

Aproximación a Reddit: Reflexión de su posible aplicación en investigaciones pragmáticas

Yuezhe Du - Universitat de València
yuezdu@alumni.uv.es

Rebut / Received: 17/03/2025

Acceptat / Accepted: 08/07/2025

Resum. Aproximació a Reddit: Reflexió sobre la seva possible aplicació en investigacions pragmàtiques. Aquest estudi examina Reddit, una pàgina de col·lecció de fòrums de debat, amb l'objectiu d'analitzar el seu potencial com a font de dades per a investigacions pragmàtiques i aportar metodològicament a l'estudi de l'anàlisi pragmàtica en entorns digitals. S'hi examinen característiques com l'anonimat dels usuaris, el sistema de votació, la Reddiquette i les normes dels subreddits, entre altres. Mitjançant una revisió bibliogràfica i una anàlisi crítica, el treball reflexiona sobre el potencial, els defectes i les consideracions ètiques d'aquesta plataforma per a estudis pragmàtics, oferint una contribució metodològica a l'anàlisi pragmàtica en contextos digitals.

Paraules clau: *pragmàtica, discurs digital, xarxes socials, Reddit, metodologia lingüística.*

Abstract. Approach to Reddit: Reflection on its potential application in pragmatic research.

This study examines Reddit, a collection of discussion forums, with the aim of analyzing its potential as a data source for pragmatic research and contributing methodologically to the study of pragmatic analysis in digital environments. It explores features such as user anonymity, the voting system, Reddiquette, and subreddit rules, among others. Through a literature review and critical analysis this work reflects on the platform's potential, shortcomings, and ethical considerations for pragmatic studies, offering a methodological contribution to pragmatic analysis in digital contexts.

Keywords: *pragmatics, digital discourse, social media, Reddit, linguistic methodology.*

1. Introducción

Reddit es una colección de foros de debate y una de las redes sociales más populares. Sin embargo, pese a su creciente relevancia, ha recibido menos atención que Twitter¹ y Facebook en estudios lingüísticos (Lu, 2024). Aunque en inglés existen varias investigaciones pragmáticas sobre Reddit como los trabajos de Lu (2025) y Schneider (2025), hasta donde sepamos en el ámbito lingüístico hispanohablante sigue siendo un terreno inexplorado. Este trabajo analiza Reddit como fuente de datos para estudios pragmáticos, explora su potencial y propone enfoques metodológicos para el análisis en contextos digitales.

La estructura es la siguiente: primero, se describen Reddit y sus características principales, como la *Reddiqueta*, las normas de las comunidades, el anonimato y el sistema de votación. Luego, se hace una revisión bibliográfica sobre su aplicación en investigaciones científicas, con énfasis en la pragmática lingüística. En la tercera parte, se evalúan ventajas y limitaciones de Reddit para estudios pragmáticos. Finalmente, se discuten aspectos éticos clave al trabajar con sus datos. A pesar de desafíos como el anonimato, la falta de datos demográficos y la heterogeneidad de los *subreddits*, se argumenta que Reddit es una fuente valiosa para estudios pragmáticos.

2. ¿Qué es Reddit?

Reddit, fundado en 2005 por Alexis Ohanian y Steve Huffman, fue la sexta página más visitada en 2024 (Statista, 2025) y gana popularidad en países hispanohablantes. Se define como una red de comunidades donde los usuarios exploran intereses y se autodenomina “*the front page of the internet*” (Reddit, 2024). Funciona como una colección de foros. Los foros se llaman *subreddits*, y cada uno está centrado en un tema específico (Lu, 2024).

Proferes et al. (2021) señalan que Reddit comparte características con Twitter, como la naturaleza pública de su contenido, una API² abierta y la rapidez de reacción ante eventos, lo que lo hace valioso para la investigación. No obstante, en comparación con otras redes sociales populares, como Facebook y Twitter, Reddit sigue estando relativamente poco estudiada desde una perspectiva lingüística (Lu, 2024, p. 31).

Reddit es accesible sin registro, pero solo los usuarios registrados (*redditores*) pueden publicar contenidos en la plataforma (Leuckert y Leuckert, 2020). Según Panek (2021), los *redditores* pueden ser contribuidores (publican o comentan) o electores (votan

1. Twitter cambió su nombre a “X” en julio de 2023, tras la adquisición de la empresa por Elon Musk. Sin embargo, el término “Twitter” sigue siendo ampliamente utilizado para referirse a la plataforma.

