las características comunes entre los datos de una misma categoría (103) y entrena el modelo de clasificación para la identificación de estos patrones (97).

Acto seguido, el investigador o la investigadora evalúa el rendimiento del modelo en los datos de prueba (102) y ajusta sus parámetros según los errores detectados. (103)

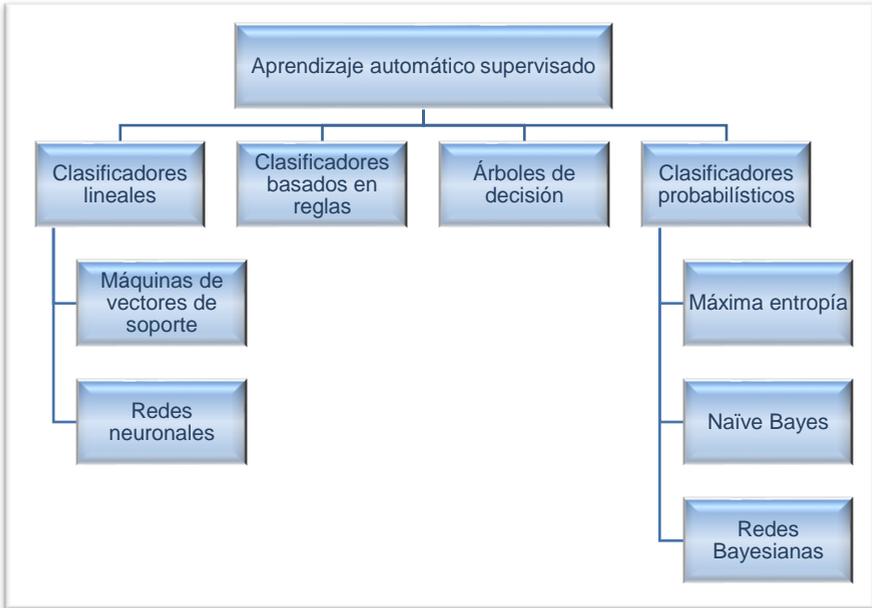
Estos pasos se repiten consecutivamente hasta que el clasificador proporcione la precisión deseada. Alcanzada esta meta, el modelo puede clasificar por sí mismo nuevos datos sin intervención humana. (104)



Fuente: modificado de Nguyen et al. (104)

Figura 6. Diagrama de flujo del aprendizaje supervisado.

El AS mediante aprendizaje supervisado puede realizarse mediante cuatro tipos de **enfoques** (Figura 7).

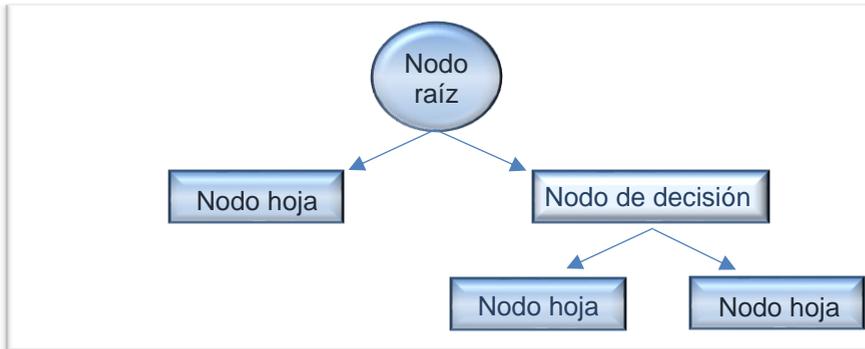


Fuente: modificado de Chauhan et al. (95)

Figura 7. Enfoques de aprendizaje automático supervisado.

1) Árboles de decisión: Funciones no lineales, en forma de árbol. Están formados por nodos de decisión (criterios de clasificación), conectados jerárquicamente mediante ramas que indican la decisión tomada. El algoritmo inicia la clasificación del texto desde el nodo raíz y va avanzando entre los nodos de decisión a través de las ramas hasta alcanzar un nodo hoja, que indica la categoría a la que pertenece el texto (Figura 8). (95)

Los clasificadores que utilizan conjuntos de árboles de decisión se denominan bosques aleatorios (*random forest*). (105)



Fuente: modificado de Syachrani et al. (106)

Figura 8. Árbol de decisión.

2) Clasificadores basados en reglas: Establecen una serie de reglas condicionales (en formato si - entonces) y las analizan individualmente. Si se cumple una condición, entonces clasifican los datos en una categoría. Si no se cumple, estudian otra regla. Estos dos pasos se repiten consecutivamente hasta obtener la clasificación final de todos los datos (Figura 9). (107)

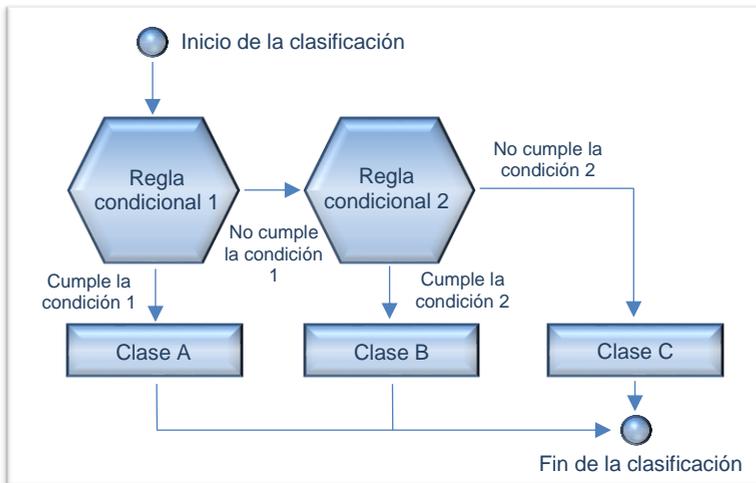


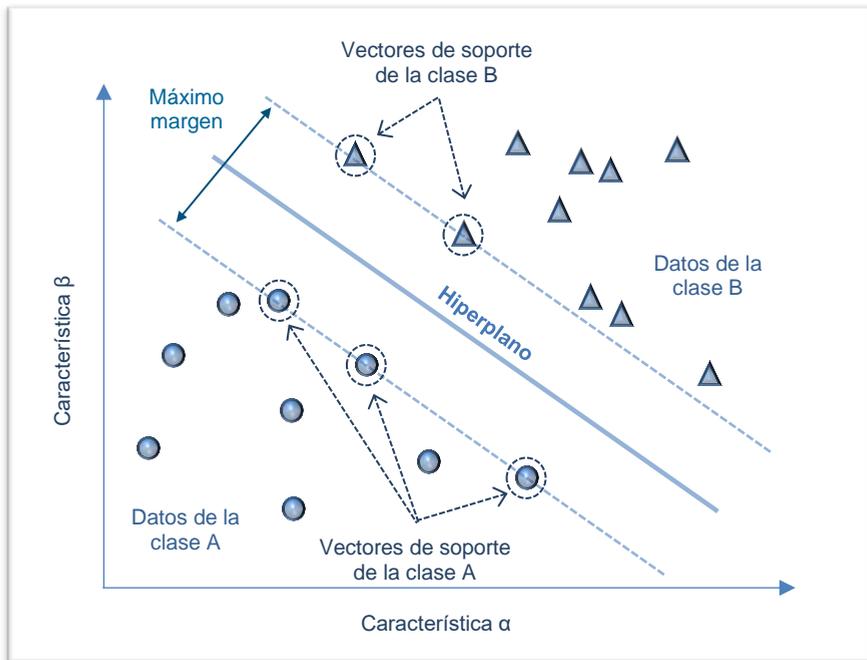
Figura 9. Diagrama de flujo de la clasificación mediante reglas.

3) Clasificadores lineales:

Se subdividen en dos tipos de técnicas: máquinas de soporte vectorial (MSV) y redes neuronales. (95)

MSV

Las **MSV** distribuyen los datos en un espacio n - dimensional y utilizan matrices de puntos de datos para determinar el plano que divide los datos en varias categorías con la máxima separación entre las clases y el mínimo error de clasificación (Figura 10). (108) Es la técnica de aprendizaje supervisado más precisa. (95)

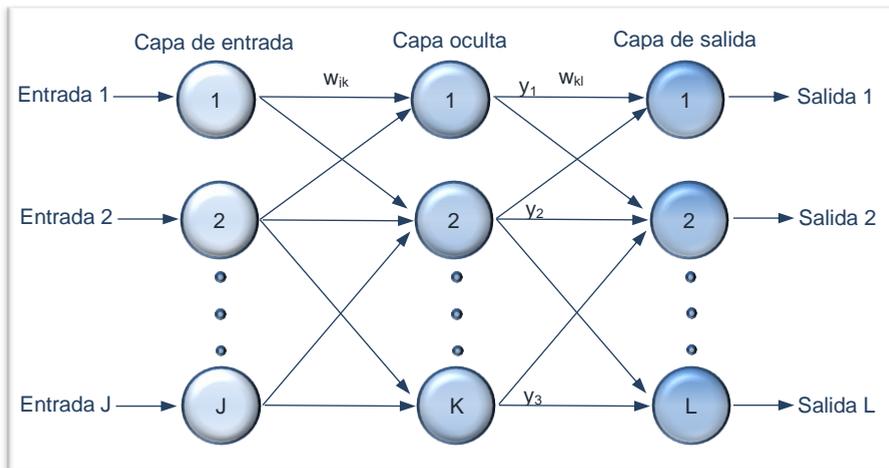


Fuente: modificado de Wang et al. (108)

Figura 10. Estructura de las máquinas de soporte vectorial.

Redes neuronales

Las **redes neuronales** son conjuntos de neuronas interconectadas organizadas en tres tipos de capas: la capa de entrada, que recibe los datos que queremos clasificar; la capa intermedia u oculta, que realiza los cálculos; y la capa de salida, que presenta los resultados (Figura 11). (109)



Fuente: modificado de Al-saffar et al. (109)

Figura 11. Arquitectura de una red neuronal.

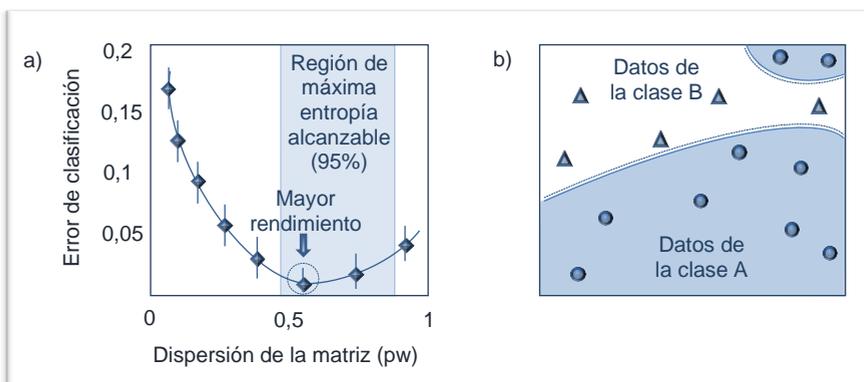
4) Clasificadores probabilísticos:

Los clasificadores probabilísticos calculan la probabilidad de que los datos de entrada pertenezcan a cada clase (110) y los clasifican en la clase más probable (111). Existen tres tipos de clasificadores probabilísticos: máxima entropía, Naïve Bayes y redes bayesianas. (110)

Máxima entropía (ME)

La **entropía** se utiliza para calcular la aleatoriedad, homogeneidad o incertidumbre de una variable. (112) El principio de máxima entropía afirma que debemos asumir la máxima incertidumbre sobre lo desconocido. (113)

Ante un grupo de datos de clasificación desconocida con un número finito de categorías excluyentes, existen múltiples opciones de distribución de categorías que podrían ser válidas. De ellas, el clasificador de ME selecciona la distribución de probabilidad con menor error y mayor incertidumbre y diseña a partir de ella un modelo predictivo (Figura 12). (114)



Notas: a) El modelo de ME con mayor rendimiento (menor error) en el área del 95% de ME se obtiene con $pw=0,6$; b) Representación gráfica de clasificación binomial basada en ME.

Fuente: a) Modificado de Qin et al. (115); b) Modificado de Jaakkola et al. (116)

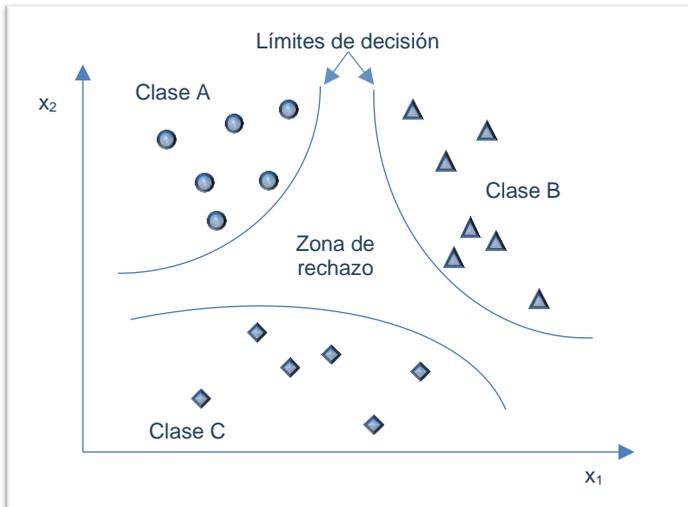
Figura 12. Clasificación basada en la máxima entropía.

Naïve Bayes

Los clasificadores tipo **Naïve Bayes** calculan la probabilidad de que una característica pertenezca a una clase según la distribución de las palabras en el documento. (117) Se basan en el teorema de Bayes. Este teorema aplicado al AS en una frase es el siguiente (111):

$$P(\text{Sentimiento} | \text{Frase}) = \frac{P(\text{Frase} | \text{Sentimiento}) \times P(\text{Sentimiento})}{P(\text{Frase})}$$

A partir de las probabilidades de que cada entidad pertenezca a cada una de las clases del estudio, el algoritmo divide el espacio en varias regiones de decisión, separadas por límites de decisión (Figura 13), y realiza la clasificación. (118)

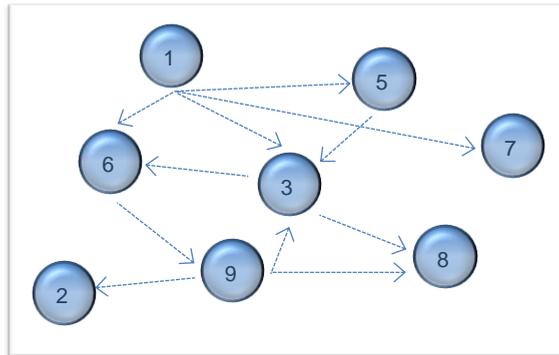


Fuente: modificado de Li et al. (118)

Figura 13. Regiones y límites de decisión.

Redes bayesianas

Las **redes bayesianas** (Figura 14) son grafos acíclicos que representan la relación entre un grupo de variables aleatorias según su distribución probabilística conjunta. Son útiles para realizar predicciones en conjuntos de datos incompletos. (119)



Fuente: modificado de Sesen et al. (119)

Figura 14. Ejemplo de red bayesiana.

3. Técnicas híbridas

Las **técnicas híbridas** calculan una puntuación según la polaridad de las unidades léxicas del texto y utilizan esta puntuación como uno de los atributos de los algoritmos de aprendizaje automático. Consiguen resultados más precisos que la aplicación individual de los métodos. (120)

D. Modelado de Tópicos

El **modelado de tópicos** es una herramienta estadística que permite determinar cuáles son los conceptos discutidos en un documento y agrupar los textos según la temática. (121)

El modelo de **Asignación Latente de Dirichlet (ALD)**, es el más usado para el modelado de tópicos. Representa los documentos como mezclas aleatorias de tópicos, donde cada tópico se caracteriza por tener una distribución de palabras determinada. (122)

Los algoritmos de modelado de tópicos **clásicos** utilizan técnicas de aprendizaje no supervisado que aplican el modelo ALD. En ellas, los algoritmos examinan el vocabulario de los documentos, determinan cuáles son las palabras que tienen más probabilidad de encontrarse agrupadas y, por tanto, constituir un tópico. (123) Posteriormente, los investigadores deben examinar los conjuntos de palabras agrupados por el algoritmo no supervisado, identificar cuáles son los conjuntos relevantes y nombrarlos.

Existen numerosas **alternativas** al modelo tradicional, también basadas en ALD. (124) La principal desventaja de estas técnicas es que, al no ser supervisadas, no garantizan que el algoritmo vaya a identificar los temas que los investigadores desean estudiar. (125) Por ello, su principal aplicación es el descubrimiento de tópicos en los textos. (122)

En cambio, si el objetivo es estudiar la presencia de temas predefinidos, como es el caso de la presente tesis, la aplicación de las **técnicas de AS** descritas en el apartado anterior permite obtener una clasificación temática más precisa que las técnicas basadas en ALD.

E. Evaluación de los Métodos de Clasificación Automática de Textos

La **efectividad** de un clasificador para identificar **cada una de las categorías** incluidas en la clasificación se evalúa en el conjunto de datos de prueba mediante las siguientes medidas (120,126):

Precisión.

Proporción de textos clasificados correctamente en una categoría (verdaderos positivos) respecto al número total de documentos clasificados en esa categoría (verdaderos positivos más falsos positivos). Indica el porcentaje de predicciones correctas que el algoritmo realiza en una categoría.

Sensibilidad, exhaustividad (*recall*).

Proporción de textos clasificados correctamente en una categoría (verdaderos positivos) respecto a todos los textos que pertenecen a esa categoría y deberían haber sido incluidos en ella (verdaderos positivos más falsos negativos). Indica el porcentaje de casos de una categoría que el algoritmo es capaz de detectar.

Puntuación F1 (*F1-score*).

Es la “media armónica de la precisión y la exhaustividad”. Constituye la medida más adecuada para la evaluación de clases con sensibilidad muy elevada (muchos verdaderos positivos con pocos falsos negativos) pero con precisión muy baja (muchos falsos positivos). Su fórmula es:

$$\text{Puntuación F1} = \frac{2 * \text{Precisión} * \text{Sensibilidad}}{\text{Precisión} + \text{Sensibilidad}}$$

Un clasificador suele tener sensibilidades, precisiones y puntuaciones F1 distintas para cada categoría evaluada. La media aritmética (**media macro**) y la armónica (**media micro**) de cada una de estas métricas en todas las categorías permite observar la **efectividad global** del clasificador para **todas las categorías** incluidas en un estudio.

Las medias macro dan igual peso a cada clase, mientras que las medias micro igualan los pesos de cada observación, por lo que las clases mayoritarias tienen más peso en las medias micro. Las medias macro son mayores que las medias micro si el algoritmo tiene un rendimiento bajo en las clases mayoritarias. Por el contrario, las medias micro son mayores que las medias macro si el rendimiento del algoritmo es deficiente en las clases minoritarias.

En las clasificaciones con categorías excluyentes, las medias micro de la sensibilidad, la precisión y la puntuación F1 son iguales y equivalen a la **exactitud** (*accuracy*) del algoritmo.

V. INFORMACIÓN Y DESINFORMACIÓN SOBRE VACUNAS EN TWITTER

Debido a la heterogeneidad de la privacidad de los mensajes en las redes sociales, los investigadores tienen acceso limitado a la información intercambiada en estos sitios *web*. (127) Dentro de las redes sociales, Twitter es la que ofrece **mayor contenido público**. Por ello, la mayoría de los proyectos sobre opiniones en redes sociales extraen sus datos de Twitter. (128)

A. Twitter

1. Generalidades sobre Twitter

El *microblogging* es un tipo de red social que permite a los usuarios intercambiar pequeñas cantidades de contenido. (129) Twitter es una plataforma de **microblogging** gratuita que permite a los usuarios intercambiar mensajes públicos y privados desde 2006. (130) Es la **decimoséptima** red social con más usuarios activos en el mundo, con una media de 353 millones de **usuarios activos** cada mes en 2020. Su página *web* es la onceava más visitada en Internet, con un billón de **visitantes** únicos a diario. (131)

Sus usuarios producen una media de 500 millones de **tuits diarios**. (131,132) Se estima que el 23% de los adultos que utilizan Internet también son usuarios de Twitter. La **franja de edad** con mayor número de usuarios es la comprendida entre 19 - 29 años (37%), seguida del grupo de edad entre 30 - 49 años (25%). (133)

En Twitter, cada usuario dispone de un espacio personal donde puede publicar mensajes llamados **tuits**. Los tuits pueden contener imágenes, enlaces a páginas *web* y textos con una restricción de 140

caracteres máximos desde 2006 hasta octubre de 2017 y de 280 caracteres desde noviembre de 2017 hasta la actualidad. (134) Cada tuit tiene un número de identificación único.

Cuando un usuario se registra en la plataforma, Twitter le asigna automáticamente un **número de identificación** único y permanente. Además, el usuario puede elegir un nombre de usuario, formado por el símbolo @ seguido de un máximo de 15 caracteres. El nombre de usuario también es único y el usuario puede cambiarlo tantas veces como quiera, sin que esto modifique su número de identificación.

Los usuarios de esta plataforma pueden convertirse en **seguidores** de otras cuentas con las que comparten intereses para recibir notificaciones cada vez que éstas publiquen nuevos tuits. Pueden valorar el contenido de los tuits propios y ajenos marcándolos como **“me gusta”** (**favoritos** hasta noviembre de 2015), escribir **respuestas** a otros tuits y retuitear, es decir, volver a publicar mensajes propios o de otros usuarios para ampliar su difusión. A los mensajes retuiteados se les denomina **retuits**. (135)

Además, pueden indexar sus mensajes por grupos temáticos con etiquetas llamadas **hashtags**, compuestas por el símbolo de la almohadilla (#) seguido de varios caracteres – letras o letras y números - sin espacios entre ellos. (136) Cuando un usuario publica un tuit con un *hashtag*, el *hashtag* se convierte en un enlace a la lista de todos los tuits que contienen el mismo *hashtag*. Los *hashtags* no están predefinidos por la plataforma, sino que los usuarios pueden crearlos libremente.

Los usuarios de Twitter pueden configurar sus cuentas como **privadas** o públicas y cambiar esta configuración cuando quieran. Si

deciden que su cuenta sea privada, sólo sus seguidores podrán ver el contenido de éstas. Por el contrario, cualquier persona puede consultar el contenido de las cuentas públicas sin necesidad de registrarse en Twitter.

2. Twitter como herramienta de investigación en salud

Numerosos grupos de investigación utilizan Twitter para extraer datos sobre las **percepciones** y **preocupaciones** en salud de los usuarios. A pesar de que Twitter no es la red social con mayor número de usuarios activos, presenta numerosas **ventajas** para este tipo de investigación en comparación con otros sitios sociales.

En primer lugar, como hemos mencionado anteriormente, Twitter permite **acceso público** a mayor cantidad de información que otras redes sociales. (128) En el caso de las **vacunas**, Twitter continúa aplicando **políticas flexibles** frente a la desinformación sobre vacunas, a diferencia de otros sitios sociales en la red que ya han limitado la visibilidad del contenido contrario a las vacunas. (137) Esto hace que los usuarios de esta red social estén expuestos a mayor cantidad de **desinformación**, pero a su vez convierte a esta plataforma en una fuente de datos más completa para la investigación sobre la reticencia vacunal.

Mientras que otras redes sociales están diseñadas para el establecimiento y mantenimiento de las conexiones sociales, la **función principal** de Twitter es permitir a los usuarios intercambiar **información** y **opiniones** sobre **cualquier tema**. (134,138) De hecho, podemos encontrar conversaciones sobre más de 379 aspectos relacionados con la salud en esta plataforma. (139)

Algunos de los **temas** multidisciplinares que han sido objeto de investigación en Twitter son:

- 📍 Adicciones: alcohol (140), cannabis (141,142) o tabaco (143,144).
- 📍 Endocrinología y nutrición: alimentación (145), diabetes (146).
- 📍 Farmacología: uso de antibióticos (147), abuso de medicamentos sujetos a prescripción médica (148) y farmacovigilancia (149).
- 📍 Psiquiatría: salud mental (150,151), trastornos de la conducta alimentaria (152) o el suicidio (153).
- 📍 Promoción de la salud: entornos saludables (154), prevención del cáncer (155,156) o evaluación del impacto de las campañas de educación para la salud (157).
- 📍 Epidemiología de las enfermedades transmisibles: *chikungunya* (158), ébola (159), zika (160,161) y la vigilancia de la gripe (132).
- 📍 Otros: el aborto (162), la demencia (163) o el dolor crónico (164).

Cabe destacar que las opiniones en Twitter están generadas en un contexto en el que el usuario se expresa libremente, sin el **sesgo** del entrevistador que se produce en la investigación cualitativa convencional. (133)

Los investigadores tienen la posibilidad de realizar un **seguimiento** del contenido publicado en **tiempo real**, además de implantar **intervenciones** dirigidas a la comunidad online y **reclutar** usuarios para posteriores estudios. (128,133)

El **procesamiento** de la información para su posterior análisis es menos laborioso que en otras fuentes debido a que la mayoría del contenido ya está proporcionado en forma de texto y estos textos tienen un número de caracteres limitados.

Adicionalmente, la descarga de los datos trae asociados numerosos **metadatos** de gran utilidad en investigación como el idioma, el nombre del autor, los datos públicos de su perfil, su geolocalización, su lista de contactos y el número y nombre de usuarios que han retuiteado el mensaje original o lo han marcado como favorito. (133) Toda esta información permite **vigilar** las **tendencias** en las conversaciones referentes a la salud en cada **región** y estudiar las **redes de contactos** para esclarecer cómo se produce el flujo de la información entre los usuarios. (138)

Por otra parte, Twitter permite a sus usuarios **citar** otros tuits y publicaciones realizadas en cualquier sitio de la red. Algunos estudios sugieren que las citas en Twitter pueden ayudar a predecir el impacto que tendrá un artículo fuera Twitter. (138)

Por todo ello, la mayor parte de la investigación cualitativa y cuantitativa en salud en redes sociales está desarrollada en Twitter, especialmente a partir de **2013**. (128,133)

3. Aplicaciones de Twitter en Salud Pública

Los profesionales de salud pública y los organismos e instituciones sanitarias utilizan Twitter para **tres tipos de actividades**: desarrollo profesional, investigación, educación y difusión de información. (165,166)

Desarrollo profesional

Twitter permite a los profesionales difundir **actualizaciones** sobre la investigación que están llevando a cabo y **contactar** con otros investigadores. (165,166)

Educación y difusión de información

Diversos profesionales de salud pública utilizan Twitter para difundir información dirigida al **público más joven** que recurre a las redes sociales, y no a los medios de comunicación tradicionales, para disipar sus dudas en salud. (166)

La limitación de caracteres exige que los textos sean **breves**, claros y centrados en el tema, lo que aumenta la probabilidad de que los usuarios los lean y entiendan. (165,166)

Los profesionales también suelen utilizar los tuits para explicar los documentos **técnicos** en lenguaje comprensible para la población general, desmentir **mitos** y compartir **enlaces** de fuentes fiables. (165,166)

Investigación:

Los investigadores pueden recopilar los tuits para analizar su **contenido** y establecer la influencia de las **redes de contactos** en las opiniones expresadas por los usuarios. (128,138)

En el ámbito de la salud pública, las conclusiones extraídas son útiles para (128,133,165,166):

-  Monitorizar las **tendencias** en salud según la **geolocalización**.
-  Realizar seguimiento a tiempo real de las **catástrofes** y las emergencias humanitarias.
-  Realizar **vigilancia** epidemiológica de enfermedades comunicables y no comunicables.
-  Diseñar e implementar **intervenciones** comunitarias.

Complementariamente, Twitter facilita el contacto entre los investigadores y la población general, por lo que constituye una herramienta útil en el proceso de **reclutamiento** de pacientes para encuestas sobre salud o cualquier otro tipo de estudios. (167)

B. Estudios sobre la Opinión acerca de las Vacunas en Twitter

Los tuits constituyen una potencial fuente de información sobre la opinión acerca de las vacunas. Los estudios realizados por otros autores sobre este tema pueden dividirse en dos grupos según las técnicas utilizadas para el análisis del contenido:

Técnicas manuales:

Destacan los estudios realizados por Bello-Orgaz et al. (168), Cuesta-Cambra et al. (169), Martin et al. (170) y Mollema et al. (171).

Técnicas automatizadas:

En la Tabla 5 resumimos los principales artículos publicados entre 2015 - 2020 que han aplicado técnicas automatizadas para realizar AS o modelado de tópicos, ordenados por tipo de vacuna estudiada y por fecha de inicio del periodo monitorizado.

Otros autores han empleado técnicas automatizadas para el análisis del comportamiento de los *bots* (62), el análisis semántico de los tuits (172) o el estudio de las comunidades de usuarios y sus interacciones (77).

Tabla 5. Trabajos relacionados

Referencia Año de publicación	Descripción	Periodo Localización Idioma	Diseño Método de análisis	Nº usuarios Nº tuits	Resultados
Vacuna contra la influenza					
McNeil et al. (173) 2017	Análisis de contenido, los patrones de retuits, las <i>URLs</i> y las redes de usuarios durante la pandemia por H1N1.	01/04/2009 - 01/05/2010. Reino Unido. Inglés.	Retrospectivo. Temas: aprendizaje no supervisado.	- 12 711 tuits.	Evolución: Aumento de tuits en octubre y noviembre, tras la comercialización de la vacuna en Reino Unido. Contenido: Mayoritariamente descriptivo: algunos tuits criticaban a las autoridades sanitarias y otros demostraban escepticismo sobre los beneficiarios económicos de la vacunación. Otros: Predominaron los retuits y las <i>URLs</i> de fuentes oficiales.
Huang et al. (174) 2017	Estudio de la aceptación de la vacunación frente a la influenza. Comparación de clasificadores.	Temporadas 2013 -2016. Mundial. Estudio geográfico sólo en EE. UU. Inglés.	Prospectivo AS: aprendizaje supervisado (Naïve Bayes, MSV y bosques aleatorios).	- 1 007 582 tuits recogidos, 10 000 tuits analizados.	AS: Resultados descartados por bajo rendimiento de los clasificadores (máximas puntuaciones F1 de 0,42 y 0,62 para los sentimientos positivo y negativo respectivamente), debido al bajo grado de acuerdo entre los investigadores en la clasificación manual. Otros: Con el mejor clasificador, correlación temporal de 0,9 y geográfica de 0,67 entre las encuestas sobre vacunación de los Centros para el Control y la Prevención de Enfermedades (CDC) y los mensajes en Twitter sobre las intenciones de los usuarios de vacunarse o haberse vacunado ya.
Vacuna SRP					
Deiner et al. (175) 2019	Análisis de las conversaciones en Twitter y Facebook sobre el sarampión en el contexto de brotes mediáticos.	2009-2015. EE. UU. Filtro: palabras clave en inglés.	Retrospectivo. AS: aprendizaje supervisado mediante regresiones logísticas.	- 58 078 publicaciones en Facebook. 82 993 tuits.	AS: a) Publicaciones en Facebook: 22 306 en contra de la vacunación, 17 928 a favor y 2 431 inciertos; b) Tuits: 43 259 en contra, 16 297 a favor y 12 222 inciertos. Otros: Correlación entre las publicaciones a favor de las vacunas y los casos semanales de sarampión (Rho de Spearman: Facebook 0,22, Twitter: 0,21), ausencia de correlación con las publicaciones con duda vacunal (Facebook: 0,01, Twitter: 0,001).

* *URL*: localizador de recursos uniforme, *Uniform Resource Locator*.

Tabla 5. Trabajos relacionados

Referencia Año de publicación	Descripción	Periodo Localización Idioma	Diseño Método de análisis	Nº usuarios Nº tuits	Resultados
Du et al. (176) 2018	Comparación de métodos de AS hacia las vacunas en Twitter durante un brote de sarampión.	01/12/2014 - 30/04/2015. Mundial. Filtro: palabras clave en inglés.	Prospectivo. AS y temas: aprendizaje supervisado (MSV, Naïve Bayes, redes neuronales y bosques aleatorios).	- 1 154 156 tuits recogidos, 1 151 analizados.	AS: a) Polaridad: 3,13% en contra de las vacunas, 79,32% neutros, 17,55% a favor, b) Emociones (sobre el sarampión): 79,84% preocupación, 9,47% humor o sarcasmo. Temas (sobre el sarampión): 62,38% noticias/información médica, 29,89% opiniones, 1,82% experiencias personales, 1,73% preguntas. Otros: Los modelos de redes neuronales convolucionales (CNN) fueron los que demostraron mayor rendimiento. El modelo de CNN con la combinación de dos incrustaciones obtuvo la mejor puntuación en la clasificación de temas y emociones (F1 micro: 0,7811 y 0,8592, respectivamente), mientras que el modelo de CNN con incrustación de Stanford logró el mejor desempeño en el análisis de la polaridad de sentimientos (F1 micro: 0,8642).
Vacuna contra el VPH					
Luo et al. (177) 2019	Evaluación de la opinión sobre la vacuna contra el VPH en Twitter a lo largo de 10 años mediante PLN.	01/01/2008 - 31/12/2017. Mundial. Inglés.	No especificado. Clasificación: aprendizaje supervisado (red neuronal).	- 287 100 tuits.	Evolución y AS (a nivel de frase): Pico de tuits en 2013. Más tuits negativos entre 2008-2011 y 2015-2016. Las localizaciones con más tuits negativos variaron cada año. Temas: Los más frecuentes en los tuits negativos fueron “lesiones”, “muertes”, “escándalos”, “problemas de seguridad” y “efectos adversos secundarios”, mientras que los principales temas en los tuits positivos fueron “cánceres de cuello uterino”, “pruebas de detección de cuello uterino”, “previene” y “campañas de vacunación”. Otros: Mención de la FDA, CDC y Merck en tuits positivos y negativos en todos los años de estudio, excepto 2010.

Tabla 5. Trabajos relacionados

Referencia Año de publicación	Descripción	Periodo Localización Idioma	Diseño Método de análisis	Nº usuarios Nº tuits	Resultados
Bahk et al. (178) 2016	Monitorización con la herramienta <i>Vaccine Sentimeter</i> del contenido en la red, incluido Twitter, sobre la vacunación contra la polio y el VPH.	10/2012-11/2014. Mundial. Tuits en inglés. Noticias en inglés, francés y castellano.	Prospectivo. AS: método no especificado.	- Polio: 39 308 tuits, AS en 1 534; VPH: 16 781 tuits, AS en 3 595.	Contenido: Aumento de los tuits sobre cada vacuna tras publicarse noticias sobre eventos relacionados con ellas. En el caso de la vacuna de la polio, el interés en Twitter fue menos duradero que en las noticias. En el caso de la vacuna frente al VPH, aunque el evento era negativo, las noticias y los tuits fueron predominantemente positivos hacia la vacuna (95,8% y 77,5% respectivamente), aunque las conversaciones negativas persistieron más que las positivas.
Dunn et al. (179) 2015	Estudio de la asociación entre la exposición y la expresión de opiniones negativas sobre las vacunas contra el VPH en Twitter.	01/10/2013-01/04/2014. Mundial. Inglés.	Retrospectivo. AS: aprendizaje supervisado (4 clasificadores basados en conexiones sociales).	30 621 usuarios. 83 551 tuits.	AS: Exactitud de los clasificadores entre 0,88 – 0,94. 25,13% tuits negativos y 74,87% positivos/neutros. Otros: La probabilidad de que un usuario publique un tuit negativo después de exponerse mayoritariamente a opiniones negativas fue del 37,78%, frente al 10,92% de probabilidad para los usuarios expuestos mayoritariamente a opiniones positivas y neutras.
Shapiro et al. (180) 2017	Análisis de las preocupaciones sobre la vacuna VPH en tres países.	01/2014 - 04/2016. Australia, Canadá y Reino Unido. Inglés.	Prospectivo. Clasificación: aprendizaje supervisado (MSV).	16 789 usuarios. 43 852 tuits.	Descripción: Usuarios: 16,6% de Australia, 43,1% de Canadá y 40,3% de Reino Unido. Porcentaje similar de tuits ($\pm 0,1 - 0,2\%$) Clasificación de los tuits: Precisión y exactitud del 90%. % de tuits que expresaban preocupaciones sobre la vacuna en cada país: el 14,9% de Canadá, el 19,4% de Australia y el 22,6% de Reino Unido. Sin diferencias en los tipos de preocupaciones entre países. Las inquietudes sobre las "barreras percibidas" fueron las más frecuentes. Otros: Los usuarios que expresan preocupación sobre la vacuna tenían una mayor proporción de seguidores internacionales que también expresaban preocupación.

Tabla 5. Trabajos relacionados

Referencia Año de publicación	Descripción	Periodo Localización Idioma	Diseño Método de análisis	Nº usuarios Nº tuits	Resultados
Massey et al. (181) 2016	AS y temas de las conversaciones sobre la vacuna contra el VPH en Twitter.	01/08/2014- 31/07/2015. Mundial. Inglés.	Retrospectivo. AS: aprendizaje supervisado (clasificador lineal binario y árboles de decisión multicategoría).	78 643 usuarios. 193 379 tuits.	AS: AUC = 0,918 para la clasificación del sentimiento hacia las vacunas. 38,99% tuits positivos, 25,31% negativos, 12,98% neutros, 22,72% no mencionaban la vacuna. Temas: Tuits sobre efectos adversos (AUC=0,739): el 46,44% de los negativos, el 17,14% de los positivos y el 15,08% de los neutros. Tuits sobre prevención o protección (AUC=0,774): el 27,09% de los positivos, el 25,79% de los neutros y el 11,5% de los negativos.
Du et al. (182) 2017	AS sobre las vacunas contra el VPH en Twitter.	02/11/2015 - 28/03/2016. Mundial. Inglés.	Prospectivo. AS: aprendizaje supervisado (MSV).	- 184 214 tuits, 110 778 relacionados con el tema de estudio.	AS: Puntuación F1 media micro de 0,786 y media macro de 0,7081. 32,0% negativos, 32,1% neutros, 35,8% positivos. Temas (tuits negativos): 79,22% seguridad, 20,44% otros, 0,35% eficacia, ninguno sobre costes. Otros: Asociaciones entre algunos sentimientos y diferentes días de la semana.
Vacunación (general)					
Tomeny et al. (183) 2017	Análisis de los tuits sobre vacunas y trastornos del espectro autista por distribución geográfica en Twitter.	01/01/2009 - 21/08/2015. Mundial. Estudio geográfico sólo en EE. UU. Filtro: palabras clave en inglés.	Retrospectivo. AS: aprendizaje automático supervisado (método no especificado).	47 236 usuarios. 549 972 tuits.	AS: Exactitud del clasificador: 0,86. 49,56% tuits negativos, 51,44% otros. El volumen de tuits negativos aumentó después de la publicación de noticias relacionadas con las vacunas. Menos negatividad en los tuits geolocalizados (31,02%). Otros: Más tuits en California, <i>Connecticut</i> , <i>Massachusetts</i> , Nueva York y Pensilvania. El 67% de la varianza en la agrupación geográfica de los tuits negativos se explicaba por las características demográficas (población más grande y con más % de: púerperas, hogares con niveles altos de ingresos, hombres de 40 - 44 años y hombres sin estudios superiores).

Tabla 5. Trabajos relacionados

Referencia Año de publicación	Descripción	Periodo Localización Idioma	Diseño Método de análisis	Nº usuarios Nº tuits	Resultados
Gunaratne et al. (184) 2019	Caracterización de del discurso a favor y en contra de las vacunas en Twitter según la polaridad de los <i>haghtags</i> .	01/01/2010 - 01/01/2019. Mundial. Inglés.	Retrospectivo. AS: automatizada basada en léxico (<i>hashtags</i>).	291 747 usuarios. 1 637 712 tuits.	Evolución: Aumento progresivo de los tuits positivos. Disminución de los negativos desde 2014, pero aumento del porcentaje de usuarios negativos (de 8,1% en 2015 al 16% en 2018). AS: Clasificación de los tuits según la polaridad media de sus <i>hashtags</i> : 35% tuits negativos, 64% positivos, 1% ambiguos. El 1,8% de los usuarios usaron <i>hashtags</i> tanto a favor como en contra de las vacunas, el 12% sólo en contra y el 86% sólo a favor. Los <i>hashtags</i> antivacunas más utilizados fueron: #cdcwhistleblower, (280 779 tuits), #vaxxed (123 382 tuits), #hearthiswell (44 426 tuits), #novax (32 424 tuits) y #cdcfraud (21 750 tuits). Otros: Poca interacción entre los usuarios de la misma polaridad (4,8%), mínima intercomunicación entre comunidades con polaridades opuestas (0,2%).
Dredze et al. (185) 2017	Análisis del discurso sobre vacunas en Twitter durante la campaña presidencial de 2016 en EE. UU.	01/12/2014 - 01/12/2016. Mundial. Filtro: palabras clave en inglés.	Prospectivo. AS: basado en léxico (<i>hashtags</i>). Temas: manual.	- 538 000 tuits recopilados y analizados para AS. 100 tuits clasificados por temas.	AS: Clasificaron como negativos los tuits que contenían alguno de los 25 <i>hashtags</i> identificados como negativos por los autores. 11 144 tuits eran negativos y mencionaban a algún candidato a las elecciones presidenciales de los EE. UU. Temas: Los tuits negativos comúnmente hablaban sobre investigar a los CDC, permitir exenciones de vacunación, eliminar mandatos o derogar la Ley Nacional de Lesiones por Vacunas en la Niñez, que disminuye el riesgo legal y riesgo financiero para los fabricantes de las vacunas. Otros: Los ocho períodos con picos de tuits sucedieron en momentos en los que se produjeron eventos relacionados con las vacunas.

Tabla 5. Trabajos relacionados

Referencia Año de publicación	Descripción	Periodo Localización Idioma	Diseño Método de análisis	Nº usuarios Nº tuits	Resultados
D'Andrea et al. (186) 2019	Análisis en tiempo real de la opinión pública sobre la vacunación en Twitter durante cinco meses en Italia.	01/09/2016 - 31/01/2017 Italia. Italiano.	Prospectivo. AS: comparación de varios sistemas de aprendizaje supervisado.	- 112 397 tuits.	AS: El mejor clasificador fue el MSV (exactitud del 64,84%). 64% tuits neutros (máximo 82% en diciembre 2016), 19% positivos (máximo 27% en mayo 2017), 17% negativos (máximo 26% en mayo 2017). La mayoría de los tuits neutros eran noticias. Otros: Pocos tuits en las épocas sin eventos relacionados (cambios legislativos, discursos, documentales...). Aumento de tuits inmediatamente después de las noticias virales y horas o días después del resto de eventos.
Tavoschi et al. (187) 2020	Análisis en tiempo real de la opinión pública sobre la vacunación en Twitter durante un año en Italia.	09/2016 - 08/2017 Italia. Italiano.	Prospectivo. AS: comparación de varios sistemas de aprendizaje supervisado.	- 180 620 tuits.	Evolución: Tendencia creciente del número de tuits ($R^2 = 0,27$, modelo lineal). Polarización progresiva del discurso. AS: El mejor clasificador fue el MSV (exactitud del 64,8%). 60% tuits neutros, 17% positivos, 23% negativos. Otros: Aumento de tuits cuando se produjeron sucesos relacionados con las vacunas, sobre todo ante la aprobación del decreto de vacunación obligatoria frente a determinadas enfermedades infantiles. Mayor presencia de <i>hashtags</i> (sobre todo <i>#novaccines</i>) y de la palabra "autismo" en los tuits en contra de la vacunación que en los tuits a favor de las vacunas. En cambio, en los tuits positivos había más insultos hacia los activistas anti-vacunas y más referencias sobre política. En ambos conjuntos de datos, la principal enfermedad inmunoprevenible fue el sarampión.

