

Resource Summary Report

Generated by SPARC SAWG on Aug 9, 2026

OpenRefine

RRID:SCR_021305

Type: Tool

Proper Citation

OpenRefine (RRID:SCR_021305)

Resource Information

URL: <https://openrefine.org/>

Proper Citation: OpenRefine (RRID:SCR_021305)

Description: Open source software Java tool for working with data. Used for loading data, cleaning up data, transforming data from one format into another and extending it with web services and external data.

Synonyms: openrefine, Google Refine

Resource Type: software resource, software application, data processing software

Defining Citation: [DOI:10.3163/1536-5050.101.3.020](https://doi.org/10.3163/1536-5050.101.3.020)

Keywords: Wikidata, data wrangling, loading data, cleaning up data, transforming data

Availability: Free, Available for download, Freely available

Resource Name: OpenRefine

Resource ID: SCR_021305

Alternate IDs: Wikidata_Q5583871

Alternate URLs: <https://www.wikidata.org/wiki/Q5583871>,
<https://github.com/OpenRefine/OpenRefine>

License: BSD licenses

Record Creation Time: 20220129T080354+0000

Record Last Update: 20260808T190327+0000

Ratings and Alerts

No rating or validation information has been found for OpenRefine.

No alerts have been found for OpenRefine.

Data and Source Information

Source: [SciCrunch Registry](#)

Usage and Citation Metrics

We found 50 mentions in open access literature.

Listed below are recent publications. The full list is available at [SPARC SAWG](#).

Nge FJ, et al. (2025) Polyploidy linked with species richness but not diversification rates or niche breadth in Australian Pomaderreae (Rhamnaceae). *Annals of botany*, 135(3), 531.

Grilo C, et al. (2025) Global Roadkill Data: a dataset on terrestrial vertebrate mortality caused by collision with vehicles. *Scientific data*, 12(1), 505.

Xu J, et al. (2025) Exploring the nanobody patent landscape: a focus on BCMA sequences and structural analysis. *mAbs*, 17(1), 2560893.

Ceccarelli S, et al. (2025) Triatomines outside the Americas: a comprehensive dataset for the global surveillance of Chagas disease vectors. *GigaByte (Hong Kong, China)*, 2025, gigabyte163.

Hart LV, et al. (2025) The stoneflies (Plecoptera) of Arkansas: a checklist compiled from museum specimen data. *Biodiversity data journal*, 13, e145637.

Yao Z, et al. (2025) Valence-based biases in collective temporal thought: The role of question framing, culture, and age. *Memory & cognition*, 53(6), 1738.

Farias GB, et al. (2025) The state of the art in Blue Amazon biodiversity: Protocol for a systematic review. *MethodsX*, 15, 103561.

Skubera M, et al. (2025) International initiatives to enhance awareness and uptake of open research in psychology: a systematic mapping review. *Royal Society open science*, 12(3), 241726.

Fernández-Gómez RA, et al. (2025) Understanding the role of ecological divergence in the evolution of isolated populations in the *Arremonops rufivirgatus* species complex across Mesoamerica. *BMC ecology and evolution*, 25(1), 34.

Nge FJ, et al. (2025) Diversification patterns of the southwest Australian biodiversity hotspot reveal a novel macroevolutionary pathway to plant hyperdiversity. *The New phytologist*, 248(2), 953.

Romero-Martínez C, et al. (2025) Cardiac Simulator Technologies and Design for Medical Education and Auscultation Training: A Systematic Review. *Bioengineering (Basel)*,

Switzerland), 12(7).

Chingagwe BT, et al. (2025) Protected Areas Are Important for the Conservation of *Disa engleriana*, an Edible Orchid in Malawi. *Ecology and evolution*, 15(7), e71778.

Yim MW, et al. (2024) A comprehensive synthesis of dung beetle records (Coleoptera, Scarabaeidae, Scarabaeinae) from Sabah, Malaysia. *Biodiversity data journal*, 12, e126697.

Wood TJ, et al. (2024) The InBIO Barcoding Initiative Database: DNA barcodes of Iberian Bees. *Biodiversity data journal*, 12, e117172.

Maspeke PNS, et al. (2024) A bibliometric analysis of ohmic heating on food processing in the last two decades. *Heliyon*, 10(20), e39315.

De Smedt S, et al. (2024) ?Ten lessons learned from the mass digitisation of a herbarium collection. *PhytoKeys*, 244, 23.

Hertz MI, et al. (2024) Eleven quick tips for properly handling tabular data. *PLoS computational biology*, 20(11), e1012604.

Potapov AM, et al. (2024) Global fine-resolution data on springtail abundance and community structure. *Scientific data*, 11(1), 22.

Vynokurov D, et al. (2024) Ukrainian Plant Trait Database: UkrTrait v. 1.0. *Biodiversity data journal*, 12, e118128.

Tejan YS, et al. (2024) The association between non-communicable diseases and COVID-19 severity and mortality among infected hospitalized healthcare workers in 29 countries: a cohort study. *F1000Research*, 13, 624.