

Citation Evidence Report

EB-1A Petition — Original Contributions of Major Significance

8 CFR § 204.5(h)(3)(v) · Criterion 5

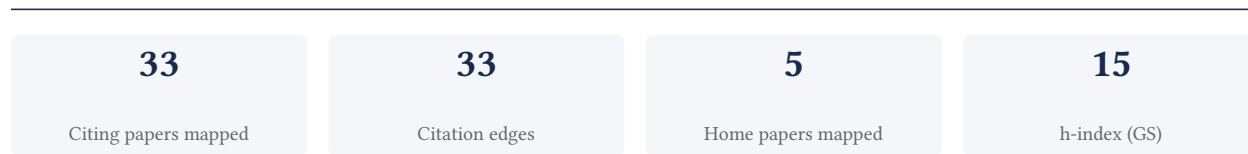
Michael V Arnold

University of Vermont

[Google Scholar profile](#)

Generated 2026-05-21 by CiteMap. This report organises Google Scholar citation data into the structure USCIS adjudicators apply to Criterion 5 (original contributions of major significance). 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



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.

78.8% independent of 33 classified citing papers

Citation type	Count
Independent	26
Self-citation	4
Co-author	2
Same-institution	1

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 multilingual framework for analyzing global pandemic attention using Twitter n-gram time series across 24 languages.

The researcher’s contribution centers on the 2021 paper titled 'How the world’s collective attention is being paid to a pandemic: COVID-19 related n-gram time series for 24 languages on Twitter.' This work appears to provide a systematic method for tracking public discourse during the pandemic through linguistic analysis of social media data across multiple languages.

This line of work addresses the need for granular, multilingual insights into global public sentiment during a crisis. By focusing on n-gram time series, the research suggests a novel approach to quantifying how attention shifts over time, moving beyond single-language or aggregate metrics to capture nuanced, cross-cultural patterns in digital communication.

The significance of this contribution is evidenced by its citation record, with 94 citations indicating substantial uptake in the field. Notably, 81.8% of these citations originate from independent researchers, suggesting that the methodology and findings have been widely adopted and validated by the broader academic community outside the researcher’s immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 10

CORE PAPER

[How the world’s collective attention is being paid to a pandemic: COVID-19 related n-gram time series for 24 languages on Twitter](#)

2021 · 94 citations (GS)

Field-normalised: 67 Semantic Scholar citations place it in the top 10% of Computer Science papers from 2021 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Exploring the challenges of remote work on Twitter users' sentiments: From digital technology development to a post-pandemic era (2022)	Rey Juan Carlos University, University of Alcalá	Spain	—
2	Political polarization drives online conversations about COVID-19 in the United States (2020)	University of Southern California	United States	—
3	The use of Twitter by state leaders and its impact on the public during the COVID-19 pandemic (2020)	—	—	—
4	What Types of COVID-19 Conspiracies Are Populated by Twitter Bots? (2020)	University of Southern California	United States	—
5	Exploratory Analysis of Covid-19 Tweets using Topic Modeling, UMAP, and DiGraphs (2020)	—	—	—
6	Online mental health services in Indonesia during the COVID-19 outbreak (2020)	Universitas Negeri Padang, Universitas Pendidikan Ganesha	Indonesia	—
7	COVID-19: Detecting Government Pandemic Measures and Public Concerns from Twitter Arabic Data Using Distributed Machine Learning (2021)	King Abdulaziz University	Saudi Arabia	—
8	Understanding the sentiment associated with cultural ecosystem services using images and text from social media (2024)	Ecole Polytechnique Fédérale de Lausanne, Inria, Wageningen University	France, Netherlands, Switzerland	—

No.	Citing paper	Citing institution(s)	Country	S2
9	Navigating Post-COVID-19 Social-Spatial Inequity: Unravelling the Nexus between Community Conditions, Social Perception, and Spatial Differentiation (2024)	Institute of Urban Environment, Chinese Academy of Sciences, Peking University	China	—
10	#COVID-19 on Twitter: Bots, Conspiracies, and Social Media Activism (2020)	—	—	—

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 Storywrangler, a large-scale exploratorium for analyzing sociolinguistic and political timelines using Twitter data, published in Science Advances.

The researcher’s core contribution is the development of Storywrangler, a massive exploratorium designed for analyzing sociolinguistic, cultural, socioeconomic, and political timelines using Twitter data. This work was published in Science Advances in 2021 and stands as a seminal piece in this specific line of inquiry, with no subsequent follow-up papers by the same researcher building directly upon it.

This line of work appears to address the need for comprehensive tools to explore complex social dynamics through social media data. By creating a dedicated exploratorium, the researcher provided a novel framework for examining how narratives evolve across different societal dimensions, suggesting a shift toward more integrated, large-scale analytical platforms in digital sociology.