2. API (Interfaz de Programación de Aplicaciones) es un conjunto de protocolos y herramientas predefinidas que permite la comunicación entre sistemas o componentes de *software*, facilitando el intercambio de datos y funcionalidades sin necesidad de acceder a su código interno.

publicaciones o comentarios). Los administradores, empleados de Reddit, gestionan el sitio, mientras que los moderadores, voluntarios, administran los *subreddits*. Aunque Reddit fomenta el anonimato, los *redditores* deben crear un nombre de usuario y un perfil público, en el que queda visible el historial de publicaciones y comentarios de los usuarios.

La página principal destaca hilos según votos, comentarios e interacciones. Los usuarios pueden personalizar su contenido según suscriptores e historial. La actividad principal en Reddit ocurre dentro de *subreddits*, comunidades temáticas gestionadas por moderadores bajo la convención de nombres “/r/” seguido del tema (Glassman y Kuznetcova, 2022). Hasta octubre de 2023, Reddit contaba con más de 100.000 *subreddits* activos (Reddit, 2024), cada uno con sus propias normas y regulaciones, cuyo cumplimiento es supervisado por moderadores que eliminan contenido inapropiado. Como se muestra en la Figura 1, la estructura de Reddit se organiza en diferentes niveles: en primer lugar, la página principal muestra una lista de publicaciones destacadas que pueden pertenecer a distintos *subreddits*; luego, cada *subreddit* agrupa publicaciones relacionadas con su temática específica. Finalmente, dentro de cada *post*, los comentarios se organizan en un árbol de discusión, en el que cada respuesta inicia una nueva rama, creando una estructura jerárquica de mensajes. Según Vepsäläinen (2022, p. 145), esta disposición facilita el seguimiento de las conversaciones, ya que permite visualizar de manera clara la relación entre un comentario y sus respuestas.

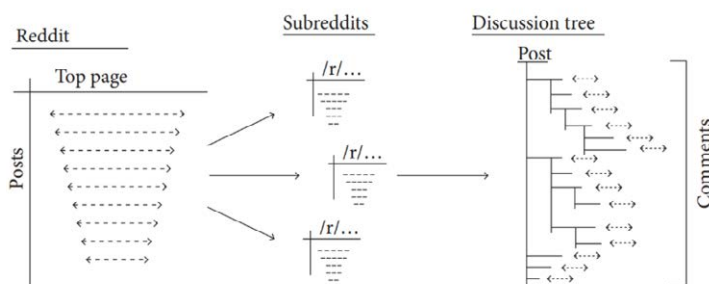


FIGURA 1. ESTRUCTURA DE REDDIT (MEDVEDEV ET AL. 2019, P.85).

Reddit sigue una política de contenido estricta basada en la *Reddiqueta* (de Reddit + etiqueta), similar a la *netiqueta* (Internet + etiqueta). Esta guía no oficial promueve el respeto, la participación constructiva y el cumplimiento de normas específicas de cada *subreddit*. Sus principios incluyen evitar *spam*, atribuir fuentes, votar con criterio y tratar a otros con respeto, recordando que detrás de cada comentario hay personas reales. Además, cada *subreddit* tiene reglas propias establecidas por moderadores, quienes

supervisan contenido, gestionan usuarios y pueden eliminar publicaciones que infrinjan la *Reddiqueta* o normas internas (Panek, 2021).

A diferencia de otras redes sociales, Reddit incorpora un sistema de votación que permite a los usuarios emitir votos positivos o negativos tanto sobre publicaciones como sobre comentarios. Según Lu (2024, p. 31), este mecanismo de votación equivale a expresar una actitud favorable o desfavorable hacia el contenido evaluado.

3. Situación general de investigación con Reddit

Proferes et al. (2021) analizan el uso de Reddit en la investigación científica, identificando 727 estudios en inglés entre 2010 y abril de 2020. Los resultados muestran un crecimiento significativo: de 4 estudios en 2010 a 230 en 2019. Sin embargo, solo 8 pertenecen a la lingüística, mientras que su uso más frecuente es en informática, ciencias médicas y sociología. Geográficamente, más de la mitad de los trabajos (427) se realizan en Norteamérica, 153 en Europa y 85 en Asia, lo que indica que Reddit sigue poco explorado en la lingüística y en Europa.

En el ámbito lingüístico, diversos estudios de la pragmática lingüística han explorado Reddit desde perspectivas y metodologías variadas. Los trabajos recientes han abordado fenómenos como los marcadores discursivos, la formulación de consejos, las dinámicas de identidad en comunidades digitales y la (des)cortesía, lo que evidencia la riqueza discursiva de esta plataforma. Aunque todos los trabajos revisados están redactados en inglés, algunos se centran en datos en lenguas distintas, como el griego (Teneketzi, 2022) o el rumano (Alecú, 2024), lo que indica un creciente interés de lingüistas de diferentes contextos culturales. A continuación, se revisan estudios sobre pragmática lingüística en Reddit.

