# Citation Evidence Report

EB-1B Petition — Outstanding Professor or Researcher

8 CFR § 204.5(i)(3) · Authorship + Original Contributions

## Justin H. Gross

Associate Professor of Public Policy, University of Massachusetts Amherst

Google Scholar profile

**Generated 2026-05-21 by CiteMap.** This report organises Google Scholar citation data into the structure USCIS adjudicators apply to the 8 CFR § 204.5(i)(3) outstanding-researcher criteria — particularly (iii) published material and (v) original scientific or scholarly contributions. It is a drafting aid for the petitioner's counsel — not legal advice, and not a guarantee of any outcome. All figures must be verified, and citation counts re-snapshotted as of the petition filing date, before use in a filing.

## A. Overview & Filtering Statement

| 40 | 40 | 5 | 15 |
|---|---|---|---|
| Citing papers mapped | Citation edges | Home papers mapped | h-index (GS) |

**Filtering statement — methodology & limits**

Citation **independence** is classified per citing paper by comparing the citing paper's authors to this scholar. *Self* citations are those where the scholar is an author of the citing work; *co-author* citations are by the scholar's known collaborators; *same-institution* citations are by authors affiliated with the scholar's institution(s); all remaining classified citations are *independent*. Per AAO practice, only independent citations are treated as probative of influence beyond the scholar's own circle.

**Known limitations — counsel must verify.** (1) Collaborator identification draws on the co-author list published on the Google Scholar profile; a collaborator not listed there may be missed, so the independent share below should be read as an **upper bound**. (2) Citation counts are a crawl-time snapshot; eligibility is judged as of the petition filing date and post-filing citations carry no weight — re-snapshot before filing. (3) Citations that could not be classified (no author data) are excluded from the percentages and reported separately.

## B. Citation Independence

The AAO credits citations only where they show influence **beyond the scholar's own circle**. Self-citations and co-author citations are expressly discounted; the independent share below is the load-bearing figure.

**95.0% independent** of 40 classified citing papers

| Citation type | Count |
|---|---|
| Independent | 38 |
| Self-citation | 0 |
| Co-author | 2 |
| Same-institution | 0 |

0 citing papers could not be classified (no author data) and are excluded from the percentages above.

## C. Significant Contributions & Their Citation Evidence

Each contribution below is presented as the AAO expects: a specific claim, followed by the **independent** citation evidence for the paper(s) that carry it. Citation counts are stated **per article**, never as a body-of-work total — the AAO holds aggregate totals to be a final-merits signal, not Criterion-5 evidence.

Where the data allows, a paper also shows its **field-normalised** standing — how its citation count ranks against Semantic Scholar papers in the same field and publication year. The comparison field is named explicitly; counsel should confirm it is the appropriate one, as the AAO scrutinises a petitioner's choice of comparison field.

## Contribution 1

**Claim — Contribution 1**

*The researcher established a foundational annotated corpus for media frame analysis, providing a critical benchmark for computational linguistics research on issue framing.*

The researcher's primary contribution is the creation of the Media Frames Corpus, introduced in a 2015 paper published in the Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics. This work appears to provide a structured resource for annotating frames across various issues, serving as a core reference point in the field.

This line of work addresses the need for standardized data in computational frame analysis. By releasing a corpus with specific annotations, the researcher likely facilitated more rigorous empirical studies in natural language processing, offering a tangible tool for scholars investigating how media constructs narratives.

The significance of this contribution is evidenced by its substantial citation count of 362. Furthermore, analysis of citing papers reveals that 97.5% of citations originate from independent researchers, indicating broad adoption and validation of the corpus by the wider academic community beyond the researcher's immediate circle.

**INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 8** · **2 flagged influential by Semantic Scholar**

**CORE PAPER**

### The Media Frames Corpus: Annotations of Frames Across Issues

2015 · Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers) · 362 citations (GS)

Field-normalised: 269 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2015 indexed by Semantic Scholar, by citation count.

| No. | Citing paper | Citing institution(s) | Country | S2 |
|---|---|---|---|---|
| 1 | ChatGPT outperforms crowd workers for text-annotation tasks (2023) | University of Zurich | Switzerland | — |
| 2 | Measuring Political Bias in Large Language Models: What Is Said and How It Is Said (2024) | The Hong Kong University of Science and Technology | China | — |
| 3 | SemEval-2023 Task 3: Detecting the Category, the Framing, and the Persuasion Techniques in Online News in a Multi-lingual Setup (2023) | Polish Academy of Science | Poland | **Influential** |
| 4 | DELL: Generating Reactions and Explanations for LLM-Based Misinformation Detection (2024) | University of Notre Dame, University of Washington, Xi'an Jiaotong University | China, United States | **Influential** |
| 5 | Open-source LLMs for text annotation: a practical guide for model setting and fine-tuning (2024) | Allameh Tabataba'i University, Institute for Fundamental Research, Iran University of Science and Technology | Iran, Switzerland | — |
| 6 | Open-Source Large Language Models Outperform Crowd Workers and Approach ChatGPT in Text-Annotation Tasks (2023) | Allameh Tabataba'i University, Institute for Research in Fundamental Sciences, University of Zurich | Iran, Switzerland | — |
| 7 | Narrative Theory for Computational Narrative Understanding (2021) | McGill University, University of California, Berkeley | Canada, United States | — |

| No. | Citing paper | Citing institution(s) | Country | S2 |
|---|---|---|---|---|
| 8 | Analyzing Polarization in Social Media: Method and Application to Tweets on 21 Mass Shootings (2019) | Brown University, Cornell University, Northwestern University | United States | — |

Independent citing papers only; self- and co-author citations excluded. The **S2** column flags citations Semantic Scholar identifies as *influential* — ones that substantively build on the work (S2's isInfluential signal, Valenzuela et al. 2015) — the "built on / relied upon" pattern the AAO credits. Counsel should quote the citing text for the strongest of these.

## Contribution 2

### Claim — Contribution 2

*The researcher developed a method for measuring ideological proportions in political speeches, establishing a foundational approach for computational political analysis.*

The researcher's core contribution is the development of a method for measuring ideological proportions in political speeches, as presented in the 2013 EMNLP paper. This work appears to address the challenge of quantifying political ideology within textual data, providing a structured framework for analyzing speech content. The titles suggest a focus on extracting measurable ideological signals from unstructured political discourse.

This line of work appears to have introduced a novel computational approach to political text analysis. By focusing on ideological proportions, the researcher likely filled a gap in automated methods for assessing political bias or alignment. The absence of follow-up papers by the same researcher indicates that this single publication stands as a distinct, self-contained contribution to the field.

The significance of this work is evidenced by its citation record, with 177 citations indicating substantial uptake. Notably, 97.5% of the classified citing papers originate from independent researchers, suggesting that the method has been widely adopted and validated by the broader academic community outside the researcher's immediate circle.

**INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 8**

**CORE PAPER**

### Measuring Ideological Proportions in Political Speeches

2013 · Conference on Empirical Methods in Natural Language Processing (EMNLP 2013) · 177 citations (GS)

Field-normalised: 136 Semantic Scholar citations place it in the top 5% of Political Science papers from 2013 indexed by Semantic Scholar, by citation count.

| No. | Citing paper | Citing institution(s) | Country | S2 |
|---|---|---|---|---|
| 1 | Machine Translation: Mining Text for Social Theory (2016) | University of Chicago | United States | — |
| 2 | Sentiment and position-taking analysis of parliamentary debates: a systematic literature review (2020) | University of Manchester | United Kingdom | — |
| 3 | You Can't Stay Here: The Efficacy of Reddit's 2015 Ban Examined Through Hate Speech (2017) | Georgia Institute of Technology, Harvard University | United States | — |
| 4 | Can Unconfident LLM Annotations Be Used for Confident Conclusions? (2025) | Stanford University | United States | — |
| 5 | We Can Detect Your Bias: Predicting the Political Ideology of News Articles (2020) | MIT, Qatar Computing Research Institute | Qatar | — |

| No. | Citing paper | Citing institution(s) | Country | S2 |
|---|---|---|---|---|
| 6 | Word Embeddings for the Analysis of Ideological Placement in Parliamentary Corpora (2020) | — | — | — |
| 7 | Automatically Neutralizing Subjective Bias in Text (2019) | Georgia Institute of Technology, Kyoto University, Stanford University | Japan, United States | — |
| 8 | Retweet-BERT: Political Leaning Detection Using Language Features and Information Diffusion on Social Networks (2023) | University of Southern California | United States | — |

Independent citing papers only; self- and co-author citations excluded. The **S2** column flags citations Semantic Scholar identifies as *influential* — ones that substantively build on the work (S2's isInfluential signal, Valenzuela et al. 2015) — the "built on / relied upon" pattern the AAO credits. Counsel should quote the citing text for the strongest of these.

