



9 September 2024

9<sup>th</sup> International Collaboration for the Automation of Systematic Reviews (ICASR) Workshop  
Global Evidence Summit 2024, Prague, Czech Republic

# Automated Indexing of Biomedical Literature using Multi-Tagger 1.0

**Puranjani Das<sup>1</sup>, Jodi Schneider<sup>1\*</sup>, Halil Kilicoglu<sup>1</sup>, Joe D. Menke<sup>1</sup>, Arthur W. Holt<sup>2</sup>, Neil R. Smalheiser<sup>2</sup>**

<sup>1</sup> School of Information Sciences, University of Illinois at Urbana-Champaign, USA

<sup>2</sup> Department of Psychiatry, University of Illinois at Chicago, USA

\* **Contact: [jodi@illinois.edu](mailto:jodi@illinois.edu)**

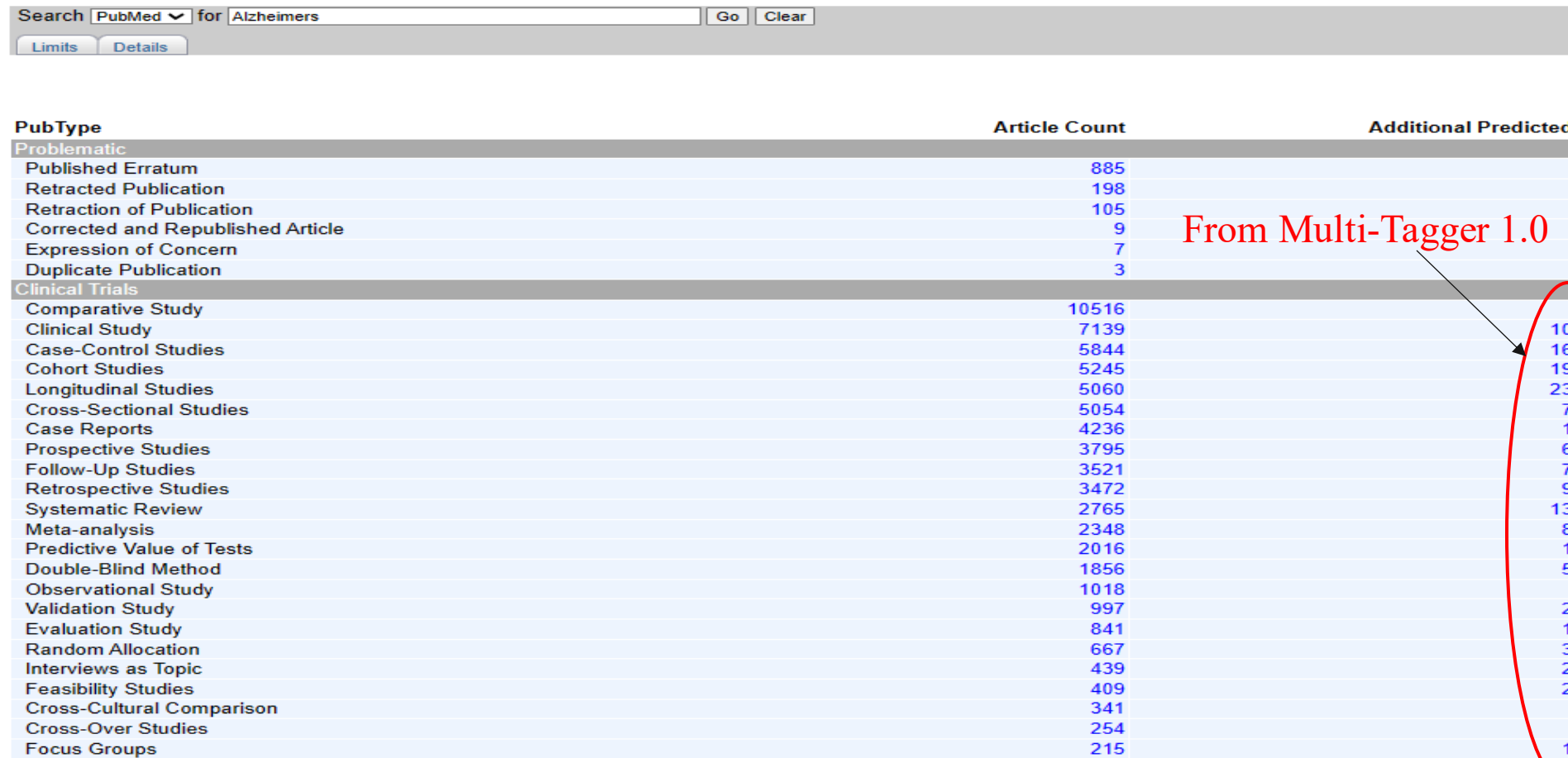
# Projects Funded by the US National Library of Medicine

