Citation Evidence Report

EB-1A Petition — Original Contributions of Major Significance

8 CFR § 204.5(h)(3)(v) · Criterion 5

Yonghui Wu

Head of Research, ByteDance Seed

Google Scholar profile

Generated 2026-06-08 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

2,780	2,821	30	85
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.

93.5% independent of 2,702 classified citing papers

Citation type	Count
Independent	2,527
Self-citation	16
Co-author	159
Same-institution	0

78 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 advanced neural machine translation and scalable deep learning infrastructure, bridging human-machine translation gaps and enabling efficient training of giant models.

The researcher’s core contribution rests on the 2016 paper 'Google's neural machine translation system: Bridging the gap between human and machine translation,' which appears to have established a significant benchmark in the field. This work is supported by a trajectory of follow-up research addressing scalability and efficiency in large-scale neural networks.

This line of work appears to address the challenge of deploying and training increasingly complex neural architectures. The 2019 paper 'Gpipe: Efficient training of giant neural networks using pipeline parallelism' suggests a focus on infrastructure efficiency, while the 2022 paper 'Scaling autoregressive models for content-rich text-to-image generation' indicates an extension of these scaling principles to multimodal generative tasks.

The significance of this research is evidenced by substantial citation counts, with the core paper accumulating 10,919 citations. Furthermore, analysis of 2,702 citing papers reveals that 93.5% originate from independent researchers, indicating broad adoption and impact beyond the researcher’s immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 759 · **40 flagged influential by Semantic Scholar**

CORE PAPER

Google's neural machine translation system: Bridging the gap between human and machine translation

2016 · 10,919 citations (GS)

Field-normalised: 7,223 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2016 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Alpacafarm: A simulation framework for methods that learn from human feedback	Cornell University, Stanford, Stanford University	Canada, United States	—
2	Unlocking the black box: an in-depth review on interpretability, explainability, and reliability in deep learning	Dumlupinar University	Turkey	—
3	Text as data	Stanford University, University of Chicago, Yale University	United States	—
4	Vision-language models for medical report generation and visual question answering: A review	H. Lee Moffitt Cancer Center and Research Institute	United States	—
5	Five sources of bias in natural language processing	Bocconi University, Carnegie Mellon University	Italy, United States	—
6	A review on large language models: Architectures, applications, taxonomies, open issues and challenges	Bangladesh University of Engineering and Technology, Charles Darwin University, Lappeenranta-Lahti University of Technology	Australia, Bangladesh, Finland	—
7	Combining machine learning and computational chemistry for predictive insights into chemical systems	Technische Universität Berlin, University of Cambridge, University of Luxembourg	Germany, Luxembourg, United Kingdom	—

No.	Citing paper	Citing institution(s)	Country	S2
8	Financial time series forecasting with deep learning: A systematic literature review: 2005–2019	TOBB University of Economics and Technology	Turkey	—
9	Deep reinforcement learning: An overview	—	—	—
10	A machine learning Automated Recommendation Tool for synthetic biology	—	—	—
11	Photonics for artificial intelligence and neuromorphic computing	Princeton University, Queen's University, University of Exeter	Canada, Germany, United Kingdom	—
12	Edge learning for B5G networks with distributed signal processing: Semantic communication, edge computing, and wireless sensing	Mohamed bin Zayed University of Artificial Intelligence, Southeast University, The University of New South Wales	Australia, China, Israel	—
13	Neural networks and deep learning	Brown University, University of Pennsylvania	United States	—
14	Review of deep learning algorithms and architectures	University of Bridgeport	United States	Influential
15	A survey on neural network interpretability	Southern University of Science and Technology, University of Birmingham	China, United Kingdom	—
16	Recent trends in deep learning based natural language processing	Nanyang Technological University, National University of Singapore, Singapore University of Technology and Design	Singapore	Influential
17	A state-of-the-art survey on deep learning theory and architectures	Comcast, Lawrence Livermore National Laboratory, Saint Louis University	United States	—
18	Exploring the frontiers of deep learning and natural language processing: A comprehensive overview of key challenges and emerging trends	Abdul Wali Khan University Mardan, International Islamic University, Islamabad, Tianjin University	China, Pakistan	—
19	Utilizing bert for information retrieval: Survey, applications, resources, and challenges	Central China Normal University, Hangzhou Dianzi University, Henan University of Technology	Canada, China	—
20	Stand-alone self-attention in vision models	Google Research	United States	—
21	FinBERT: A large language model for extracting information from financial text	HKUST, Hong Kong University of Science and Technology, Renmin University of China	China, Hong Kong	—
22	Opportunities and obstacles for deep learning in biology and medicine	Broad Institute, Brown University, Carnegie Mellon University	Canada, Germany, United Kingdom	—
23	A comprehensive survey of large language models and multimodal large language models in medicine	Chongqing University of Technology	China	—

