

¿Evolución o estancamiento? Un análisis de la evolución de los datos en abierto sobre drogodependencias depositados en *Figshare* en los años 2014 y 2023

Lydia García-Gómez

Universitat de València

<https://orcid.org/0009-0003-1719-8623>

lydia.garcia@uv.es

Yiming Liu

Universitat de València

<https://orcid.org/0009-0006-3592-0175>

yiming.liu@uv.es

Antonio Vidal-Infer

Universitat de València

<https://orcid.org/0000-0002-7860-8652>

antonio-vidal.infer@uv.es

García-Gómez, Lydia; Liu, Yiming; Vidal-Infer, Antonio (2026). “¿Evolución o estancamiento? Un análisis de la evolución de los datos en abierto sobre drogodependencias depositados en *Figshare* en los años 2014 y 2023”. En: Orduña-Malea, Enrique; Olea, Isabel (coords). *El valor de las revistas científicas* (pp. 140-144). Ediciones Profesionales de la Información. <https://doi.org/10.3145/vrc.18>

Resumen

El depósito de datasets sobre drogodependencias en repositorios de acceso abierto es escaso y, en muchos casos, no responde a los estándares de datos brutos de investigación. Este trabajo analiza la evolución del uso, formato y visibilidad de los registros depositados en *Figshare* en los años 2014 y 2023, así como su correspondencia con datos reales de investigación. Se identificaron 1.002 registros mediante búsqueda en campos de título y palabras clave, de los cuales se seleccionaron 574 tras una selección manual. Se analizaron variables como tipo de archivo, contenido (tablas, texto, gráficos) y métricas bibliográficas (visualizaciones, descargas y citas) extraídas de *PlumX* y *Scopus*. Los resultados revelan una caída significativa en visualizaciones y citas entre ambos años, así como una proporción muy baja de datasets que puedan considerarse datos en bruto. Estos resultados obtenidos ponen de manifiesto limitaciones en la indexación y aprovechamiento de los datos abiertos en adicciones, y refuerzan la necesidad de mejorar los procesos de formación, catalogación y uso efectivo de los repositorios científicos.

Palabras clave

Dataset; Datos brutos; Drogodependencias; Repositorio; *Figshare*; Métricas.

Abstract

The deposit of datasets on drug dependence in open access repositories is limited and, in many cases, does not meet the standards of raw research data. This study analyses the evolution of usage, format, and visibility of records deposited in *Figshare* in the years 2014 and 2023, as well as their alignment with actual research data. A total of 1,002 records were identified through a search of title and keyword fields, of which 574 were selected after manual screening. Variables such as file type, content (tables, text, graphs), and bibliometric indicators (views, downloads, and citations) from *PlumX* and *Scopus* were analysed. The results reveal a significant drop in views and citations between the two years, along with a very low proportion of datasets that can be considered raw data. These findings highlight limitations in the indexing and effective use of open data in the field of addiction research and underscore the need to improve training, cataloging, and the strategic use of scientific repositories.

Keywords

Dataset; Raw data; Drug dependence; Repository; *Figshare*; Metrics.

1. Introducción

La adopción de la ciencia abierta ha generado un cambio profundo en la manera de entender la producción y difusión de la evidencia científica. Este paradigma, basado en los principios de transparencia, reproducibilidad y accesibilidad, promueve el libre acceso no sólo a publicaciones, sino también a los datos que sustentan los hallazgos. En el ámbito de las adicciones, esta transición resulta especialmente relevante, dado que la complejidad y multidimensionalidad del fenómeno exige evidencia verificable, reutilizable y accesible para diseñar intervenciones eficaces y sostenibles (Scheibein *et al.*, 2022; Eben *et al.*, 2023).

Pese a estos avances, múltiples estudios han revelado que la cultura de los datos abiertos sigue en una fase inicial dentro de la investigación en drogodependencias. Por ejemplo, Vassar *et al.* (2020) identificaron que ninguno de los ensayos clínicos aleatorizados sobre adicciones analizados compartía sus datos públicamente, incluso cuando habían sido financiados con fondos públicos. Esta reticencia se asocia, entre otros factores, a la ausencia de incentivos institucionales, al desconocimiento sobre las plataformas adecuadas y al temor a la pérdida de autoría o malinterpretación de los datos (Aleixandre-Benavent *et al.*, 2020). En ese sentido, la adopción de prácticas de ciencia abierta en investigación sobre adicciones sigue en su infancia (Eben *et al.*, 2023).