**2017-2021** Text Mining Pipeline to Accelerate Systematic Reviews in Evidence-Based Medicine  
PIs: Neil R. Smalheiser, Aaron M. Cohen  
Output: RCT Tagger, Human Tagger, Aggregator, Multi-Tagger 1.0, Trials to Publications

**August 2023-July 2026** Automated Indexing for Publication Types and Study Design  
PI: Neil R. Smalheiser  
Co-PIs: Halil Kilicoglu, Jodi Schneider  
Output: Multi-Tagger 2.0

# Past Results

Anne O’Tate is a web-based search engine that, for any PubMed query, takes the set of retrieved articles and allows the user to mine them further in a variety of ways. It can display articles using both Multi-Tagger 1.0 predicted Publication Types and NLM-indexing [2,3].



Search PubMed for Alzheimers [Go] [Clear]

Limits Details

PubType	Article Count	Additional Predicted
<b>Problematic</b>		
Published Erratum	885	
Retracted Publication	198	
Retraction of Publication	105	
Corrected and Republished Article	9	
Expression of Concern	7	
Duplicate Publication	3	
<b>Clinical Trials</b>		
Comparative Study	10516	
Clinical Study	7139	10
Case-Control Studies	5844	16
Cohort Studies	5245	19
Longitudinal Studies	5060	23
Cross-Sectional Studies	5054	7
Case Reports	4236	1
Prospective Studies	3795	6
Follow-Up Studies	3521	7
Retrospective Studies	3472	9
Systematic Review	2765	13
Meta-analysis	2348	8
Predictive Value of Tests	2016	1
Double-Blind Method	1856	5
Observational Study	1018	
Validation Study	997	2
Evaluation Study	841	1
Random Allocation	667	3
Interviews as Topic	439	2
Feasibility Studies	409	2
Cross-Cultural Comparison	341	
Cross-Over Studies	254	
Focus Groups	215	1

From Multi-Tagger 1.0

Multi-Tagger 1.0 scores are publicly available for download [9].

Screenshot taken from the Anne O’Tate website [8].

# Past Results: Real-life Evaluation

DERP Report	# in initial retrieval set	# filtered out by our strategy	% work savings	# of included articles	# of included articles removed by our strategy	% recall
Anticoagulants-Original-Report	1766	659	37.32	82	0	100
Asthma-COPD	1964	497	25.31	28	0	100
Benzodiazepines-Summary-Review	581	302	51.98	12	0	100
Hepatitis-C-Update-2	4917	1417	28.82	75	2	97.33
Long-Acting-Insulins	1086	301	27.72	37	1	97.3
Long-Acting-Opioids-Update-7	503	60	11.93	13	0	100
MS-Drugs-Update-3	1849	825	44.62	45	3	93.33
Newer-Diabetes-Meds-Update-2	1065	400	37.56	21	1	95.24
PCSK9	75	32	42.67	13	0	100
Second-Generation-Antipsychotics-Update-5	1110	314	28.29	37	0	100

We tested Multi-Tagger 1.0 on 10 systematic reviews about drug effectiveness in collaboration with OHSU's Evidence-based Practice Center. **The mean work savings was 33.6% and the mean recall was 98.3% [5].**

# Current Work: Multi-Tagger 1.0 API

Get job results

Job ID: 191

Get Results

Data sent to Multi\_Tagger/taggerapi.cgi:

```
form{"mode": "results"; "jobid": "191"; }
```

```
{
  "success": true,
  "status": "completed",
  "jobid": "191",
  "message": "retrieved data for 2 pmids",
  "results": {
    "30013321": {
      "Autobiography": [
        0,
        0
      ],
      "Bibliography": [
        0,
        0
      ],
      "Biography": [
        0.0000077,
        0
      ],
      "Case-Control Studies": [
        0.00463804,
        0
      ],
      "Case Reports": [
        0.00032666,
        0
      ],
      "Clinical Studies as Topic": [
        0.28856383,
        0
      ],
      "Clinical study": [
        0.79083799,
        1
      ],
      "cohort studies": [
        0.00235718,
        0
      ],
      "Comment": [
        0.00014996,
        0
      ],
      "Congress": [
        0.00019176,
        0
      ],
      "Consensus Development Conference": [
        0.00002673,
        0
      ]
    }
  }
}
```

The Multi-Tagger 1.0 API gives scores for articles which may or may not be in PubMed [10]. The user can choose which Publication Types' scores they want to see.