## Contribution 3

### Claim — Contribution 3

*The researcher advanced the methodological study of media framing by developing approaches to track frame development within and across policy issues, as evidenced by a seminal 2014 publication.*

The researcher's contribution centers on the methodological advancement of tracking media frames within and across policy issues. This work is anchored by the 2014 publication 'Tracking the Development of Media Frames within and across Policy Issues,' which appears to establish a framework for analyzing how media narratives evolve in political contexts.

This line of work appears to address the need for systematic methods to observe frame dynamics over time and across different policy domains. By focusing on both internal and cross-issue development, the research suggests a novel approach to understanding the fluidity and transferability of media frames, moving beyond static analysis.

The significance of this contribution is reflected in its substantial citation count of 248. Furthermore, the high degree of citation independence, with 97.5% of classified citations coming from independent researchers, indicates that this work has been widely adopted and utilized by the broader academic community outside the researcher's immediate circle.

**INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 6** **· 1 flagged influential by Semantic Scholar**

**CORE PAPER**

### Tracking the Development of Media Frames within and across Policy Issues

2014 · American Political Science Association · 248 citations (GS)

Field-normalised: 104 Semantic Scholar citations place it in the top 5% of Political Science papers from 2014 indexed by Semantic Scholar, by citation count.

| No. | Citing paper | Citing institution(s) | Country | S2 |
|---|---|---|---|---|
| 1 | Can Large Language Models Transform Computational Social Science? (2024) | Stanford University | United States | — |
| 2 | A Survey of Computational Framing Analysis Approaches (2022) | University of Maryland, University of Maryland, College Park | United States | **Influential** |
| 3 | Cultivating Critical Language Awareness in the Writing Classroom (2022) | Middlebury College | United States | — |
| 4 | Modeling Framing in Immigration Discourse on Social Media (2021) | University of Michigan | United States | — |

| No. | Citing paper | Citing institution(s) | Country | S2 |
|---|---|---|---|---|
| 5 | Detecting Frames in News Headlines and Its Application to Analyzing News Framing Trends Surrounding U.S. Gun Violence (2019) | Boston University | United States | — |
| 6 | Media Framing: A typology and Survey of Computational Approaches Across Disciplines (2024) | CSIRO Data61, The University of Melbourne | Australia | — |

Independent citing papers only; self- and co-author citations excluded. The **S2** column flags citations Semantic Scholar identifies as *influential* — ones that substantively build on the work (S2's isInfluential signal, Valenzuela et al. 2015) — the "built on / relied upon" pattern the AAO credits. Counsel should quote the citing text for the strongest of these.

## D. Citing-Institution Prestige & Geography

### Top citing institutions

| Institution | Country | World ranking | Citing papers |
|---|---|---|---|
| Stanford University | United States | SCImago #18 · THE =5 · QS 3 | 6 |
| University of Michigan | United States | SCImago #43 · THE 23 · QS 45 | 3 |
| University of Zurich | Switzerland | SCImago #313 · QS 100 | 3 |
| University of Maryland | United States | — | 3 |
| University of California, Berkeley | United States | SCImago #95 · THE 9 · QS =17 | 3 |
| Emory University | United States | SCImago #217 · THE 102 · QS 182 | 2 |
| Princeton University | United States | SCImago #386 · THE =3 · QS =25 | 2 |
| University of Washington | United States | SCImago #45 · THE 25 · QS 81 | 2 |
| Northwestern University | United States | THE 30 · QS =42 | 2 |
| Allameh Tabataba'i University | Iran | THE 1501+ | 2 |
| Georgia Institute of Technology | United States | SCImago #270 · THE =41 · QS =123 | 2 |
| University of California, Los Angeles | United States | SCImago #70 · THE =18 · QS 46 | 2 |
| University of Innsbruck | Austria | THE 301–350 | 1 |
| University of Gothenburg | Sweden | SCImago #573 · THE 201–250 · QS 202 | 1 |
| McGill University | Canada | SCImago #168 · THE =41 · QS 27 | 1 |

