

CIENCIA DE DATOS EN ESPAÑA

DATASCIENCE SPAIN



Proyecto financiado por la Fundación Española para la Ciencia y la Tecnología (FECYT) en la Convocatoria de ayudas para el fomento de la cultura científica, tecnológica y de la innovación 2019-2020. [FCT-18-13437]



VNIVERSIDAD
D SALAMANCA



DATASCIENCE SPAIN Project



Proyecto financiado por la Fundación Española para la Ciencia y la Tecnología (FECYT) en la Convocatoria de ayudas para el fomento de la cultura científica, tecnológica y de la innovación 2019-2020. [FCT-18-13437]

OBSERVATORIO DE LOS CONTENIDOS AUDIOVISUALES UNIVERSIDAD DE SALAMANCA

Facultad de Ciencias Sociales. Campus Miguel de Unamuno. Edificio F.E.S. Paseo de Francisco Tomás y Valiente, s/n. 37007. Salamanca. España
www.ocal.es/datasciencespain

Enero 2021.

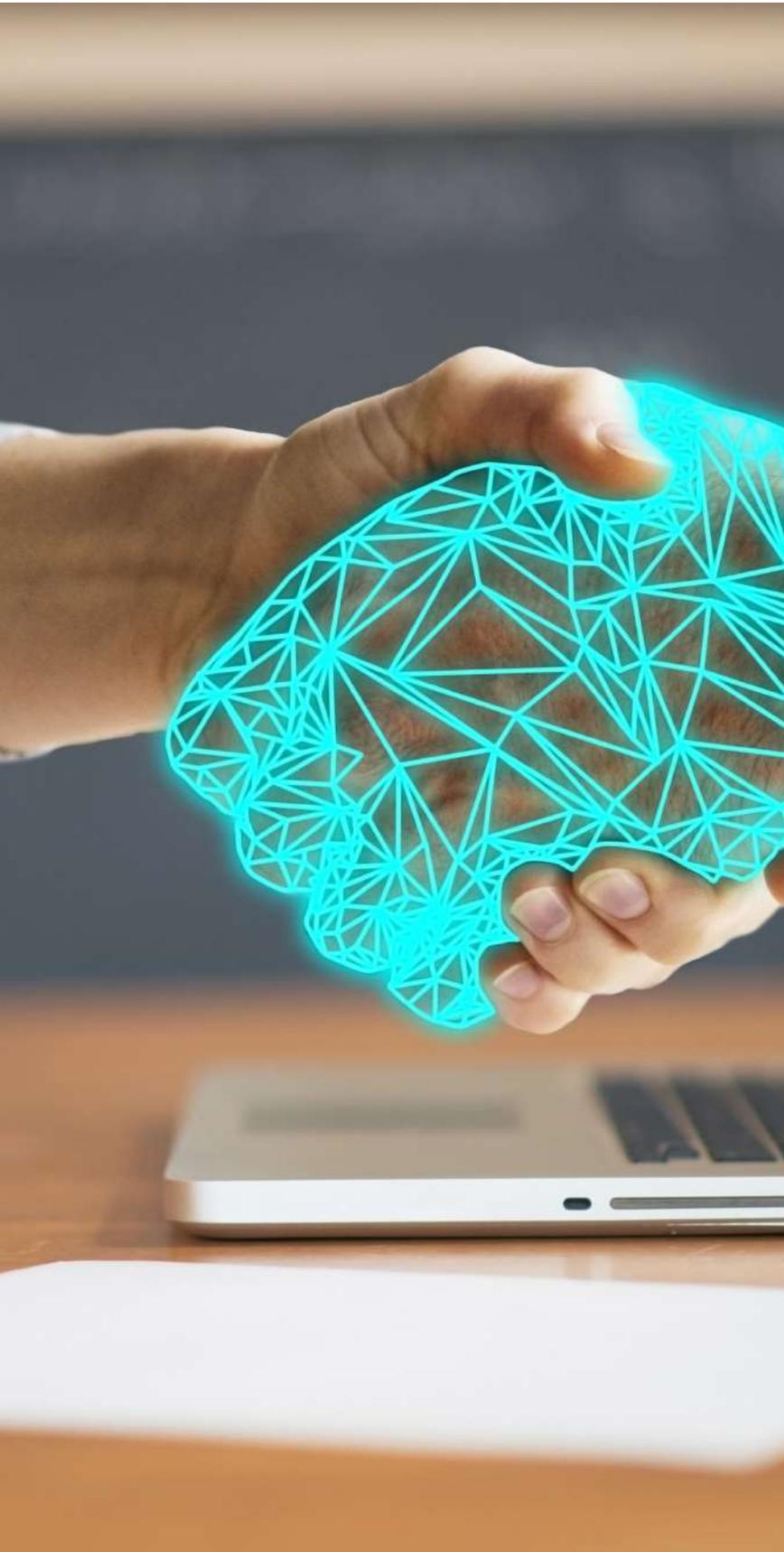
ISBN: 978-84-09-28167-1

Diseño base e imágenes de Canva.

This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. Esta licencia permite a otros entremezclar, ajustar y construir a partir de su obra con fines no comerciales, y aunque en sus nuevas creaciones deban reconocerle su autoría y no puedan ser utilizadas de manera comercial, no tienen que estar bajo una licencia con los mismos términos. <https://creativecommons.org/licenses/by-nc/4.0/legalcode.es>



CONTENIDO



01

CONOCIMIENTO
PÚBLICO Y ACTITUDES
HACIA LA
CIENCIA DE DATOS EN
ESPAÑA

02

CONOCIMIENTOS Y
PRÁCTICAS DE LOS
PERIODISTAS
PARA COMUNICAR
CIENCIA DE DATOS

03

COBERTURA DE LA
CIENCIA DE DATOS
POR LOS MEDIOS
DIGITALES
EN ESPAÑA

EQUIPO DE TRABAJO

Investigador principal:
Dr. Carlos Arcila Calderón
carcila@usal.es

Investigadores:
Dr. Juan José Igartua Perosanz
Dr. Félix Ortega Mohedano
Dra. Beatriz González de Garay Domínguez
Dr. Francisco Javier Frutos Esteban
Dra. María Marcos Ramos
Dra. Lifen Cheng Lee
Patricia Sánchez Holgado
Francisco Javier Jiménez Amores
David Blanco Herrero
Maximiliano Frías Vázquez

Colaboradores:
Juan Pablo Latorre Jiménez

CIENCIA DE DATOS EN ESPAÑA

CONOCIMIENTO PÚBLICO Y ACTITUDES HACIA LA CIENCIA DE DATOS EN ESPAÑA

DATASCIENCE SPAIN



Proyecto financiado por la Fundación Española para la Ciencia y la Tecnología (FECYT) en la Convocatoria de ayudas para el fomento de la cultura científica, tecnológica y de la innovación 2019-2020. [FCT-18-13437]



01

INTRODUCCIÓN

Existe un creciente interés de los ciudadanos en el impacto que tiene la Ciencia de Datos en su vida diaria, principalmente debido a la enorme atención que los medios prestan a temas relacionados con el Big Data o la Inteligencia Artificial. Este interés se traduce en grandes expectativas sobre las potenciales aplicaciones de estas tecnologías, pero también en serias preocupaciones, como el uso de datos personales, los procesos de toma de decisiones dirigidos por máquinas o la desaparición de empleos.

Junto a esas preocupaciones, también existe una falta de comprensión pública sobre unas tecnologías aún nuevas y cambiantes. La comunicación de la Ciencia de Datos, y particularmente del Big Data y la Inteligencia Artificial, se lleva a cabo en un momento en que los medios no son los únicos encargados de transmitir información sobre la ciencia, sino que otros actores entran en juego para comunicarse directamente con la sociedad a través de canales digitales, de manera que se hace más necesario garantizar la calidad de los procesos de comunicación científica.

01

INTRODUCCIÓN

El objetivo de este trabajo es encuestar el conocimiento, la comprensión y las actitudes del público hacia estas materias.

Hasta ahora, se han llevado a cabo encuestas en diferentes contextos para estudiar el nivel de conocimiento de la ciudadanía sobre ciencia. De hecho, las actitudes y la comprensión de la ciencia por la sociedad es un tema de estudio frecuente, pero los enfoques centrados en la Ciencia de Datos son todavía escasos, especialmente en el entorno español, donde el estudio más cercano es la Encuesta de Percepción Social de la Ciencia y la Tecnología, elaborada por la Fundación Española para la Ciencia y Tecnología (FECYT), cada dos años.

Para completar este vacío de conocimiento, se ha realizado una encuesta con preguntas cerradas a una muestra representativa de 684 ciudadanos españoles utilizando la plataforma *Qualtrics*.

01

DESAFÍOS

“¿Cuál es el grado de conocimiento de la Ciencia de Datos, especialmente el Big Data y la Inteligencia Artificial, entre los españoles?”

”

```
this.each(function() {  
    VERSION="3.3.7"  
    "href"),d=d&#amp;#amp;  
    },g=a.Event(  
    st("li"),c),t  
    }},c.prototype  
    ata-toggle=""  
    in")):b.remove  
    ",!0),e&&e()  
    e("bsTransit  
    ,a.fn.tab.no  
    [data-toggle=  
    n(){var d=a(  
    ons=a.extend  
    .affix.data-  
    ;c.VERSION=  
    lTop(),f=thi  
    npin<=f.top)  
    prototype.ge  
    ollTop(),b=t  
    .checkPositi
```

01

FACT SHEET

-  El grado de conocimiento de la Ciencia de Datos es moderado.
-  El interés de los Ciudadanos es mayor que su grado de familiarización.
-  Las actitudes que muestran los ciudadanos son generalmente positivas.
-  Los ciudadanos consideran que la Ciencia de Datos tiene una relación más directa con su conducta de consumidor y usuario.
-  La Inteligencia Artificial obtiene una valoración más positiva que el Big Data.
-  Las personas de más edad tienen menor conocimiento de estas tecnologías.
-  Las personas más jóvenes (de 18 a 24 años) muestran menor interés que las categorías de edad adulta.

METODOLOGÍA

El presente estudio se apoya en una encuesta con preguntas cerradas realizadas a una muestra representativa de la ciudadanía española compuesta por 684 personas.

El instrumento fue diseñado *ad hoc* para esta investigación, pero apoyado en los cuestionarios de Percepción Social de la Ciencia y la Tecnología (FECYT, 2018)(1), en información procedente del informe Future Trends Forum (2019) de la Fundación Bankinter (2), de los informes estratégicos de la Comisión Europa sobre Big Data e Inteligencia Artificial (European Commission, 2020)(3) y de la Iniciativa Interplataformas PlanetiC en España (Sáez & Costa-Soria, 2019)(4).

La muestra es n=684

Nivel de confianza 95%

Margen de error 5%

La distribución por edad es equilibrada y representativa.

(1) Fundación Española para la Ciencia y la Tecnología, FECYT (2018). IX Encuesta de Percepción Social de la Ciencia y la Tecnología 2018.

(2) Paniagua, E. (2019). Inteligencia Artificial. In Future Trends Forum (Vol. 52, Issue 55). Fundación Bankinter.

(3) European Commission. (2014). Special Eurobarometer 419: Public perceptions of science, research and innovation (Issue October). European Commission. <https://doi.org/10.2777/95599>

European Commission. (2017a). Special Eurobarometer 460: Attitudes towards the impact of digitisation and automation on daily life. <https://doi.org/10.2759/835661>

(4) Sáez, D., & Costa-Soria, C. (2019). Whitepaper: Análisis de la estrategia Big Data e Inteligencia Artificial en España.

01

MEDIDAS

Salvo las preguntas de género y edad (cinco tramos de edad: 18 a 24 años, entre 25 y 34 años, entre 35 y 49 años, entre 50 y 64 años y 65 o más años), las medidas utilizadas estuvieron formadas por escalas tipo *Likert* de cinco puntos. Así, las medidas del estudio fueron:

Familiarización con el concepto Big Data y con el de Inteligencia Artificial: entre 1 (nada familiarizado, no conozco nada del tema) y 5 (muy familiarizado, comprendo su uso y lo utilizo), mide hasta qué punto los encuestados conocen la materia y la utilizan.

Interés en el Big Data y en la Inteligencia Artificial, su conocimiento y/o su uso: entre 1 (nada interesado) y 5 (muy interesado), mide el grado de interés de los ciudadanos por la materia.

Beneficios del Big Data y la Inteligencia Artificial: entre 1 (ningún beneficio) y 5 (muchos beneficios), se midió el nivel de beneficios percibidos por los encuestados en una serie de aplicaciones destacadas.

Riesgos del Big Data y la Inteligencia Artificial: entre 1 (ningún riesgo) y 5 (muchos riesgos), se midió el nivel de riesgos percibidos por los encuestados en una serie de aplicaciones destacadas.

01

MEDIDAS

Las respuestas fueron anonimizadas, y posteriormente se aplicaron pruebas estadísticas descriptivas (análisis de frecuencias y de medias).

Igualmente, en el estudio sobre el conocimiento de estas tecnologías, se aplicaron pruebas de estadística inferencial: pruebas T de Student para muestras independientes para investigar si existen diferencias en función del género, pruebas ANOVA de un factor para investigar las posibles diferencias entre grupos de edad y correlaciones bivariadas para comprobar la relación entre las distintas medidas utilizadas para estudiar el conocimiento de estas tecnologías.

