



# La Participación de la Audiencia en los Metadatos de TikTok

## Audience Participation in Tiktok's Metadata

Dra. Amparo Huertas-Bailén. Profesora titular, Universidad Autónoma de Barcelona (España)

(amparo.huertas@uab.cat) (<https://orcid.org/0000-0002-8851-5417>)

Dra. Natalia Quintas-Froufe. Profesora titular, Universidade da Coruña (España) (n.quintas.froufe@udc.es)

(<https://orcid.org/0000-0001-7597-6516>)

Dra. Ana González-Neira. Profesora titular, Universidade da Coruña (España) (ana.gneira@udc.es)

(<https://orcid.org/0000-0002-6369-0323>)

### RESUMEN

Con la expansión de la cultura digital, es necesaria una reflexión profunda sobre cómo investigar las audiencias. Si antes el individuo era ubicado en una categoría social determinante de los gustos culturales, ahora la tecnología identifica pautas de comportamiento a partir del registro directo de sus acciones. Este texto explora el tipo de conocimiento que puede obtenerse sobre la participación de la audiencia en TikTok. Se propone una metodología que consiste en el análisis de los metadatos de uso visibles. El trabajo de campo se ha centrado en «A2ality», cuenta de información con 4,4 millones de seguidores en España. Se han analizado todos los vídeos compartidos durante seis semanas del primer trimestre del año 2023 (n=173). La finalidad era conocer (a) el grado de correspondencia lineal entre los metadatos de una misma pieza y (b) la existencia de correlaciones entre metadatos y tipo de vídeo/contenido. Para cada metadato disponible en abierto (comentarios, «me gusta», guardados, compartidos y reproducción), se han establecido cuatro niveles de actividad (bajo, medio, alto y muy alto). La tendencia mayoritaria señala que los niveles obtenidos por los metadatos de un mismo contenido no son coincidentes, es decir, un vídeo tendrá más o menos alcance según el metadato observado. La homogeneidad de los vídeos hace que solo se detecten correlaciones claras entre tema y metadatos. Los temas con menor presencia pueden alcanzar niveles de actividad altos.

### ABSTRACT

With the expansion of the digital culture, deep reflection on how to research audiences is necessary. If before the individual was placed in a social category determining cultural tastes, now technology identifies patterns of behavior from the direct record of their actions. This text explores the type of knowledge that can be gained on audience participation in TikTok. We propose a methodology that consist in the analyse of usage metadata. Fieldwork has focused on «A2ality», an information account with 4.4 million followers in Spain. We analyzed all videos shared during six weeks of the first quarter of 2023 (n=173). The purpose was to know (a) the degree of the linear correlation between the metadata of the same piece and (b) the existence of correlations between metadata and type of video/content. For each metadata available in open (comments, likes, saves, shares and very high), four activity levels have been established (low, medium, high and top). The majority trend indicates that the levels obtained by the metadata of the same content are not coincident, that is, a video will have more or less scope according to the observed metadata. The homogeneity of the videos has caused that only clear correlations between topic and metadata are detected. Topics with less presence can reach high levels of activity.

### PALABRAS CLAVE | KEYWORDS

Audiencia, Metadatos, Participación, TikTok, Redes Sociales, Vídeo.

Audience, Metadata, Participation, TikTok, Social Media, Video.

## 1. Introducción y Estado de la Cuestión

Los continuos cambios en los modos de producción y distribución de los contenidos mediáticos obligan a redefinir constantemente cuál es la mejor manera de investigar las audiencias. Si nos centramos en el área de la medición, observamos que el valor concedido a los tradicionales estudios sobre muestras poblacionales representativas se está desdibujando. Por un lado, la creciente hiperconexión global ha ido debilitando la utilidad de trabajar con universos poblacionales cuya representatividad estadística se delimita a territorios marcados por fronteras físicas y, por otro, la extrema personalización del consumo digital ha propiciado un aumento de consumos minoritarios cuyo registro queda fuera del alcance de los grandes sondeos (García-Orosa et al., 2023; Huertas, 2002, 2015; Livingstone, 2019; Wolton, 1999). Con el desarrollo de la tecnología digital, se han reavivado viejas polémicas: la utilidad de las clasificaciones sociodemográficas para entender los consumos mediáticos vuelve a ponerse en duda ya que los rastros digitales se perfilan como la mejor fuente de información sobre el comportamiento de las audiencias (Andersen, 2020; Boyd, 2020; Fisher & Mehozay, 2019; Fulgoni & Lipsman, 2017; Hallinan & Striphas, 2016; Huertas, 2015; Jensen, 2019; Livingstone, 2019; Reder, 2012; Webster, 2016). Nunca han estado exentas de críticas, a pesar de que los criterios de clasificación se han ido ajustando en cada momento de acuerdo con los cambios socio-económicos. Recordemos, por su relevancia, la discusión entre quienes defienden la noción de «clases sociales», definidas como el «superhéroe teórico sin discusión» (Hartley, 2000: 37), y quienes creen más pertinente hablar de «estilos de vida» (Huertas, 2002). Pero, ahora, las variables sociodemográficas se cuestionan en su totalidad. Hay quienes consideran que la homogeneidad de los grupos sociales tradicionales era solo un espejismo (Jones, 2023). Ante este panorama, es ineludible pensar y experimentar nuevas maneras de investigar las audiencias.

En el entorno de la cultura digital, cada vez tiene más fuerza la idea de que un mismo contenido puede atraer a personas diversas y que, por lo tanto, basta con localizar las «pautas de afinidad». La epistemología algorítmica rechaza la adscripción social y clasifica al individuo en categorías predictivas de consumo (Fisher & Mehozay, 2019). Los algoritmos actúan a partir del registro de las interacciones de los usuarios, a quienes siguen y cuyas decisiones calculan (Fisher, 2015; García-Marín & Serrano-Contreras, 2023; Riemer & Peter, 2021). Los usuarios alimentan los algoritmos y eso es justamente lo que les hace funcionar (Andersen, 2020; Jensen, 2019). Esta mirada consolida la importancia de una relación empática entre el público y las personas creadoras de los contenidos que siguen, con quienes han de sentir que comparten gustos, intereses y deseos (Fisher & Mehozay, 2019) y, además, la posibilidad de que la audiencia pueda llegar a tomar consciencia de la continua monitorización a la que es sometida hace que podamos hablar de la actividad digital como un acto político, en el que es posible llegar a crear «audiencias imaginadas algorítmicamente», o contrapúblicos, a partir del control de los algoritmos (Bucher, 2017; Fisher & Mehozay, 2019; Warner, 2012).

