Citation Evidence Report

EB-2 NIW Petition — National Interest Waiver

Matter of Dhanasar · Prong 2 (well-positioned)

Yejin Choi

Stanford University / NVIDIA

[Google Scholar profile]

Generated 2026-06-10 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

800	800	30	139
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.

87.4% independent of 779 classified citing papers

Citation type	Count
Independent	681
Self-citation	24
Co-author	74
Same-institution	0

21 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 pioneered the analysis of neural text degeneration, establishing a foundational framework for evaluating toxic outputs in language models through highly cited seminal and follow-up work.

The researcher's contribution centers on identifying and analyzing neural text degeneration, anchored by the 2019 paper 'The curious case of neural text degeneration.' This core work established a critical lens for understanding how neural models generate undesirable or degenerate text, a problem that became increasingly relevant as language models grew in capability and deployment.

Originality in this line of work appears to stem from framing text degeneration as a distinct, systematic failure mode rather than isolated errors. The 2020 follow-up, 'Realtoxicityprompts: Evaluating neural toxic degeneration in language models,' suggests the researcher extended this framework to specifically evaluate toxicity, indicating a progression from general degeneration to measurable harmful outputs. This chronological development implies a novel methodological approach to assessing model safety and reliability.

The significance of this contribution is evidenced by substantial citation counts, with the core paper accumulating 5047 citations and the follow-up 2042 citations. Furthermore, citation independence analysis reveals that 87.4% of citing papers originate from independent researchers, demonstrating that this work has been widely adopted and validated by the broader scientific community beyond the researcher's immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 681 · **129 flagged influential by Semantic Scholar**

CORE PAPER

The curious case of neural text degeneration

2019 · 5,047 citations (GS)

Field-normalised: 4,081 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2019 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Risk taxonomy, mitigation, and assessment benchmarks of large language model systems	Ant Group, Institute of Information Engineering, Chinese Academy of Sciences, Tsinghua University	China	—
2	A survey of large language models	Alibaba Group, Renmin University of China, Université de Montréal	Canada, China, United States	—
3	Large language models are zero-shot time series forecasters	Carnegie Mellon University, New York University	United States	—
4	Ai alignment: A comprehensive survey	Peking University	China	—
5	Unified-io 2: Scaling autoregressive multimodal models with vision language audio and action	Allen Institute for AI, Allen Institute for Artificial Intelligence, University of Illinois at Urbana-Champaign	United States	—
6	A survey on large language models for code generation	NAVER Cloud, The Hong Kong University of Science and Technology, The Hong Kong University of Science and Technology (Guangzhou)	China, South Korea	—

No.	Citing paper	Citing institution(s)	Country	S2
7	Unleashing the potential of prompt engineering for large language models	Beijing Normal University, BNU-HKBU United International College	China	—
8	A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions	Harbin Institute of Technology, Huawei Inc., Zhejiang University	China	—
9	Illuminating protein space with a programmable generative model	Generate Biomedicines	United States	**Influential**
10	Taxonomy of risks posed by language models	DeepMind, Google DeepMind, Google DeepMind (United Kingdom)	United Kingdom	—
11	Large language model inference acceleration: A comprehensive hardware perspective	Infinigence-AI, Shanghai Jiao Tong University, Tsinghua University	China	—
12	Large language models generate functional protein sequences across diverse families	Salesforce Research	United States	—
13	A survey of the usages of deep learning for natural language processing	University of Colorado Colorado Springs	United States	—
14	Pre-trained language models for text generation: A survey	Alibaba Group, Renmin University of China, University of Montreal	Canada, China	**Influential**
15	Recurrent neural networks: A comprehensive review of architectures, variants, and applications	University of California, Irvine Medical Center, University of Johannesburg	South Africa, United States	—
16	A unified sequence interface for vision tasks	Google DeepMind, Google Research, NVIDIA	United Kingdom, United States	—
17	Lift: Language-interfaced fine-tuning for non-language machine learning tasks	University of Wisconsin-Madison	United States	—
18	Can llm be a personalized judge?	University of Cambridge	United Kingdom	**Influential**
19	Longlamp: A benchmark for personalized long-form text generation	Adobe Research, University of Massachusetts Amherst	United States	**Influential**
20	Interactive natural language processing	01.AI, Alibaba Group, Beihang University	Australia, Canada, China	—
21	Lamp-qa: A benchmark for personalized long-form question answering	University of Massachusetts Amherst	United States	**Influential**
22	Layouttransformer: Layout generation and completion with self-attention	Amazon AWS, University of California, Irvine Medical Center, University of Maryland, College Park	United States	—
23	Ccvs: Context-aware controllable video synthesis	ENS, Inria; New York University, Univ. Grenoble Alpes	France	—
24	A survey on llm-generated text detection: Necessity, methods, and future directions	KAUST, Peking University, University of Macau	China	—
25	Scalable watermarking for identifying large language model outputs	Google, Google DeepMind	United Kingdom, United States	—