Todas estas pruebas se realizaron en el paquete SPSS de IBM, versión 25.

01

RESULTADOS

“Conocimiento y Actitudes de la ciudadanía española hacia la Ciencia de Datos, el big data y la inteligencia artificial.”

```
this.each(fun  
VERSION="3.3.7  
"href"),d=d&  
}),g=a.Event(  
st("li"),c),t  
}}},c.protot  
ata-toggle=""  
in")):b.remov  
",!0),e&&e()  
e("bsTransit  
,a.fn.tab.no  
[data-toggle=  
n(){var d=a(  
ons=a.extend  
.affix.data-  
;c.VERSION=  
lTop(),f=thi  
npin<=f.top)  
prototype.ge  
ollTop(),b=t  
.checkPositi
```

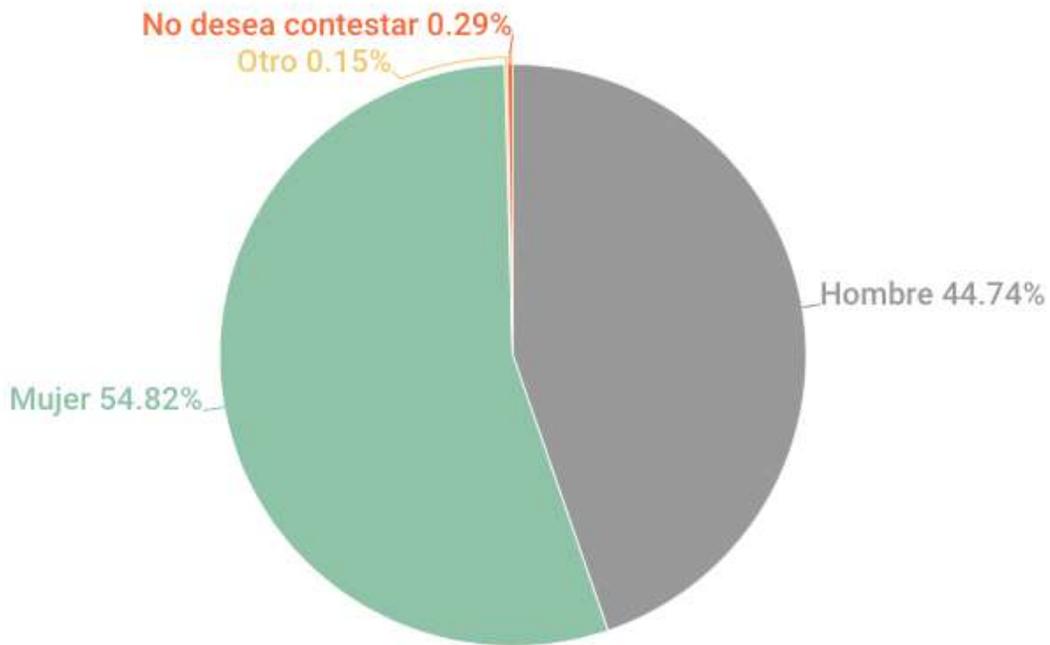
CONOCIMIENTO Y USO

ACTITUDES Y PERCEPCIONES

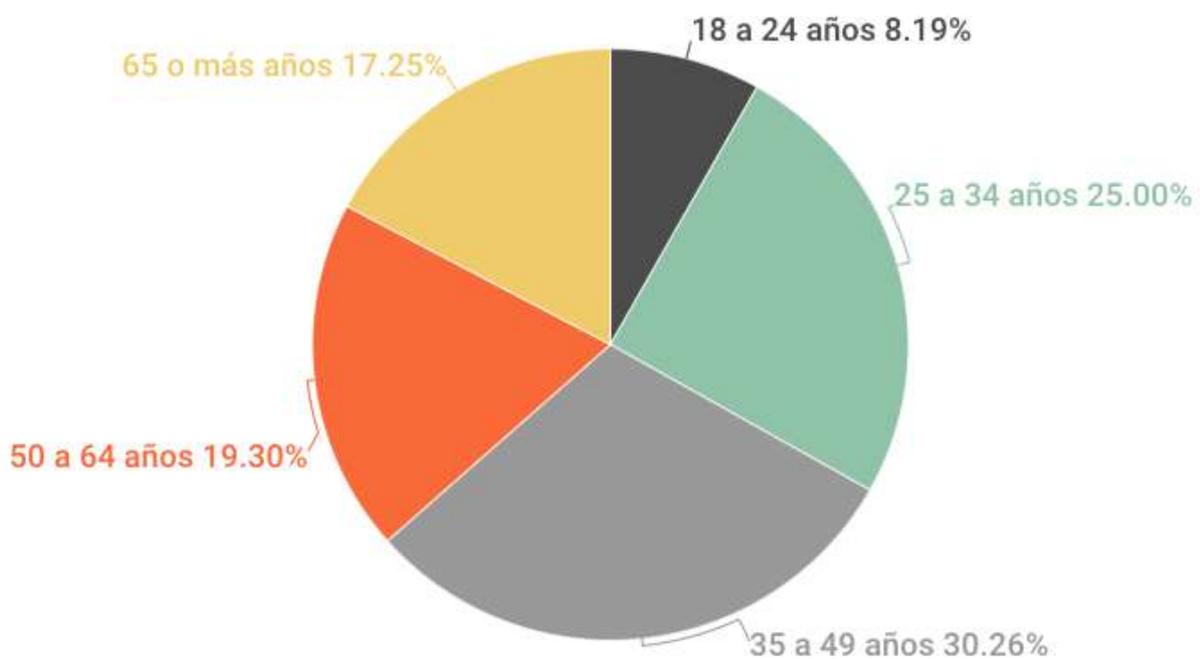
INFORMACIÓN

01

DISTRIBUCIÓN DE LA MUESTRA POR GÉNERO

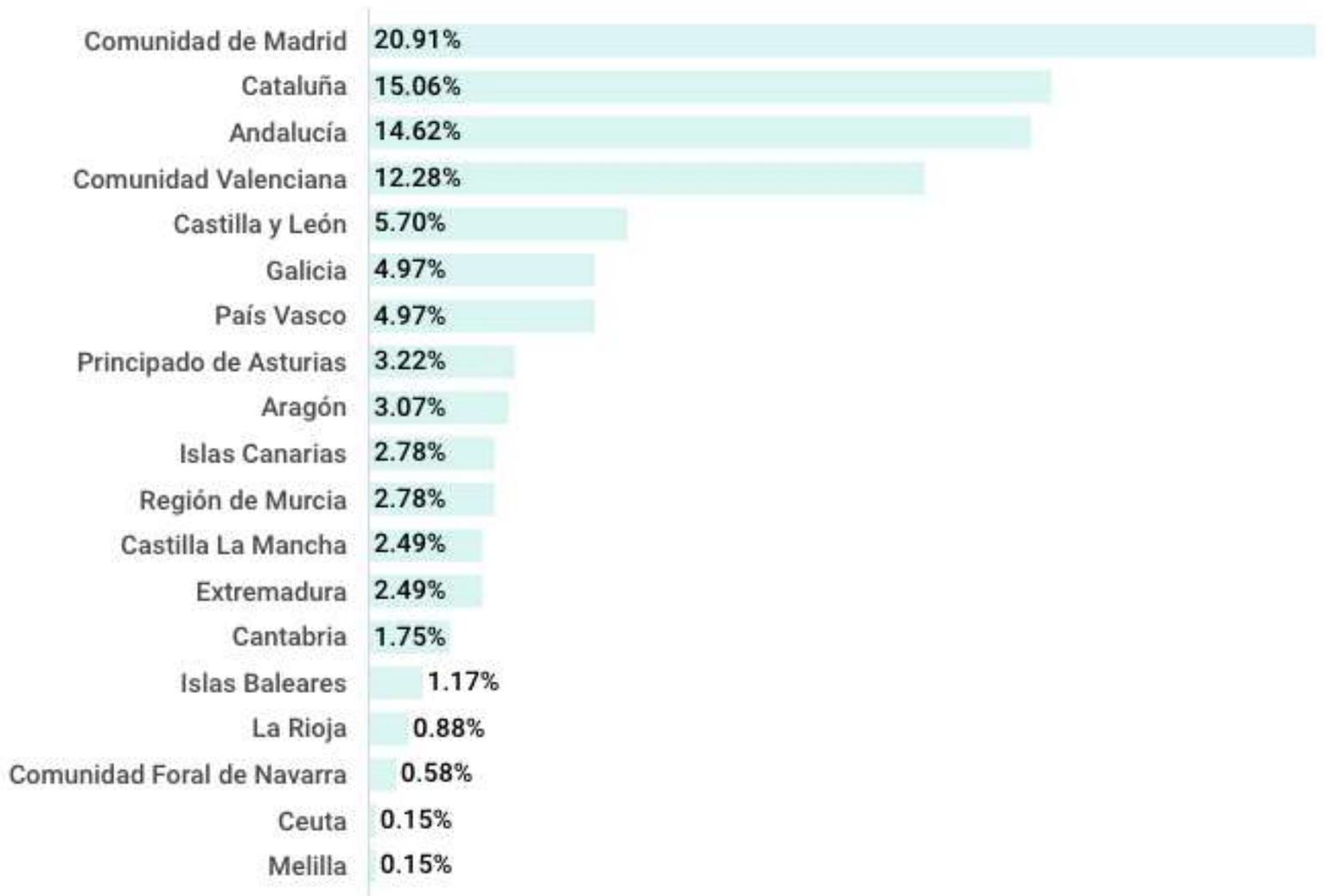


DISTRIBUCION DE LA MUESTRA POR EDAD



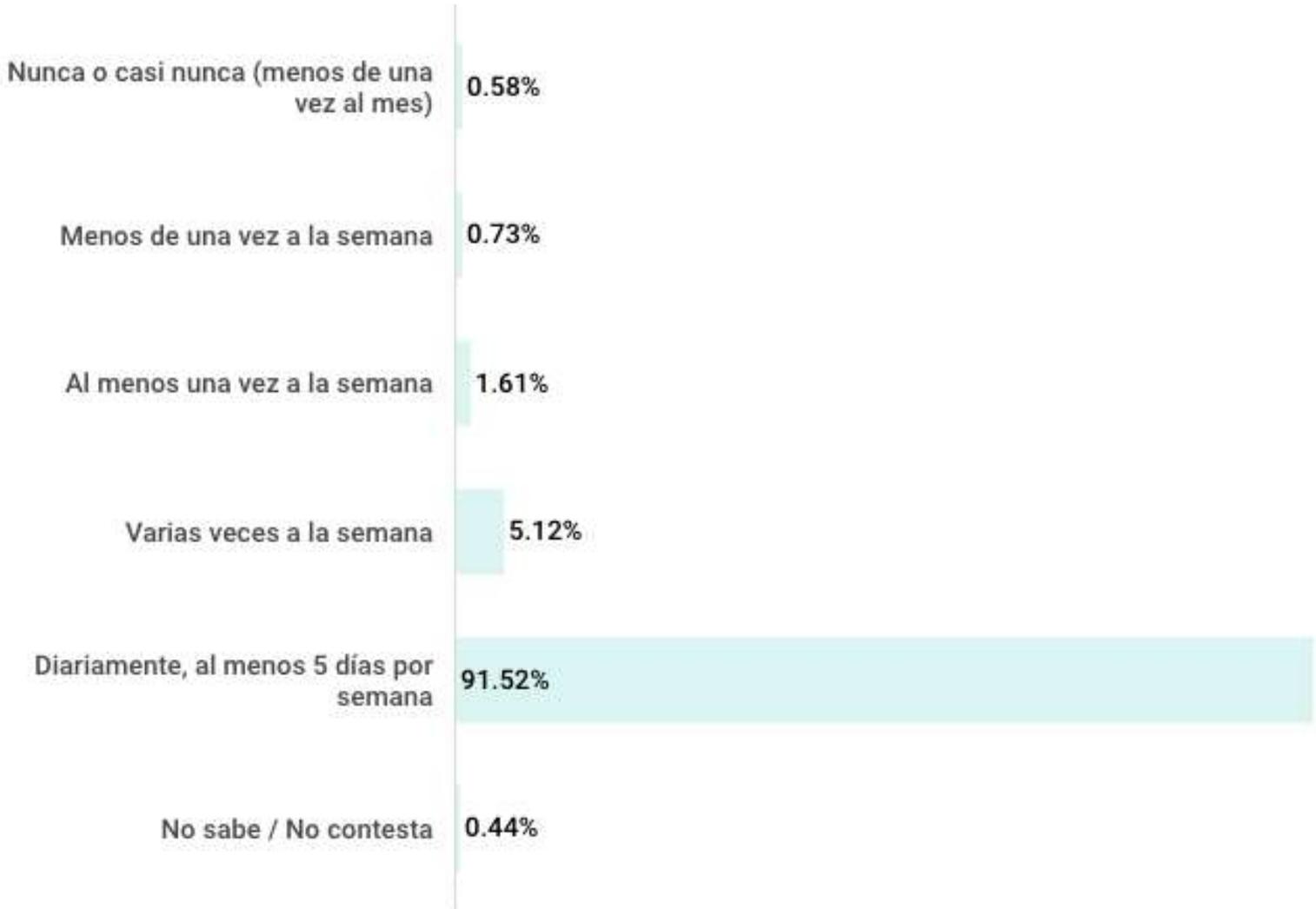
01

COMUNIDAD AUTÓNOMA DE PROCEDENCIA DE LA MUESTRA



01

FRECUENCIA DE CONEXIÓN A INTERNET DE LA MUESTRA



01

CONOCIMIENTO Y USO DEL BIG DATA Y LA INTELIGENCIA ARTIFICIAL



01

FAMILIARIZACIÓN CON EL BIG DATA Y LA INTELIGENCIA ARTIFICIAL



MEDIA BIG DATA

2.85

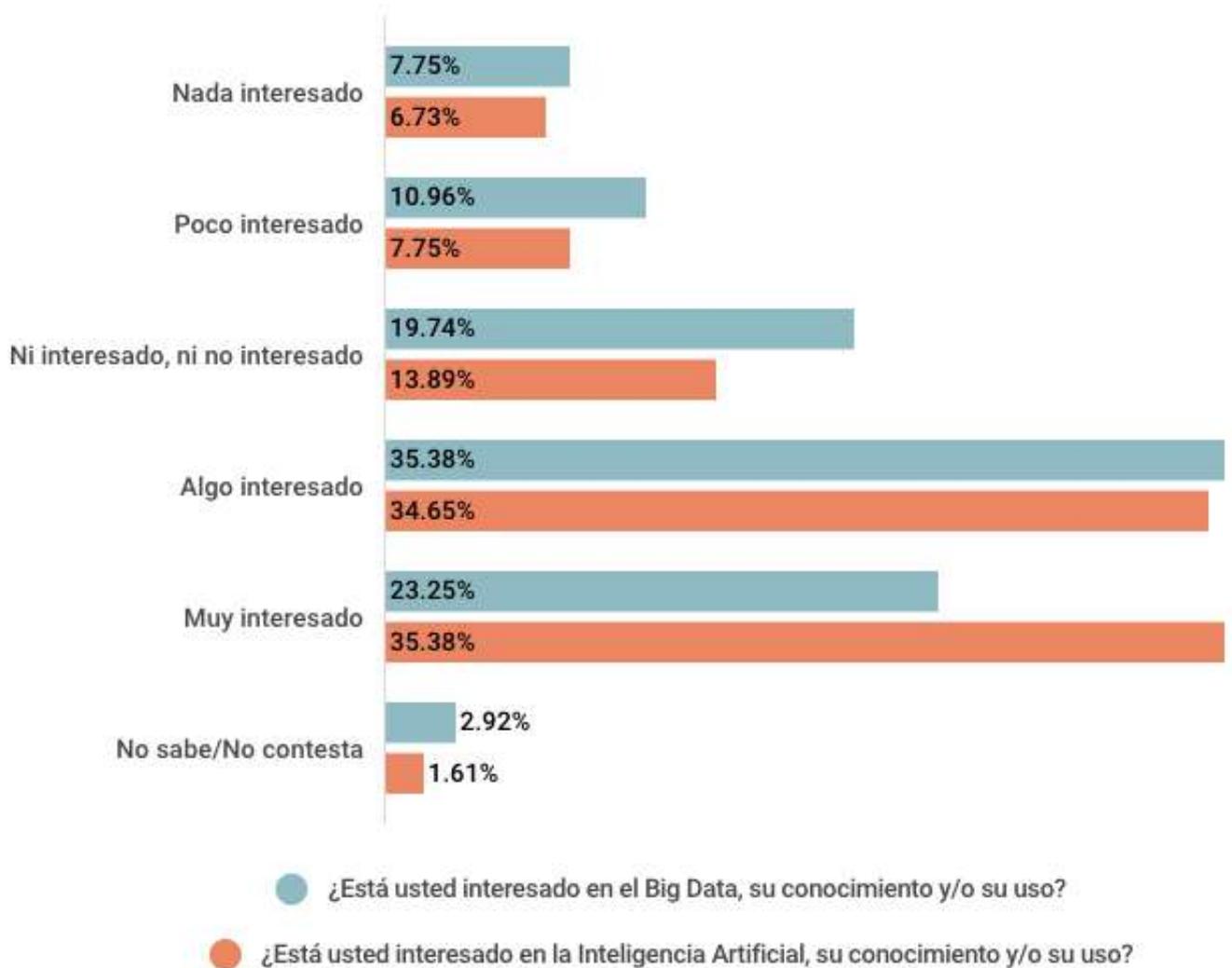
MEDIA INTELIGENCIA ARTIFICIAL

3.43

La población encuestada está más familiarizada con el concepto de Inteligencia Artificial (M=3.43; DT=1.276) que con el de Big Data (M=2.85; DT=1.445).

01

