# Veriscite:
# A Resource for Scientific Claim Verification

JDSE 2025

Lucía Catalán Gris, Kim Gerdes
25 September, 2025

LISN, Université Paris-Saclay, Orsay 91400, France

# Table of contents

# 1. Context and motivation

The task of assessing whether a factual claim is valid based on evidence.
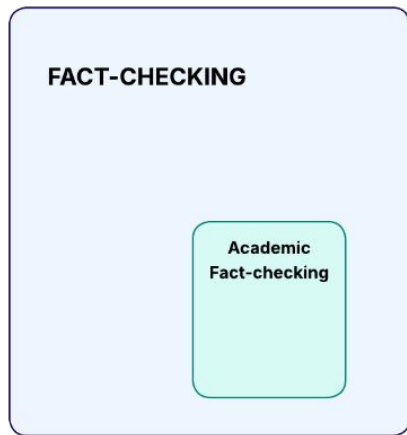
Examples:

1. Newspaper's articles

2. COVID-19 vaccination tweets

3. Climate change tweets

4. Presidential candidates' statements

FACT-CHECKING

The task of assessing whether a factual claim is valid based on evidence.

Examples:

1. Newspaper's articles

2. COVID-19 vaccination tweets

3. Climate change tweets

4. Presidential candidates' statements



FACT-CHECKING

Academic
Fact-checking

Assessing whether a claim from an **academic source** is valid based on evidence.

Types of claims in a paper:

1. Scientific statements
2. Logic reasoning
3. Citation claims

**FACT-CHECKING**

Academic Fact-checking
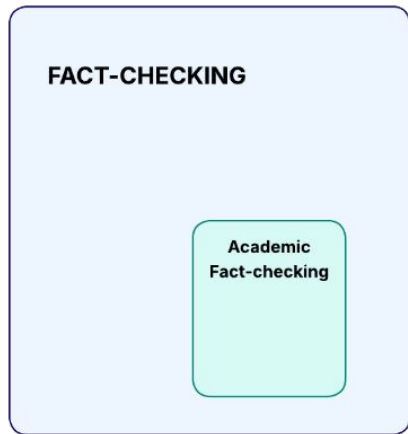
Assessing whether a claim from an **academic source** is valid based on evidence.

Types of claims in a paper:

1. Scientific statements
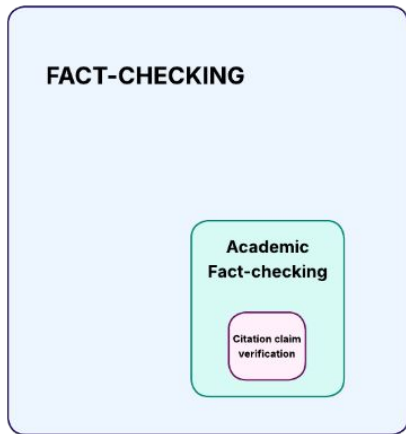2. Logic reasoning
3. Citation claims

## *Scientific statement Verification*

1. Evidence: more than one relevant academic articles

2. Verify if the meaning of the claim itself it's true or not.

## *Citation claim Verification*

1. Evidence: the referenced article (appears in the references section)

2. Verify if the cited author meant to claim that or not.

Is what the author says about the cited article verified in the

referenced article?
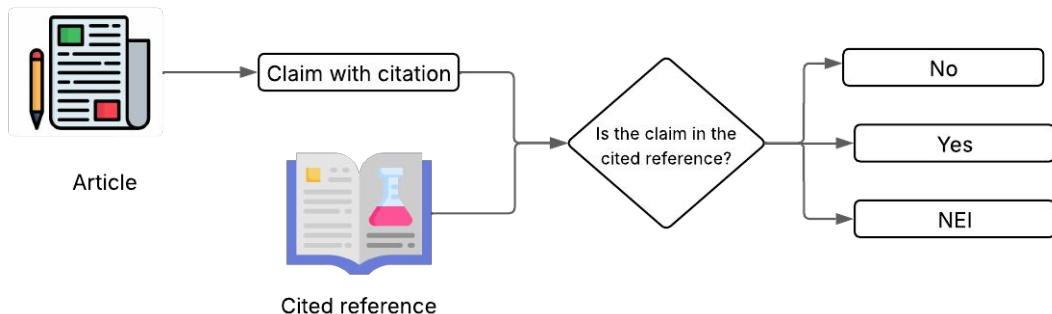
- Scientists are required to **explicitly cite all the claims** in their publications that were **inspired or used in earlier works** (Nicolaisen, 2007).

