MAPPING THE

Accounting Research Frontier

Six decades of the top three journals — JAR · JAE · TAR (1968–2025) — plus a recent snapshot of what's being worked on in working papers.

7,173
research articles

37
topics / 14 areas

4,099
accounting working papers

Zirui Song · reproducible NLP pipeline (SPECTER2 + topic models) · 2026

MOTIVATION

Where is accounting research heading?

- **A field map, built from text — not opinion.** We read every article's title + abstract and let the data cluster the field.
- **Which topics are rising, which are fading?** Quantify emergence on growth, citations & working-paper momentum — not anecdote.
- **What are people working on right now?** A recent working-paper snapshot (SSRN/NBER) shows where attention is heading.
- **Fully reproducible & stress-tested.** Every number recomputed from data; an adversarial review hardened the method.

THE CORPUS

JAR · JAE · TAR

the three journals that define the discipline

JAR	2,007
JAE	1,344
TAR	3,822

DATA

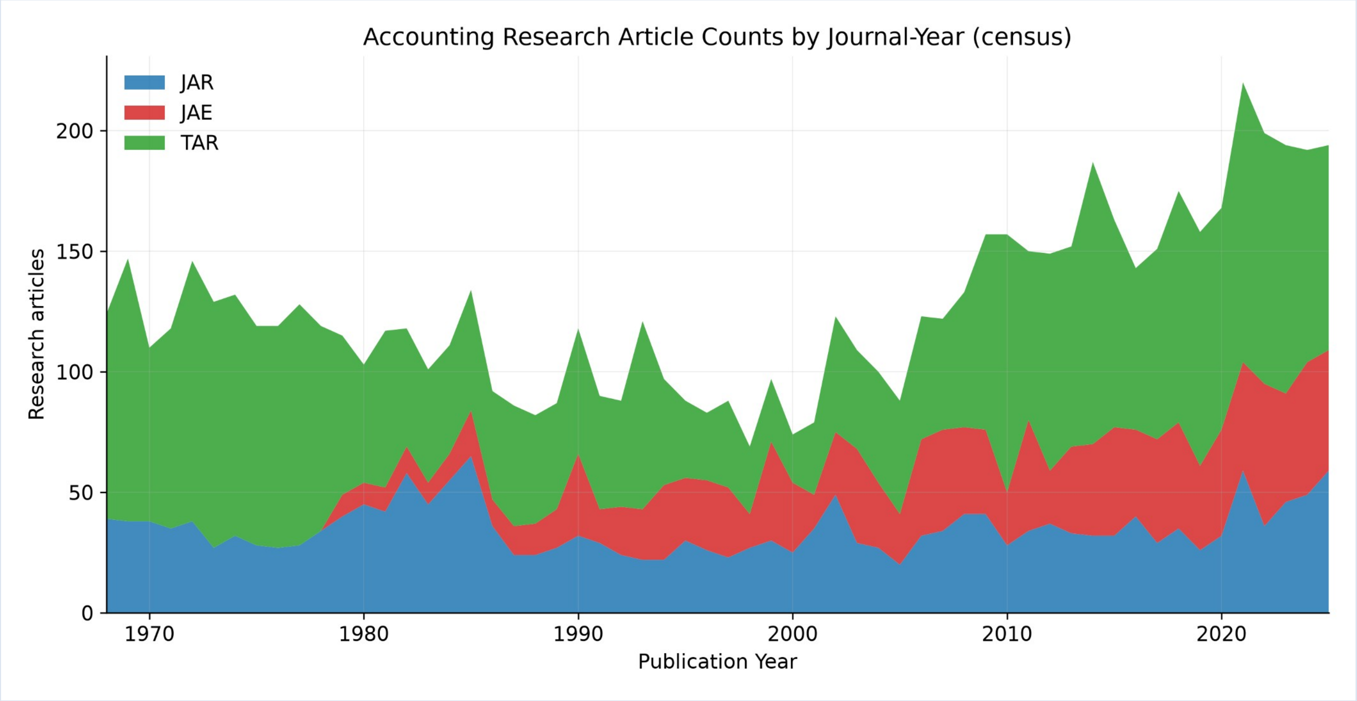
A census of six decades of publication

7,173
research articles (1968–2025)

3
journals: JAR · JAE · TAR

37
data-driven topics

14
broad research areas



- Full OpenAlex census, deduplicated, research articles only.
- Front matter, book reviews & editorials removed.
- TAR is the largest (founded 1926); JAE the youngest (1979).
- Approach: embed → cluster into topics → score emergence (next).