INTERÉS EN EL BIG DATA Y LA INTELIGENCIA ARTIFICIAL



MEDIA BIG DATA

3.64

MEDIA INTELIGENCIA ARTIFICIAL

3.89

La población encuestada está más interesada en conocer o usar la Inteligencia Artificial (M=3.89; DT=1.209) que el Big Data (M=3.64; DT=1.248).

01

ÁREAS DE LA VIDA DIRECTAMENTE RELACIONADAS



CONDUCTA COMO CONSUMIDOR Y USUARIO

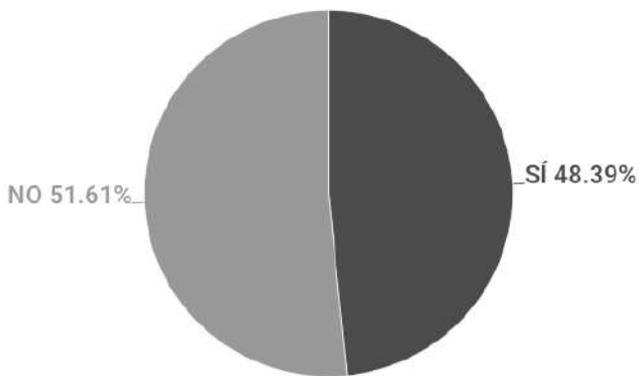
3.93

La población encuestada considera que el área que se relaciona más en su vida diaria con la Inteligencia Artificial y el Big Data es su conducta como consumidor y usuario (M=3.93; DT=1.15), seguido de su ámbito laboral (M=3.77; DT=1.26)

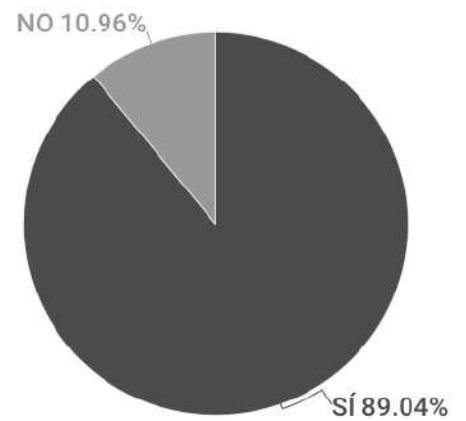
01

USO DE APLICACIONES CON BIG DATA E INTELIGENCIA ARTIFICIAL

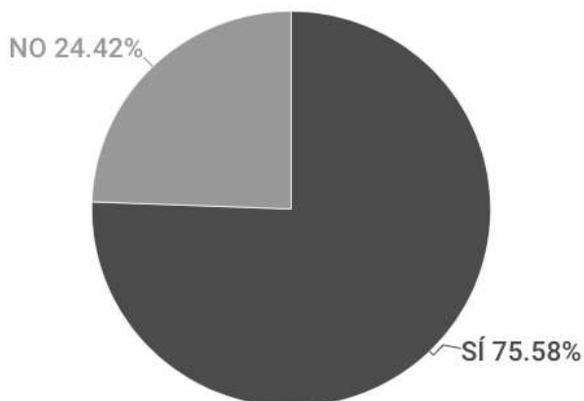
ASISTENTES VIRTUALES
(COMO SIRI, ALEXA, AURA O CORTANA)



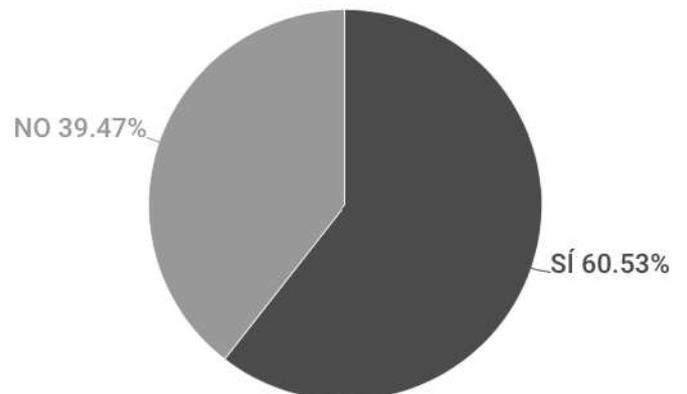
REDES SOCIALES
(COMO FACEBOOK O TWITTER)



PLATAFORMAS DE VIDEO STREAMING
(COMO NETFLIX , HBO, MOVISTAR)



PLATAFORMAS DE AUDIO STREAMING
(COMO SPOTIFY, IVOOX)



01

**ACTITUDES Y
PERCEPCIONES
HACIA EL BIG DATA Y LA
INTELIGENCIA ARTIFICIAL**



01

BENEFICIOS PERCIBIDOS DEL BIG DATA



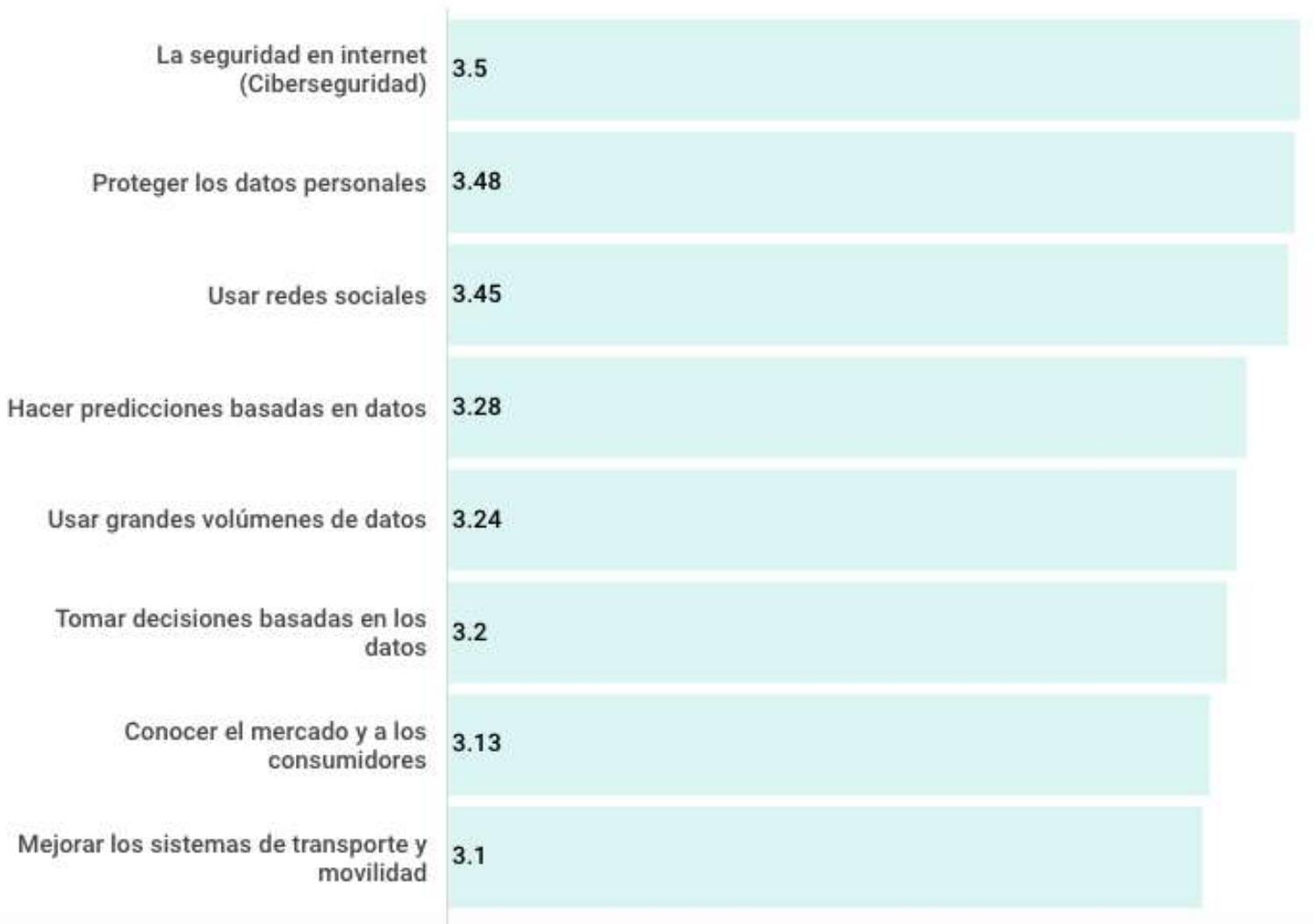
SISTEMAS DE TRANSPORTE Y MOVILIDAD

3.89

El mayor beneficio del Big Data percibido por la población encuestada es la mejora de los sistemas de transportes y movilidad (M=3.89 DT=1.26)