- That claim should be **connected to a reference** in the bibliography section of the article.

| Dataset | # Claims | Claim Origin | Evidence Source | Domain |
|---------|----------|--------------|-----------------|--------|
| SCIFACT (Wadden et al., 2020) | 1,409 | Researchers | Research papers | Biomedical |
| PUBHEALTH (Kotonya and Toni, 2020b) | 11,832 | Fact-checkers | Fact-checking sites | Public health |
| CLIMATE-FEVER (Diggelmann et al., 2020) | 1,535 | News articles | Wikipedia articles | Climate change |
| HEALTHVER (Sarrouti et al., 2021) | 1,855 | Search queries | Research papers | Health |
| COVID-FACT (Saakyan et al., 2021) | 4,086 | Reddit posts | Research, news | COVID-19 |
| COVERT (Mohr et al., 2022) | 300 | Twitter posts | Research, news | Biomedical |

Table 1: Datasets for the task of scientific fact-checking and claim verification

Peer Review:

➔ Near universal application for evaluating scientific articles before publication.

➔ It's mostly a manual process.

**Problems:**

1. Explosion of submissions co-written by ChatGPT.

2. It's very slow.

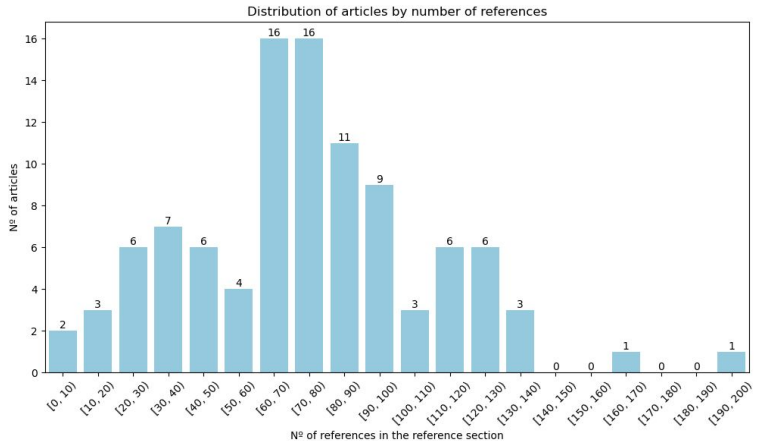3. Most reviewers don't have time to check all the citations inside the paper.

1. We know it.

2. It's largely unexplored; most academic fact-checking is conducted in the biomedical domain.

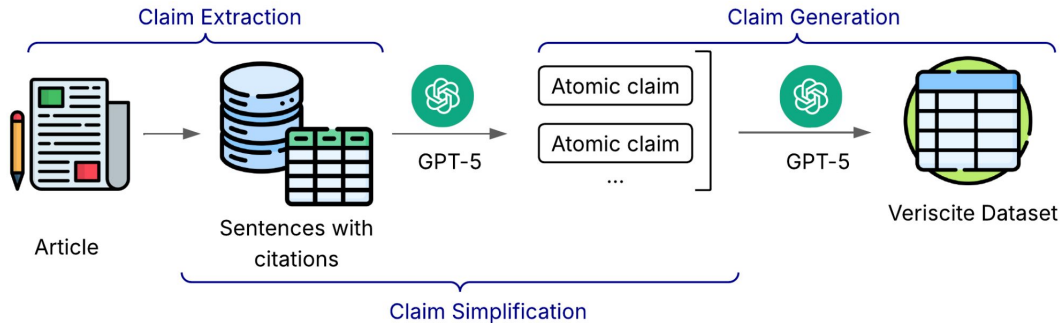3. There are fewer open-source databases available for it compared to other fields.