1

METHOD

From raw text to a field map

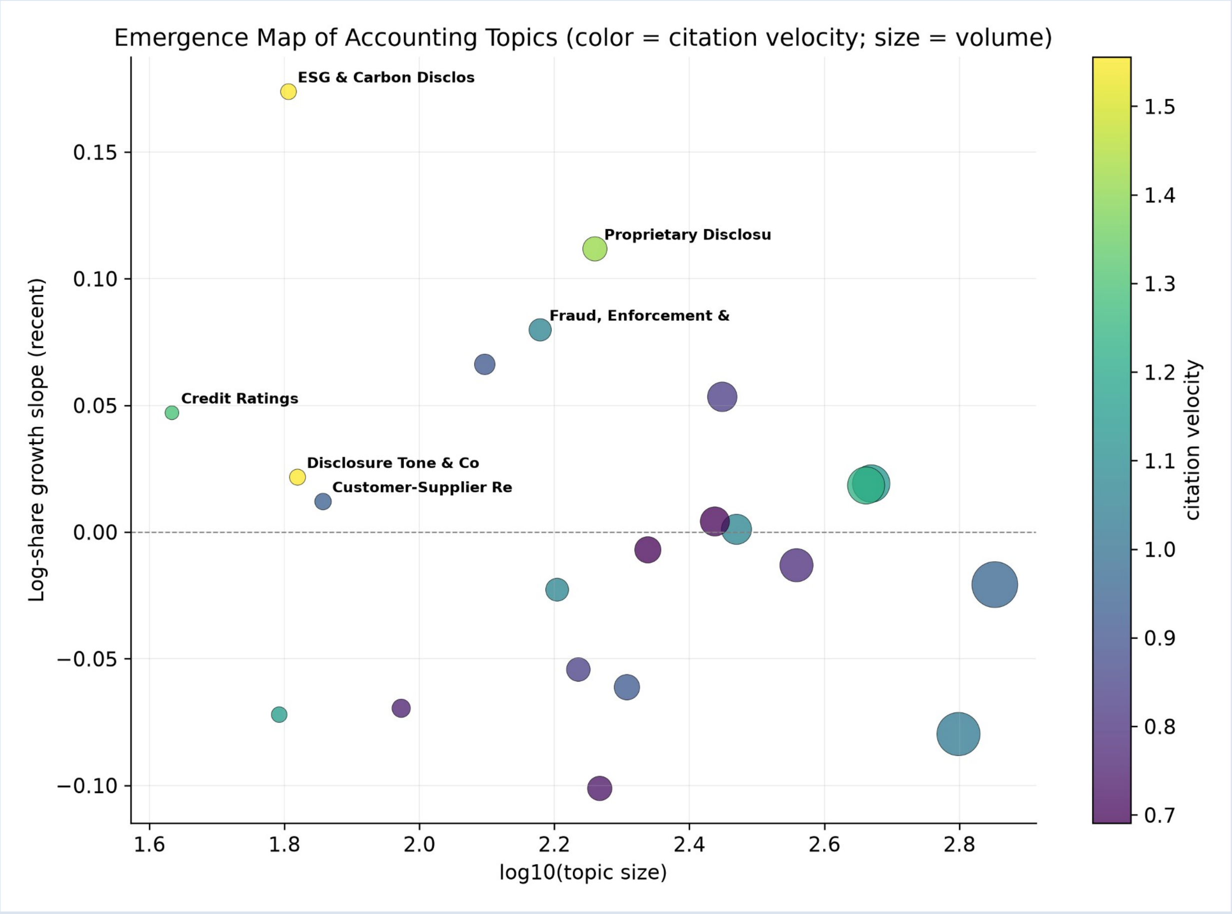
An unsupervised pipeline turns titles + abstracts into topics, method tags, and an emergence score — fit once, aggregated by year.