En cuanto al uso de marcadores discursivos en Reddit, Alecú (2024) se enfoca en la función discursiva de tres préstamos del inglés —welp (‘well’, bueno), TBH (‘to be honest’, para ser honesto), y TBF (‘to be fair’, para ser justo)— como marcadores discursivos digitales en interacciones de *r/Romania*. Williams (2021), por su parte, investiga el uso del operador de marca registrada (™) en Reddit, comparándolo con el *hashtag* (#) en funciones como evaluación, afiliación y crítica.

Un segundo grupo de estudios se centra en la práctica del consejo en Reddit. Lu (2025, p. 250) afirma que el consejo es uno de los tipos de interacción más populares en la plataforma, lo que puede justificar el interés de los investigadores en este fenómeno. Lu (2024) explora la práctica de dar consejos en Reddit mediante un análisis de movimientos y modelos de *regresión de Poisson*, e identifica desalineaciones en los objetivos interactivos entre consejeros, receptores y observadores, lo que evidencia la complejidad contextual de esta práctica en línea. En la misma línea, el mismo autor (Lu, 2025) adopta una metodología experimental basada en escenarios para explorar cómo diferentes movimientos retóricos influyen en la percepción de los consejos por parte de distintas audiencias. A través de un enfoque metapragmático, este estudio revela la

importancia de considerar tanto a los interlocutores directos como a los participantes no ratificados (oyentes y observadores) en la evaluación del discurso de consejo. Además, Vepsäläinen (2022) examina las respuestas que no ofrecen consejo directo, como relatos personales o expresiones empáticas, y destaca su papel en la cohesión comunicativa y el apoyo mutuo dentro de la interacción en Reddit.

El análisis de la identidad y el duelo en esta red social también ha sido objeto de estudio. Schneider (2025) examina cómo los cuidadores de personas con demencia participan en discusiones sobre la muerte en *r/Dx*. A través del análisis de referencias deícticas y no deícticas, el estudio revela cómo los usuarios negocian su identidad, conceptualizan las transiciones temporales del duelo y redefinen su pertenencia a la comunidad incluso después del fallecimiento del ser querido. El *subreddit* funciona como una Comunidad de Práctica digital, donde se mantiene el vínculo emocional y discursivo post-pérdida. Asimismo, Signorelli (2022) analiza publicaciones en *r/FatPeopleStories* y se centra en el acto de *venting*, entendido como una forma de desahogo emocional, lo cual permite a los usuarios no solo compartir experiencias difíciles, sino también subvertir discursos sociales dominantes sobre el cuerpo y fortalecer el apoyo mutuo y la identidad grupal.

Otros trabajos se aproximan a fenómenos relacionados con el humor y la descortesía. Dynel y Poppi (2019) analizan el fenómeno *RoastMe*, originado en Reddit, como un *insulto ritual* enmarcado en el humor, y destacan su creatividad y descortesía simulada. En esta línea, Dynel (2020) estudia el humor vigilante en *r/IncelTears*, explorando su función crítica y comunitaria, aunque advierte dificultades metodológicas derivadas del anonimato y el *trolling*. A su vez, Teneketzi (2022) compara la descortesía en Reddit y YouTube en debates sobre los incendios de Ática, y concluye que YouTube favorece la descortesía directa, mientras que Reddit tiende a formas más sutiles e implícitas.

Finalmente, en el campo de la argumentación digital, Dayter y Messerli (2022) investigan el lenguaje persuasivo en *r/ChangeMyView*. Aunque no logran identificar diferencias sistemáticas en cuanto al nivel de formalidad entre publicaciones persuasivas y no persuasivas, destacan el valor del discurso basado en hechos y evidencia.

En resumen, estos estudios confirman la relevancia de Reddit como fuente de datos para la investigación pragmática, especialmente en lengua inglesa. Desde el humor y la descortesía hasta el lenguaje persuasivo y el acto de consejo, las investigaciones revisadas ponen de manifiesto la diversidad metodológica aplicada a esta plataforma, al tiempo que señalan su potencial y los desafíos que plantea para estudios futuros.

