

# Resource Summary Report

Generated by SPARC SAWG on Sep 20, 2026

## SoftCite

RRID:SCR\_024411

Type: Tool

---

### Proper Citation

SoftCite (RRID:SCR\_024411)

---

### Resource Information

**URL:** <https://github.com/howisonlab/softcite-dataset>

**Proper Citation:** SoftCite (RRID:SCR\_024411)

**Description:** Gold standard dataset of software mentions in research publications. Provides dataset of annotated software mentions from full text academic literature in biomedicine and economics directly converted from published PDFs with reproducible infrastructure. Includes provenance, and is formatted for immediately usefulness in NLP. Useful for supervised learning at scale.

**Synonyms:** Software Citation

**Resource Type:** data or information resource, knowledge base, portal, project portal, software resource

**Defining Citation:** [DOI:10.1002/asi.24454](https://doi.org/10.1002/asi.24454)

**Keywords:** Gold standard dataset, software mentions, research publications, annotated software mentions, full text academic literature,

**Funding:** Sloan Foundation

**Availability:** Free, Available for download, Freely available

**Resource Name:** SoftCite

**Resource ID:** SCR\_024411

**Alternate URLs:** [https://github.com/softcite/softcite\\_dataset\\_v2](https://github.com/softcite/softcite_dataset_v2),  
<https://zenodo.org/record/7995565>

**Record Creation Time:** 20230913T050234+0000

**Record Last Update:** 20260919T195543+0000

---

## Ratings and Alerts

No rating or validation information has been found for SoftCite.

No alerts have been found for SoftCite.

---

## Data and Source Information

**Source:** [SciCrunch Registry](#)

---

## Usage and Citation Metrics

We have not found any literature mentions for this resource.