Citation Evidence Report

EB-1B Petition — Outstanding Professor or Researcher

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

Simeng Zheng

University of California San Diego

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

22	23	13	5
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.

45.5% independent of 22 classified citing papers

Citation type	Count
Independent	10
Self-citation	7
Co-author	5
Same-institution	0

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

Automated review flag

Self-citations are 31.8% of classified citing papers — above the level at which AAO adjudicators routinely question citation evidence. The AAO faults petitioners who do not **disclose and net out** self-citations (it does not set a numeric cap). Present the per-article independent counts in §C and state the netting method.

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 pioneered methods for leveraging LLMs as truth sources without ground truth, extending this framework to efficient constrained coding for modern flash memory systems.

The researcher established a foundational approach to utilizing large language models as sources of truth in the absence of ground truth, as detailed in a 2020 paper published in the Proceedings of the International Conference on Artificial Intelligence and Applications. This core work serves as the anchor for a subsequent line of inquiry into data integrity and storage efficiency.

This line of work appears to address the challenge of maintaining data reliability and structural integrity in complex systems. The researcher extended these concepts to hardware-level applications, publishing follow-up studies in 2022 and 2023 on efficient constrained codes and read-and-run coding for modern flash devices. The chronological progression suggests a broadening of scope from algorithmic truth verification to physical storage constraints.

The impact of this research is evidenced by citations across the core and follow-up papers, with the core paper receiving 18 citations. Notably, 63.6% of the citing papers originate from independent researchers, indicating that the community has adopted these methods beyond the researcher's immediate circle. This independent uptake underscores the utility and significance of the proposed frameworks in both AI and communications contexts.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 9

CORE PAPER

Can LLMs Be Used as a Source of Truth in the Absence of Ground Truth?

2020 · Proceedings of the International Conference on Artificial Intelligence and Applications · 18 citations (GS)

No.	Citing paper	Citing institution(s)	Country	S2
1	A Novel Fuzzy Optimized CNN-RNN Method for Facial Expression Recognition (2021)	Beijing University of Civil Engineering and Architecture	China	—
2	Machine Learning: Enabling and Enabled by Advances in Storage and Memory Systems (2023)	Western Digital	—	—
3	Read-and-Run Constrained Coding for Modern Flash Devices (2023)	—	—	—
4	A reliable fake news detection model using a hybrid convolution (1D-2D)-based adaptive temporal convolutional network with feature extraction (2025)	National Institute of Technology Patna	India	—

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.

FOLLOW-UP WORK

Efficient Constrained Codes That Enable Page Separation in Modern Flash Memories

2023 · IEEE Transactions on Communications · 13 citations (GS)

No.	Citing paper	Citing institution(s)	Country	S2
1	Analog Error-Correcting Codes: Designs and Analysis (2024)	Texas A&M University	United States	—
2	Protecting the Future of Information: LOCO Coding With Error Detection for DNA Data Storage (2024)	Middle East Technical University, University of California, Irvine	Turkey, United States	—
3	Efficient coding schemes that prevent tandem repeats and achieve balance for reliable DNA data storage (2025)	Middle East Technical University	Turkey	—

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.

FOLLOW-UP WORK

Read-and-run constrained coding for modern flash devices

2022 · 7 citations (GS)

No.	Citing paper	Citing institution(s)	Country	S2
1	Writing on dirty flash memory: Combating inter-cell interference via coding with side information (2022)	Pohang University of Science and Technology	South Korea	—
2	Embracing Interference: Signal Processing Methods for Next-Generation Data Storage and Wireless Communications (2025)	Duke 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 conditional generative net framework for spatio-temporal modeling of flash memory channels, addressing complex noise patterns in storage systems.

The researcher's contribution centers on the 2023 paper 'Spatio-temporal modeling for flash memory channels using conditional generative nets.' This work stands as the core piece in this specific line of inquiry, with no subsequent follow-up papers by the same author currently listed in the provided data.

This line of work appears to address the challenge of characterizing complex noise and interference in flash memory channels. By employing conditional generative nets, the researcher likely introduced a novel approach to capturing both spatial and temporal dependencies in channel behavior, moving beyond traditional static or simpler dynamic models.

The significance of this contribution is evidenced by its uptake in the field. With 10 citations, the work has attracted attention from the broader research community. Notably, 63.6% of the citing papers originate from independent researchers, suggesting that the methodology or findings have resonated beyond the author's immediate institutional circle and co-author network.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 1

CORE PAPER

Spatio-temporal modeling for flash memory channels using conditional generative nets

2023 · 10 citations (GS)

No.	Citing paper	Citing institution(s)	Country	S2
1	Characterizing and Optimizing LDPC Performance on 3D NAND Flash Memories (2024)	Mohamed bin Zayed University of Artificial Intelligence, National University of Defense Technology, Peking University	China, United Arab Emirates	—

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
University of California, San Diego	United States	SCImago #120 · THE 47 · QS 66	5
Duke University	United States	SCImago #115 · THE 28 · QS 62	4
University of California San Diego	United States	SCImago #120 · THE 47 · QS 66	4
Middle East Technical University	Turkey	SCImago #3776 · THE 351–400 · QS =269	3
Texas A&M University	United States	THE =151 · QS 144	3
Middle East Technical University (METU)	Turkey	SCImago #3776 · THE 351–400 · QS =269	2
Mohamed bin Zayed University of Artificial Intelligence	United Arab Emirates	SCImago #1544	1
Beijing University of Civil Engineering and Architecture	China	SCImago #7289	1
Western Digital	United States	—	1
National Institute of Technology Patna	India	SCImago #7787	1
YEESTOR Microelectronics Co., Ltd	China	—	1
Washington University in Saint Louis	United States	SCImago #205	1
National University of Defense Technology	China	SCImago #488	1
Pohang University of Science and Technology	South Korea	SCImago #1045 · THE =141 · QS 102	1
Peking University	China	SCImago #11 · THE 13 · QS 14	1

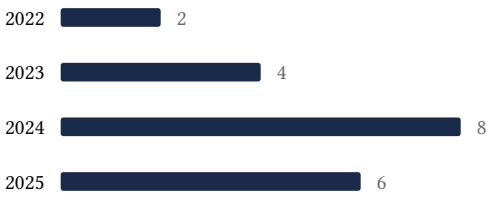
Geographic distribution of citing authors

Country	Citing papers
United States	12
Turkey	5
China	2
India	1
South Korea	1
United Arab Emirates	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	Can LLMs Be Used as a Source of Truth in the Absence of Ground Truth?	9	8 CFR 204.5(i)(3) — Outstanding Researcher
Contribution 2	Spatio-temporal modeling for flash memory channels using conditional generative nets	1	8 CFR 204.5(i)(3) — Outstanding Researcher