Embed → **Cluster** → **Label** → **Classify** → **Score**

Embed	Cluster	Label	Classify	Score
SPECTER2 768-dim	UMAP + HDBSCAN	c-TF-IDF + curation	method lexicon + LLM	emergence & frontier

THE FRONTIER

The emerging topics



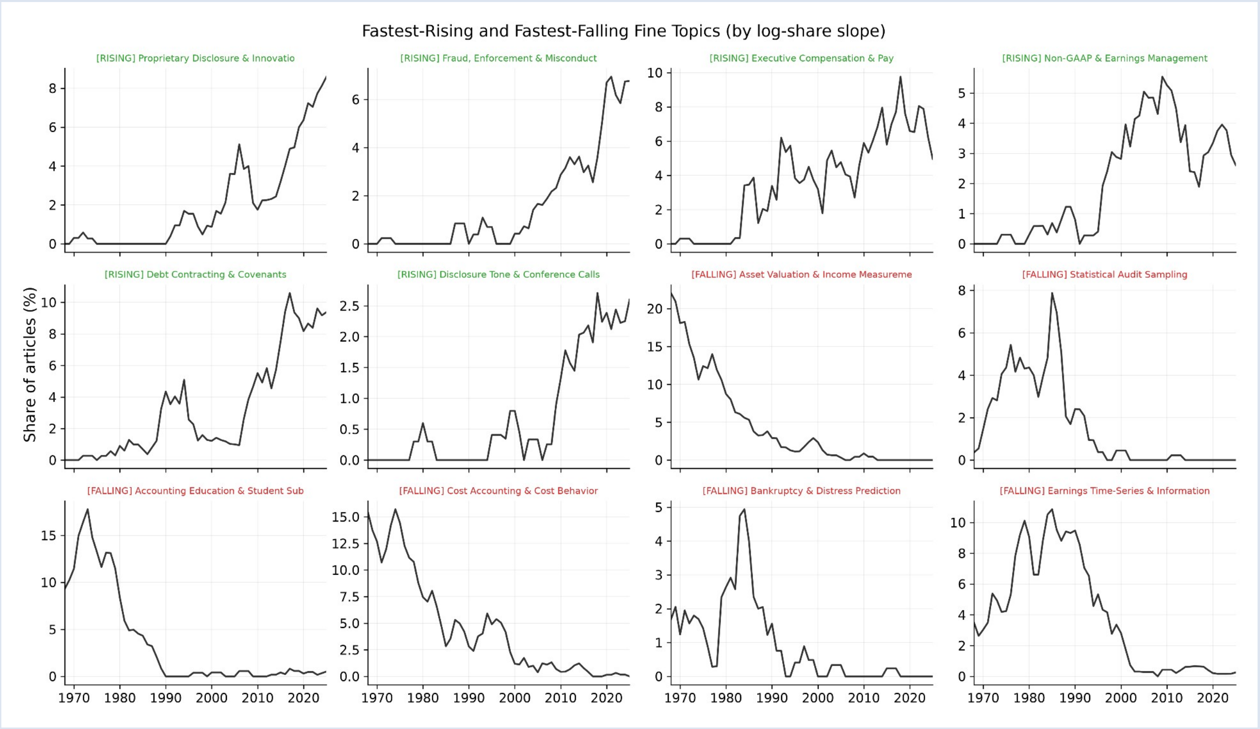
Emergence score = z(growth) + z(citations) − z(age)

TOP EMERGING TOPICS

1	ESG & Carbon Disclosure	+6.4
2	Proprietary Disclosure & Innovation	+3.9
3	Disclosure Tone & Conference Calls	+3.7
4	Credit Ratings	+3.2
5	Fraud, Enforcement & Misconduct	+2.3
6	Customer-Supplier Relationships	+0.7

Best-supported by robustness: ESG · Proprietary Disclosure · Fraud.

