

Campo / Field	Contenido / Content
Tipo de contribución / Article type	Artículo original de investigación
Título en español	Concentración de actividad y escala de participación en chats de streaming: reanálisis de 694.728 canales de Twitch
Title in English	Activity Concentration and Participation Scaling in Livestream Chats: A Reanalysis of 694,728 Twitch Channels
Autorías / Authors	Myriam Cecilia García-Enríquez — UNIVERSIDAD DE GUAYAQUIL, Ecuador — ORCID: https://orcid.org/0009-0005-9851-7925
Correspondencia / Correspondence	Myriam Cecilia García-Enríquez — myriam.garcia@revista.asesoreseducativos-ec.com

Resumen

El chat sincrónico hace visible una parte de la participación en plataformas de streaming, pero los promedios pueden ocultar desigualdades muy pronunciadas entre canales. Este estudio cuantificó la concentración de mensajes y participantes activos, y estimó cómo escala la actividad acumulada con el número de participantes activos. Se reanalizó un conjunto público anonimizado de Twitch con 694.728 canales observados entre el 26 de agosto y el 10 de noviembre de 2014. Se calcularon descriptivos, coeficientes de Gini, participaciones acumuladas, estratos de tamaño y una regresión robusta $\log_{10}(\text{mensajes}) \sim \log_{10}(\text{participantes activos})$, con una sensibilidad que excluyó el 0,1 % superior por mensajes. Los canales reunieron 496.580.654 mensajes; la mediana fue 7 mensajes y 2 participantes activos, frente a medias de 714,78 y 36,42. El Gini alcanzó 0,9653 para mensajes y 0,9313 para participantes activos. El 1 % superior concentró 71,00 % de los mensajes y 71,60 % de los recuentos de participantes; el 50 % inferior aportó 0,16 % y 1,64 %. La pendiente de escala fue 1,5301 (IC 95 %: 1,5280–1,5322; $R^2 = 0,8229$) y permaneció en 1,5407 tras excluir el 0,1 % superior. La actividad y la participación activa estuvieron fuertemente concentradas, y el volumen de mensajes mostró una asociación superlineal con el número de participantes activos. El diseño no identifica causalidad: los acumulados mezclan exposición, duración y frecuencia, y no representan audiencia pasiva, retención ni calidad comunitaria.

Palabras clave: streaming en vivo; Twitch; chat sincrónico; participación digital; concentración; escalamiento

Abstract

Synchronous chat makes one component of participation visible on livestreaming platforms, yet averages may conceal extreme inequalities across channels. This study quantified the concentration of messages and active chatters and estimated how accumulated activity scales with the number of active chatters. We reanalyzed an anonymized public Twitch dataset comprising 694,728 channels observed from 26 August to 10 November 2014. Descriptive statistics, Gini coefficients, cumulative shares, size strata, and a robust $\log_{10}(\text{messages}) \sim \log_{10}(\text{active chatters})$ regression were calculated, with a sensitivity analysis excluding the top 0.1% of channels by messages. The channels accumulated 496,580,654 messages. Medians were 7 messages and 2 active chatters, compared with means of 714.78 and 36.42. Gini coefficients were 0.9653 for messages and 0.9313 for active-chatter counts. The top 1% accounted for 71.00% of messages and 71.60% of active-chatter counts, whereas the bottom 50% accounted for 0.16% and 1.64%. The scaling slope was 1.5301 (95% CI: 1.5280–1.5322; $R^2 = 0.8229$) and remained 1.5407 after excluding the top 0.1%. Activity and active participation were highly concentrated, and message volume had a superlinear association with active-chatter counts. The design does not identify causality: accumulated counts combine exposure, duration, and frequency and do not measure passive audiences, retention, or community quality.

Keywords: livestreaming; Twitch; synchronous chat; digital participation; concentration; scaling

1. Introducción

El streaming en vivo combina transmisión audiovisual, copresencia temporal e interacción textual. En Twitch, el chat permite que espectadores, moderadores y creadores produzcan señales públicas mientras el contenido se desarrolla. Esa infraestructura favorece conversaciones, bromas compartidas, normas locales y formas de pertenencia que no se reducen a mirar un video (Hamilton et al., 2014; Cai et al., 2021). La actividad observable, sin embargo, se distribuye en un ecosistema compuesto por canales de escalas muy diferentes. Un promedio global puede describir poco tanto al canal típico como a los canales que concentran grandes audiencias y ritmos intensos de mensajes.

La participación por chat es relevante porque conecta conducta y estructura comunitaria. Estudios contemporáneos han asociado distintas formas de involucramiento con capital social, compromiso financiero, suscripción y apoyo entre participantes (Küper & Krämer, 2021; Bründl et al., 2022; Wolff & Shen, 2022; Yuan & Lee, 2022; You et al., 2023). Otros trabajos muestran que las reglas de microcomunidades, la moderación y el sentido de comunidad condicionan qué interacciones resultan posibles o legítimas (Cai et al., 2021; Kairam et al., 2022). Estas relaciones no convierten el número de mensajes en una medida completa de comunidad, pero explican por qué su distribución constituye un objeto de investigación propio.