01

RIESGOS PERCIBIDOS DEL BIG DATA



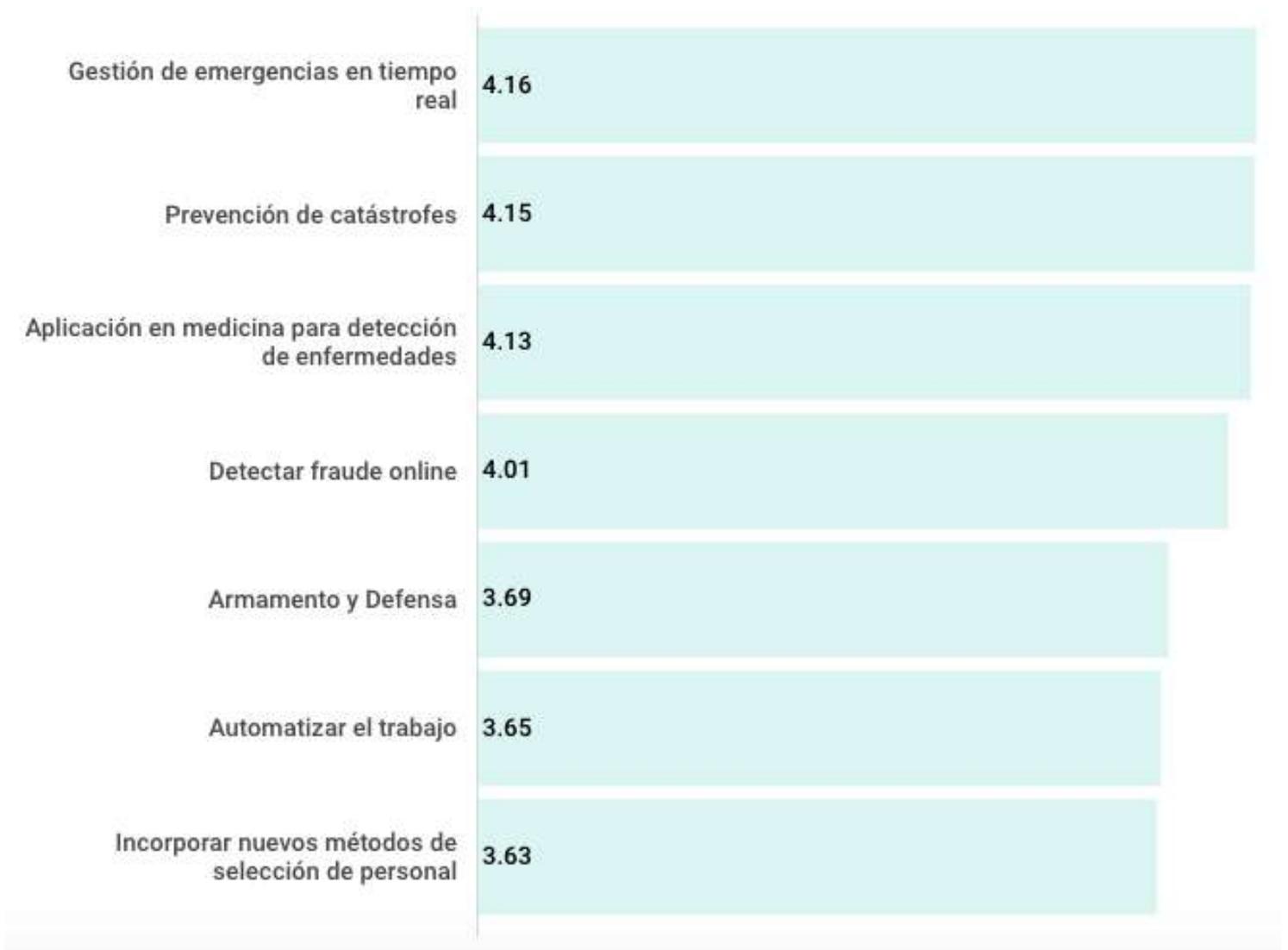
LA SEGURIDAD EN INTERNET

3.5

El mayor riesgo del Big Data percibido por la población encuestada es la seguridad en internet (M=3.5 DT=1.42), seguido por la protección de datos personales (M=3.48 ; DT=1.41)

01

BENEFICIOS PERCIBIDOS DE LA INTELIGENCIA ARTIFICIAL



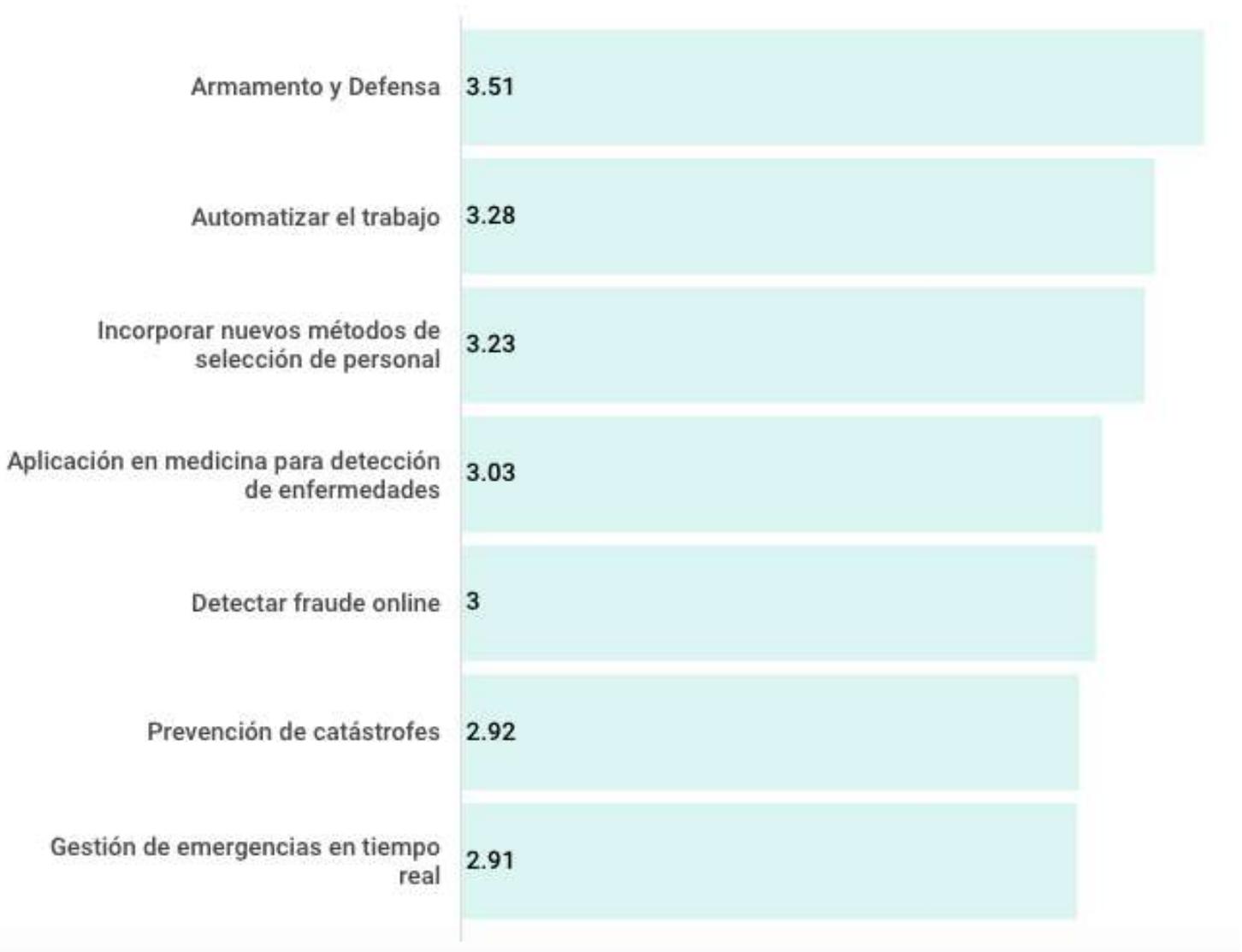
GESTIÓN DE EMERGENCIAS EN TIEMPO REAL

4.16

El mayor beneficio de la Inteligencia Artificial percibido por la población encuestada es la gestión de emergencias en tiempo real (M=4.16 DT=1.17)

01

RIESGOS PERCIBIDOS DE LA INTELIGENCIA ARTIFICIAL



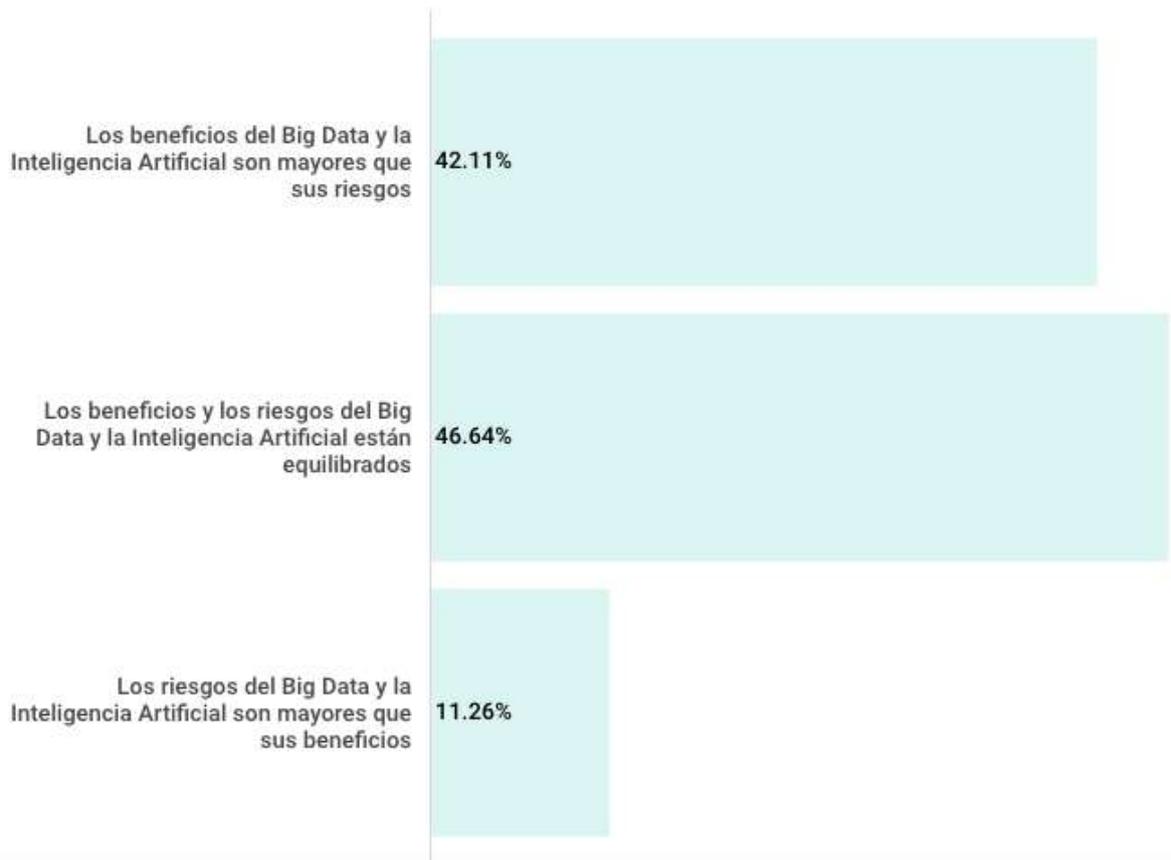
ARMAMENTO Y DEFENSA

3.51

El mayor riesgo de la Inteligencia Artificial percibido por la población encuestada es el armamento y la defensa (M=3.51 DT=1.42)