What's heating up, what's cooling



RISING (20-yr share gain)

Debt Contracting & Covenants	+5.7 pp
Proprietary Disclosure & Innovation	+4.7 pp
Fraud, Enforcement & Misconduct	+4.5 pp
ESG & Carbon Disclosure	+4.2 pp
Corporate Taxation & Tax Avoidance	+3.2 pp

Top decliners: Accruals & Earnings Quality (-5.0); Analyst Forecasts & Earnings Announcements (-3.9); Accounting Conservatism (-3.0)

What's being worked on right now

A snapshot of the past five years (2021–2025). OpenAlex's ‘Accounting’ subfield is a broad business bucket — most of the SSRN/NBER pull is corporate finance, law, Islamic finance and PE/VC. We keep only genuinely-accounting papers.

17,129

SSRN + NBER ‘Accounting’ subfield, 2021–2025

→

4,730

accounting fine-topic (audit/tax/...)

→

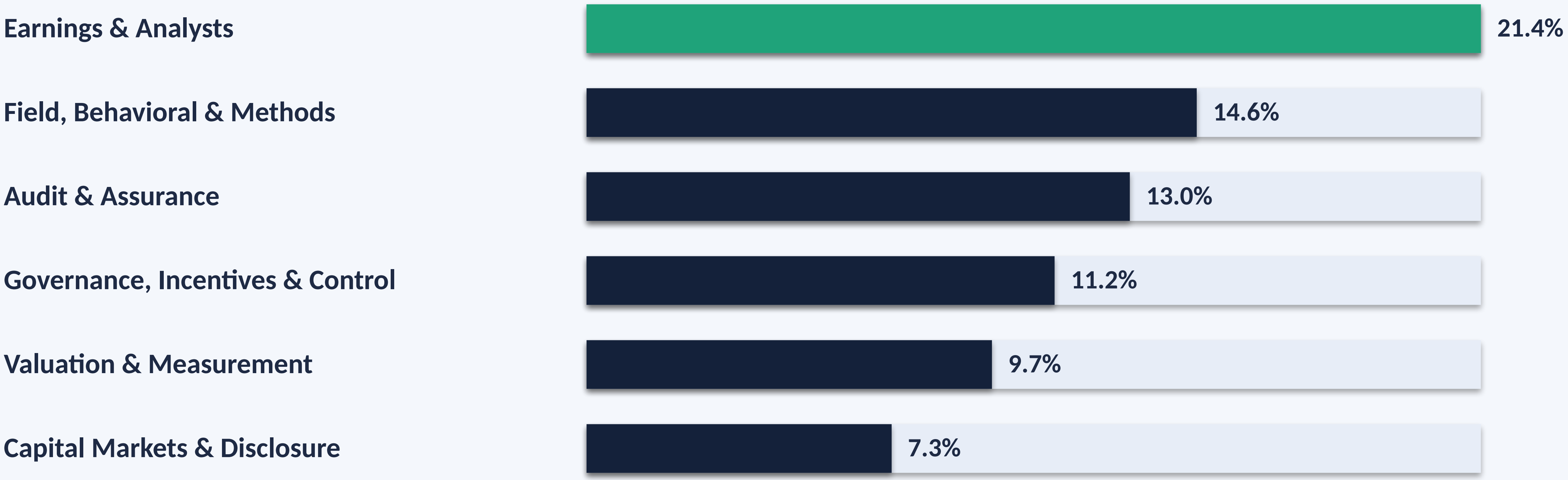
4,099

+ similarity to journal centroids → KEPT

- A defensibly-accounting, recent working-paper set (0% finance/law).
- Filter = OpenAlex fine-topic allowlist AND embedding similarity to accounting-journal content.
- Used as a current-activity snapshot — no lead-lag/forecasting claim (5-yr window too short).