El tamaño de una transmisión puede modificar la experiencia del chat. Nematzadeh et al. (2019), autores del conjunto reutilizado, analizaron la transición desde conversaciones manejables hacia condiciones de sobrecarga. Cuando muchas personas intervienen de manera simultánea, la atención disponible se fragmenta y el intercambio puede aproximarse a una cacofonía. La sobrecarga también aparece en investigaciones posteriores sobre consumo de información en transmisiones, aunque en contextos y con medidas diferentes (Zhang et al., 2023). En consecuencia, más actividad no debe interpretarse automáticamente como mejor conversación, mayor cohesión o participación más equitativa.

La literatura reciente también ha ampliado el objeto de estudio más allá del videojuego. Se han examinado transmisiones de la vida cotidiana, comercio, turismo, conciertos y realidad virtual (Onderdijk et al., 2021; Lee et al., 2023; Wu et al., 2023; Luo et al., 2023). La presencia social contribuye a explicar por qué una transmisión puede sentirse compartida y estimular intenciones de participación (Chen & Liao, 2022). No obstante, presencia percibida, participación textual y permanencia son constructos distintos. Un conjunto agregado por canal solo observa a quienes escribieron, no a quienes miraron silenciosamente ni la interpretación subjetiva de la experiencia.

En términos distributivos, hay al menos dos preguntas. La primera es cuánto de la actividad total está concentrado en una fracción pequeña de canales. La segunda es si el volumen de mensajes aumenta de forma proporcional, subproporcional o superproporcional respecto del número de participantes activos. Una relación proporcional implicaría que duplicar participantes activos se asocia, en promedio, con duplicar mensajes acumulados. Una pendiente logarítmica superior a uno describiría una asociación superlineal: los canales con más participantes activos acumulan mensajes a un ritmo mayor que el proporcional. Pero esa pendiente puede reflejar exposición temporal, frecuencia de emisión, selección de canales, repetición de participantes u otros factores no medidos.

La concentración tiene consecuencias metodológicas. Si pocos canales aportan la mayoría de observaciones de actividad, las medias pueden quedar dominadas por casos extremos y los análisis de una muestra de transmisiones populares no representarían el ecosistema completo. Del mismo modo, los canales pequeños son numerosos pero aportan una proporción reducida del tráfico. Distinguir ambas escalas permite interpretar mejor estudios de comportamiento, diseñar muestreos y evitar generalizaciones desde una parte visible hacia el conjunto.

El objetivo general fue cuantificar la concentración de mensajes y participantes activos entre canales de Twitch y estimar la asociación de escala entre ambas variables. Los objetivos específicos fueron describir las distribuciones; calcular Gini y participaciones de las colas; comparar estratos de tamaño; estimar una pendiente log-log robusta; y evaluar su sensibilidad a los canales con mayor volumen. Se formuló como expectativa descriptiva que la actividad estaría fuertemente concentrada y que la relación entre mensajes y participantes activos sería superlineal. No se plantearon hipótesis causales.

2. Metodología

2.1. Diseño y fuente de datos

Se realizó un análisis secundario cuantitativo, transversal y descriptivo. La fuente fue el conjunto público asociado con *Information overload in group communication: from conversation to cacophony in the Twitch chat* (Nematzadeh et al., 2019), depositado en Zenodo con DOI 10.5281/zenodo.1182793 y licencia CC BY 2.0. El archivo stats.h5 ocupa 217.936.312 bytes; su MD5, b850c6227691e3473cce381204584c6a, coincidió con el publicado por el repositorio.

Los datos proceden de observaciones realizadas entre el 26 de agosto y el 10 de noviembre de 2014. El contenedor HDF5 incluye estadísticas agregadas por canal y por usuario anonimizado. Para este artículo se utilizó la tabla por canal /chan_stats; /user_stats se auditó para comprender la estructura, pero no se empleó en los modelos presentados. La antigüedad de la observación se conserva explícitamente: la literatura hasta el 30 de junio de 2025 contextualiza los resultados, pero no transforma el conjunto en una medición contemporánea de Twitch.

2.2. Unidad de análisis y variables

La unidad fue el canal con al menos un mensaje y un participante activo registrado. Se analizaron 694.728 canales. messages representa mensajes acumulados por canal y users, participantes activos acumulados por canal. La suma de users entre canales no equivale a personas únicas en la plataforma, porque una misma identidad anonimizada puede participar en varios canales y porque la tabla resume cada canal por separado. Se calculó messages_per_active_chatter como cociente descriptivo.