01

BALANCE DE BENEFICIOS Y RIESGOS



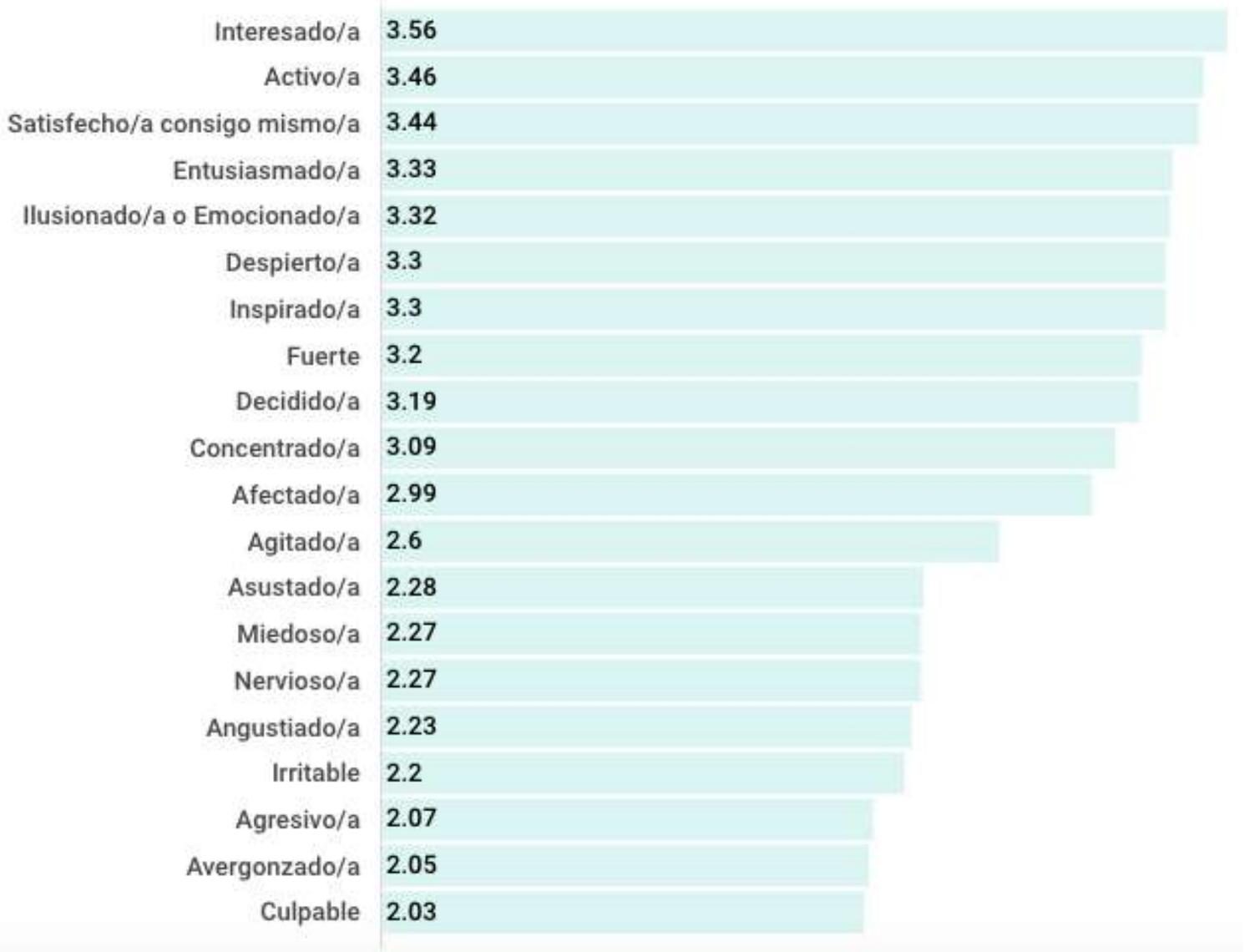
BENEFICIOS Y RIESGOS EQUILIBRADOS

46.6%

Un 46,64% de los encuestados piensan que los beneficios y riesgos del Big Data y la Inteligencia Artificial están equilibrados, seguido de un 42.11% que piensa que los beneficios son mayores que los riesgos.

01

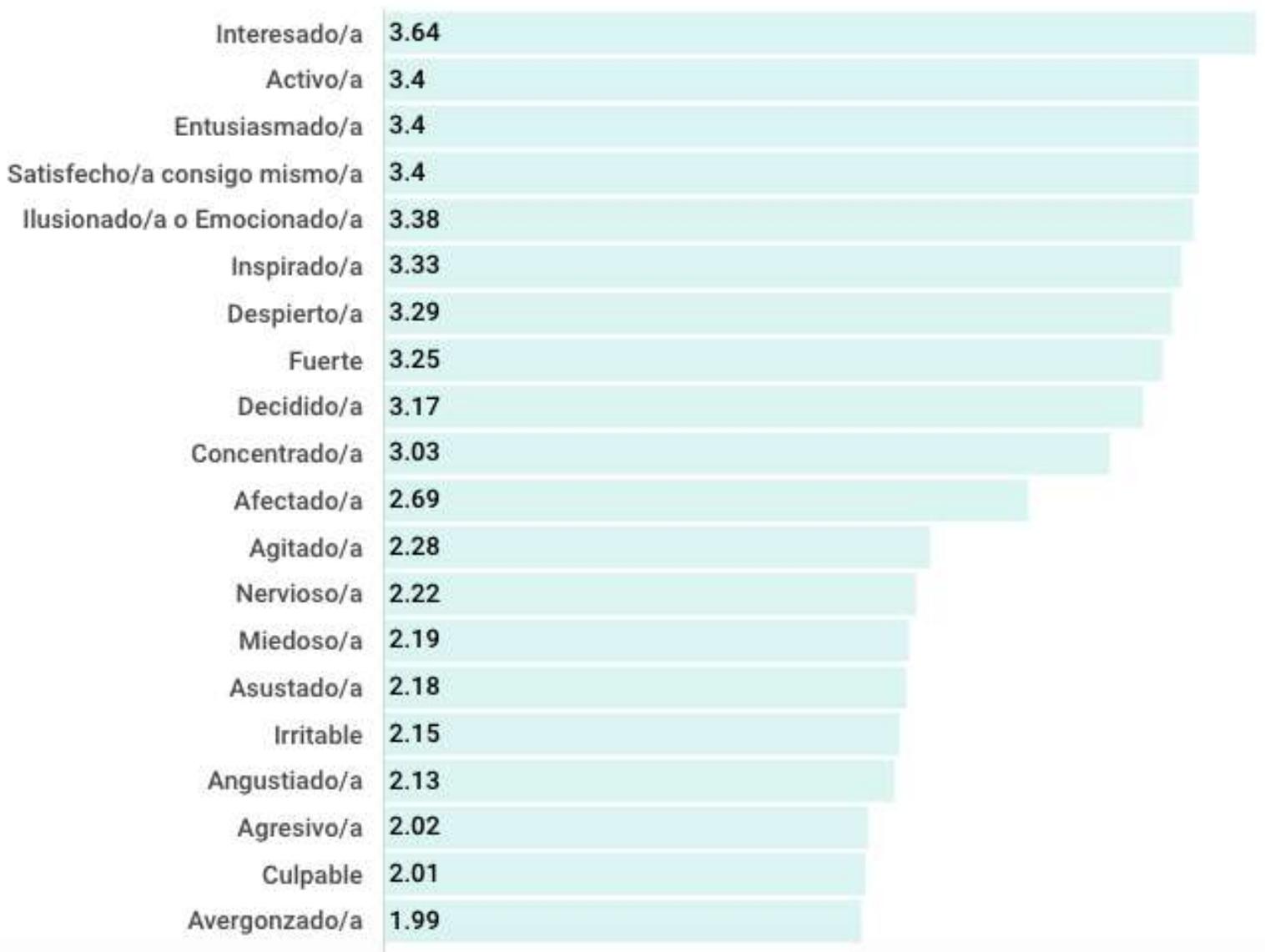
EMOCIONES HACIA EL BIG DATA



Entre las emociones que inspira a los encuestados el Big Data, destaca el positivismo con: Interesado/a, Activo/a y Satisfecho/a

01

EMOCIONES HACIA LA INTELIGENCIA ARTIFICIAL



Entre las emociones que inspira a los encuestados la Inteligencia Artificial, destaca igualmente el positivismo con:
Interesado/a, Activo/a y Entusiasmado/a

01

**MEDIOS A TRAVÉS DE LOS
QUE SE INFORMAN SOBRE
BIG DATA E INTELIGENCIA
ARTIFICIAL**



01

MEDIOS A TRAVÉS DE LOS QUE SE INFORMAN



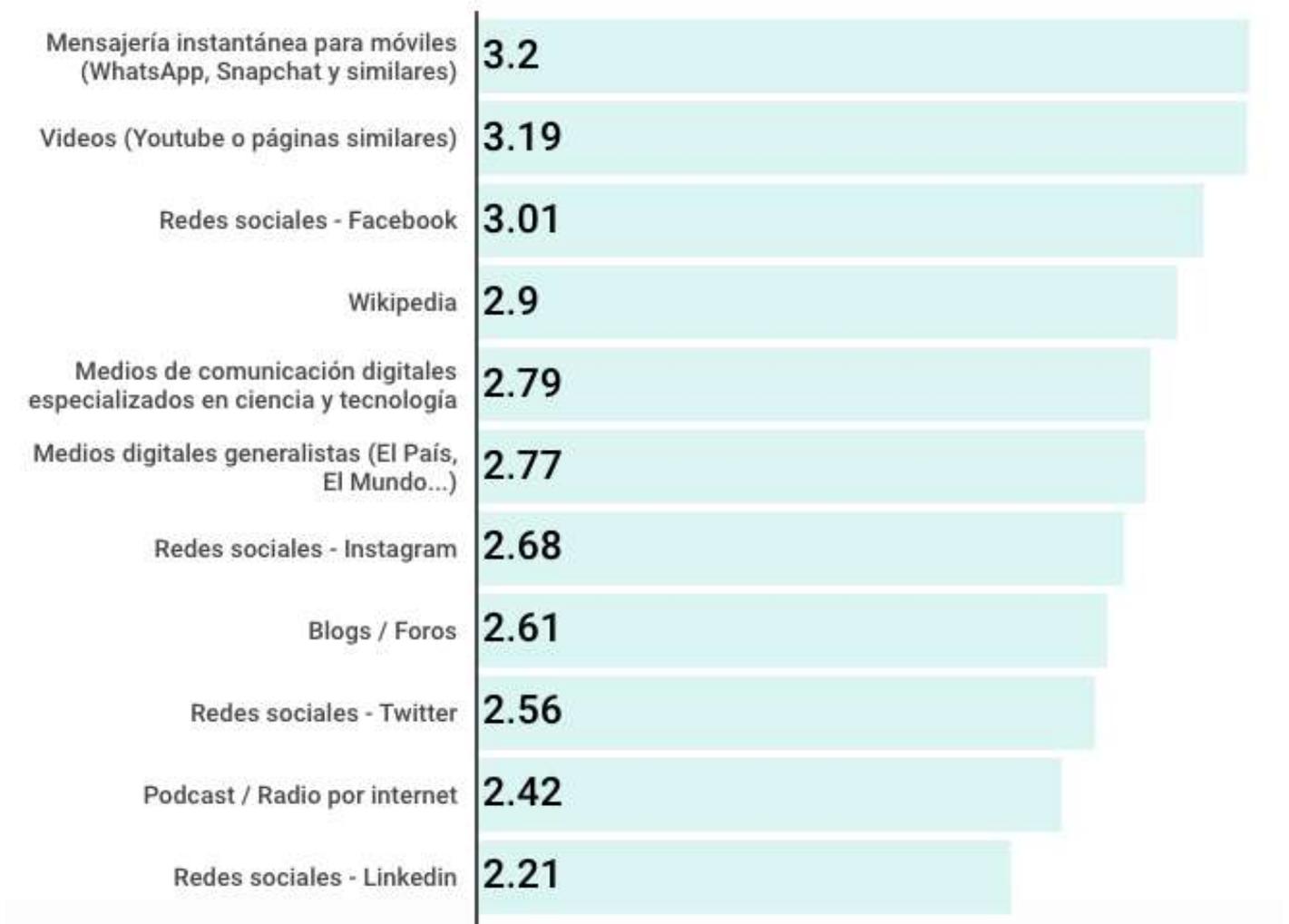
INTERNET

3.33

El medio preferente es Internet, seguido del entorno personal, la televisión y el entorno laboral (M=3.33; DT=1.20).

