SciClaimEval

Cross-Modal Claim Verification in Scientific Papers

Pilot Task @ NTCIR-19 sciclaimeval.github.io

Akiko **Aizawa** *(National Institute of Informatics)*
Xanh **Ho** *(National Institute of Informatics)*
André **Greiner-Petter** *(University of Göttingen, Germany)*
Florian **Boudin** *(Nantes University, France)*

Why Scientific Claim Verification?

01 Overwhelming surge in scientific submissions without automated quality assurance metrics

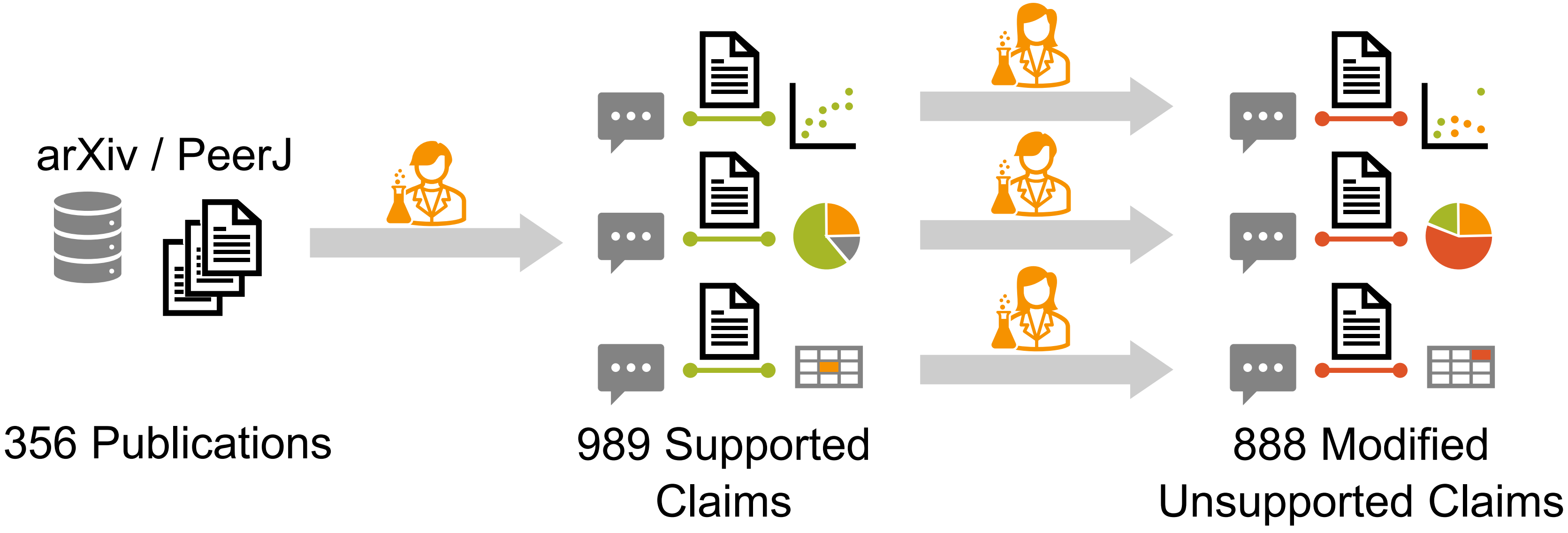
02 A possible quality insurance metric are the number of supported / unsupported claims