La expresión “participante activo” designa exclusivamente a una cuenta que envió mensajes. No incluye espectadores pasivos, seguidores, suscriptores ni audiencia simultánea. Tampoco permite distinguir personas, cuentas automatizadas o funciones de moderación. Esta precisión terminológica se aplicó a tablas, figuras e interpretación.

2.3. Auditoría y preparación

Se verificaron nombres de tablas, tipos, dimensiones, valores ausentes, duplicados, mínimos, máximos y consistencia aritmética. Los registros analizados tuvieron valores positivos en mensajes y participantes activos, por lo que no fue necesario añadir constantes antes de la transformación logarítmica. No se imputaron datos ni se eliminaron casos del análisis principal. El conjunto derivado conserva una fila por canal y las tres variables analíticas.

2.4. Análisis descriptivo y concentración

Para cada variable se calcularon n, media, desviación estándar, mínimo, percentiles 25, 50, 75, 90, 95 y 99, y máximo. La desigualdad se estimó mediante el coeficiente de Gini, donde cero indica igualdad perfecta y uno una concentración límite. También se calculó la participación acumulada del 0,1 %, 1 %, 5 % y 10 % superiores y del 50 % inferior. Las curvas de Lorenz representan la fracción acumulada de canales frente a la fracción acumulada de la variable, ordenada de menor a mayor.

Los canales se agruparon a priori según participantes activos: 1; 2–5; 6–20; 21–100; 101–500; y más de 500. Dentro de cada estrato se informaron número de canales, medianas de participantes, mensajes y mensajes por participante activo, y rango intercuartílico de mensajes. Esta estratificación no supone categorías naturales de Twitch; funciona como resumen interpretable de una distribución continua.

2.5. Modelo de escala y sensibilidad

Se ajustó por mínimos cuadrados el modelo $\log_{10}(\text{messages}) = a + b \times \log_{10}(\text{users})$. La pendiente b resume elasticidad descriptiva: un incremento de 1 % en participantes activos se asocia con b % en mensajes dentro de la relación ajustada. Se calcularon errores estándar HC3 e intervalos del 95 %, apropiados frente a heterocedasticidad. El R^2 cuantificó la variación explicada en escala logarítmica.

Como análisis de sensibilidad se excluyó el 0,1 % de canales con más mensajes, definido por el percentil 99,9 (93.578,926 mensajes), y se repitió el ajuste. El propósito no fue preferir una estimación recortada, sino comprobar si la dirección y magnitud dependían exclusivamente de los casos más extremos. La figura de escala utiliza medianas por intervalos de participantes y muestra la recta solo dentro del soporte observado, sin extrapolar.

2.6. Ética, licencia y reproducibilidad

El estudio no reclutó participantes ni intentó reidentificar cuentas. Reutilizó recuentos públicos anonimizados con atribución al repositorio y al artículo fuente. Por las características descritas, no se requirió una nueva evaluación ética para el presente reanálisis. Código, resultados tabulares, figuras y huellas de integridad se conservan en el expediente científico. La fecha efectiva y el corte bibliográfico fueron el 30 de junio de 2025.

3. Resultados

3.1. Distribuciones de actividad y participación

Los 694.728 canales acumularon 496.580.654 mensajes. La suma de recuentos de participantes activos por canal fue 25.300.073, sin interpretarse como personas únicas. El canal mediano tuvo 7 mensajes y 2 participantes activos, mientras que las medias fueron 714,78 y 36,42. La gran separación entre mediana y media, junto con máximos de 5.428.026 mensajes y 139.286 participantes activos, muestra una cola derecha extrema.

Tabla 1. Distribución de mensajes, participantes activos y cociente por canal

Variable	Media (DE)	P25	Mediana	P75	P90	P95	P99	Máximo
Mensajes	714,78 (13.503,48)	2	7	52	341	1.162	12.288,65	5.428.026
Participantes activos	36,42 (681,13)	1	2	5	20	54	452	139.286
Mensajes por participante activo	25,81 (8.505,94)	1,00	3,00	9,83	23,21	36,67	83,67	5.428.026

Nota. $n = 694.728$ canales. El cociente extremo corresponde a un canal con un participante activo y no caracteriza al canal típico.

El 40,16 % de los canales tuvo un solo participante activo y otro 35,12 % tuvo entre dos y cinco. Así, tres cuartas partes del conjunto se concentraron en los dos estratos inferiores. Solo 6.312 canales, 0,91 % del total, superaron 500 participantes activos.

3.2. Concentración entre canales

El Gini fue 0,9653 para mensajes y 0,9313 para participantes activos. Ambas curvas de Lorenz se apartaron ampliamente de la diagonal de igualdad. La de mensajes mostró una concentración ligeramente mayor en la mayor

parte de la distribución, mientras que la fracción extrema de canales reunió una proporción particularmente alta de participantes activos.