01

MEDIOS DE INTERNET A TRAVÉS DE LOS QUE SE INFORMAN



MENSAJERÍA INSTANTÁNEA

3.2

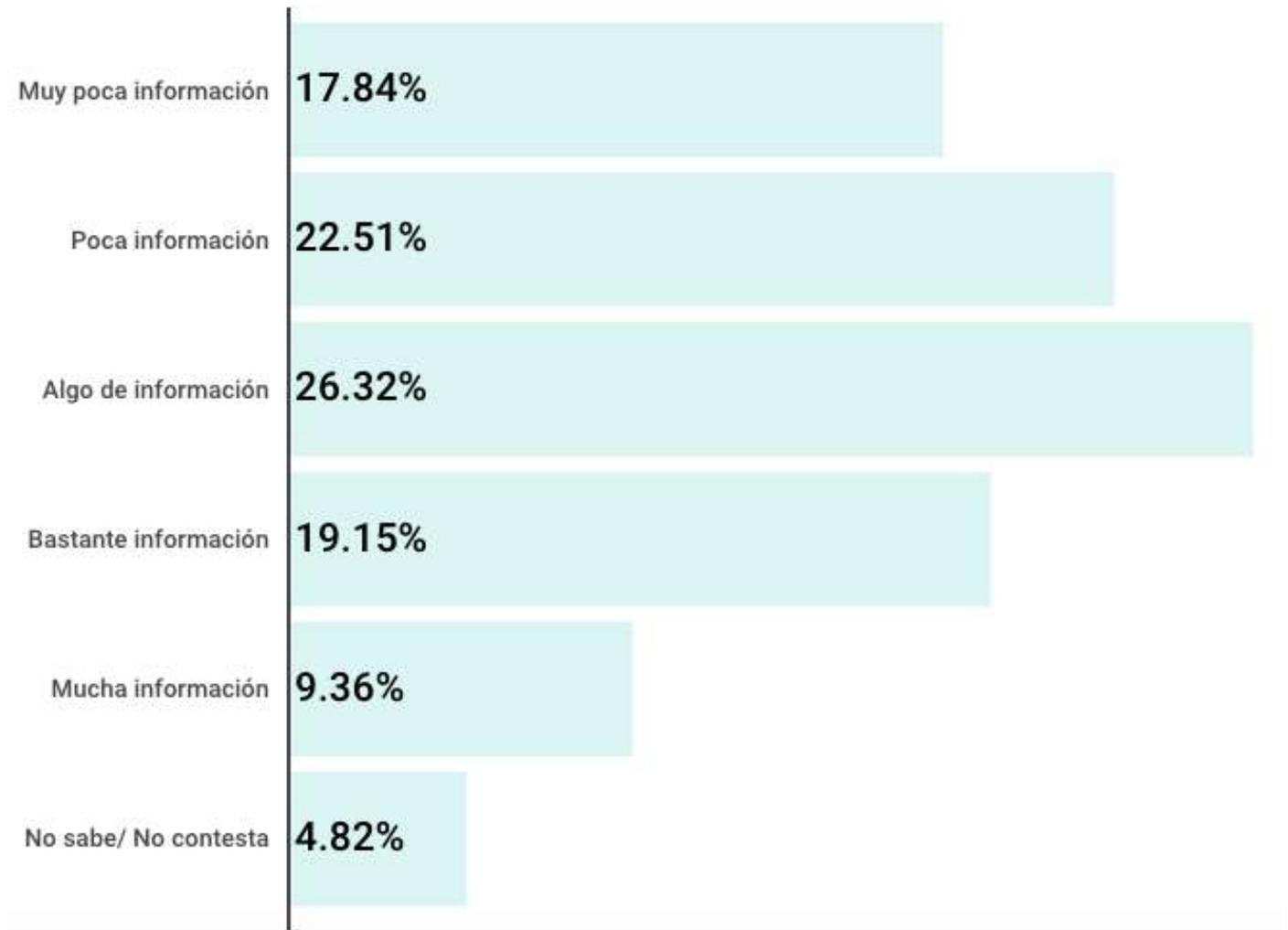
VÍDEOS ON LINE

3.19

Entre las personas que se informan a través de Internet, el medio preferido es la mensajería instantánea (M=3.2; DT=1.36), seguido de cerca por los vídeos online (M=3.19; DT=1.19)

01

CANTIDAD DE INFORMACIÓN QUE RECIBE



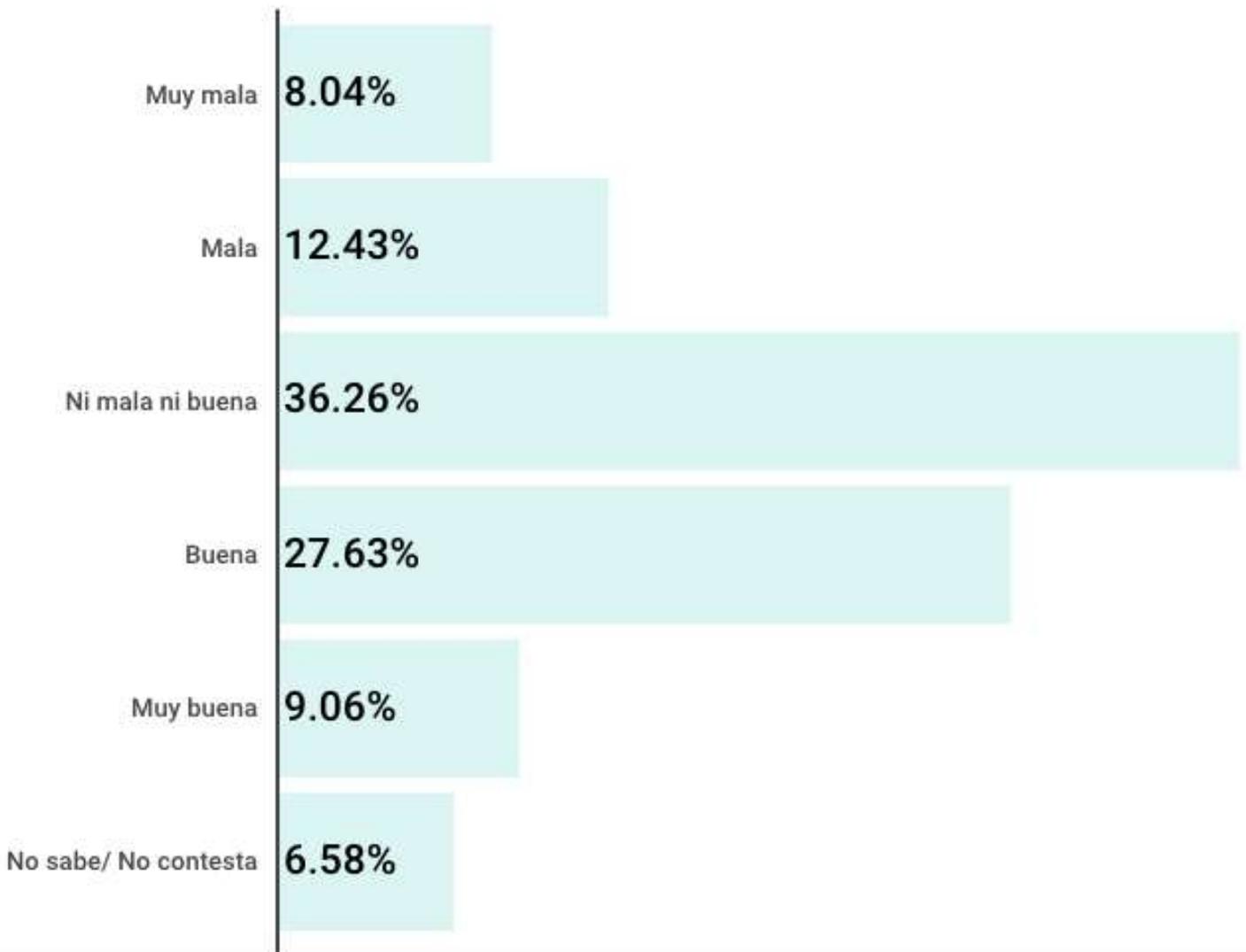
MEDIA

2.94

La población encuestada considera que recibe una información sobre Big Data e Inteligencia Artificial aceptable (M=2.94 DT=1.39).

01

CALIDAD DE LA INFORMACIÓN QUE RECIBE



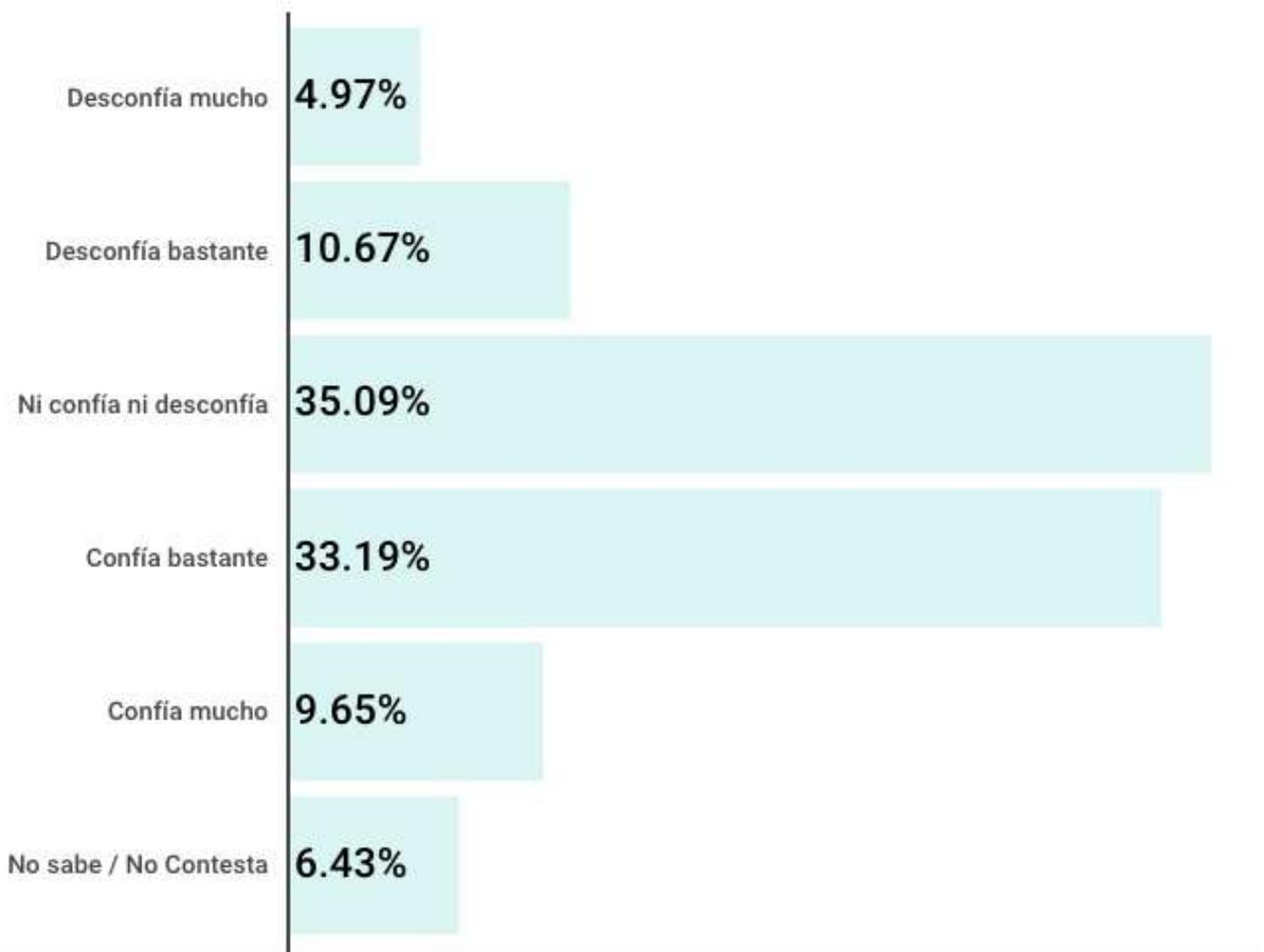
MEDIA

3.37

La población encuestada considera que la calidad de la información que reciben sobre Big Data o Inteligencia Artificial es aceptable o buena (M=3.37; DT=1.24)

01

CONFIANZA EN LA INFORMACIÓN QUE RECIBE



MEDIA

3.51

Los encuestados muestran bastante confianza en la información que reciben sobre Big Data o Inteligencia Artificial (M=3.51; DT=1.16)

01

CLAVES

CONOCIMIENTO



El grado de conocimiento de la Ciencia de Datos es moderado.



En la actualidad son tecnologías muy relevantes, por lo que convendría que hubiera mejoras en ese conocimiento por parte de los ciudadanos.



El interés de los ciudadanos es mayor que su grado de familiarización.



Es preciso que la ciudadanía no especializada tenga acceso a información sobre estas tecnologías, por lo que hay que poner el foco sobre los medios y los comunicadores.

01

El Interés que despiertan el Big Data y la Inteligencia Artificial tienen grandes contrastes

MUY INTERESADO

NADA INTERESADO

Big Data

23.2%

7.7%

Inteligencia Artificial

35.4%

6.7%

El grado de familiarización que presentan el Big Data y la Inteligencia Artificial también presenta contraste

MUY INTERESADO

NADA INTERESADO

Big Data

10%

25%

Inteligencia Artificial

16.8%

9.2%

01

CLAVES

ACTITUDES



Las actitudes que muestran los ciudadanos son positivas.



Se perciben mayores beneficios que riesgos en todas las aplicaciones analizadas.



Los hombres perciben más beneficios que las mujeres.



Las personas de mayor edad perciben más beneficios en las diferentes aplicaciones tanto del Big Data como de la Inteligencia Artificial.



Los ciudadanos consideran que la ciencia de datos tiene relación directa con su conducta de consumidor y usuario.

01

CLAVES

ACTITUDES



La Inteligencia Artificial obtiene una valoración más positiva que el Big Data.



Los ciudadanos muestran un conocimiento moderado y un interés alto.



La actitud hacia las aplicaciones del Big Data y la Inteligencia Artificial es más escéptica entre los rangos de edad jóvenes (18 a 24 y 25 a 34 años), sin embargo son los más familiarizados y los que más conocimientos tienen sobre ellas.

01

46.6%

considera que los beneficios y los riesgos del Big Data y la Inteligencia Artificial están equilibrados

INTERÉS

es la emoción más repetida que provoca el Big Data y la Inteligencia Artificial

01

CLAVES

INFORMACIÓN



El acceso prioritario a la información sobre Big Data e Inteligencia Artificial se realiza a través de Internet.



Hay un volumen elevado de personas que no se informan sobre el tema en ningún medio.



La cantidad de información que reciben es considerada suficiente.



La calidad de la información que reciben es buena.



La confianza en la información que se recibe es bastante.

01

17,8%

consideran que la cantidad de información que reciben sobre Big Data e Inteligencia Artificial es muy poca.

38,6%

no se informan acerca del Big Data y la Inteligencia Artificial.