C. Justificación del Estudio

Internet y las redes sociales desempeñan un papel importante en la difusión de la desinformación sobre vacunas, ya que permiten a la comunidad reticente compartir sus teorías sobre el riesgo de las vacunas instantáneamente desde cualquier lugar del mundo y coordinar las acciones contra las políticas que favorecen la vacunación. (188)

Aunque Wang et al. (189) exponen que la cantidad de desinformación sobre vacunas en las redes sociales es menor que la información correcta, también han detectado que la desinformación alcanza mayor popularidad. Shah et al. (190) llegan a conclusiones similares tras analizar los enlaces a páginas *web* sobre vacunación compartidos en Twitter.

Dado que la exposición a la desinformación sobre vacunas incrementa la duda vacunal y el riesgo de epidemias, (62) es importante que los profesionales de salud pública incluyan intervenciones sobre las redes sociales en el abordaje de la reticencia vacunal. Para ello, primero deben monitorizar las conversaciones y determinar cuál es la naturaleza y la magnitud del problema de desinformación en estas plataformas.

Si bien el estudio manual de la opinión aporta resultados muy precisos y fiables, el análisis manual de grandes cantidades de datos es lento y requiere la inversión de muchos recursos humanos y económicos. En cambio, los métodos automatizados permiten analizar un gran volumen de datos en menor cantidad de tiempo y con menor coste.

A pesar de que otros autores han realizado AS y modelado de tópicos sobre vacunas en Twitter mediante métodos automatizados, como se ha descrito en el apartado anterior, estos análisis suelen estar limitados a conjuntos pequeños de datos, de períodos de tiempo cortos y en raras ocasiones estudian el origen geográfico de la información y desinformación.

Por otra parte, aunque las **técnicas híbridas** de AS son las de mayor rendimiento (120), ninguno de estos autores ha aplicado estas técnicas, sino que se limitan a utilizar AS basado en léxico o técnicas de aprendizaje automático supervisado o no supervisado.

Por ello, planteamos un estudio del contenido sobre vacunas en Twitter mediante técnicas de inteligencia artificial, consistente en: un análisis semántico, un AS basado en aspectos mediante técnicas híbridas y un modelado de tópicos mediante técnicas de AS. Analizamos un conjunto de datos de gran escala (*big data*), de un periodo de ocho años, segmentado según la localización geográfica de los autores de los tuits.

CAPÍTULO II

HIPÓTESIS Y OBJETIVOS

I. HIPÓTESIS

Twitter es una fuente de datos desestructurados que podemos estructurar y analizar mediante técnicas de inteligencia artificial con el fin de obtener información relevante acerca de la opinión sobre las vacunas y los motivos de la duda vacunal entre los usuarios de esta plataforma social.

II. OBJETIVOS

A. Objetivo General

Analizar el discurso sobre vacunas de los usuarios de Twitter desde junio de 2011 hasta abril de 2019, en un conjunto de macrodatos, mediante algoritmos automatizados diseñados específicamente para este estudio.

B. Objetivos Específicos

-  Describir las características básicas de los tuits del estudio: frecuencia, idioma, localización y evolución temporal.
-  Enumerar los unigramas y bigramas a nivel de palabra más frecuentes en los tuits escritos en castellano y en inglés.
-  Evaluar el rendimiento de nuestro algoritmo híbrido para la clasificación de la polaridad de los sentimientos hacia las vacunas.
-  Calcular la prevalencia y la tendencia del contenido negativo, neutro y positivo hacia las vacunas en el conjunto de tuits.

-  Determinar si hubo diferencias en el nivel de interacción social que generaron los tuits de distintas polaridades de sentimiento.
-  Caracterizar a los usuarios que escribieron contenido sobre vacunas en Twitter en el período de estudio.
-  Identificar si hubo una variación en la polaridad del discurso de los usuarios de Twitter después de las campañas de promoción de las vacunas en este medio.
-  Definir los temas más frecuentes en el discurso sobre vacunas en Twitter en el período de estudio, con énfasis en los tuits que expresaron duda vacunal.
-  Comprobar si hubo relación entre las coberturas vacunales contra el VPH y el número de tuits promotores de esta vacuna entre 2012 - 2018 en los seis países con más tuits sobre esta vacuna.
-  Examinar si hubo relación entre el número de casos de sarampión y el número de tuits reticentes hacia la vacuna SRP entre enero de 2012 y abril de 2019 en los países con más de 100 tuits negativos sobre esta vacuna.

CAPÍTULO III

MATERIAL Y MÉTODOS

I. DISEÑO DEL ESTUDIO

Realizamos un estudio observacional, longitudinal, retrospectivo, descriptivo con componentes analíticos sobre la opinión de los usuarios de Twitter respecto a las vacunas.

II. FUENTE DE INFORMACIÓN

 La principal fuente de información que utilizamos para extraer los textos sobre vacunas fue **Twitter**. Los motivos de elección de Twitter y no otras redes sociales son los expuestos en el capítulo I, que incluyen:

 **Menor número de caracteres de los mensajes:** A diferencia de otras redes sociales, los mensajes de Twitter tienen una restricción de caracteres (máximo 140 caracteres desde 2006 hasta octubre de 2017 y 280 caracteres desde noviembre de 2017 hasta la actualidad). Los algoritmos de AS que se aplicarán en el presente estudio han sido validados en textos con número reducido de caracteres, como los mensajes de Twitter, pero no hay suficientes estudios que demuestren su fiabilidad y validez en textos más largos.

 **Mayor diversidad de contenido sobre las vacunas:** Twitter permite a sus usuarios escribir mensajes tanto a favor como en contra de las vacunas con menos limitaciones que otras redes sociales. Pinterest ha bloqueado las búsquedas de todos los mensajes que hacen referencia a las vacunas para evitar que su red se emplee como medio de difusión de información falsa acerca de las vacunas. Por la misma razón, Instagram ha eliminado *hashtags* antivacunas y, junto con Facebook, ha

decidido ocultar publicaciones que manifiestan rechazo a las vacunas, lo que dificulta el acceso a esta información.

 **Mayor acceso a la información:** Twitter facilita la descarga de la información de su plataforma a través de varias interfaces de programación de aplicaciones (API).

 Adicionalmente, utilizamos los casos de sarampión y las coberturas vacunales de VPH estimadas por la **OMS/UNICEF** en cada país entre 2011 - 2019, disponibles en <https://www.who.int>.

III. MARCO TEMPORAL. CRITERIOS DE INCLUSIÓN Y EXCLUSIÓN

Realizamos la descarga de los tuits en mayo de 2019. Incluimos tuits escritos en castellano o en inglés, entre el 1 de junio de 2011 y el 30 de abril de 2019.

Los **criterios de exclusión** fueron:

-  Mensajes compuestos únicamente de: *URLs*, *hashtags*, símbolos o caracteres especiales.
-  Textos en idioma distinto al castellano o inglés.
-  Mensajes sobre la vacunación de animales.

IV. ENTORNO DE TRABAJO

A. Servidor para la Base de Datos

Utilizamos como servidor un ordenador con 64 GB de memoria RAM y un procesador i5-9600k con CPU 3,70GHz (Figura 15).

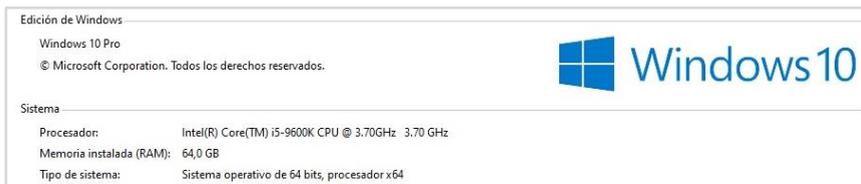


Figura 15. Características del servidor para la base de datos.

B. Estación de Trabajo

Utilizamos como estación de trabajo un ordenador con 32 GB de memoria RAM y un procesador i7-4790 con CPU 3,60GHz (Figura 16).

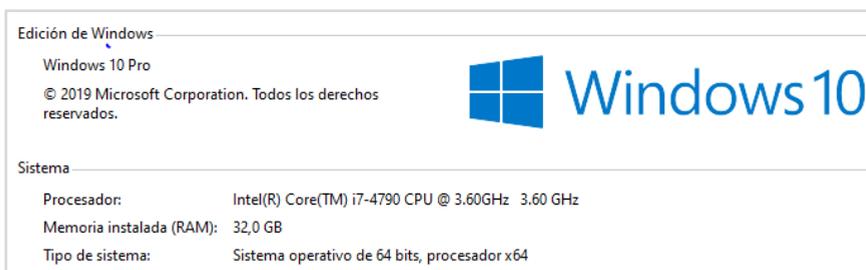


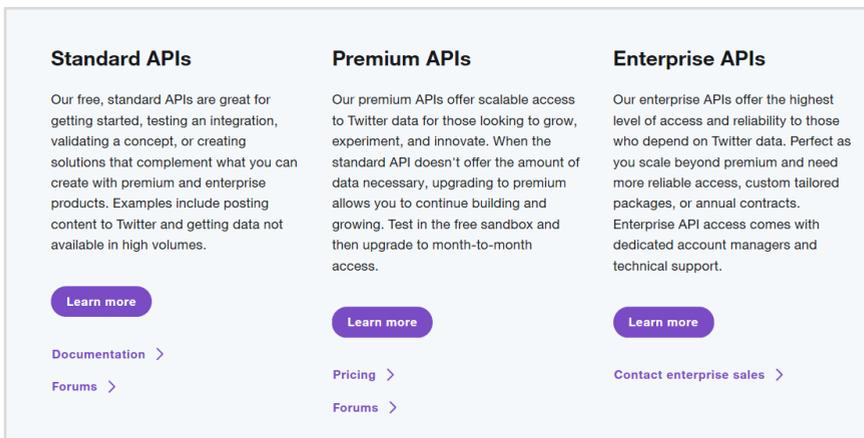
Figura 16. Características de la estación de trabajo.

V. TECNOLOGÍAS Y HERRAMIENTAS UTILIZADAS

Para llevar a cabo el proceso de extracción y explotación de los datos, empleamos diferentes tecnologías de información. En este apartado se describen los **lenguajes de programación, herramientas y gestores de bases de datos** que han facilitado esta labor. Así mismo, se da una breve explicación de la aplicación específica de cada herramienta dentro del marco de este proyecto de investigación.

A. Aplicación para la Extracción de los Datos de Twitter

Existen diferentes métodos para extraer la información publicada en Twitter. Twitter dispone de varias API que permiten a cualquier usuario descargar los mensajes publicados en esta plataforma y sus metadatos asociados. De las tres versiones disponibles (Figura 17), la *Standard API* es la versión gratuita, que sólo permite obtener los tuits publicados en los últimos 7 días, mientras que las versiones *Premium* y *Enterprise* son de pago y permiten extraer los datos históricos. Utilizamos la *Standard API* para el estudio piloto previo al inicio de esta investigación.



Standard APIs	Premium APIs	Enterprise APIs
Our free, standard APIs are great for getting started, testing an integration, validating a concept, or creating solutions that complement what you can create with premium and enterprise products. Examples include posting content to Twitter and getting data not available in high volumes.	Our premium APIs offer scalable access to Twitter data for those looking to grow, experiment, and innovate. When the standard API doesn't offer the amount of data necessary, upgrading to premium allows you to continue building and growing. Test in the free sandbox and then upgrade to month-to-month access.	Our enterprise APIs offer the highest level of access and reliability to those who depend on Twitter data. Perfect as you scale beyond premium and need more reliable access, custom tailored packages, or annual contracts. Enterprise API access comes with dedicated account managers and technical support.
Learn more	Learn more	Learn more
Documentation >	Pricing >	Contact enterprise sales >
Forums >	Forums >	

Fuente: <https://developer.twitter.com/en.html>

Figura 17. APIs de Twitter.

Como alternativa a las API de Twitter, varias empresas y desarrolladores externos comercializan APIs propias. Para la extracción de los datos de Twitter en el estudio principal, empleamos una aplicación desarrollada específicamente para este proyecto por un ingeniero de *software*.

B. Python



Python es uno de los lenguajes de programación más usados para el procesamiento del *big data*, dada la simplicidad de su sintaxis, su compatibilidad con diversas plataformas y la gran cantidad de librerías especializadas de las que dispone para el tratamiento de grandes volúmenes de información. En este proyecto empleamos Python para programar la extracción y explotación de los datos. Las librerías que utilizamos se resumen en la Tabla 6.

Tabla 6. Librerías de Python utilizadas en este estudio	
Librería y referencia	Descripción
<p>GeoPy</p>  <p>https://geopy.readthedocs.io</p>	<p>Permite determinar las coordenadas (latitud y longitud) de las localizaciones (direcciones, ciudades, regiones...), haciendo uso de servicios de geolocalización como Google Maps, Bing Maps o Nominatim.</p>
<p>JavaScript Object Notation (JSON)</p>  <p>https://www.json.org/json-en.html</p>	<p>Incluye las funciones necesarias para trabajar con archivos en el formato de datos JSON. La utilizamos para guardar la información de todos los tuits, transformar estos datos a otros formatos para su tratamiento y exportarlos.</p>
<p>Natural Language Toolkit (NLTK)</p>  <p>https://www.nltk.org</p>	<p>Permite realizar PLN. Utilizamos las herramientas de NLTK para preprocesar los datos, construir y entrenar el clasificador de temas e importar el analizador de sentimientos VADER *.</p>
<p>scikit-learn</p>  <p>https://scikit-learn.org</p>	<p>Librería utilizada para el aprendizaje automático en Python. La usamos para crear y entrenar el clasificador híbrido de sentimientos y para calcular las métricas de nuestro algoritmo y las de VADER.</p>

Tabla 6. Librerías de Python utilizadas en este estudio

Librería y referencia	Descripción
<p>TextBlob</p>  <p>TextBlob</p> <p>https://textblob.readthedocs.io/en/dev</p>	<p>Librería utilizada para el tratamiento de datos de texto. Permite realizar tareas de procesamiento de lenguaje natural muy comunes como el etiquetado gramatical (<i>part-of-speech tagging</i>, POST), la extracción de frases o el AS, entre otros. En este proyecto realizamos POST mediante <i>TextBlob</i>.</p>

* **Valence Aware Dictionary and sEntiment Reasoner (VADER)**
(191):

Herramienta de código abierto especializada en AS en redes sociales. Utiliza un modelo basado en lexicón y en reglas para dar a cada texto una puntuación de polaridad compuesta, con valores entre -1 y 1.

Arbitrariamente, se suele considerar que los textos con una puntuación compuesta menor de -0,05 son negativos, entre -0,05 y 0,05 son neutros y mayor de 0,05 son positivos.

C. Visual Studio Code

Editor de código fuente gratuito desarrollado por Microsoft. Proporciona una gran cantidad de herramientas para el desarrollo de aplicaciones utilizando diversos lenguajes, entre ellos Python. En este proyecto se utilizó Visual Studio Code con Python para el desarrollo de la aplicación de extracción y explotación de datos descargados desde Twitter.

D. Visual Studio (C#)

Visual Studio C# es otro editor de código fuente de Microsoft, utilizado para el desarrollo profesional de aplicaciones. En este trabajo utilizamos el lenguaje C# para crear una aplicación específica en la que los investigadores pudieron visualizar los tuits del estudio y clasificarlos por temas y polaridad de sentimientos (Figura 18).

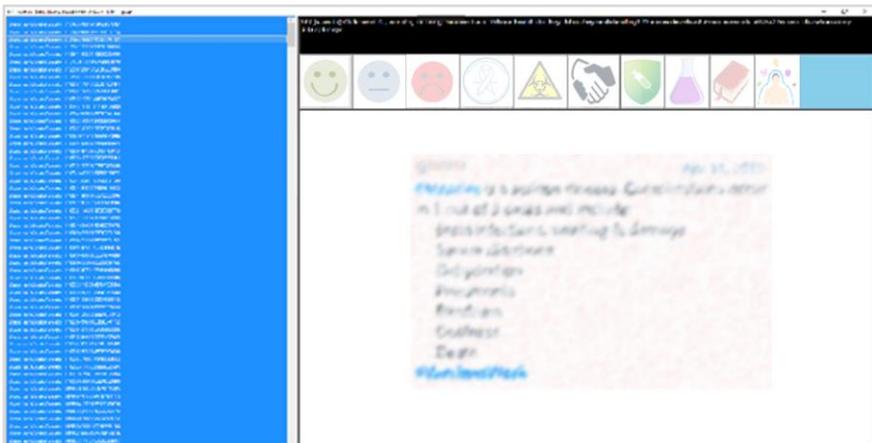


Figura 18. Aplicación para la clasificación manual de tuits.

E. Botometer



Botometer (70) es una herramienta de detección de *bots* desarrollada por el Observatorio en Redes Sociales (*Observatory on Social Media, OSoMe*) de la Universidad de Indiana. Extrae múltiples atributos de cada cuenta, tales como sentimientos, metadatos del usuario, de sus amigos y sus dispositivos, o la frecuencia y los patrones de tuits, retuits y menciones. Además, incluyen características lingüísticas en el análisis para la evaluación de usuarios que escriben en lengua inglesa.

A partir de estos atributos, un algoritmo de aprendizaje supervisado de bosques aleatorios asigna una puntuación entre 0–5 a cada cuenta. A mayor puntuación, la cuenta analizada presenta más características en común con las cuentas controladas por *bots* con las que se entrenó el algoritmo y menos con las cuentas controladas por humanos. Además de la puntuación de *bot* general, Botometer también asigna puntuaciones según la similitud con varios subtipos de *bots*. Otro algoritmo aplica el teorema de Bayes para calcular la probabilidad completa de automatización (CAP, por sus siglas en inglés), esto es, la probabilidad de que la cuenta esté realmente controlada por un *bot* (Figura 19).

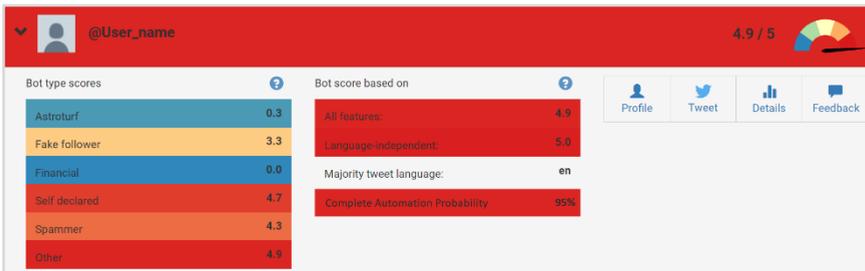


Figura 19. Clasificación mediante Botometer.

VI. PROCEDIMIENTO DE OBTENCIÓN DE DATOS

En primer lugar, realizamos un **estudio piloto** para decidir los criterios de selección de textos para la investigación principal. Para ello, mediante la *Standard API* de Twitter, extrajimos prospectivamente los tuits escritos entre el 18 de marzo de 2019 y el 15 de abril de 2019 que contenían alguno de los siguientes términos: *vaccin** OR *immuni** OR *antivax** OR *getvax* OR *vax** OR *vacuna** OR *inmuniza** OR *antivacun**. Se obtuvieron 27 000 tuits.

Examinamos manualmente estos mensajes y detectamos que el tema principal de algunos de ellos no era la vacunación. Con el fin de acotar la búsqueda de datos al tema de interés, decidimos que para la investigación principal analizaríamos únicamente mensajes con *hashtags* asociados. Seleccionamos los **28 *hashtags*** más comunes relativos a la vacunación, en inglés y en castellano:

👉 **20 *hashtags* en inglés:** *#antivaxx*, *#antivaxxer*, *#getvax*, *#immunization*, *#learntherisk*, *#vaccinate*, *#vaccinateyourkids*, *#vaccination*, *#vaccinations*, *#vaccine*, *#vaccinefailure*, *#vaccinehesitancy*, *#vaccineinjury*, *#vaccinesafety*, *#vaccineskill*, *#vaccines*, *#vaccinesdontwork*, *#vaccinessavelives*, *#vaccineswork* y *#vaxxed*.

👉 **8 *hashtags* en castellano:** *#antivacunas*, *#inmunizacion*, *#lasvacunasfuncionan*, *#lasvacunassalvanvidas*, *#vacuna*, *#vacunas*, *#vacunacion* y *#vacunate*.

Una vez finalizado el estudio piloto, iniciamos la fase de descarga de la información de Twitter para la **investigación principal**. Con la aplicación desarrollada, realizamos una descarga retrospectiva de todos los tuits públicos originales (no retuits), disponibles en mayo de

2019, escritos entre el 1 de junio de 2011 y el 30 de abril de 2019, que mencionaban alguno de los *hashtags* seleccionados en el estudio piloto. Almacenamos los siguientes **datos de cada tuit**:

- 📄 Contenido del tuit: sólo el texto, descartamos las imágenes.
- 📄 Número de identificación del tuit.
- 📄 Fecha de publicación de cada mensaje: día, mes y año.
- 📄 Número de retuits, favoritos/me gusta y respuestas.
- 📄 Idioma.
- 📄 Número de identificación del usuario autor del mensaje.
- 📄 Localización de los usuarios en mayo de 2019: geolocalización, es caso de ser pública, y ubicación escrita por el usuario en su perfil.

Obtuvimos 1 652 156 tuits escritos por 307 682 usuarios distintos.

VII. ALMACENAMIENTO DE DATOS

Los tuits descargados estaban en formato de texto JSON y los almacenamos en una base de datos SQL. Por otra parte, generamos una base de datos de todos los *hashtags* contenidos en estos tuits.

VIII. DEPURACIÓN DE LOS DATOS

Tras la recolección de datos, eliminamos los tuits que sólo contenían *URLs*, *hashtags*, símbolos o caracteres especiales; los mensajes escritos en idiomas distintos al castellano o al inglés y los tuits sobre la vacunación de animales. Después de este proceso, quedaron 1 499 227 tuits originales de 278 858 usuarios.

IX. ANONIMIZACIÓN DE LOS TUIITS

Anonimizamos los números de identificación de los tuits mediante enmascaramiento de datos. Reemplazamos los nombres de los usuarios mencionados en los tuits por *user_mention*.

X. GEOLOCALIZACIÓN

Dado que la mayoría de los tuits no disponían de localización propia, procedimos a asignarles la localización del **usuario** en el momento de la descarga de datos (mayo de 2019). En los casos en los que la **geolocalización** estaba disponible, la seleccionamos y descartamos la localización del perfil, por ser menos fiable.

En los individuos sin geolocalización, revisamos las localizaciones extraídas de los **perfiles**. Eliminamos las ubicaciones irreales y las localizaciones imprecisas (por ejemplo, mundo, global, nombres de continentes o municipios con el mismo nombre en distintos países si el usuario no especificaba el país).

A continuación, mediante la librería de Python GeoPy, geocodificamos las localizaciones, es decir, transformamos las direcciones en coordenadas (Figura 20). Posteriormente, codificamos las coordenadas por países para homogeneizar los datos.

```

Geolocalizationpy > web
25 a = 1
26 for line in file:
27     person_dict = json.loads(line)
28     type(file)
29     time.sleep(1)
30     location = web(1, "https://twitter.com/"+person_dict["screen_name"])
31     locationlonlat = geolocator.geocode(location)
32     if locationlonlat != None:
33         jsonlocation = json.dumps({'ID': person_dict["ID"], 'location': location, 'lat': locationlonlat.latitude, 'long': locationlonlat.longitude})
34         data["loc"].append(jsonlocation)
35         print(data)
36         a = a + 1
37     else:
38         jsonlocation = json.dumps({'ID': person_dict["ID"], 'location': location, 'lat': 0, 'long': 0})
39         data["loc"].append(jsonlocation)
40         print(data)
41         a = a + 1
42
43
44 if a > 3:
45     break
46
47 with open('userlocatios.txt', 'w') as outfile:
48     json.dump(data, outfile)

```

Figura 20. Geocodificación mediante Python.

XI. PREPROCESAMIENTO DE LOS DATOS

Para que un algoritmo de aprendizaje automático procese los tuits, es necesario transformar previamente los textos en entidades normalizadas que contengan la información relevante y eliminar el ruido. A este proceso se le llama **preprocesamiento** de datos.

En el presente estudio, realizamos la preparación de datos mediante las librerías NLTK y *TextBlob*. El proceso consistió en:

-  Eliminar los signos de puntuación, las URLs, los emoticonos, los símbolos – incluido el de la almohadilla (#) – y los caracteres especiales.
-  Convertir todas las letras a minúsculas.
-  Reemplazar dos o más espacios con un solo espacio.
-  *Tokenizar*: dividir las frases en palabras.
-  Corrección ortográfica.
-  Clasificar las palabras en clases: adjetivos, adverbios, artículos, conjunciones, interjecciones, preposiciones, pronombres, sustantivos y verbos.

- 🖱 Eliminar las “palabras vacías” (artículos, conjunciones, preposiciones y pronombres personales).
- 🖱 Lematizar: sustituir las palabras en forma flexionada por su lema, esto es, su forma canónica o de diccionario (por ejemplo, infinitivo para los verbos o masculino singular para los sustantivos).

XII. ANÁLISIS SEMÁNTICO

Después de eliminar las palabras vacías y previamente a la lematización, utilizamos TextBlob para extraer los unigramas y bigramas a nivel de palabra de los tuits. A continuación, analizamos la frecuencia de estos términos.

XIII. CLASIFICACIÓN DE TUI TS SEGÚN LA POLARIDAD DE SENTIMIENTOS

Definimos tres categorías excluyentes de tuits según su polaridad hacia las vacunas:

- 🖱 **Positivos:** Tuits a favor de la vacunación o que contienen mensajes orientados a disminuir la duda vacunal.
- 🖱 **Negativos:** Tuits que expresan reticencia vacunal, total o parcial.
- 🖱 **Neutros:** Tuits que no cumplen los criterios para ser clasificados como positivos o negativos. Esta categoría incluye tuits objetivos y tuits cuya polaridad frente a las vacunas no se puede entender sin un contexto adicional.

Realizamos el análisis de la polaridad de los sentimientos aplicando un **modelo híbrido**, derivado de la combinación de un enfoque basado en léxico y un enfoque de aprendizaje automático supervisado.

En primer lugar, confeccionamos un **lexicón propio**, específico para el estudio de la polaridad de los sentimientos hacia las vacunas. Dos codificadores independientes asignaron manualmente una polaridad frente a las vacunas a los *hashtags* de la base de datos creada antes del preprocesamiento y a los unigramas y bigramas a nivel de palabra más frecuentes.

Por otra parte, los dos **codificadores** independientes clasificaron manualmente la polaridad frente a las vacunas de una muestra aleatoria simple de tuits. Los codificadores utilizaron la aplicación programada en C# específicamente para este estudio para visualizar los mensajes (textos depurados, anonimizados y sin preprocesar) y registrar la clasificación que asignaban a cada tuit. El proceso finalizó cuando el programa detectó que se habían identificado un mínimo de 5 000 tuits de cada clase.

La **confiabilidad intercodificador** fue alta (puntuación Kappa de Cohen $> 0,82$ para cada categoría), tanto para la clasificación de la polaridad de los tuits como para la clasificación de la polaridad de las palabras del lexicón. En ambos casos, los dos investigadores discutieron las discrepancias en las clasificaciones hasta alcanzar un acuerdo.

Seleccionamos mediante **muestreo aleatorio simple** 5 000 tuits de cada polaridad de la muestra codificada manualmente. Dividimos

estos datos en un conjunto de entrenamiento (3 500 tuits de cada clase) y un conjunto de prueba (2 500 tuits de cada clase).

Utilizamos el **conjunto de entrenamiento** y el **lexicón** para entrenar un clasificador automático **MSV** de la polaridad de sentimientos respecto a las vacunas de los tuits.

Evaluamos el **rendimiento** del algoritmo en el **conjunto de prueba**. Corregimos los errores identificados y llevamos a cabo el proceso de forma iterativa hasta que el algoritmo alcanzó una precisión $> 0,80$, una exhaustividad $> 0,80$ y una puntuación F1 $> 0,85$ en la clasificación de cada una de las tres categorías.

Finalmente, el **algoritmo entrenado** clasificó la polaridad de sentimientos frente a las vacunas de todo el conjunto de datos (1 499 227 tuits).

Adicionalmente, clasificamos el conjunto de datos mediante **VADER**. Calculamos sus métricas de rendimiento en el conjunto de datos de prueba y las comparamos con las de nuestro algoritmo.

XIV. CLASIFICACIÓN DE TUI TS POR TEMAS

Dos investigadores realizaron independientemente una **clasificación multitemática** de los **hashtags** de la base de datos de *hashtags* creada antes del preprocesamiento y de los unigramas y bigramas a nivel de palabra más frecuentes. Resolvieron las discrepancias mediante discusión hasta que se llegó a un consenso. La **confiabilidad intercodificador** fue alta (puntuación Kappa de Cohen $> 0,80$ para cada tema).

Utilizamos los **términos clasificados manualmente** por temas como atributos de entrada para **entrenar** mediante las herramientas de

NLTK un clasificador temático multicategoría basado en **diccionarios**.

Obtuvimos un **algoritmo automático** que clasificaba los tuits en categorías temáticas no excluyentes (Tabla 7) según los n - gramas y los *hashtags* presentes en el texto. Aplicamos el algoritmo para clasificar todo el conjunto de datos (1 499 227 tuits).

Tabla 7. Temas	
Categoría	Descripción
Campañas a favor o en contra de las vacunas	
Campañas	Tuits que promocionan la vacunación o animan a no vacunarse.
Determinantes de reticencia vacunal	
Accesibilidad	Tuits sobre las barreras para acceder a las vacunas.
Ciencia	Tuits que contienen comentarios sobre la evidencia científica disponible sobre las vacunas.
Confianza	Tuits sobre la confianza o desconfianza en las vacunas y conspiraciones sobre la vacunación.
Efectividad	Tuits sobre la efectividad de las vacunas para prevenir enfermedades.
Legislación y autonomía	Tuits sobre aspectos políticos, éticos y legales de la obligatoriedad de la vacunación.
Religión / cultura	Tuits sobre aspectos religiosos o socioculturales que influyen en la vacunación.
Seguridad	Tuits sobre el riesgo de vacunar/no vacunar o la morbimortalidad prevenida o atribuida a la vacunación.
Miscelánea	Tuits que no cumplen con los criterios para ser etiquetados como "accesibilidad", "ciencia", "confianza", "efectividad", "legislación y autonomía", "religión / cultura" o "seguridad". Incluye tuits generales que hacen referencia a la vacunación (por ejemplo, "vacúnate") y tuits sobre temas minoritarios (<10 000 tuits) como el coste o el calendario vacunal.

Tabla 7. Temas

Categoría	Descripción
Población diana de la vacunación	
Adultos	Tuits sobre la vacunación del adulto, excepto embarazadas y viajeros.
Embarazadas	Tuits sobre la vacunación de embarazadas.
Niños	Tuits sobre la vacunación infantil.
Viajeros	Tuits sobre la vacunación del viajero.
Tipo de vacuna	
Cólera	Tuits sobre la vacuna contra el cólera.
Ébola	Tuits sobre la vacuna contra el ébola.
Fiebre amarilla	Tuits sobre la vacuna de la fiebre amarilla.
Fiebre tifoidea	Tuits sobre la vacuna de la fiebre tifoidea.
Influenza	Tuits sobre la vacuna de la gripe.
Hepatitis	Tuits sobre la vacuna contra la hepatitis A, B o C (experimental).
Neumococo	Tuits sobre la vacuna contra el neumococo.
Meningitis	Tuits sobre la vacuna contra la meningitis.
Polio	Tuits sobre la vacuna contra la polio.
Rabia	Tuits sobre la vacuna contra la rabia.
Rotavirus	Tuits sobre la vacuna contra el rotavirus.
SRP	Tuits sobre la vacuna SRP.
Tosferina	Tuits sobre cualquier vacuna con el componente <i>pertussis</i> .
Tuberculosis	Tuits sobre la vacuna contra la tuberculosis.
Varicela	Tuits sobre la vacuna contra la varicela.
VIH	Tuits sobre la vacuna contra el VIH (experimental).
Viruela	Tuits sobre la vacuna contra la viruela.
VPH	Tuits sobre la vacuna contra el VPH.

Tabla 7. Temas

Categoría	Descripción
Conceptos relacionados con la reticencia vacunal	
Autismo	Tuits sobre vacunas y autismo.
Cáncer	Tuits sobre vacunas y cáncer.
Componentes	Tuits sobre los componentes de las vacunas, tanto reales como atribuidos por los usuarios.
Homeopatía	Tuits sobre vacunas y productos homeopáticos.
Instituciones públicas o privadas mencionados	
Agencias de salud pública	Tuits de vacunas que mencionen alguna agencia de salud pública.
Gobierno	Tuits de vacunas que mencionen algún órgano de gobierno, excluyendo los dedicados a la salud pública.
Industria farmacéutica	Tuits de vacunas que mencionen a la industria farmacéutica.
Instituciones sanitarias asistenciales	Tuits de vacunas que mencionen hospitales o centros de salud.

XV. CLASIFICACIÓN DE LOS USUARIOS

Confeccionamos una lista de **cuentas institucionales** conocidas y la comparamos con nuestra lista de usuarios. Acto seguido, revisamos las cuentas no identificadas como institucionales en el primer cribado y las clasificamos en “institucionales” o “no institucionales” según el nombre del usuario, los datos del perfil y el tipo de tuits publicados.

De cada **idioma**, seleccionamos los **10 usuarios** con mayor número absoluto de tuits sin distinción de polaridad, los 10 usuarios con mayor número absoluto de tuits positivos y los 10 usuarios con mayor número absoluto de tuits negativos

Examinamos estas cuentas con **Botometer**. Decidimos que para clasificar una cuenta como *bot* debía cumplir al menos uno de los siguientes criterios: a) puntuación de *bot* autodeclarado igual o mayor a 4,6, b) puntuación de *bot* general mayor o igual a 4,8, equivalente a un CAP del 92%. Todas las cuentas que clasificamos como *bots* cumplían los dos criterios.

Revisamos las cuentas que no fueron detectadas como *bots* y las subclasificamos según el **tipo de usuario** en las siguientes categorías excluyentes: “asociación profesional o sociedad científica”, “asociación de usuarios”, “medio de comunicación”, “padre o madre”, “profesional sanitario”, “otro_no institucional” y “otras instituciones”.

En esta última categoría agrupamos a los usuarios clasificados como “organismo gubernamental”, “organización no gubernamental”, “institución académica”, “institución sanitaria”, “empresa farmacéutica” y “otras instituciones”, ya que detectamos 2 o menos cuentas de cada uno de estos grupos.

Finalmente, anonimizamos los números de identificación de los usuarios mediante **enmascaramiento** de datos.

XVI. VARIABLES DEL ESTUDIO

La base de datos final contiene las variables de la Tabla 8.

Tabla 8. Variables del estudio			
Variable	Descripción	Tipo de variable	Valores
N.º del tuit	Número de identificación del tuit enmascarado.	Cualitativa nominal no dicotómica.	Secuencia única de caracteres (números y letras) aleatorios.
N.º de usuario	Número de identificación enmascarado del usuario autor del tuit.	Cualitativa nominal no dicotómica.	Secuencia única de caracteres (números y letras) aleatorios.
Tipo de usuario	Clasificación del usuario autor del tuit según su cuenta sea o no institucional.	Cualitativa nominal dicotómica.	{institucional, no_institucional}
Año	Año de publicación del tuit.	Cuantitativa discreta.	{2011 – 2019}
Mes	Mes de publicación del tuit.	Cualitativa nominal no dicotómica.	{enero – diciembre}
Día semana	Día de la semana en el que se publicó el tuit.	Cualitativa nominal no dicotómica.	{lunes- domingo}
Idioma	Idioma de tuit.	Cualitativa nominal dicotómica.	{castellano, inglés}
Localización	Localización del autor del tuit en mayo de 2019.	Cualitativa nominal no dicotómica.	Países. “sin loc.” para usuarios sin localización.
Es_respuesta	Tipo de tuit: respuesta o no respuesta.	Cualitativa nominal dicotómica.	{sí, no}

Tabla 8. Variables del estudio

Variable	Descripción	Tipo de variable	Valores
N.º_retuits	Número de retuits del tuit en mayo de 2019.	Cuantitativa discreta.	{0 – ∞}
N.º_respuestas	Número de respuestas al tuit en mayo de 2019.	Cuantitativa discreta.	{0 – ∞}
N.º_me gusta	Número de “me gusta” del tuit en mayo de 2019.	Cuantitativa discreta.	{0 – ∞}
Polaridad	Polaridad del sentimiento frente a las vacunas del tuit.	Cualitativa nominal no dicotómica.	{negativo, neutro, positivo}
Temas	Una variable para cada uno de los temas de la Tabla 7. Indicamos si el tuit pertenece o no a estas categorías.	Cualitativa nominal dicotómica.	{sí, no}

Paralelamente, creamos una base de datos con los usuarios con mayor número absoluto de tuits totales, negativos y positivos en castellano y en inglés (Tabla 9).

Tabla 9. Variables del estudio de los usuarios con más tuits			
Variable	Descripción	Tipo de variable	Valores
N.º de usuario	Número de identificación enmascarado del usuario autor del tuit.	Cualitativa nominal no dicotómica.	Secuencia única de caracteres (números y letras) aleatorios.
Tipo de usuario	Clasificación del usuario según sea o no <i>bot</i> . Si no es <i>bot</i> , se clasifica en alguna de las otras categorías.	Cualitativa nominal no dicotómica.	{ <i>bot</i> , asociación profesional-científica, asociación de usuarios, medio de comunicación, padre o madre, profesional sanitario, otras instituciones, otro_no_institucional}
Seguidores	Número de seguidores del usuario en mayo de 2019.	Cuantitativa discreta.	{0 – ∞}
Idioma	Indicamos si el usuario está en la lista de autores con más número de tuits en castellano o en inglés.	Cualitativa nominal dicotómica.	{castellano, inglés}
Categoría	Indicamos si el usuario está en la lista de autores con más tuits totales, positivos o negativos, del idioma de la variable anterior.	Cualitativa nominal no dicotómica.	{total, positivo, negativo}
N.º tuits en castellano	Número de tuits en castellano escritos por el usuario.	Cuantitativa discreta.	{0 – ∞}
N.º tuits negativos castellano	Número de tuits negativos en castellano escritos por el usuario.	Cuantitativa discreta.	{0 – ∞}

Tabla 9. Variables del estudio de los usuarios con más tuits

Variable	Descripción	Tipo de variable	Valores
N.º tuits neutros castellano	Número de tuits neutros en castellano escritos por el usuario.	Cuantitativa discreta.	$\{0 - \infty\}$
N.º tuits positivos castellano	Número de tuits positivos en castellano escritos por el usuario.	Cuantitativa discreta.	$\{0 - \infty\}$
N.º tuits en inglés	Número de tuits en inglés escritos por el usuario.	Cuantitativa discreta.	$\{0 - \infty\}$
N.º tuits negativos inglés	Número de tuits negativos en inglés escritos por el usuario.	Cuantitativa discreta.	$\{0 - \infty\}$
N.º tuits neutros castellano	Número de tuits neutros en inglés escritos por el usuario.	Cuantitativa discreta.	$\{0 - \infty\}$
N.º tuits positivos inglés	Número de tuits positivos en inglés escritos por el usuario.	Cuantitativa discreta.	$\{0 - \infty\}$

VII. ANÁLISIS ESTADÍSTICO

Describimos las **variables cualitativas** mediante frecuencias absolutas y relativas. Las representamos mediante diagramas de barras y diagramas de sectores. Creamos tablas de contingencia a partir de los datos y usamos la prueba Z para comparar las proporciones.

Comprobamos que las **variables cuantitativas** no presentaban una distribución normal mediante la prueba de Kolmogórov-Smirnov. Calculamos la mediana como medida de tendencia central y los

percentiles 25 y 75 y el rango intercuartílico (RI) como medidas de dispersión. Representamos estos datos mediante gráficos de cajas y bigotes.

Comparamos las medianas mediante la prueba de Kruskal-Wallis. Estudiamos el grado de asociación entre las variables cuantitativas mediante el coeficiente de correlación de Spearman.

Empleamos gráficos de **serie temporal** para visualizar los cambios en el número de tuits (total, por polaridad o por temas) desde el 1 de junio de 2011 hasta el 30 de abril de 2019.

Para evaluar la tendencia de cada serie, ajustamos el componente ciclo-tendencia a funciones lineales, logarítmicas, cuadráticas y exponenciales. Seleccionamos el modelo de regresión con mejor ajuste según el coeficiente de determinación (R^2) ajustado y el p-valor de la prueba F. Calculamos el intervalo de predicción al 95% del mejor modelo.

Evaluamos la estacionariedad de las series temporales mediante funciones de autocorrelación y el grado de relación lineal entre las series mediante funciones de correlación cruzada.

Realizamos el análisis estadístico mediante el paquete estadístico *Statistical Package for Social Sciences (SPSS)* para Windows versión 23.0 (IBM Corp., Armonk, NY, EE. UU.).

En todos los casos, planteamos contrastes bilaterales y consideramos como estadísticamente significativos los valores de **p** menores de 0,05.

XVIII. CONSIDERACIONES ÉTICAS

El proyecto de investigación recibió la aprobación del Comité de Ética en Investigación Experimental de la Universidad de Valencia (Anexo).

Twitter dispone de herramientas para que cualquier persona descargue y analice la información publicada en su plataforma y, al darse de alta, los usuarios aceptan la política de privacidad, en la que autorizan la visualización y el análisis de la información por terceros.

Adicionalmente, los usuarios tienen la opción de proteger sus mensajes para que sólo puedan ser vistos por sus seguidores. Todos los mensajes incluidos en este estudio son públicos y, por lo tanto, visibles para cualquier persona en la red.

Dado el ámbito y amplitud del estudio, no es posible contactar con los 278 858 autores de todos los mensajes incluidos en este estudio para informarles ni solicitar su consentimiento.