4. Potencial y defectos de Reddit como herramienta para investigaciones pragmáticas

Reddit, como Twitter, es una plataforma pública con una API abierta y usuarios que reaccionan rápidamente, lo que la hace valiosa para la investigación (Proferes et al., 2021). Chi y Chen (2023) destacan que el anonimato, la popularidad de Reddit, su organización en *subreddits*, la extensión de comentarios y publicaciones, la facilidad

para recopilar datos y el sistema de votos crean condiciones propicias para su uso como fuente de datos. Sin embargo, Lu (2024) señala que los estudios lingüísticos en Reddit siguen siendo limitados. Además, Proferes et al. (2021) advierten sobre desafíos como la diversidad de normas y culturas entre *subreddits*, el anonimato de los usuarios, el uso de cuentas temporales y la prevalencia de temas sensibles. En este contexto, la presente sección tiene como objetivo analizar el potencial y las limitaciones de Reddit como herramienta para la investigación, con un énfasis particular en los estudios pragmáticos.

4.1 Potencial de Reddit como herramienta para investigaciones pragmáticas

Para empezar, la accesibilidad de los datos en Reddit representa una ventaja clave para las investigaciones pragmáticas. Primero, sus contenidos son públicos, lo que permite el acceso libre a cualquier internauta (Proferes et al., 2021). Además, su API facilita la recopilación de datos (Chi y Chen, 2023), complementada por herramientas como PullPush Reddit API³, que optimizan la búsqueda avanzada mediante filtros por *subreddit*, fecha o usuario.

Segundo, el anonimato en Reddit permite a los usuarios expresarse con mayor sinceridad mientras protege su privacidad. A diferencia de Instagram y Facebook, donde predominan los nombres reales, en Reddit se utilizan apodos, generando una *cultura del anonimato* (Dyner, 2020). Esto fomenta la libertad de expresión, incluso en temas sensibles, facilitando la recopilación de datos difíciles de obtener en otras plataformas (Proferes et al., 2021). Además, protege la privacidad de los participantes, a la vez que reduce dilemas éticos en la investigación.

Tercero, la organización temática de Reddit facilita la filtración y clasificación de información, lo que representa una ventaja clave para la investigación. Esta organización se basa en *subreddits*, espacios dedicados a temas específicos, cada uno con su propio enfoque y normas (Lu, 2024). Por ejemplo, el *subreddit r/esGaming* es una comunidad española centrada en videojuegos, donde se prohíbe el uso de otros idiomas, lo que lo convierte en un espacio ideal para investigar discursos relacionados con videojuegos. De manera similar, *r/SpainPolitics* se especializa en discusiones políticas, ofreciendo una fuente relevante para investigaciones pragmáticas sobre discurso político. Además, muchos países hispanohablantes cuentan con *subreddits* propios, como *r/AskArgentina*, *r/mexico* o *r/PERU*, formados principalmente por usuarios locales. Estas comunidades brindan un contexto valioso para explorar las variedades del español y analizar fenómenos pragmáticos en contextos culturales específicos.

Cuarto, los comentarios y publicaciones en Reddit se caracterizan por su extensión y la posibilidad de ofrecer contextos completos y perspectivas diversas, lo que resulta especialmente valioso para estudios pragmáticos del discurso digital. A diferencia de

3. PullPush Reddit API ofrece un servicio para la indexación y recuperación de contenido enviado a Reddit por sus usuarios.

otras redes sociales, donde las interacciones suelen ser breves e informales, Reddit permite comentarios y publicaciones de hasta 40.000 caracteres, lo que favorece debates detallados y estructurados (Proferes et al., 2021). Asimismo, las normas de la plataforma promueven el uso correcto de gramática y ortografía. En la *Reddiqueta*, Reddit (2024, s.p.) recomienda: “asegúrate de [...] usar las reglas de gramática y ortografía correctas. El discurso inteligente requiere un sistema estándar de comunicación. Disponte a aplicar correcciones sutiles”. En conjunto, el formato extensivo y normado de Reddit le confiere un valor particular para el estudio lingüístico de las interacciones en redes sociales.

Cabe señalar que las ventajas mencionadas anteriormente son aplicables no solo a la investigación pragmática, sino también a otras ramas de la lingüística. No obstante, las dos ventajas que se presentan a continuación son específicas del análisis pragmático.

La primera de estas ventajas es que el sistema de votos en Reddit constituye un recurso valioso para los estudios pragmáticos, al permitir observar cómo los usuarios evalúan enunciados según su adecuación comunicativa. Aunque los contenidos de la plataforma son de acceso libre, solo los usuarios registrados pueden emitir votos positivos o negativos. Según Lu (2024), estos votos reflejan las actitudes de los usuarios hacia publicaciones y comentarios: por lo general, se otorgan votos positivos a aquellos contenidos que generan acuerdo o simpatía, y negativos a los que se perciben como inapropiados o con los que se discrepa. Así, el sistema de votos actúa como una *paralinguistic digital affordance* que comunica juicios o emociones de manera no verbal (Lu, 2024). Además, influye en las estrategias lingüísticas empleadas por los participantes, como el uso de mensajes alentadores o experiencias compartidas (Panek, 2021; Teneketzi, 2021). En este sentido, insistimos en que este sistema tiene un gran potencial para la aplicación de la teoría del trabajo relacional⁴ (*relational work*) de Locher y Watts (2005) que ofrece una perspectiva novedosa en el estudio de la cortesía.