27.6%

consideran que la calidad de la información que reciben sobre Big Data e Inteligencia Artificial es buena

33.2%

confían bastante en la información que reciben sobre Big Data e Inteligencia Artificial

ENCUESTA DE CONOCIMIENTO Y PERCEPCIÓN SOCIAL DE LA CIENCIA DE DATOS, BIG DATA E INTELIGENCIA ARTIFICIAL

FICHA TÉCNICA

Técnica empleada:

Cuestionario estructurado distribuido a través de la plataforma online *Qualtrics*.

Población:

Personas de 18 años de edad en adelante, residentes en España.

Muestra:

Se han obtenido 684 respuestas, distribuidas de forma proporcional a la población de cada una de las 17 comunidades autónomas, teniendo en consideración el tamaño de hábitat.

Procedimiento de muestreo:

Selección de unidades por rutas aleatorias y cuotas de sexo y edad.

Error muestral:

Margen de error de ± 0.5 puntos porcentuales para un nivel de confianza del 95,5%, 2σ y $p=q$, con el supuesto de muestreo aleatorio simple, calculado considerando muestras no proporcionales.

Trabajo de campo:

Del 10 al 24 de Enero de 2020.

Ámbito:

Todo el territorio nacional (Península, Islas Baleares e Islas Canarias).

Empresa encargada del trabajo de campo:

Qualtrics



Proyecto financiado por la Fundación Española para la Ciencia y la Tecnología (FECYT) en la Convocatoria de ayudas para el fomento de la cultura científica, tecnológica y de la innovación 2019-2020. [FCT-18-13437]

OBSERVATORIO DE LOS CONTENIDOS AUDIOVISUALES UNIVERSIDAD DE SALAMANCA

Facultad de Ciencias Sociales. Campus Miguel de Unamuno. Edificio F.E.S. Paseo de Francisco Tomás y Valiente, s/n. 37007. Salamanca. España
www.ocausal.es/datasciencespain

Enero 2021. This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. Diseño base e imágenes de Canva.



CIENCIA DE DATOS EN ESPAÑA

CONOCIMIENTOS Y PRÁCTICAS DE LOS PERIODISTAS PARA COMUNICAR CIENCIA DE DATOS

DATASCIENCE SPAIN



Proyecto financiado por la Fundación Española para la Ciencia y la Tecnología (FECYT) en la Convocatoria de ayudas para el fomento de la cultura científica, tecnológica y de la innovación 2019-2020. [FCT-18-13437]



UNIVERSIDAD
DE SALAMANCA

02

INTRODUCCIÓN

La comunicación de la Ciencia de Datos es uno de los temas más relevantes en las sociedades actuales, pero también uno que todavía está en parte rodeado por la falta de comprensión. Los resultados del primer estudio sobre "Conocimiento y percepción de la Ciencia de Datos, el Big Data y la Inteligencia Artificial" muestran que los niveles de conocimiento de la ciudadanía son moderados, pero el interés es alto.

Al tratarse de tecnologías todavía novedosas y que generan escepticismo en ciertos sectores de la población, el papel de la comunicación científica es de particular relevancia. Así, a pesar de la ruptura del monopolio comunicacional que hasta ahora poseían los periodistas, su rol sigue siendo determinante en este aspecto.

Por este motivo, este trabajo estudia el conocimiento, las prácticas y las percepciones de periodistas especializados en ciencia y tecnología, revisando y analizando su proceso personal de creación y publicación de contenidos relacionados con la Ciencia de Datos, el Big Data o la Inteligencia Artificial.

02

¿CUÁLES SON LOS DESAFÍOS?



"¿Cuál es el conocimiento y la percepción de los periodistas que escriben sobre Ciencia de Datos?"

¿Cómo es el perfil del periodista que elabora estos contenidos y cuál es su rutina de trabajo?

¿Cómo entiende el periodista o la periodista la comprensibilidad de un contenido de ciencia de datos?

¿Qué beneficios y riesgos perciben los periodistas en el big data y la inteligencia artificial?

¿Cómo es el proceso de trabajo de un artículo, sus fuentes y sus características?

02

FACT SHEET

- El perfil de periodistas que escribe sobre temas de ciencia de datos es de profesionales con amplia experiencia y mediana edad, la mayoría centrados en medios digitales.
- Las publicaciones están muy ligadas a la actualidad.
- Las principales fuentes son del sector profesional y empresarial.
- Se percibe que el público está muy influenciado por la ciencia ficción.
- Fomentan el uso de un lenguaje claro, asequible y riguroso, sin tecnicismos pero explicando cada término y acompañándolo con representaciones gráficas.
- Los periodistas consideran que la sociedad percibe muchos riesgos en el Big Data y la Inteligencia Artificial.
- Los periodistas aportan una visión optimista y una posición informativa neutral.

02

METODOLOGÍA

El estudio se ha desarrollado mediante entrevistas en profundidad realizadas a una muestra aleatoria de 10 periodistas que tenían publicaciones recientes de contenidos informativos relacionados con la Ciencia de Datos, Big Data o Inteligencia Artificial.

La muestra final estuvo formada por 2 mujeres y 8 hombres, todos ellos profesionales con residencia en España, que colaboran habitualmente con medios de comunicación y que elaboran contenidos relacionados con la Ciencia de Datos, el Big Data, o la Inteligencia Artificial desde cualquier perspectiva.

La entrevista se estructuró en preguntas abiertas y preguntas de reconstrucción a partir de uno de los artículos publicados en medios en línea.

Las preguntas se estructuraron en 4 bloques:

- El periodista y su trabajo
- Comprensibilidad de los contenidos
- Beneficios y riesgos percibidos
- Entrevista de reconstrucción con un artículo recientemente publicado

RESULTADOS

“¿Cómo es el perfil del periodista que elabora estos contenidos y cuál es su rutina de trabajo?”



Los entrevistados realizan contenidos habitualmente para medios digitales, sobre todo generalistas, como periodistas independientes o por cuenta ajena. Tienen una experiencia de 10 a 20 años y su dedicación es mayoritariamente completa.

Obtienen gran parte de la información siguiendo la actualidad, a otros medios especializados, a expertos y asistiendo a eventos de interés. Con base en ello, sienten libertad de proponer temas, pero los responsables pueden realizar un filtrado o priorizar.

Las fuentes más habituales son profesionales y expertos que publican trabajos y comparten sus resultados, empresas que desarrollan tecnología o informes institucionales.

Para mejorar la presencia de estos contenidos en medios, destacan aspectos como: la formación de los periodistas y su actualización, aumentar la calidad y establecer una relación entre las aplicaciones de la tecnología y los intereses de los lectores.

02

“Yo siempre digo que las noticias, nunca están en los teletipos; mis principales fuentes son básicamente las propias empresas en el mundo de la tecnología, que se comunican directamente conmigo, muchas notas de prensa que llegan al correo, también los blogs y medios internacionales”

02

CLAVES

PERFIL DEL PROFESIONAL



El perfil de periodistas que escribe sobre temas de Ciencia de Datos es de profesionales con una amplia experiencia y mediana edad, la mayoría centrados en medios digitales.



Las publicaciones están muy ligadas a la actualidad y los eventos y los periodistas están en constante búsqueda de nuevas ideas, especialmente temas que tengan un impacto en la sociedad.



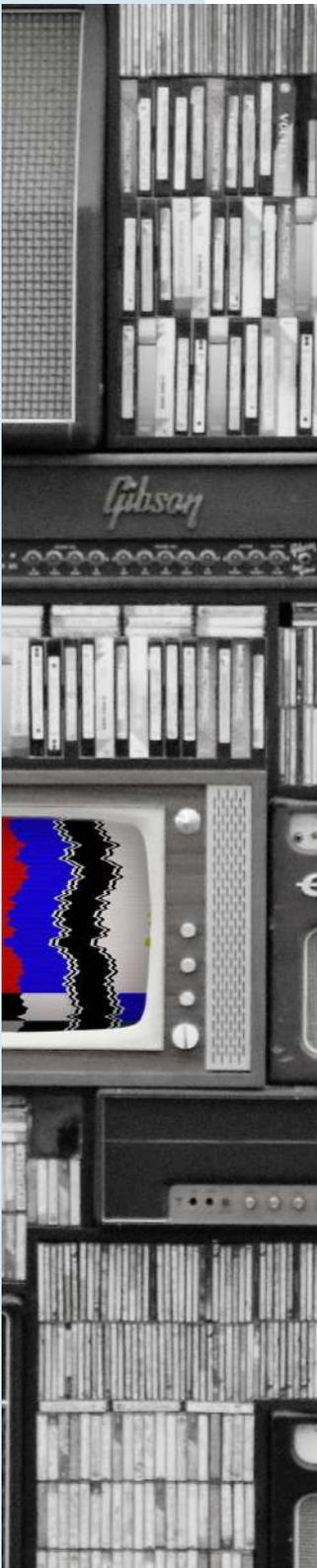
Se apoyan en sus contactos y en expertos del sector profesional, pero escasean las fuentes de académicos.

RESULTADOS

“¿Cómo entiende el o la periodista la comprensibilidad de un contenido de Ciencia de Datos?”

Los entrevistados consideran que existe una necesidad de mejorar los conocimientos básicos que los ciudadanos tienen sobre Ciencia de Datos.

Lo que entienden por contenidos comprensibles incluye: el uso de un lenguaje claro y asequible, pero a la vez riguroso y sin tecnicismos, apoyándose en expresión gráfica cuando sea posible. A la vez, ejemplificar los contenidos con historias reales y la aplicación práctica de estas tecnologías en la sociedad.



02

“Me parece que para que alguien te lea tienes que escribir de una forma clara y a la vez rigurosa. Para mí es la combinación ideal. Que no use tecnicismos, que sea un lenguaje llano, pero que no sea simplista, es decir, que admita complejidades dentro de una explicación no técnica”



02

CLAVES

COMPENSIBILIDAD



Existe una necesidad de mejorar los conocimientos básicos que los ciudadanos tienen sobre Ciencia de Datos.



Perciben que el público está muy influenciado por la ciencia ficción.



Para ser comprensible destaca el uso de un lenguaje claro, asequible y riguroso, sin tecnicismos pero explicando cada término y usando representaciones gráficas.



El mejor enfoque es mostrar soluciones a problemas cotidianos y no perder de vista la ética.

RESULTADOS

“¿Qué beneficios y riesgos perciben los periodistas en el Big Data y la Inteligencia Artificial?”

Los beneficios percibidos más destacados han sido facilitar tareas diarias, optimizar recursos y ser más eficientes. En los sectores de aplicación destaca la automatización de procesos industriales, la medicina y la salud (prevención y análisis de datos), la movilidad urbana o aérea y la economía.

Los riesgos percibidos se centran en la pérdida de empleos, la seguridad en la red (privacidad, fraudes y amenazas) y la exposición selectiva a los medios.



02

“

“Uno de los riesgos que yo creo que la sociedad percibe principalmente es algo que ocurría también en la Segunda Revolución Industrial, la posibilidad de sustituir al trabajador por una máquina, lo que puede terminar con el trabajo de muchas personas. Los empleos van a cambiar en el futuro y creo que el principal miedo de la gente es quedarse sin trabajo, tener que reinventarse, buscarse la vida de otra manera.”

”

02

CLAVES

BENEFICIOS PERCIBIDOS



La mayoría de los periodistas entrevistados consideran que la Ciencia de Datos facilita tareas, optimiza recursos y les hace eficientes.



Consideran sectores de interés para su aplicación: la automatización de procesos industriales, medicina y salud, movilidad urbana o aérea y economía.



En cuanto a usos más personales destacan: las comunicaciones humanas, las redes sociales y la geolocalización.

02

CLAVES

RIESGOS PERCIBIDOS

Los periodistas consideran que la sociedad percibe muchos riesgos en el Big Data y la Inteligencia Artificial:



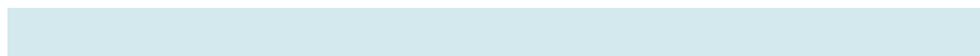
La pérdida de empleos por la sustitución de personas por máquinas debido al vínculo que las personas han interiorizado entre estas tecnologías y la ciencia ficción.



La seguridad en la red y todo lo que afecta a la privacidad, el uso de datos personales, los fraudes o las ciberamenazas.



La exposición selectiva vinculada a las redes sociales, al igual que la proliferación de noticias falsas y el exceso de vanidad.



RESULTADOS

“¿Cómo es el proceso de trabajo de un artículo, sus fuentes y sus características?”

”

En su rutina de trabajo diaria coinciden en unos pasos básicos de trabajo: informarse, investigar, acudir a fuentes y expertos, redactar y revisar resultados.

Las aplicaciones y técnicas de big data de las que más se habla han sido: el uso de grandes volúmenes de datos, hacer predicciones basadas en datos y la toma de decisiones, seguidos por la ciberseguridad y conocer al mercado y los consumidores, para dejar en último lugar la protección de datos personales, la movilidad y las redes sociales.

Las aplicaciones y técnicas de la inteligencia artificial de las que más se habla son: la medicina para la detección de enfermedades y la automatización del trabajo, seguido por la gestión de emergencias en tiempo real, los nuevos métodos de selección de personal, la detección del fraude en línea y la prevención de catástrofes, y en última posición el armamento y defensa.



RESULTADOS

“¿Cómo es el proceso de trabajo de un artículo, sus fuentes y sus características?”

”

El tratamiento de los artículos se realiza prestando más atención a los beneficios y otorgando menos visibilidad a los riesgos.