La tecnología digital propicia procesos de auto-segmentación y esto genera un conocimiento distinto sobre la audiencia, que tampoco hay que considerar que sea más exacto que el proporcionado por las muestras representativas (Fisher & Mehozay, 2019). Rifkin (2000) pronosticó que acabaríamos hablando de tecnologías de relación en lugar de tecnologías de la información y la teoría del omnivorismo cultural, basada en el rechazo de la clase social como determinante de los gustos culturales, también puede ser retomada para entender este nuevo tipo de consumo mediático (Fernández Rodríguez & Heikkilä, 2011).

Una pregunta teórica clave es si podemos llegar a conocer el funcionamiento de los algoritmos. Andersen (2020) afirma que, al igual que sucede con muchos otros artefactos culturales, resulta imposible tener acceso directo al sentido de los algoritmos. Este autor explica que, aun en el caso de poder llegar a conocer su diseño, no podría llegarse a entender su funcionamiento, pues los algoritmos se caracterizan por ser mutantes, van variando como efecto de la constante información que genera la continua interacción de los usuarios. Conscientes, entonces, de la dificultad que supone asumir el reto de conocer la dinámica de los algoritmos, el objetivo de este texto es realizar una primera exploración a partir del análisis de los metadatos de uso visibles. Se parte de la tipología elaborada por Fisher (2015), quien distinguió entre datos públicos (contribuciones de los usuarios directas y visibles en abierto), privados (información compartida confidencialmente) y metadatos (información que generan los algoritmos automáticamente). Pero, además, también se tiene en cuenta el concepto metadato de uso. Desarrollado en el área de la archivística, este se refiere a la información sobre el acceso y consulta de documentos (Méndez Rodríguez, 2003).

No obstante, hay que resaltar que el estudio de los algoritmos ya tiene un recorrido. En términos generales, la opacidad técnica, la falta de transparencia del proceso y la presencia de sesgos éticamente cuestionables son

los aspectos que más atención están recibiendo (Boyd, 2020; Noble, 2018; Pasquale, 2015; Thomas et al., 2018). En cambio, cuestiones como la parcialidad y la descontextualización de la información que aportan, apuntadas por Livingstone (2019), apenas están siendo atendidas. Se distinguen dos líneas de investigación: una ahonda en cómo los algoritmos se construyen, entrenan y ajustan (Dourish, 2016; Flyverbom & Murray, 2018) y la otra pone el foco en los usuarios. En el marco de esta segunda vía, aunque despuntan los estudios basados en los datos generados por los propios algoritmos (Andersen, 2020; Bucher, 2017), también aparecen trabajos empíricos sobre muestras poblacionales. Un ejemplo de estos últimos es la propuesta de Lomborg and Kapsch (2020), quienes analizan la experiencia algorítmica (conocimiento, sentimientos que genera y evaluación). Estos autores acaban negando que los algoritmos sepan más de los usuarios de lo que ellos saben sobre sí mismos.

La pregunta concreta que nos guía a la hora de asumir nuestro reto es qué tipo de información puede aportarnos el análisis de los metadatos y, más concretamente, los metadatos de uso. Aunque se trata de una cuestión cuya respuesta precisa de un trabajo empírico y reflexivo continuado, nos proponemos iniciar ese camino y, para ello, se ha escogido la cuenta de «A2ality», disponible en TikTok, como estudio de caso. Diversos estudios científicos se han centrado en el estudio de las redes sociales, incluida TikTok (Estrada-García & Gamir-Ríos, 2023; García-Marín & Salvat-Martinrey, 2022; Pérez-Escolar et al., 2023; Vázquez-Herrero et al., 2021, entre otras), pero la mayoría se centran en la producción y la oferta, y no abordan en profundidad qué maneras de estudiar las audiencias se están configurando.

TikTok, que apareció en 2017, es la versión internacional de una red social de origen chino basada exclusivamente en la difusión de bailes y música, lanzada el año anterior (Abidin, 2021). En TikTok los usuarios comparten vídeos cortos, a los que se pueden añadir fondos animados, efectos de sonido y diversos filtros visuales. Al inicio, su duración era inferior a un minuto y, desde 2022, pueden alcanzar los 10 minutos. Su reproducción es en bucle, pero también se pueden guardar. Las interacciones posibles son en forma de comentarios, compartir y «me gusta». En 2020, TikTok anunció que su sistema de recomendación estaba altamente individualizado a partir del conocimiento de las interacciones de los usuarios, de los vídeos que se siguen y de los datos sobre la conexión (Jones, 2023). Su diseño está pensado para atrapar a los usuarios (Lovink, 2019). TikTok es una de las redes preferidas por la juventud, con marcada penetración en Latinoamérica (Martínez-Estrella et al., 2023). En España, uno de los países con mayor penetración de esta red (Newman, 2022), el nivel de afinidad más alto se concentra en la franja de edad de los 18 a los 24 años, en la que el promedio diario de visitas supera las 1.500; se sitúa así en niveles próximos a Youtube e Instagram, muy por encima de Twitter o Twitch, donde la media de visitas diarias ronda las 500 en esta franja de edad (Comscore, 2023). Los datos de IAB (2023) también reflejan el éxito de TikTok en España: es la red que presenta mayor incremento en los últimos tres años (las visualizaciones han aumentado un +109% en un año), se encuentra entre las cinco preferidas y, tras Instagram, es la que genera más interacción.

El motivo que nos llevó a elegir TikTok fue esa mayor precisión en las recomendaciones individualizadas, además de su éxito en España. Pero era necesario acotar los elementos analizables y, para ello, optamos por «A2ality», compañía española fundada el año 2020. En esta elección se ha tenido en cuenta su éxito y el hecho de que esta se adentra en el terreno informativo, pues este artículo se enmarca en un proyecto de investigación sobre el seguimiento de la actualidad política por parte de la juventud en España. El principal contenido del canal de «A2ality» en TikTok, dirigido al público joven, lo constituyen vídeos que resumen una selección de noticias publicadas en medios de información ajenos. Comscore (2023) sitúa su número de seguidores en España en cerca de 4,4 millones y Atresmedia acaba de firmar un acuerdo con la compañía para entrar a formar parte de su accionariado (Martínez, 2023). El Instituto Reuters la ha clasificado como «medio informativo» y la ubica como la tercera cuenta informativa de TikTok con mayor media de visionado por vídeo del mundo, tras «NBC News» y «G1-Globo» (Newman, 2022).

