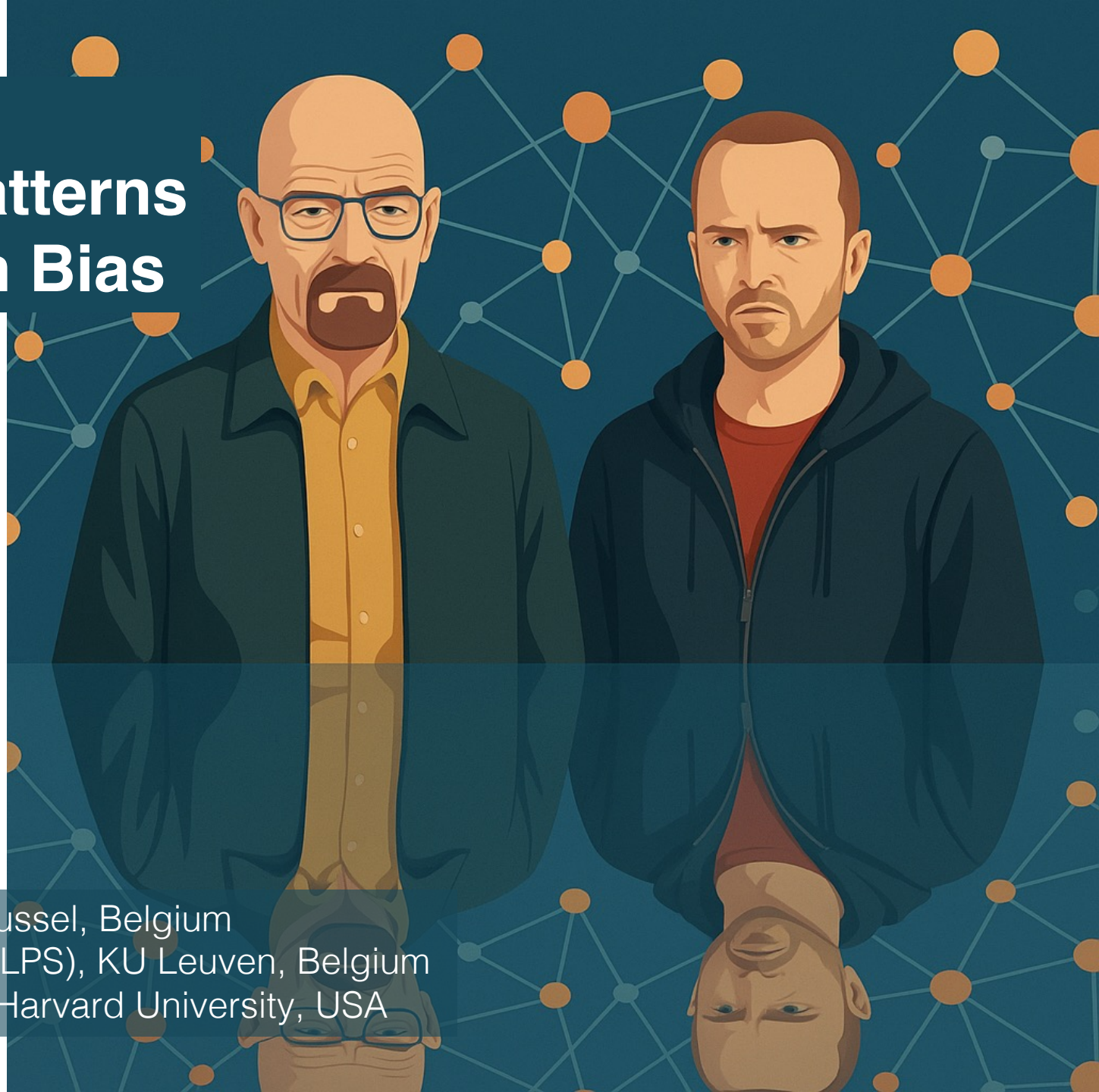
Large Language Models Reflect Human Citation Patterns with a Heightened Citation Bias