El trabajo relacional, una noción poco explorada en el ámbito hispanico, surge como revisión crítica de la propuesta de Brown y Levinson (1987). Locher y Watts (2005) utilizan esta noción para referirse al trabajo comunicativo que las personas realizan para negociar, construir, mantener o redefinir sus relaciones interpersonales durante la interacción. Asimismo, estos autores entienden la comunicación como un *continuo* que abarca conductas verbales y no verbales, y proponen que las conductas comunicativas pueden ser marcadas o no marcadas. Las no marcadas responden a expectativas de los interlocutores, por lo que se perciben como neutras (ni corteses ni descorteses), diplomáticas y apropiadas. En cambio, las marcadas pueden ser positivas (corteses) o negativas (descorteses o *hipercorteses*), y su evaluación depende del juicio de los destinatarios.

4. La terminología empleada relacionada con el trabajo relacional en esta investigación (por ejemplo, “*apropiado*”, “*trabajo relacional*”, “*diplomático*”) sigue las traducciones propuestas por Albelda Marco (2021), quien introduce por primera vez en español la teoría del *relational work* formulada por Locher y Watts (2005).

Diversos autores han señalado limitaciones en la aplicación de esta teoría. Por ejemplo, Mills (2011) destaca la dificultad para identificar los juicios de los interlocutores, mientras que Culpeper (2008) y Hu (2021) subrayan la complejidad derivada de las diferencias en normas contextuales, estructuras cognitivas y experiencias individuales, que dificultan establecer un marco normativo común.

A pesar de estas dificultades, Reddit ofrece un contexto propicio para su aplicación. Por un lado, los votos reflejan juicios concretos y observables sobre los enunciados. El balance de votos (diferencia entre positivos y negativos) puede interpretarse como un indicador de evaluación comunicativa. Así, un balance positivo sugiere una conducta marcada positivamente, es decir, percibida como apropiada, cortés y diplomática; mientras que un balance negativo implica una conducta marcada negativamente, valorada como inapropiada y poco diplomática, asociada a la descortesía o *hipercortesía*.

Por otro lado, la existencia de la *Reddiquette* y de normas en cada *subreddit* establecen un marco normativo compartido que facilita la emisión de juicios consistentes. En este contexto, es posible inferir que los contenidos con balance positivo son percibidos como apropiados y corteses, mientras que aquellos con balance negativo se interpretan como descorteses o *hipercorteses*, conforme a la teoría del trabajo relacional.

Por último, varias características de la interacción en Reddit ofrecen condiciones favorables para investigaciones pragmáticas sobre fenómenos específicos. En primer lugar, el predominio del humor en los intercambios comunicativos (Tenekezi, 2021) convierte a esta plataforma en una fuente valiosa para estudios sobre el humor desde una perspectiva pragmática (Dynel, 2020). En segundo lugar, la abundancia de interacciones centradas en consejos y recomendaciones (Vepsäläinen, 2022; Lu, 2025) facilita el análisis de este tipo de dinámicas comunicativas. En tercer lugar, Reddit permite explorar procesos de construcción de identidad, ya que el anonimato y la presencia de narrativas personales ofrecen datos significativos para observar cómo los usuarios negocian su pertenencia y posición en comunidades específicas (Signorelli, 2022; Schneider, 2025). Finalmente, esta plataforma es reconocida como un espacio en el que emergen nuevos fenómenos lingüísticos (Dynel y Poppi, 2019), lo que la hace especialmente relevante para investigaciones centradas en innovaciones del uso del lenguaje en entornos digitales.

4.2 Defectos de Reddit como herramienta para investigaciones pragmáticas

Pese a las ventajas previamente señaladas, Reddit presenta ciertas limitaciones que pueden afectar las investigaciones lingüísticas, especialmente en el ámbito pragmático. Para comenzar, aunque el anonimato en Reddit beneficia las investigaciones pragmáticas, también presenta desafíos, especialmente en estudios socioculturales. Por un lado, la falta de datos personales dificulta la recopilación de información sociocultural, lo que supone un obstáculo importante. Por otro, Dynel (2020) señala que el anonimato complica la identificación de la autenticidad, el origen y las intenciones de los contenidos, facilitando el *trolling* y la difusión de interacciones engañosas o paródicas.

