Citation Evidence Report

EB-1A Petition — Original Contributions of Major Significance

8 CFR § 204.5(h)(3)(v) · Criterion 5

Jane Lydia Adams

PhD Student, Khoury College of Computer Sciences, Northeastern University, Boston Massachusetts

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

31	31	4	13
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.

83.9% independent of 31 classified citing papers

Citation type	Count
Independent	26
Self-citation	1
Co-author	3
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 via Twitter n-gram time series across 24 languages, providing a foundational dataset for digital epidemiology.

The researcher's core contribution rests on the 2021 PLOS ONE 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 introduce a systematic approach to tracking public discourse during the pandemic using linguistic data from social media platforms.

This line of work addresses the need for granular, multilingual insights into global public sentiment during health crises. By focusing on n-gram time series across 24 languages, the research suggests a novel method for capturing cross-cultural variations in attention and discourse, moving beyond single-language or aggregate analyses.

The significance of this contribution is evidenced by its citation record, with 94 citations indicating substantial uptake in the field. Notably, 90.3% of these citations originate from independent researchers, suggesting that the work has served as a widely adopted reference point for scholars outside the researcher's immediate network, thereby demonstrating broad independent impact.

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 · PLOS ONE · 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	The use of Twitter by state leaders and its impact on the public during the COVID-19 pandemic (2020)	—	—	—
3	What Types of COVID-19 Conspiracies Are Populated by Twitter Bots? (2020)	University of Southern California	United States	**Result**
4	Exploratory Analysis of Covid-19 Tweets using Topic Modeling, UMAP, and DiGraphs (2020)	—	—	—
5	Online mental health services in Indonesia during the COVID-19 outbreak (2020)	Universitas Negeri Padang, Universitas Pendidikan Ganesha	Indonesia	—
6	COVID-19: Detecting Government Pandemic Measures and Public Concerns from Twitter Arabic Data Using Distributed Machine Learning (2021)	King Abdulaziz University	Saudi Arabia	**Methodology**

No.	Citing paper	Citing institution(s)	Country	S2
7	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	—
8	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	—
9	#COVID-19 on Twitter: Bots, Conspiracies, and Social Media Activism (2020)	—	—	Result
10	ArCOV-19: The First Arabic COVID-19 Twitter Dataset with Propagation Networks (2021)	—	—	—

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology / Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

Citing-text excerpts — how the field used this work

RESULT What Types of COVID-19 Conspiracies Are Populated by Twitter Bots?

"As of the time of this writing (mid-April 2020), the vast majority of these studies are pre-print papers that provide a timely, yet partial, characterization of online discussion and issues revolving around COVID-19 (Alshaabi et al., 2020; Chen et al., 2020; Cinelli et al., 2020; Gallotti et al., 2020; Gao et al., 2020; Kleinberg et al., 2020; Li et al., 2020; Pennycook et al., 2020; Schild et al., 2020; Singh et al., 2020)."

METHODOLOGY COVID-19: Detecting Government Pandemic Measures and Public Concerns from Twitter Arabic Data Using Distributed Machine Learning

"The second method is based on Twitter data we have, where it can give us the detailed information in addition to space and time information, particularly if it was posted by an official news account such as @spagov or the account of Ministry of Health."

RESULT #COVID-19 on Twitter: Bots, Conspiracies, and Social Media Activism

"As of the time of this writing (mid-April 2020), the vast majority of these studies are pre-print papers that provide a timely, yet partial, characterization of online discussion and issues revolving around COVID-19 (Alshaabi et al., 2020; Chen et al., 2020; Cinelli et al., 2020; Gallotti et al., 2020; Gao et al., 2020; Kleinberg et al., 2020; Li et al., 2020; Pennycook et al., 2020; Schild et al., 2020; Singh et al., 2020)."

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 contribution centers on 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 standalone core contribution without direct follow-up papers by the same author in the provided record.

This line of work appears to address the need for scalable tools to explore complex social dynamics within large-scale social media datasets. By framing the contribution as an 'exploratorium,' the research suggests a novel approach to enabling broad, interactive analysis of temporal patterns in public discourse, moving beyond static metrics to dynamic timeline exploration.