Nuestro objetivo radica en comprobar el tipo de conocimiento que puede obtenerse acerca de la participación de la audiencia de TikTok a partir de una propuesta metodológica que consiste en el análisis de los metadatos de uso disponibles en abierto. Partimos de dos hipótesis: 1) aparece correlación lineal entre todos los metadatos, es decir, todos los metadatos de una misma pieza se alinean en niveles próximos, ya sea en la gama de valores altos o bajos; 2) los niveles de participación reflejados por los metadatos dependen del contenido del material visionado.

## 2. Material y Métodos

El corpus reunido para el análisis de los metadatos de «A2ality» está formado por todos los vídeos

subidos a su cuenta en TikTok durante 42 días del primer trimestre del año 2023 (n=173). Entre el 3 y el 15 de abril de 2023, se extrajeron los vídeos publicados en enero (durante la tercera y cuarta semanas), febrero (la primera y segunda semanas) y marzo (la segunda y tercera semanas). Dado que «A2ality» centra su contenido en la actualidad informativa, fueron escogidas estas fechas para evitar en la medida de lo posible momentos excepcionales, como períodos festivos u organización de eventos mediáticos.

Para el diseño de la ficha de análisis (definición de variables y categorías de análisis) se realizó una primera observación, de carácter exploratorio, sobre una muestra de veinticinco vídeos. El material fue visionado por las tres autoras de este documento, lo que ayudó a trabajar en una ficha consensuada y definir los criterios de clasificación para garantizar la objetividad en la recogida de información del corpus definitivo. La ficha de análisis final se estructuró de este modo:

- Datos identificativos: asignación de un número identificativo y registro de la fecha y día de la semana de publicación, además de la duración total expresada en segundos.
- Formato. Se definieron cuatro categorías: resumen (vídeo que reúne cinco titulares de la actualidad), noticia (vídeo sobre un hecho noticioso), publicidad (vídeo publicitario) y clip internet (vídeo extraído de Internet con edición ajena), además de otros (vídeos no admitidos en las categorías anteriores).
- Presencia del equipo responsable. Se observó el modo en que las personas identificadas como parte del equipo de «A2ality» aparecían en el vídeo, lo que nos permitió definir cuatro categorías: al inicio, al inicio y al cierre, de forma constante (en más del 80% del vídeo o al completo) y ausencia.
- Uso de «hashtags». Se registró el número empleado en cada pieza y, para tener recogida la lista de hashtags y saber la periodicidad individual de aparición, también se copió literalmente cada uno de ellos.
- El uso de menciones a otros perfiles. Se registró el número empleado en cada pieza y también se copió literalmente cada una de las menciones.
- Registro de metadatos de uso: (número visible de) «me gusta», comentarios, reproducciones, veces que fue compartido y veces que fue guardado.

Tabla 1: Características del corpus de análisis (n=173).

Variable	Categoría	N	% (sobre total corpus)
Formato	Noticia	110	63,6
	Resumen	30	17,4
	Clips	21	12,1
	Publicidad	5	2,9
	Otros	7	4
Presencia equipo responsable	Inicio	79	45,7
	Inicio y cierre	20	11,5
	Constante	42	24,3
	Ausentes	32	18,5
«Hashtags»	Sí	136	78,6
	No	37	21,4
Menciones a perfiles en redes	Sí	10	5,7
	No	163	94,3
Día de emisión	Días laborables	159	92
	Fin de semana	14	8
Duración	< 20"	23	13,3
	Entre 21" i 1'	142	82
	> 1'	8	4,7

Además, para abordar con mayor profundidad los aspectos específicos de los vídeos con formato noticia, sobre esta porción de la muestra se recogieron dos elementos más (tabla 2):

- Tema. A partir del visionado de cada vídeo y de la lectura de los textos que los acompañan, especialmente de los hashtags empleados para identificar el tema, se definieron nueve categorías temáticas: internacional; vida privada de personajes públicos; ciencia y salud; tecnología; política (España); violencia de género; economía, deporte y cultura. No fue necesario crear la categoría otros, pues todas las piezas de la muestra pudieron ser clasificadas.

- Medio de información citado como origen de la noticia. Se registró el nombre de las cabeceras y cadenas mencionadas, lo que permitió, además de tener una lista completa de los medios referenciados, establecer cuatro categorías: prensa, televisión, medios digitales y radio, además de otros (donde aparecen, por ejemplo, «newsletters»).

**Tabla 2: Características de las noticias que forman parte del corpus de análisis (n=110).**

Variable	Categoría	N (citaciones)	% (sobre total noticias)
Tema	Internacional	32	29
	Vida privada personajes públicos	21	19
	Ciencia y Salud	12	11
	Tecnología	12	11
	Política (España)	11	10
	Violencia de género	7	6,3
	Economía	6	5,5
	Deporte	5	4,5
	Cultura	4	3,7
Variable	Categoría	N (citaciones)	% (sobre total citaciones)
Tipos de medios enunciados como origen de la información*	Prensa	52	49,5
	Televisión	31	29,5
	Medios digitales	19	18
	Radio	1	1
	Otros	2	2

Notas: \*En la cuantificación de las citaciones no se han eliminado duplicaciones de un mismo medio, pues interesaba conocer el peso total según categoría; 27 noticias (24,5%) no indican ningún medio como origen de la información tratada.

Previamente, la ficha de análisis fue testada por dos de las investigadoras que firman este artículo y, a continuación, se procedió a la recogida de datos de forma manual. Cabe tener presente que TikTok no ofrecía acceso a una API académica en el momento en que se realizó el trabajo de campo (TikTok, 2022). Pero, sobre todo, es preciso hacer énfasis en que era necesario un visionado atento de todo el material para poder identificar todas las variables y establecer las categorías correspondientes. La información reunida fue sistematizada en un documento Excel para gestionar su análisis posterior.