Dentro de este ecosistema, *Figshare* se ha consolidado como uno de los repositorios internacionales más empleados para el depósito de material complementario y conjuntos de datos. Sin embargo, tal como han apuntado estudios recientes, el hecho de etiquetar un archivo como “dataset” no garantiza que cumpla los criterios de datos brutos reutilizables ni los principios FAIR (*Findable, Accessible, Interoperable, Reusable*) (Sixto-Costoya *et al.*, 2025). Esta falta de estandarización y control de calidad limita el potencial de los repositorios como fuente de datos abiertos útiles para la investigación en salud pública.

El presente trabajo analiza la evolución del uso, la visibilidad y la clasificación de registros relacionados con drogodependencias en *Figshare* en los años 2014 y 2023.

2. Metodología

Para la realización de este trabajo se analizaron los registros depositados en el repositorio *Figshare* en los años 2014 y 2023, que contenían información relacionada con las drogodependencias. Para su identificación, la estrategia de búsqueda que se utilizó fue:

:title: Alcohol OR :title: Cocaine OR :title: Cannabis OR :title: Marihuana OR :title: Marijuana

Esta estrategia proporcionó un total de 1.002 registros, de los cuales se seleccionaron 574 tras una depuración manual que excluyó documentos no pertinentes o duplicados. A partir de los 574 registros resultantes, para ofrecer un diagnóstico sobre el estado actual del depósito de datos abiertos en adicciones y señalar las carencias estructurales que impiden su aprovechamiento, se examinaron variables como:

- Tipo de archivo (.xls, .docx, entre otros);
- contenido (tablas, texto, gráficos);
- clasificación como datos brutos: evaluada manualmente según criterios de apertura, estructura y reutilización (Vidal-Infer *et al.*, 2019);
- métricas de impacto a través de *PlumX* (visualizaciones, descargas, citas) y *Scopus* (citas).

El análisis estadístico se realizó con el software *IBM SPSS*, versión 29.0 (*IBM Corp.*, 2022), utilizando pruebas de comparación de medias para muestras independientes.

3. Resultados

Los tipos de archivo más frecuentes fueron hojas de cálculo en formatos xls y xlsx (294 registros) y de procesador de texto en formatos doc y docx (118). En cuanto al contenido, predominan las tablas (252), seguidas de textos (73) y gráficos (11), mientras que fueron identificados como datos brutos de investigación un 4,35% de los registros en 2014 y un 0,29% en 2023.

En términos de impacto, los registros de 2014 mostraron una media de 140,23 visualizaciones en *PlumX*, frente a las 15,55 de los registros de 2023 ($p = 0,0019$). En cuanto a las citas en *PlumX*, no se hallaron diferencias significativas entre ambos años (2014: media = 77,06; 2023: media = 3,34; $p = 0,1611$). Respecto a los datos recogidos en *Scopus*, los registros de 2014 presentaron una media de 42,04 citas, significativamente superior a las 3,54 citas de 2023 ($p < 0,0001$). No se observaron diferencias sustanciales en relación con otros indicadores de impacto.

4. Discusión y conclusión

El análisis realizado muestra que los formatos predominantes entre los registros depositados en *Figshare* sobre drogodependencias son hojas de cálculo y documentos de texto, lo cual en principio facilitaría la reutilización. Sin embargo, la disponibilidad de archivos en formatos adecuados no implica necesariamente la existencia de datos brutos o estructurados, tal y como se describe en estudios anteriores (*Vidal-Infer et al.*, 2019). Esta discordancia entre formato y utilidad limita el potencial real de los registros para ser reaprovechados por otros investigadores.

En cuanto al contenido de los archivos, aunque predominan las tablas, en muchos casos estas aparecen descontextualizadas o sin estructura reutilizable, lo que impide su uso como *datasets* propiamente dichos. Tal como subrayan *Alexandre-Benavent et al.* (2020), los investigadores tienden a compartir archivos que, si bien se etiquetan como conjuntos de datos, no cumplen con los requisitos mínimos para su análisis secundario, como ocurre con documentos cerrados.