Con el fin de asegurar que los usuarios no puedan ser identificados, presentamos los resultados agregados y anonimizados. No mencionamos los nombres ni los números de identificación de los usuarios. Tampoco reproducimos total ni parcialmente el contenido de los tuits en este trabajo ni en las publicaciones derivadas. Sólo la investigadora principal tiene acceso a la base de datos generada tras la descarga de datos y que contiene los tuits sin anonimizar.

CAPÍTULO IV

RESULTADOS

I. ANÁLISIS GLOBAL DE LOS TUIITS

A. Características Generales de los Tuits

Incluimos en el estudio 1 499 227 tuits originales (no retuits), de los cuales 159 628 (10,65%) tuits eran respuestas a otros tuits y 1 339 599 (89,35%) no eran respuestas. 1 295 823 (86,43%) tuits estaban escritos en inglés, mientras que los 203 404 (13,57%) tuits restantes estaban en castellano (Figura 21).

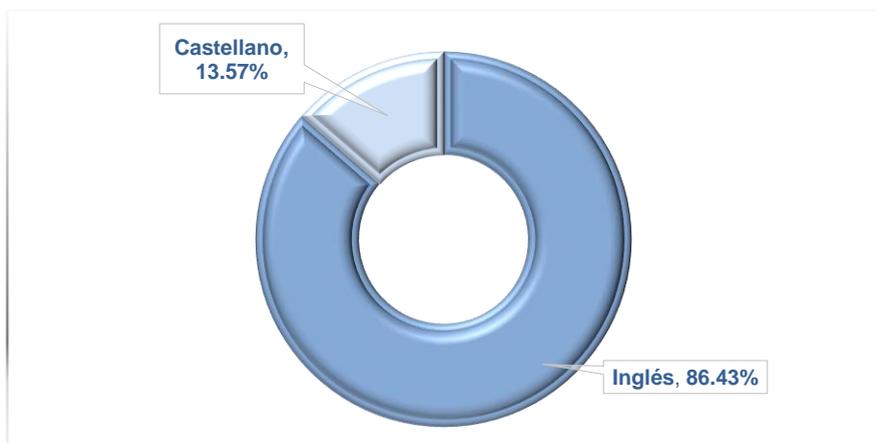
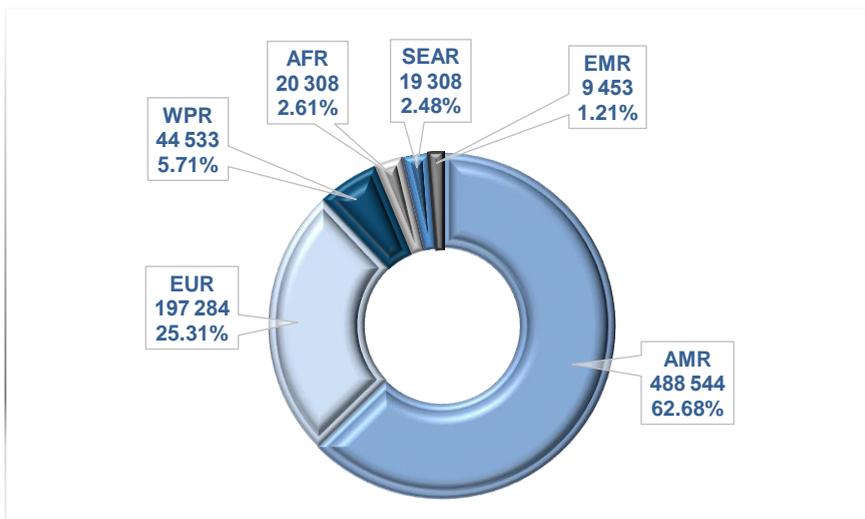


Figura 21. Tuits por idioma.

Los tuits recopilados fueron escritos por 278 858 usuarios. Etiquetamos 134 207 (48,13%) usuarios como “sin loc.”. Geocalizamos 144 651 (51,87%) usuarios, autores de 779 430 (51,99%) tuits sobre vacunas de nuestra base de datos.

El 62,68% de los tuits geocalizados procedían de la Región de las Américas de la OMS (AMR) y el 25,31% de la Región de Europa (EUR) (Figura 22).



Nota: Regiones de la OMS: Región de África (AFR), Región de las Américas (AMR), Región del Mediterráneo Oriental (EMR), Región de Europa (EUR), Región de Asia Suroriental (SEAR) y Región del Pacífico Occidental (WPR).

Figura 22. Frecuencia de tuits en las regiones de la OMS.

Identificamos usuarios de todos los países de la Organización de las Naciones Unidas (ONU) excepto Georgia, Libia y Surinam. Los cinco países con más tuits sobre vacunas fueron EE. UU. (372 457 tuits, 24,84%), Reino Unido (82 268 tuits, 5,49%), España (45 101 tuits, 3,01%), Canadá (44 215 tuits, 2,95%) y Australia (20 315 tuits, 1,36%).

B. Evolución en el Número de Tuits (junio 2011 - abril 2019)

En la Figura 23 representamos el número absoluto mensual de tuits, total y por idiomas, desde junio de 2011 hasta abril de 2019. En el primer mes del periodo analizado (junio de 2011), identificamos 2 700 tuits sobre vacunas.

Desde junio de 2011 hasta abril de 2015, el número mensual de tuits presentó una tendencia creciente ($p < 0,01$, $R^2 = 0,77$, modelo exponencial), con un pico de 57 544 tuits en febrero de 2015. A continuación, hubo una tendencia decreciente hasta diciembre de 2015 ($p < 0,01$, $R^2 = 0,81$, modelo exponencial). Posteriormente, hasta diciembre de 2018, el número mensual de tuits osciló entre 13 569 y 22 296, con picos cada abril.

Desde diciembre de 2018 hasta abril de 2019, observamos un aumento medio de 8 952 tuits por mes (IC 95% 2 632-15 272 tuits por mes; tendencia creciente, $p < 0,01$, $R^2 = 0,99$, modelo exponencial), hasta alcanzar el pico más alto de tuits (57 667 tuits) en el último mes del estudio (abril de 2019).

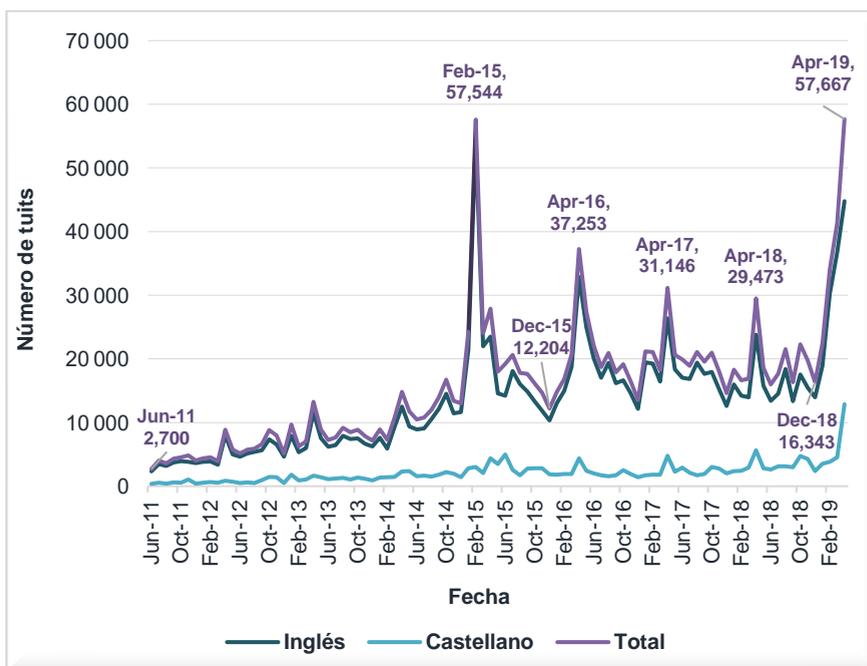


Figura 23. Número de tuits por idioma (junio 2011 – abril 2019).

II. ANÁLISIS SEMÁNTICO

A. Unigramas a Nivel de Palabra

Las 30 palabras más frecuentes en los tuits en inglés y las 30 palabras más frecuentes en los tuits en castellano se representan en la Figura 24. En el análisis excluimos las palabras vacías y los términos de búsqueda.

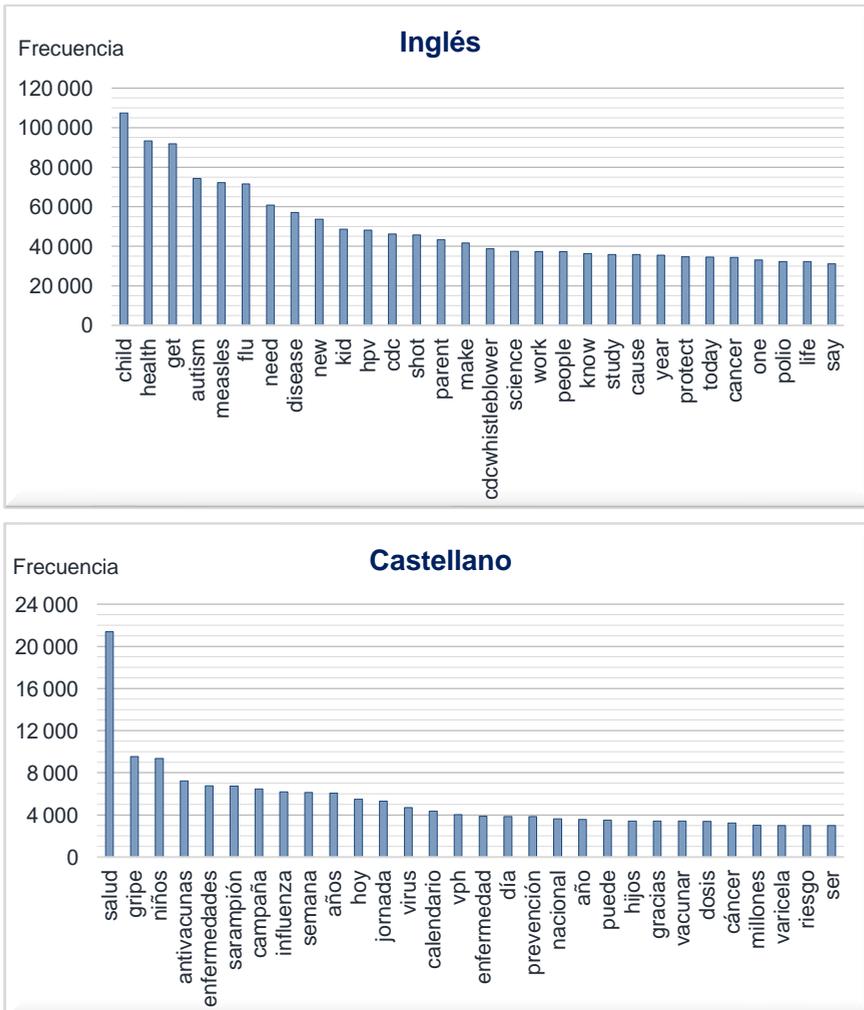


Figura 24. 30 palabras más frecuentes por idioma.

B. Bigramas a Nivel de Palabra

En la Figura 25 representamos los 30 bigramas a nivel de palabra más frecuentes en ambos idiomas. En el análisis excluimos las palabras vacías.

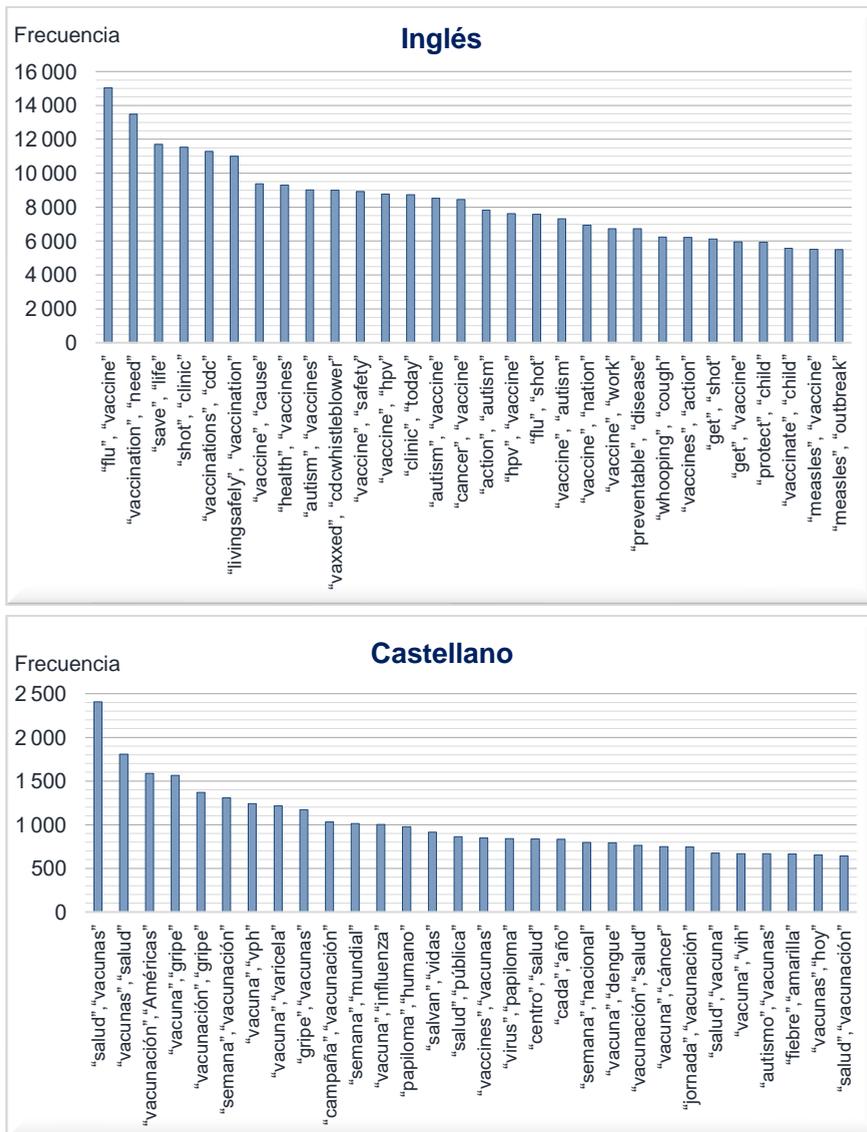


Figura 25. 30 bigramas más frecuentes por idioma.

III. ANÁLISIS DE LA POLARIDAD DE SENTIMIENTOS

A. Métricas y Comparación con el AS de Vader

El algoritmo Vader clasificó 1 120 069 (74,71%) tuits como neutros, 203 584 (13,58%) tuits como positivos y 175 574 (11,71%) tuits como negativos. El algoritmo Vader y nuestro algoritmo clasificaron 947 858 (63,22%) tuits en las mismas categorías y difirieron en la clasificación de 551 369 (36,78%) tuits.

Evaluamos en el conjunto de datos de prueba, la precisión, exhaustividad y la puntuación F1 de ambos algoritmos para la clasificación de la polaridad de sentimientos de los tuits, global y por categorías (Tabla 10).

Tabla 10. Métricas de los algoritmos de clasificación		
Métrica	Algoritmo	
	Vader	Híbrido
Clasificación de tuits negativos		
Precisión	0,116	0,809
Exhaustividad	0,274	0,939
Puntuación F1	0,163	0,869
Clasificación de tuits neutros		
Precisión	0,686	0,902
Exhaustividad	0,881	0,957
Puntuación F1	0,771	0,929

Tabla 10. Métricas de los algoritmos de clasificación				
Métrica	Algoritmo			
	Vader		Híbrido	
Clasificación de tuits positivos				
Precisión	0,949		0,977	
Exhaustividad	0,580		0,882	
Puntuación F1	0,720		0,927	
Global	Media macro	Media micro	Media macro	Media micro
Precisión	0,584	0,689	0,896	0,921
Exhaustividad	0,578	0,689	0,926	0,921
Puntuación F1	0,551	0,689	0,908	0,921

B. Polaridad de Sentimientos por Idioma

Nuestro algoritmo clasificó 1 039 864 (69,36%) tuits como neutros, 326 497 (21,78%) tuits como positivos y 132 866 (8,86%) tuits como negativos. Evidenciamos diferencias estadísticamente significativas en la distribución de las tres categorías de polaridad según el idioma (Figura 26):

- 👉 De los 203 404 tuits escritos en castellano: 160 883 (79,10%) eran neutros, 40 476 (19,90%) positivos y 2 045 (1,01%) negativos.
- 👉 De los 1 295 823 tuits escritos en inglés, 878 981 (67,83%) eran neutros, 286 021 (22,07%) positivos y 130 821 (10,10%) negativos.

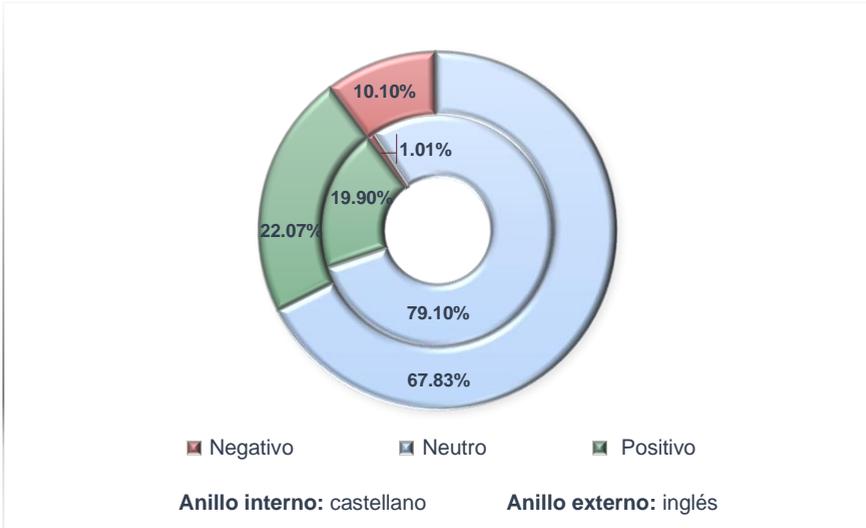


Figura 26. Porcentaje de tuits por polaridad e idioma.

C. Evolución Temporal de la Polaridad de Sentimientos

El número mensual de tuits por polaridad de sentimientos a lo largo del periodo de estudio se muestra en la Figura 27.

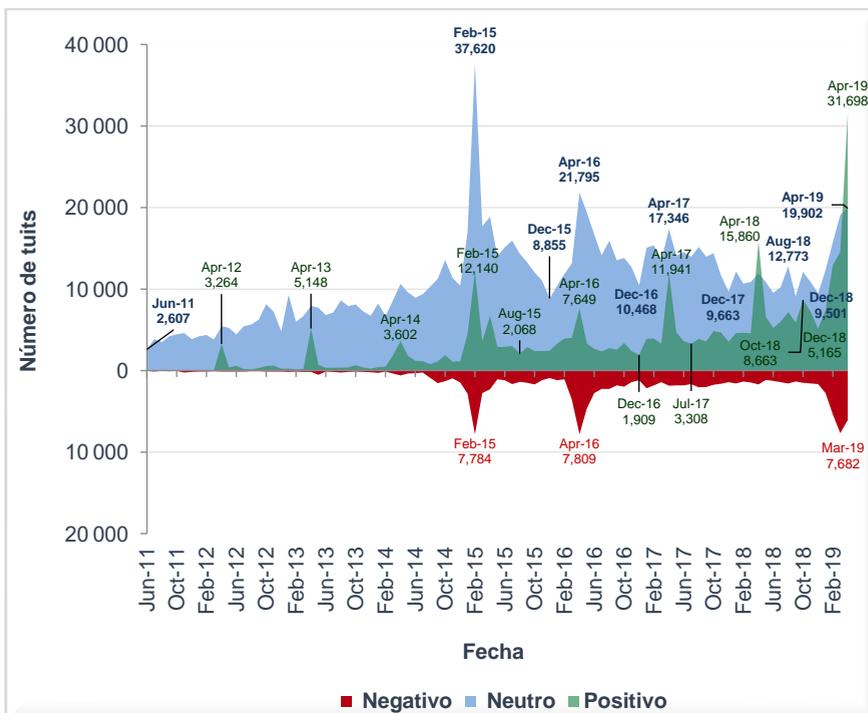


Figura 27. Número mensual de tuits por polaridad de sentimiento (junio 2011 - abril de 2019).

El número de tuits neutros fue mayor que el de tuits positivos durante todo el periodo de estudio, excepto en abril de 2018 y abril de 2019. El número de tuits negativos siempre fue inferior al de los tuits neutros.

Los tuits negativos fueron más frecuentes que los tuits positivos en los siguientes periodos:

-  Julio de 2011 - marzo de 2012 (diferencia de 13 - 201 tuits).
-  Agosto y septiembre de 2014 (diferencia de 127 y 424 tuits respectivamente).
-  Diciembre de 2014 (325 tuits más).
-  Abril - junio de 2016 (diferencia de 81 - 1 371 tuits).

La frecuencia relativa de tuits neutros (Figura 28 a) tuvo una tendencia decreciente en el período de estudio ($p < 0,01$, $R^2 = 0,75$, modelo lineal). En el primer mes de estudio (junio de 2011) los tuits neutros constituían el 96,55% de la muestra, mientras que en el último mes de estudio (abril de 2019) descendieron hasta un mínimo de 34,51%.

El porcentaje de tuits positivos (Figura 28 b) fue del 1,81% al inicio del estudio y ascendió a su máximo (54,97%) en abril de 2019, con una tendencia global creciente ($p < 0,01$, $R^2 = 0,69$, modelo exponencial).

El porcentaje de tuits negativos (Figura 28 c) también presentó una tendencia global creciente ($p < 0,01$, $R^2 = 0,62$, modelo exponencial). El 1,63% de los tuits eran negativos en junio de 2011, mientras que el 18,59% y el 10,52% de los tuits eran negativos en marzo y abril de 2019 respectivamente. El máximo porcentaje de tuits negativos se produjo en abril de 2016 (20,96%).

Resultados

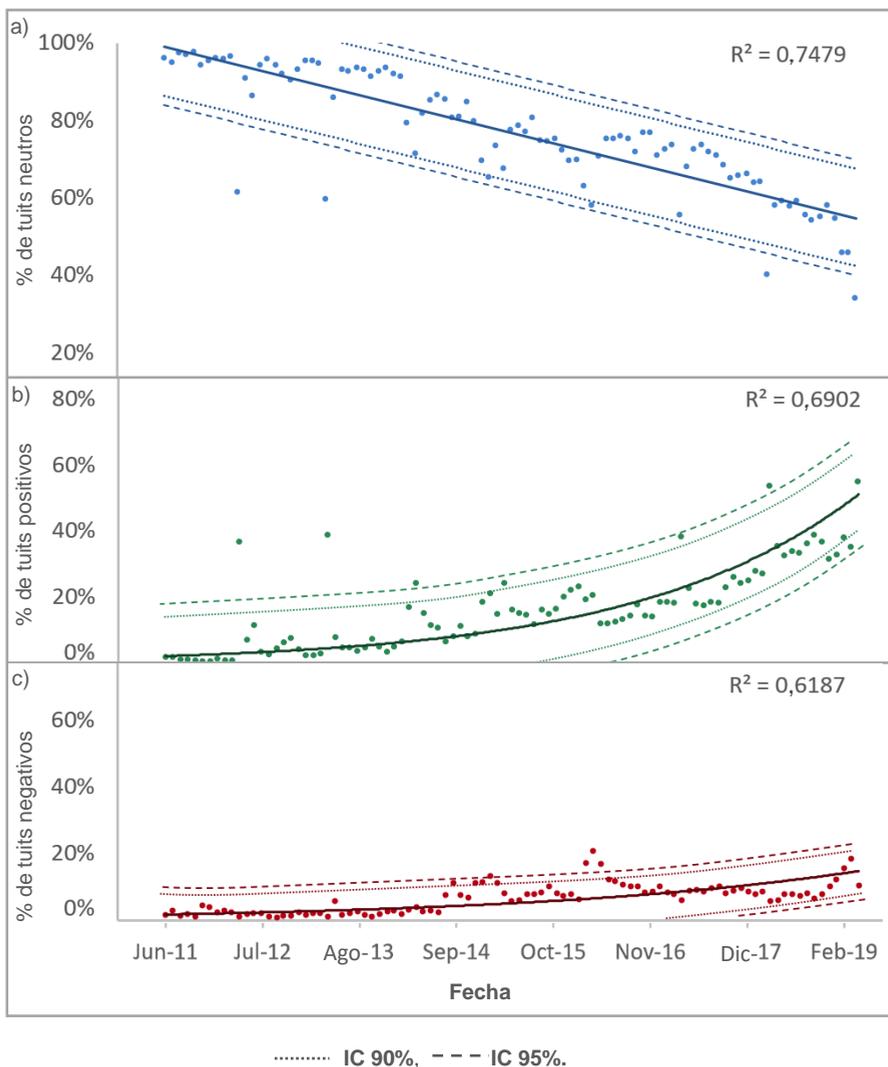


Figura 28. Frecuencia relativa de los tuits clasificados por polaridad de sentimiento frente a las vacunas (junio 2011 - abril 2019).

D. Polaridad de Sentimientos por Mes

1. Frecuencia absoluta mensual de tuits por polaridad

La frecuencia absoluta de tuits por meses presentaba una distribución distinta a la normal. Observamos una alta dispersión de los datos (Figura 29). La mediana de tuits, global y por categorías de polaridad de sentimiento, no presentaba diferencias estadísticamente significativas según los meses del año.

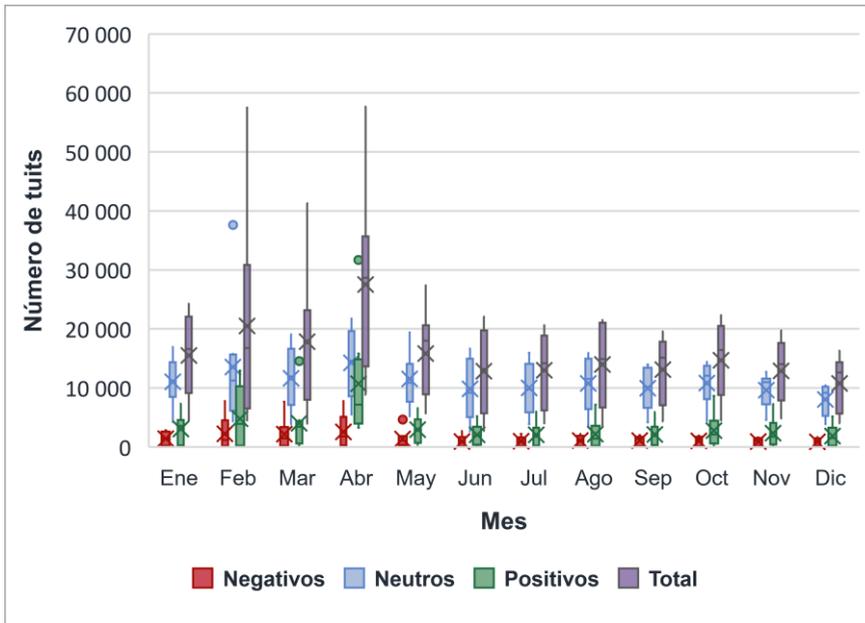


Figura 29. Número de tuits por polaridad de sentimiento y mes.

2. Frecuencia relativa mensual de tuits por polaridad

El porcentaje mensual de tuits neutros fue significativamente menor en abril (56,27%) en comparación con el resto de los meses, mientras que el porcentaje de tuits positivos fue significativamente mayor (36,47%). No hubo diferencias estadísticamente significativas en el porcentaje mensual de tuits negativos (Figura 30).

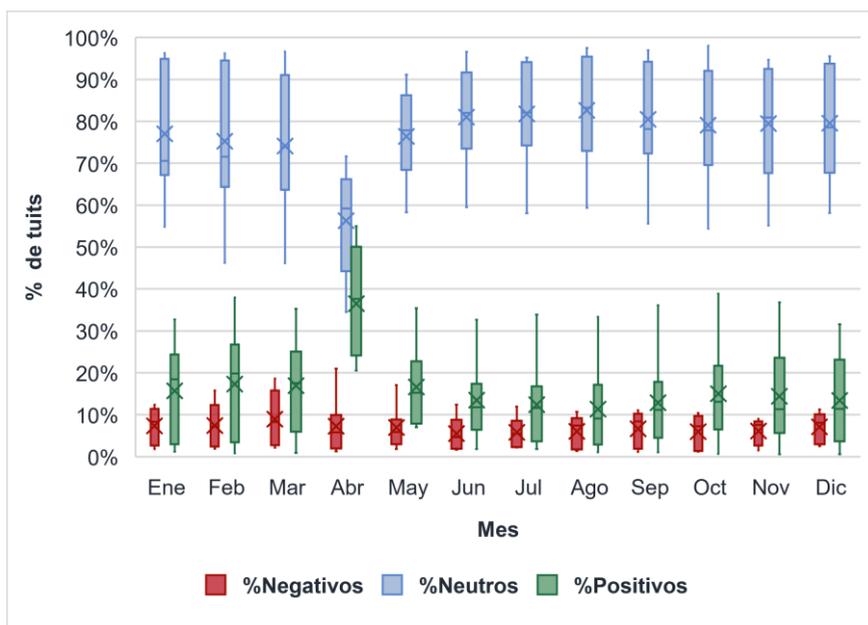


Figura 30. Porcentaje de tuits por polaridad y mes.

E. Polaridad de Sentimientos por Día de la Semana

1. Frecuencia absoluta de tuits por polaridad y día de la semana

En la comparación por pares, la mediana del total de tuits fue significativamente mayor ($p < 0,05$) los lunes, martes, miércoles, jueves y viernes que los sábados y domingos (Figura 31). Lo mismo ocurre con la mediana de tuits positivos, negativos y neutros.

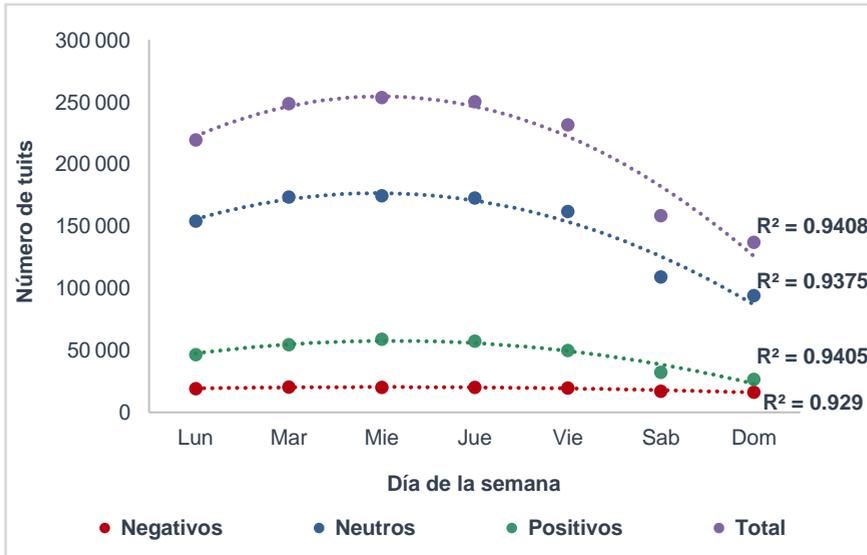


Figura 31. Porcentaje de tuits por polaridad y día de la semana.

2. Frecuencia relativa de tuits por polaridad y día de la semana

La proporción de tuits **negativos** fue significativamente menor los miércoles ($P50 = 6,39\%$, $RI = 2,60\% - 8,89\%$) y los jueves ($P50 = 6,29\%$, $RI = 2,56\% - 9,00\%$) y mayor los sábados ($P50 = 7,99\%$, $RI = 3,48\% - 11,06\%$) y los domingos ($P50 = 9,09\%$, $RI = 4,44\% - 12,35\%$).

Por el contrario, la proporción de tuits **positivos** fue significativamente mayor los miércoles (P50 = 15,31%, RI = 6,75% - 23,30%) y los jueves (P50 = 15,37%, RI = 7,31% - 23,40%) y menor los sábados (P50 = 13,04%, RI = 5,50% - 22,52%) y los domingos (P50 = 12,66%, RI = 4,44% - 21,81%). La proporción de tuits **neutros** osciló entre el 68,66% y el 70,14%, sin variaciones significativas entre semana respecto al fin de semana (p = 0,07).

F. Polaridad de Sentimientos por Países

Las dos **regiones** con mayor porcentaje de tuits negativos fueron AMR (9,22%) y WPR (7,27%). La negatividad en el resto de las regiones osciló entre 2,20% - 3,60%. El porcentaje de tuits positivos se situó entre el 20 - 30%, sólo superado en AFR (43,99%) (Figura 32).

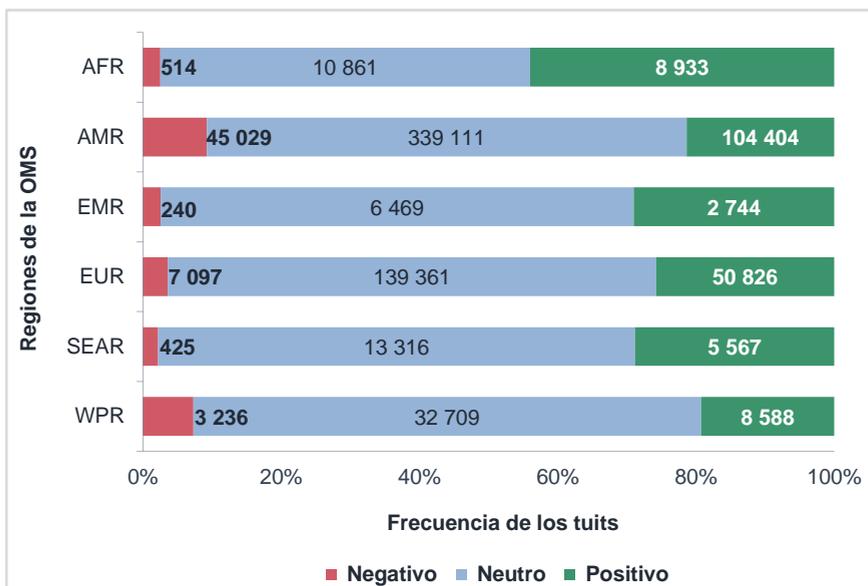


Figura 32. Frecuencia de tuits en las regiones de la OMS.

Identificamos reticencia vacunal en 165 países de la ONU. De los 40 **países** de la ONU con un número igual o mayor a 1 000 tuits:

-  Las localizaciones con mayor porcentaje de tuits **negativos** fueron Países Bajos ($n_{\text{negativos}} = 897$, 15,53%), Canadá ($n_{\text{negativos}} = 5\ 005$, 11,32%), Japón ($n_{\text{negativos}} = 1\ 007$, 10,74%) y EE. UU. ($n_{\text{negativos}} = 39\ 064$, 10,49%).
-  La proporción de tuits **positivos** fue mayor en Suiza ($n_{\text{positivos}} = 11\ 621$, 71,43%), Kenia ($n_{\text{positivos}} = 2\ 030$, 53,55%), Nigeria ($n_{\text{positivos}} = 3\ 765$, 42,46%), Irlanda ($n_{\text{positivos}} = 3\ 594$, 43,10%), Puerto Rico ($n_{\text{positivos}} = 850$, 39,89%), Pakistán ($n_{\text{positivos}} = 671$, 36,87%), Bélgica ($n_{\text{positivos}} = 3\ 429$, 34,95%) y Turquía ($n_{\text{positivos}} = 824$, 34,28%).

A continuación, presentamos los mapas con los países con 1 000 o más tuits coloreados según el porcentaje de polaridad negativa (Figura 33) y positiva (Figura 34).

Resultados

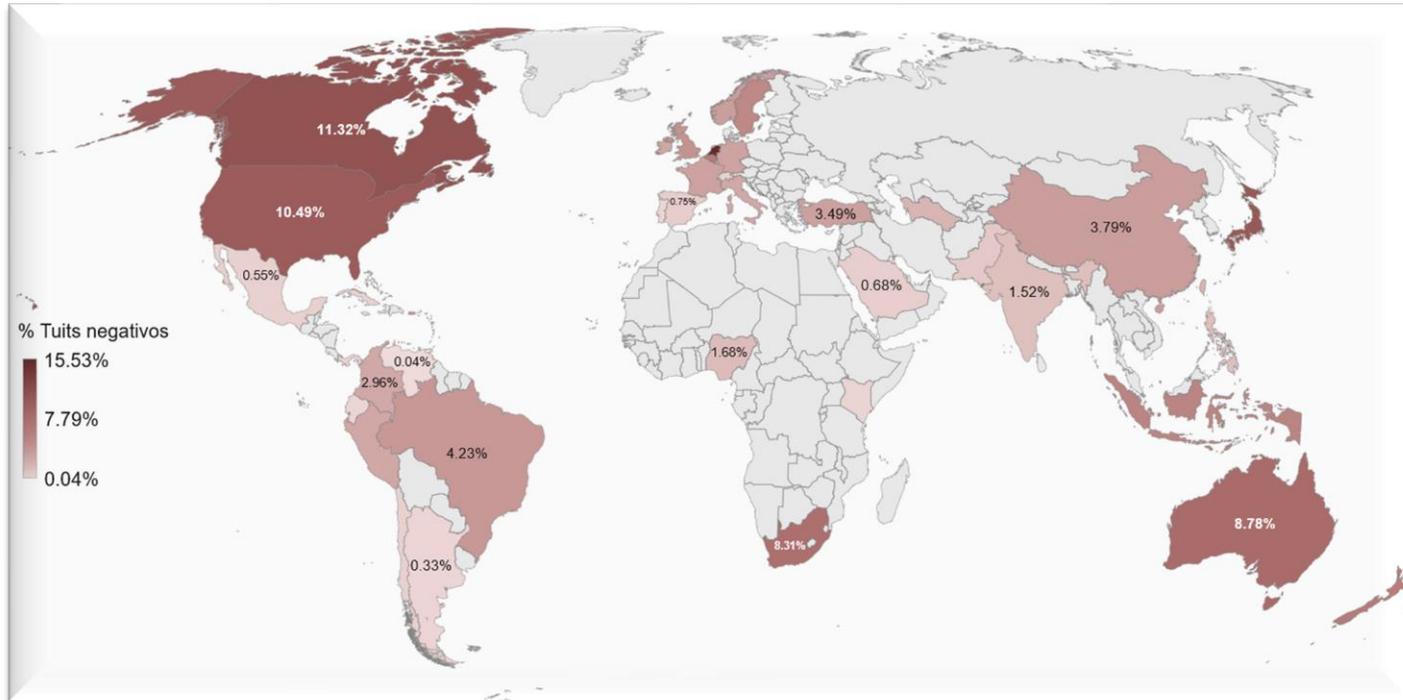


Figura 33. Porcentaje de tuits negativos (países $\geq 1\ 000$ tuits).

Valoración de las vacunas desde una plataforma social con aplicaciones de uso en promoción de la salud.

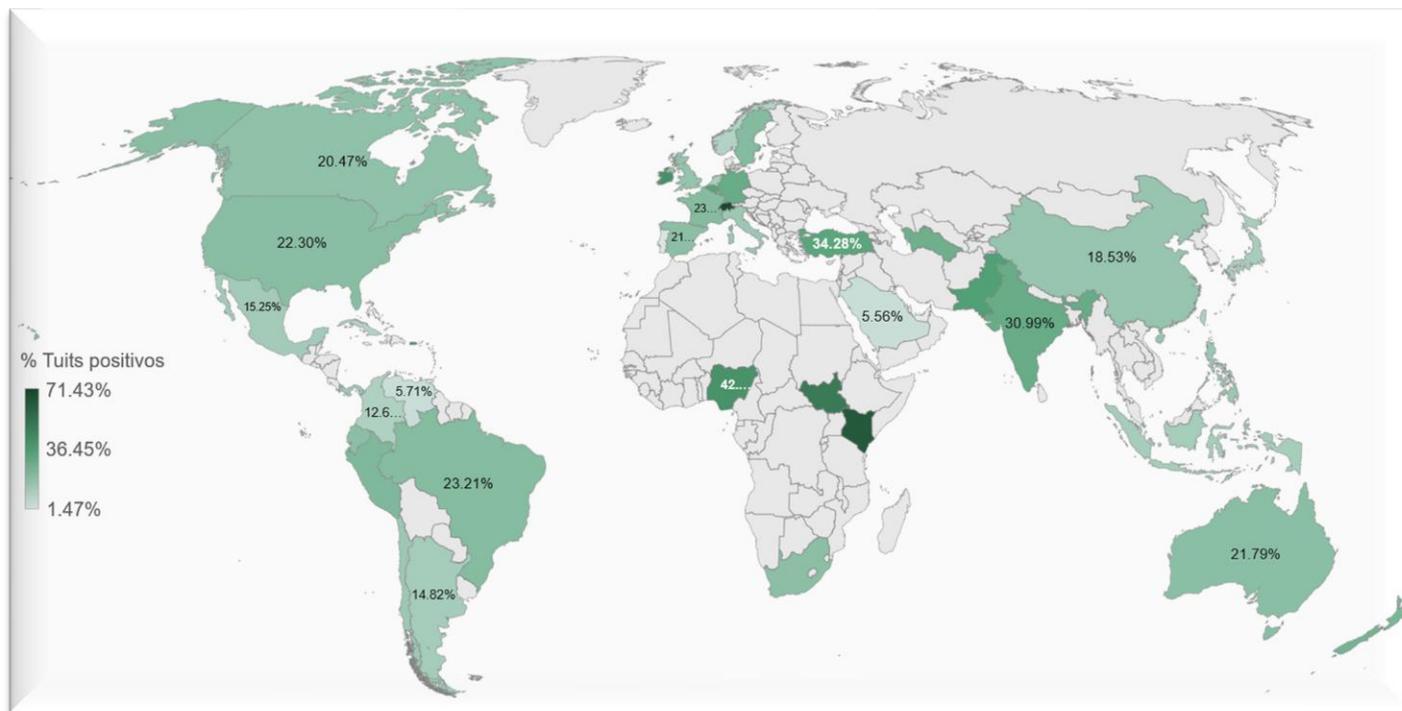


Figura 34. Porcentaje de tuits positius (països $\geq 1\ 000$ tuits).

G. Interacción Social Generada según la Polaridad de los Tuits

Las variables “número de retuits”, “número de favoritos/me gusta” y “número de respuestas” presentaron una distribución no normal. No hubo diferencias estadísticamente significativas en la mediana global de los tres elementos por categorías de polaridad de sentimiento (Tabla 11).