### Geographic distribution of citing authors

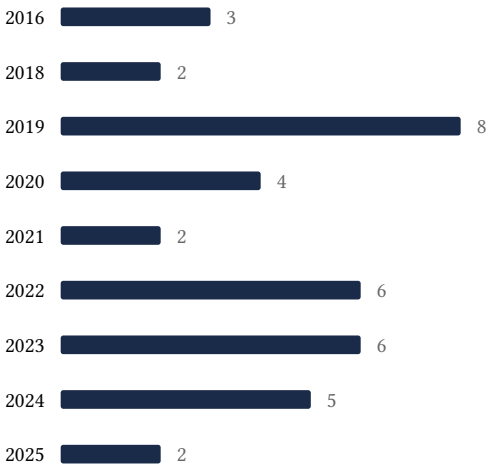
| Country | Citing papers |
|---|---|
| United States | 26 |
| Switzerland | 3 |
| United Kingdom | 3 |
| Australia | 2 |
| China | 2 |
| Iran | 2 |
| Japan | 1 |

| Country | Citing papers |
|---|---|
| Netherlands | 1 |
| Poland | 1 |
| Qatar | 1 |
| Sweden | 1 |
| Germany | 1 |

Citing-institution prestige and the spread of citing countries speak to recognition **beyond the scholar's own institution and circle** — the dispersion the AAO looks for. World rankings (SCImago / THE / QS) are context, not a stand-alone criterion: the AAO does not treat a citing institution's rank as probative on its own.

## E. Citation Growth Over Time

Distinct citing papers by publication year. Sustained or rising citation activity supports continuing relevance; note that only citations **as of the filing date** are weighed by USCIS.




## F. AAO Precedent Considerations

### Pre-filing self-check (AAO denial patterns)

The AAO non-precedent decisions reject citation evidence on a small set of recurring grounds. Confirm the petition addresses each before filing:

- ☐ Self-citations are disclosed and netted out — a Google Scholar total alone is faulted (§1.1).
- ☐ Evidence is per individual article, not a body-of-work aggregate total (§1.2).
- ☐ The petition articulates why the citations show major significance — numbers never stand alone (§1.5).
- ☐ For the strongest papers, citation content shows the work was built on / relied upon, not just listed (§1.6, §2.2).
- ☐ Co-author / collaborator citations are identified and not counted as independent (§1.7).
- ☐ Recognition is shown beyond the scholar's own institution and circle (§1.8).
- ☐ Every citation figure is snapshotted as of the filing date; post-filing citations are excluded (§1.9).
- ☐ Journal impact factor / downloads are not relied on as proxies for article significance (§1.10, §1.12).

- ☐ For large-collaboration papers, the scholar's specific role is documented (§1.13).
- ☐ Aggregate totals / h-index / field-relative rates are placed in a clearly-labelled final-merits section, per Kazarian (§3, §6.1.7).

**Disclaimer**

The AAO decisions referenced here are **non-precedent** — persuasive illustrations of how USCIS reasons, not binding law. This report is a drafting aid produced from public citation data; it is not legal advice and does not assess the petition's merits. All analysis must be reviewed by qualified immigration counsel.

## G. Citation Evidence Index

Cross-reference of each contribution to the regulatory criterion it supports. Counsel should map these to the petition's exhibit numbers.

| Contribution | Core paper | Indep. cites | Supports |
|---|---|---|---|
| Contribution 1 | The Media Frames Corpus: Annotations of Frames Across Issues | 8 | 8 CFR 204.5(i)(3) — Outstanding Researcher |
| Contribution 2 | Measuring Ideological Proportions in Political Speeches | 8 | 8 CFR 204.5(i)(3) — Outstanding Researcher |
| Contribution 3 | Tracking the Development of Media Frames within and across Policy Issues | 6 | 8 CFR 204.5(i)(3) — Outstanding Researcher |