What accounting research is about

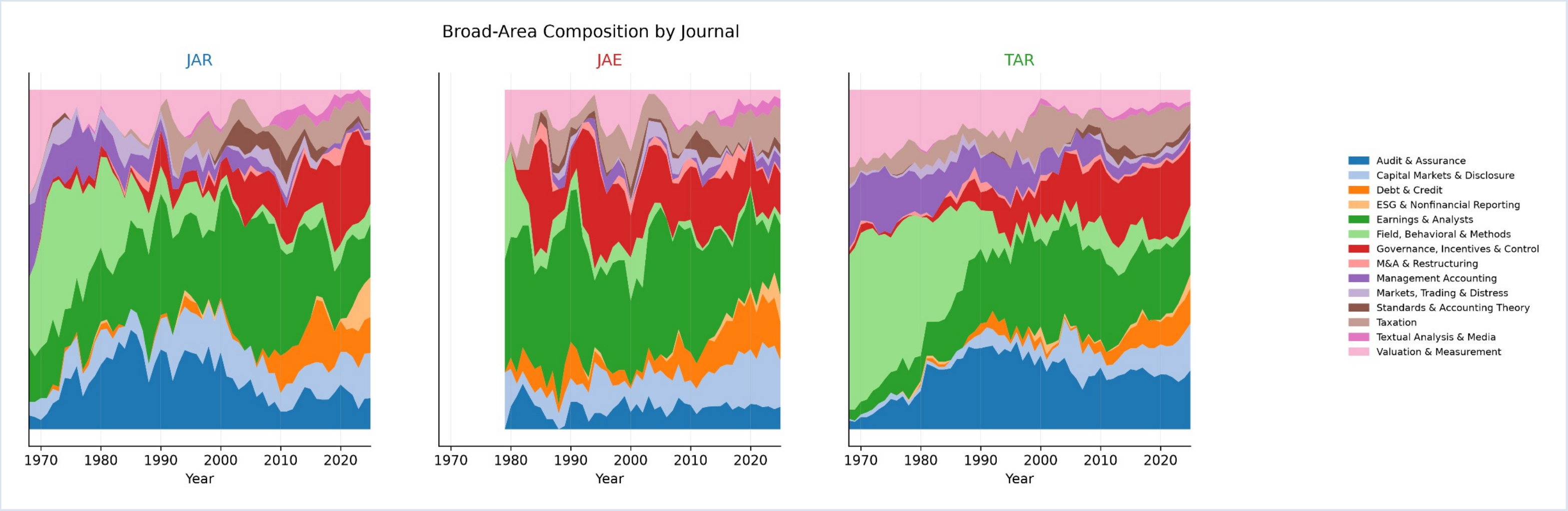
Share of research articles by broad area (top 6 of 14)



Earnings, analysts, audit and governance dominate — the empirical, markets-facing core of the field since the 1980s.