En la planificación inicial estaba previsto realizar el análisis diferenciando entre los vídeos con más antigüedad y los más recientes, dado que el tiempo de disponibilidad de un contenido se considera determinante de su consumo. Inicialmente estaba prevista la clasificación del corpus según el mes de publicación (enero: 73 piezas, febrero: 41 piezas y marzo: 59 piezas). No obstante, una vez realizada esta clasificación, se observó que el resultado no brindaba submuestras con datos homogéneos acordes con la temporalidad. Para este proceso se habían establecido cinco niveles en cada metadato (bajo, medio, intermedio, alto y muy alto) y el caso más destacado, en cuanto a la escasa diferencia en los metadatos según tiempo de exposición, se dio en el número de comentarios. Aquí, las diferencias eran mínimas. Por ejemplo, el 39% de las piezas de enero, el 41% de las de febrero y el 42% de las de marzo se situaban en el nivel bajo.

Esto nos llevó a descartar la idea de subdividir el corpus según tiempo transcurrido desde su publicación y se pasó a trabajar con el corpus de forma conjunta. Ahora bien, ese proceso de clasificación y análisis según mes de publicación acabó siendo muy útil para afinar los rangos definitivos de los niveles de cada metadato.

**Tabla 3: Rangos de los niveles de la audiencia.**

Actividad	D	C	B	A
Comentarios	9-299	300-999	1000-4.999	+ 5000
«Me gusta»	5,720-49,999	50,000-149,999	150,000-399,999	+ 400.000
Guardados	97-999	1000-4,999	5000-19999	+ 20,000
Compartidos	11-499	500-1499	1500-5900	+ 6,000
Reproducción	45,000-299,000	300,000-699,999	700,000-1,999,999	+ 2 millones

En el diseño metodológico se planteó trabajar con cuatro niveles: bajo (D), medio (C), alto (B) y muy alto (A). Es decir, optamos por evitar un rango intermedio para poder detectar con mayor claridad tendencias en

sentido positivo o negativo. Además, era necesario definir cada rango de forma que quedara garantizado un número significativo de piezas en cada uno de ellos. Finalmente, para adaptar la metodología al caso analizado, pasamos a delimitar los valores iniciales y finales de cada rango a partir del análisis de la participación de la audiencia en «A2ality». En el nivel D, el valor inicial vino determinado por el metadato más bajo registrado.

### 3. Análisis y Resultados

Los resultados se han agrupado en dos bloques a partir de las hipótesis planteadas: 3.1) presencia de correlación lineal entre los metadatos que obtiene cada pieza y 3.2) presencia de correlación entre metadatos y tipo de vídeo/contenido.

#### 3.1. Presencia de Correlación Lineal Entre los Metadatos Que Obtiene Cada Pieza

En la tabla 4, donde se recoge la información del total del corpus, puede comprobarse la diversidad de porcentajes que aparecen dentro de cada nivel. Si los niveles recogidos por los metadatos de cada pieza de forma individual tendieran a comportarse de forma homogénea, ya fuera en el marco del nivel bajo, medio, alto o muy alto, los porcentajes de cada columna deberían ser similares. Al detectar que tanto el número de piezas como el porcentaje que alcanzan los niveles alto (B) y muy alto (A) en «compartidos» y «reproducción» son similares, se ha comprobado en qué medida esos metadatos se corresponden con las mismas piezas, pues esa información nos hace pensar que ahí sí se cumple con nuestra primera hipótesis. Sin embargo, el resultado nos indica que esa correlación lineal en esos dos metadatos es poco significativa: el número de piezas que coinciden en el nivel B en ambos metadatos es 25 (y no 46, como la tabla 4 nos puede hacer pensar) y las que llegan al nivel A en esos dos metadatos son un total de 8 (y no las 13 que pueden calificarse como potencialmente coincidentes en la tabla 4).

Si se observan los datos porcentuales de cada nivel, en el más bajo (D) aparece el 41% de la muestra en relación al metadato que cuantifica los comentarios; el 45,7%, en el referente a los «me gusta»; el 30%, en «guardados»; el 37,8%, en «compartidos» y el 34,3 %, en «reproducción». Es decir, la diferencia en puntos porcentuales dentro del nivel D de todos los metadatos entre el valor más alto y el valor más bajo es de 15,7, siendo el porcentaje más alto el de «me gusta» y el más bajo el de «guardados». En el caso del nivel medio (C), es donde se da la diferencia porcentual mayor, de 22,3 puntos: el valor más alto aparece en «guardados» (48,5%) y el más bajo en «compartidos» (26,2%). En el nivel alto (B), la diferencia es de 15,9 puntos porcentuales: el valor más elevado está en «compartidos» (28%) y el más bajo, en «me gusta» (12,1%). Sobre el nivel muy alto, la diferencia es de 5,1 puntos porcentuales. Esta es una diferencia menor, pero se ha de tener en cuenta que las cifras son mucho más bajas en este rango. Aquí, el valor más alto aparece en «reproducción» (8,2%) y el más bajo, en «comentarios» (2,9%). Sobre esta ausencia de correlación lineal de los metadatos según niveles de participación, llama también la atención comprobar que los índices más altos de cada nivel aparezcan en metadatos diferentes y lo mismo sucede con los valores más bajos. Es decir, tampoco se detecta la tendencia a que un mismo metadato destaque en uno u otro sentido.

Tabla 4: Distribución del corpus según nivel de la audiencia (n=173).

Actividad	D		C		B		A	
	N	%	N	%	N	%	N	%
Comentarios	71	41	62	35,8	35	20,2	5	2,9
«Me gusta»	79	45,7	66	38,1	21	12,1	7	4
Guardados	52	30	84	48,5	29	16,7	8	4,6
Compartidos*	62	37,8	43	26,2	46	28	13	8
Reproducción**	58	34,3	51	30,1	46	27,2	14	8,2

Notas: \*Solo se dispone de esta información en 164 vídeos del corpus (el % es sobre este total). \*\*Solo se dispone de esta información en 169 vídeos del corpus (el % es sobre este total).