No.	Citing paper	Citing institution(s)	Country	S2
24	Model compression and hardware acceleration for neural networks: A comprehensive survey	Massachusetts Institute of Technology, Tsinghua University, University of California, Irvine Medical Center	China, United States	—
25	Deep learning: new computational modelling techniques for genomics	Helmholtz Munich, Helmholtz Zentrum München	Germany	—
26	Towards better understanding of gradient-based attribution methods for deep neural networks	ETH Zurich, ETH Zürich, University of Zurich and ETH Zurich	Switzerland	—
27	Neural sign language translation	RWTH Aachen University, University of Surrey	Germany, United Kingdom	—
28	Deep neural networks for the evaluation and design of photonic devices	Stanford University	United States	—
29	Recurrent neural networks: A comprehensive review of architectures, variants, and applications	University of California, Irvine Medical Center, University of Johannesburg	South Africa, United States	—
30	A survey on kolmogorov-arnold network	Texas State University	United States	—

Showing the 30 most-cited of 759 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

Scaling autoregressive models for content-rich text-to-image generation

2022 · 1,779 citations (GS)

Field-normalised: 1,454 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2022 indexed by Semantic Scholar, by citation count.

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

FOLLOW-UP WORK

Gpipe: Efficient training of giant neural networks using pipeline parallelism

2019 · 2,707 citations (GS)

Field-normalised: 2,054 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 2

Claim — Contribution 2

The researcher pioneered the Gemini family of multimodal models, establishing a foundational framework for high-capability AI that evolved to support massive context windows and advanced agentic reasoning.

The researcher's primary contribution is the development of the Gemini family of highly capable multimodal models, anchored by the seminal 2023 paper. This work established a core architectural approach that has since been iteratively refined in subsequent publications, demonstrating a sustained and coherent research trajectory in multimodal artificial intelligence.

This line of work appears to address the challenge of integrating diverse data modalities within a unified model structure. The progression from the initial 2023 release to the 2024 and 2025 follow-ups suggests a deliberate expansion of capabilities, specifically targeting the limitations of context length and reasoning complexity. The titles indicate a shift from general multimodal understanding to handling millions of tokens and enabling next-generation agentic behaviors, marking a significant evolution in the field's technical frontier.

The significance of this contribution is evidenced by substantial citation metrics, with the core paper accumulating over 9,000 citations and subsequent versions garnering thousands more. Crucially, analysis of citing literature reveals that approximately 93.5% of citations originate from independent researchers, indicating that this work has been widely adopted and validated by the broader scientific community rather than relying on self-citation or institutional bias.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 959 · **29 flagged influential by Semantic Scholar**

CORE PAPER

Gemini: a family of highly capable multimodal models

2023 · 9,105 citations (GS)