Tabla 11. Estadísticos descriptivos de los retuits, favoritos/me gusta y respuestas según la polaridad de sentimiento de los tuits						
	Media	P25	P50	P75	Mínimo	Máximo
Retuits						
Negativo	2,17	0	0	2	0	1 730
Neutro	1,37	0	0	1	0	16 734
Positivo	4,99	0	0	2	0	37 351
Total	2,23	0	0	1	0	37 351
Favoritos/me gusta						
Negativo	2,34	0	0	2	0	7 608
Neutro	1,52	0	0	1	0	13 143
Positivo	7,97	0	1	3	0	36 170
Total	2,99	0	0	1	0	36 170
Respuestas						
Negativo	0,39	0	0	0	0	4 152
Neutro	0,19	0	0	0	0	1 663
Positivo	0,42	0	0	0	0	2 457
Total	0,26	0	0	0	0	4 152

El porcentaje de tuits que habían sido **retuiteados** al menos una vez y el porcentaje de tuits que habían sido marcados al menos una vez como **favoritos/me gusta** fueron mayores en el grupo de tuits positivos, seguido del grupo de los negativos y en último lugar los neutros. En cambio, el porcentaje de tuits negativos que recibió alguna **respuesta** fue mayor que el de los tuits positivos. A su vez, todos los porcentajes fueron menores en el grupo de tuits neutros. Todas las diferencias fueron estadísticamente significativas. Las proporciones de tuits **sin** retuits, favoritos/me gusta y respuestas fueron mayores que las proporciones de tuits con alguno de estos atributos en la mayoría de los grupos (Figura 35).

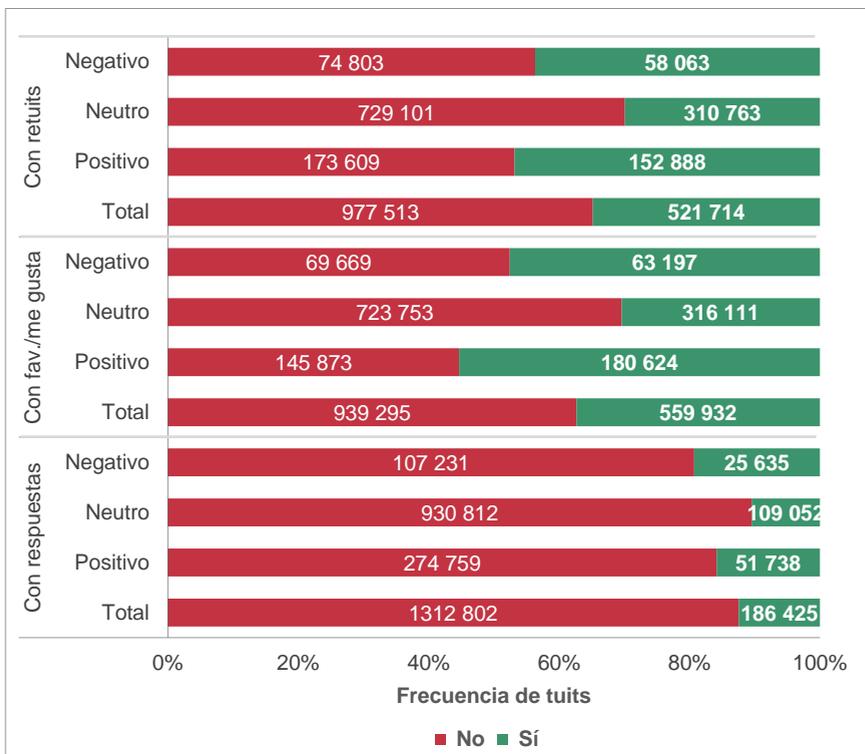


Figura 35. Porcentaje de tuits por polaridad de sentimiento según tengan o no retuits, favoritos/ me gusta y respuestas.

IV. ESTUDIO DE LOS USUARIOS

A. Idioma de los Usuarios

De los 278 858 usuarios incluidos en la base de datos, 224 232 (80,41%) usuarios escribieron solamente mensajes en inglés, 47 580 (17,06%) usuarios escribieron solamente mensajes en castellano y 7 046 (2,53%) usuarios escribieron mensajes en castellano y en inglés.

B. Clasificación de los Usuarios

Identificamos 90 263 (32,37%) cuentas institucionales y 188 595 (67,63%) cuentas no institucionales. Los usuarios institucionales escribieron el 16,44% (n = 246 435) de los tuits recopilados.

C. Número de Tuits por Usuario

La distribución del número de publicaciones por usuario fue no normal. La mediana de publicaciones por usuario fue de 1, con un rango de 1 - 21 329 tuits. 210 057 (75,33%) usuarios escribieron sólo un tuit con los *hashtags* estudiados, mientras que 68 801 (24,67%) usuarios tuitearon dos o más mensajes sobre vacunas.

D. Polaridad de Sentimiento según el Usuario

De los 68 801 (24,67%) usuarios que escribieron dos o más tuits de la base de datos:

 El 91,82% (n = 63 181) de los usuarios expresaron la misma polaridad de sentimiento hacia la vacunación en todos sus tuits:

52 020 (75,60%) usuarios fueron totalmente neutros, 9 626 (13,99%) totalmente positivos y 1 535 (2,30%) totalmente negativos.

-  El 8,17% (5 620) de los usuarios escribieron mensajes con diversas polaridades.
-  El 6,92% (4 764) de los usuarios escribieron tanto tuits positivos como negativos respecto a las vacunas. Este grupo escribió el 31,55% (n = 472 959) de los tuits del estudio.

La proporción media de tuits positivos y neutros fue mayor en las cuentas institucionales (24,24 ± 13,52% y 71,01 ± 16,13% respectivamente) en comparación con las cuentas no institucionales (21,39 ± 15,83% y 69,03 ± 13,99% respectivamente), aunque las diferencias no fueron significativas (p = 0,16 y p = 0,25 respectivamente). La proporción media de tuits negativos fue significativamente mayor (p < 0,01) en las cuentas no institucionales (9,67 ± 2,72%) que en las cuentas institucionales (4,74 ± 1,93%).

E. Usuarios con más Tuits

Analizamos el conjunto de los 10 usuarios con más tuits en inglés y los 10 usuarios con más tuits en castellano, sin distinción de polaridad y por polaridades (positiva o negativa). Dividimos los usuarios en dos grupos en los que esperábamos encontrar diferencias en la polaridad hacia las vacunas:

-  **Grupo 1:** profesionales sanitarios y usuarios institucionales, exceptuando asociaciones no científicas (medio de comunicación, asociación profesional o sociedad científica, otras instituciones).

 **Grupo 2:** *bots*, asociaciones de usuarios y cuentas personales (padre o madre, otro_personal), exceptuando cuentas de profesionales sanitarios.

Los 10 usuarios con más tuits negativos en inglés pertenecían al grupo 2, mientras que 6 de los 10 usuarios con más tuits positivos en inglés pertenecían al grupo 1 (Tabla 12). Las diferencias entre grupos fueron significativas en los dos casos ($p < 0,01$).

Tabla 12. Clasificación de los usuarios con más tuits en inglés				
Tipo de usuario	N.º de usuarios con más tuits			Total
	Negativos	Positivos	Sin distinción de polaridad	
Grupo 1				
Medio de comunicación	0	1	1	2
Asociación profesional o sociedad científica	0	0	0	0
Otras instituciones	0	2	2	4
Profesional sanitario	0	3	1	4
Total usuarios grupo 1	0	6	4	10
Grupo 2				
<i>Bot</i>	2	1	2	5
Asociación de usuarios	1	0	0	1
Padre o madre	6	0	1	7
Otro_personal	1	3	3	7
Total usuarios grupo 2	10	4	6	20
Total	10	10	10	30

Los 10 usuarios con más tuits positivos pertenecían al grupo 1, mientras que 9 de los 10 usuarios con más tuits negativos pertenecían al grupo 2 (Tabla 13). Las diferencias entre grupos fueron significativas en los dos casos ($p < 0,01$).

Tabla 13. Clasificación de los usuarios con más tuits en castellano				
Tipo de usuario	N.º de usuarios con más tuits			
	Negativos	Positivos	Sin distinción de polaridad	Total
Grupo 1				
Medio de comunicación	0	0	3	3
Asociación profesional o sociedad científica	0	2	3	5
Otras instituciones	0	1	2	3
Profesional sanitario	1	7	2	10
Total usuarios grupo 1	1	10	10	21
Grupo 2				
<i>Bot</i>	1	0	0	1
Asociación de usuarios	3	0	0	3
Padre o madre	3	0	0	3
Otro_personal	2	0	0	2
Total usuarios grupo 2	9	0	0	9
Total	10	10	10	30

El número de seguidores osciló entre 0 – 732 200 000, sin diferencias significativas en la mediana de seguidores entre los usuarios con más tuits positivos y los usuarios con más tuits negativos ($p = 0,24$).

V. ANÁLISIS DE TEMAS

A. Campañas de Promoción de las Vacunas

Dividimos los datos en cuatro grupos: campañas de promoción de las vacunas, otros tuits positivos/neutros, campañas contra las vacunas y otros tuits reticentes. 949 557 (63,34%) tuits no pertenecían a campañas a favor ni en contra de las vacunas, de los cuales el 92,92% (n = 882 329 tuits) tenía una orientación neutra. 471 782 tuits (31,47%) formaban parte de campañas de promoción de las vacunas y 77 888 (5,19%) tuits formaban parte de campañas contra las vacunas.

La mayor parte de los tuits positivos (96,25%, n = 314 247 tuits) pertenecían a campañas de promoción de las vacunas, mientras que el 58,62% (n = 77 888) de los tuits negativos pertenecían a campañas contra las vacunas (Figura 36).

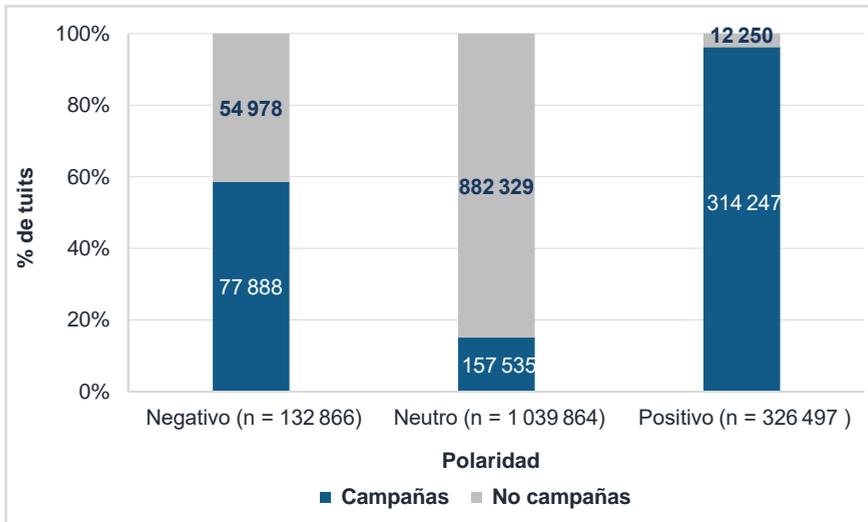


Figura 36. Porcentaje de tuits de cada polaridad según la variable “campañas”.

Estudiamos las series temporales de los cuatro grupos (Figura 37).

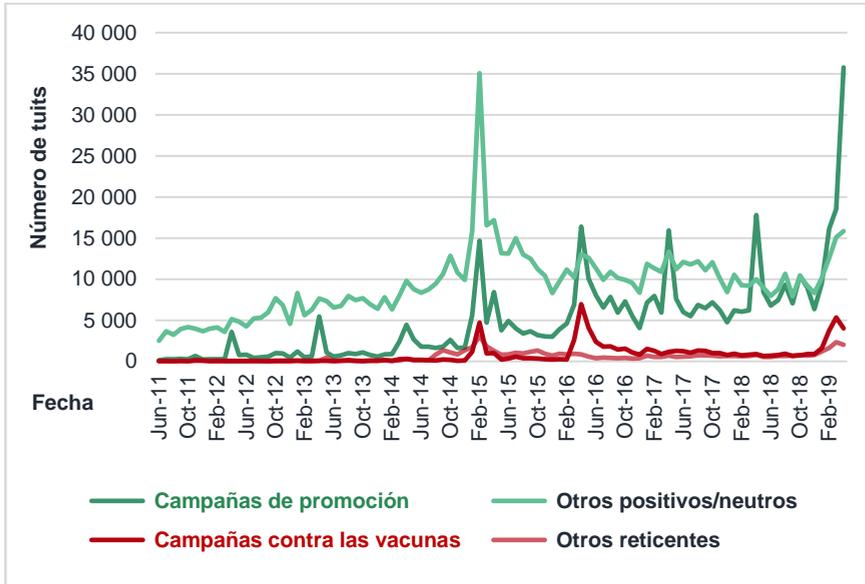


Figura 37. Evolución del número de tuits de cada polaridad según la variable “campañas”.

Mediante el análisis de correlación, comprobamos la “no estacionariedad” de las series ni de sus residuales. Al tratarse de series con tendencia estocástica, calculamos las series de las primeras diferencias, que fueron estacionarias. A continuación, empleamos estas series para estudiar las correlaciones cruzadas entre los cuatro grupos con una periodicidad de doce meses. Hubo correlación significativa en el retardo 0 entre las campañas de promoción y el resto de las series: campañas contra las vacunas ($R^2 = 0,506$), otros tuits positivos/neutros ($R^2 = 0,567$) y otros tuits reticentes ($R^2 = 0,741$). También hubo correlación significativa entre las series de las campañas contra las vacunas y otros tuits reticentes ($R^2 = 0,619$). Las correlaciones no fueron significativas en el resto de los retardos en ningún caso.

B. Determinantes de Reticencia Vacunal

Agrupamos 737 064 (49,16%) tuits bajo la etiqueta “miscelánea” y 762 163 (50,84%) tuits en al menos uno de los siete determinantes de reticencia vacunal analizados. Los determinantes más frecuentes en la base de datos fueron: 1º “ciencia” (n = 452 322 tuits, 30,17%), 2º “seguridad” (n = 288 279, 19,23%) y 3º “efectividad” (n = 87 063 tuits, 5,81%). Los dos primeros determinantes también fueron los más frecuentes dentro del grupo de polaridad positiva y dentro del grupo de polaridad negativa (Figura 38).

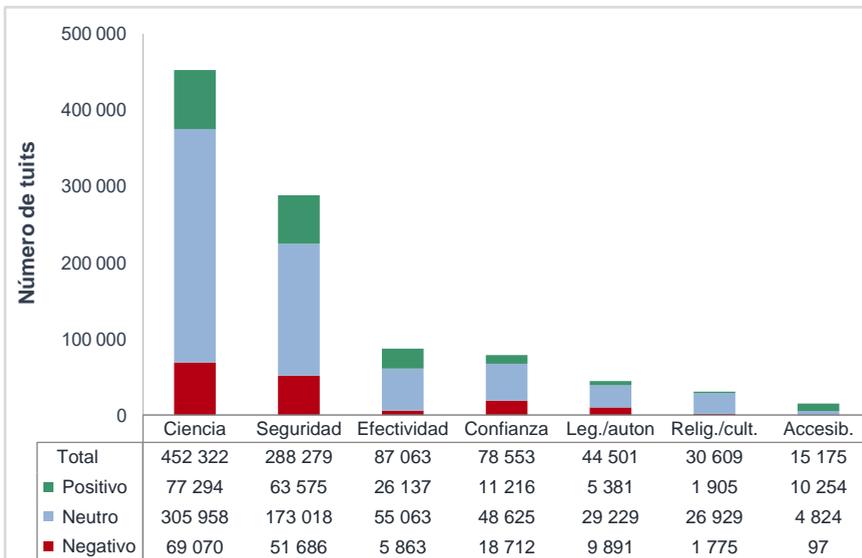


Figura 38. Número de tuits por polaridad y tema.

Observamos que el interés por cada uno de los temas presentó variaciones a lo largo de los años completos del estudio (2012 - 2018) (Figura 39).

Valoración de las vacunas desde una plataforma social con aplicaciones de uso en promoción de la salud.

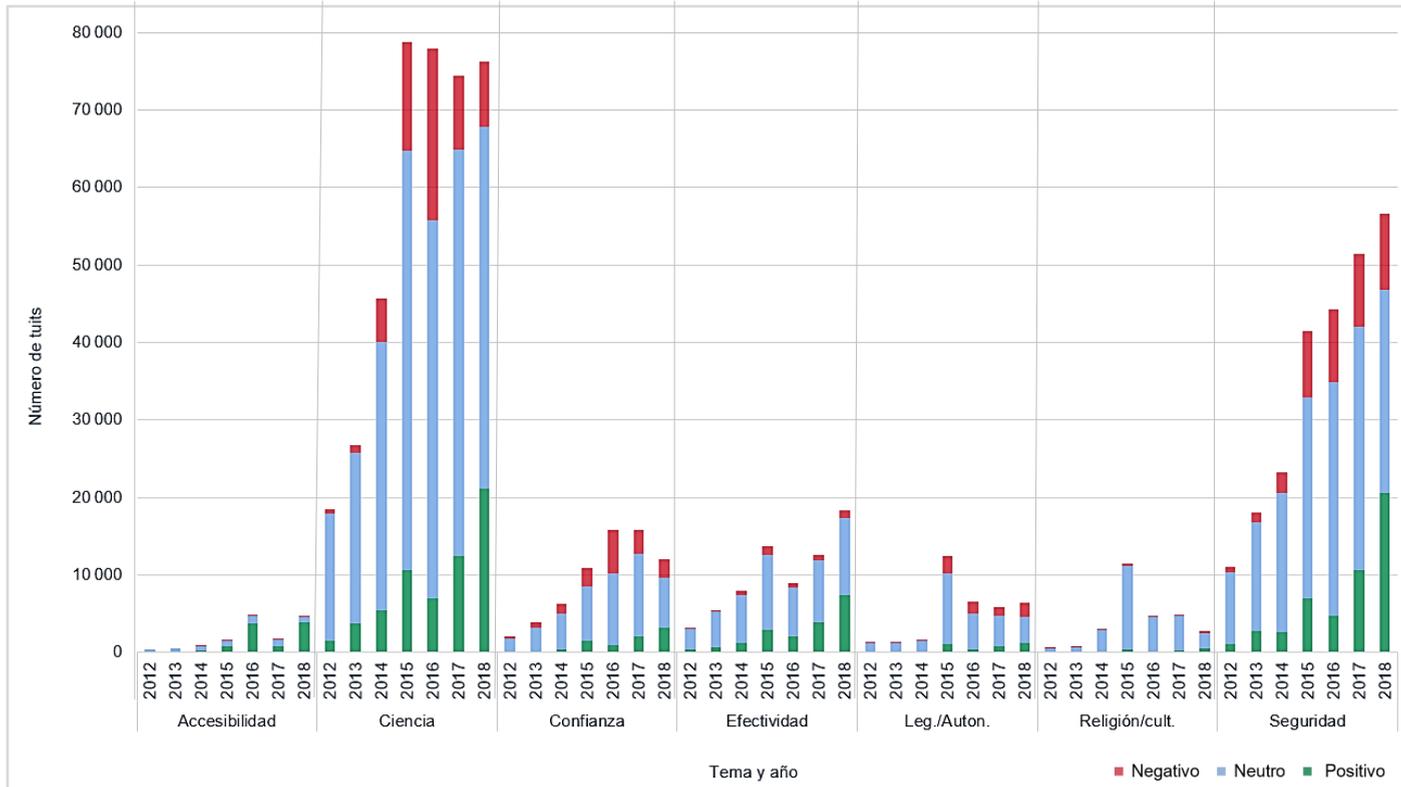


Figura 39. Número de tuits por polaridad, tema y año (2012 - 2018).

El número de tuits sobre los tres temas más frecuentes siguió una tendencia global creciente. En 2015 y 2016, los tuits negativos sobre la evidencia científica y la seguridad de las vacunas superaron a los tuits positivos. Durante el resto de los años completos de estudio, la positividad hacia ambos temas fue mayor que la reticencia. El mayor número de tuits positivos de estos dos temas se registró en 2018, año en el que las campañas de promoción de las vacunas disminuyeron su contenido neutro e incrementaron su contenido positivo sobre ambos temas.

Hubo diferencias estadísticamente significativas en la distribución de la polaridad de sentimientos según el tema. En la Figura 40, representamos el porcentaje de tuits de cada polaridad dentro de cada tema.

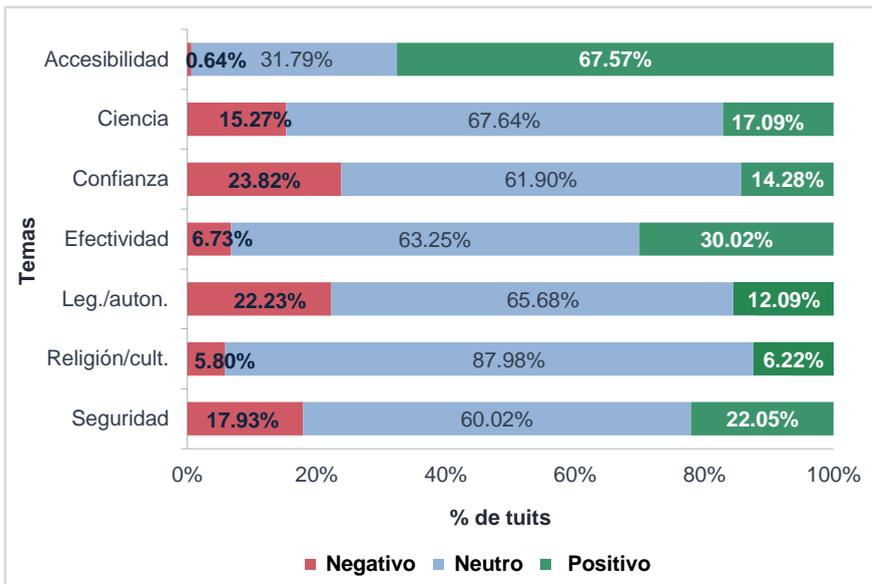


Figura 40. Porcentaje de tuits por polaridad y tema.

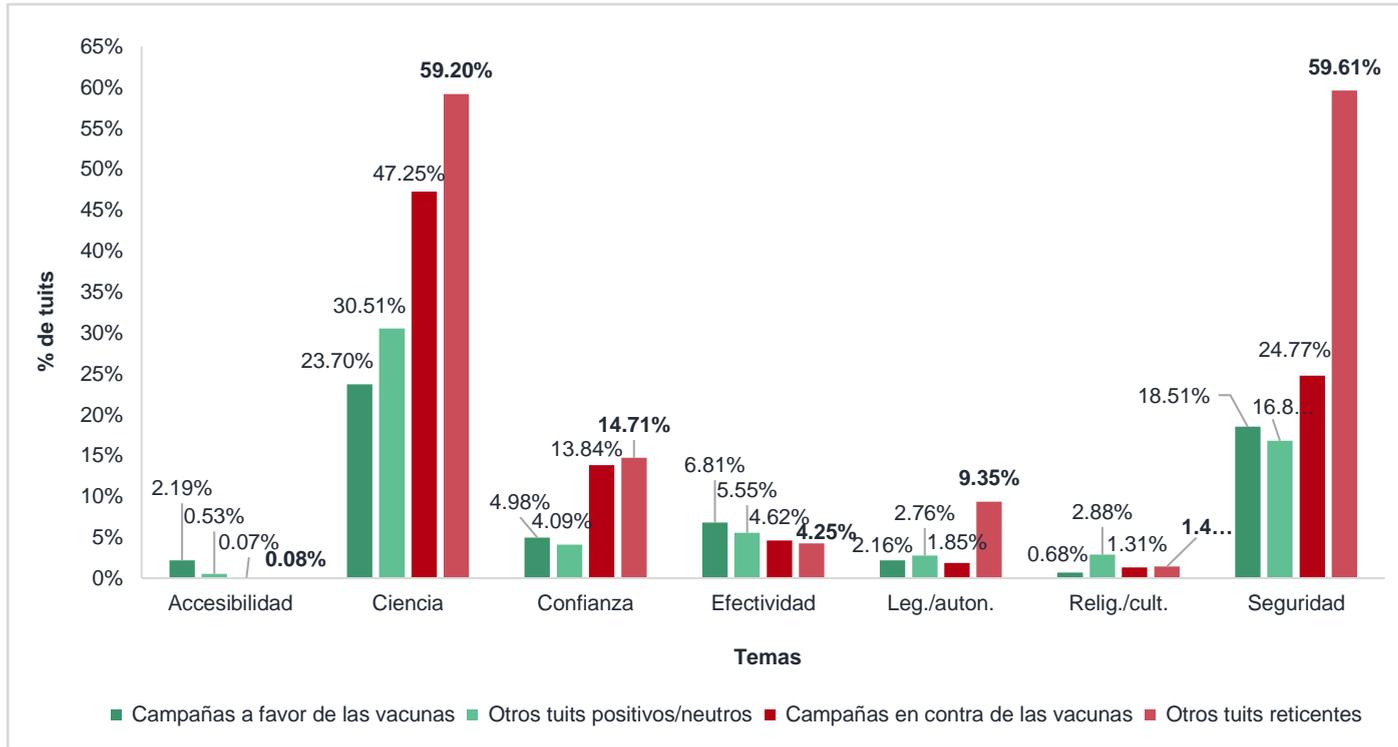
La accesibilidad fue el único tema con predominio de tuits positivos. Aunque los tuits sobre este tema fueron minoritarios a lo largo de todos los años del estudio, observamos que cuando los usuarios escribían sobre accesibilidad, solían hacerlo para abogar por la eliminación de las barreras que limitan la vacunación (67,57% de los tuits sobre accesibilidad), principalmente en el marco de campañas de promoción de la vacunación en Twitter (80,51% de los tuits positivos sobre accesibilidad). Por el contrario, sólo el 0,64% de los tuits sobre este tema expresaban duda vacunal.

En el resto de los temas estudiados, el grupo neutro agrupó más del 50% de los tuits. Los temas con mayor porcentaje de tuits negativos fueron: “confianza” (23,82%), “legislación y autonomía” (22,23%) y “seguridad” (17,93%).

En los cuatro grupos según la variable “campaña” y la polaridad, los dos temas de mayor interés continúan siendo los mismos que en el estudio global de la base datos sin distinción de polaridad (evidencia científica y seguridad) (Figura 41). Aunque ambos temas también fueron los más comunes en los tuits positivos de las campañas de promoción de las vacunas (23,70% y 18,51% respectivamente), fueron menos frecuentes que en el grupo de polaridad negativa que no pertenecía a campañas (59,20% y 59,61% respectivamente).

Asimismo, un menor porcentaje de los tuits positivos de las campañas promocionales abordaban directamente la existencia de la desconfianza vacunal por teorías conspiratorias (4,98%) o trataban sobre los aspectos éticos, políticos y legales de la vacunación (2,16%). No obstante, estos fueron el tercer y cuarto tema más presentes en los tuits reticentes que no pertenecían a campañas (9,35% y 14,71% de los tuits de su grupo respectivamente).

Resultados



Nota: Diferencias significativas en el % de temas entre los cuatro grupos excepto: 1) accesibilidad: otros positivos/neutros, campañas contra las vacunas y otros tuits reticentes, 2) religión / cultura: campañas contra las vacunas y otros tuits reticentes.

Figura 41. Porcentaje de tuits de cada tema según la polaridad y la variable “campaña”.

C. Determinantes de Reticencia Vacunal por Localización

Estudiamos la frecuencia de los temas en los tuits negativos de las seis **regiones** de la OMS (Figura 42).

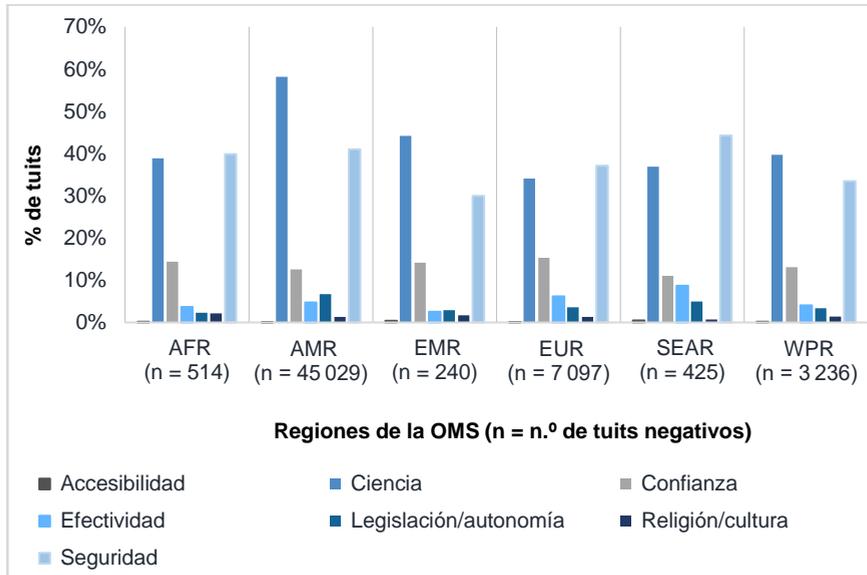


Figura 42. Tuits negativos de cada región clasificados por temas.

-  **Accesibilidad:** tema minoritario en todas las regiones (<1%).
-  **Ciencia:** significativamente más frecuente en AMR (58,20%, n = 26 194 tuits) y en EMR (44,20%, n = 106 tuits).
-  **Confianza:** sólo diferencias significativas entre AMR (12,55%, n = 5 653 tuits) y EUR (15,27%, n = 1 084 tuits).
-  **Efectividad:** significativamente más frecuente en SEAR (8,71%, n = 37 tuits).
-  **Legislación/ autonomía:** más frecuente en AMR (6,70%, n =3 019 tuits), diferencias significativas respecto a todas las regiones excepto SEAR (4,94%, n = 21 tuits).

-  **Religión / cultura:** entre el 0,71% (SEAR) y el 2,14% (AFR) de los tuits negativos. Diferencias no significativas entre regiones.
-  **Seguridad:** significativamente más frecuente en AMR (41,02%, n = 18 472 tuits) y SEAR (44,24% n = 188 tuits) respecto a las otras cuatro regiones. La diferencia no es significativa entre AMR y SEAR.

De los 40 **países** de la ONU con un número igual o mayor a 1 000 tuits, clasificamos los tuits negativos de los 19 países que presentaban un porcentaje de tuits negativos $\geq 1\%$ (Figura 43):

-  **Accesibilidad:** <1% de los tuits reticentes de cada país.
-  **Ciencia:** tema comentado en más del 50% de los tuits negativos de Nigeria (53,69%, n = 234), Japón (56,70%, n = 571) y EE. UU. (60,21%, n = 23 522).
-  **Confianza:** el algoritmo etiquetó en esta categoría menos del 20% de los tuits de todos los países, excepto España (69,44%, n = 234).
-  **Efectividad:** tema presente en más del 10% de los tuits negativos en Indonesia (10,56%, n = 17) y Brasil (13,71%, n = 17).
-  **Legislación/ autonomía:** tema tratado en menos del 5% de los tuits de todos los países, excepto Colombia (6,91%, n = 13) y EE. UU. (6,95%, n = 1 588).
-  **Religión / cultura:** entre el 0,39% (Irlanda) y el 3,36% (Nigeria) de los tuits negativos, diferencias no significativas entre países.
-  **Seguridad:** tema presente en más del 50% de los tuits negativos de Nigeria (62,42%, n = 93), Colombia (60,64%, n = 114), Indonesia (66,46%, n = 107) y Bélgica (78,42%, n = 516).

Valoración de las vacunas desde una plataforma social con aplicaciones de uso en promoción de la salud.

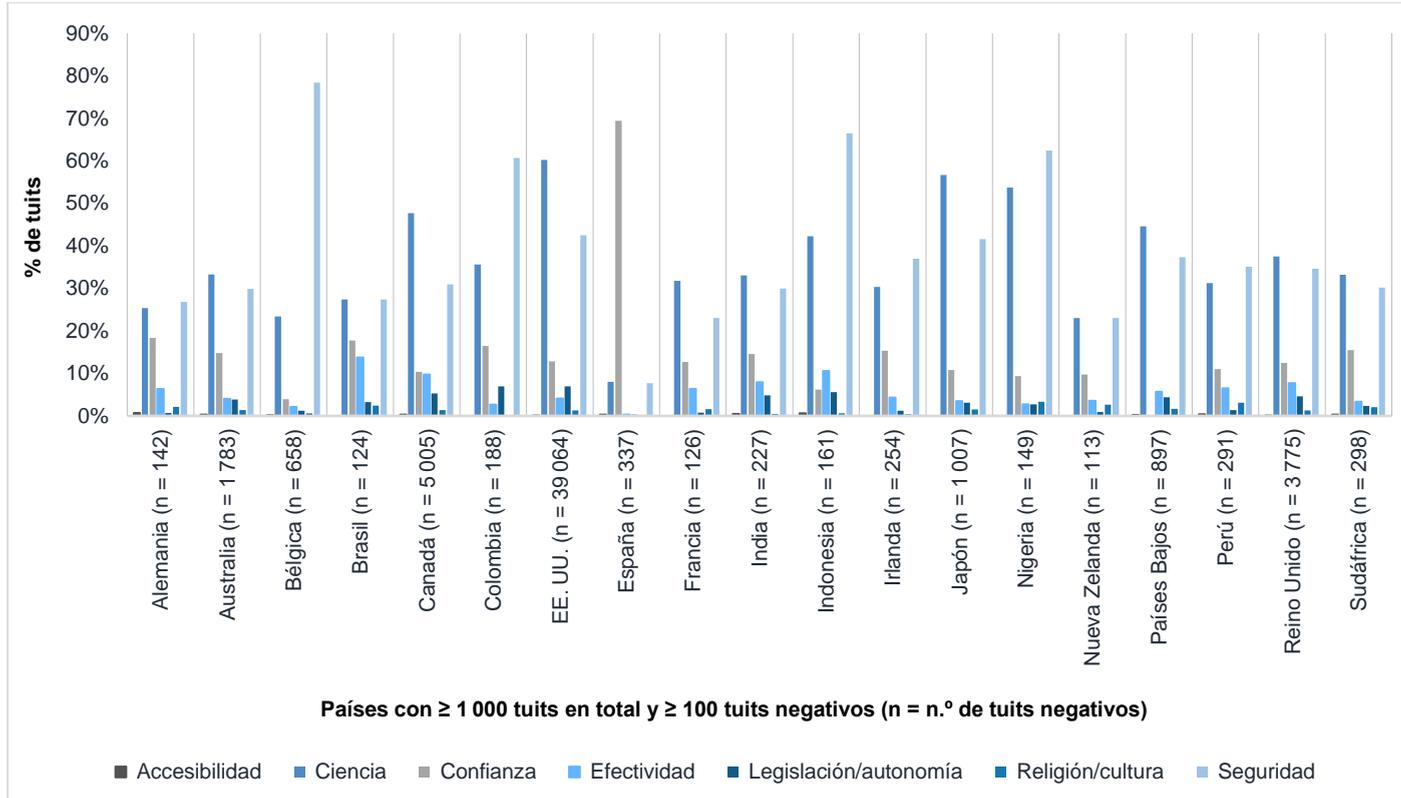


Figura 43. Tuits negativos de cada país clasificados por temas.

D. Tipo de Vacuna

402 124 (26,82%) tuits hablaban sobre alguna vacuna específica – comercializada o experimental. La vacuna frente a la influenza fue la más mencionada en general (26,88% de los tuits sobre alguna vacuna específica, n = 108 090 tuits) (Figura 44) y en el desglose por años (Figura 45), excepto en el 2015, año en el que predominó el interés por la vacuna SRP.

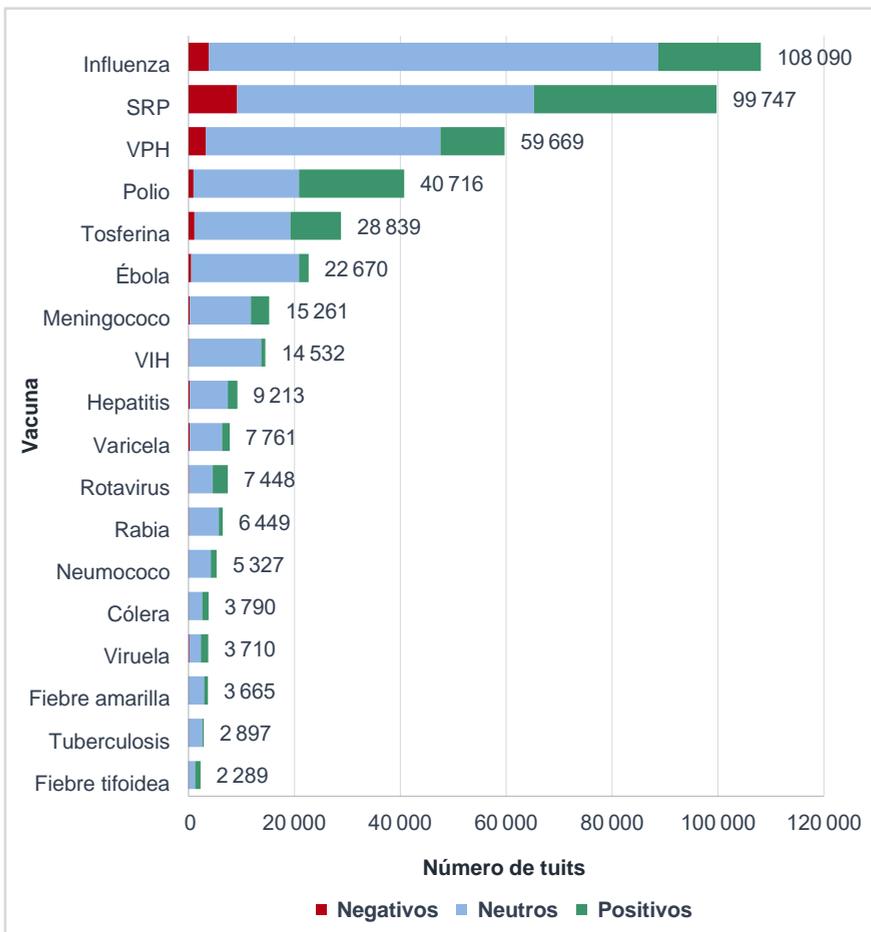


Figura 44. Número de tuits por polaridad y tipo de vacuna.

Valoración de las vacunas desde una plataforma social con aplicaciones de uso en promoción de la salud.

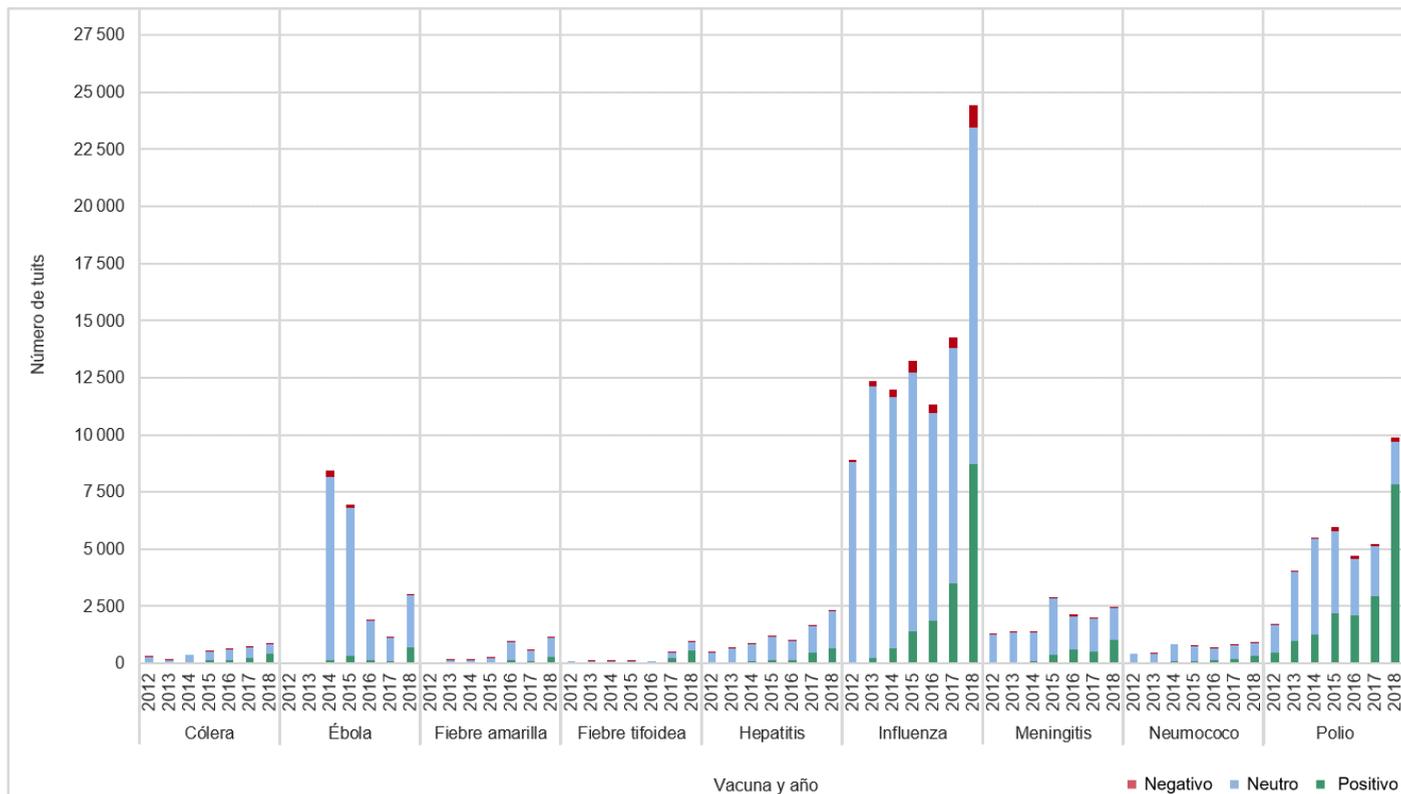


Figura 45. Número de tuits por polaridad, tipo de vacuna y año (2012 -2018) (I).