Si se comprueba cuántas piezas sitúan todos los metadatos en el mismo nivel, aparecen 43 (25% de la muestra). La mayoría, 31 (18%), se queda en el nivel bajo, que es el más numeroso. Este grupo reúne piezas de todos los formatos: resúmenes (15), noticias (10), otros (3), publicidad (2) y clips (1). Cabe prestar atención a que aquí se sitúa el 50% de los resúmenes registrados en nuestro trabajo de campo, que es justamente el formato que «A2ality» más utiliza para promocionarse. El resto se reparte de la siguiente

manera: 8 (4,6%) colocan todos sus metadatos en el nivel C; 2 (1,1%), en el nivel B y otras 2 (1,1%), en el nivel A. Estas dos últimas son dos noticias (pieza 7: «Clara Chía sobre la canción de Shakira» y pieza 39: «El núcleo de la tierra se ha detenido»).

Finalmente, también se ha comprobado el grado de dispersión y la tendencia mayoritaria es a situarse en dos niveles contiguos. Así sucede con 91 piezas (52,6%): 44 (25,4%) presentan metadatos en D y C; 36 (20,8%), en C y B, y 11 (6,3%) en B y A. Solo 24 piezas aparecen en 3 niveles yuxtapuestos: 17 (9,8%) tienen metadatos en niveles D, C y B, y 7 (4%), en C, B y A. Únicamente la pieza 21 (clip de internet: «Brote de tiña en España») presenta la mayor dispersión posible (muy alto: «compartidos»; alto: «reproducción»; medio: «guardados» y «me gusta»; bajo: «comentarios»). Del mismo modo que solo la pieza 59 (noticia: «El reality show de Anna Sorokin») presenta metadatos en niveles no contiguos, en A («comentarios») y en C (el resto de metadatos).

### 3.2. Presencia de Correlación Entre Metadatos y Contenido

Pasamos ahora a analizar el comportamiento de los metadatos según el formato y la presencia en los vídeos de las personas identificadas como miembros del equipo de «A2ality». Analizadas el total de variables recogidas en la ficha de análisis, estas dos son las únicas que permiten trabajar con cierta profundidad. En general, el resto de variables se ha descartado por tener un efecto homogeneizador en el corpus (al destacar de forma muy pronunciada una de las categorías sobre el resto): «A2ality» concentró su actividad durante los días laborables (92%); respecto a la duración del corpus, aunque la pieza más corta dura 4" y la más larga 1'50" (ambas son clips de internet), el 82% se sitúa entre los 21" y los 60"; sobre los hashtags, el 79,6% del corpus incluye al menos uno y, aunque se ha podido registrar que en el 48% de los casos sirve para identificar el tema y se ha elaborado una lista de los más frecuentes (#explicamelofacil: en el 52,6% de las piezas; #aprendiendoconTikTok: 23,7%; #noticias1minuto2023: 22%; #NoticiasTikTok: 18% y #NoticiasFáciles: 8,6%), no se ha detectado ningún criterio en su elección, por lo que solo se ha podido establecer con rigor científico la categoría presencia/ausencia. La excepción a esta lógica homogeneizadora se da en las menciones a otros perfiles en «A2ality». Este recurso solo es empleado por el 5,7% del corpus.

Si se observan los niveles alcanzados por los metadatos de los tres formatos más habituales en el corpus, lo primero que llama la atención es que las noticias están presentes en todos los niveles, tendiendo a aparecer los porcentajes más altos en los rangos intermedios (C y B); excepto en el caso de «me gusta», donde estos aparecen en los niveles bajo (D) y medio (C). Esta presencia expandida en todos los niveles resulta lógica, dado que representa el 63,6% del corpus. A continuación, aparecen los clips de internet, que, excepto en «comentarios» (donde solo se alcanza niveles D, C y B), en el resto de metadatos logran ubicar piezas en todos los rangos. Pero, además, este formato es el que consigue colocar un mayor porcentaje de piezas en el nivel muy alto (A) cuando llega a él (por ejemplo, el 19% en el caso de «reproducción»). Por último, los metadatos de los resúmenes tienden a ubicar este contenido en los niveles más bajos: bajo (D) y medio (C). Solo en dos de los cinco metadatos («reproducción» y «compartidos») algunas de estas piezas alcanzan el nivel alto (B).

Tabla 5: Nivel de interacción con los vídeos (% en horizontal).

Categoría (total)	Comentarios				Me gusta			
	D	C	B	A	D	C	B	A
Según el formato								
Noticia (110)	24,6	45,4	25,4	4,6	36,4	41,8	17,2	4,6
Resumen (30)	90	10	0	0	63,4	36,6	0	0
Clips (21)	38	28,5	33,5	0	52,5	28,5	9,5	9,5
Publicidad (5)	100	0	0	0	100	0	0	0
Otros (7)	57,2	42,8	0	0	57,2	42,8	0	0
Según presencia equipo responsable								
Inicio (79)	52	33	14	1	50,7	36,7	10,1	2,5
Inicio y cierre (20)	30	25	35	10	30	50	20	0
Constante (42)	26,2	47,7	21,4	4,7	38	40,6	14,3	7,1
Ausentes (32)	40,7	34,3	25	0	53,2	31,3	9,3	6,2

La observación del alcance de los metadatos en función del modo en que aparecen las personas identificadas como miembros de «A2ality» no permite indicar de una forma clara que unos estimulen más la participación

que otros. Tan solo se puede decir que aparecer al inicio y al cierre tiende a generar un mejor resultado en los metadatos relacionados con el uso del vídeo, pero, aun así, estos quedan mayoritariamente en los niveles medio (C) y bajo (B). Eso sí, esta opción se perfila con algo más de potencial para generar comentarios: el 10% llega al nivel muy alto, donde solo aparece el 4,7% de las piezas con presencia constante y el 1% de las que la presencia se concentra al inicio. Ahora bien, estas dos últimas opciones sí logran colocar alguna pieza en el nivel más alto en todos los metadatos. Desde el punto de vista de la actividad de la audiencia, la opción más desaconsejable sería la ausencia, sin embargo, la reacción también es dispar según el metadato que se observe. En la interacción con la audiencia los valores dominan en el nivel más bajo («comentarios»: 40,7%; «me gusta»: 53,2%). Pero, en cambio, en relación al uso del vídeo, el metadato de las reproducciones de material donde no aparece personal de «A2ality» sí logra destacar en el nivel más alto (12,5%).

Tabla 6: Nivel de uso de los vídeos (% en horizontal).