En segundo lugar, las comunidades en Reddit son relativamente autónomas, con reglas y culturas propias, lo que genera desafíos adicionales. Los *subreddits* operan como foros independientes con normativas específicas impuestas por moderadores y esto limita la autenticidad de los hallazgos lingüísticos. Por lo tanto, los investigadores enfrentan dos retos, deben familiarizarse con la dinámica de cada comunidad y emplear metodologías adaptadas a cada *subreddit* (Proferes et al., 2021).

Por último, aunque desde una perspectiva netnográfica (*netnography*⁵) el anonimato ha sido interpretado como una estrategia generalizada para la protección de la privacidad en entornos digitales (Kozinets y Gambetti, 2021), también plantea desafíos importantes para la investigación académica. Esta característica no es exclusiva de Reddit, sino común a muchas plataformas sociales, donde la falta de datos demográficos puede dificultar la identificación de usuarios vulnerables, como menores de edad (Proferes et al., 2021), lo que genera posibles riesgos éticos. Además, la ausencia de directrices estandarizadas sobre el uso de datos de plataformas como Reddit complica la garantía de una protección adecuada de la privacidad de los participantes (Fiesler et al., 2024). Dado que las cuestiones éticas constituyen un eje fundamental en la investigación con estos datos, la siguiente sección se centrará en analizar estas consideraciones en mayor profundidad, con especial énfasis en el caso de Reddit.

5. Consideraciones éticas en las investigaciones con redes sociales

La consideración ética es un aspecto ineludible en las investigaciones que utilizan datos de redes sociales. Sin embargo, esta consideración resulta particularmente compleja, ya que, como señala López Escarcena (2021, p. 242), la mayoría de los principios éticos tradicionales son formulados para contextos presenciales, no digitales. En la actualidad, aún no existe un marco normativo común ni un conjunto de principios organizadores que aborden adecuadamente los desafíos éticos vinculados al estudio de redes sociales.

En el ámbito lingüístico, López Escarcena (2021) identifica dos factores clave en esta problemática: la ambigüedad entre lo público y lo privado, y la falta de consenso sobre el consentimiento informado. Si bien gran parte de los contenidos en estas plataformas son de acceso público, Willis (2017) destaca la coexistencia de espacios públicos y privados, evidenciada en la presencia de conversaciones privadas dentro de entornos abiertos. Además, los usuarios suelen esperar cierto grado de protección de sus datos, por ello, en algunos casos recurren al anonimato para preservar su privacidad (Kozinets y Gambetti, 2021). Adicionalmente, al compartir contenido, no pueden prever con certeza si este será recopilado con fines de investigación (Kantanen y Manninen, 2016).

5. La netnografía (en inglés *netnography*) es una metodología cualitativa que adapta técnicas etnográficas al estudio de comunidades y culturas en línea, en respuesta al auge de Internet y los nuevos patrones de interacción digital.

Además, el consentimiento informado también plantea desafíos significativos. Tradicionalmente, se considera esencial obtenerlo de todos los participantes (Sveningsson, 2004), pero en redes sociales esto resulta complejo debido al volumen y dispersión de los usuarios. Willis (2017) propone eximir de este requisito cuando los datos sean públicos y de naturaleza textual, postura que comparten los investigadores de AoIR⁶ (Franzke et al., 2020). No obstante, la ambigüedad entre lo público y lo privado dificulta determinar si los datos cumplen este criterio, mientras que la multimodalidad de las plataformas añade otra capa de complejidad. Hasta la fecha, no se ha alcanzado un consenso sobre cómo abordar el consentimiento informado en investigaciones basadas en datos de redes sociales.

5.1 Consideraciones éticas en las investigaciones con Reddit

En el caso de Reddit, la ética sigue siendo un tema de debate sin acuerdo generalizado. Fiesler et al. (2024) revisan 184 estudios que utilizan datos de la plataforma y revelan discrepancias en aspectos como la discusión ética, la naturaleza pública de los datos, la revisión ética, el consentimiento informado y la protección de la privacidad de los usuarios.

En la misma línea, Adams (2024) identifica varias consideraciones éticas clave en investigaciones con datos de Reddit: el consentimiento y autonomía de los participantes, el contexto de la investigación, la vulnerabilidad de los usuarios y el riesgo de identificación. En primer lugar, aunque el acuerdo de usuario de Reddit permite la recopilación de datos, acceder a publicaciones sin el consentimiento explícito de los usuarios puede contravenir sus expectativas de privacidad y eliminar su control sobre el contenido. Incluso cuando los usuarios eliminan sus publicaciones, estas pueden seguir en posesión de los investigadores, lo que plantea riesgos éticos y posibles daños. Además, muchas comunidades de Reddit se identifican como *espacios seguros*, por lo que extraer datos de ellas puede descontextualizar publicaciones y exponer a usuarios vulnerables fuera de su entorno de confianza sin su consentimiento. Por último, el uso de citas textuales en investigaciones académicas puede facilitar búsquedas inversas en internet, comprometiendo el anonimato de los usuarios y permitiendo vincular comentarios con sus autores originales, lo que representa una amenaza directa a su privacidad (Adams, 2024, pp. 51-55).