La clasificación manual realizada en este estudio confirma esta tendencia: sólo un 4,35% de los archivos de 2014 y un 0,29% de los de 2023 pueden considerarse datos brutos reales. Este dato se alinea con los resultados de *Sixto-Costoya et al.* (2025), quienes hallaron que menos del 30% de los más de 5.900 conjuntos de datos evaluados en el ámbito de las adicciones cumplían los principios FAIR. *Scheibein et al.* (2022) también destacan que, aunque la ciencia abierta ha sido ampliamente promovida en el campo, su adopción sigue en una fase inicial.

Las métricas de impacto evidencian diferencias importantes entre los años analizados: los registros de 2014 presentan una media significativamente superior tanto en visualizaciones como en citas, en comparación con los de 2023. Esta caída puede atribuirse a varios factores, entre ellos la falta de estandarización en la catalogación, la escasa calidad de los contenidos o un uso incorrecto del etiquetado “dataset”, lo que afecta negativamente a su visibilidad y aprovechamiento posterior.

Desde una perspectiva bibliométrica, estos resultados son coherentes con lo señalado por *Catone* (2023), quien advierte que el aumento cuantitativo de la producción científica en *open*

data no siempre implica una mejora en su accesibilidad o calidad. En conjunto, los hallazgos de este trabajo ponen de relieve una brecha persistente entre las políticas de ciencia abierta y su implementación real en el ámbito de las adicciones.

Cerrar esta brecha requiere reforzar la formación en buenas prácticas de datos abiertos, mejorar los mecanismos de catalogación en los repositorios y fomentar una cultura científica que priorice no sólo la publicación, sino también la calidad, transparencia y reutilización efectiva de los datos compartidos.

5. Referencias

Aleixandre-Benavent, Rafael; Vidal-Infer, Antonio; Alonso-Arroyo, Adolfo; Peset, Fernanda; Ferrer-Sapena, Antonia (2020). "Research data sharing in Spain: Exploring determinants, practices, and perceptions". *Data*, v. 5, n. 2.

<https://doi.org/10.3390/data5020029>

Eben, Charlotte; Bóthe, Beáta; Brevers, Damien; Clark, Luke; Grubbs, Joshua B.; Heirene, Robert; Kräplin, Anja; Lewczuk, Karol; Palmer, Lucas; Perales, Jose C.; Peters, Jan; Van Holst, Ruth J.; Billieux, Joël (2023). "The landscape of open science in behavioral addiction research: Current practices and future directions". *Journal of Behavioral Addictions*, v. 12, n. 4, pp. 862–870.

<https://doi.org/10.1556/2006.2023.00052>

IBM Corp. (2022). *IBM SPSS Statistics for Windows, Version 29.0*. Armonk, NY: IBM Corp

Scheibein, Florian; Donnelly, William; Wells, John S.G. (2022). "Assessing open science and citizen science in addictions and substance use research: A scoping review". *International journal of drug policy*, v. 100.

<https://doi.org/10.1016/j.drugpo.2021.103505>

Sixto-Costoya, Andrea; Ferrer-Sapena, Antonia; Aleixandre-Benavent, Rafael; Peset, Fernanda; Valderrama-Zurián, Juan C.; Petrosyan, Luiza (2025). "The compliance to FAIR principles of shared data in addiction research". *Scientometrics*, v. 130, pp. 763–779.

<https://doi.org/10.1007/s11192-024-05227-5>

Vassar, Matt; Jellison, Sam; Wendelbo, Hannah; Wayant, Cole (2020). "Data sharing practices in randomized trials of addiction interventions". *Addictive behaviors*, v. 102.

<https://doi.org/10.1016/j.addbeh.2019.106193>

Vidal-Infer, Antonio; Aleixandre-Benavent, Rafael; Lucas-Domínguez, Rut; Sixto-Costoya, Andrea (2019). "The availability of raw data in substance abuse scientific journals". *Journal of substance use*, v. 24, n. 1.

<https://doi.org/10.1080/14659891.2018.1489905>