Categoría (total)	Guardados				Compartidos				Reproducción			
	D	C	B	A	D	C	B	A	D	C	B	A
<b>Según el formato</b>												
Noticia (110)	21	51	22,6	5,4	28,1	29	31,9	9	27,2	30	31,9	9
Resumen (30)	63,4	36,6	0	0	73,4	13,4	10	0	53	33,5	13,5	0
Clips (21)	23,8	52,5	14,2	9,5	9,5	23,8	23,8	14,2	23,8	23,8	28,5	19
Publicidad (5)	40	40	20	0	40	40	20	0	80	20	0	0
Otros (7)	42,8	57,2	0	0	71,5	0	28,5	0	57,2	28,5	14,3	0
<b>Según presencia equipo responsable</b>												
Inicio (79)	36,7	48,1	11,4	3,8	46,8	27,8	16,5	6,3	39,2	29,1	25,3	5
Inicio y cierre (20)	5	65	30	0	15	25	55	5	15	35	45	5
Constante (42)	28,4	40,6	23,9	7,1	31	26,1	31	9,5	35,6	31	21,4	12
Ausentes (32)	31,3	50	12,5	6,2	28,1	15,6	28,1	9,3	28,1	25	25	12,5

Sobre las noticias, se han detectado correlaciones interesantes entre metadatos y tema. En cambio, la aplicación de la variable de los medios citados como fuente de origen no permitió un análisis en profundidad. La diversidad de los medios mencionados hizo que dominaran las apariciones puntuales, lo que impidió establecer correlaciones claras.

Las noticias que tratan temas internacionales, que son aquellas con mayor peso en el conjunto de noticias (32,29%), tienden a situar sus metadatos en el nivel medio (C): aparecen 18 en «comentarios» y «me gusta» respectivamente y 21, en «guardados». Como hechos diferenciales, cabe destacar que 10 sí llegan al nivel alto (B) en «reproducción» y que 1 alcanza el muy alto (A) en «comentarios».

Los metadatos de las noticias sobre la vida privada de personajes públicos, la segunda temática en volumen (21,19%), están presentes en todos los niveles y, además, de una forma bastante equilibrada. En comentarios y «me gusta» aparece un número similar (entre 5 y 8) en D, C y B. En «guardados», la mayor proporción (10) se da en el nivel medio (C); en «compartidos», aparecen 12 en el nivel bajo (D) y, en reproducción, 9 en nivel alto (B). Pero, en todos los metadatos, siempre hay piezas que logran el nivel más alto: 1, (en «guardados»); 2, (en «comentarios», «me gusta» y «compartidos», respectivamente y 3, en «reproducción».

Los metadatos de las noticias sobre ciencia y salud, a pesar de suponer tan solo el 11%, con 12 piezas, también tienden a estar presentes en todos los niveles, pero, en comparación con las noticias sobre la vida privada de personajes públicos, suelen ubicarse en rangos más altos, entre C y A. En cuanto a la interacción, el mayor número de piezas se acumula en el nivel medio (5: «comentarios» y 6: «me gusta»). En los metadatos sobre el uso del vídeo, el mayor número de piezas aparece en el rango alto (5: «guardados»; 4: «compartidos» y 6: «reproducción»). Pero, quizá, lo más llamativo es que 5 piezas alcanzan el nivel más alto en «compartidos»; 2 lo logran en 3 metadatos («me gusta», «guardados» y «reproducción») y 1 lo consigue en «comentarios».

Las noticias sobre tecnología, con el mismo peso en el corpus que la temática ciencia y salud, también tienden a aparecer en todos los niveles de todos los metadatos. Sin embargo, aunque también llegan al nivel más alto (1, en «me gusta»; 3, en «guardados» y 2, en «compartidos» y «reproducción»), la tendencia es a quedarse en los niveles más bajos (D y C).

Los metadatos de las 10 noticias sobre política (España), que suponen el 10% de las noticias, tienden a concentrarse en el nivel medio (C). Aquí aparecen entre 5 y 7 piezas en los metadatos «comentarios», «me gusta», «guardados» y «reproducción». Solo una pieza alcanza el nivel más alto en un metadato («compartidos»).

El resto de temáticas tienen un escaso peso en el corpus. De su análisis, se puede destacar que solo llega al nivel muy alto (A) una pieza sobre deportes (en «comentarios») y 2 de cultura (en «reproducción»), que son justamente las temáticas con menor peso en el conjunto de noticias (4,5 % y 3,7%, respectivamente). Sobre el resto, los metadatos de las noticias en torno a agresiones de violencia de género tienden a situarse en los niveles intermedios (C y B) y, en el caso de economía, en el más bajo (D).

#### 4. Discusión y Conclusiones

Aunque es imposible acceder de forma directa al sentido de los algoritmos (Andersen, 2020), consideramos que nuestra apuesta metodológica nos ha sido muy útil. Nos ha permitido obtener información de gran interés para reflexionar sobre el modo en que las audiencias interactúan con contenidos elaborados por personas que se presentan como semejantes (Fisher & Mehozay, 2019; Rifkin, 2000). En una primera fase del trabajo de campo, al observar que el tiempo de permanencia en abierto de un contenido no influye en los metadatos, ya constatamos que el consumo de «A2ality» en TikTok es efímero. Posteriormente, durante el desarrollo del trabajo de campo, el análisis a partir de la definición «ad hoc» de rangos para determinar los niveles de participación registrados por cada metadato de uso nos ha sido muy enriquecedor.

A partir de la refutación de la primera hipótesis, queda claro que no es una cuestión baladí el metadato que se escoge como parámetro para medir el éxito en TikTok, dado que los metadatos de uso de un mismo contenido no tienden a coincidir en el mismo nivel de actividad (bajo, medio, alto y muy alto). La correlación lineal solo se ha detectado en el 25% de la muestra y se presenta mayoritariamente en niveles bajos, que son aquellos que más piezas reúnen. Por tanto, a la hora de analizar el alcance de un contenido, debe tenerse muy en cuenta el metadato que se observa y, en su interpretación, no dar por hecho que se da una correlación lineal, especialmente cuando el metadato que se maneja indica alta actividad. En segundo lugar, a partir del trabajo de comprobación de la segunda hipótesis, hemos detectado fenómenos cuyo conocimiento es muy útil para el diseño de contenidos en red. El análisis del caso de «A2ality» nos ha permitido observar, por un lado, que los contenidos con más presencia no siempre consiguen los mejores niveles de participación, así como tampoco aquellos contenidos que el canal emplea para promocionarse identificándolos como distintivos o únicos –nos referimos aquí a los resúmenes de noticias– y, por otro lado, que la ausencia de personas en los vídeos que representen el canal no impide obtener buenos resultados. De este modo, y para ilustrar esta reflexión, asuntos apenas tratados en «A2ality» pueden alcanzar niveles altos de actividad, como se constató con las noticias sobre ciencia y salud, o, aunque la ausencia de la figura del presentador sí parece reducir la interacción, esto no impide lograr niveles altos en el metadato de reproducción.