Resultados

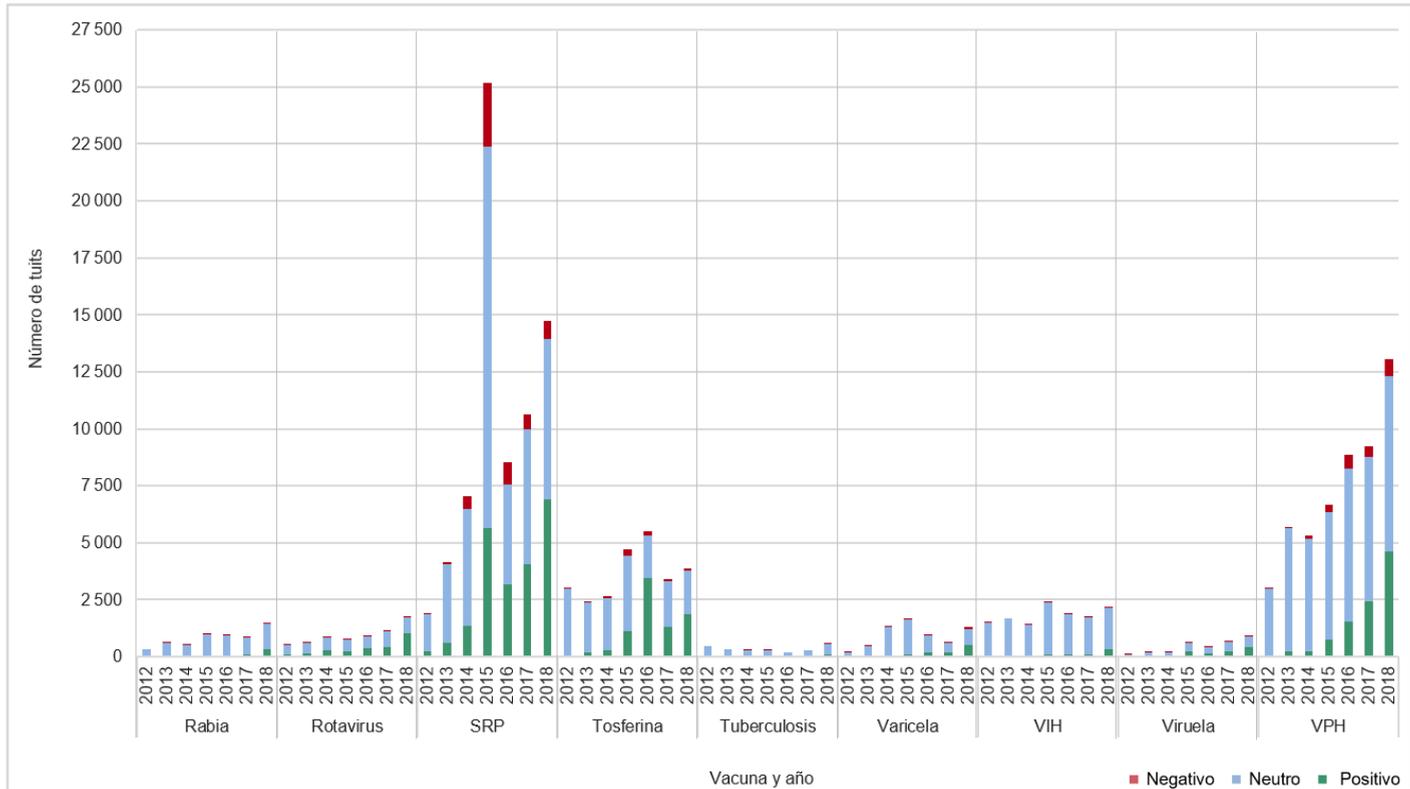


Figura 45. Número de tuits por polaridad, tipo de vacuna y año (2012 - 2018) (II).

Los usuarios expresaron mayor reticencia frente a la vacuna SRP (9,22%), la vacuna contra el VPH (5,43%) y la vacuna contra la influenza (3,58%) (Figura 46).

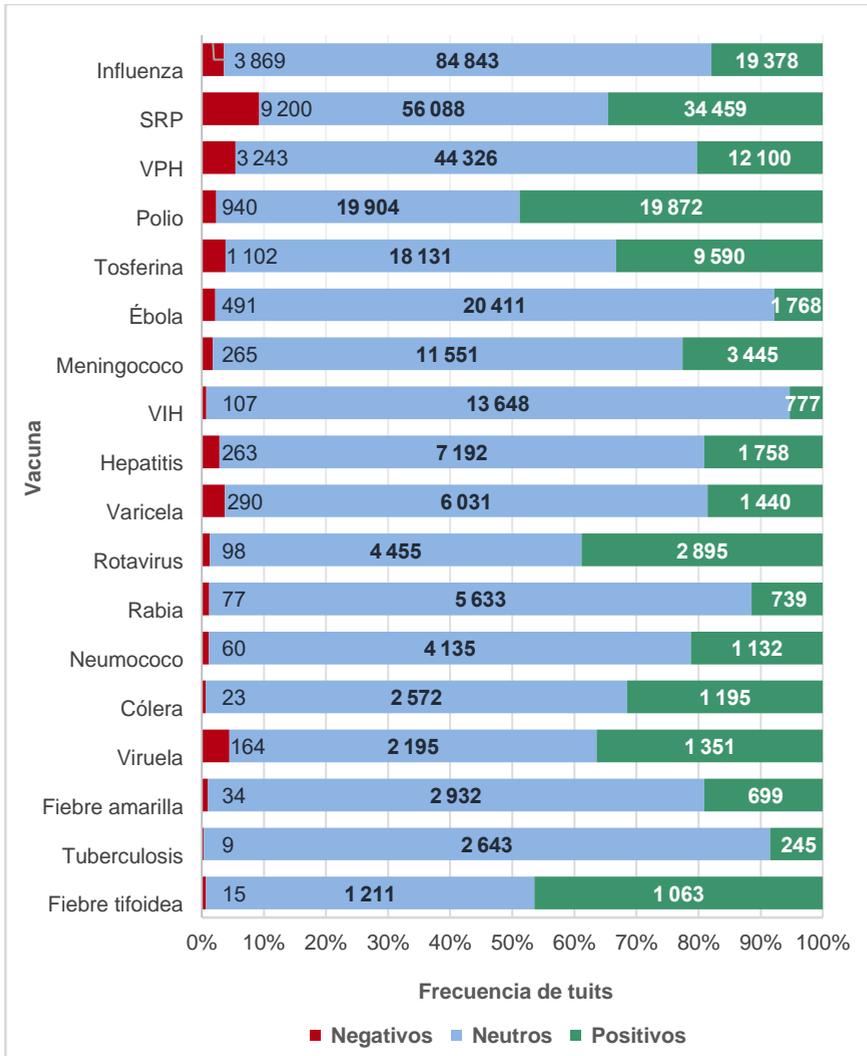


Figura 46. Porcentaje de tuits por polaridad y tipo de vacuna.

De las cinco vacunas comercializadas que generaron más reticencia, destacó la preocupación por la seguridad de la vacuna frente al VPH y de la vacuna frente a la influenza (Tabla 14).

Tabla 14. Porcentaje de tuits negativos de cada tema según la vacuna

Tema	Vacuna				
	SRP	VPH	Tosferina	Varicela	Influenza
Accesibilidad	0,10%	0,03%	0,00%	0,00%	0,10%
Confianza	15,40%*	9,93%	4,17%	5,07%	9,63%
Religión/ cultura	1,09%	0,95%	0,82%	5,41%*	2,42%*
Efectividad	20,98%*	2,76%	17,42%*	18,92%*	8,96%
Ciencia	47,12%	48,51%	42,20%	38,51%	54,20%*
Legislación y autonomía	7,34%	8,89%*	5,26%	10,47%*	8,66%*
Seguridad	28,22%	73,67%*	34,66%*	42,23%*	61,80%*

* Aumento significativo entre los tuits sobre la vacuna de la columna y los tuits que no hablan de esta vacuna.

E. Población Diana

La población diana de vacunación más mencionada en los tuits de la base de datos fueron los “niños” (258 195 tuits, 17,22%). 55 928 (3,73%) mensajes trataban sobre la vacunación de viajeros, adultos o embarazadas (Figura 47).

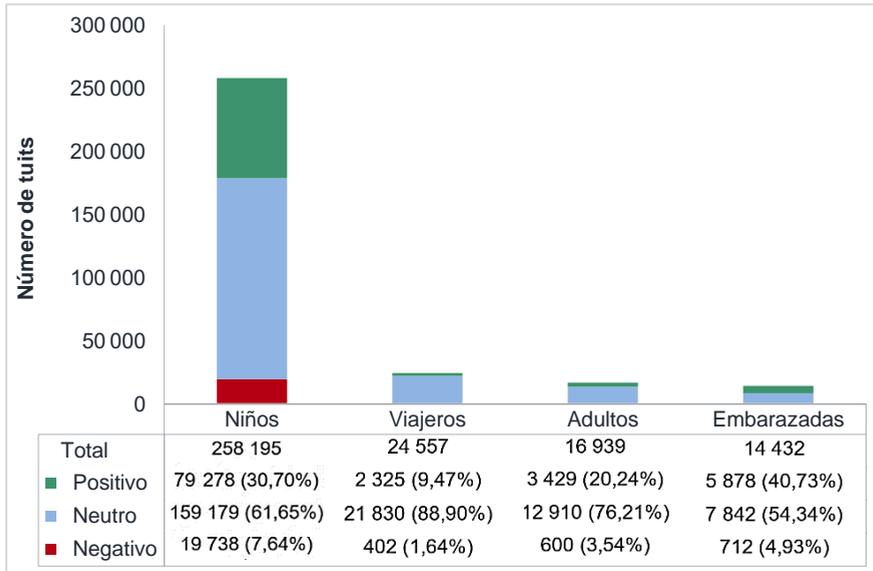


Figura 47. Número de tuits por polaridad y grupo poblacional.

Los temas y las vacunas mencionados con mayor frecuencia en los tuits negativos sobre la vacunación en alguno de estos grupos poblacionales se encuentran en la Tabla 15 y en la Tabla 16 respectivamente. La seguridad y la evidencia científica de la vacunación fueron los aspectos que más preocupaban a los usuarios en todos los grupos poblacionales. Los usuarios expresaron mayor reticencia hacia la administración de la vacuna antigripal en embarazadas y hacia la vacuna SRP en el resto de los grupos.

Tabla 15. Porcentaje de tuits negativos de cada tema según la población diana

Tema	Población diana			
	Niños	Viajeros	Adultos	Embarazadas
Accesibilidad	0,10%	0,60%*	0,00%	0,00%
Confianza	11,06%	6,59%	7,29%	12,12%
Religión/cultura	2,06%	0,30%	1,15%	11,21%*
Efectividad	5,42%	6,89%	14,20%*	3,64%
Ciencia	53,98%	79,34%*	60,84%	62,27%*
Legislación y autonomía	11,06%	3,59%	22,07%*	20,61%*
Seguridad	68,98%*	22,46%	56,81%*	72,12%*

* Aumento significativo en los tuits sobre vacunación en el grupo poblacional de la columna respecto los tuits que no hablan de este grupo poblacional.

Tabla 16. Porcentaje de tuits negativos de cada vacuna según la población diana

Vacuna	Población diana			
	Niños	Viajeros	Adultos	Embarazadas
Tosferina	8,40%	18,18%*	7,87%	14,38%*
Ébola	0,50%	6,06%*	0,50%	0,34%
Fiebre amarilla	0,00%	4,55%*	0,00%	0,00%
Influenza	13,85%	7,58%	24,74%*	60,27%*
Hepatitis	3,30%	9,09%*	2,10%	1,40%
Neumococo	0,66%	0,00%	4,74%*	0,00%
Meningitis	1,94%	9,09%*	2,10%	0,00%*
Polio	5,93%*	4,55%*	0,50%	0,30%
SRP	49,68%*	54,55%*	56,36%*	2,70%
Varicela	3,30%*	0,00%	3,68%*	5,14%*
VPH	16,70%	4,55%	10,00%	14,04%

* Aumento significativo en los tuits sobre vacunación en el grupo poblacional de la columna respecto los tuits que no hablan de este grupo poblacional.

F. Tipo de Instituciones Mencionadas

En la Figura 48, representamos la frecuencia absoluta y relativa por polaridad de los tuits que mencionan alguna agencia de salud pública (9,20%), alguna institución sanitaria asistencial (4,39%), a la industria farmacéutica (3,78%) o algún órgano del gobierno (2,83%).

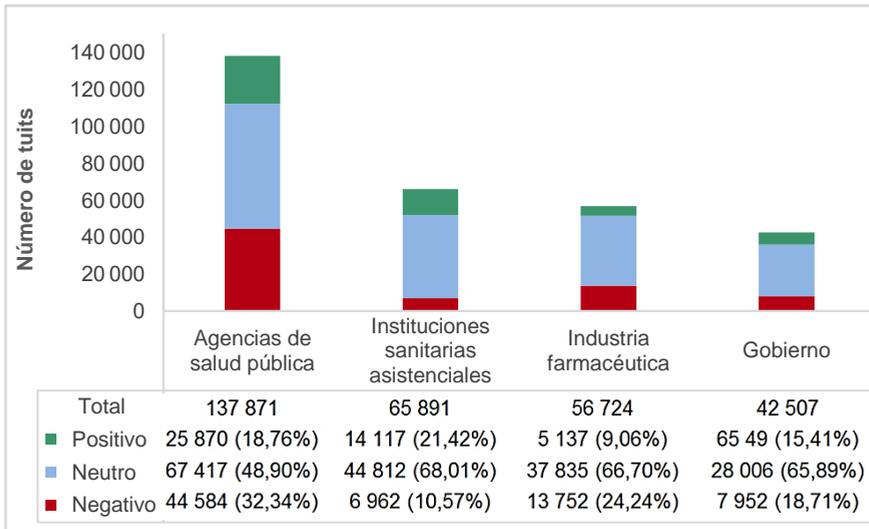


Figura 48. Número de tuits por polaridad e institución mencionada.

Destaca que, dentro de los tuits negativos, las instituciones más nombradas fueron las agencias de salud pública ($n = 44\,584$ tuits). Por el contrario, las instituciones menos citadas en los tuits reticentes fueron las instituciones sanitarias asistenciales ($n = 6\,962$ tuits).

G. Otros Conceptos Relacionados con la Reticencia Vacunal

1. Autismo

81 435 (5,43%) tuits mencionaban el autismo. De ellos, 18 893 (23,20%) eran negativos, 55 625 (68,31%) neutros y 6 917 (8,49%) positivos. Constituían el 14,22%, el 5,35% y el 2,12% respectivamente de todos los tuits de la base de datos con su misma polaridad de sentimiento. Los países con más tuits sobre autismo y vacunas fueron EE. UU. (33 555 tuits, 41,20%), Reino Unido (2 623 tuits, 3,22%) y Canadá (2 522 tuits, 3,10%).

El número mensual de tuits negativos superó al de positivos durante todo el estudio, excepto en agosto de 2018 y en marzo de 2019 (Figura 49). Los años con más tuits negativos sobre autismo fueron 2014 (2 850 tuits, 15,08%), 2015 (4 310 tuits, 22,81%) y 2016 (4 548 tuits, 24,07%).

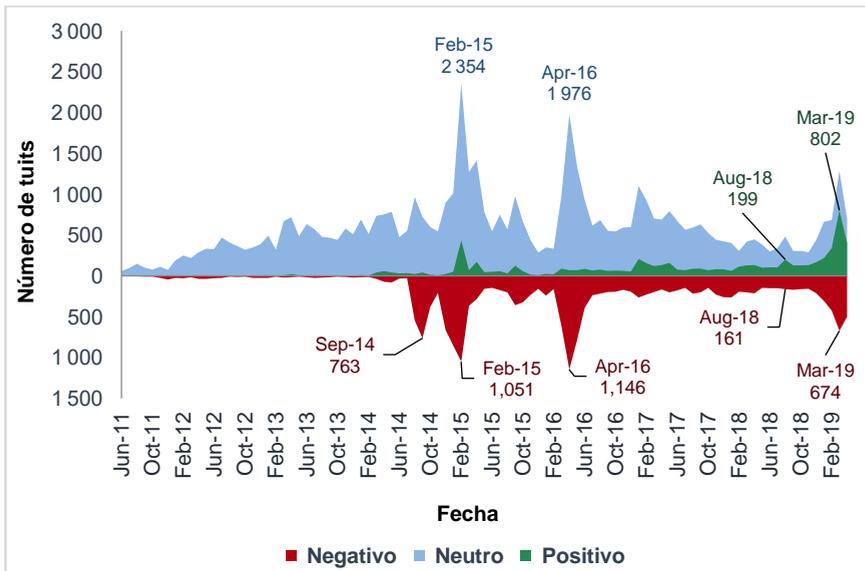


Figura 49. Número de tuits sobre autismo por polaridad.

En los tuits negativos sobre vacunas y autismo:

-  Los temas más frecuentes fueron la “ciencia” (12 789 tuits, 67,69%) y la “seguridad” (6611 tuits, 34,99%). El algoritmo clasificó 8 474 tuits negativos (44,85%) como “campañas” reticentes.
-  2 300 (12,17%) tuits relacionaban el autismo con una o varias vacunas concretas. De ellos, 2 058 (89,48%) tuits trataban sobre la vacuna SRP y 245 (10,65%) tuits sobre la vacuna frente a la influenza.
-  10 179 (53,88%) tuits mencionaban alguna institución de salud pública, 1 374 (7,27%) tuits mencionaban al gobierno y 1 311 (6,94%) tuits mencionaban a la industria farmacéutica.
-  El 50,38% (n = 9 519) fueron etiquetados como “sin loc.” y el 22,95% (n = 4 335) eran de EE. UU. Las demás localizaciones agrupaban menos del 5% de los tuits cada una.

2. Cáncer

40 934 (2,73%) tuits hacían referencia al cáncer. De este grupo, 1 762 (4,30%) tuits eran negativos, 32 339 (79,00%) neutros y 6 833 (16,69%) positivos. Esto equivale al 1,33%, al 3,11% y a 2,09% de todos los tuits de la base de datos de su polaridad respectivamente.

El 15,72% (n = 277) de los tuits negativos, el 30,85% (n = 9 975) de los tuits neutros y el 62,32% (n = 4 258) de los tuits positivos hablaban sobre la vacuna VPH.

De los tuits negativos sobre vacunas y cáncer:

-  El 74,46% (n = 1 312) hablaban sobre las vacunas en general, sin mencionar ningún producto concreto. La única vacuna nombrada en más del 5% de los tuits negativos fue la vacuna frente al VPH.
-  El 45,57% (n = 803) y el 55,11% (n = 971) manifestaban preocupación sobre la seguridad y la evidencia científica respectivamente.
-  El 32,92% (n = 580) relacionaban el cáncer con algún componente de las vacunas.
-  El 51,99% (n = 916) fueron etiquetados como “sin loc.”, el 31,27% (n = 551) eran de EE. UU. y el 10,05% (n = 177) de Canadá. Las demás localizaciones agrupaban menos del 5% de los tuits cada una.
-  El 25,54% (n = 450) pertenecían a campañas contra las vacunas.
-  El 24,29% (n = 428) y el 18,27% (n = 322) mencionaban alguna institución de salud pública o a la industria farmacéutica respectivamente.

3. Composición de las vacunas

46 770 (3,12%) tuits hablaban sobre los componentes de las vacunas, de los cuales 15 735 (33,64%) eran negativos, 29 432 (62,93%) neutros y 1 603 (3,43%) positivos. El 11,84% de todos los tuits negativos de la base de datos, el 2,83% de los neutros y el 0,49% de los positivos trataban sobre este tema.

Los países con más tuits fueron EE. UU. (n = 11 298, 24,16%), Portugal (n = 4 566, 9,76%) y Canadá (n = 1 776 tuits, 3,80%). Las localizaciones con mayor porcentaje de tuits negativos fueron EE. UU. (n = 4 335, 38,37%) y Canadá (n = 579, 32,60%).

El número de tuits negativos fue mayor que el número de tuits positivos durante todo el estudio. No observamos picos relevantes de tuits polarizados en ninguno de los años estudiados.

7 438 (15,90%) tuits trataban sobre la vacunación en algún grupo poblacional concreto. De ellos, 6 725 (90,41%) tuits hacían referencia a la vacunación infantil, 450 (6,05%) tuits a la vacunación en embarazadas, 207 (2,78%) tuits a la vacunación del adulto y 56 (0,75%) tuits a la vacunación del viajero.

Todos los tuits negativos sobre los componentes de las vacunas evidenciaban preocupación sobre su seguridad. Otros temas frecuentes en estos tuits fueron la evidencia científica (7 380 tuits, 46,90%) y la confianza (3 710 tuits, 23,58%). Identificamos 7 398 tuits negativos (47,02% de los tuits negativos) pertenecientes a campañas contra las vacunas.

6 270 (13,41%) tuits trataban sobre los componentes de alguna vacuna en concreto. En los 2 029 (32,36%) tuits negativos de este grupo, las vacunas más mencionadas fueron la vacuna frente a la influenza (1 072 tuits, 52,83%), la vacuna frente al VPH (581 tuits, 28,63%) y la vacuna SRP (376 tuits, 18,53%).

4. Homeopatía

398 tuits negativos (0,30% de los tuits reticentes) y 354 tuits positivos (0,11% de los tuits positivos) contenían mensajes sobre la homeopatía. Los temas más frecuentes en los tuits negativos fueron “seguridad” (180 tuits, 45,23%) y “ciencia” (148 tuits, 37,19%).

VI. RELACIÓN CON DATOS DE LA OMS

A. Coberturas Vacunales frente al VPH y Campañas en Twitter

Estudiamos si hubo relación entre las coberturas vacunales anuales contra el VPH estimadas por la OMS y el número de tuits de nuestra base de datos pertenecientes a campañas de promoción de la vacuna VPH entre 2012 - 2018 en los seis países con más tuits sobre esta vacuna. Excluimos los años 2011 y 2019 debido a que no disponíamos de todos los tuits de estos años completos, sólo la segunda mitad de 2011 y los cuatro primeros meses de 2019. Hallamos correlaciones significativas en España, EE. UU. y Suiza (Tabla 17).

Tabla 17. Correlación entre las coberturas vacunales y el número de tuits sobre VPH		
País	Rho de Spearman	p
Australia	0,679	0,094
Canadá	0,754	0,084
España	0,955	0,001
EE. UU.	0,919	0,003
Irlanda	-0,577	0,175
Suiza	0,756	0,049

B. Casos de Sarampión y Reticencia Vacunal en Twitter

Estudiamos mediante funciones de correlación cruzada la relación entre el número mensual de casos de sarampión declarados a la OMS - variable independiente - y la frecuencia relativa y absoluta mensual de tuits negativos sobre la vacuna SRP - variable dependiente - entre enero de 2012 y abril de 2019, en los tres países con más de 100 tuits negativos sobre esta vacuna en nuestra base de datos: EE.UU. (n = 2 629), Canadá (n = 440) y Reino Unido (n = 414).

También estudiamos la relación entre el número de mensual de casos de sarampión en estos tres países y el número mensual de tuits sobre la vacuna SRP, sin distinción de la polaridad, de cada localización, en este mismo período de tiempo. Excluimos los tuits escritos en 2011 debido a que la base de datos de la OMS no informaba sobre el número mensual de casos antes del año 2012 ni en EE.UU. ni en Canadá.

Hallamos correlaciones positivas estadísticamente significativas en todos los análisis (Figura 50), excepto en el estudio de correlación cruzada entre los casos mensuales de sarampión en Reino Unido y:

-  La frecuencia absoluta mensual de tuits sobre la vacuna SRP localizados en este país.
-  La frecuencia absoluta mensual de tuits negativos sobre la vacuna SRP localizados en este país.

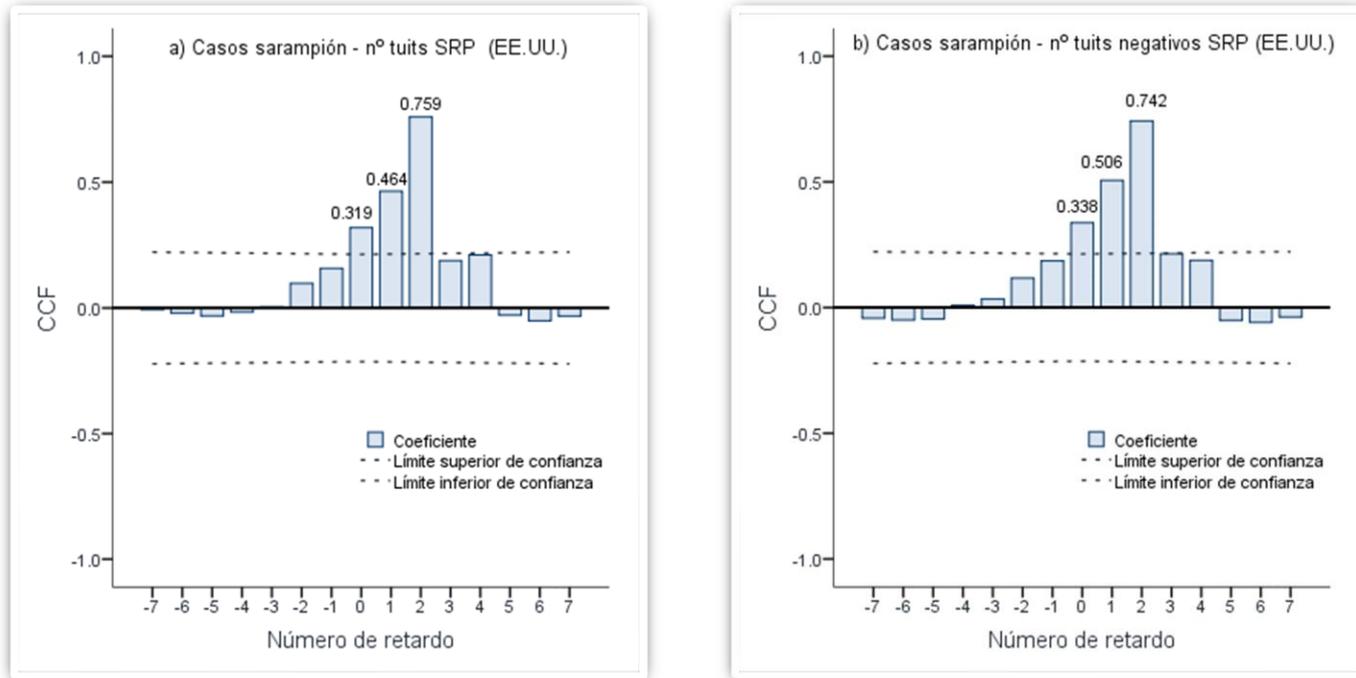


Figura 50. Correlogramas cruzados entre las series “casos mensuales de sarampión” y “tuits sobre SRP” (I).

Valoración de las vacunas desde una plataforma social con aplicaciones de uso en promoción de la salud.

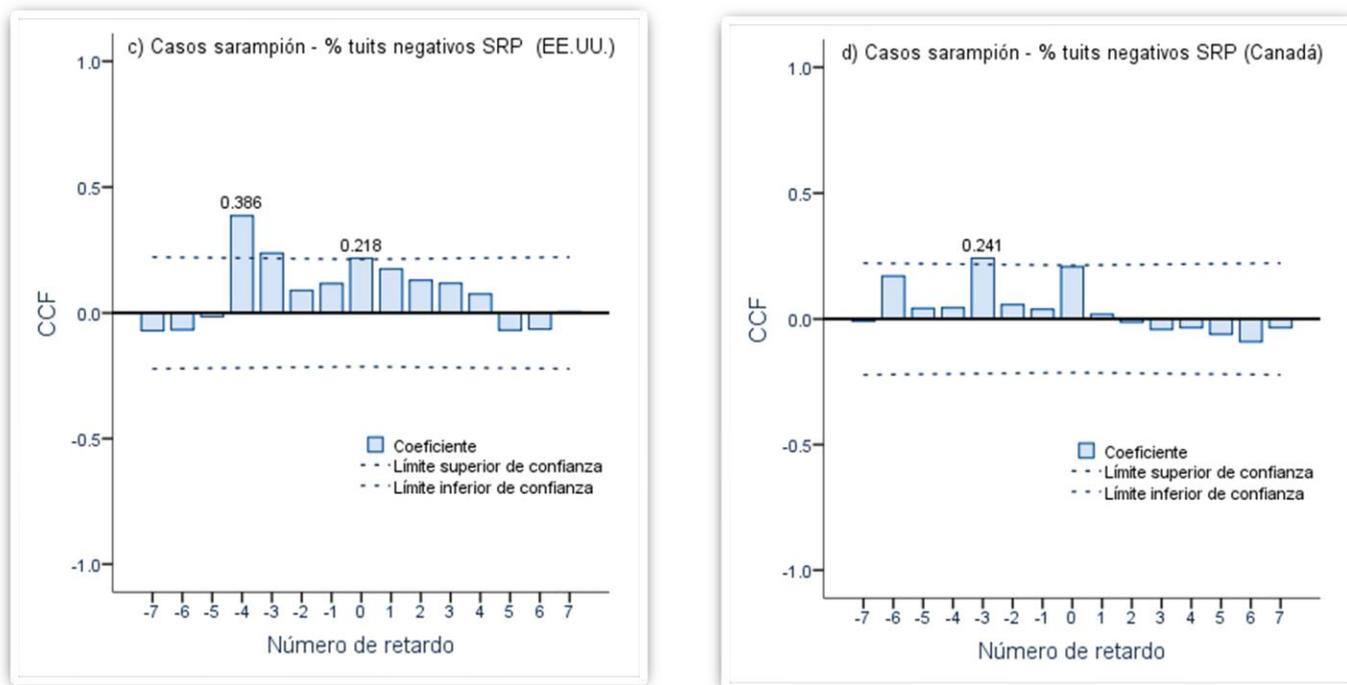


Figura 50. Correlogramas cruzados entre las series “casos mensuales de sarampión” y “tuits sobre SRP” (II).

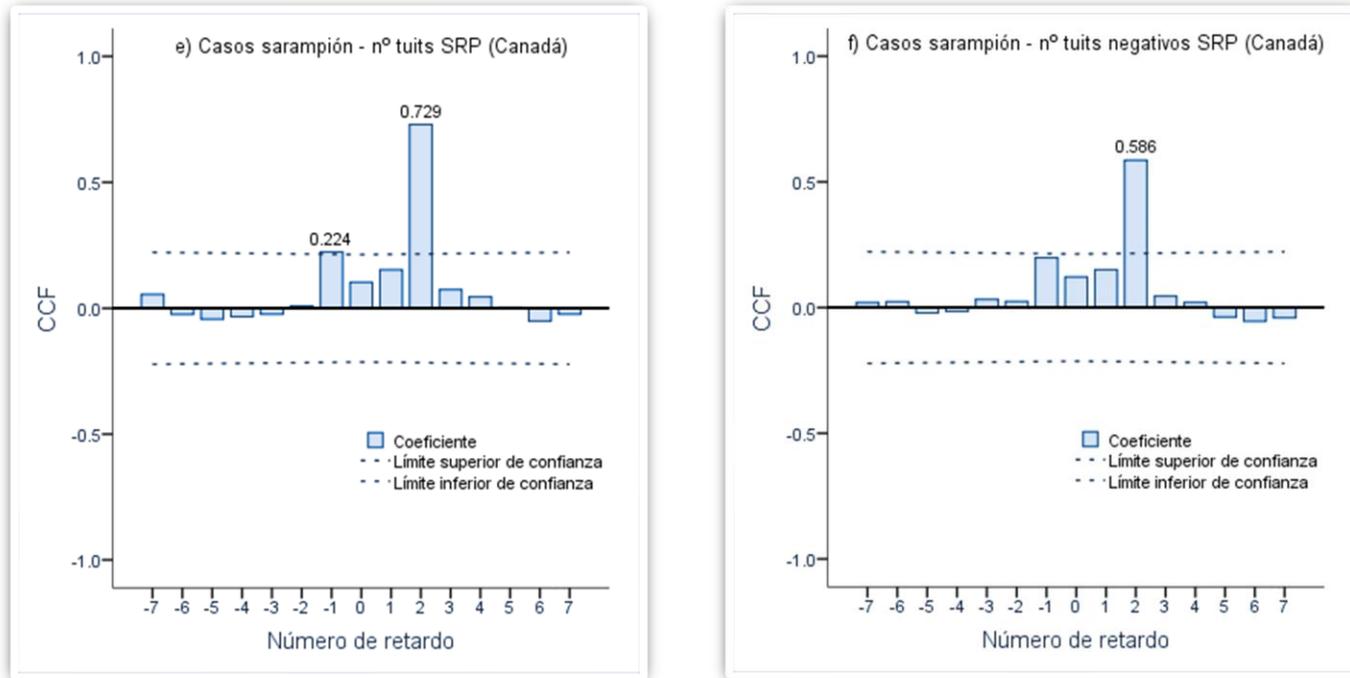


Figura 50. Correlogramas cruzados entre las series “casos mensuales de sarampión” y “tuits sobre SRP” (III).

Valoración de las vacunas desde una plataforma social con aplicaciones de uso en promoción de la salud.

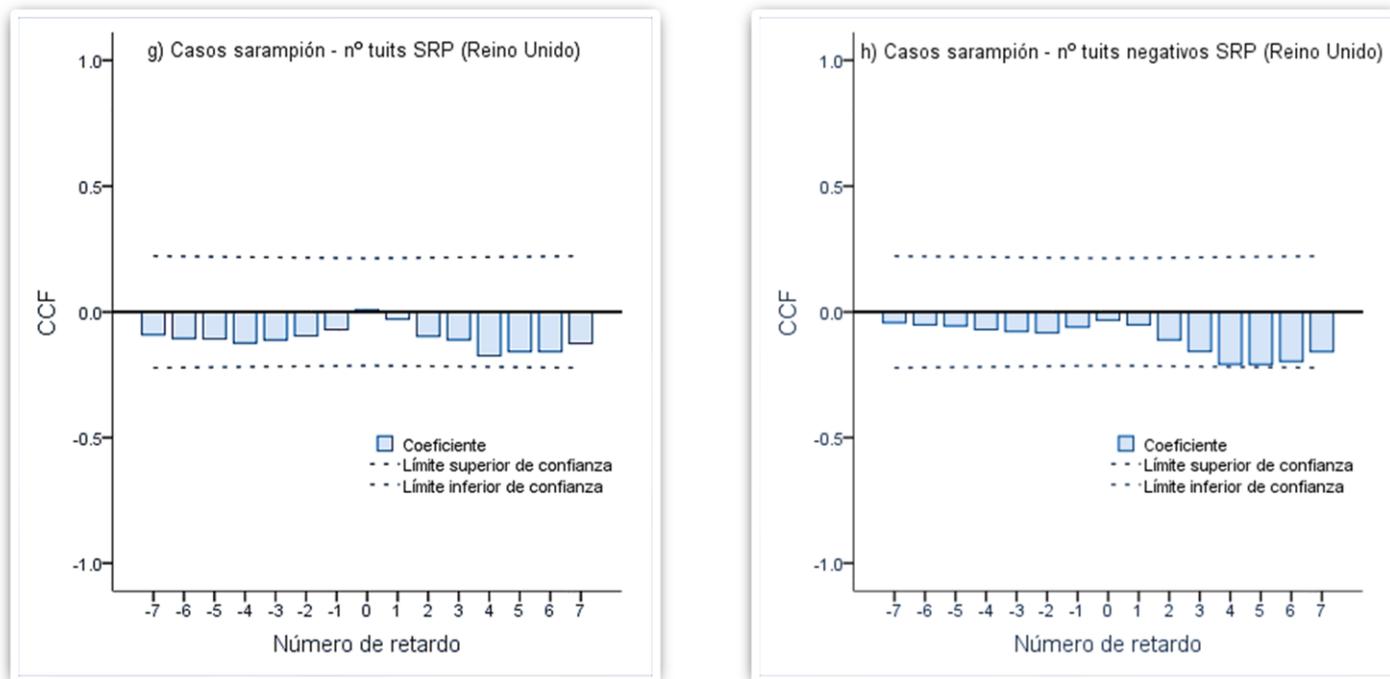


Figura 50. Correlogramas cruzados entre las series “casos mensuales de sarampión” y “tuits sobre SRP” (IV).

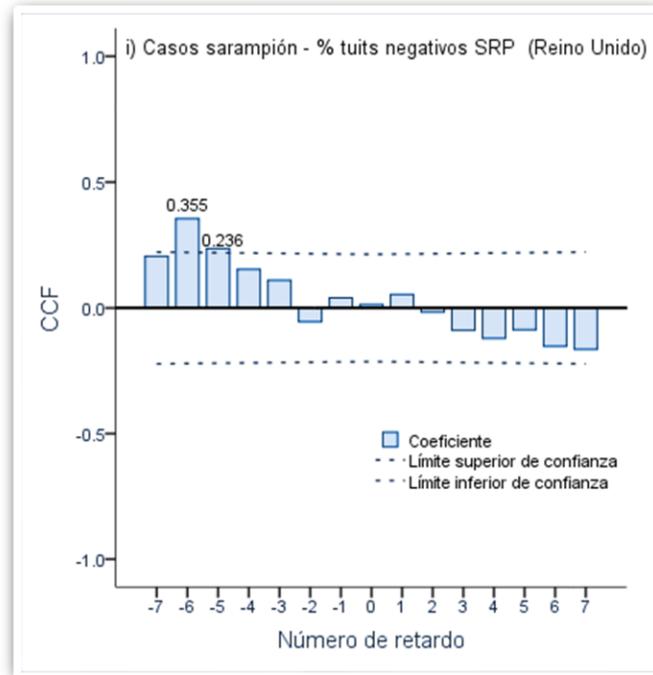


Figura 50. Correlogramas cruzados entre las series “casos mensuales de sarampión” y “tuits sobre SRP” (V).

Valoramos la fuerza de las correlaciones significativas según los criterios de Cohen. Ninguna correlación era muy alta ($0,80 \leq r < 1$) o perfecta ($r = 1$). Las siguientes correlaciones eran moderadas ($0,40 \leq r < 0,60$) o altas ($0,60 \leq r < 0,80$):

-  Correlación moderada entre el número de casos de sarampión en **EE. UU.** y el número de tuits sobre la vacuna SRP en este país en el desfase 1 ($r = 0,464$) y alta en el desfase 2 ($r = 0,759$).
-  Correlación moderada entre el número de casos de sarampión en **EE. UU.** y el número de tuits negativos sobre la vacuna SRP en este país en el desfase 1 ($r = 0,506$) y alta en el desfase 2 ($r = 0,742$).
-  Correlación alta entre el número de casos de sarampión en **Canadá** y el número de tuits sobre la vacuna SRP en este país en el desfase 2 ($r = 0,729$).
-  Correlación moderada entre el número de casos de sarampión en **Canadá** y el número de tuits negativos sobre la vacuna SRP en este país en el desfase 2 ($r = 0,586$).
-  Ninguna correlación moderada o alta en **Reino Unido**.

CAPÍTULO V

DISCUSIÓN

Y

CONCLUSIONES

I. DISCUSIÓN

En el presente capítulo discutiremos los principales hallazgos de esta investigación siguiendo la misma estructura que en el apartado de resultados, haciendo énfasis en las implicaciones de estos resultados en el campo de la salud pública.

A. Resultados del Análisis Global de los Tuits

1. Características generales de los tuits

Realizamos un análisis de la opinión sobre las vacunas en un conjunto de macrodatos de Twitter mediante técnicas de inteligencia artificial.

Aunque incluimos términos de búsqueda en **castellano e inglés** y decidimos incluir en nuestro estudio sólo los tuits escritos en alguno de estos dos idiomas, sólo el 13,57% de los tuits de la base de datos final estaban escritos en castellano. Y de ellos, sólo el 1,01% expresaban duda vacunal. Por ello, el resto del estudio decidimos realizarlo en la base de datos completa sin distinción del idioma.

Creemos que el predominio de los tuits en **inglés** en nuestra base de datos se debe principalmente a dos motivos: 1) según Bello-Orgaz et al. (168), los usuarios de EE. UU., Irlanda, Reino Unido, Canadá y Australia participan más en las conversaciones sobre vacunas que los usuarios de otros países; 2) Twitter tiene más usuarios de las localizaciones anteriores que de países de habla hispana (192). De hecho, un cuarto de los tuits de nuestro estudio procedía de EE. UU.

A pesar de no haber incluido todos los posibles idiomas, recopilamos tuits sobre vacunas del 98,44% (n = 190) de los **países** de la ONU.

La exclusión de idiomas distintos al inglés y al castellano explica que el 87,99% de los tuits de nuestro estudio fueran de las regiones de la Américas de la OMS y la de Europa.

2. Evolución en el número de tuits (junio 2011 - abril 2019)

Observamos un interés creciente en las vacunas a lo largo de los **años** de estudio, explicado parcialmente por el crecimiento del número de usuarios de Twitter de 2011 a 2019.

Por **meses**, objetivamos un incremento de tuits cada abril, coincidiendo con la Semana Mundial de la Inmunización. De hecho, el pico mensual más alto de tuits se produjo en abril de 2019. El segundo pico más alto lo identificamos en febrero de 2015, coincidiendo con un debate político mediático sobre la vacunación obligatoria en los EE. UU.

B. Resultados del Análisis Semántico

En nuestro análisis de **unigramas** observamos que, tanto en castellano como en inglés, los usuarios mencionaron frecuentemente en sus conversaciones sobre vacunas los términos “niños” o “hijos”, “salud”, “enfermedad”, “sarampión”, “gripe” o “influenza” y “cáncer”.

Aunque en este análisis no dividimos los tuits por polaridad de sentimiento hacia las vacunas, posteriormente aplicamos algoritmos automatizados de AS y modelado de tópicos para identificar cuáles fueron los grupos poblaciones y las vacunas más comentadas según la polaridad de los tuits. Confirmamos que los niños fueron la población diana de la vacunación sobre la que los usuarios mostraron

mayor interés tanto en los tuits a favor de las vacunas como en los tuits reticentes. Asimismo, confirmamos que las vacunas más discutidas fueron la vacuna SRP y la vacuna antigripal en ambos grupos de polaridad.

Por otra parte, concluimos a partir del análisis de **bigramas** que los usuarios centraron sus conversaciones en la relación de las vacunas con la salud, el autismo, el cáncer, la necesidad y la efectividad de las vacunas y la promoción de las vacunas. Además de la vacuna frente al sarampión y la influenza, otras vacunas nombradas con frecuencia fueron la vacuna frente al VPH (en los tuits de ambos idiomas), la vacuna contra la tosferina (en los tuits en inglés), la vacuna frente al VIH y la fiebre amarilla (en los tuits en castellano).

Dado que los tuits en **castellano** fueron minoritarios, las conversaciones sobre las dos últimas vacunas no constaban entre las más frecuentes en el análisis temático global, mientras que los tuits sobre las vacunas frente el VPH y la tosferina fueron el tercer y quinto tema más frecuente respectivamente.

Por tanto, el **análisis semántico** permite rápidamente obtener una aproximación sobre cuáles son los aspectos que generan más interés en las conversaciones sobre vacunas en Twitter. Podemos confirmar las conclusiones obtenidas en el análisis semántico mediante estudios de modelado de tópicos combinados con técnicas de AS.

Llama la atención que veintiuno de los treinta unigramas en lengua inglesa más frecuentes en nuestro conjunto de datos coinciden con los comunicados por Raghupathi et al. (172) en su estudio sobre la reticencia vacunal en Twitter entre enero y abril de 2019. Asimismo, nuestra lista de los treinta bigramas más frecuentes incluye ocho de

los diez bigramas más frecuentes en la base de datos de estos autores. Nuestro estudio abarcó un período de tiempo mayor y nuestra estrategia de búsqueda fue distinta, pero encontramos que el discurso de los usuarios tuvo una alta similitud en ambos estudios.