Field-normalised: 3,480 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2023 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Mm-llms: Recent advances in multimodal large language models	Kyoto University, Mohamed bin Zayed University of Artificial Intelligence, Tencent	China, Japan, United Arab Emirates	—
2	Visual autoregressive modeling: Scalable image generation via next-scale prediction	ByteDance, ByteDance (China), ByteDance Inc	China, United States	—
3	Visual cot: Advancing multi-modal language models with a comprehensive dataset and benchmark for chain-of-thought reasoning	SenseTime, Sensetime (China), SenseTime Research	Canada, China, Hong Kong	—
4	Empowering biomedical discovery with AI agents	Broad Institute, Harvard Medical School, Harvard University	United Kingdom, United States	—
5	A comprehensive survey of large AI models for future communications: Foundations, applications and challenges	Brunel University of London, Hunan Normal University, Hunan University of Technology and Business	China, Singapore, United Arab Emirates	—
6	An agentic system for rare disease diagnosis with traceable reasoning	Harvard Medical School, Shanghai Artificial Intelligence Laboratory, Shanghai Jiao Tong University	China, United States	—
7	Sapiens: Foundation for human vision models	Meta	United States	—
8	Continual learning of large language models: A comprehensive survey	Google, Rutgers University	United States	—
9	Foundational challenges in assuring alignment and safety of large language models	—	—	—
10	Large language models: A survey	Cologne University of Applied Sciences, Czech Technical University in Prague, Expedia Group (United States)	Czech Republic, Germany, Singapore	—

No.	Citing paper	Citing institution(s)	Country	S2
11	Unleashing the potential of prompt engineering for large language models	Beijing Normal University, BNU-HKBU United International College	China	—
12	A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions	Harbin Institute of Technology, Huawei Inc., Zhejiang University	China	—
13	Parameter-efficient fine-tuning for large models: A comprehensive survey	Harvard University, New York University, Northeastern University	United States	—
14	A comprehensive overview of large language models	Australian National University, Commonwealth Scientific and Industrial Research Organisation, King Fahd University of Petroleum and Minerals	Australia, China, Pakistan	—
15	Safety alignment should be made more than just a few tokens deep	Google DeepMind, Princeton University, Tsinghua University	China, United Kingdom, United States	—
16	Model merging in llms, mllms, and beyond: Methods, theories, applications, and opportunities	Nanyang Technological University, Northeastern University, Sun Yat-sen University	China, Singapore, United States	—
17	: A Vision-Language-Action Flow Model for General Robot Control	—	—	—
18	Understanding transformer reasoning capabilities via graph algorithms	Columbia University, Google, Google DeepMind (United Kingdom)	Canada, United Kingdom, United States	—
19	Large language models meet text-centric multimodal sentiment analysis: A survey	Chinese Academy of Sciences, Harbin Institute of Technology	China	—
20	Aligning vision to language: Annotation-free multimodal knowledge graph construction for enhanced llms reasoning	East China Normal University, New York University, Shanghai Artificial Intelligence Laboratory	China, United States	—
21	Flipattack: Jailbreak llms via flipping	National University of Singapore	Singapore	—
22	From llm reasoning to autonomous ai agents: A comprehensive review	Khalifa University, Technology Innovation Institute, United Arab Emirates University	United Arab Emirates	—
23	A survey on benchmarks of multimodal large language models	Peking University, Tencent	China	—
24	Spatialbot: Precise spatial understanding with vision language models	Peking University, Shanghai Jiao Tong University, Southeast University	China, United Kingdom	—
25	When semantics mislead vision: Mitigating large multimodal models hallucinations in scene text spotting and understanding	Hong Kong University of Science and Technology, Nanjing University of Science and Technology, Nankai University	China, Hong Kong, Italy	—
26	A survey on federated fine-tuning of large language models	The Hong Kong University of Science and Technology	China	—