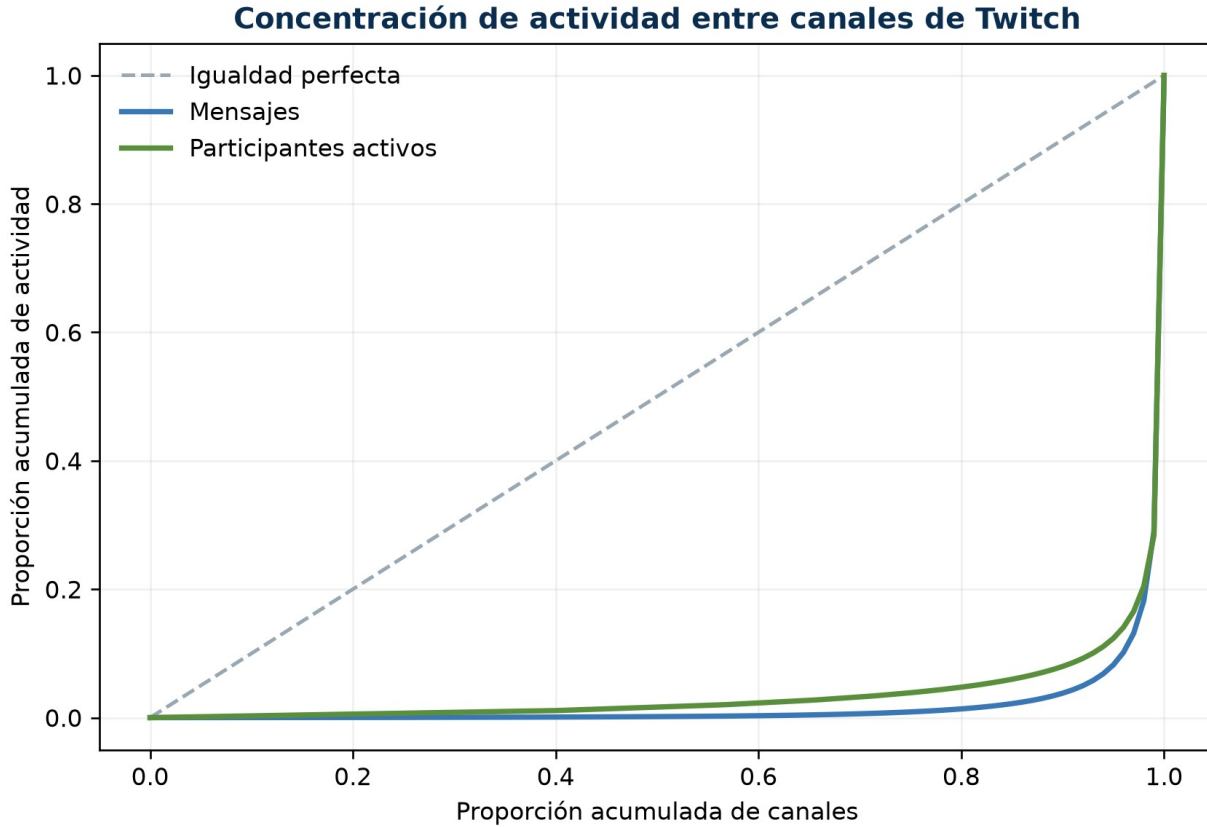


Figura 1. Curvas de Lorenz de mensajes y participantes activos acumulados entre canales.

Tabla 2. Indicadores de concentración entre canales

Variable	Gini	0.1 % superior	1 % superior	5 % superior	10 % superior	50 % inferior
Mensajes	0,9653	32,41 %	71,00 %	91,78 %	96,19 %	0,16 %
Participantes activos	0,9313	39,93 %	71,60 %	87,65 %	92,03 %	1,64 %

El 1 % superior concentró alrededor de siete décimas partes de ambos totales. En contraste, la mitad inferior de canales produjo 0,16 % de los mensajes y reunió 1,64 % de los recuentos de participantes activos. La distribución, por tanto, no puede resumirse adecuadamente con un “canal promedio”.

3.3. Estratos de participantes activos

La mediana de mensajes aumentó de 1 en canales con un participante activo a 25.247 en los que superaron 500. El cociente mediano pasó de 1 a 20,38 mensajes por participante activo en el estrato 101–500 y descendió a 18,42 en el estrato superior. Esta no monotonicidad advierte que la relación global no equivale a un incremento uniforme del comportamiento individual.