Creemos que este trabajo puede ser de gran ayuda para abordar el estudio de las audiencias con el nuevo enfoque que el actual contexto socio-tecnológico digital exige. El crecimiento de la actividad de la audiencia en las redes está provocando una ampliación de la gama disponible de metadatos de uso y, por tanto, cabría pensar en la utilidad de llegar a un consenso sobre cuál emplear para determinar el posicionamiento alcanzado por un contenido. También creemos interesante indagar sobre la pertinencia de distinguir entre audiencias y seguidores. Esta idea nos surge después de comprobar que es más fácil conseguir altos niveles de participación en aquellos metadatos que miden el uso del vídeo (guardar, compartir y reproducir) que en aquellos que registran la interacción directa con el mismo (comentar y «me gusta»). Espectadores serían aquellos que solo utilizan el vídeo sin interactuar con él, mientras que los seguidores serían aquellos que, además, se implican dialogando con el contenido y señalando su aceptación. Por tanto, es lógico que la mayor actividad se dé en los metadatos de uso del vídeo, donde coinciden ambos perfiles.

No obstante, la realización de este estudio también nos ha generado muchas preguntas. Por limitación de espacio, plantearé aquí solo algunas de ellas. En primer lugar, ¿de qué manera los algoritmos pueden influir a la hora de (a) potenciar un tipo de participación sobre otro; (b) un contenido sobre otro o (c) un formato sobre otro? Y, en segundo lugar, ¿es posible que los algoritmos estimulen la participación sobre contenidos y estrategias cuyo éxito ya viene marcado por los medios tradicionales? La aparición de noticias sobre la vida privada de personas o la elaboración de textos sensacionalistas para llamar la atención son contenidos y estrategias que ya funcionan fuera de las redes y que, al menos en el caso de «A2ality», también parecen dar buen resultado.

Ahora bien, para contrarrestar los problemas que Livingstone (2019) ya apuntó de forma muy acertada, la parcialidad y la descontextualización de gran parte de los estudios disponibles, consideramos que es imprescindible complementar este tipo de trabajos con muestras poblacionales representativas. Por ejemplo,

los resultados y conclusiones obtenidos aquí son un buen punto de partida para el desarrollo de sondeos específicos sobre muestras de usuarios de «A2ality» en TikTok.

### Contribución de Autores

Idea, A.H., N.Q., A.G.; Revisión de literatura (estado del arte), A.H.; Metodología, A.H., N.Q., A.G.; Análisis de datos, A.H., N.Q., A.G.; Resultados, A.H.; Discusión y conclusiones, A.H., N.Q., A.G.; Redacción (borrador original), A.H.; Revisiones finales, N.Q., A.G.; Diseño del Proyecto, A.H., N.Q., A.G.

### Apoyos

Este estudio es una acción fruto del proyecto «Polarización social e interculturalidad: el seguimiento de la actualidad política por parte de la juventud migrante y autóctona desde una mirada interseccional (POINTAP)» (PID2021-125032OB-I00), aprobado en la convocatoria 2021 «PROYECTOS DE GENERACIÓN DE CONOCIMIENTO», en el marco del Programa Estatal para impulsar la Investigación Científico-Técnica y su Transferencia, Plan Estatal de Investigación Científica, Técnica y de Innovación 2021-2023.