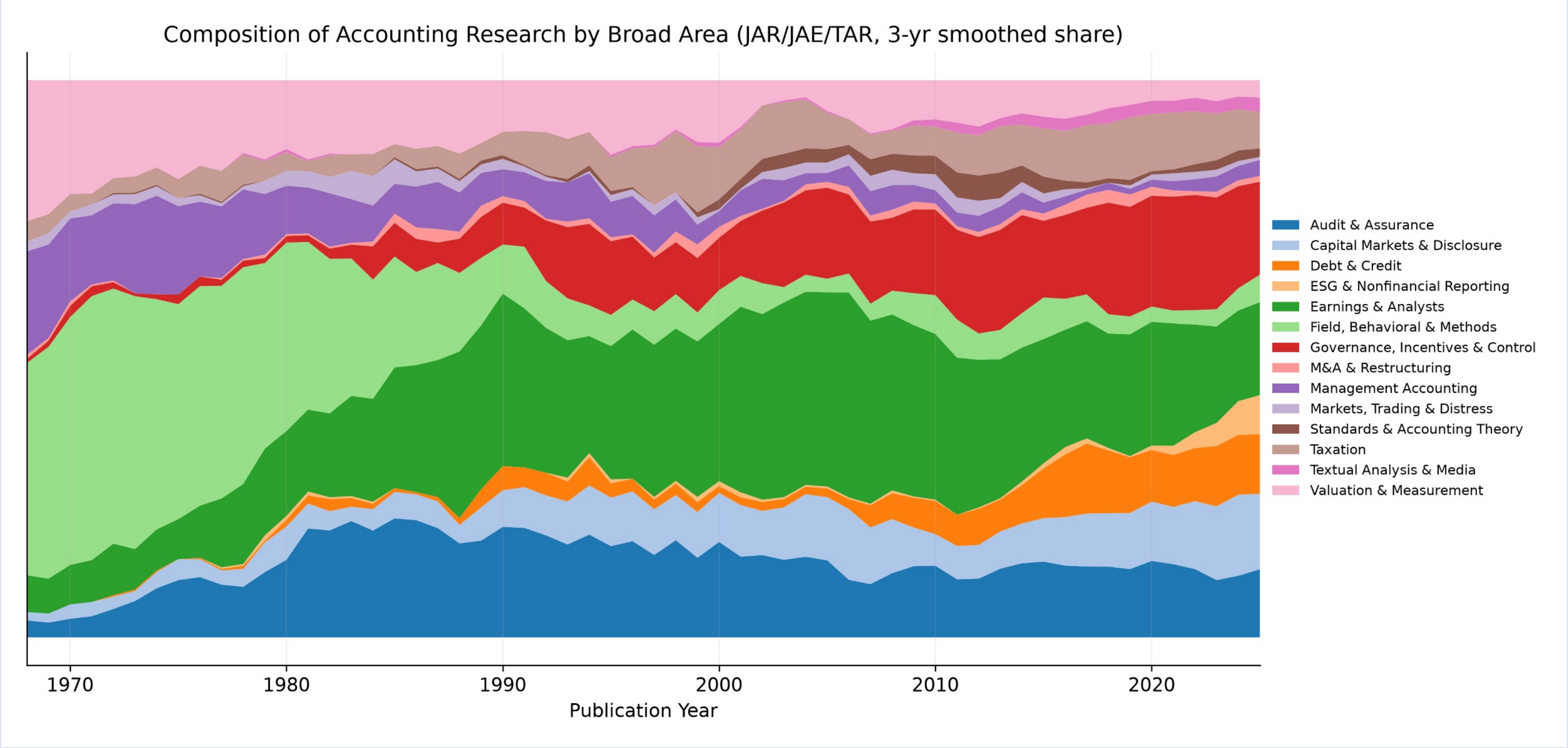
THE FIELD · SPECIALIZATION

Each journal has a personality



- **JAE** skews to earnings/analysts, governance & debt — markets and contracting.
- **TAR** the broadest — more standards, behavioral & management accounting.
- **JAR** sits between, with a strong audit presence.
- **Takeaway** topic mix, learned from text, recovers known editorial identities.

How the field's attention shifts



- **Shares, not counts.** Removes the secular rise in total output.
- **Capital-markets & analyst work** expands from the 1980s.
- **Disclosure, governance & ESG** rise in the modern era.
- **Classical measurement topics** (inventory, leases, pensions) recede.