Por tanto, aunque los resultados de los análisis de la opinión sobre vacunas en las redes sociales no pueden generalizarse y los promotores de salud deben realizar el análisis de los datos del período de tiempo, localización y vacuna o subtema de interés, podemos tener en cuenta las conclusiones del análisis semántico de otros autores para decidir en qué áreas vamos a enfocar la monitorización de la reticencia vacunal en nuestros proyectos.

C. Resultados del Análisis de la Polaridad de Sentimientos

1. Métricas y comparación con el AS de Vader

Nuestro algoritmo híbrido entrenado para la clasificación de la polaridad de sentimientos hacia las vacunas demostró un alto **rendimiento** en el conjunto de datos de prueba (puntuación F1 global: media macro = 0,908 y media micro = 0,921).

El menor rendimiento lo obtuvo en la clasificación de los tuits **negativos** (puntuación F1 = 0,869), debido a que presentó mayor dificultad para detectar los tuits pertenecientes a esta clase (precisión = 0,809), aunque la mayoría de los tuits que clasificó como negativos realmente lo eran (exhaustividad = 0,939). Ahora bien, aunque no se ha establecido un punto de corte universal para la puntuación F1, los valores iguales o superiores a 0,80 suelen considerarse adecuados. Asimismo, nuestro algoritmo presentó un rendimiento superior al de otros autores (174,182,186).

Con el fin de comprobar si había valido la pena entrenar un algoritmo específico para nuestro estudio, decidimos comparar sus métricas con un algoritmo preexistente. Seleccionamos **VADER**, ya que es el *gold-standard* para la clasificación de sentimientos en plataformas de *microblogging* como Twitter, con una precisión de 0,99, una exhaustividad de 0,94 y una puntuación F1 de 0,96 (medias macro) para la clasificación de los tuits según sus creadores (191).

En nuestro conjunto de datos de prueba, VADER tuvo un rendimiento menor (puntuación F1 global: media macro = 0,551 y media micro = 0,689) que nuestro algoritmo, con un desempeño especialmente pobre en la clasificación de los tuits negativos (puntuación F1 = 0,163).

Por ende, es conveniente entrenar algoritmos específicos para cada campo de estudio, aunque sea un proceso largo y laborioso, ya que los resultados son más exactos que los arrojados por algoritmos preentrenados en conjuntos de datos no especializados o pertenecientes a otras disciplinas de estudio.

2. Evolución temporal de la polaridad de sentimientos y resultados por meses del año

En consonancia con los estudios de D'Andrea et al. (186) y Tavoschi et al. (187), observamos una tendencia hacia la polarización del discurso de los usuarios a lo largo del estudio, aunque los mensajes sobre vacunas fueron mayoritariamente **neutros** (69,36%).

Al igual que en la población general (193), el sentimiento **positivo** predominó sobre el negativo. La proporción de tuits negativos se mantuvo por debajo del 20% durante todo el período de estudio, excepto en abril de 2016, tras el anuncio de la retirada de un documental dirigido por Wakefield de la programación del festival Tribeca, en el que acusaba a los CDC de haber omitido información que demostraba la relación causal entre las vacunas y el autismo en su informe de 2004.

Identificamos que los **picos** de tuits sobre vacunas observados en los meses de abril se debieron a una polarización del discurso a favor del contenido positivo, tal y como era de esperar en el contexto de la Semana Mundial de la Inmunización. Sin embargo, no evidenciamos cambios significativos en la frecuencia de los tuits negativos en abril respecto al resto de los meses.

Una posible explicación es el efecto de **cámara de eco** (*echo-chamber*). Este fenómeno, propio de los medios de comunicación y las redes sociales, facilita que los individuos reticentes sólo se expongan a información de su misma ideología e interactúen mayoritariamente con personas que comparten su misma opinión. (194)

Destacan dos **estrategias** para disminuir el efecto de las cámaras de eco. Por una parte, podemos **mencionar** en nuestros tuits a los usuarios reticentes, (195) aunque de esta manera sólo aumentamos la difusión de los mensajes a un pequeño porcentaje de los usuarios reticentes, por lo que podemos reservar esta estrategia para dirigirnos a los usuarios reticentes más influyentes (con más seguidores y que generan más mensajes negativos).

Por otra parte, podemos incluir los *hashtags* y las **palabras clave** que los usuarios reticentes suelen buscar en las redes sociales en nuestros mensajes promotores de las vacunas para aumentar la probabilidad de que los usuarios reticentes encuentren estos mensajes positivos cuando naveguen por Twitter. (195)

En nuestro esfuerzo por disminuir el efecto de cámara de eco, debemos tener en cuenta que los individuos reticentes suelen ignorar los mensajes a favor de las vacunas y, cuando interactúan con personas con una opinión contraria, refuerzan sus ideas previas sobre la vacunación. (196) Por ello, algunos autores (194,197,198) afirman que las intervenciones de salud pública actuales enfocadas a aumentar la aceptación de la vacunación pueden no ser efectivas, e incluso, pueden fortalecer la reticencia vacunal.

En su lugar, podemos entablar **conversaciones** con las personas reticentes encaminadas a comprender los motivos de sus dudas. Durante estas interacciones, debemos evitar utilizar un discurso extremadamente polarizado y no debemos realizar críticas de las opiniones reticentes (199), ya que el uso de frases culpabilizadoras y los estereotipos genera mayor rechazo hacia a la vacunación. (200)

En numerosas ocasiones, las conversaciones sobre vacunas en las redes sociales se producen en **entornos** hostiles, impulsados tanto por usuarios a favor como en contra de las vacunas. (57)

Las organizaciones de salud pública deben crear entornos seguros en las redes sociales en los que los usuarios con diversidad de opiniones puedan plantear sus dudas y debatir sobre las vacunas desde el respeto. Para ello, pueden establecer **alianzas** con miembros de la comunidad no reticente. Esta colectividad, a través de las redes sociales, puede abogar por las vacunas desde una perspectiva humanizada y libre de conflicto de intereses, además de colaborar en la moderación y la monitorización de las conversaciones. (57)

3. Polaridad de sentimientos por día de la semana

Otro aspecto que queríamos esclarecer es si las polaridades de sentimientos hacia las vacunas variaron entre los días de la semana. El número absoluto de tuits sobre vacunas, total y por categorías de polaridad de sentimientos, fue mayor entre semana, especialmente los **miércoles y jueves**, en comparación con los fines de semana.

El descenso medio en el **número absoluto** de tuits los fines de semana fue mayor en el grupo de polaridad positiva, seguido del grupo de polaridad neutra y el grupo de polaridad negativa. En consecuencia, el porcentaje de tuits neutros no varió significativamente según el día de la semana, el porcentaje de tuits positivos fue mayor los miércoles y jueves y menor los sábados y domingos, y el porcentaje de tuits negativos mostró la tendencia opuesta a los tuits positivos.

Este patrón de variación de **frecuencia relativa** de los tuits por día de la semana es el mismo descrito por Du et al. (182) en su estudio de sentimientos en Twitter hacia la vacuna frente al VPH. Estos autores proponían a los profesionales de salud pública que aumentaran los mensajes de promoción de las vacunas durante los fines de semana, ya que los sábados y domingos aumenta el porcentaje de mensajes con polaridad negativa.

Pero Du et al. (182) no describieron la variación en la frecuencia absoluta de tuits según el día de la semana y, en nuestro estudio, realmente el aumento en la frecuencia relativa de tuits negativos los sábados y domingos se debe a que la actividad de los usuarios reticentes disminuye los fines de semana en menor medida que la actividad de los usuarios no reticentes. Por lo tanto, creemos que los profesionales de salud pública pueden lograr una **mayor interacción** con los usuarios reticentes en los días de la **mitad de la semana**.

4. Polaridad de sentimientos por países

WPR fue la segunda **región** con mayor reticencia vacunal, tanto en nuestro estudio en Twitter como en la población general (193). El primer puesto lo ocupó AMR en nuestro estudio, frente a EUR en la población general (193).

Identificamos reticencia vacunal en un porcentaje ligeramente menor de **países** (86,84%, n = 165) que Lane et al. (201) en su análisis de reticencia vacunal mundial en la población general entre 2014 – 2016 (92,75%, n = 179). Al igual que ocurre en la población general (193,201), el sentimiento negativo no fue predominante en ninguno de los países.

De los cuatro países con mayor **reticencia** vacunal en nuestro estudio, Japón es el único que también está en la lista con mayor duda vacunal fuera de las redes sociales. (193) Los otros tres países (Países Bajos, Canadá y EE. UU.) concentran una mayor reticencia vacunal entre la población con un nivel educativo más alto. (193) Dado que los usuarios de Twitter suelen tener un mayor nivel educativo que la población general, (202) entendemos por qué estos tres países fueron líderes en reticencia vacunal en Twitter, pero no en los estudios de encuestas tradicionales.

En cuanto al sentimiento **positivo**, creemos que la mayor proporción de tuits positivos en Suiza podría deberse a que muchas organizaciones internacionales tienen su sede en este país y utilizan Twitter activamente para promover las vacunas.

5. Interacción social según la polaridad de sentimiento

Para medir el compromiso o **nivel de interacción social** (*engagement*), analizamos la cantidad de retuits, favoritos/me gusta y respuestas que recibieron los tuits según su polaridad de sentimiento. Observamos que el compromiso generado por los tuits sobre vacunas fue muy bajo ($P50 < 1$ en todas las categorías).

Las **respuestas** fueron el mecanismo de interacción menos frecuente. De las tres categorías de sentimiento, el grupo de tuits negativos fue el que englobó un mayor porcentaje de tuits con al menos una respuesta (19,30%), mientras que sólo el 15,80% de los tuits positivos obtuvo alguna respuesta.

Debido a que no recogimos los textos de las respuestas a los tuits de la base de datos que no contenían los *hashtag* seleccionados, no

podimos determinar si la polaridad de sentimientos de las respuestas estaba en consonancia con la polaridad del tuit al que respondían los usuarios. Pero el bajo porcentaje de tuits de todas las categorías de sentimientos con alguna respuesta refuerza la idea expresada en párrafos anteriores sobre la necesidad de aumentar las conversaciones entre los usuarios promotores de las vacunas y los usuarios reticentes.

En cuanto a los tuits con algún **favorito/me gusta**, fueron más frecuentes que los tuits con algún retuit en todas las polaridades de sentimiento. Esto se explica en parte por la distinta funcionalidad de ambos elementos. Los usuarios no sólo marcan los tuits como favoritos/me gusta para indicar acuerdo con su contenido, sino también para almacenar tuits, para promocionar a otros usuarios o a sí mismos, o como respuestas a tuits en los que son mencionados. (203)

En cambio, los **retuits**, además de indicar acuerdo, amplifican la visibilidad de los mensajes y sus autores, por lo que son el principal mecanismo de difusión de información en Twitter. (204) La mediana de cero retuits por tuit en todas las categorías de sentimientos coincide con la descrita por Blankenship et al. (205)

Observamos que el contenido neutro tuvo menor probabilidad de ser retuiteado que el contenido polarizado. Por lo tanto, si queremos difundir información acerca de las vacunas, tendremos más probabilidad de que el contenido sea compartido si añadimos alguna frase positiva que si escribimos mensajes totalmente neutros. Otras estrategias para aumentar los retuits de nuestros mensajes incluyen añadir *URLs*, *hashtags* y mencionar a otros usuarios en los textos. (135)

D. Resultados del Análisis de los Usuarios

1. Resultados del estudio de todos los usuarios

Los profesionales sanitarios y las organizaciones de salud pública utilizamos las redes sociales para enviar mensajes promotores de la vacunación y corregir la desinformación que los usuarios comparten en las redes. (64)

Al diseñar nuestras campañas de promoción de las vacunas, es indispensable que tengamos en cuenta que el **idioma** supone una barrera de acceso a la información sobre vacunas. Los usuarios multilingües (2,53% en nuestro estudio) favorecen el flujo de la información entre personas de diferentes comunidades (206), con lo que contribuyen a la disminución de las desigualdades en salud. Por ello, es importante incluir en nuestras campañas de promoción de las vacunas, portavoces multilingües que sirvan de enlace con los miembros de todas las culturas y naciones que componen nuestra población diana.

La gran cantidad de **usuarios no institucionales** (2/3 del total) que enviaron algún tuit sobre vacunas y la mayor reticencia en este tipo de usuarios sugieren que Twitter es una plataforma adecuada para detectar personas reticentes, con las que los profesionales de salud pública podemos contactar y dialogar con el fin de aumentar su confianza vacunal. Asimismo, podemos diseñar intervenciones colectivas para el abordaje de la reticencia vacunal si logramos establecer alianzas con las cuentas institucionales que expresan reticencia vacunal.

No obstante, objetivamos que la participación de los usuarios en las conversaciones sobre vacunación fue fundamentalmente

esporádica, ya que sólo un cuarto de ellos escribió más de un mensaje sobre el tema. De este pequeño grupo de usuarios que muestran un interés más constante en la vacunación, podemos seleccionar a los usuarios reticentes como población diana de las intervenciones y entrenar a los usuarios no reticentes para que ejerzan como promotores de las vacunas en Twitter.

De los usuarios de nuestro estudio que escribieron más de un tuit acerca de las vacunas, la mayoría (91,83%) mantuvo una **polaridad homogénea** en todo su discurso, tal y como describieron Schmidt et al. (196) en Facebook. Es decir, la mayor parte de los usuarios podían clasificarse en uno de los siguientes **grupos**: usuarios totalmente a **favor** de las vacunas, usuarios con un discurso totalmente **negativo** y usuarios totalmente **neutros**.

Los usuarios con opiniones extremas sobre las vacunas son los que escriben con mayor frecuencia sobre este tema en Twitter (207), por lo que el contenido de las redes sociales no representa el **continuum** de reticencia vacunal descrito en la población general por el grupo SAGE (20).

2. Resultados del estudio de los usuarios con más tuits

Dado el amplio rango de tuits por usuario, consideramos importante estudiar los **perfiles** de los usuarios que generaron más contenido sobre vacunas.

Evidenciamos que, en ambos idiomas, los profesionales sanitarios y los usuarios institucionales lideraron la lista de usuarios con más tuits **positivos**, mientras que los usuarios que enviaron más mensajes

reticentes eran *bots*, asociaciones de usuarios o cuentas personales que no pertenecían a profesionales sanitarios.

No obstante, resulta inquietante que uno de los diez usuarios con más tuits negativos en castellano fuera un **profesional sanitario** con 110 600 000 seguidores en el momento del estudio. Los adultos jóvenes consideran que los profesionales sanitarios son la fuente más fiable de información sobre salud. (208) Por ello, es indispensable que los profesionales de salud pública identifiquen quiénes son los profesionales sanitarios reticentes que participan activamente en las redes sociales y realicen intervenciones dirigidas específicamente a este colectivo.

Identificamos que eran **bots**: dos de los diez usuarios con más tuits negativos en inglés, uno de los diez usuarios con más tuits positivos en inglés y dos de los diez usuarios con más tuits en inglés sin distinción de la polaridad. Estos dos últimos *bots* compartieron contenido predominantemente neutro. Es decir, no sólo se utilizaron *bots* para difundir contenido negativo sobre vacunas, sino también para compartir contenido positivo y neutro sobre vacunas.

Broniatowski et al. (62) investigaron el comportamiento de los *bots* en un grupo de tuits sobre vacunas entre julio de 2014 y septiembre de 2017. Identificaron que los *bots*:

-  Se usan frecuentemente para compartir contenido neutro sobre las vacunas, como por ejemplo noticias.
-  Tuitean con mayor frecuencia contenido polarizado que los usuarios que no son *bots*.
-  Aumentan el efecto de las cámaras de eco.

- ✍ En total, generan igual cantidad de contenido positivo que negativo, probablemente para promover el debate entre los usuarios reticentes y los usuarios provacunas.
- ✍ Los *bots* que difunden contenido negativo respecto a las vacunas tienen mayor probabilidad de contener enlaces maliciosos.

Para luchar contra la desinformación que propagan los *bots*, podemos utilizar herramientas como Botometer para distinguir a los *bots* que difunden mensajes negativos, denunciar estas cuentas a Twitter para su suspensión y avisar desde las cuentas oficiales del riesgo que conlleva acceder a los enlaces que comparten este tipo de *bots*.

E. Resultados del Análisis de Temas

1. Tuits según la polaridad y la variable “campana”

Los usuarios prefirieron expresar sus **ideas polarizadas** para participar en campañas. Fuera de las campañas a favor o en contra de las vacunas, el discurso de los usuarios fue predominantemente neutro (92,92%). Teniendo en cuenta que los mensajes desligados de campañas constituyeron el 63,34% de los tuits recopilados y que el 33,39% de los tuits de las campañas de promoción de las vacunas utilizaron mensajes neutros, entendemos por qué los tuits neutros predominaron en la base de datos.

Mientras que el 58,62% de los tuits con **polaridad negativa** pertenecían a campañas contra las vacunas, los usuarios no reticentes realizaron el 96,25% de sus aportaciones con polaridad positiva dentro de campañas de promoción de las vacunas.

Es decir, los individuos reticentes tuvieron una menor inclinación por asociar su opinión sobre las vacunas con campañas.

Desde el punto de vista de salud pública, es favorable que los tuits negativos estén asociados en menor medida a campañas contra las vacunas, ya que es probable que estos mensajes tengan una menor difusión a través de Twitter, a menos que su autor sea un usuario con muchos seguidores. De la misma manera, asociar a los tuits positivos **hashtags** o términos de campañas de promoción de las vacunas ayuda a aumentar su visibilidad.

Con el fin de comprobar si las campañas de promoción de las vacunas iban seguidas de una disminución de los tuits reticentes y un aumento de los positivos no incluidos en campañas, estudiamos la correlación entre las **series temporales**.

Observamos que los tuits positivos que no pertenecían a **campañas** aumentaban en los meses en los que había más tuits asociados a campañas de promoción de las vacunas. Pero simultáneamente, también aumentaban las campañas contra las vacunas y los tuits reticentes que no pertenecían a campañas. Sin embargo, no hubo cambios significativos mantenidos en el tiempo. Por tanto, el impacto de las campañas de promoción de las vacunas en el contenido polarizado sobre vacunas en Twitter parece ser mínimo.

No obstante, no podemos valorar el **impacto real** que tuvieron las campañas de promoción de las vacunas en la confianza y la aceptación de la vacunación por parte de los lectores de estos mensajes únicamente con los datos obtenidos directamente de Twitter. Para poder hacerlo, deberíamos realizar encuestas sobre la

duda vacunal a los usuarios de Twitter antes y después de realizar campañas de promoción de las vacunas.

Las **estrategias** para abordar la reticencia vacunal se agrupan en tres categorías (209):

-  Estrategias basadas en **incentivos no económicos** (comida o bienes) a los individuos de la comunidad que se vacunen.
-  Estrategias basadas en **recordatorios**: llamadas telefónicas, mensajes de texto o cartas a la población diana para recordarles que tienen la posibilidad de vacunarse.
-  Estrategias basadas en el **diálogo**: incluyen el uso de medios de comunicación, páginas *web*, redes sociales y la difusión de información a través de los líderes de la comunidad.

La evidencia sobre la **efectividad** de estas estrategias es limitada, ya que la mayoría de los estudios investigan la reticencia hacia la vacuna antigripal, la vacuna del VPH y las inmunizaciones infantiles en EE. UU., o la reticencia hacia las vacunas frente a la tosferina o la poliomielitis en países con ingresos medios o bajos. En estos contextos, las estrategias que han demostrado mayor efectividad son las multicomponente y las basadas en el diálogo. (209)

Cairns et al. (210) sugieren que las **estrategias multicomponente** que incluyen el uso de Internet para la promoción de la vacunación son efectivas. En cuanto a las estrategias de diálogo realizadas exclusivamente en las redes sociales, la evidencia para determinar su efectividad es insuficiente y de baja calidad. (209,211)

Los estudios que se han realizado sobre las **estrategias implantadas en redes sociales** sugieren que son útiles en usuarios habituales de este tipo de redes que tienen dudas sobre las vacunas,

pero no las rechazan totalmente, y que ya han iniciado el plan de vacunación. (209)

2. Determinantes de reticencia vacunal

Los **temas más frecuentes** en nuestra base de datos fueron la evidencia científica (30,17%) y la seguridad (19,23%), con un mayor porcentaje de positividad que de negatividad hacia ambos temas en la mayoría de los años.

Llama la atención el predominio de la **reticencia** sobre la positividad hacia estos temas en **2015** y **2016**. El aumento de la negatividad hacia ambos temas durante estos dos años pudo estar influido por varios factores:

-  El brote de sarampión en el parque Disney de California entre diciembre de 2014 y febrero de 2015 (212).
-  Los debates sobre el autismo y las vacunas que se produjeron entre los candidatos a las elecciones presidenciales en EE. UU. en el último cuatrimestre de 2015.
-  Las controversias en 2016 por el documental de Wakefield y por el proyecto de ley 277 del Senado de California (213), que pretendía eliminar las exenciones de vacunación por causas no médicas en este Estado y que, finalmente, fue aprobado.

Aunque el **orden** de los temas más frecuentes fue similar entre los tuits de las campañas de promoción de las vacunas, los tuits de las campañas contra las vacunas, los tuits positivos que no pertenecían a campañas y los tuits negativos que no pertenecían a campañas; los usuarios reticentes dieron mayor importancia a la evidencia científica,

la seguridad, la confianza y los aspectos ético-legales de la vacunación.

En base a estos resultados, creemos que sería adecuado aumentar el contenido sobre estos cuatro temas en las campañas de promoción de las vacunas, especialmente sobre la evidencia científica.

Estos mensajes pueden ir o no asociados a **experiencias personales** de miembros de la comunidad. Según Steffens et al. (57), los mensajes promotores de vacunas más eficaces son aquellos que combinan evidencia científica con elementos que permiten a los usuarios generar asociaciones emocionales, tales como historias reales.

En contraposición, Cuesta-Cambra et al. (169) alegan que los **procesos cognitivos**, y no los emocionales, son los responsables de los cambios de opinión de los individuos tras exponerse a mensajes polarizados sobre vacunas. Por tanto, es necesario continuar estudiando si realmente el contenido emocional aumenta la efectividad de los mensajes promotores de las vacunas.

Por otra parte, según un informe realizado por Lane et al. (201), los tres motivos más frecuentes de desconfianza vacunal en el **mundo** entre 2014 – 2016 fueron:

-  **Primero:** las dudas sobre el riesgo-beneficio y la evidencia científica (22% en 2014, 23% en 2015 y 2016).
-  **Segundo:** el nivel de conocimiento insuficiente y la falta de conciencia sobre la importancia de la vacunación (15% en 2014, 13% en 2015 y 10% en 2016).



Tercero: los aspectos religiosos, culturales, de género y socioeconómicos (10% en 2014, 9% en 2015 y 12% en 2016).

En nuestro estudio, los tres **temas** más comentados en los tuits **reticentes** fueron la evidencia científica sobre las vacunas (n = 69 070, 51,98%), la seguridad de la vacunación (n = 51 686, 38,90%) y los aspectos que influyen en la autonomía de los individuos para aceptar o rechazar la vacunación (n = 9 891, 7,44%). Por tanto, los dos determinantes de reticencia vacunal más prevalentes entre los usuarios de Twitter en el periodo de estudio coinciden en orden de importancia con lo informado por Lane et al. (201).

En cuanto a la evaluación del nivel de conocimientos de los usuarios y su grado de concienciación sobre la importancia de la vacunación, no fueron objeto de nuestro estudio.

Identificamos una menor duda determinada por aspectos religiosos y culturales (n = 1 775, 1,34% de los tuits reticentes) y una mayor reticencia relacionada con los aspectos éticos, políticos y legales de la vacunación. Estas **diferencias** podrían deberse al hecho de que los usuarios de Twitter tienen un perfil socio-cultural y demográfico distinto al de la población general (202) y a las desigualdades en el acceso a Internet (65) que limitan el uso de las redes sociales y, por tanto, impiden que toda la población esté representada en nuestro estudio.

Los usuarios de Twitter suelen tener un mayor interés por **temas políticos** (202) que la población general, lo que justifica que el porcentaje de reticencia vacunal relacionada con aspectos políticos, éticos y legislativos sea mayor en nuestro estudio.

En cuanto a la reticencia vacunal por **creencias religiosas** en la población general, es más frecuente en personas mayores de 65 años e individuos de países de WPR como Mongolia (50,5%), Tailandia (44,4%) y Vietnam (31,8%). (193)

La infraestimación de la reticencia vacunal mundial por motivos religiosos en nuestro estudio probablemente se deba a:

-  La baja mediana de edad de los usuarios de Twitter (133).
-  No incluimos los idiomas oficiales de la mayoría de los países de la región WPR, por lo que sólo el 2,91% de los tuits de nuestra base de datos procedían de esta zona.

La **heterogeneidad** de los motivos de la reticencia vacunal entre las distintas localizaciones, tanto en la población general (193,201,214) como en Twitter, pone en manifiesto la complejidad de la **reticencia vacunal** y la necesidad de evaluarla considerando el contexto socio-cultural.

Por ejemplo, podemos entender que los usuarios de Twitter de la región AMR, especialmente de **EE. UU.**, hayan expresado una mayor reticencia debido a los **aspectos éticos** y **legales** de la vacunación, si tenemos en cuenta que en EE. UU., la vacunación es obligatoria para poder acceder a la escolarización, lo cual ha generado numerosos debates políticos y mediáticos a lo largo de la historia (215).

Se ha evidenciado una mayor reticencia vacunal en países con programas de **vacunación obligatoria**, lo que podría sugerir que la implantación obligatoria de la vacunación reduce la confianza de la sociedad en este programa sanitario. (193)

En **España**, el cumplimiento del calendario vacunal oficial no es obligatorio, sino que la legislación actual transfiere a los padres o tutores legales de los niños la decisión de aceptar o rechazar la vacunación. Así y todo, las coberturas vacunales en España son superiores a las de algunos países con vacunación obligatoria. (216)

Aunque se han publicado pocos datos sobre la **prevalencia** de la reticencia vacunal en España y su impacto, (200) Larson et al. (193) sugieren que la duda vacunal en nuestro país es baja. En 2016, de 67 países ordenados de mayor a menor confianza en las vacunas, **España** ocupó el puesto 37 respecto a la creencia en la importancia de la vacunación infantil y entre el puesto 25 - 28 respecto a la confianza en la seguridad y la efectividad de las vacunas, así como la compatibilidad de las vacunas con las creencias religiosas de los encuestados.

En nuestro estudio, España ocupó el puesto número **30** en cuanto a **porcentaje de personas reticentes** entre los 40 países de la ONU con un número igual o mayor a 1 000 tuits. Ahora bien, los usuarios reticentes de España fueron más proclives a confirmar su desconfianza hacia la vacunación en general, mientras que los usuarios de otros países prefirieron manifestar sus dudas hacia aspectos concretos de la vacunación. De esta forma, comprobamos como los aspectos socioculturales influyen en la manera de expresar la reticencia vacunal.

A pesar de la existencia de reticencia vacunal en **España**, las **coberturas vacunales** alcanzan los objetivos establecidos por el Ministerio (217), excepto para la vacuna frente a la **varicela**, la vacuna frente al **VPH** y la vacuna **SRP**.

El Comité de Bioética de la Asociación Española de Pediatría (216) atribuye la **alta adherencia** a la vacunación en España a las siguientes razones:

-  La gratuidad de las vacunas incluidas en el calendario de vacunación sistemático.
-  La participación del personal sanitario de Atención Primaria, teniendo especial relevancia la labor del colectivo de enfermería.
-  La implicación de las sociedades científicas en la formación y difusión de información entre los facultativos de pediatría y la población general.
-  Las campañas de promoción de la vacunación.
-  La gran implicación social junto con un movimiento antivacunas escaso.

Si en el **futuro** se produjera un descenso de las tasas de vacunación sistemática como consecuencia de la desconfianza vacunal, el Comité de Bioética de España (218) contempla la posibilidad de implantar la vacunación obligatoria.

No obstante, la vacunación obligatoria genera **conflictos éticos** entre el principio de autonomía – referente al derecho de los padres a vacunar o no a sus hijos según sus creencias - y el principio de justicia – al poner en riesgo la inmunidad de grupo al ejercer el derecho de no recibir vacunas que confieren protección colectiva. (216) Por ello, el Comité de Bioética de España (218) apoya el mantenimiento de la vacunación no obligatoria por el momento y, en su lugar, recurrir a estrategias basadas en el empoderamiento de los ciudadanos.

3. Tipo de vacuna

Consideramos **importante** desglosar la reticencia por vacunas, ya que, según la propia definición de reticencia vacunal, (20) la duda vacunal varía según la vacuna. Una persona puede rechazar un tipo de vacuna concreta, pero aceptar otras.

En la población general, las vacunas que más **escepticismo** generan son aquellas en las que ha habido más **debate científico**, tales como la vacuna contra el VPH, la vacuna contra la influenza o la vacuna SRP. (200,219) Por el contrario, las vacunas frente al tétanos y la polio son las más aceptadas, debido a que se perciben como **enfermedades de mayor riesgo**, aunque no sean las más frecuentes. (200)

En nuestro estudio, hallamos el mismo **patrón** de reticencia y aceptación entre las vacunas incluidas en el calendario vacunal. Por tanto, a pesar de las diferencias en las particularidades de la reticencia vacunal entre los usuarios de Twitter y la población general debido a la disparidad entre sus características sociodemográficas, el sentimiento predominante hacia cada vacuna está alineado. Esto nos indica la **utilidad** de la minería de opinión con *big data* en redes sociales como herramienta rápida y costo-efectiva para la **monitorización** de la reticencia vacunal.

La vacuna que generó más dudas fue la **SRP**, probablemente condicionado por los antecedentes históricos que comentamos en la introducción de este trabajo y a los distintos debates mediáticos sobre ella que ha habido a lo largo de los años.

En el caso de la vacuna contra a la **influenza**, observamos que la seguridad fue el motivo más frecuente de duda vacunal, coincidiendo

con las preocupaciones expresadas por las personas encuestadas en distintas partes del mundo en otros estudios (220–222) fuera de las redes sociales.

Respecto a la vacuna contra el **VPH**, junto con la vacuna contra la hepatitis B, son las únicas que protegen frente a infecciones por virus potencialmente carcinógenos. (223) No obstante, la falta de conocimiento sobre ésta y el miedo a sus posibles efectos adversos constituyen una gran barrera en la vacunación, (224) de modo que las coberturas vacunales en la mayoría de países son subóptimas (225).

Las personas que usan las **redes sociales** son más conscientes de la existencia del VPH y su vacuna que aquellos que no las usan, lo que puede ser indicativo de una mayor exposición a información sobre ambos temas a través de este medio. (226)

Pero el conocimiento sobre el cáncer por VPH (226) y la aceptación de la vacunación no varían (227), probablemente debido a la **desinformación** que circula en las distintas plataformas sociales (181,228–230).

Walker et al. (231) identificaron que tanto las redes sociales como los medios de comunicación contribuyen a **aumentar** la duda sobre la vacuna VPH. Según Dunn et al., (179) los usuarios **expuestos** a mensajes que cuestionan la seguridad de la vacuna contra el VPH tienen mayor probabilidad de tuitear mensajes similares durante los seis meses siguientes que los usuarios que han leído mensajes positivos o neutros. Por su parte, Margolis et al. (232) describieron que la exposición a historias **negativas** sobre la vacuna VPH en redes sociales aumenta el **rechazo** a la vacunación.

En cambio, el aumento del **conocimiento** sobre la infección por VPH, el cáncer y la vacuna del VPH generan una mayor **aceptación** de la vacunación. (233) Por ello, es crucial incluir elementos educativos en el abordaje de la duda hacia la vacuna contra el VPH.

4. Población diana

La vacunación del **infantil** fue un tema diez veces más frecuente que el segundo tópico más nombrado, la **vacunación del viajero**. A pesar de que varias enfermedades del viajero son prevenibles - como la fiebre amarilla, la fiebre tifoidea o el cólera – la mayoría de los viajeros desconocen los riesgos asociados a los desplazamientos a zonas endémicas y la posibilidad de acudir a la consulta del viajero (234), lo que justifica la gran diferencia entre el número de tuits sobre la vacunación del viajero y la vacunación infantil.

Asimismo, aunque entre el 35 -52% de los **viajeros** que desean información sobre los riesgos sanitarios derivados del viaje utilizan **Internet** como fuente de información, (234,235) sólo uno de cada seis de ellos recurre a redes sociales (235).

Si bien los viajeros que buscan información en las redes sociales tiene más probabilidad de **rechazar** la vacunación, (235) la información disponible en Twitter sobre la vacunación del viajero es predominantemente neutra o positiva. Sólo el 1,64% de los tuits escritos entre 2011 - 2019 que mencionaban la vacunación del viajero eran negativos. La **ignorancia** sobre la **existencia** de la vacunación del viajero por parte de la población general (234) y la **búsqueda** de una **protección individual** en los pacientes que solicitan consejo antes del viaje (236), hacen que los usuarios que escriben sobre este

tema en Twitter sean los que tienen mayor **conocimiento** sobre el mismo y menor reticencia.

Las **vacunas** del **viajero** que generaron mayor reticencia entre los usuarios de Twitter no son las más prescritas en la consulta previa al viaje, (237,238) sino las que también generan mayor reticencia en la población general. Un ejemplo de ello es la **mayor reticencia** frente a la vacuna SRP que frente a la vacuna contra la fiebre amarilla en los tuits sobre la vacunación del viajero, a pesar de ambas son vacunas atenuadas y la vacuna frente a la fiebre amarilla se prescribe con mayor frecuencia a los viajeros que la vacuna SRP (237,238).

En un estudio observacional de cinco meses de duración realizado por Alves et al., (239) aproximadamente el 40% de los viajeros que demandaron atención médica antes del viaje lo hicieron para solicitar la vacunación frente a la **fiebre amarilla**, aunque el 20-30% de ellos no iban a viajar a una zona endémica. Por tanto, vemos como la percepción social sobre cada vacuna y la percepción del riesgo frente a cada enfermedad inmunoprevenible por parte de los pacientes parecen influir en el grado de aceptación y demanda de cada vacuna.

Es probable que el **desconocimiento** de la población y de los **profesionales sanitarios** sobre las indicaciones de la vacunación del **adulto** (240,241) también condicionen el bajo número de tuits sobre este tema.

Los estudios sobre la reticencia hacia la vacunación del **adulto** se centran en la vacuna contra la influenza (242–246) o la vacuna antineumocócica (247–250). Pocos autores han realizado **estudios** sobre los motivos de la duda vacunal en general (251,252) o hacia otras vacunas (253,254) en este grupo poblacional, por lo que la

investigación en las redes sociales puede contribuir en gran medida a ampliar los conocimientos sobre este campo aún poco explorado.

En cuanto a la vacunación frente a enfermedades que pueden afectar al feto o al recién nacido, la vacuna SRP, la vacuna contra la varicela y la vacuna contra el VPH están recomendadas **antes del embarazo**. (255) De ellas, esta última es la que generó más dudas, debido a que a las mujeres les inquieta que pueda tener efectos negativos sobre su fertilidad y la seguridad de esta vacuna durante el embarazo (256).

En cambio, la vacunación contra la influenza y la tosferina está recomendada **durante el embarazo**. (255) Aunque varios estudios demuestran la seguridad de la inmunización con ambas vacunas durante el embarazo (257–260), este tema continúa preocupando a los usuarios de las redes sociales (170) y a la población general (261,262), lo cual se refleja en una cobertura vacunal mundial subóptima (263–266).

Martin et al. (170) analizaron las conversaciones sobre la vacunación de las **embarazadas** en Twitter y otros sitios sociales entre el 1 de noviembre de 2018 y el 30 de abril de 2019. Identificaron que la negatividad hacia la vacuna contra la influenza empeoró debido a la difusión de un **rumor** acerca de que la FDA había admitido tener insuficientes datos sobre la vacunación prenatal. También detectaron que los usuarios **extrapolan** los efectos adversos de las vacunas contraindicadas durante el embarazo a las vacunas recomendadas.

Por otra parte, Ford et al. (267) encuestaron a 305 mujeres usuarias de las redes sociales, de las cuales 64 (20,8%) habían **buscado** información sobre vacunas en los sitios sociales. Este grupo presentó

una probabilidad 58% **menor** de aceptar la vacunación contra la tosferina durante el embarazo.

Así pues, es esencial tomar medidas para disminuir la difusión y el impacto de la información **negativa** sobre la vacunación en embarazadas en las redes sociales.

5. Instituciones mencionadas

Comprobamos que las instituciones más mencionadas por los usuarios reticentes fueron las **agencias de salud pública**. Si los **usuarios** no confían en los emisores de los mensajes a favor de las vacunas, **difícilmente** escucharán sus argumentos, con lo cual no modificarán positivamente su percepción acerca de las vacunas. Por ello, sería de interés que cada institución identificara cuál es la **percepción** de los usuarios reticentes acerca de ella y, en caso de ser negativa, averiguar las **causas** e implementar medidas correctoras adaptadas a cada una de ellas.

En nuestro análisis semántico, observamos una gran **desconfianza** hacia los **CDC**, que además fueron la **institución** más mencionada en los tuits en inglés. El decimosexto unigrama más frecuente fue "*cdcwhistleblower*", término que suele usarse en los tuits negativos y que raramente aparece en los tuits positivos (62,207).

El origen de este **término**, y la desconfianza inherente de las personas que lo usan, se remonta a **2014**, tras la filtración de unas conversaciones telefónicas en las que, supuestamente, el Dr. Thompson, científico de los **CDC**, reconocía que ocultaron **datos** que demostraban la **relación** causal entre la inmunización con la vacuna SRP y el **autismo** en los niños afroamericanos. (268) Si en un futuro los **CDC** logran convencer a la población de la **falsedad** de las

acusaciones, es probable que sus intervenciones contra la **reticencia** vacunal tengan mayor **efectividad**.

Es alentador que las instituciones menos nombradas en los tuits reticentes fueran las **instituciones sanitarias asistenciales**, ya que como hemos comentado antes, los profesionales sanitarios son considerados por la población general (63,64,208) como la **fuentes** de información sobre salud más **confiable**. Por este motivo, es imprescindible que el abordaje de la reticencia vacunal no se realice sólo a nivel poblacional, sino también de manera **individualizada** en el ámbito asistencial.

6. Otros conceptos relacionados con la reticencia vacunal

a) Autismo

Numerosos estudios epidemiológicos (38–44) han demostrado que **no** existe **relación causal** entre la inmunización con la vacuna SRP y el desarrollo de autismo en los niños, sino una **coincidencia** temporal. Debido a la evolución natural de la enfermedad, el diagnóstico de autismo se produce en la infancia temprana, misma época en la que se administra la mayoría de la vacunación sistemática.

Pese a ello, la publicación del artículo de **Wakefield et al.** (34) en el que sugería una relación causal entre la inmunización con la vacuna SRP y el autismo continúa teniendo un efecto negativo sobre la confianza en esta vacuna, especialmente en **EE. UU.** La gran repercusión mediática de la controversia generada en torno al artículo ha contribuido en gran medida a que los padres continúen evitando la administración de la vacuna SRP a sus hijos por el **miedo** a que desarrollen **autismo** (269).

Esto motivó a Hviid et al. (270) a actualizar la evidencia científica sobre este tema en **2019**. Realizaron un estudio de cohortes retrospectivo de los niños daneses nacidos entre el 1 de enero de 1999 y el 31 de diciembre de 2010, en el que no hallaron aumento de riesgo de autismo por la inmunización con la vacuna SRP.

Sin embargo, los usuarios reticentes de nuestro estudio manifestaron una gran desconfianza en la **evidencia científica** que desvincula al autismo de la vacunación, por lo que probablemente realizar más investigaciones sobre este tema no sea la solución más costo-efectiva (271).

Ciertamente, el **apoyo** de los resultados de Wakefield et al. (34) por parte de diversas personalidades públicas, incluso después de su retracción, (272) y las **teorías conspiratorias** sobre el encubrimiento por parte del CDC de la relación causal entre la vacuna SRP y el autismo (268) suponen un gran **obstáculo** en nuestra lucha contra la desconfianza en la información científica que contradice los hallazgos de Wakefield et al. (34).

Por consiguiente, aumentar el **conocimiento** de la población sobre cómo se produce la evidencia científica acerca de la **seguridad** de las **vacunas** y mejorar la imagen de las instituciones podría ser una estrategia más efectiva que invertir más recursos en la investigación de un tema sobre el que ya hay suficiente evidencia científica, en la que los usuarios no creen.

Por otra parte, al igual que en el estudio de Tomeny et al. (183), observamos que los **picos** de tuits **negativos** sobre autismo y vacunas coincidieron temporalmente con épocas en las que se

publicaron un gran número de **noticias** sobre eventos relacionados con la vacunación.

Se ha demostrado que el **interés** de los usuarios de Twitter y Facebook sobre algunas vacunas, incluida la del sarampión, aumenta cuando se publican más **noticias** sobre ellas. (169) Por ello, creemos que es necesario realizar estudios enfocados a establecer si hay un vínculo entre el volumen de noticias y el número de tuits sobre autismo y vacunas, y, si se confirma, reforzar la monitorización de las conversaciones en Twitter cuando los medios de comunicación publiquen más noticias sobre este tema.

Finalmente, aunque no sea un tema muy mediático, no debemos olvidar la existencia del **temor** de las embarazadas ante un posible desarrollo de autismo en los fetos por la vacunación durante el **embarazo** (170). Este miedo es mayor frente a la vacuna contra la **influenza**, a pesar de que varios autores (273,274) han demostrado la seguridad de esta vacuna.

b) Cáncer. Composición de las vacunas.

Gracias al análisis temático, conocemos con mayor profundidad las particularidades de **dudas** de los usuarios reticentes acerca de los componentes de las vacunas y la creencia errónea de su potencial cancerígeno.

Podemos utilizar esta información para diseñar **intervenciones educativas** dirigidas especialmente a los usuarios de las localizaciones con más dudas acerca de estos temas, enfocadas a las vacunas cuyos componentes generan mayor desconfianza, haciendo énfasis en la seguridad y la evidencia científica.

c) Homeopatía.

Los padres reticentes suelen **apoyar** el uso de la homeopatía. (219) Pero el apoyo a las medicinas alternativas no implica el **rechazo** de la medicina científica y la vacunación, como observamos en nuestro conjunto de datos.

Otro ejemplo lo aportan Eizayaga et al. (275) Preguntaron a 512 médicos homeópatas sobre su opinión acerca de las vacunas y el 75,60% de ellos **afirmaron** que “la vacunación es segura, efectiva y necesaria”.

