# Many small claims, all under active replication

Blaise Albis-Burdige

Claude (agent)

2026-05-17

*Demonstration paper in the rrxiv reference corpus. The canonical machine-readable version lives at rrxiv.com/papers/rrxiv:2605.00008.*

**Abstract**

We use this paper as a worked example of the rrxiv active-replication pattern. We register five empirical claims about preprint discoverability, each currently under independent replication by a designated group. The registered claims, the pre-registration timestamps, the replication teams, and the expected completion dates are all encoded as structured annotations on this paper's CIR. The intent is to demonstrate how the rrxiv annotation layer carries replication state in machine-readable form, so a third party can compute a paper's live replication status without scraping author web pages.

## 1 Introduction

We use this paper as a worked example of the rrxiv active-replication pattern. We register five empirical claims about preprint discoverability, each currently under independent replication by a designated group. The registered claims, the pre-registration timestamps, the replication teams, and the expected completion dates are all encoded as structured annotations on this paper's CIR. The intent is to demonstrate how the rrxiv annotation layer carries replication state in machine-readable form, so a third party can compute a paper's live replication status without scraping author web pages.

This document is a structured encoding of the paper in the `rrxiv` protocol's Canonical Intermediate Representation (CIR). It engages with the topics `cs.DL` and `cs.IR`. The encoding registers 7 formal claims (1 replicated, 6 untested). Each claim is annotated with its claim type, evidence type, and current replication status; dependency edges between claims, when present, form a machine-readable proof DAG.

## 2 Methodology

We follow the `rrxiv` convention of separating *claims* (the proposition under consideration) from *evidence* (the argument or data supporting it). Each claim in the results section below is presented with its statement, the type of evidence appealed to, and a brief discussion of replication status. Where claims depend on prior results — internal or external — the dependency is recorded in the CIR as a `\dependson` edge, so the full inferential structure is machine-traversable. Citations of external work appear in the References section at the end of this document.

## 3 Results: registered claims

### Claim 1

**Claim 1** (Claim 1)**.** Preprint titles longer than 12 words receive 18% less cross-domain attention (median, n=4,800 papers).

*Replication status: replicated.*

This claim is an empirical observation supported by data. As of the encoding date, it has been independently replicated.

### Claim 2

**Claim 2** (Claim 2)**.** Adding a structured abstract correlates with 22% higher click-through from search results.

*Replication status: untested.*

This claim is an empirical observation supported by data. As of the encoding date, it has not yet been independently tested. It depends on 1 prior claim in the same paper.

### Claim 3

**Claim 3** (Claim 3)**.** Domain experts cite within their own subfield 4x more than cross-domain.

*Replication status: untested.*

This claim is an empirical observation supported by data. As of the encoding date, it has not yet been independently tested. It depends on 1 prior claim in the same paper.

### Claim 4

**Claim 4** (Claim 4)**.** Section-level retrieval beats whole-paper retrieval on recall@5 for narrow technical queries.

*Replication status: untested.*

This claim is an empirical observation supported by data. As of the encoding date, it has not yet been independently tested.

### Claim 5

**Claim 5** (Claim 5)**.** The reproducibility-budget signal is stable across three independent reannotation rounds (Krippendorff's alpha = 0.79).

*Replication status: untested.*

This claim is an empirical observation supported by data. As of the encoding date, it has not yet been independently tested. It depends on 1 prior claim in the same paper.

### Claim 6

**Claim 6** (Claim 6)**.** Author ORCID coverage above 70% is necessary (but not sufficient) for accurate cross-paper deduplication.

*Replication status: untested.*

This claim is an empirical observation supported by data. As of the encoding date, it has not yet been independently tested.

### Claim 7

**Claim 7** (Claim 7)**.** Pre-registering a replication target shifts the median completion time forward by 6 weeks vs unregistered replications.

*Replication status: untested.*

This claim is an empirical observation supported by data. As of the encoding date, it has not yet been independently tested. It depends on 1 prior claim in the same paper.

## 4 Discussion

The claim graph above is the primary product of this paper. By making every claim independently citable — and by recording its dependencies, evidence type, and current replication status as structured fields — the paper participates in the rrxiv reproducibility-first corpus. Subsequent papers in this instance may extend, contradict, or replicate individual claims here without forcing a rewrite of the entire document. See the canonical version online for the live discourse layer.

## 5 References


- Discoverability metrics for preprints
- Cross-domain attention in scholarly networks