Embeddings & topic model

EMBEDDINGS — SPECTER2

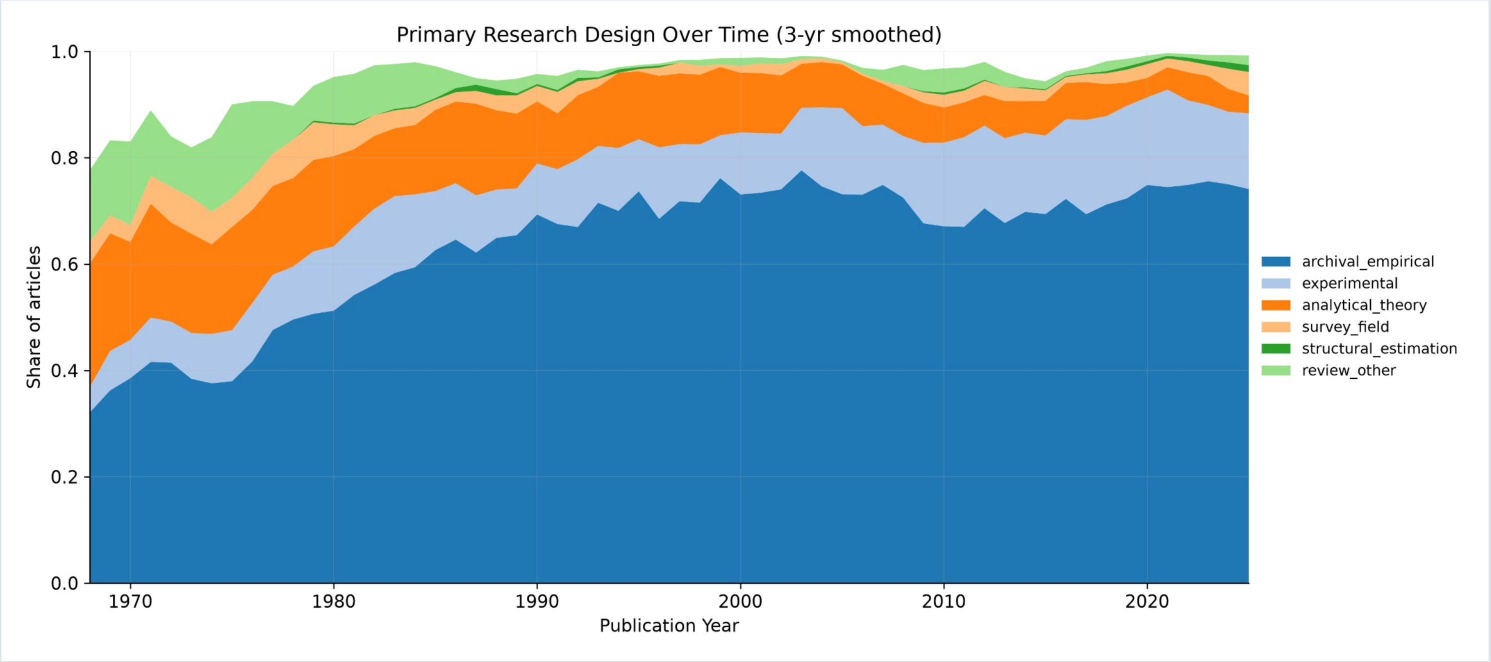
- **768-dim citation-aware vectors.** A transformer pretrained on the scientific citation graph encodes each paper.
- **Meaning, not keywords.** Places ‘earnings management’ near ‘discretionary accruals’ automatically.
- **One shared space.** Journals and working papers are directly comparable.

TOPIC MODEL

- **UMAP → HDBSCAN.** Density-cluster into fine topics; no preset number.
- **c-TF-IDF + curation.** Distinctive terms name each topic; all hand-checked, paratext flagged.
- **Agglomerative broad areas.** Fine topics nest into 14 interpretable areas.
- **Fit once.** Trends aggregate fixed labels by year — never refitting.