Tabla 3. Actividad según estratos de participantes activos

Estrato	Canales	Mediana de participantes	Mensajes, mediana [P25-P75]	Mediana de mensajes por participante
1	279.004	1	1 [1-2]	1,00
2-5	244.007	3	10 [5-28]	3,67
6-20	104.450	9	93 [42-199]	9,63
21-100	45.054	37	637 [323-1.260]	15,97
101-500	15.901	184	4.029 [2.154-7.622]	20,38
Más de 500	6.312	1.102	25.247 [12.972,5-50.632]	18,42

3.4. Asociación de escala

En el conjunto completo, la pendiente log-log fue 1,5301 (EE HC3 = 0,0011; IC 95 %: 1,5280–1,5322) y el R^2 fue 0,8229. Una diferencia de 1 % en participantes activos se asoció, dentro del modelo, con una diferencia aproximada de 1,53 % en mensajes. Al ser mayor que uno, el patrón fue superlineal.

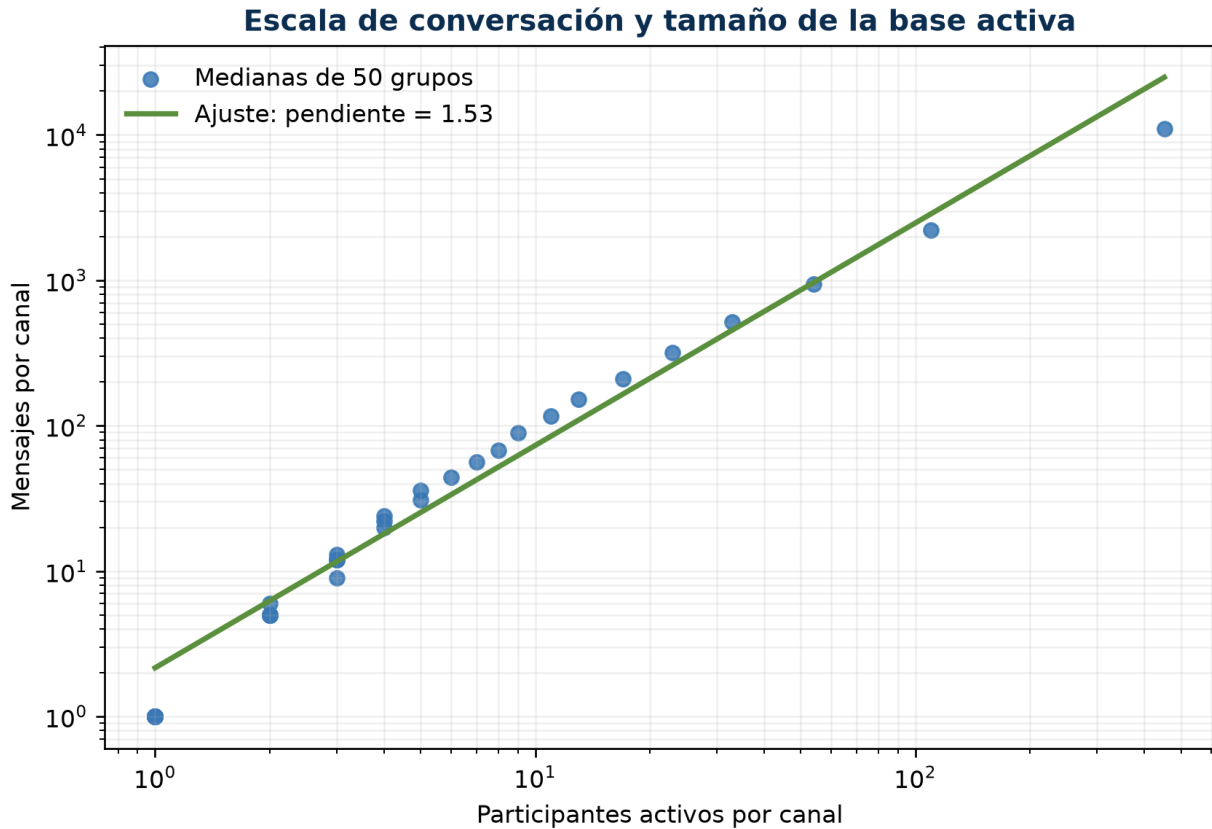


Figura 2. Escala de mensajes acumulados respecto de participantes activos: medianas por intervalos y ajuste log-log.

Después de excluir los 695 canales del 0,1 % superior por mensajes, la pendiente fue 1,5407 (EE HC3 = 0,0010; IC 95 %: 1,5387–1,5427; R^2 = 0,8214; n = 694.033). La magnitud no disminuyó y la conclusión descriptiva no dependió de la cola más extrema. Sin embargo, el modelo no separa intensidad individual, duración total observada, frecuencia de transmisión o composición de participantes.

4. Discusión

Este reanálisis muestra dos rasgos simultáneos del ecosistema observado: concentración extrema y asociación superlineal. La mayoría de los canales reunió pocos participantes activos y mensajes, mientras una fracción pequeña concentró la mayor parte de la actividad. Además, los canales con más participantes activos acumularon mensajes a una tasa mayor que la proporcional. La sensibilidad confirma que este patrón no fue creado únicamente por el 0,1 % de canales con más mensajes.