No.	Citing paper	Citing institution(s)	Country	S2
26	Provable robust watermarking for ai-generated text	Carnegie Mellon University, UC Berkeley, UC Santa Barbara	United States	—
27	Machine-generated text: A comprehensive survey of threat models and detection methods	American University, University of Ottawa	Canada, United States	Influential
28	Dna-gpt: Divergent n-gram analysis for training-free detection of gpt-generated text	Nanjing University of Aeronautics and Astronautics, NEC, University of California, Irvine Medical Center	China, United States	—
29	Authorship attribution in the era of llms: Problems, methodologies, and challenges	Emory University, Northwestern University	United States	—
30	ChatGPT and academic integrity concerns: Detecting artificial intelligence generated content	Bursa Uludag University	Turkey	—

Showing the 30 most-cited of 681 independent citing papers.

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).

FOLLOW-UP WORK

Realtoxicityprompts: Evaluating neural toxic degeneration in language models

2020 · 2,042 citations (GS)

Field-normalised: 1,626 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2020 indexed by Semantic Scholar, by citation count.

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

Contribution 2

Claim — Contribution 2

The researcher developed Hellaswag, a highly cited benchmark that rigorously evaluates whether machines can genuinely complete sentences, establishing a critical standard for commonsense reasoning in language models.

CLAIM: The researcher's primary contribution is the creation of Hellaswag, a seminal benchmark introduced in 2019 that challenges the ability of machines to finish sentences in a way that reflects genuine understanding rather than superficial pattern matching. This work stands as a foundational piece in the evaluation of natural language processing systems.

ORIGINALITY: The title suggests a critical inquiry into the depth of machine comprehension, addressing a gap where existing metrics may have failed to distinguish between true commonsense reasoning and statistical coincidence. By framing the task as a test of whether a machine can 'really' finish a sentence, the researcher appears to have introduced a more rigorous standard for assessing contextual understanding in AI models.

SIGNIFICANCE: With 4,430 citations, this work has achieved substantial recognition within the academic community. Notably, 87.4% of the citing papers originate from independent researchers, indicating that the benchmark has been widely adopted and validated by the broader scientific community beyond the researcher's immediate circle, underscoring its broad impact and utility in advancing the field.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 0

CORE PAPER

[Hellaswag: Can a machine really finish your sentence?]

2019 · 4,430 citations (GS)

Field-normalised: 3,988 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2019 indexed by Semantic Scholar, by citation count.

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

Contribution 3

Claim — Contribution 3

The researcher developed Winogrande, a large-scale adversarial dataset for evaluating commonsense reasoning in natural language understanding systems.

The researcher’s primary contribution is the creation of Winogrande, a seminal dataset introduced in 2021 that scales the Winograd schema challenge to test adversarial robustness in language models. This work addresses the critical gap in evaluating whether AI systems truly understand commonsense reasoning or merely rely on superficial statistical patterns, as suggested by the dataset's adversarial design. The high citation count indicates that this resource has become a standard benchmark in the field, widely adopted by the research community to assess model capabilities. Furthermore, the fact that the vast majority of citations come from independent researchers demonstrates that the work has significantly influenced external scientific inquiry and established a new standard for rigorous evaluation in natural language processing.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 0

CORE PAPER

Winogrande: An adversarial winograd schema challenge at scale

2021 · 3,472 citations (GS)

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 Washington	United States	SCImago #45 · THE 25 · QS 81	52
University of California, Irvine Medical Center	United States	—	46
Tsinghua University	PR China	SCImago #8 · THE 12 · QS =17	45
Carnegie Mellon University	United States	SCImago #266 · THE 24 · QS 52	43
Google DeepMind	United Kingdom	SCImago #90	43
Stanford University	United States	SCImago #18 · THE =5 · QS 3	42
New York University	United States	SCImago #116 · THE =31 · QS 55	31
Allen Institute for AI	United States	—	27
Google Research	United States	—	26
Peking University	China	SCImago #11 · THE 13 · QS 14	25

Institution	Country	World ranking	Citing papers
UC Berkeley	United States	—	22
Meta AI	United States	—	22
Google	United States	—	21
NVIDIA	United States	—	20
Nanyang Technological University	Singapore	SCImago #137	20

Geographic distribution of citing authors

Country	Citing papers
United States	468
China	199
United Kingdom	101
Singapore	39
Canada	36
Switzerland	33
Hong Kong	26
Germany	23
France	23
Australia	22
South Korea	22
India	15

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.

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 curious case of neural text degeneration	681	Dhanasar — Prong 2 (well-positioned)
Contribution 2	Hellaswag: Can a machine really finish your sentence?	0	Dhanasar — Prong 2 (well-positioned)
Contribution 3	Winogrande: An adversarial winograd schema challenge at scale	0	Dhanasar — Prong 2 (well-positioned)