Ante esta situación, proponemos algunas recomendaciones éticas para investigaciones con datos de Reddit. En primer lugar, es fundamental preservar el anonimato de los usuarios, tal como se sugiere también para otras redes sociales (Mancera Rueda y Pano Alamán, 2014). Para ello, puede aplicarse la estrategia del *disfraz leve* de López Escarcena (2021), que consiste en reemplazar los nombres de usuario por seudónimos.

6. La Asociación de Investigadores de Internet (AoIR) es una organización académica dedicada al avance del campo interdisciplinario de los estudios sobre Internet.

Esta medida no solo protege la identidad, sino que permite reconocer intervenciones en contextos de coconstrucción narrativa. Dado que obtener el consentimiento informado de todos los *redditors* es inviable, se recomienda solicitar autorización a los moderadores, especialmente en comunidades con usuarios vulnerables. Además, para evitar la identificación de estos usuarios a través de citas textuales, se aconseja recurrir a la cita indirecta en investigaciones sobre temas sensibles.

6. Conclusiones

El presente estudio ha examinado el potencial de los datos extraídos de Reddit para investigaciones pragmáticas, destacando las características de la plataforma y realizando una revisión bibliográfica. Aunque Reddit es una plataforma popular, sus datos han sido poco aprovechados en el ámbito pragmático, con estudios limitados que abordan temas como el humor y la (des)cortesía y ninguno de ellos son sobre el español. Sostenemos que Reddit es una plataforma ideal para investigaciones pragmáticas debido a la accesibilidad de los datos, la organización temática, la extensión de los textos y su sistema de votación, entre otros factores. Sin embargo, dado que no existe un consenso claro sobre las consideraciones éticas en el uso de estos datos, es esencial adoptar medidas como mantener el anonimato de los usuarios para proteger la privacidad de ellos. Consideramos que este trabajo contribuye a visibilizar el valor de Reddit como recurso para investigaciones pragmáticas y a promover un uso ético y efectivo de la plataforma en futuros estudios.

Referencias

- Adams, Nicholas Norman (2024). “Scraping” reddit posts for academic research? addressing some blurred lines of consent in growing internet-based research trend during the time of covid-19. *International Journal of Social Research Methodology*, 27(1), 47–62.
- Albelda Marco, Marta (2021). “Las relaciones entre el hablante y el oyente: La cortesía verbal”. En O. Loureda Lamas y A. Schrott (Eds.), *Manual de lingüística del hablar* (pp. 221–242). De Gruyter.
- Alecu, Bianca (2024). “Digital discourse markers in Romanian forum communication: The case of welp, TBH (to be honest) and TBF (to be fair)”. En C. M. Popescu y O. A. Duță (Eds.), *Discourse markers in romance languages. Crosslinguistic approaches in romance and beyond* (pp. 185–205). Peter Lang.
- Chi, Yu y Chen, Huai-yu (2023). Investigating substance use via reddit: Systematic scoping review. *Journal of Medical Internet Research*, 25(1).