The API supports 3 input formats:

- 1. list of PMIDs
- 2. RIS bibliographic format
- 3. JSON

The documentation page [11] shows examples for each format.

A Multi-Tagger 2.0 API will be available when the development is complete.

Result Response:

Row	Autobiography	Bibliography	Biography	Case-Control Studies	Case Reports	Clinical Studies as Topic	Clinical Study	Cohort Studies	Comment	Congress	Consensus Development Conference	Cross-Cultural Comparison	Cross-Over Studies	Cross-Sectional Studies	Diagnostic Test Accuracy	Double-Blind Method	E
30013321	0,0	0,0	0.0000077,0	0.00463804,0	0.00032666,0	0.28856383,	0.79083799,1	.00235718,0	0.00014996,0	0.00019176,0	0.00002673,0	0.0000197,0	0.0000734,0	0.00063101,0	0,0	0.31770833,0	0.00

# Current Work: Planning & Designing Multi-Tagger 2.0

- We are talking to potential stakeholders to close the gap between research production & research use by
  - Improving the quality of the Multi-Tagger
  - Increasing the relevance of the Multi-Tagger
- We are collaborating with stakeholders who have sets of annotated articles in order to evaluate the performance of Multi-Tagger against the annotation.

# Current Work: Multi-Tagger 2.0 Development

## **Improvements:**

- Using state-of-the-art transformer-based models based on PubMedBERT for automatic tagging of Publication Types [6]
- Training with a multi-label classification approach [6]
- Exploring undersampling, feature verbalization, and contrastive learning [6]

## **Results:**

- PubMedBERT provides a strong baseline for Publication Type indexing [6]
- Undersampling, feature verbalization, and unsupervised contrastive loss have a positive impact on performance [6]
- Supervised contrastive learning degrades the performance [6]

# Current Work: Multi-Tagger 2.0 Development

1. Expanding the list of Publication Types to encompass the full set of NLM-indexed Publication Types and clinical trial subtypes.
2. Introducing threshold value customization so that a diverse user base can customize to sensitivity and specificity
3. Seek collaborators to help evaluate Multi-Tagger in real life scenarios.



# You can help!

We are seeking:

- Help checking the definitions & prioritization of the Publication Types, especially when no consensus exists (e.g. DTAs)
- Annotated data to compare against Multi-Tagger scores, to check the relevance and performance of the models and to help train new PTs (e.g., DTAs, case series)
- Feedback on the usefulness of the Multi-Tagger API, both now when it accepts metadata as input and later when it will accept upload of full-text PDFs as input

# References

1. Cohen AM, Schneider J, Fu Y, McDonagh MS, Das P, Holt AW, Smalheiser NR. Fifty Ways to Tag Your Pubtypes: Multi-Tagger, a Set of Probabilistic Publication Type and Study Design Taggers to Support Biomedical Indexing and Evidence-Based Medicine. medRxiv [Preprint]. 2021 Jul 13: <http://doi.org/2021.07.13.21260468>
2. Smalheiser NR, Fragnito DP, Tirk EE (2021) Anne O'Tate: Value-added PubMed search engine for analysis and text mining. PLOS ONE 16(3): e0248335. <https://doi.org/10.1371/journal.pone.0248335>
3. Smalheiser, N.R., Zhou, W. & Torvik, V.I. Anne O'Tate: A tool to support user-driven summarization, drill-down and browsing of PubMed search results. J Biomed Discov Collaboration 3, 2 (2008). <https://doi.org/10.1186/1747-5333-3-2>
4. Schneider J, Hoang L, Kansara Y, Cohen AM, Smalheiser NR. Evaluation of publication type tagging as a strategy to screen randomized controlled trial articles in preparing systematic reviews. JAMIA Open. 2022 Mar 30;5(1):ooac015. <http://doi.org/10.1093/jamiaopen/ooac015>
5. Proescholdt R, Hsiao TK, Schneider J, Cohen AM, McDonagh MS, Smalheiser NR. Testing a filtering strategy for systematic reviews: evaluating work savings and recall. AMIA Jt Summits Transl Sci Proc. 2022 May 23;2022:406-413. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9285169/>
6. Menke JD, Halil Kilicoglu H, Smalheiser NR. Publication type tagging using transformer models and multi-label classification. AMIA Annu Symp Proc. 2024.