37 fine topics · 14 broad areas · model fit once on the full census

How accounting is studied



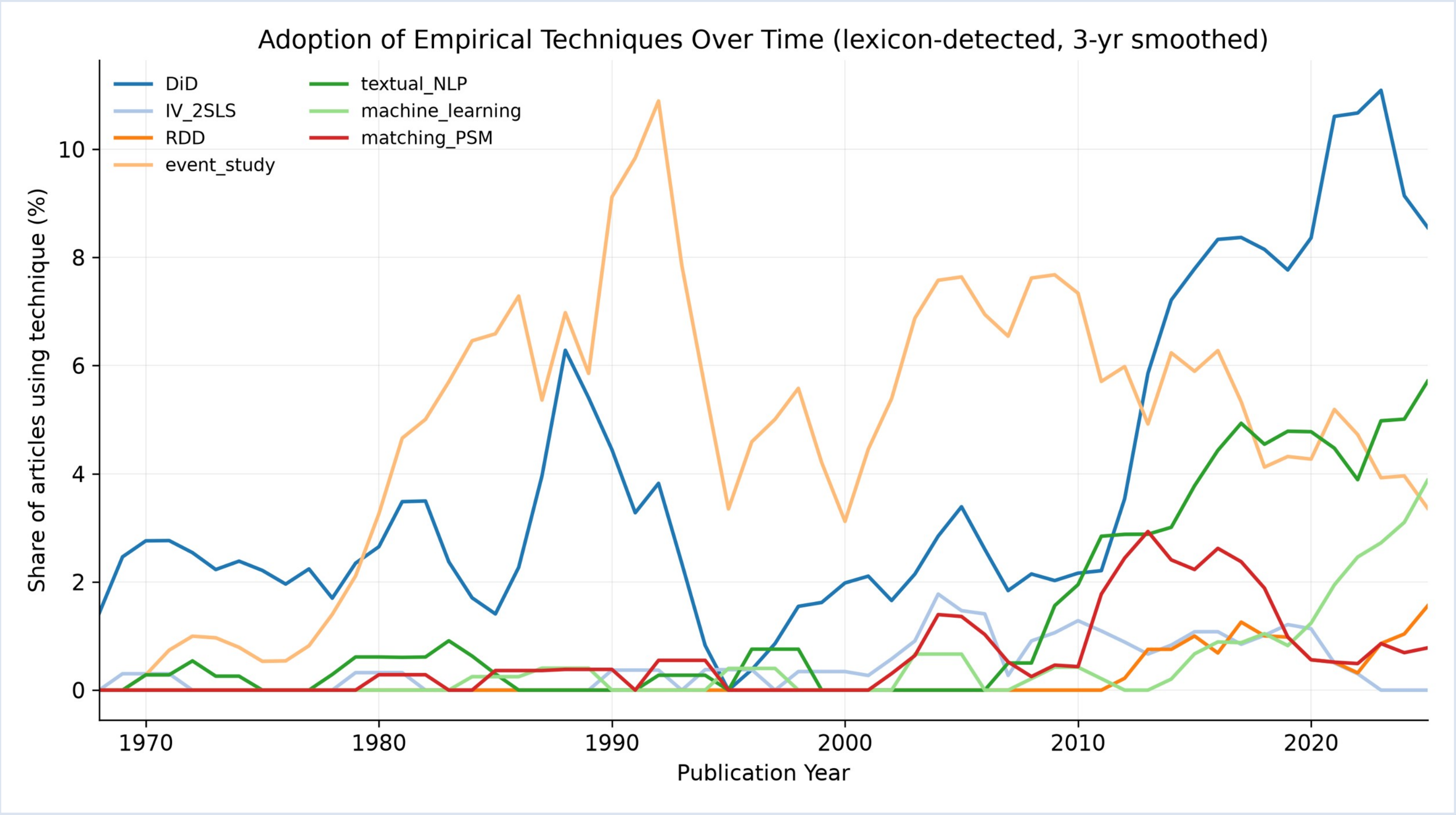
65% Archival empirical

13% Experimental

11% Analytical / theory

Classified by a transparent keyword lexicon, then an LLM pass on the hard cases — only 3% remain unclassified.

The empirical toolkit is changing



- **Causal-inference designs** (DiD, IV, RDD) climb sharply post-2005.
- **Textual analysis & NLP** emerge as a distinct modern toolkit.
- **Machine learning** appears in the most recent decade.
- **Tags are a lower bound** — keyword-based, so conservative.

CREDIBILITY

Stress-tested by an adversarial review

Method mix was a mirage

60% silently defaulted → fixed recall + LLM pass

‘ESG #1’ was an artifact

degenerate age feature → recomputed over full history

Rankings could be noise

permutation null + held-out validation back ESG, Proprietary Disclosure & Fraud

Working papers weren't accounting

finance/law bulk → filtered to a clean accounting set

‘WPs lead’ was overclaimed

WPs scoped to a 5-yr activity snapshot; no lead-lag claim

Frontier ranking is robust

stable to leave-one-out & every fix above

What the map tells us

- **ESG, disclosure tone & proprietary disclosure are the frontier.** Young, fast-growing, citation-rich.
- **Accruals, conservatism & classic valuation are cooling.** Once-dominant, now declining shares.
- **The field is still ~65% archival — but the toolkit is shifting.** Causal inference & textual analysis rising fast.
- **Working papers show where attention is heading now.** A recent (5-yr) accounting WP snapshot — a current-activity layer, not a forecast.
- **The method is transparent, reproducible & stress-tested.** Every claim recomputed; an adversarial review hardened it.

AT A GLANCE

7,173
articles mapped

37
topics, 14 areas

#1
ESG & carbon

4,099
recent WPs (5-yr)

Thank you

Questions & discussion welcome.

A fully interactive HTML report — every figure, table and the adversarial review — accompanies this deck.

7,173	**37**	**4,099**	**#1 ESG**
articles	topics	working papers	frontier topic