The significance of this work is evidenced by its citation record, with 67 citations indicating strong uptake in the field. Notably, 90.3% of the classified citing papers originate from independent researchers, suggesting that the tool has been widely adopted and utilized by the broader academic community outside the researcher's immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 5

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	Doomscrolling during COVID-19: The negative association between daily social and traditional media consumption and mental health symptoms during the COVID-19 pandemic (2022)	University of Vermont	United States	Methodology
2	Evolving linguistic divergence on polarizing social media (2024)	Tallinn University	Estonia	—
3	Entropy and type-token ratio in gigaword corpora (2025)	Institute for Cross-Disciplinary Physics and Complex Systems (IFISC)	Spain	—
4	Evol project: a comprehensive online platform for quantitative analysis of ancient literature (2024)	Peking University	China	—
5	Disentangling the cultural evolution of ancient China: a digital humanities perspective (2023)	—	—	—

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology / Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

Citing-text excerpts — how the field used this work

METHODOLOGY Doomscrolling during COVID-19: The negative association between daily social and traditional media consumption and mental health symptoms during the COVID-19 pandemic

"Data obtained from the Storywrangler project (Alshaabi et al., 2020): https://storywrangling."

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 seminal 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 focusing on over 150 languages and a decade of data, the research suggests a move beyond limited, single-language studies to capture the global dynamics of social media amplification. The absence of follow-up papers by the same researcher indicates this core publication stands as a definitive, self-contained contribution to the field.

The significance of this work is evidenced by its 95 citations, with 90.3% originating from independent researchers. This high rate of independent uptake suggests the methodology or findings have been widely adopted and trusted by the broader academic community, validating the research as a foundational reference for studying cross-linguistic social dynamics.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 5

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.	Citing paper	Citing institution(s)	Country	S2
1	TurkishBERTweet: Fast and Reliable Large Language Model for Social Media Analysis (2023)	Sabanci University	Turkey	Background
2	A perspective on friction interventions to curb the spread of misinformation (2025)	Indiana University	—	—
3	Fostering YouTube followers' stickiness through social contagion: The role of digital influencer' characteristics and followers' compensation psychology (2024)	National Chung Hsing University	—	—
4	The short life of the European Super League: a case study on institutional tensions in sport industries (2022)	Loughborough University	United Kingdom	—
5	Who can verify this? Finding authorities for rumor verification in Twitter (2023)	—	—	—

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology / Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
University of Vermont	United States	SCImago #2315 · QS 1001-1200	3
Inria	France	—	2
Peking University	China	SCImago #11 · THE 13 · QS 14	2
Northeastern University	United States	QS 384	2
Universitas Pendidikan Ganesha	Indonesia	—	1
MassMutual	United States	—	1
Institute for Cross-Disciplinary Physics and Complex Systems (IFISC)	Spain	—	1
University of A Coruña	Spain	THE 1201–1500	1

Institution	Country	World ranking	Citing papers
Massachusetts Institute of Technology	United States	SCImago #41 · THE 2 · QS 1	1
The Hong Kong University of Science and Technology (Guangzhou)	China	SCImago #483 · THE =58 · QS 44	1
University of Extremadura	Spain	—	1
Brown University	United States	SCImago #553 · THE 65 · QS 69	1
University of Southern California	United States	SCImago #192 · THE =73 · QS 146	1
Indiana University	United States	THE =198	1
Institute of Urban Environment, Chinese Academy of Sciences	China	SCImago #2490	1

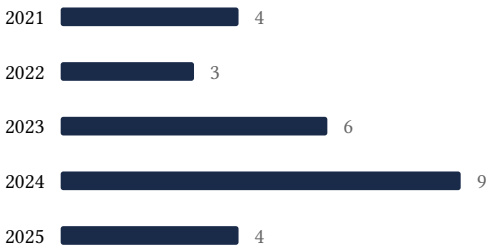
Geographic distribution of citing authors

Country	Citing papers
United States	8
China	4
Spain	4
France	2
Australia	1
Saudi Arabia	1
Switzerland	1
Turkey	1
United Kingdom	1
People's Republic of China	1
Estonia	1
Indonesia	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	5	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	5	8 CFR 204.5(h)(3)(v) — Criterion 5