# Related Resources

7. Text Mining Pipeline to Accelerate Systematic Reviews in Evidence-Based Medicine –

[https://arrowsmith.psych.uic.edu/evidence\\_based\\_medicine/index.html](https://arrowsmith.psych.uic.edu/evidence_based_medicine/index.html)

8. Anne O’Tate – [https://arrowsmith.psych.uic.edu/cgi-bin/arrowsmith\\_uic/AnneOTate\\_summarize.cgi](https://arrowsmith.psych.uic.edu/cgi-bin/arrowsmith_uic/AnneOTate_summarize.cgi)

9. Multi-Tagger Files – [https://arrowsmith.psych.uic.edu/evidence\\_based\\_medicine/mt\\_download.html](https://arrowsmith.psych.uic.edu/evidence_based_medicine/mt_download.html)

10. Multi-Tagger 1.0 API – [https://arrowsmith.psych.uic.edu/cgi-bin/arrowsmith\\_uic/Multi\\_Tagger/taggerapi\\_testpage.php](https://arrowsmith.psych.uic.edu/cgi-bin/arrowsmith_uic/Multi_Tagger/taggerapi_testpage.php)

11. Multi-Tagger 1.0 API documentation –

[https://arrowsmith.psych.uic.edu/cgi-bin/arrowsmith\\_uic/Multi\\_Tagger/taggerapi\\_documentation.php](https://arrowsmith.psych.uic.edu/cgi-bin/arrowsmith_uic/Multi_Tagger/taggerapi_documentation.php)

# Appendix

# Current Work: Multi-Tagger 1.0 API

## MultiTagger API Test Page - model version 1.0

### Initial Request Post

Mode:

pmid ▾

PMID List (comma-separated or one each line):

31894029, 30013321

Submit

Data sent to Multi\_Tagger/taggerapi.cgi: `form{"mode": "pmid"; "pmidlist": "31894029, 30013321"; }`

Response: `{"success":true,"status":"job started","jobid":"191","message":"job created"}`

# Current Work: Multi-Tagger 1.0 API

Data sent to Multi\_Tagger/taggerapi.cgi: `form{"mode": "pmid"; "pmidlist": "31894029, 30013321"; }`

Response: `{"success":true,"status":"job started","jobid":"191","message":"job created"}`

## Poll for Job Status

Job ID:

Data sent to Multi\_Tagger/taggerapi.cgi: `form{"mode": "status"; "jobid": "191"; }`

Poll Response: `{"success":true,"status":"processing","jobid":"191" "message":"getting data for 2 pmids"}`

Upon completion

Data sent to Multi\_Tagger/taggerapi.cgi: `form{"mode": "pmid"; "pmidlist": "31894029, 30013321"; }`

Response: `{"success":true,"status":"job started","jobid":"191","message":"job created"}`

## Poll for Job Status

Job ID:

Data sent to Multi\_Tagger/taggerapi.cgi: `form{"mode": "status"; "jobid": "191"; }`

Poll Response: `{"success":true,"status":"completed","jobid":"191" "message":"retrieved data for 2 pmids"}`

# The 49 Publication Types

Autobiography

Bibliography

Biography

Case-Control Studies

Case Reports

Clinical Studies as Topic

Clinical Study

Cohort Studies

Comment

Congress

Consensus Development

Conference

Cross-Cultural Comparison

Cross-Over Studies

Cross-Sectional Studies

Diagnostic Test Accuracy

Double-Blind Method

Editorial

Evaluation Studies as Topic

Evaluation Study

Feasibility Studies

Focus Groups

Follow-Up Studies

Genome-Wide Association  
Study

Historical Article

Human Experimentation  
interview

Interviews as Topic

Lecture

Legal Case

Letter

Longitudinal Studies

Matched-Pair Analysis

Meta-analysis

Multicenter Study

News

Personal Narrative

Portrait

Practice Guideline

Predictive Value of Tests

Prospective Studies

Random Allocation

Randomized Controlled Trial

Reproducibility of Results

Retrospective Studies

Review

Systematic Reviews as Topic

Systematic Review

Twin Study

Validation Study