La concentración tiene una lectura sustantiva y otra metodológica. Sustantivamente, indica que la visibilidad de la conversación pública estuvo dominada por un conjunto reducido de espacios. Metodológicamente, implica que una muestra basada en tráfico, tendencias o canales conocidos sobreponderaría experiencias atípicas. La mediana de 7 mensajes describe un ecosistema muy distinto del promedio de 714,78. Estudios que examinen moderación, normas o interacción deberían definir si su población de interés son los numerosos canales pequeños, los pocos canales de gran actividad o una mezcla ponderada.

La pendiente superlineal es compatible con mecanismos de interacción grupal: más participantes generan más oportunidades de respuesta, imitación, coordinación o señalización. La investigación sobre participación sincrónica sugiere que las conductas de coparticipantes pueden influir en compromiso y suscripción (Bründl et al., 2022), y la literatura sobre capital social relaciona diferentes patrones de involucramiento con recursos relacionales (Küper & Krämer, 2021; Yuan & Lee, 2022). Sin embargo, estos trabajos no identifican la causa de la pendiente aquí estimada. Los datos agregados carecen de orden temporal y no distinguen mensajes iniciadores de respuestas.

También es plausible que la exposición común impulse ambas variables. Un canal que transmite durante más horas o días tiene más oportunidades de acumular participantes y mensajes. La frecuencia de emisión, la popularidad previa, los eventos, las categorías de contenido y la promoción externa pueden producir una asociación aun si el comportamiento individual no cambia con el tamaño. El descenso del cociente mediano en el estrato superior refuerza la cautela: superlinealidad agregada no significa que cada participante escriba progresivamente más en los canales más grandes.

Los resultados dialogan con la tesis de sobrecarga de Nematzadeh et al. (2019). El crecimiento intenso del volumen puede hacer más difícil seguir conversaciones, reconocer destinatarios o sostener turnos. Mihailova (2020) mostró, en un estudio cualitativo de chats, que la negatividad y su interpretación son contextuales. Ringer et al. (2020) subrayaron la necesidad de conjuntos adecuados para estudiar lenguaje de chat. En conjunto, estas contribuciones muestran que contar mensajes es un punto de partida; comprender conversación requiere texto, secuencia, moderación y contexto.

La actividad tampoco equivale a calidad comunitaria. Las reglas locales y su aplicación estructuran microcomunidades (Cai et al., 2021), mientras que el sentido de comunidad depende de factores individuales, relacionales y ambientales (Kairam et al., 2022). Un flujo abundante puede contener apoyo, celebración y aprendizaje, pero también repetición, spam, exclusión o mensajes imposibles de procesar. Por eso, los indicadores de volumen deberían combinarse con medidas de reciprocidad, diversidad de voces, moderación, permanencia y experiencia subjetiva.

La literatura reciente confirma que “streaming” no designa una situación homogénea. En transmisiones de la vida cotidiana, la presencia social puede apoyar construcción comunitaria (Lee et al., 2023); en realidad virtual, la interacción atraviesa pantallas y espacios (Wu et al., 2023); en conciertos, presencia y contexto social se relacionan de forma diferenciada con conexión (Onderdijk et al., 2021); y en comercio coexisten información, persuasión y sobrecarga (Luo et al., 2023; Zhang et al., 2023). Estas fuentes ayudan a delimitar el alcance: un patrón hallado en Twitch en 2014 no debe transferirse automáticamente a turismo, comercio u otras plataformas.

La participación activa es además una selección de la audiencia. Chen y Liao (2022) vincularon presencia social con motivaciones para mirar transmisiones, pero mirar no exige escribir. Speed et al. (2023) mostraron que el disfrute de Twitch comprende razones que exceden la intervención textual. Por consiguiente, los 25,3 millones de recuentos canal-participante no describen la audiencia total y no permiten calcular una tasa de participación. Para ello se requerirían espectadores únicos o concurrentes con ventanas temporales compatibles.

Los hallazgos sí ofrecen implicaciones operativas prudentes. Primero, los análisis de plataforma deberían informar medianas, percentiles y concentración, no solo promedios. Segundo, la selección de canales debería estratificarse por tamaño y justificar ponderaciones. Tercero, las herramientas de moderación y visualización pueden necesitar diseños diferentes para conversaciones escasas, medianas y masivas. Cuarto, cualquier indicador de “engagement” basado en mensajes debe declarar a quién excluye y qué periodo acumula.

Finalmente, la concentración observada no debe naturalizarse como una propiedad inmutable. Corresponde a una ventana histórica, una plataforma y un procedimiento de recolección específicos. Twitch, sus interfaces y prácticas comunitarias han evolucionado; publicaciones como Wolff y Shen (2022) y Chaudhry et al. (2024) examinan mecanismos que no pueden retroproyectarse al archivo. El valor del conjunto reside en permitir un análisis masivo y reproducible de aquella ventana, no en representar sin reservas el presente.