De hecho, la duda vacunal está más relacionada con la **desconfianza** en la medicina científica que con la confianza en las medicinas alternativas, incluida la homeopatía. Por tanto, **no** debemos considerar a las medicinas alternativas como **obstáculos** para la aceptación de la vacunación. En su lugar, debemos centrarnos en **mejorar** la confianza en la medicina tradicional. (276)

F. Resultados del Análisis de la Relación con datos de la OMS

1. Coberturas vacunales frente al VPH y campañas en Twitter

El objetivo final de las campañas de promoción de las vacunas en las redes sociales debe de ser lograr una mayor **aceptación** de la vacunación y que esto se traduzca en un aumento de las **coberturas** vacunales. Aunque la información negativa suele tener mayor **impacto** que la positiva en la opinión de las personas reticentes, la exposición a información positiva en la red ha demostrado aumentar la **aceptación** de la vacunación. (277)

Por ello, aunque en nuestro estudio no observamos cambios significativos en la positividad hacia las vacunas en Twitter en los meses posteriores a las campañas de vacunación en este medio, consideramos que la **efectividad** de estas campañas debe medirse según las variaciones en las coberturas vacunales.

A modo de ejemplo, analizamos el caso de la vacuna contra el **VPH**. Seleccionamos esta vacuna porque es la que tiene coberturas más bajas en la mayoría de los países y la que presenta más variaciones a lo largo de los años.

Entre los seis países estudiados, hallamos una **correlación** significativa, alta en Suiza y muy alta en España y en EE. UU., entre las coberturas vacunales anuales de VPH estimadas por la OMS y el número de tuits de nuestra base de datos pertenecientes a campañas de promoción de la vacuna VPH entre 2012 -2018. Es decir, la **cobertura vacunal** frente a VPH **augmentó** en los años en los que hubo más tuits promotores de esta vacuna en estos tres países.

No obstante, estos resultados deben **interpretarse** con cautela, ya que no podemos descartar la existencia de factores de confusión que hayan podido alterar estos resultados, como la **exposición** a información negativa sobre la vacuna contra el VPH (278), las campañas de promoción de las vacunas en otros medios o los cambios en la accesibilidad a esta vacuna. Debido a que éramos conscientes de que los resultados no eran necesariamente concluyentes por el diseño del estudio, decidimos no ampliar el estudio a otras localizaciones ni a otras vacunas.

Para poder asegurar que una **campaña** de vacunación en Twitter es responsable del aumento de la cobertura vacunal de un tipo de

vacuna concreta en una determinada localización, deberíamos **identificar** previamente a las personas reticentes que se han expuesto a la campaña, **preguntarles** por sus intenciones de vacunarse antes y después de la campaña y realizar un **seguimiento** para comprobar si finalmente se vacunan.

Pese a las limitaciones, nuestros resultados animan a continuar explorando el posible efecto positivo de las campañas en redes sociales sobre las coberturas vacunales.

2. Casos de sarampión y reticencia vacunal en Twitter

Teniendo en consideración que la **incidencia** de sarampión aumentó en todas las regiones de la OMS entre 2017 y 2019 debido a las bajas coberturas vacunales (279), y que la vacuna **SRP** fue la que generó **mayor** reticencia entre los usuarios de nuestro estudio, decidimos **estudiar** si hubo variaciones en el interés y la duda hacia esta vacuna según el número de casos mensuales de sarampión en los tres países con más tuits reticentes respecto a la vacuna SRP (EE. UU., Canadá y Reino Unido).

Algunos autores habían estudiado en el pasado la **correlación** entre el número de casos de sarampión en una localización concreta y el volumen de mensajes en redes sociales, con diversas limitaciones y resultados opuestos. Por ejemplo, **Deiner et al.** (175) no hallaron correlación entre el número de casos semanales de sarampión en EE. UU. y el número de mensajes que expresaban reticencia vacunal en Twitter ni en Facebook entre el 4 de enero de 2009 y el 27 de agosto de 2016. No obstante, no focalizaron el estudio hacia la reticencia frente a la vacuna SRP.

Por su parte, **Mollema et al.** (171) comunicaron una correlación moderada significativa entre el número de casos semanales de sarampión en Países Bajos y el número semanal de tuits. Sin embargo, limitaron el estudio a treinta y una semanas en el contexto de un brote por esta enfermedad inmunoprevenible en 2013.

En **dos de los tres países** estudiados en la presente tesis ((EE. UU. y Canadá), evidenciamos que el **interés** general por la vacuna SRP y la **negatividad** hacia ella aumentaron o disminuyeron durante los tres primeros meses tras la variación directamente proporcional del número de casos nacionales de sarampión. No disponíamos de la geolocalización de aproximadamente la mitad de los tuits totales y negativos sobre la vacuna SRP, lo cual limitó la interpretación de nuestros resultados y podría explicar la ausencia de correlación en Reino Unido.

Asimismo, es probable que los **resultados** hayan estado influidos por la cobertura de los brotes de sarampión por parte de los medios de comunicación. (171,175) Cuando el conocimiento sobre el aumento de casos se vuelve de dominio público, los usuarios a **favor** de las vacunas recurren a las **redes sociales** para culpar a los usuarios que rechazan la vacunación por motivos no médicos. (175) Por el contrario, los usuarios reticentes suelen aumentar sus intervenciones durante los brotes para expresar su desconfianza entorno a la seguridad y la efectividad de las vacunas. (175)

Por ello, es importante **monitorizar** las conversaciones sobre las vacunas en las redes sociales y aprovechar el aumento en el interés por éstas en el contexto de brotes para:

-  Detectar la desinformación y tomar medidas para detener su difusión.
-  Identificar comunidades reticentes y conocer sus opiniones, para realizar un abordaje contra la reticencia vacunal individualizado a corto plazo.
-  Aumentar la exposición de los usuarios a información veraz y contrastada sobre la seguridad y la evidencia científica de las vacunas.

G. Limitaciones del Estudio

Una de las principales limitaciones del estudio se debe al diseño **retrospectivo** del estudio. Los usuarios de Twitter pueden eliminar los mensajes de su autoría en cualquier momento después de publicarlos y pueden borrar/desactivar sus cuentas o cambiar la configuración de visibilidad de sus perfiles (públicos o privados) fácilmente cuando lo consideren conveniente. Además, Twitter se reserva el derecho de suspender temporal o indefinidamente las cuentas que incumplan sus reglas de funcionamiento.

Todo ello puede ocasionar que el conjunto de datos descargado retrospectivamente no sea exactamente el mismo que si se hubiera obtenido la información a tiempo real, ya sea por la desaparición de los mensajes borrados o privatizados a lo largo del tiempo o por la inclusión de mensajes que actualmente son públicos, pero no lo eran

cuando fueron escritos. En consecuencia, no podemos evaluar con exactitud de manera retrospectiva la reticencia vacunal en Twitter en un momento concreto del tiempo. Con el fin de reducir las consecuencias de esta limitación en este estudio, hemos analizado un periodo de tiempo prolongado y un gran volumen de datos.

La segunda limitación la constituye el procedimiento de **localización** de los tuits. Twitter no permite acceder al histórico de localizaciones de los usuarios. Al no recoger los datos en tiempo real, no pudimos registrar las localizaciones de los usuarios en el momento en el que escribieron cada mensaje. En su lugar, utilizamos la localización de los usuarios en el momento de la descarga de los datos (mayo de 2019), que puede coincidir o no con la localización del usuario en el momento de la publicación de un tuit concreto.

Asimismo, el análisis por países probablemente esté sesgado debido a los **idiomas** seleccionados (sólo inglés y castellano) y el método de recuperación de la ubicación (localización autodeclarada en el perfil del usuario para aquellos usuarios que no disponían de geolocalización pública). La localización del perfil y la localización real de los usuarios pueden diferir y, en numerosas ocasiones, no es posible distinguir entre los datos verdaderos y los datos falsos escritos en este campo. Un análisis exhaustivo de la opinión en Twitter en un país específico debe considerar el idioma local y utilizar métodos de geolocalización basados en los parámetros de GPS e IP de los usuarios en tiempo real.

Otra gran limitación del estudio es inherente a la clasificación de textos. La identificación de la **ironía** y el **sarcasmo** continúa siendo un desafío, no sólo para los métodos de clasificación automáticos, sino también para los manuales, por lo que los resultados pueden no

ser exactos. El uso de un léxico clasificado manualmente como complemento al aprendizaje automático en nuestro estudio reduce el efecto de esta limitación en los resultados generales.

Cabe destacar que la mayoría de los algoritmos basados en aprendizaje automático, incluido el nuestro, cuentan con la dificultad añadida de no disponer del **contexto** para realizar la clasificación, ya que se suelen eliminar los emoticonos, las imágenes y las *URLs* asociadas a los textos. Estos elementos podrían agregarse al análisis para aumentar el rendimiento de la clasificación.

No obstante, en el caso de los emoticonos, aunque son útiles en el AS a nivel de frase, pueden disminuir la precisión y exactitud del AS basado en aspectos. Por ejemplo, la frase “el documental *Vaxxed* es interesante, feliz de que se haya distribuido”, acompañada por un emoticono de felicidad, tiene una polaridad positiva en el AS a nivel de frase, pero negativa en el AS sobre la vacunación. (186) Por ello, en este estudio prescindimos de los emoticonos.

Por otra parte, no pudimos **evaluar** la **precisión** del **algoritmo** de clasificación por **temas**. Al tratarse de una clasificación basada exclusivamente en léxico, no disponíamos de un conjunto de pruebas para evaluar los datos, a diferencia del algoritmo híbrido que se creó para clasificar la polaridad. La evaluación de la precisión de los métodos de clasificación basados en léxico requiere de un conjunto de tuits clasificados mediante otro método, manual o automático, validado previamente en los datos de estudio (no disponible en este caso).

Restringimos nuestro análisis a los tuits etiquetados con alguno de los **hashtags** preseleccionados, por lo que desconocemos si los resultados son extrapolables o no a conjuntos de tuits sobre vacunas no etiquetados o etiquetados con otros **hashtags**.

Por esta misma razón, es probable que nuestra base de datos no contuviera todos los tuits escritos sobre cada vacuna. En consecuencia, no podemos asegurar si hubo relación entre todos los tuits sobre una vacuna y sus coberturas vacunales o el número de casos de la enfermedad infecciosa frente a la que protege.

Este es uno de los motivos por los que limitamos el apartado VI de los resultados a dos casos de estudio en unas pocas localizaciones, a modo de ejemplos de los posibles efectos beneficiosos de las campañas de promoción de las vacunas en Twitter sobre las coberturas vacunales y de las posibles repercusiones de la reticencia vacunal en las redes sociales sobre los casos de una enfermedad inmunoprevenible.

Por último, es posible que los usuarios de Twitter no sean **representativos** de la población en general. Además de las diferencias demográficas, se debe tener en cuenta las desigualdades sociales que limitan el acceso a Internet (65). Por tanto, los resultados no son extrapolables. El análisis de redes sociales no debe aplicarse como una herramienta independiente para abordar la duda ante las vacunas, sino que debe ser un componente de estrategias más amplias que incluyan tanto intervenciones digitales como tradicionales.

H. Fortalezas del Estudio y Futuras Líneas de Investigación

Tras una amplia revisión de la literatura publicada sobre la reticencia vacunal en Twitter, identificamos que éste es el **único estudio** que analiza la polaridad de sentimientos hacia las vacunas **a nivel de oración** utilizando un enfoque **híbrido** en más de un **millón** de tuits escritos a lo largo de varios **años**, con distinción de la **localización**.

Adicionalmente, es una de las pocas investigaciones que incluye un **análisis temático automatizado** de un conjunto de **datos masivos** del campo de la vacunología.

Dado que las redes sociales se han convertido en una fuente común de **información** y **desinformación** sobre vacunas, es responsabilidad de los profesionales de salud pública, junto con los responsables de estas plataformas, realizar un control del contenido publicado en ellas.

Twitter ha sido hasta el momento una de las redes sociales con políticas más flexibles frente a la desinformación sobre vacunas. Ante la proliferación de la desinformación sobre las vacunas contra la COVID-19 a finales de 2020, decidieron **eliminar** los tuits que contenían **información falsa** sobre los efectos adversos de estas vacunas y la ausencia de necesidad de su administración. En el primer cuatrimestre de 2021, anunciaron la implantación de **etiquetas** de alerta en los tuits que contengan **desinformación** sobre estas vacunas.

Desde el inicio de la pandemia de **COVID-19**, hemos estado recopilando prospectivamente los tuits relativos a la vacuna frente al SARS-CoV2 con el **fin** de realizar un **análisis** de la opinión de los

usuarios de Twitter sobre cada una de las vacunas comercializadas y en investigación en los distintos momentos de la pandemia.

Asimismo, estamos **perfeccionando** nuestros **algoritmos** para superar las limitaciones expresadas en el apartado anterior.

Por otra parte, tenemos previsto ampliar retrospectiva y prospectivamente el estudio actual a **otras redes sociales**. Adicionalmente, incluiremos **encuestas tradicionales** para valorar el efecto de la exposición a los mensajes positivos, negativos y neutros sobre vacunas publicados en las redes sociales.

II. CONCLUSIONES

En base a los objetivos propuestos, los resultados obtenidos en el desarrollo de la presente investigación permiten formular las siguientes conclusiones:

1. Es posible extraer, estructurar y analizar de forma rápida, económica y precisa grandes cantidades de datos de las conversaciones de Twitter aplicando técnicas de inteligencia artificial. La información derivada de este tipo de análisis puede ser útil para mejorar las estrategias de salud pública destinadas a disminuir la duda vacunal.
2. El conjunto de datos consistió en 1 499 227 tuits originales, mayoritariamente en inglés y de la región AMR de la OMS. La tendencia creciente en el número mensual de tuits sobre vacunas a lo largo del estudio sugiere un aumento en el interés hacia la vacunación.
3. Los unigramas más frecuentes en inglés fueron: *“child”*, *“health”*, *“get”*, *“autism”* y *“measles”*; en castellano: “salud”, “gripe”, “niños”, “antivacunas” y “enfermedades”. Los bigramas más frecuentes en inglés fueron: *“flu-vaccine”*, *“vaccination-need”*, *“save-life”*, *“shot-clinic”* y *“vaccinations-cdc”*; en castellano: “salud-vacunas”, “vacunas-salud”, “vacunación-Américas”, “vacuna-gripe” y “semana-vacunación”. El análisis semántico aportó una visión preliminar de los temas más nombrados en los tuits según el modelado de tópicos.
4. Nuestro algoritmo híbrido presentó altas puntuaciones de precisión, exhaustividad y F1 para la clasificación de la polaridad de sentimientos hacia las vacunas en las tres categorías

analizadas (neutro, negativo y positivo) en el conjunto de datos de prueba. El rendimiento fue mayor que el del algoritmo VADER, no entrenado específicamente en el conjunto de datos de interés. La diferencia entre las métricas de ambos algoritmos sustenta la conveniencia de entrenar algoritmos específicos para cada campo de estudio.

5. En contra de la creencia popular de que el contenido sobre vacunas de Twitter es predominantemente negativo, detectamos que el discurso neutro predominó entre 2011 - 2019. No obstante, también evidenciamos una tendencia creciente hacia la polarización del discurso, sobre todo hacia la positividad. Adicionalmente, identificamos que los miércoles y jueves fueron los mejores días para entablar una conversación directa sobre vacunas con los usuarios de Twitter.

El análisis de los resultados de la clasificación de la polaridad de los sentimientos de los tuits hacia las vacunas según su geolocalización permitió detectar las regiones geográficas con mayor reticencia hacia las vacunas. Destacan AMR y AFR como las regiones con mayor porcentaje de tuits negativos y positivos respectivamente, así como Países Bajos y Suiza por ser los países con una proporción más alta de tuits negativos y positivos respectivamente. Los usuarios de las localizaciones con mayor negatividad hacia las vacunas podrían beneficiarse en mayor medida de intervenciones de salud pública en las redes sociales destinadas a aumentar la confianza en las vacunas.

6. Aunque la mayoría de los tuits no generaron ningún tipo de interacción social (retuits, favoritos/me gusta o respuestas), los tuits positivos fueron difundidos en mayor medida que los tuits

negativos y neutros. Por tanto, es más probable que los usuarios compartan los tuits promotores de las vacunas si el contenido es positivo en lugar de neutro.

7. Estudiamos las características de los 278 858 usuarios. Destaca el mayor número de usuarios no institucionales, sobre todo entre los autores de los tuits negativos. A diferencia de la población general, la mayoría de los usuarios expresaron opiniones extremas sobre las vacunas (totalmente positivas, negativas o neutras). Entre los usuarios con más tuits, encontramos grandes diferencias en la polaridad predominante según el tipo de usuario. Los usuarios con más tuits positivos eran sobre todo profesionales sanitarios o usuarios institucionales, mientras que la mayoría de los usuarios con más tuits negativos eran *bots*, asociaciones de usuarios y cuentas personales (excluyendo las de los profesionales sanitarios).
8. Las campañas de promoción de las vacunas en Twitter no lograron cambios mantenidos en el tiempo en la opinión sobre las vacunas de los usuarios de esta red de *microblogging*. No obstante, la información derivada de Twitter aislada no permite valorar la asociación causal ni los cambios fuera de la plataforma que generan este tipo de campañas.
9. Los estudios temáticos mediante técnicas de AS permiten conocer cuáles son las áreas de la vacunación que suscitan mayor interés y los aspectos que generan mayor duda vacunal. La información obtenida de este análisis permite personalizar las medidas de abordaje de la reticencia vacunal por localización, por tipo de vacuna y por población diana de la vacunación.

En base a nuestros resultados, creemos que es conveniente que los profesionales de salud pública centren sus intervenciones de promoción de las vacunas en Twitter hacia la lucha contra la desinformación sobre la evidencia científica y la seguridad de las vacunas, sobre todo de: la vacuna SRP en niños, adultos y viajeros; la vacuna contra el VPH en la población general y la vacuna contra la influenza en adultos y embarazadas, temas de candente interés especialmente en AMR. En concreto, hay que hacer hincapié en desmentir los mitos que vinculan el autismo con la vacuna SRP y los mitos sobre la inseguridad de los componentes de la vacunación infantil.

10. Comprobamos que las coberturas vacunales frente al VPH y el número de tuits promotores de la vacuna contra el VPH entre 2012 - 2018 presentaron una correlación significativa fuerte y positiva en tres de los seis países analizados (España, EE. UU. y Suiza). Aunque no descartamos la presencia de factores de confusión, los hallazgos sugieren un posible efecto positivo de las campañas en redes sociales sobre las coberturas vacunales.
11. En EE. UU. y Canadá, los tuits que expresaban reticencia hacia la vacuna SRP, entre enero de 2012 y abril de 2019, aumentaron o disminuyeron 0, 1 y 2 meses después de que el número de casos de sarampión variara en el mismo sentido. Hallamos el mismo patrón en el análisis de los tuits sin distinción de la polaridad. La variación en el número tuits sobre una vacuna en una localización concreta podría ser un indicador para incluir en la vigilancia periódica de los casos de las enfermedades inmunoprevenibles en la población de cada área geográfica.

CAPÍTULO VI

BIBLIOGRAFÍA

BIBLIOGRAFÍA

1. OMS, UNICEF, Banco Mundial. Vacunas e inmunización: Situación mundial. 3ª Ed. Ginebra, Organización Mundial de la Salud; 2010.
2. World Health Organization. Ten threats to global health in 2019. [Internet]. 2019 [cited 2020 Sep 12]. Available from: <https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019>
3. Ahmed N, Quinn SC, Hancock GR, Freimuth VS, Jamison A. Social media use and influenza vaccine uptake among White and African American adults. *Vaccine*. 2018;36(49):7556–61.
4. Arrazola Martínez MP, de Juanes Pardo JR, García de Codes Ilario A. Conceptos generales. Calendarios de vacunación sistemática del niño y del adulto en España. Impacto de los programas de vacunación. *Enferm Infecc Microbiol Clin*. 2015 Jan 1;33(1):58–65.
5. Centers for Disease Control and Prevention. Principles of Vaccination. In: Hamborsky J, Kroger A, Wolfe S, editors. *Epidemiology and Prevention of Vaccine-Preventable Diseases*. 13th Ed. Washington D.C: Public Health Foundation; 2015. p. 1–8.
6. World Health Organization. Vacunas [Internet]. Vacunas. 2016 [cited 2019 Apr 29]. Available from: <https://www.who.int/topics/vaccines/es/>
7. Sociedad Española de Geriátría y Gerontología. Vacunación en adultos y mayores: Algoritmos y Protocolos de actuación. SEGG, editor. Madrid; 2018.
8. Teresa Criado M, Sánchez S, Ferreirós CM. Vacunología clásica y nuevas tecnologías en el diseño de vacunas. *Enferm Infecc Microbiol Clin*. 2008 Nov 1;26(9):564–72.
9. Ministry of Health. General immunisation principles. In: *Immunisation Handbook*. 2nd Ed. Wellington: Ministry of Health; 2018. p. 15–38.
10. Comité Asesor de Vacunas de la AEP. Manual de Vacunas en línea de la Asociación Española de Pediatría. Madrid: AEP; 2014.
11. Salleras Sanmartí L. Concepto, clasificación y características generales de las vacunas. In: Salleras Sanmartí L, editor. *Vacunaciones preventivas: Principios y aplicaciones*. 2ª Ed. Barcelona: Masson; 2003. p. 17–37.
12. Salleras L, Domínguez A, Borrás E, Soldevila N. Eficacia protectora de las vacunas y efectividad de las vacunaciones: Introducción a la medición de la protección directa e indirecta. *Vacunas*. 2011 Oct;12(4):136–46.
13. Giglio N, Bakir J, Gentile A. Eficacia, efectividad e impacto en vacunas: ¿es lo mismo? *Rev del Hosp Niños (Buenos Aires)*. 2018;60(268):34–41.
14. Su JR, Duffy J, Shimabukuro TT. Essentials of Vaccine Safety. In: Poland GA, Whitaker JA, editors. *Vaccinations*. 1st Ed. Elsevier; 2018.
15. Centers for Disease Control and Prevention. General Recommendations on Immunization. In: Hamborsky J, Kroger A, Wolfe S, editors. *Epidemiology and Prevention of Vaccine-Preventable Diseases*. 13th Ed. Washington D.C: Public Health Foundation; 2015. p. 9–31.
16. Organización Mundial de la Salud. Cobertura vacunal [Internet]. 2020 [cited 2020 Dec 8]. Available from: <https://www.who.int/es/news-room/fact-sheets/detail/immunization-coverage>
17. WHO-UNICEF. Global and regional immunization profile [Internet]. 2020 [cited 2020 Jan 24]. Available from: https://www.who.int/immunization/monitoring_surveillance/data/g_s_gloprofile.pdf
18. Eskola J, MacDonald NE. How to deal with vaccine hesitancy? *Vaccine*. 2015 Aug 14;33(34):4215–7.

19. Dubé E, Gagnon D, MacDonald NE. Strategies intended to address vaccine hesitancy: Review of published reviews. *Vaccine*. 2015 Aug 14;33(34):4191–203.
20. MacDonald NE. Vaccine hesitancy: Definition, scope and determinants. *Vaccine*. 2015 Aug 14;33(34):4161–4.
21. Larson HJ, Jarrett C, Schulz WS, Chaudhuri M, Zhou Y, Dube E, et al. Measuring vaccine hesitancy: The development of a survey tool. *Vaccine*. 2015 Aug 14;33(34):4165–75.
22. Strategic Advisory Group of Experts on Immunization. Report of the Sage Working Group on Vaccine Hesitancy. 2014 [cited 2020 Jan 9]. Available from: https://www.who.int/immunization/sage/meetings/2014/october/1_Report_WORKING_GROUP_vaccine_hesitancy_final.pdf
23. Gangarosa EJ, Galazka AM, Wolfe CR, Phillips LM, Gangarosa RE, Miller E, et al. Impact of anti-vaccine movements on pertussis control: the untold story. *Lancet (London, England)*. 1998;351(9099):356–61.
24. Hickler B, MacDonald NE, Senouci K, Schuh HB. Efforts to monitor Global progress on individual and community demand for immunization: Development of definitions and indicators for the Global Vaccine Action Plan Strategic Objective 2. *Vaccine*. 2017 Jun 16;35(28):3515–9.
25. Organización Mundial de la Salud. Plan de acción mundial sobre vacunas 2011–2020. 2013 [cited 2020 Jan 9]. Available from: https://apps.who.int/iris/bitstream/handle/10665/85398/9789243504988_spa.pdf
26. González Cárdenas DM, Reyes González ME, Fomental Hidalgo B, Gómez de Molina Iglesias M. La variolación y el advenimiento de la vacuna contra la viruela. *Revista Habanera de Ciencias Médicas*. 2007;6(2).
27. Tuells J. Historia de las vacunas. In: Álvarez Pasquín MJ, Velasco Muñoz C, editors. *Vacuna a vacuna: Manual de información sobre vacunas online*. 3ª Edición. Zaragoza: Amazing Books; 2018. p. 27–47.
28. Poland GA, Jacobson RM. The Age-Old Struggle against the Antivaccinationists. *N Engl J Med*. 2011;364(2):97–9.
29. Wolfe RM, Sharp LK. Anti-vaccinationists past and present. *BMJ*. 2002 Aug 24;325(7361):430–2.
30. Swales JD. The Leicester anti-vaccination movement. *Lancet (London, England)*. 1992 Oct 24;340(8826):1019–21.
31. Nelson MC, Rogers J. The Right to Die? Anti-vaccination Activity and the 1874 Smallpox Epidemic in Stockholm. *Soc Hist Med*. 1992;5(3):369–88.
32. Kulenkampff M, Schwartzman JS, Wilson J. Neurological complications of pertussis inoculation. Vol. 49, *Archives of Disease in Childhood*. 1974.
33. Gómez Marco JJ, Zamanillo Rojo I. Grupos anti-vacunas. Análisis de sus causas y consecuencias. Situación en España y resto de países. *Rev Clínica Electrónica en Atención Primaria*. 2005;8:1–6.
34. Wakefield AJ, Murch SH, Anthony A, Linnell J, Casson DM, Malik M, et al. Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *Lancet*. 1998;351(9103):637–41.
35. Godlee F, Smith J, Marcovitch H. Wakefield's article linking MMR vaccine and autism was fraudulent. *BMJ*. 2011 Jan 5;342:c7452.
36. Dyer C. *Lancet* retracts Wakefield's MMR paper. *BMJ*. 2010 Feb 2;340:c696.
37. Rao TSS, Andrade C. The MMR vaccine and autism: Sensation, refutation, retraction, and fraud. *Indian J Psychiatry*. 2011 Apr;53(2):95–6.
38. Peltola H, Patja A, Leinikki P, Valle M, Davidkin I, Paunio M. No evidence for measles, mumps, and rubella vaccine-associated inflammatory bowel disease or autism in a 14-year prospective study. *Lancet*. 1998 May 2;351(9112):1327–8.

39. DeStefano F, Thompson WW. MMR vaccine and autism: An update of the scientific evidence. Vol. 3, Expert Review of Vaccines. Taylor & Francis; 2004. p. 19–22.
40. Farrington CP, Miller E, Taylor B. MMR and autism: Further evidence against a causal association. *Vaccine*. 2001 Jun 14;19(27):3632–5.
41. Fombonne E, Chakrabarti S. No evidence for a new variant of measles-mumps-rubella-induced autism. *Pediatrics*. 2001 Oct 1;108(4):e58–e58.
42. Madsen KM, Hviid A, Vestergaard M, Schendel D, Wohlfahrt J, Thorsen P, et al. A Population-Based Study of Measles, Mumps, and Rubella Vaccination and Autism. *N Engl J Med*. 2002 Nov 7;347(19):1477–82.
43. Taylor B, Miller E, Farrington CP, Petropoulos MC, Favot-Mayaud I, Li J, et al. Autism and measles, mumps, and rubella vaccine: No epidemiological evidence for a causal association. *Lancet*. 1999 Jun 12;353(9169):2026–9.
44. Honda H, Shimizu Y, Rutter M. No effect of MMR withdrawal on the incidence of autism: a total population study. *J Child Psychol Psychiatry*. 2005 Jun 1;46(6):572–9.
45. Hussain A, Ali S, Ahmed M, Hussain S. The Anti-vaccination Movement: A Regression in Modern Medicine. *Cureus*. 2018 Jul 3;10(7).
46. Benecke O, DeYoung SE. Anti-Vaccine Decision-Making and Measles Resurgence in the United States. *Glob Pediatr Heal*. 2019;6:2333794X19862949.
47. Larson HJ, Ghinai I. Lessons from polio eradication. *Nature*. 2011 May 25;473(7348):446–7.
48. Limia A, Pachón I. Coverage of human papillomavirus vaccination during the first year of its introduction in Spain. *Euro Surveill*. 2011;16(21):pii=19873.
49. Trilla A. Vacunación sistemática: convencidos, indecisos y radicales Systematic vaccination: Convinced, hesitant and radicals. *Med Clin*. 2015;145(4):160–2.
50. Real Academia Española. Diccionario de la lengua española. 23 Ed. 2014.
51. Buchanan T, Benson V. Spreading Disinformation on Facebook: Do Trust in Message Source, Risk Propensity, or Personality Affect the Organic Reach of “Fake News”? *Soc Media + Soc*. 2019 Oct 17;5(4):205630511988865.
52. Rosselli R, Martini M, Bragazzi NL. The old and the new: vaccine hesitancy in the era of the Web 2.0. Challenges and opportunities. *J Prev Med Hyg*. 2016;57:E47-50.
53. Sofroniou A. World Wide Web. In: *Surfing The Internet, Then, Now, Later*. Lulu Com; 2014. p. 47–81.
54. Jošanov B, Pucihar A, Jošanov-Vrgović I, Maroščan Z. Opinions and Behavior of Students About Abuse of Internet in Social Involvements: Gender Analysis. *Bus Sch*. 2016;1:11–21.
55. Sbaifi L, Rowley J. Trust and Credibility in Web-Based Health Information: A Review and Agenda for Future Research. *J Med Internet Res*. 2017;19(6):e218.
56. Sun Y, Zhang Y, Gwizdzka J, Trace CB. Consumer Evaluation of the Quality of Online Health Information: Systematic Literature Review of Relevant Criteria and Indicators. *J Med Internet Res*. 2019;21(5):e12522.
57. Steffens MS, Dunn AG, Wiley KE, Leask J. How organisations promoting vaccination respond to misinformation on social media: a qualitative investigation. *BMC Public Health*. 2019 Dec 23;19(1):1348.
58. Jacobs W, Amuta AO, Jeon KC. Health information seeking in the digital age: An analysis of health information seeking behavior among US adults. Alvares C, editor. *Cogent Soc Sci*. 2017 Mar 13;3(1).
59. TNS Political & Social. European citizens' digital health literacy. A report to the European Commission. 2014.

60. Asociación Española de Vacunología. La Asociación Española de Vacunología (AEV) recuerda la importancia de consultar páginas de información fiables sobre vacunas [Internet]. 2017 [cited 2020 Feb 12]. Available from: <https://www.vacunas.org/la-asociacion-espanola-de-vacunologia-aev-recuerda-la-importancia-de-consultar-paginas-de-informacion-fiables-sobre-vacunas/>
61. Mohd Azizi FS, Kew Y, Moy FM. Vaccine hesitancy among parents in a multi-ethnic country, Malaysia. *Vaccine*. 2017 May 19;35(22):2955–61.
62. Broniatowski DA, Jamison AM, Qi SH, AlKulaib L, Chen T, Benton A, et al. Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate. *Am J Public Health*. 2018;108(10):1378–84.
63. Alber JM, Cohen C, Nguyen GT, Ghazvini SF, Tolentino BT. Exploring Communication Strategies for Promoting Hepatitis B Prevention among Young Asian American Adults. *J Health Commun*. 2018;23(12):977–83.
64. Stephens AB, Wynn CS, Stockwell MS. Understanding the use of digital technology to promote human papillomavirus vaccination—A RE-AIM framework approach. *Hum Vaccines Immunother*. 2019;15(7–8):1549–61.
65. Brown-Johnson CG, Boeckman LM, White AH, Burbank AD, Paulson S, Beebe LA. Trust in Health Information Sources: Survey Analysis of Variation by Sociodemographic and Tobacco Use Status in Oklahoma. *JMIR public Heal Surveill*. 2018 Feb 12;4(1):e8.
66. Loeb S, Sengupta S, Butaney M, Macaluso JN, Czarniecki SW, Robbins R, et al. Dissemination of Misinformative and Biased Information about Prostate Cancer on YouTube. *Eur Urol*. 2019 Apr 1;75(4):564–7.
67. Oxford Internet Institute. *The Use of Social Media for Research and Analysis: A Feasibility Study*. 2014.
68. Boyd DM, Ellison NB. Social Network Sites: Definition, History, and Scholarship. *J Comput Commun*. 2007;13(1):210–30.
69. Karduni A. Human-Misinformation interaction: Understanding the interdisciplinary approach needed to computationally combat false information. *arXiv Prepr*. 2019;
70. Yang K-C, Varol O, Davis CA, Ferrara E, Flammini A, Menczer F. Arming the public with artificial intelligence to counter social bots. *Hum Behav Emerg Technol*. 2019;1(1):48–61.
71. Zhao Z, Resnick P, Mei Q. Enquiring Minds: Early Detection of Rumors in Social Media from Enquiry Posts. In: *Proceedings of the 24th International Conference Companion on World Wide Web*. 2015. p. 1395–1405.
72. Gupta A, Kumaraguru P, Castillo C. TweetCred : Real-Time Credibility Assessment. In: *International Conference on Social Informatics*. 2014. p. 2012–3.
73. Finn S, Metaxas P, Mustafaraj E. Investigating Rumor Propagation with TwitterTrails. *CComputation Journal* . 2014;
74. Shao C, Ciampaglia GL, Flammini A, Menczer F. Hoaxy: A Platform for Tracking Online Misinformation. In: *Association for Computing Machinery, editor. Proceedings of the 25th International Conference Companion on World Wide Web*. Montreal; 2016. p. 745–750.
75. Lazer DMJ, Baum MA, Benkler Y, Berinsky AJ, Greenhill KM, Menczer F, et al. The science of fake news: Addressing fake news requires a multidisciplinary effort. *Science (80-)*. 2018;359(6380):1094–6.
76. Wonodi CB, Privor-Dumm L, Aina M, Pate AM, Reis R, Gadhoke P, et al. Using social network analysis to examine the decision-making process on new vaccine introduction in Nigeria. *Health Policy Plan*. 2012;27:ii27–38.
77. Himelboim I, Xiao X, Lee DKL, Wang MY, Borah P. A Social Networks Approach to Understanding Vaccine Conversations on Twitter: Network Clusters, Sentiment, and Certainty in HPV Social Networks. *Health Commun*. 2019;00(00):1–9.

78. Tao D, Yang P, Feng H. Utilization of text mining as a big data analysis tool for food science and nutrition. *Compr Rev Food Sci Food Saf*. 2020 Feb 16;1541-4337.12540.
79. Cortez Vásquez A, Vega Huerta H, Pariona Quispe J. Procesamiento de lenguaje natural. *Rev Ing Sist e Informática*. 2009;6(2).
80. Dale R. Classical Approaches to Natural Language Processing. In: Indurkha N, Damerau FJ, editors. *Handbook of Natural Language Processing*. 2nd Ed. Boca Raton: Chapman & Hall/CRC; 2010. p. 3–8.
81. Kalyanathaya KP, Akila D, Rajesh P. Advances in natural language processing – A survey of current research trends, development tools and industry applications. *Int J Recent Technol Eng*. 2019;7(5):199–201.
82. Costa-Jussà MR, Fonollosa JAR. Deep learning technologies for speech and audio processing [DeepVoice: Tecnologías de Aprendizaje Profundo aplicadas al Procesado de Voz y Audio]. *Proces Leng Nat*. 2017;59:117–20.
83. Etchegoyhen T, Garcia EM, Azpeitia A, Alegria I, Labaka G, Otegi A, et al. QUALES: Machine translation quality estimation via supervised and unsupervised machine learning. *Proces Leng Nat*. 2018;61:143–6.
84. Bhirud NS, Bhavsar RP, Pawar B V. Grammar checkers for natural languages: A review. *Int J Nat Lang Comput*. 2017;6(4).
85. Villena F, Dunstan J, Villena F, Dunstan J. Obtención automática de palabras clave en textos clínicos: una aplicación de procesamiento del lenguaje natural a datos masivos de sospecha diagnóstica en Chile. *Rev Med Chil*. 2019 Oct;147(10):1229–38.
86. Valle C, Simón-cuevas A, Valladares-valdés E, Olivas A, Romero FP, Tecnológica U, et al. Generación de resúmenes extractivos de múltiples documentos usando grafos semánticos. *Proces del Leng Nat*. 2019;63:103–10.
87. Acosta O, Aguilar C, Araya F. QuarryMeaning: A topic model application focused on Spanish documents. *Proces Leng Nat*. 2018;61:197–200.
88. Apolinario Ó, Medina-Moreira J, Lagos-Ortiz K, Luna-Aveiga H, García-Díaz JA, Valencia-García R. Tecnologías inteligentes para la autogestión de la salud. *Proces Leng Nat*. 2018;61:159–62.
89. Yadollahi A, Shahraiki AG, Zaiane OR. Current state of text sentiment analysis from opinion to emotion mining. *ACM Comput Surv*. 2017;50(2).
90. Henríquez Miranda C, Pla Santamaría F, Hurtado Oliver LF, Guzmán J. Análisis de sentimientos a nivel de aspecto usando ontologías y aprendizaje automático. *Proces del Leng Nat*. 2017;59:49–56.
91. Cambria E. Affective Computing and Sentiment Analysis. *IEEE Intell Syst*. 2016 Mar;31(2):102–7.
92. Colón-Ruiz C, Segura-Bedmar I, Martínez, Paloma. Análisis de Sentimiento en el dominio salud: Analizando comentarios sobre fármacos. *Proces del Leng Nat*. 2019;63:15–22.
93. Mäntylä M V., Graziotin D, Kuutila M. The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. *Comput Sci Rev*. 2018 Feb 1;27:16–32.
94. Behdenna S, Barigou F, Belalem G. Document Level Sentiment Analysis: A survey. *EAI Endorsed Trans Context Syst Appl*. 2018 Mar 14;4(13):154339.
95. Chauhan P, Singh A. Sentiment Analysis: A Comparative Study of Supervised Machine Learning Algorithms Using Rapid miner. *Int J Res Appl Sci Eng Technol*. 2017;V(XI):80–9.
96. Almatarneh S, Gamallo P. A lexicon based method to search for extreme opinions. *PLoS One*. 2018;13(5):1–19.
97. Kamble SS, Itkikar PAR. Study of supervised machine learning approaches for sentiment analysis. *Int Res J Eng Technol*. 2018;3045–7.

98. Medhat W, Hassan A, Korashy H. Sentiment analysis algorithms and applications: A survey. *Ain Shams Eng J.* 2014 Dec 1;5(4):1093–113.
99. Wiemken TL, Kelley RR. Machine Learning in Epidemiology and Health Outcomes Research. *Annu Rev Public Health.* 2020;41(1):1–16.
100. Souza VMA, Rossi RG, Batista G, Rezende SO. Unsupervised active learning techniques for labeling training sets: An experimental evaluation on sequential data. *Intell Data Anal.* 2017;21(5):1061–95.
101. Zhao Y, Ball R, Mosesian J, De Palma JF, Lehman B. Graph-based semi-supervised learning for fault detection and classification in solar photovoltaic arrays. *IEEE Trans Power Electron.* 2015 May 1;30(5):2848–58.
102. Samal BR, Behera AK, Panda M. Performance analysis of supervised machine learning techniques for sentiment analysis. *Proc 2017 3rd IEEE Int Conf Sensing, Signal Process Secur ICSSS 2017.* 2017;(October):128–33.
103. Hernández M, Gómez J. Aplicaciones de Procesamiento de Lenguaje Natural. *Rev Politécnica.* 2013;32(1):87–96.
104. Nguyen D, Nguyen C, Duong-Ba T, Nguyen H, Nguyen A, Tran T. Joint network coding and machine learning for error-prone wireless broadcast. In: 2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC). 2017.
105. Uddin S, Khan A, Hossain ME, Moni MA. Comparing different supervised machine learning algorithms for disease prediction. *BMC Med Inform Decis Mak.* 2019 Dec 21;19:281.
106. Syachrani S, Jeong HS “David”, Chung CS. Decision Tree–Based Deterioration Model for Buried Wastewater Pipelines. *J Perform Constr Facil.* 2013 Oct 1;27(5):633–45.
107. Fürnkranz J, Kliegr T. A Brief Overview of Rule Learning. In: Bassiliades N, G G, Sadri F, Paschke A, Roman D, editors. *Rule Technologies: Foundations, Tools, and Applications RuleML 2015 Lecture Notes in Computer Science.* Cham: Springer; 2015. p. 54–69.
108. Wang H, Zheng B, Yoon SW, Ko HS. A support vector machine-based ensemble algorithm for breast cancer diagnosis. *Eur J Oper Res.* 2018;267(2):687–99.
109. Al-saffar RZ, Khattab SI, Yousif ST. Prediction of Soil’s Compaction Parameter Using Artificial Neural Network. *AI-Rafidain Eng.* 2013;21(3):15–27.
110. Mehta P, Pandya S. A Review On Sentiment Analysis Methodologies, Practices And Applications. *Int J Sci Technol Res.* 2020;9:2.
111. Chaudhari M, Govilkar S. A Survey of Machine Learning Techniques for Sentiment Classification. *Int J Comput Sci Appl.* 2015;5(3):13–23.
112. Obulesu O, Mahendra M, Thrilokreddy M. Machine Learning Techniques and Tools: A Survey. 2018 *Int Conf Inven Res Comput Appl.* 2018;605–11.
113. Gerniers A. Maximum entropy method for multi-label classification. *École polytechnique de Louvain;* 2018.
114. Tsatsaronis G, Macari N, Torge S, Dietze H, Schroeder M. A Maximum-Entropy approach for accurate document annotation in the biomedical domain. *J Biomed Semantics.* 2012;3(1):S2.
115. Qin S, Li Q, Tang C, Tu Y. Optimal compressed sensing strategies for an array of nonlinear olfactory receptor neurons with and without spontaneous activity. *Proc Natl Acad Sci U S A.* 2019 Oct 8;116(41):20286–95.
116. Jaakkola T, Meila M, Jebara T. Maximum entropy discrimination. In: Solla S, Leen T, Müller K, editors. *Advances in Neural Information Processing Systems 12.* MIT press; 2000. p. 470–6.
117. Osama MR, Ahmad KH, Dana AAQ. Sentiment analysis as a way of web optimization. *Sci Res Essays.* 2016 Apr 30;11(8):90–6.