### Referencias

- Abidin, C. (2021). Mapping internet celebrity on TikTok: Exploring attention economies and visibility labours. *Cultural Science Journal*, 12(1), 77-103. <https://doi.org/10.5334/csci.140>
- Andersen, J. (2020). Understanding and Interpreting Algorithms: Toward a Hermeneutics of Algorithms. *Media, Culture & Society*, 42(7-8), 1479-1494. <https://doi.org/10.1177/0163443720919373>
- Boyd, D. (2020). Questioning the legitimacy of data. *Information Services & Use*, 40(3), 259-272. <https://doi.org/10.3233/ISU-200098>
- Bucher, T. (2017). The Algorithmic Imaginary: Exploring the Ordinary Affects of Facebook Algorithms. *Information, Communication & Society*, 20(1), 30-44. <https://doi.org/10.1080/1369118X.2016.1154086>
- Comscore. (2023, March 15). *Comscore Insights: nuevos tiempos, nuevos consumos, perspectiva 2023*. <https://bit.ly/43wDsuZ>
- Dourish, P. (2016). Algorithms and Their Others: Algorithmic Culture in Context. *Big Data & Society*, 3(2), 1-11. <https://doi.org/10.1177/2053951716665128>
- Estrada-García, S., & Gamir-Ríos, J. (2023). Soft News in Original Videos. Adaptation to Tiktok of the Main Spanish Online Media. *Profesional De La Información*, 32(2), 320222. <https://doi.org/10.3145/epi.2023.mar.22>
- Fernández Rodríguez, C. J., & Heikkilä, R. (2011). El Debate Sobre El Omnivorismo Cultural. Una Aproximación a Nuevas Tendencias en Sociología Del Consumo. *Revista Internacional De Sociología*, 69(3), 585-606. <https://doi.org/10.3989/ris.2010.04.15>
- Fisher, E. (2015). 'You Media': Audiencing as Marketing in Social Media. *Media, Culture & Society*, 37(1), 50-67. <https://doi.org/10.1177/0163443714549088>
- Fisher, E., & Mehozay, Y. (2019). How Algorithms See Their Audience: Media Epistemes and the Changing Conception of the Individual. *Media, Culture & Society*, 41(8), 1176-1191. <https://doi.org/10.1177/0163443719831598>
- Flyverbom, M., & Murray, J. (2018). Datastructuring—organizing and Curating Digital Traces Into Action. *Big Data & Society*, 5(2), 1-12. <https://doi.org/10.1177/2053951718799114>
- Fulgoni, G. M., & Lipsman, A. (2017). Measuring Television in the Programmatic Age: Why Television Measurement Methods Are Shifting Toward Digital. *Journal of Advertising Research*, 57(1), 10-14. <https://doi.org/10.2501/JAR-2017-009>
- García-Marín, D., & Salvat-Martinrey, G. (2022). Viralizar la verdad. Factores predictivos del engagement en el contenido verificado en TikTok. *Profesional De La Información*, 31(2), 310210. <https://doi.org/10.3145/epi.2022.mar.10>
- García-Marín, J., & Serrano-Contreras, I.-J. (2023). (Un) Founded Fear Towards the Algorithm: Youtube Recommendations and Polarisation. *Comunicar*, 31(74), 61-70. <https://doi.org/10.3916/C74-2023-05>
- García-Orosa, B., Canavilhas, J., & Vázquez-Herrero, J. (2023). Algorithms and Communication: a Systematized Literature Review. *Comunicar*, 31(74), 9-21. <https://doi.org/10.3916/C74-2023-01>
- Hallinan, B., & Striphias, T. (2016). Recommended for You: the Netflix Prize and the Production of Algorithmic Culture. *New Media & Society*, 18(1), 117-137. <https://doi.org/10.1177/1461444814538646>
- Hartley, J. (2000). *Los Usos De La Televisión*. Paidós. <https://go.revistacomunicar.com/vccDVk>
- Huertas, A. (2002). *La Audiencia Investigada*. Gedisa. <https://go.revistacomunicar.com/2leYgE>
- Huertas, A. (2015). *Yo Soy Audiencia. Ciudadanía, Público Y Mercado*. UOC Press. <https://go.revistacomunicar.com/dJwgaX>
- IAB. (2023, May 10). *Estudio De Redes Sociales*. IAB Spain. <https://go.revistacomunicar.com/ILOSmV>
- Jensen, K. B. (2019). The Double Hermeneutics of Audience Research. *Television & New media*, 20(2), 142-154. <https://doi.org/10.1177/1527476418811103>
- Jones, C. (2023). How to train your algorithm: the struggle for public control over private audience commodities on Tiktok. *Media, Culture & Society*, 45(6). <https://doi.org/10.1177/01634437231159555>
- Livingstone, S. (2019). Audiences in an Age of Datafication: Critical Questions for Media Research. *Television & New Media*, 20(2), 170-183. <https://doi.org/10.1177/1527476418811118>
- Lomborg, S., & Kapsch, P. H. (2020). Decoding Algorithms. *Media, Culture & Society*, 42(5), 745-761. <https://doi.org/10.1177/0163443719855301>
- Lovink, G. (2019). *Tristes Por Diseño: Las Redes Sociales Como Ideología*. Consonni. <https://doi.org/https://doi.org/10.7203/rd.v1i7.208>

- Martínez-Estrella, E.-C., Samacá-Salamanca, E., García-Rivero, A., & Cifuentes-Ambra, C. (2023). Generation Z in Chile, Colombia, México, and Panama: Interests and new digital consumption habits. Their use of Instagram and TikTok. *Profesional De La Información*, 32(2), 1-15. <https://doi.org/10.3145/epi.2023.mar.18>
- Martínez, J. (2023, May 5). *Atresmedia Entra en El Canal De Tiktok Ac2uality a Cambio De Un Pacto Comercial*. La Información. <https://go.revistacomunicar.com/M6CABa>
- Méndez Rodríguez, E. M. (2003). La descripción de documentos electrónicos a través de metadatos: una visión para la Archivística desde la nueva e-Administración. *Revista d'Arxius*, 2003, 47-82. <https://go.revistacomunicar.com/dOUT60>
- Newman, N. (2022). *How Publishers Are Learning to Create and Distribute News on TikTok* (1914566041). Reuters Institute for the Study of Journalism. <https://bit.ly/3OJ2RO0>
- Noble, S. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press. <https://doi.org/10.2307/j.ctt1pwt9w5>
- Pasquale, F. (2015). *The Black Box Society: the Secret Algorithms That Control Money and Information*. Harvard University Press. <https://doi.org/10.4159/harvard.9780674736061.c8>
- Pérez-Escolar, M., Alcaide-Pulido, P., & Del Toro, A. (2023). Nuevos referentes informativos de la generación Z: Estudio del rol de los y las influencers en TikTok como divulgadores/as de contenidos. *Prisma Social: Revista De Investigación Social*, (40), 262-288. <https://go.revistacomunicar.com/RWmfH4>
- Reder, M. (2012). *Globalización Y Filosofía*. Herder. <https://doi.org/10.2307/j.ctvt9k0dr>
- Riemer, K., & Peter, S. (2021). Algorithmic Audiencing: Why We Need to Rethink Free Speech on Social Media. *Journal of Information Technology*, 36(4), 409-426. <https://doi.org/10.1177/02683962211013358>
- Rifkin, J. (2000). *La Era Del Acceso. La Revolución De La Nueva Economía*. Paidós. <https://go.revistacomunicar.com/7Ahqlz>
- Thomas, S. L., Nafus, D., & Sherman, J. (2018). Algorithms as Fetish: Faith and Possibility in Algorithmic Work. *Big Data & Society*, 5(1), 1-11. <https://doi.org/10.1177/2053951717751552>
- TikTok. (2022, Nov 17). *An update on our platform API for researchers*. <https://bit.ly/3IGWBIO>
- Vázquez-Herrero, J., Negreira-Rey, M.-C., & Rodríguez-Vázquez, A.-I. (2021). Intersections Between TikTok and Tv: Channels and Programmes Thinking Outside the Box. *Journalism and Media*, 2(1), 1-13. <https://doi.org/10.3390/journalmedia2010001>
- Warner, M. (2012). *Público, Públicos, Contrapúblicos*. Fondo De Cultura Económica (FCE). <https://go.revistacomunicar.com/YdSkEH>
- Webster, J. G. (2016). Why Study Measures of Exposure? From Exposure to Attention. *Communication Methods and Measures*, 10(2-3), 179-180. <https://doi.org/10.1080/19312458.2016.1150974>
- Wolton, D. (1999). *Sobre La Comunicación*. Acento Editorial. <https://go.revistacomunicar.com/b94hh5>