5. Conclusiones

En 694.728 canales de Twitch observados en 2014, los mensajes y los recuentos de participantes activos estuvieron extremadamente concentrados. El 1 % superior reunió aproximadamente 71 % de ambos totales, mientras la mitad inferior aportó una fracción mínima. El canal típico fue pequeño: 7 mensajes y 2 participantes activos.

El volumen de mensajes mostró una asociación superlineal robusta con el número de participantes activos. La pendiente de 1,53 permaneció después de excluir el 0,1 % superior por actividad. Esta relación describe la covariación de acumulados; no demuestra que aumentar el tamaño cause más actividad individual ni que un chat más activo sea más saludable, inclusivo o valioso.

Para investigación y gestión, el principal aprendizaje es evitar que el promedio o los canales más visibles sustituyan al ecosistema completo. El análisis de chats de streaming requiere reconocer concentración, escala, exposición y selección de participantes, y complementar el volumen con evidencia temporal, conversacional y comunitaria.

6. Limitaciones

Los datos corresponden a agosto–noviembre de 2014 y no representan automáticamente el Twitch de 2025. Las variables son agregadas por canal y no permiten reconstruir sesiones, duración de transmisiones, concurrencia, categorías, idioma, país, género, edad, moderación o eventos externos. La suma de participantes activos por canal contiene posibles repeticiones entre canales.

La medición excluye a espectadores pasivos y no permite estimar audiencia, conversión, retención o propensión a escribir. Tampoco distingue cuentas humanas, automatizadas o de moderación. El cociente de mensajes por participante es una razón agregada y no la media exacta de mensajes individuales.

El modelo log-log es descriptivo y transversal. La precisión estadística se beneficia del gran n , pero no corrige confusión, dependencia entre canales, sesgo de cobertura o error de medición. La heterocedasticidad se atendió con HC3; no se modeló autocorrelación temporal porque el archivo analítico no conserva una serie por canal.

La concentración se calculó sobre canales incluidos con actividad positiva; no incorpora canales existentes sin mensajes. Los umbrales de estratificación son analíticos, no categorías oficiales. Finalmente, la literatura posterior contextualiza posibles mecanismos sin probar que operaran en la ventana histórica analizada.

Agradecimientos

Se reconoce a Azadeh Nematzadeh, Giovanni Luca Ciampaglia, Yong-Yeol Ahn y Alessandro Flammini por publicar el conjunto y la documentación científica asociada. Este reconocimiento no implica participación en el presente reanálisis ni aprobación de sus conclusiones.

Declaraciones

Campo / Field	Contenido / Content
Contribuciones CRediT	Conceptualización; metodología; software; validación; análisis formal; investigación; curación de datos; visualización; redacción del borrador original; redacción, revisión y edición; administración del proyecto.
Financiamiento	La investigación no recibió financiación externa específica.
Conflictos de interés	La persona autora declara no tener conflictos de interés.
Aprobación ética	Reanálisis de estadísticas públicas anonimizadas, sin reclutamiento ni intento de reidentificación. Por estas características no se requirió una nueva evaluación ética.
Consentimiento	No aplicable; no se reclutaron participantes ni se recopilaban datos nuevos.
Disponibilidad de datos	Zenodo: https://doi.org/10.5281/zenodo.1182793 . Código y resultados derivados en el expediente reproducible.
Uso de inteligencia artificial	Codex de OpenAI apoyó la programación reproducible, la organización del manuscrito, la traducción académica y los controles editoriales. La persona autora revisó el contenido y asume la responsabilidad íntegra por la exactitud, originalidad, interpretación y versión final del artículo.

Referencias

- Bründl, S., Matt, C., Hess, T., & Engert, S. (2022). How synchronous participation affects the willingness to subscribe to social live streaming services: The role of co-interactive behavior on Twitch. *European Journal of Information Systems*, 32(5), 800–817. <https://doi.org/10.1080/0960085X.2022.2062468>
- Cai, J., Guanlao, C., & Wohn, D. Y. (2021). Understanding rules in live streaming micro communities on Twitch. *ACM International Conference on Interactive Media Experiences*, 290–295. <https://doi.org/10.1145/3452918.3465491>
- Chaudhry, A., Wang, Y., & Ouyang, E. (2024). Community gifting and recipients' engagement with peers and influencers: Evidence from natural experiments on Twitch. *Journal of Interactive Marketing*, 60(1), 83–99. <https://doi.org/10.1177/10949968241257891>
- Chen, J., & Liao, J. (2022). Antecedents of viewers' live streaming watching: A perspective of social presence theory. *Frontiers in Psychology*, 13, 839629. <https://doi.org/10.3389/fpsyg.2022.839629>
- Hamilton, W. A., Garretson, O., & Kerne, A. (2014). Streaming on Twitch: Fostering participatory communities of play within live mixed media. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1315–1324. <https://doi.org/10.1145/2556288.2557048>
- Kairam, S. R., Mercado, M. C., & Sumner, S. A. (2022). A social-ecological approach to modeling sense of virtual community (SOVC) in livestreaming communities. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2), 1–35. <https://doi.org/10.1145/3555081>
- Küper, A., & Krämer, N. C. (2021). Influencing factors for building social capital on live streaming websites. *Entertainment Computing*, 39, 100444. <https://doi.org/10.1016/j.entcom.2021.100444>