No.	Citing paper	Citing institution(s)	Country	S2
		(Guangzhou), University of Macau		
27	Sensorlm: Learning the language of wearable sensors	Google DeepMind, Google Research, University of Cambridge	United Kingdom, United States	—
28	Core: Benchmarking LLMs' code reasoning capabilities through static analysis tasks	Purdue University	United States	—
29	Mobile edge intelligence for large language models: A contemporary survey	University of Hong Kong	China	—
30	From system 1 to system 2: A survey of reasoning large language models	Alibaba Group, Chinese Academy of Sciences, East China Normal University	China, United Arab Emirates, United Kingdom	—

Showing the 30 most-cited of 890 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

Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context

2024 · 4,332 citations (GS)

Field-normalised: 3,480 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2024 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	An agentic system for rare disease diagnosis with traceable reasoning	Harvard Medical School, Shanghai Artificial Intelligence Laboratory, Shanghai Jiao Tong University	China, United States	—
2	Foundational challenges in assuring alignment and safety of large language models	—	—	—
3	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	—
4	A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions	Harbin Institute of Technology, Huawei Inc., Zhejiang University	China	—
5	A comprehensive overview of large language models	Australian National University, Commonwealth Scientific and Industrial Research Organisation, King Fahd University of Petroleum and Minerals	Australia, China, Pakistan	—
6	From system 1 to system 2: A survey of reasoning large language models	Alibaba Group, Chinese Academy of Sciences, East China Normal University	China, United Arab Emirates, United Kingdom	—

No.	Citing paper	Citing institution(s)	Country	S2
7	From Google Gemini to OpenAI Q*(Q-Star): a survey on reshaping the generative artificial intelligence (AI) research landscape	Academies Australasia Polytechnic, Cyberonomy Pty Ltd, Massey University	Australia, New Zealand	—
8	Towards reasoning era: A survey of long chain-of-thought for reasoning large language models	Central South University, Fudan University, Harbin Institute of Technology	China	—
9	Long context tuning for video generation	ByteDance, Sea AI Lab, The Chinese University of Hong Kong	China	—
10	Hourvideo: 1-hour video-language understanding	Stanford University	United States	**Influential**
11	Cambrian-s: Towards spatial supersensing in video	New York University, Stanford University	United States	—
12	Textgrad: Automatic" differentiation" via text	Chan Zuckerberg Biohub San Francisco, Huazhong University of Science and Technology, Stanford University	China, United States	—
13	Expanding performance boundaries of open-source multimodal models with model, data, and test-time scaling	Nanjing University	China	—
14	What matters when building vision-language models?	Hugging Face, Sorbonne Université	France, United States	—
15	Internvl3: Exploring advanced training and test-time recipes for open-source multimodal models	Nanjing University, Shanghai AI Laboratory	China	—
16	Thinking with videos: Multimodal tool-augmented reinforcement learning for long video reasoning	ByteDance, Tsinghua University	China	—
17	Humanity's last exam	Carnegie Mellon University, Scale AI	United States	—
18	Video-mme: The first-ever comprehensive evaluation benchmark of multi-modal llms in video analysis	Carnegie Mellon University, Institute of Automation, Nanjing University	China, United States	**Influential**
19	Thinking in space: How multimodal large language models see, remember, and recall spaces	New York University, Stanford University, Yale University	United States	**Influential**
20	Chain-of-retrieval augmented generation	Microsoft, Microsoft Research	United States	**Influential**
21	Visionllm v2: An end-to-end generalist multimodal large language model for hundreds of vision-language tasks	Beijing Institute of Technology, CUHK, Nanjing University	China, Hong Kong	—
22	Med-r1: Reinforcement learning for generalizable medical reasoning in vision-language models	Emory University, Johns Hopkins University, The Wallace H. Coulter Department of Biomedical Engineering	United States	—
23	Video-xl: Extra-long vision language model for hour-scale video understanding	Beijing Academy of Artificial Intelligence, Chinese Acad-	China, Italy	—