118. Li P, Li J, Huang Z, Gao C-Z, Chen W-B, Chen K. Privacy-preserving outsourced classification in cloud computing. *Cluster Comput.* 2018 Mar 8;21(1):277–86.
119. Sesen MB, Nicholson AE, Banares-Alcantara R, Kadir T, Brady M. Bayesian Networks for Clinical Decision Support in Lung Cancer Care. *PLoS One.* 2013 Dec 6;8(12):e82349.
120. Kolchyna O, Souza TTP, Treleaven P, Aste T. Twitter Sentiment Analysis: Lexicon Method, Machine Learning Method and Their Combination. In: Mitra G, Yu X, editors. *Handbook of Sentiment Analysis in Finance.* 1st Ed. Albury Books; 2016.
121. Vayansky I, Kumar SAP. A review of topic modeling methods. *Inf Syst.* 2020 Dec 1;94:101582.
122. Jelodar H, Wang Y, Yuan C, Feng X, Jiang X, Li Y, et al. Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimed Tools Appl.* 2019;78(11):15169–211.
123. Liu L, Tang L, Dong W, Yao S, Zhou W. An overview of topic modeling and its current applications in bioinformatics. Vol. 5, SpringerPlus. SpringerOpen; 2016. p. 1608.
124. Kherwa P, Bansal P. Topic Modeling: A Comprehensive Review. *ICST Trans Scalable Inf Syst.* 2019;7(24):e2.
125. Smith MC, Dredze M, Quinn SC, Broniatowski DA. Monitoring Real-time Spatial Public Health Discussions in the Context of Vaccine Hesitancy. In: *SMM4H@AMIA.* 2017.
126. Van Asch V. Macro-and micro-averaged evaluation measures [[basic draft]]. *Comput Linguist Psycholinguist.* 2013 Sep 9;1–27.
127. Pagoto S, Waring ME, Xu R. A Call for a Public Health Agenda for Social Media Research. *J Med Internet Res.* 2019;21(12):1–12.
128. Sinnenberg L, Buttenheim AM, Padrez K, Mancheno C, Ungar L, Merchant RM. Twitter as a Tool for Health Research: A Systematic Review. *Am J Public Health.* 2017;107(1):e1–8.
129. Donelle L, Booth RG. Health tweets: An exploration of health promotion on twitter. *Online J Issues Nurs.* 2012;17(3):2011.
130. Gomes C, Coustasse A. Tweeting and Treating: How Hospitals Use Twitter to Improve Care. *Health Care Manag (Frederick).* 2015;34(3):203–4.
131. Clement J. Most popular social networks worldwide as of October 2020, ranked by number of active users [Internet]. 2020 [cited 2021 Jan 6]. Available from: <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>
132. Kagashe I, Yan Z, Suheryani I. Enhancing Seasonal Influenza Surveillance: Topic Analysis of Widely Used Medicinal Drugs Using Twitter Data. *J Med Internet Res.* 2017;19(9):e315.
133. Zhang AJ, Albrecht L, Scott SD. Using Twitter for Data Collection With Health-Care Consumers. *Int J Qual Methods.* 2018 Dec 17;17(1):160940691775078.
134. Talbot C V., O'Dwyer ST, Clare L, Heaton J, Anderson J. How people with dementia use twitter: A qualitative analysis. *Comput Human Behav.* 2020 Jan 1;102:112–9.
135. Bhattacharya S, Srinivasan P, Polgreen P. Engagement with Health Agencies on Twitter. Sullivan PS, editor. *PLoS One.* 2014 Nov 7;9(11):e112235.
136. Katz MS, Anderson PF, Thompson MA, Salmi L, Freeman-Daily J, Utengen A, et al. Organizing Online Health Content: Developing Hashtag Collections for Healthier Internet-Based People and Communities. *JCO Clin cancer informatics.* 2019;3(3):1–10.
137. Burki T. Vaccine misinformation and social media. *Lancet Digit Heal.* 2019;1(6):e258–9.
138. Kloth YM, Deutsch KM, Danielson KA, Strack J, Law C. What Twitter teaches us about patient-provider communication on pain. *PLoS One.* 2019;14(12):e0226321.

139. Zhang Z, Ahmed W. A comparison of information sharing behaviours across 379 health conditions on Twitter. *Int J Public Health*. 2019 Apr 26;64(3):431–40.
140. Litt DM, Lewis MA, Spiro ES, Aulck L, Waldron KA, Head-Corliss MK, et al. #drunktwitter: Examining the relations between alcohol-related Twitter content and alcohol willingness and use among underage young adults. *Drug Alcohol Depend*. 2018;193:75–82.
141. Cavazos-Rehg P, Krauss M, Fisher SL, Salyer P, Grucza RA, Jean Bierut L. Twitter Chatter about Marijuana NIH Public Access. *J Adolesc Heal*. 2015;56(2):139–45.
142. Dakkak H, Brown R, Twynstra J, Charbonneau K, Seabrook JA. The perception of pre- and post-natal marijuana exposure on health outcomes: A content analysis of Twitter messages. *J Neonatal Perinatal Med*. 2018;11(4):409–15.
143. Lazard AJ, Wilcox GB, Tuttle HM, Glowacki EM, Pikowski J. Public reactions to e-cigarette regulations on Twitter: a text mining analysis. *Tob Control*. 2017;26(e2):e112–6.
144. Onezi H Al, Khalifa M, El-Metwally A, Househ M. The impact of social media-based support groups on smoking relapse prevention in Saudi Arabia. *Comput Methods Programs Biomed*. 2018;159:135–43.
145. Zhou J, Liu F, Zhou H. Understanding health food messages on Twitter for health literacy promotion. *Perspect Public Health*. 2018 May 7;138(3):173–9.
146. Liu Y, Mei Q, Hanauer DA, Zheng K, Lee JM. Use of Social Media in the Diabetes Community: An Exploratory Analysis of Diabetes-Related Tweets. *JMIR diabetes*. 2016 Nov 7;1(2):e4.
147. Kendra RL, Karki S, Eickholt JL, Gandy L. Characterizing the Discussion of Antibiotics in the Twittersphere: What is the Bigger Picture? *J Med Internet Res*. 2015;17(6):e154.
148. Hoang T, Liu J, Pratt N, Zheng VW, Chang KC, Roughead E, et al. Authenticity and credibility aware detection of adverse drug events from social media. *Int J Med Inform*. 2018 Dec 1;120:157–71.
149. Hu H, Phan N, Geller J, Iezzi S, Vo H, Dou D, et al. An Ensemble Deep Learning Model for Drug Abuse Detection in Sparse Twitter-Sphere. *Stud Health Technol Inform*. 2019 Aug 21;264:163–7.
150. Hswen Y, Naslund JA, Brownstein JS, Hawkins JB. Online Communication about Depression and Anxiety among Twitter Users with Schizophrenia: Preliminary Findings to Inform a Digital Phenotype Using Social Media. *Psychiatr Q*. 2018;89(3):569–80.
151. Robinson P, Turk D, Jilka S, Cella M. Measuring attitudes towards mental health using social media: investigating stigma and trivialisation. *Soc Psychiatry Psychiatr Epidemiol*. 2019;54(1):51–8.
152. Wang T, Brede M, Ianni A, Mentzakis E. Social interactions in online eating disorder communities: A network perspective. Tang M, editor. *PLoS One*. 2018 Jul 30;13(7):e0200800.
153. Lee SY, Kwon Y. Twitter as a place where people meet to make suicide pacts. *Public Health*. 2018;159:21–6.
154. Reyes-Menendez A, Saura JR, Alvarez-Alonso C. Understanding #WorldEnvironmentDay User Opinions in Twitter: A Topic-Based Sentiment Analysis Approach. *Int J Environ Res Public Health*. 2018;15(11).
155. Lenoir P, Moulahi B, Azé J, Bringay S, Mercier G, Carbonnel F. Raising Awareness About Cervical Cancer Using Twitter: Content Analysis of the 2015 #SmearForSmear Campaign. *J Med Internet Res*. 2017;19(10):e344.
156. Teoh D, Shaikh R, Vogel RI, Zoellner T, Carson L, Kulasingam S, et al. A Cross-Sectional Review of Cervical Cancer Messages on Twitter During Cervical Cancer Awareness Month. *J Low Genit Tract Dis*. 2018;22(1):8–12.

157. Fung IC-H, Jackson AM, Ahweyevu JO, Grizzle JH, Yin J, Ho Tse ZT, et al. #Globalhealth Twitter Conversations on #Malaria, #HIV, #TB, #NCDS, and #NTDS: a Cross-Sectional Analysis. *Ann Glob Heal*. 2017;83(3–4):682–90.
158. Mahroum N, Adawi M, Sharif K, Waknin R, Mahagna H, Bisharat B, et al. Public reaction to Chikungunya outbreaks in Italy—Insights from an extensive novel data streams-based structural equation modeling analysis. *PLoS One*. 2018;13(5):e0197337.
159. van Lent LG, Sungur H, Kunneman FA, van de Velde B, Das E. Too Far to Care? Measuring Public Attention and Fear for Ebola Using Twitter. *J Med Internet Res*. 2017;19(6):e193.
160. Gui X, Wang Y, Kou Y, Reynolds TL, Chen Y, Mei Q, et al. Understanding the Patterns of Health Information Dissemination on Social Media during the Zika Outbreak. *AMIA . Annu Symp proceedings AMIA Symp*. 2017;2017:820–9.
161. Stefanidis A, Vraga E, Lamprianidis G, Radzikowski J, Delamater PL, Jacobsen KH, et al. Zika in Twitter: Temporal Variations of Locations, Actors, and Concepts. *JMIR Public Heal Surveill*. 2017;3(2):e22.
162. Ahmed W. Public health implications of #ShoutYourAbortion. *Public Health*. 2018 Oct 1;163:35–41.
163. Robillard JM, Johnson TW, Hennessey C, Beattie BL, Illes J. Aging 2.0: Health Information about Dementia on Twitter. *PLoS One*. 2013;8(7):69861.
164. Parsons CF, Breckons M, Durham J. Twitter: a viable medium for daily pain diaries in chronic orofacial pain? *Br Dent J*. 2015 Jul 24;219(2):75–8.
165. Bartlett C, Wurtz R. Twitter and Public Health. *J Public Heal Manag Pract*. 2015;21(4):375–83.
166. Hart M, Stetten N, Islam S, Pizarro K. Twitter and Public Health (Part 2): Qualitative Analysis of How Individual Health Professionals Outside Organizations Use Microblogging to Promote and Disseminate Health-Related Information. *JMIR public Heal Surveill*. 2017 Oct 4;3(4):e54.
167. Wasilewski MB, Stinson JN, Webster F, Cameron JI. Using Twitter to recruit participants for health research: An example from a caregiving study. *Health Informatics J*. 2019 Dec 30;25(4):1485–97.
168. Bello-Orgaz G, Hernandez-Castro J, Camacho D. Detecting discussion communities on vaccination in twitter. *Futur Gener Comput Syst*. 2016;66(2017):125–36.
169. Cuesta-Cambra U, Martínez-Martínez L, Niño-González J-I. An analysis of pro-vaccine and anti-vaccine information on social networks and the internet: Visual and emotional patterns. *El Prof la Inf*. 2019 Mar 25;28(2):e280217.
170. Martin S, Kilich E, Dada S, Kummervold PE, Denny C, Paterson P, et al. “Vaccines for pregnant women...?! Absurd” – Mapping maternal vaccination discourse and stance on social media over six months. *Vaccine*. 2020;38(42):6627–37.
171. Mollema L, Harmsen IA, Broekhuizen E, Clijnk R, De Melker H, Paulussen T, et al. Disease detection or public opinion reflection? content analysis of tweets, other social media, and online newspapers during the measles outbreak in the Netherlands in 2013. *J Med Internet Res*. 2015 May 1;17(5):e128.
172. Raghupathi V, Ren J, Raghupathi W. Studying public perception about vaccination: A sentiment analysis of tweets. *Int J Environ Res Public Health*. 2020;17(10):3464.
173. McNeill A, Harris PR, Briggs P. Twitter Influence on UK Vaccination and Antiviral Uptake during the 2009 H1N1 Pandemic. *Front Public Heal*. 2016 Feb 22;4:22.
174. Huang X, Smith MC, Paul MJ, Ryzhkov D, Quinn SC, Broniatowski DA, et al. Examining patterns of influenza vaccination in social media. In: *AAAI Workshops. North America*; 2017. p. 542–6.

175. Deiner MS, Fathy C, Kim J, Niemeyer K, Ramirez D, Ackley SF, et al. Facebook and Twitter vaccine sentiment in response to measles outbreaks. *Health Informatics J*. 2019 Sep 17;25(3):1116–32.
176. Du J, Tang L, Xiang Y, Zhi D, Xu J, Song H-Y, et al. Public Perception Analysis of Tweets During the 2015 Measles Outbreak: Comparative Study Using Convolutional Neural Network Models. *J Med Internet Res*. 2018;20(7):e236.
177. Luo X, Zimet G, Shah S. A natural language processing framework to analyse the opinions on HPV vaccination reflected in twitter over 10 years (2008 - 2017). *Hum Vaccin Immunother*. 2019 Aug 3;15(7–8):1496–504.
178. Bahk CY, Cumming M, Paushter L, Madoff LC, Thomson A, Brownstein JS. Publicly Available Online Tool Facilitates Real-Time Monitoring Of Vaccine Conversations And Sentiments. *Health Aff*. 2016 Feb 2;35(2):341–7.
179. Dunn AG, Leask J, Zhou X, Mandl KD, Coiera E. Associations Between Exposure to and Expression of Negative Opinions About Human Papillomavirus Vaccines on Social Media: An Observational Study. *J Med Internet Res*. 2015;17(6):e144.
180. Shapiro GK, Surian D, Dunn AG, Perry R, Kelaher M. Comparing human papillomavirus vaccine concerns on Twitter: a cross-sectional study of users in Australia, Canada and the UK. *BMJ Open*. 2017 Oct 5;7(10):e016869.
181. Massey PM, Leader A, Yom-Tov E, Budenz A, Fisher K, Klassen AC. Applying Multiple Data Collection Tools to Quantify Human Papillomavirus Vaccine Communication on Twitter. *J Med Internet Res*. 2016 Dec 5;18(12):e318.
182. Du J, Xu J, Song H-Y, Tao C. Leveraging machine learning-based approaches to assess human papillomavirus vaccination sentiment trends with Twitter data. *BMC Med Inform Decis Mak*. 2017;17(Suppl. 2):69.
183. Tomeny TS, Vargo CJ, El-Toukhy S. Geographic and demographic correlates of autism-related anti-vaccine beliefs on Twitter, 2009-15. *Soc Sci Med*. 2017;191:168–75.
184. Gunaratne K, Coomes EA, Haghbayan H. Temporal trends in anti-vaccine discourse on Twitter. *Vaccine*. 2019;37(35):4867–71.
185. Dredze M, Wood-Doughty Z, Quinn SC, Broniatowski DA. Vaccine opponents' use of Twitter during the 2016 US presidential election: Implications for practice and policy. *Vaccine*. 2017;35(36):4670–2.
186. D'Andrea E, Ducange P, Bechini A, Renda A, Marcelloni F. Monitoring the public opinion about the vaccination topic from tweets analysis. In: *Expert Systems with Applications*. Pergamon; 2019. p. 209–26.
187. Tavošchi L, Quattrone F, D'Andrea E, Ducange P, Vabanesi M, Marcelloni F, et al. Twitter as a sentinel tool to monitor public opinion on vaccination: an opinion mining analysis from September 2016 to August 2017 in Italy. *Hum Vaccines Immunother*. 2020;16(5):1062–9.
188. Schwartz JL. New media, old messages: Themes in the history of vaccine hesitancy and refusal. *Am Med Assoc J Ethics*. 2012;14(1):50–5.
189. Wang Y, McKee M, Torbica A, Stuckler D. Systematic Literature Review on the Spread of Health-related Misinformation on Social Media. *Soc Sci Med*. 2019;240(2019):112552.
190. Shah Z, Surian D, Dyda A, Coiera E, Mandl KD, Dunn AG. Automatically Appraising the Credibility of Vaccine-Related Web Pages Shared on Social Media: A Twitter Surveillance Study. *Journal Med Internet Res*. 2019;21(11):e14007.
191. Hutto C, Gilbert E. VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. In: *Eighth International AAAI Conference on Weblogs and Social Media (ICWSM-14)*. Ann Arbor, Michigan; 2014. p. 216–25.
192. Countries with the most Twitter users [Internet]. 2021 [cited 2021 Mar 18]. Available from: <https://www.statista.com/>

193. Larson HJ, de Figueiredo A, Xiaohong Z, Schulz WS, Verger P, Johnston IG, et al. The State of Vaccine Confidence 2016: Global Insights Through a 67-Country Survey. *EBioMedicine*. 2016;12:295–301.
194. Meyer SB, Violette R, Aggarwal R, Simeoni M, MacDougall H, Waite N. Vaccine hesitancy and Web 2.0: Exploring how attitudes and beliefs about influenza vaccination are exchanged in online threaded user comments. *Vaccine*. 2019;37(13):1769–74.
195. Massey PM, Budenz A, Leader A, Fisher K, Klassen AC, Yom-Tov E. What Drives Health Professionals to Tweet About #HPVvaccine? Identifying Strategies for Effective Communication. *Prev Chronic Dis*. 2018;15:E26.
196. Schmidt AL, Zollo F, Scala A, Betsch C, Quattrocioni W. Polarization of the vaccination debate on Facebook. *Vaccine*. 2018;36(25):3606–12.
197. Nyhan B, Reifler J, Richey S, Freed GL. Effective messages in vaccine promotion: A randomized trial. *Pediatrics*. 2014;133(4).
198. Nyhan B, Reifler J. Does correcting myths about the flu vaccine work? An experimental evaluation of the effects of corrective information. *Vaccine*. 2015;33(3):459–64.
199. Betsch C, Sachse K. Debunking Vaccination Myths: Strong Risk Negations Can Increase Perceive Vaccination Risks. *Heal Psychol*. 2013;32(2):146–155.
200. Cruz Piqueras M, Rodríguez García de Cortazar A, Hortal Carmona J, Padilla Bernáldez J. Reticencia vacunal: análisis del discurso de madres y padres con rechazo total o parcial a las vacunas. *Gac Sanit*. 2019;33(1):53–9.
201. Lane S, MacDonald NE, Marti M, Dumolard L. Vaccine hesitancy around the globe: Analysis of three years of WHO/UNICEF Joint Reporting Form data-2015–2017. *Vaccine*. 2018;36(26):3861–7.
202. Mellon J, Prosser C. Twitter and Facebook are not representative of the general population: Political attitudes and demographics of British social media users. *Res Polit*. 2017 Jul 13;4(3):205316801772000.
203. Gorrell G, Bontcheva K. Classifying Twitter favorites: Like, bookmark, or Thanks? *J Assoc Inf Sci Technol*. 2016;67(1):17–25.
204. Park CS, Kaye BK. Expanding Visibility on Twitter: Author and Message Characteristics and Retweeting. *Soc Media + Soc*. 2019;5(2):205630511983459.
205. Blankenship EB, Goff ME, Yin J, Tse ZTH, Fu K-W, Liang H, et al. Sentiment, Contents, and Retweets: A Study of Two Vaccine-Related Twitter Datasets. *Perm J*. 2018;22:17–138.
206. Becker BFH, Larson HJ, Bonhoeffer J, van Mulligen EM, Kors JA, Sturkenboom MCJM. Evaluation of a multinational, multilingual vaccine debate on Twitter. *Vaccine*. 2016;34(50):6166–71.
207. Yuan X, Schuchard RJ, Crooks AT. Examining Emergent Communities and Social Bots Within the Polarized Online Vaccination Debate in Twitter. *Soc Media + Soc*. 2019 Jul 4;5(3):205630511986546.
208. Mendes Á, Abreu L, Vilar-Correia MR, Borlido-Santos J. “That Should Be Left to Doctors, That’s What They are There For!”—Exploring the Reflexivity and Trust of Young Adults When Seeking Health Information. *Health Commun*. 2017 Sep 2;32(9):1076–81.
209. Jarrett C, Wilson R, O’Leary M, Eckersberger E, Larson HJ. Strategies for addressing vaccine hesitancy – A systematic review. *Vaccine*. 2015 Aug 14;33(34):4180–90.
210. Cairns G, MacDonald L, Angus K, Walker L, Cairns-Haylor T, Bowdler T. Systematic literature review of the evidence for effective national immunisation schedule promotional communications Insights into health communication Systematic literature review of the evidence for effective national immunisation schedule promotional communications. Stockholm; 2012.

211. Odone A, Ferrari A, Spagnoli F, Visciarelli S, Shefer A, Pasquarella C, et al. Effectiveness of interventions that apply new media to improve vaccine uptake and vaccine coverage A systematic review. *Hum Vaccin Immunother*. 2015;11(1):72–82.
212. Zipprich J, Winter K, Hacker J, Xia D, Watt J, Watt J. Measles Outbreak — California, December 2014–February 2015. Vol. 64, *MMWR Morb Mortal Wkly Rep*. 2015.
213. Mello MM, Studdert DM, Parmet WE. Shifting Vaccination Politics — The End of Personal-Belief Exemptions in California. *N Engl J Med*. 2015 Aug 27;373(9):785–7.
214. Larson PH, Figueiredo A De, Karafillakis E, Rawal M. State of vaccine confidence in the European Union in 2018 [Internet]. Luxembourg; 2018 [cited 2020 Jan 27]. Available from: https://ec.europa.eu/health/sites/health/files/vaccination/docs/2018_vaccine_confidence_en.pdf
215. Caplan AL. The Battle Over Compulsory Vaccination in the United States. *Am J Public Health*. 2018 Apr;108(4):424–5.
216. Riaño Galán I, Martínez González C, Sánchez Jacob M. Recomendaciones para la toma de decisiones ante la negativa de los padres a la vacunación de sus hijos: Análisis ético. *An Pediatr*. 2013;79(1):50.e1-50.e5.
217. Ministerio de Sanidad Servicios Sociales e Igualdad. Coberturas de vacunación. Datos estadísticos [Internet]. España. 2018 [cited 2019 Nov 13]. Available from: <https://www.msbs.gob.es/profesionales/saludPublica/prevPromocion/vacunaciones/coberturas.htm>
218. Comité de Bioética de España. Cuestiones ético-legales del rechazo a las vacunas y propuestas para un debate necesario [Internet]. 2016 [cited 2019 Nov 20]. Available from: <http://assets.comitedebioetica.es/files/documentacion/es/cuestiones-etico-legales-rechazo-vacunas-propuestas-debate-necesario.pdf>
219. Piñero Pérez R, Hernández Martín D, Carro Rodríguez MÁ, de la Parte Cancho M, Casado Verrier E, Galán Arévalo S, et al. Consulta de asesoramiento en vacunas: El encuentro es posible. *An Pediatría*. 2017 Jun 1;86(6):314–20.
220. Al Awaidy ST, Al Mayahi ZK, Kaddoura M, Mahomed O, Lahoud N, Abubakar A, et al. Influenza vaccination hesitancy among healthcare workers in south al batinah governorate in Oman: A cross-sectional study. *Vaccines*. 2020 Dec 1;8(4):1–14.
221. Owotomo O, Stritzel H, McCabe SE, Boyd CJ, Maslowsky J. Smoking intention and progression from e-cigarette use to cigarette smoking. *Pediatrics*. 2020;156(6):e2020007609.
222. Olmedo Lucerón C, Limia Sánchez A, Santamarina C. La confianza en la vacunación frente a la gripe en España: Discursos y actitudes reticentes en población general y profesionales sanitarios. *Rev Esp Salud Publica*. 2021;95:1–13.
223. Ghebre R, Berry-lawhorn JM, Souza GD. State of the Science: Screening , Surveillance , and Epidemiology of HPV-Related Malignancies. *Am Soc Clin Oncol Educ book Am Soc Clin Oncol Annu Meet*. 2021;41(41):1–12.
224. Zheng L, Wu J, Zheng M, First T, Hospital A, Diseases I, et al. Barriers to and Facilitators of Human Papillomavirus Vaccination Among People Aged 9 – 26 Years: A Systematic Review. *Sex Transm Dis*. 2021 Mar 18;Epub ahead of print.
225. Bruni L, Saura-Lázaro A, Montoliu A, Brotons M, Alemany L, Diallo MS, et al. HPV vaccination introduction worldwide and WHO and UNICEF estimates of national HPV immunization coverage 2010–2019. *Prev Med (Baltim)*. 2021 Mar 1;144:106399.
226. Lama Y, Quinn SC, Nan X, Cruz-Cano R. Social media use and human papillomavirus awareness and knowledge among adults with children in the household: examining the role of race, ethnicity, and gender. *Hum Vaccin Immunother*. 2021 Apr 3;17(4):1014–24.

227. Ortiz RR, Smith A, Coyne-Beasley T. A systematic literature review to examine the potential for social media to impact HPV vaccine uptake and awareness, knowledge, and attitudes about HPV and HPV vaccination. *Hum Vaccin Immunother.* 2019 Aug 3;15(7–8):1465–75.
228. Basch CH, MacLean SA. A content analysis of HPV related posts on instagram. *Hum Vaccines Immunother.* 2019 Aug 3;15(7–8):1476–8.
229. Ekram S, Debiec KE, Pumper MA, Moreno MA. Content and Commentary: HPV Vaccine and YouTube. *J Pediatr Adolesc Gynecol.* 2019 Apr 1;32(2):153–7.
230. Luisi MLR. From bad to worse II: Risk amplification of the HPV vaccine on Facebook. *Vaccine.* 2021 Jan 8;39(2):303–8.
231. Walker KK, Owens H, Zimet G. "We fear the unknown": Emergence, route and transfer of hesitancy and misinformation among HPV vaccine accepting mothers. *Prev Med Reports.* 2020;20:101240.
232. Margolis MA, Brewer NT, Shah PD, Calo WA, Gilkey MB. Stories about HPV vaccine in social media, traditional media, and conversations. *Prev Med (Baltim).* 2019;118(2019):251–6.
233. Sitaresmi MN, Rozanti NM, Simangunsong LB, Wahab A. Improvement of Parent's awareness, knowledge, perception, and acceptability of human papillomavirus vaccination after a structured-educational intervention. *BMC Public Health.* 2020 Dec 1;20(1):1836.
234. Al-Abri SS, Abdel-Hady DM, Al-Abaidani IS. Knowledge, attitudes, and practices regarding travel health among Muscat International Airport travelers in Oman: Identifying the gaps and addressing the challenges. *J Epidemiol Glob Health.* 2016 Jun 1;6(2):67–75.
235. Bechini A, Zanobini P, Zanella B, Ancillotti L, Moscadelli A, Bonanni P, et al. Travelers' Attitudes, Behaviors, and Practices on the Prevention of Infectious Diseases: A Study for Non-European Destinations. *Int J Environ Res Public Health.* 2021 Mar 18;18(6):3110.
236. Güdel DS, Berger CT. Update and practical advice for travel vaccinations and consultations in primary care. *Ther Umschau.* 2021 Mar 1;78(2):83–92.
237. Boubaker R, Meige P, Mialet C, Ngarambe Buffat C, Uwanyiligira M, Widmer F, et al. Travellers' profile, travel patterns and vaccine practices—a 10-year prospective study in a Swiss Travel Clinic. *J Travel Med.* 2016 Jan;23(1):tav017.
238. Weitzel T. Profile and complexity of travel medicine consultations in Chile: Unicentric cross-sectional study. *BMJ Open.* 2020;10(9):e037903.
239. Alves I, Teodósio R, Pereira F. If I told you that there is no need for yellow fever vaccine booster would you still come to the travel clinic?: a cross-sectional study. *Trop Dis Travel Med Vaccines.* 2021 Dec;7(1):7.
240. Tan L. Adult vaccination: Now is the time to realize an unfulfilled potential. *Hum Vaccin Immunother.* 2015;11(9):2158–66.
241. de Gomensoro E, Del Giudice G, Doherty TM. Challenges in adult vaccination. *Ann Med.* 2018 May;50(3):181–92.
242. Alzeer AA, Alfantoukh LA, Theneyan A, Bin Eid F, Almangour TA, Alshememry AK, et al. The influence of demographics on influenza vaccine awareness and hesitancy among adults visiting educational hospital in Saudi Arabia. *Saudi Pharm J.* 2021 Feb;29(2):188–93.
243. Strain WD, Mansi J, Boikos C, Boivin M, Fisher WA. Achieving Influenza Vaccine Uptake Target in Canada via a Pharmacy-Led Telephone Discussion during the 2019–2020 Season. *Vaccines.* 2021 Mar;9(4):312.
244. Trent MJ, Salmon DA, MacIntyre CR. Using the health belief model to identify barriers to seasonal influenza vaccination among Australian adults in 2019. *Influenza Other Respi Viruses.* 2021 Feb 15;Epub ahead of print.

245. Wei Z, Sun X, Yang Y, Zhan S, Fu C. Seasonal influenza vaccine hesitancy profiles and determinants among Chinese children's guardians and the elderly. *Expert Rev Vaccines*. 2021 Apr;Epub ahead of print.
246. Williams L, Deakin K, Gallant A, Rasmussen S, Young D, Cogan N. A mixed methods study of seasonal influenza vaccine hesitancy in adults with chronic respiratory conditions. *Influenza Other Respi Viruses*. 2021 Mar 27;Epub ahead of print.
247. Gatwood J, Shuvo S, Hohmeier KC, Hagemann T, Chiu CY, Tong R, et al. Pneumococcal vaccination in older adults: An initial analysis of social determinants of health and vaccine uptake. *Vaccine*. 2020 Jul 31;38(35):5607–17.
248. Gatwood J, Chiu CY, Shuvo S, Ramachandran S, Jadhav S, Hohmeier KC, et al. Role of social determinants of health in pneumococcal vaccination among high-risk adults. *Vaccine*. 2021 Apr 1;39(14):1951–62.
249. Gatwood J, McKnight M, Frederick K, Hohmeier K, Kapan S, Chiu CY, et al. Extent of and Reasons for Vaccine Hesitancy in Adults at High-Risk for Pneumococcal Disease. *Am J Heal Promot*. 2021 Mar 4;Epub ahead of print.
250. Shuvo S, Hagemann T, Hohmeier K, Chiu CY, Ramachandran S, Gatwood J. The role of social determinants in timely herpes zoster vaccination among older American adults. *Hum Vaccines Immunother*. 2021 Jan 30;Epub ahead of print.
251. Neufeind J, Betsch C, Habersaat KB, Eckardt M, Schmid P, Wichmann O. Barriers and drivers to adult vaccination among family physicians – Insights for tailoring the immunization program in Germany. *Vaccine*. 2020 Jun 2;38(27):4252–62.
252. Verelst F, Kessels R, Willem L, Beutels P. No Such Thing as a Free-Rider? Understanding Drivers of Childhood and Adult Vaccination through a Multicountry Discrete Choice Experiment. *Vaccines*. 2021 Mar 16;9(3):264.
253. Wong LP, Wong PF, Megat Hashim MMAA, Han L, Lin Y, Hu Z, et al. Multidimensional social and cultural norms influencing HPV vaccine hesitancy in Asia. *Hum Vaccines Immunother*. 2020 Jul 2;16(7):1611–22.
254. Robles C, Bruni L, Acera A, Riera JC, Prats L, Poljak M, et al. Determinants of Human Papillomavirus Vaccine Uptake by Adult Women Attending Cervical Cancer Screening in 9 European Countries. *Am J Prev Med*. 2021 Apr 1;60(4):478–87.
255. Barchitta M, Maugeri A, Magnano San Lio R, La Rosa M, La Mastra C, Favara G, et al. Vaccination Status of Mothers and Children from the 'Mamma & Bambino' Cohort. *Vaccines*. 2021 Feb 17;9(2):168.
256. Qaderi K, Mirmolaei ST, Geranmayeh M, Farnam F, Sheikh Hasani S. "Does HPV affect my fertility?" Reproductive concerns of HPV-positive women: a qualitative study. *Reprod Health*. 2021 Dec 1;18(1):72.
257. Regan AK, Tracey LE, Blyth CC, Richmond PC, Effler P V. A prospective cohort study assessing the reactivity of pertussis and influenza vaccines administered during pregnancy. *Vaccine*. 2016 Apr 29;34(20):2299–304.
258. Glover C, Crawford N, Leeb A, Wood N, Macartney K. Active SMS-based surveillance of adverse events following immunisation with influenza and pertussis-containing vaccines in Australian pregnant women using AusVaxSafety. *Vaccine*. 2020 Jun 26;38(31):4892–900.
259. Regan AK, Munoz FM. Efficacy and safety of influenza vaccination during pregnancy: realizing the potential of maternal influenza immunization. *Expert Rev Vaccines*. 2021 Apr 8;14760584.2021.1915138.
260. Lu QC, Zhang TY, Bundhun PK, Chen C. One "misunderstood" health issue: demonstrating and communicating the safety of influenza a vaccination in pregnancy: a systematic review and meta-analysis. *BMC Public Health*. 2021 Dec 9;21(1):703.
261. Krishnaswamy S, Cheng AC, Wallace EM, Buttery J, Giles ML. Understanding the barriers to uptake of antenatal vaccination by women from culturally and linguistically diverse backgrounds: A cross-sectional study. *Hum Vaccines Immunother*. 2018 Jul 3;14(7):1591–8.

262. SteelFisher GK, Caporello HL, Broussard CS, Schafer TJ, Ben-Porath EN, Blendon RJ. Seasonal Influenza Vaccine in Pregnant Women: Views and Experiences of Obstetrician–Gynecologists. *J Women’s Heal.* 2021 Feb 2;Epub ahead of print.
263. Laenen J, Roelants M, Devlieger R, Vandermeulen C. Influenza and pertussis vaccination coverage in pregnant women. *Vaccine.* 2015 Apr 27;33(18):2125–31.
264. Blanchard-Rohner G, Eberhardt C. Review of maternal immunisation during pregnancy: focus on pertussis and influenza. *Swiss Med Wkly.* 2017 Nov 9;147:w14526.
265. McHugh L, Andrews RM, Lambert SB, Viney KA, Wood N, Perrett KP, et al. Birth outcomes for Australian mother-infant pairs who received an influenza vaccine during pregnancy, 2012–2014: The FluMum study. *Vaccine.* 2017 Mar 7;35(10):1403–9.
266. CDC. Coverage by Season, FluVaxView, Seasonal Influenza (Flu) [Internet]. [cited 2021 Apr 10]. Available from: <https://www.cdc.gov/flu/fluview/coverage-by-season.htm>
267. Ford AJ, Alwan NA. Use of social networking sites and women’s decision to receive vaccinations during pregnancy: A cross-sectional study in the UK. *Vaccine.* 2018 Aug 23;36(35):5294–303.
268. Vaccinate Your Family. The Facts Behind the “CDC Whistleblower” Accusations Spotlighted in the film Vaxxed [Internet]. Washington; 2020 [cited 2021 Apr 5]. Available from: <https://vaccinateyourfamily.org>
269. Wilder-Smith AB, Qureshi K. Resurgence of measles in Europe: A systematic review on parental attitudes and beliefs of measles vaccine. *J Epidemiol Glob Health.* 2020 Mar 1;10(1):46–58.
270. Hviid A, Hansen JV, Frisch M, Melbye M. Measles, Mumps, Rubella Vaccination and Autism. *Ann Intern Med.* 2019 Apr 16;170(8):513.
271. Omer SB, Yildirim I. Further evidence of MMR vaccine safety: Scientific and communications considerations. Vol. 170, *Annals of Internal Medicine.* American College of Physicians; 2019. p. 567–8.
272. Eggertson L. Lancet retracts 12-year-old article linking autism to MMR vaccines. *CMAJ.* 2010 Mar 9;182(4):E199-200.
273. Ludvigsson JF, Winell H, Sandin S, Cnattingius S, Stephansson O, Pasternak B. Maternal Influenza A(H1N1) Immunization During Pregnancy and Risk for Autism Spectrum Disorder in Offspring: A Cohort Study. *Ann Intern Med.* 2020 Oct 20;173(8):597–604.
274. Zerbo O, Qian Y, Yoshida C, Fireman BH, Klein NP, Croen LA. Association between influenza infection and vaccination during pregnancy and risk of autism spectrum disorder. *JAMA Pediatr.* 2017;171(1):e163609.
275. Eizayaga JE, Waisse S. What do homeopathic doctors think of vaccines? An international online survey. *Homeopathy.* 2016;105(2):180–5.
276. Hornsey MJ, Lobera J, Díaz-Catalán C. Vaccine hesitancy is strongly associated with distrust of conventional medicine, and only weakly associated with trust in alternative medicine. *Soc Sci Med.* 2020;255(April):113019.
277. Qian M, Chou SY, Lai EK. Confirmatory bias in health decisions: Evidence from the MMR-autism controversy. *J Health Econ.* 2020 Mar 1;70:102284.
278. Dunn AG, Surian D, Leask J, Dey A, Mandl KD, Coiera E. Mapping information exposure on social media to explain differences in HPV vaccine coverage in the United States. *Vaccine.* 2017;35(23):3033–40.
279. Patel MK, Goodson JL, Alexander JP, Kretsinger K, Sodha S V., Steulet C, et al. Progress Toward Regional Measles Elimination — Worldwide, 2000–2019. *MMWR Morb Mortal Wkly Rep.* 2020;69(45):1700–5.

CAPÍTULO VII

ANEXO

INFORME DEL COMITÉ DE ÉTICA DE INVESTIGACIÓN

El comité Ético de Investigación en Humanos de la Comisión de Ética en Investigación Experimental de la Universitat de València,

CERTIFICA:

Que el Comité d'Ètica d'Investigació en Humans, en la reunió celebrada el dia , una vez estudiado el proyecto de tesis doctoral: "*Valoración de las vacunas desde una plataforma social con aplicaciones de uso en promoción de la salud.*", Cuyo/a responsable es D/Dña.

HILARY PIEDRAHITA VALDES, dirigida por D/Dña. M DEL CARMEN SAIZ SANCHEZ

ha acordado informar favorablemente el mismo.

Y para que conste, se firma el presente certificado



Av. Blasco Ibáñez, 13 tel: 963864109 vicerec.investigacio@uv.es
València 46010 fax: 963983221 www.uv.es/etivnes

Firmado digitalmente por
PEDRO JESUS PEREZ ZAFRILLA
Cargo: Presidente del Comité de Ética de la Investigación en Humanos
Fecha: 04/12/2020 12:56:49 CE



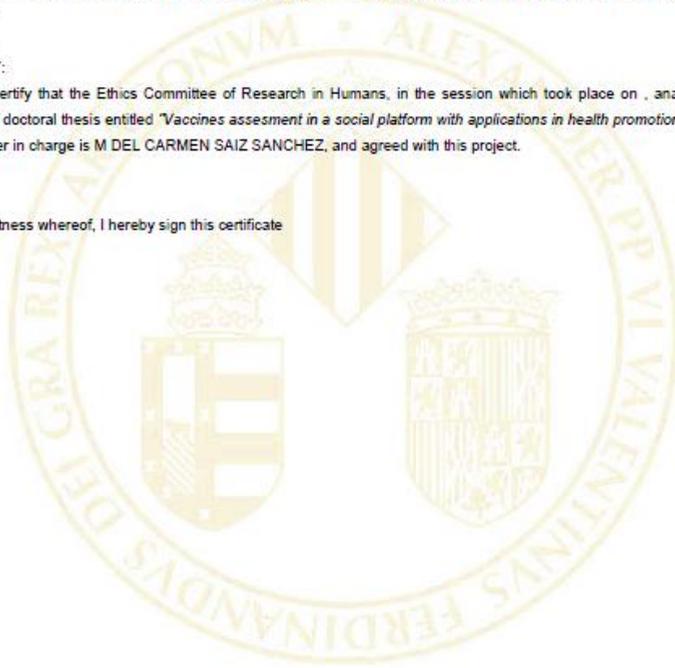
Valoración de las vacunas desde una plataforma social con aplicaciones de uso en promoción de la salud.

The Ethics Committee of Research in Humans of the Ethics Commission in Experimental Research of University of Valencia,

CERTIFY:

Hereby certify that the Ethics Committee of Research in Humans, in the session which took place on . analysed the project of doctoral thesis entitled "*Vaccines assesment in a social platform with applications in health promotion.*", whose researcher in charge is M DEL CARMEN SAIZ SANCHEZ, and agreed with this project.

And in witness whereof, I hereby sign this certificate



Av. Blasco Ibáñez, 13 tel: 963864109 vicerec.investigacio@uv.es
València 46010 fax: 963883221 www.uv.es/serinves

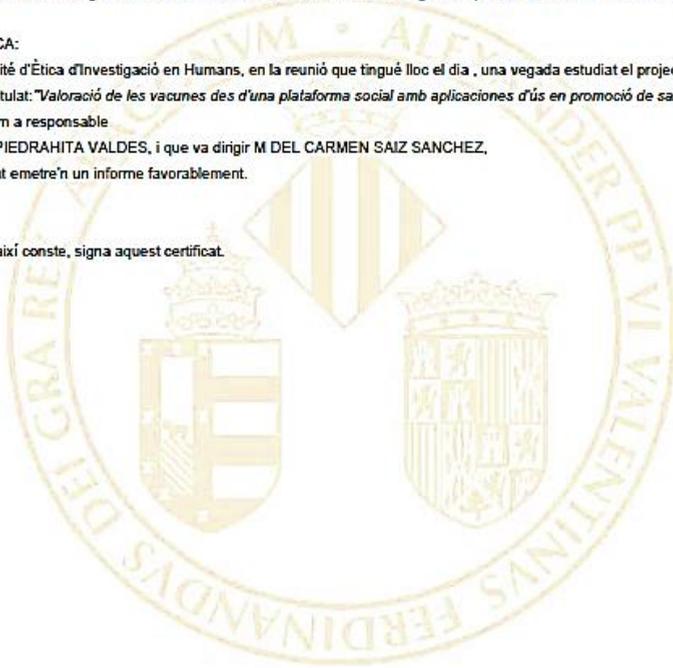
Comitè d'Ètica d'Investigació en Humans de la Comissió d'Ètica en Investigació Experimental de la Universitat de València,

CERTIFICA:

Que Comitè d'Ètica d'Investigació en Humans, en la reunió que tingué lloc el dia , una vegada estudiat el projecte de tesi doctoral titulat: "*Valoració de les vacunes des d'una plataforma social amb aplicacions d'ús en promoció de salut.*", que té com a responsable

HILARY PIEDRAHITA VALDES, i que va dirigir M DEL CARMEN SAIZ SANCHEZ, ha acordat emetre'n un informe favorablement.

I perquè així conste, signa aquest certificat.



Av. Blasco Ibáñez, 13 tel: 963864109 vicerec.investigacio@uv.es
València 46010 fax: 963983221 www.uv.es/serinves

