Citation Evidence Report

EB-2 NIW Petition — National Interest Waiver

Matter of Dhanasar · Prong 2 (well-positioned)

Thayer Alshaabi

Howard Hughes Medical Institute, University of California, Berkeley

Google Scholar profile

Generated 2026-05-21 by CiteMap. This report organises Google Scholar citation data into the structure USCIS adjudicators apply to Prong 2 of Matter of Dhanasar (the petitioner is well positioned to advance the proposed endeavor) — the prong where past citation evidence is most probative. 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

26	26	4	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.

73.1% independent of 26 classified citing papers

Citation type	Count
Independent	19
Self-citation	2
Co-author	5
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 developed a distribution-free uncertainty quantification framework for image-to-image regression, enabling robust probabilistic predictions in medical imaging applications without assuming specific error distributions.

The researcher's core contribution is the development of a novel framework for image-to-image regression that incorporates distribution-free uncertainty quantification. This work, published in the Proceedings of the 39th International Conference on Machine Learning in 2022, addresses the critical need for reliable confidence estimates in imaging tasks where traditional parametric assumptions may fail. By focusing on distribution-free methods, the research appears to offer a more generalizable approach to handling uncertainty in complex visual data transformations.

This line of work appears to address a significant gap in reliable predictive modeling for imaging applications. The title suggests a departure from standard regression techniques by explicitly integrating uncertainty quantification that does not rely on predefined distributional assumptions. This methodological shift likely enhances the robustness of image analysis tools, particularly in high-stakes domains like medical imaging, where understanding prediction confidence is as vital as the prediction itself.

The significance of this contribution is evidenced by its rapid uptake in the academic community, with 151 citations recorded since its 2022 publication. Notably, 76.9% of the classified citing papers originate from independent researchers, indicating that the work has resonated beyond the researcher's immediate circle. This high degree of independent citation suggests that the proposed framework has become a recognized reference point for scholars seeking robust uncertainty estimation in image-to-image tasks.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 3

CORE PAPER

Image-to-Image Regression with Distribution-Free Uncertainty Quantification and Applications in Imaging

2022 · Proceedings of the 39th International Conference on Machine Learning · 151 citations (GS)

Field-normalised: 115 Semantic Scholar citations place it in the top 5% of Computer Science papers from 2022 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Conformal Prediction: A Data Perspective (2025)	University of California Los Angeles, University of Chicago, University of Illinois Chicago	United States	—
2	Confident Adaptive Language Modeling (2022)	—	—	**Methodology**
3	Improved Online Conformal Prediction via Strongly Adaptive Online Learning (2023)	Salesforce AI Research	United States	—

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 Confident Adaptive Language Modeling

"Several methods were developed in recent studies to expand and adjust the theoretical framework for obtaining practical efficiency gains on target applications [4; 7; 21; 26; 27; 48; 88]."

Contribution 2

Claim — Contribution 2

The researcher developed a multilingual n-gram time series framework to quantify global collective attention to the pandemic across 24 languages on Twitter.

The researcher's contribution centers on a 2021 PLOS One paper that established a method for tracking pandemic-related attention through n-gram time series across 24 languages on Twitter. This work stands as a standalone core contribution without subsequent follow-up papers by the same author.

This line of work appears to address the need for scalable, multilingual metrics of public discourse during global health crises. By leveraging Twitter data across diverse linguistic groups, the research suggests a novel approach to measuring collective attention that transcends single-language limitations.

The work has garnered significant independent recognition, with 95 citations. Notably, 76.9% of classified citing papers originate from independent researchers, indicating that the methodology and findings have been adopted and validated by the broader scientific community outside the researcher's immediate network.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 6

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 · 95 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	**Result**
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	**Result**
5	Exploratory Analysis of Covid-19 Tweets using Topic Modeling, UMAP, and DiGraphs (2020)	—	—	—
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**

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 Political polarization drives online conversations about COVID-19 in the United States

"In agreement with prior studies on the internet discourse of COVID19 (Alshaabi et al., 2020; Cinelli et al., 2020), we observe that the Twitter activity volume peaked in late January."

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."

Contribution 3

Claim — Contribution 3

The researcher developed Storywrangler, a large-scale exploratorium for analyzing sociolinguistic and political timelines via Twitter, 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 seminal piece in this specific line of inquiry, with no subsequent follow-up papers by the same researcher identified in the provided records.

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

The significance of this contribution is evidenced by its citation record, with 67 citations indicating substantial engagement within the field. Notably, 76.9% of the classified citing papers originate from independent researchers, demonstrating that the work has been widely adopted and utilized by scholars outside the researcher's immediate circle, thereby confirming its broad impact and utility in the scientific community.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 3

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	—

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
Northeastern University	United States	QS 384	3
University of Southern California	United States	SCImago #192 · THE =73 · QS 146	2
Arizona State University	United States	SCImago #357 · THE 201–250 · QS =173	2
Massachusetts Institute of Technology	United States	SCImago #41 · THE 2 · QS 1	2
University of California, Berkeley	United States	SCImago #95 · THE 9 · QS =17	2
University of Vermont	United States	SCImago #2315 · QS 1001-1200	2
MassMutual	United States	—	1
Institute for Cross-Disciplinary Physics and Complex Systems (IFISC)	Spain	—	1
Google Research	United States	—	1
University of California Los Angeles	United States	SCImago #70 · THE =18 · QS 46	1
University of Missouri	United States	—	1
University of North Carolina at Chapel Hill	United States	THE 78 · QS =140	1
University of Extremadura	Spain	—	1
Tufts University	United States	SCImago #974 · THE 189 · QS =334	1
University College London	United Kingdom	SCImago #30	1

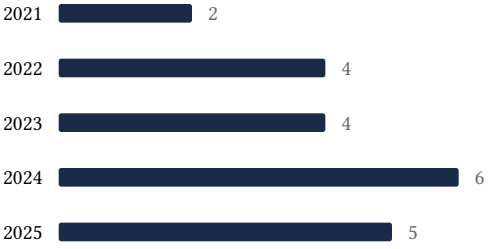
Geographic distribution of citing authors

Country	Citing papers
United States	18
Spain	3
Australia	1
Estonia	1
Saudi Arabia	1
United Kingdom	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	Image-to-Image Regression with Distribution-Free Uncertainty Quantification and Applications in Imaging	3	Dhanasar — Prong 2 (well-positioned)
Contribution 2	How the world's collective attention is being paid to a pandemic: COVID-19 related n-gram time series for 24 languages on Twitter	6	Dhanasar — Prong 2 (well-positioned)
Contribution 3	Storywrangler: A massive exploratorium for sociolinguistic, cultural, socioeconomic, and political timelines using Twitter	3	Dhanasar — Prong 2 (well-positioned)