- Lee, Y.-H., Yuan, C. W., & Bi, N. (2023). Investigating social presence in “In Real Life” streaming for community building. *New Media & Society*, 27(1), 379–403. <https://doi.org/10.1177/14614448231176769>
- Luo, X., Lim, W. M., Cheah, J.-H., Lim, X.-J., & Dwivedi, Y. K. (2023). Live streaming commerce: A review and research agenda. *Journal of Computer Information Systems*, 65(3), 376–399. <https://doi.org/10.1080/08874417.2023.2290574>
- Mihailova, T. (2020). Navigating ambiguous negativity: A case study of Twitch.tv live chats. *New Media & Society*, 24(8), 1830–1851. <https://doi.org/10.1177/1461444820978999>
- Nematzadeh, A., Ciampaglia, G. L., Ahn, Y.-Y., & Flammini, A. (2019). Information overload in group communication: From conversation to cacophony in the Twitch chat. *Royal Society Open Science*, 6(10), 191412. <https://doi.org/10.1098/rsos.191412>
- Onderdijk, K. E., Swarbrick, D., Van Kerrebroeck, B., Mantei, M., Vuoskoski, J. K., Maes, P.-J., & Leman, M. (2021). Livestream experiments: The role of COVID-19, agency, presence, and social context in facilitating social connectedness. *Frontiers in Psychology*, 12, 647929. <https://doi.org/10.3389/fpsyg.2021.647929>
- Ringer, C., Nicolaou, M., & Walker, J. (2020). TwitchChat: A dataset for exploring livestream chat. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 16(1), 259–265. <https://doi.org/10.1609/aiide.v16i1.7439>
- Sjöblom, M., & Hamari, J. (2017). Why do people watch others play video games? An empirical study on the motivations of Twitch users. *Computers in Human Behavior*, 75, 985–996. <https://doi.org/10.1016/j.chb.2016.10.019>
- Speed, A., Burnett, A., & Robinson, T., II. (2023). Beyond the game: Understanding why people enjoy viewing Twitch. *Entertainment Computing*, 45, 100545. <https://doi.org/10.1016/j.entcom.2022.100545>
- Wolff, G. H., & Shen, C. (2022). Audience size, moderator activity, gender, and content diversity: Exploring user participation and financial commitment on Twitch.tv. *New Media & Society*, 26(2), 859–881. <https://doi.org/10.1177/14614448211069996>
- Wu, L., Liu, Q., Zhao, J., & Lank, E. (2023). Interactions across displays and space: A study of virtual reality streaming practices on Twitch. *Proceedings of the ACM on Human-Computer Interaction*, 7(ISS), 242–265. <https://doi.org/10.1145/3626473>
- You, Z., Wang, M., & Shamu, Y. (2023). The impact of network social presence on live streaming viewers’ social support willingness: A moderated mediation model. *Humanities and Social Sciences Communications*, 10(1), 385. <https://doi.org/10.1057/s41599-023-01892-8>
- Yuan, C. W., & Lee, Y.-H. (2022). Examining different viewer engagement patterns for social capital on streaming communities. *Social Science Computer Review*, 41(6), 2055–2072. <https://doi.org/10.1177/08944393221131930>
- Zhang, G., Cao, J., & Liu, D. (2023). Examining the influence of information overload on consumers’ purchase in live streaming: A heuristic-systematic model perspective. *PLOS ONE*, 18(8), e0284466. <https://doi.org/10.1371/journal.pone.0284466>

Fuente de datos

Nematzadeh, A., Ciampaglia, G. L., Ahn, Y.-Y., & Flammini, A. (2018). *Replication data for: Information overload in group communication: From conversation to cacophony in the Twitch chat* [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.1182793>

Anexos

No se incluyen anexos.

Información editorial

Campo / Field	Contenido / Content
Recepción / Received	12 de mayo de 2025 / May 12, 2025
Revisión / Revised	3 de junio de 2025 / June 3, 2025
Aceptación / Accepted	28 de junio de 2025 / June 28, 2025
Publicación / Published	29 de junio de 2025 / June 29, 2025
Volumen, número y año / Volume, issue and year	Vol. 1, Núm. 1, enero-junio de 2025