03 Claim supports (evidence) is typically **multi-modal** (tables, plots, graphs, etc.)

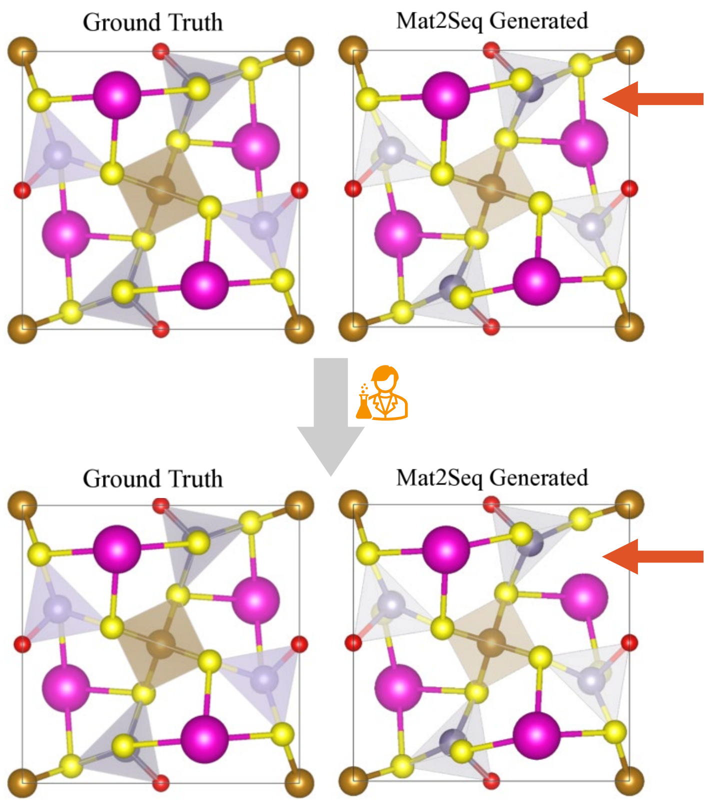
04 Existing datasets alter the textual claim into artificial and **unrealistic** negations rather than the evidence

Tempered evidence leads to **realistic** unsupported claims

Dataset Construction



Example Modification



Modification Operations

Figure Edits	Legend changes, graph manipulations, datapoint adjustments, axis scaling, fraudulent data insertion, image manipulations, and more.
Table Edits	Value changes, row/column swaps, full table altercations, value scaling, caption adjustments.

SciClaimEval @ NTCIR-19

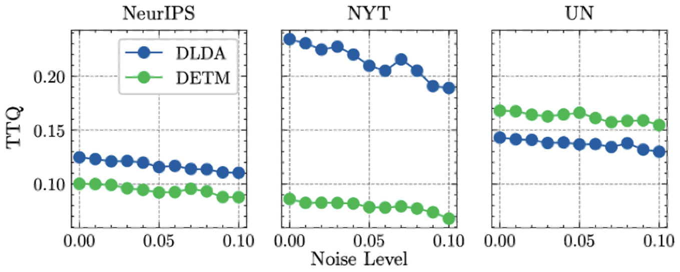
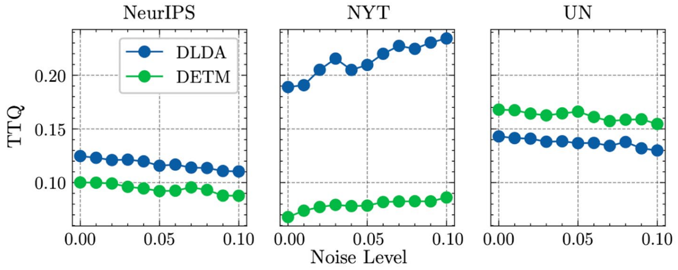
Input: Claim (text) + Evidence (tables, various figures) + Context
sourced from scientific articles from NLP, ML and medical domain

---

Output: ✅ Supported / ⭕ Unsupported

TestBed: Original Claims paired with Perturbed Evidence
→ *plausible claims but uncertain support*

Example:

Claim	Evidence	Label
The TTQ measure decreases continuously as more intrusion words are added to the topics for all three datasets.		✅ **Supported**
The TTQ measure decreases continuously as more intrusion words are added to the topics for all three datasets.		⭕ **Unsupported**

Timeline

2026

January 2026

Dataset Release

January - June ‘26

Dry Run

March - July ‘26

Formal Run

August 1, ‘26

Evaluation Results Return

August 1, ‘26

Task Overview Release

September 1, ‘26

Participants Paper Submission Deadline

December ‘26

NTCIR-19 Conference

SciClaimEval

THANK YOU

sciclaimeval.github.io