The significance of this contribution is evidenced by its citation record, with 67 citations indicating substantial uptake in the field. Notably, 81.8% of the classified citing papers originate from independent researchers, demonstrating that the work has resonated beyond the researcher’s immediate circle and has been adopted by the broader academic community for independent study and application.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 4

CORE PAPER

[Storywrangler: A massive exploratorium for sociolinguistic, cultural, socioeconomic, and political timelines using Twitter](#)

2021 · Science Advances · 67 citations (GS)

Field-normalised: 47 Semantic Scholar citations place it in the top 10% of Sociology papers from 2021 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Evolving linguistic divergence on polarizing social media (2024)	Tallinn University	Estonia	—
2	Entropy and type-token ratio in gigaword corpora (2025)	Institute for Cross-Disciplinary Physics and Complex Systems (IFISC)	Spain	—
3	Prediction of changes in war-induced population and CO2 emissions in Ukraine using social media (2024)	Northeastern University, University of Extremadura	Spain, United States	—

No.	Citing paper	Citing institution(s)	Country	S2
4	Evol project: a comprehensive online platform for quantitative analysis of ancient literature (2024)	Peking University	China	—

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 established a framework for measuring temporal and social contagion dynamics across over 150 languages on Twitter from 2009 to 2020.

The researcher’s contribution centers on a 2021 paper titled ‘The growing amplification of social media: Measuring temporal and social contagion dynamics for over 150 languages on Twitter for 2009–2020.’ This work appears to provide a comprehensive analysis of how information spreads across diverse linguistic communities on a major social platform over an extended period.

This line of work addresses the need for large-scale, multilingual empirical evidence regarding social contagion. By covering over 150 languages and a decade of data, the research suggests a novel approach to understanding the temporal dynamics of amplification, moving beyond single-language or short-term studies to capture broader global patterns.

The significance of this contribution is evidenced by its citation record, with 95 citations indicating substantial uptake in the field. Notably, 81.8% of the classified citing papers originate from independent researchers, suggesting that the work has influenced scholars outside the researcher’s immediate network and institution, thereby demonstrating broad independent recognition.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 0

CORE PAPER

[The growing amplification of social media: Measuring temporal and social contagion dynamics for over 150 languages on Twitter for 2009–2020](#)

2021 · 95 citations (GS)

Field-normalised: 72 Semantic Scholar citations place it in the top 5% of Sociology papers from 2021 indexed by Semantic Scholar, by citation count.

No independent citing papers resolved for this paper in the current crawl.

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
University of Vermont	United States	SCImago #2315 · QS 1001-1200	5
Northeastern University	United States	QS 384	4
University of Southern California	United States	SCImago #192 · THE =73 · QS 146	2
MassMutual	United States	—	2
Peking University	China	SCImago #11 · THE 13 · QS 14	2

Institution	Country	World ranking	Citing papers
Arizona State University	United States	SCImago #357 · THE 201–250 · QS =173	2
Universitas Negeri Padang	Indonesia	SCImago #3941 · THE 1501+ · QS 1401+	1
Universitas Pendidikan Ganesha	Indonesia	—	1
Institute for Cross-Disciplinary Physics and Complex Systems (IFISC)	Spain	—	1
Massachusetts Institute of Technology	United States	SCImago #41 · THE 2 · QS 1	1
University of Missouri	United States	—	1
University of Pittsburgh	United States	SCImago #212 · QS =281	1
Columbia University	United States	SCImago #65 · THE 20 · QS =38	1
University of Extremadura	Spain	—	1
Tufts University	United States	SCImago #974 · THE 189 · QS =334	1

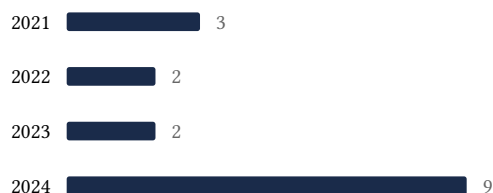
Geographic distribution of citing authors

Country	Citing papers
United States	17
Australia	3
Spain	3
China	2
Indonesia	1
Netherlands	1
Saudi Arabia	1
Switzerland	1
United Kingdom	1
France	1
Estonia	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	How the world's collective attention is being paid to a pandemic: COVID-19 related n-gram time series for 24 languages on Twitter	10	8 CFR 204.5(h)(3)(v) – Criterion 5
Contribution 2	Storywrangler: A massive exploratorium for sociolinguistic, cultural, socioeconomic, and political timelines using Twitter	4	8 CFR 204.5(h)(3)(v) – Criterion 5
Contribution 3	The growing amplification of social media: Measuring temporal and social contagion dynamics for over 150 languages on Twitter for 2009–2020	0	8 CFR 204.5(h)(3)(v) – Criterion 5