# 2. The Veriscite Dataset

100 open-access articles drawn from two journals:

➢ 50 from Glossa

➢ 50 from Computational Linguistics



Distribution of articles by number of references

Article

Citation extraction (Grobid) → Discard non-academic citations → Claim

**Claim**
- id of the claim
- context
- sentence
- referenced article

|  | **N found** | **Example** |
|---|---|---|
| Sentences (claims) with citation | 9.800 | Associative learning is a fundamental mechanism that holds strong explanatory power for general learning in both humans and other animals  (Pavlov 1949, [...]) |
| References | 7.489 | Pavlov, Ivan P. 1949. Conditioned responses. Readings in General Psychology, pages 249–267. https://doi.org/10.1037/11352-036 |

Example from Jon-And and Michaud (2024)
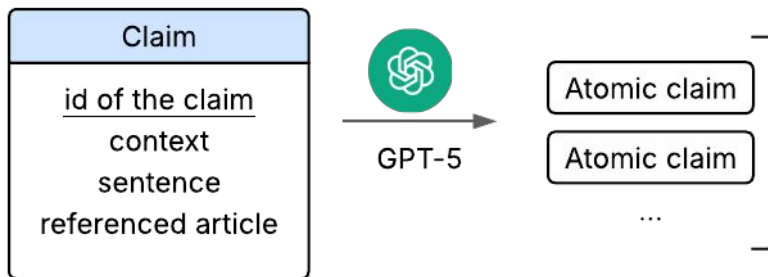
**TASK**: Each statement about a cited article is split into atomic claims.
**MODEL**: GPT-5

| | N found | Example |
|---|---|---|
| Sentences with citation | 9.800 | Associative learning is a fundamental mechanism that holds strong explanatory power for general learning in both humans and other animals (Pavlov 1949, [...]) |
| Atomic claims | 13.021 | 1. The paper supports the view that associative learning is a fundamental mechanism. <br> 2. The paper provides evidence that associative learning has strong explanatory power for general learning. <br> 3. The paper provides evidence that associative learning accounts for learning in humans. |

**TASK**: To add nuanced and false claims to the dataset, we generate 5 variants of the previous claims.

**MODEL**: GPT-5

|  | Example |
|---|---|
| Sentences with citation | Associative learning is a fundamental mechanism that holds strong explanatory power for general learning in both humans and other animals  (Pavlov 1949, [...]) |
| Atomic claims | 1. The paper supports the view that associative learning is a fundamental mechanism. <br> 2. The paper provides evidence that associative learning has strong explanatory power for general learning. <br> 3. The paper provides evidence that associative learning accounts for learning in humans. |

**Claim Generation: The paper supports the view that associative learning is a fundamental mechanism.**

|   | Description | Example |
|---|-------------|---------|
| 1 | Oposite claim | Associative learning is not a fundamental mechanism and has little explanatory power for general learning in humans or other animals. |
| 2 | More general claim (not necessarily true) | Learning processes can be explained by a small set of universal mechanisms shared across species. |
| 3 | More specific claim (a fortiori true | Classical conditioning, as a form of associative learning, provides strong explanatory power for certain types of learning in both humans and other animals. |
| 4 | Similar claim in the same domain, but cannot be true at the same time | General learning in humans and other animals is primarily explained by innate cognitive structures, not associative mechanisms. |
| 5 | Paraphrase necessarily true if the claim is true | The view presented is that associative learning underlies and explains much of how humans and other animals learn. |

New open dataset for citation verification in scientific articles:

- Focus on computational linguistics.
- 65,105 atomic claims about cited articles.
    - Labeled as True, False, or NEI (Not Enough Information)
    - Together with the reference and URL of the cited work

# 4. Conclusion and future work

➢ First step toward building a comprehensive benchmark for scientific fact-checking.

➢ Complementary challenge: given a text without references, predicting where citations should appear

1. Developing a typology for the claims.

2. We are setting up two human verification processes:

   a. Manual claim simplification step to assess the quality of atomic claims generated by GPT-5

   b. Human claim verification to critically review the assigned labels from the typology.

Nicolaisen, J. (2007), Citation analysis. Ann. Rev. Info. Sci. Tech., 41: 609-641. https://doi.org/10.1002/aris.2007.1440410120

Juraj Vladika and Florian Matthes. 2023. Scientific Fact-Checking: A Survey of Resources and Approaches. In Findings of the Association for Computational Linguistics: ACL 2023, pages 6215–6230, Toronto, Canada. Association for Computational Linguistics.

Anna Jon-And, Jérôme Michaud; Usage-based Grammar Induction from Minimal Cognitive Principles. Computational Linguistics 2024; 50 (4): 1375–1414. doi: https://doi.org/10.1162/coli_a_00528

# Thank you for listening