- Culpeper, Jonathan (2008). "Reflections on impoliteness, relational work and power". En D. Bousfield, y M. A. Locher (Eds.), *Impoliteness in language: Studies on its interplay with power in theory and practice* (pp. 22–53). Mouton de Gruyter.
- Dayter, Daria y Messerli, Thomas C. (2021). Persuasive language and features of formality on the r/ChangeMyView subreddit. *Internet Pragmatics*, 5 (1), 165-195.
- Dynel, Marta (2020). Vigilante disparaging humour at r/IncelTears: Humour as critique of incel ideology. *Language & Communication*, 74, 1–14.
- Dynel, Marta y Poppi, Fabio I.M. (2019). Risum teneatis, amici?: The socio-pragmatics of RoastMe humour. *Journal of Pragmatics*, 139, 1–21.
- Fiesler, Casey, Zimmer, Michael, Proferes, Nicholas, Gilbert, Sarah y Jones Naiyan (2024). Remember the human: A systematic review of ethical considerations in reddit research. *Proceedings of the ACM on Human-Computer Interaction*, 8(group), 1–33.
- Franzke, Aline Shakti, Bechmann, Anja, Zimmer, Michael, Ess, Charles y the Association of Internet Researchers (2020). *Internet Research: Ethical Guidelines 3.0*. <https://aoir.org/reports/ethics3.pdf>.
- Glassman, Michael. y Kuznetcova, Irina (2022). The GameStop saga: Reddit communities and the emerging conflict between new and old media. *First Monday*, 27(7).
- Hu, Yanfei (2021). Relational work in dispute: Negotiating norms, negotiating relationships. *Journal of Politeness Research*, 17(2), 133–154.
- Kantanen, Helena y Manninen, Jyri (2016). Hazy boundaries: Virtual communities and research ethics. *Media and Communication*, 4(4), 86–96.
- Kozinets, Robert V. y Gambetti, Rossella (2021). *Netnography unlimited: Understanding technoculture using qualitative social media research*. Routledge
- Leuckert, Sven y Leuckert, Martin (2020) "Towards a digital sociolinguistics: Communities of practice on reddit". En S. Rüdiger y D. Dayter (Eds.), *Corpus approaches to social media* (pp. 15–40). John Benjamins Publishing Company
- Locher, Miriam A. y Watts, Richard J. (2005). Politeness theory and relational work. *Journal of Politeness Research*, 1, 9–33.
- López Escarcena, Ignacio (2021). Desafíos éticos en el estudio de comunidades digitales: Una propuesta basada en un caso empírico. *Cuadernos. Info*, 49, 239–259.
- Lu, Rickey (2024). Audience design and pragmatic conceptions of moves and upvotes during advice-giving on reddit. *Journal of Pragmatics*, 219, 30–47.
- Lu, Rickey (2025). Understanding online advice-giving evaluations through the politeness evaluation model. *Journal of Politeness Research*, 21(1), 249–287.
- Mancera Rueda, Ana. y Pano Alamán, Ana. (2014). "Las redes sociales como corpus de estudio para el análisis del discurso mediado por ordenador". En S. López Poza, y N. Pena Sueiro (Eds.), *Humanidades Digitales: desafíos, logros y perspectivas de futuro* (pp. 305-315). Universidade da Coruña, SIELAE.
- Medvedev, Alexey N., Lambiotte, Renaud y Delvenne Jean-Charles (2019). "The anatomy of reddit: An overview of academic research". En F. Ghanbarnejad, R.

- Rishiraj Saha, F. Karimi, J.C. Delvenne, y M. Bivas(Eds.), *Dynamics on and of complex networks: III machine learning and statistical physics approaches* (pp.183-204). Springer.
- Mills, Sara (2011). "Discursive approaches to politeness and impoliteness". En Linguistic Politeness Research Group (Ed.), *Discursive approaches to politeness* (pp.19-56). Mouton de Gruyter.
- Panek, Elliot T. (2021). *Understanding reddit*. Routledge.
- Proferes, Nicholas, Jones, Nicholas, Gilbert, Sarah, Fiesler, Casey y Zimmer, Michael (2021). Studying reddit: A systematic overview of disciplines, approaches, methods, and ethics. *Social media + society*, 7(2), 1–14.
- Reddit. (2024). *Reddit inc homepage*. <https://redditinc.com>
- Schneider, Carolin (2025). Deictic and non-deictic reference at the online-offline nexus: Discussing death on reddit. *Discourse, Context and Media*, 66.
- Signorelli, Julia (2022). Of pumpkin spice lattes and hamplanets: The venting genre as support and subversion on reddit's r/FatPeopleStories. *Journal of Language Aggression and Conflict*, 10(1), 64–84.
- Statista (2025). *Most popular websites worldwide as of November 2024, by total visits* <https://www.statista.com/statistics/1201880/most-visited-websites-worldwide/>.
- Sveningsson, Malin (2004). "Ethics in internet ethnography". En E.A. Buchanan, (Ed.), *Readings in virtual research ethics: Issues and controversies* (pp.45-61). Idea Group Publishing.
- Teneketzi, Korallia (2022). Impoliteness across social media platforms: A comparative study of conflict on YouTube and reddit. *Journal of Language Aggression and Conflict*, 10 (1), 38–63.
- Vepsäläinen, Heidi (2022). Responses to advice-seeking on reddit that do not give advice. *Journal of Pragmatics*, 191, 143–155.
- Williams, Emily A.E. (2021). Pragmatic extension in computer-mediated communication: The case of “#” and “™”. *Journal of Pragmatics*, 181, 165–179.
- Willis, Roxana (2017). Observations online: Finding the ethical boundaries of Facebook research. *Research Ethics*, 15(1), 1–17.