No.	Citing paper	Citing institution(s)	Country	S2
		emy of Sciences, Huawei Technologies		
24	Tokenpacker: Efficient visual projector for multimodal llm	Ant Group, Nanjing University of Aeronautics and Astronautics, The Hong Kong Polytechnical University	China, Hong Kong	—
25	Mllms know where to look: Training-free perception of small visual details with multimodal llms	University of Southern California, Vrije Universiteit Amsterdam	Netherlands, United States	—
26	Vlm-3r: Vision-language models augmented with instruction-aligned 3d reconstruction	Meta, Texas A&M University, The University of Texas at Austin	China, United States	**Influential**
27	Kimi-vl technical report	—	—	—
28	Emma: End-to-end multimodal model for autonomous driving	Stanford University, University of California, Irvine Medical Center, Waymo LLC	United States	—
29	Chain of agents: Large language models collaborating on long-context tasks	Google, Penn State University, Soochow University	China, United States	—
30	Benchmarking and defending against indirect prompt injection attacks on large language models	Hong Kong University of Science and Technology, Microsoft, Microsoft Corporation	China, Hong Kong, United States	—

Showing the 30 most-cited of 69 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

Gemini 2.5: Pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities

2025 · 3,113 citations (GS)

Field-normalised: 2,686 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2025 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 a novel minimum spanning tree algorithm for constructing efficient and accurate genetic linkage maps, establishing a foundational computational method in genetics.

CLAIM: The researcher's primary contribution is the development of a method for the efficient and accurate construction of genetic linkage maps using the minimum spanning tree of a graph, as detailed in their 2008 paper. This work stands as a seminal core contribution in the field.

ORIGINALITY: The title suggests the researcher addressed the computational challenge of building genetic linkage maps by applying graph theory, specifically minimum spanning trees. This approach appears to offer a distinct algorithmic solution for improving both the efficiency and accuracy of map construction compared to prior methods.

SIGNIFICANCE: The core paper has accumulated 656 citations, indicating substantial uptake by the scientific community. Furthermore, citation analysis reveals that 93.5% of citing papers originate from independent researchers, demonstrating that this method has been widely adopted and utilized by the broader genetics and bioinformatics community beyond the researcher's immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 0

CORE PAPER

Efficient and accurate construction of genetic linkage maps from the minimum spanning tree of a graph

2008 · 656 citations (GS)

Field-normalised: 535 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2008 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
Google	United States	—	123
Carnegie Mellon University	United States	SCImago #266 · THE 24 · QS 52	118
Tsinghua University	China	SCImago #8 · THE 12 · QS =17	118
Google Research	United States	—	109
University of California, Irvine Medical Center	United States	—	106
Microsoft	United States	—	103
Shanghai Jiao Tong University	China	SCImago #10 · THE 40 · QS =47	101
Stanford University	United States	SCImago #18 · THE =5 · QS 3	94
Google DeepMind	United States	SCImago #90	86
The Chinese University of Hong Kong	Hong Kong	SCImago #163 · THE =41 · QS =32	83
Peking University	China	SCImago #11 · THE 13 · QS 14	82
Nanyang Technological University	Singapore	SCImago #137	79
NVIDIA	United States	—	74
Microsoft Research	United States	—	68
National University of Singapore	Singapore	SCImago #59 · THE 17 · QS 8	67

Geographic distribution of citing authors

Country	Citing papers
United States	1,292
China	1,047
United Kingdom	284

Country	Citing papers
Singapore	162
Hong Kong	137
Canada	127
Germany	114
South Korea	104
Japan	92
Australia	88
France	65
Switzerland	64

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	Google's neural machine translation system: Bridging the gap between human and machine translation	759	8 CFR 204.5(h)(3)(v) — Criterion 5
Contribution 2	Gemini: a family of highly capable multimodal models	959	8 CFR 204.5(h)(3)(v) — Criterion 5
Contribution 3	Efficient and accurate construction of genetic linkage maps from the minimum spanning tree of a graph	0	8 CFR 204.5(h)(3)(v) — Criterion 5