Cumplen lo mencionado en el resto de bloques, de manera que la mayor parte de las fuentes fueron profesionales o empresariales, al mismo tiempo, se hace hincapié en las dificultades relacionadas con el acceso y la disponibilidad de dichas fuentes. Además, si están condicionados por la actualidad, los contenidos necesitan un desarrollo ágil.

Uno de las demandas de los periodistas es la continuidad de los trabajos, que pueden ser proyectos más amplios y analizarse desde otras perspectivas, como la ética o el impacto social.



02

“Acudes a gente que ya conoces, expertos en la materia y que sabes que te pueden ayudar.”

“Mi posición es presentar esta tecnología y esbozar, en su caso, pros y contras.”

“Yo intento tener una continuidad en los temas de impacto social y ético de la tecnología. Cómo esas tecnologías están impactando a la gente o cosas que nos están contando.”

02

CLAVES

PROCESOS DE TRABAJO



Pasos básicos: informarse, investigar, acudir a fuentes o expertos, redactar y revisar resultados.



Los periodistas aportan una visión optimista.



Posición informativa neutral, mostrando hechos y evidencias para que los lectores puedan sacar sus conclusiones.



Preocupados por su formación y su actualización de conocimientos.



Buscan establecer una relación entre las aplicaciones de la tecnología y los intereses de los lectores.

02

9 DE 10

*periodistas entrevistados
considera necesario mejorar los
conocimientos básicos que tiene
la ciudadanía sobre este tema.*

7 DE 10

*periodistas entrevistados
considera que la calidad de las
publicaciones es buena.*

02

9 DE 10

*periodistas entrevistados
considera que el público confía
en la información.*

9 DE 10

*periodistas entrevistados
considera que la Ciencia de
Datos facilita tareas, optimiza
recursos y les hace eficientes.*



Proyecto financiado por la Fundación Española para la Ciencia y la Tecnología (FECYT) en la Convocatoria de ayudas para el fomento de la cultura científica, tecnológica y de la innovación 2019-2020. [FCT-18-13437]

OBSERVATORIO DE LOS CONTENIDOS AUDIOVISUALES
UNIVERSIDAD DE SALAMANCA

Facultad de Ciencias Sociales. Campus Miguel de
Unamuno. Edificio F.E.S. Paseo de Francisco Tomás y
Valiente, s/n. 37007. Salamanca. España
www.ocausal.es/datasciencespain

Enero 2021. This work is licensed under the Creative
Commons Attribution-NonCommercial-ShareAlike 4.0
International License. Diseño base e imágenes de Canva.



CIENCIA DE DATOS EN ESPAÑA

COBERTURA DE LA CIENCIA DE DATOS POR LOS MEDIOS DIGITALES EN ESPAÑA

DATASCIENCE SPAIN



Proyecto financiado por la Fundación Española para la Ciencia y la Tecnología (FECYT) en la Convocatoria de ayudas para el fomento de la cultura científica, tecnológica y de la innovación 2019-2020. [FCT-18-13437]



UNIVERSIDAD
DE SALAMANCA

03

INTRODUCCIÓN

Como se viene destacando, la Ciencia de Datos es uno de los campos de estudio más novedosos, controvertidos y con mayor proyección de la actualidad, y su inmersión en todos los ámbitos de la vida es innegable. No obstante, se viene observando que el nivel de conocimiento por parte de los ciudadanos todavía no es completo, y que los periodistas reclaman esfuerzos también en esta línea.

Debido a esto, conviene abordar también los contenidos mediáticos sobre Ciencia de Datos, una de las principales fuentes de información y conocimiento sobre la materia para los ciudadanos.

Con esto, se aborda un análisis mediático de la cobertura, el tratamiento y la divulgación de la Ciencia de Datos en los medios digitales españoles.

DESAFÍOS

“¿Cómo es el tratamiento mediático de los contenidos de Ciencia de Datos?”



03**FACT SHEET**

El tratamiento que se le da al Big Data y la Inteligencia Artificial, aunque con una magnitud débil, es positivo.



Los medios generalistas son los que más frecuentemente cubren el Big Data y la Inteligencia Artificial.



La cobertura de estas materias se realiza en gran medida mediante noticias de economía, destacando también con frecuencia los riesgos de estas tecnologías y su penetración en la esfera privada.

METODOLOGÍA

Para analizar el tratamiento mediático de la Ciencia de Datos, se diseñó un libro de códigos que, por medio de un análisis de contenido, permitió clasificar las noticias de prensa digital española sobre dicho tema.

MUESTRA

Para abordar este trabajo se extrajeron todas aquellas noticias digitales publicadas entre enero y febrero de 2020 que mencionaran los conceptos de Big Data e Inteligencia Artificial.

El total de la base de datos contenía 21486 registros de noticias que cumplían esos requisitos, pero para el análisis que aquí se reporta se utilizó una muestra aleatoria de 750 noticias de prensa digital.

Esta clasificación se realizó en la plataforma Doccano, que permite el etiquetado en línea. En este proceso se descartaron 350 noticias que no eran contenidos noticiosos directamente.

Posteriormente, las 400 noticias seleccionadas fueron codificadas en el programa SPSS, según el libro de códigos diseñado *ad hoc*.

03

MEDIDAS

El libro de códigos recoge 14 variables con el objetivo de conocer y analizar el tratamiento mediático de la Ciencia de Datos.

Se midió la fiabilidad de todas las variables, obteniendo resultados satisfactorios en el estadístico *Alpha de Krippendorff*.

Algunas variables son generales, de carácter nominal, como el Tipo de Medio que publica la noticia: Medio digital generalista, Medio digital especializado en ciencia y tecnología, Blogs o foros, Medios digitales especializados y Otros. Esta variable se utilizó con la finalidad de analizar si los medios especializados en ciencia y tecnología hablaban más de Ciencia de Datos que los medios generalistas.

La variable Temática trata de la sección o del tema en general de la noticia, en este caso se hizo una división en cuanto a las temáticas más importantes, las opciones de esta variable son: Economía, Política, Ciencia, Tecnología y Salud, Sociedad y Educación y Otros. Esta variable se elaboró con el fin de distinguir qué temáticas se vinculaban con más frecuencia a la Ciencia de Datos en la agenda mediática.

03

MEDIDAS

Otras variables son las de aspectos formales del texto y están conformadas por variables independientes y de control.

La variable Extensión de la Noticia, se emplea para analizar si las noticias de Ciencia de Datos tienen un espacio amplio dentro de la agenda mediática. Esta variable fue medida a través de una escala de 4 ítems: de 10 a 25 líneas, de 26 a 45 líneas, de 45 a 60 líneas y más de 60 líneas.

La variable Objeto, hace referencia si en la noticia se habla de Big Data, Inteligencia Artificial o habla de ambas.

La variable Visión establece si la postura de los medios frente a la Ciencia de Datos es negativa, neutra o positiva. Se trata de una variable que sirve para medir la actitud que se transmite sobre la Ciencia de Datos y la percepción de beneficios y riesgos.

La variable Posición estudia si la noticia está a favor o en contra de la tecnología en cuestión, y está codificada en términos de ataque, defensa, o no toma posición.

MEDIDAS

También se midieron dos aspectos relativos a las aplicaciones del Big Data y la Inteligencia Artificial cuyos riesgos y beneficios habían sido analizados en el primer estudio. Para ello, se indicaba si cada una de las aplicaciones seleccionadas (8 en el caso del Big Data y 7 en el caso de la Inteligencia Artificial) aparecía o no en el texto; a esto le seguía una variable que codificaba el atributo dado a esta aplicación, en términos de beneficio, riesgo o neutralo.

La variable Comprensibilidad hace referencia a si un codificar humano es capaz de entender o no un texto presentado. Se divide en 4 ítems:

No se entiende. Sin una preparación educativa suficiente en este campo o con ayuda especializada, es casi imposible entender la idea desarrollada.

Se entiende con dificultad. Una persona no puede entender estos textos por sí sola sin recurrir a ayuda o materiales auxiliares.

Se entiende relativamente bien. Una persona puede entender en general la idea principal de la noticia, aunque tenga ciertas dudas o dificultades en la comprensión en algunas partes.

Se entiende muy bien. Una persona puede entender totalmente la idea principal del texto, sin esfuerzo extra.

03

MEDIDAS

La variable Relación con la vida diaria buscaba medir si los contenidos relacionaban la tecnología en cuestión con la vida diaria, y se evaluó a través de la presencia o ausencia en la noticia de menciones a la relación de la Inteligencia Artificial y el Big Data con los siguientes aspectos: "Mi conducta como consumidor y usuario", "Mi comprensión del mundo", "Mis relaciones con otras personas", "Mi profesión o ámbito laboral" y "Mi formación de opiniones políticas y sociales".

Las variables Género y Encuadre, hacen alusión a la franja informativa de las noticias y al tipo de lenguaje periodístico que se usa en ellas.

La variable Género está codificada en tres opciones: Informativo, Interpretativo e Híbrido.

La variable Encuadre se refería al tipo de lenguaje o desarrollo que se le daba a la noticia. Dividida en: información, descripción, narración y cita de acontecimientos.

RESULTADOS

“¿Cuál es el tratamiento que proponen los medios digitales del Big Data y de la Inteligencia Artificial?”

Los resultados exploratorios evidenciaron que la Ciencia de Datos es un tema que suele tratarse de manera extensa en los medios analizados ($M=2.27$; $DT=,88$), además de hacerlo en conexión con todas la temáticas y aspectos de los diarios digitales.

Por otra parte, la mayoría de las noticias trataban temas de economía y eran emitidas por medios digitales generalistas.

Esto nos daría un primer acercamiento a los dos resultados más importantes de la investigación.



RESULTADOS

“¿Cuál es el tratamiento que proponen los medios digitales del Big Data y de la Inteligencia Artificial?”

Para responder a la pregunta se hizo una regresión lineal múltiple, donde la variable Visión actuaría como variable independiente y el resto actuarían como variables independientes.

Este análisis permitió observar que los atributos que los medios digitales utilizan para referirse a la Ciencia de Datos tienden a ser positivos cuando las noticias están ligadas a la economía, mientras que es más frecuente que sean negativos cuando están ligadas a la vida privada de los individuos.



RESULTADOS

“¿En que difiere el tratamiento del Big Data al de la Inteligencia Artificial?”

Aunque los atributos afectivos no mostraron diferencias entre el Big Data y la Inteligencia Artificial, sí se observó que el Big Data es tratado en mayor medida por medio de géneros interpretativos, jugando un papel más importante para los medios digitales, debido a su cantidad de noticias y a la extensión de éstas, y siendo un tema que no se trata o relaciona bien con aspectos de la vida cotidiana o esfera de los individuos.

Por su parte la Inteligencia Artificial goza de menos presencia en la agenda mediática y su tratamiento se hace por medio de géneros y encuadres informativos, más vinculados a la ciencia, la salud y la tecnología.



03

44%

de contenidos se publican en medios generalistas, frente a un 13% en medios especializados en ciencia y tecnología.

53.3%

La fuente más citada son expertos profesionales y empresas con un 53.3%

03

54.8%

*de los contenidos relacionan el
BD y la IA con “la profesión o
el ámbito laboral”*

72.4%

*de los textos se clasifican con un
nivel de comprensibilidad
bueno*

03



41%

Hay más volumen de contenidos publicados sobre **Big Data** que sobre **Inteligencia Artificial**.

de los textos hablan o están relacionados con **temas económicos**

03**COMPRENSIBLE****NO
COMPRENSIBLE**

Los textos informativos comprensibles:

Están escritos en tercera persona.

Visión objetiva.

No utilizan el humor.

Perspectiva optimista.

Destacan los beneficios.

Los textos informativos no comprensibles:

Presentan largas frases

Usan terminología técnica.

No aportan explicación de los tecnicismos.

No usan descripciones explicativas.

Van dirigidos a personas con conocimientos previos.



Proyecto financiado por la Fundación Española para la Ciencia y la Tecnología (FECYT) en la Convocatoria de ayudas para el fomento de la cultura científica, tecnológica y de la innovación 2019-2020. [FCT-18-13437]

OBSERVATORIO DE LOS CONTENIDOS AUDIOVISUALES
UNIVERSIDAD DE SALAMANCA

Facultad de Ciencias Sociales. Campus Miguel de Unamuno. Edificio F.E.S. Paseo de Francisco Tomás y Valiente, s/n. 37007. Salamanca. España
www.ocal.es/datasciencespain

Enero 2021. This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. Diseño base e imágenes de Canva.



OBSERVATORIO DE LOS CONTENIDOS AUDIOVISUALES
UNIVERSIDAD DE SALAMANCA

Facultad de Ciencias Sociales. Campus Miguel de
Unamuno. Edificio F.E.S. Paseo de Francisco Tomás y
Valiente, s/n. 37007. Salamanca. España
www.ocausal.es/datasciencespain

Enero 2021. This work is licensed under the Creative
Commons Attribution-NonCommercial-ShareAlike 4.0
International License. Diseño base e imágenes de Canva.



Proyecto financiado por la Fundación Española para la Ciencia y la Tecnología (FECYT) en la Convocatoria de ayudas para el fomento de la cultura científica, tecnológica y de la innovación 2019-2020. [FCT-18-13437]



UNIVERSIDAD
DE SALAMANCA