Andres Algaba[1], Carmen Mazijn[1],
Vincent Holst[1], Floriano Tori[1],
Sylvia Wenmackers[2] and Vincent Ginis[1,3]

Contact:
vincent.thorge.holst@vub.be

1 Data Analytics Laboratory, Vrije Universiteit Brussel, Belgium
2 Centre for Logic and Philosophy of Science (CLPS), KU Leuven, Belgium
3 School of Engineering and Applied Sciences, Harvard University, USA



Motivation

- Human citation behavior (~training data) has some well-documented biases (Letchford et al., 2015; Price, 1976; Wang, 2014; Wuchtly et al., 2007)

Price, D. D. S. (1976). A general theory of bibliometric and other cumulative advantage processes. Journal of the American society for Information science, 27(5), 292-306.

Wang, J. (2014). Unpacking the Matthew effect in citations. Journal of Informetrics, 8(2), 329-339.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. Science, 316(5827), 1036-1039.

Letchford, A., Moat, H. S., & Preis, T. (2015). The advantage of short paper titles. Royal Society open science, 2(8), 150266.

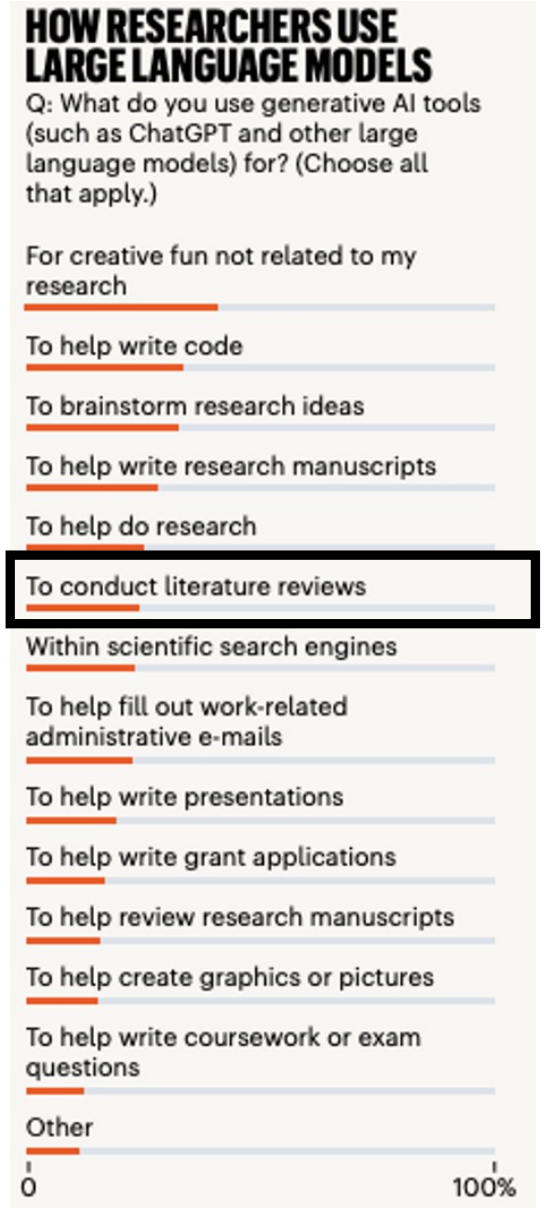
https://www.nature.com/articles/d41586-023-02980-0

Motivation

- Human citation behavior (~training data) has some well-documented biases (Letchford et al., 2015; Price, 1976; Wang, 2014; Wuchtly et al., 2007)

- Researchers are using LLMs in literature reviews

A Nature survey in September 2023



Price, D. D. S. (1976). A general theory of bibliometric and other cumulative advantage processes. Journal of the American society for Information science, 27(5), 292-306.

Wang, J. (2014). Unpacking the Matthew effect in citations. Journal of Informetrics, 8(2), 329-339.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. Science, 316(5827), 1036-1039.

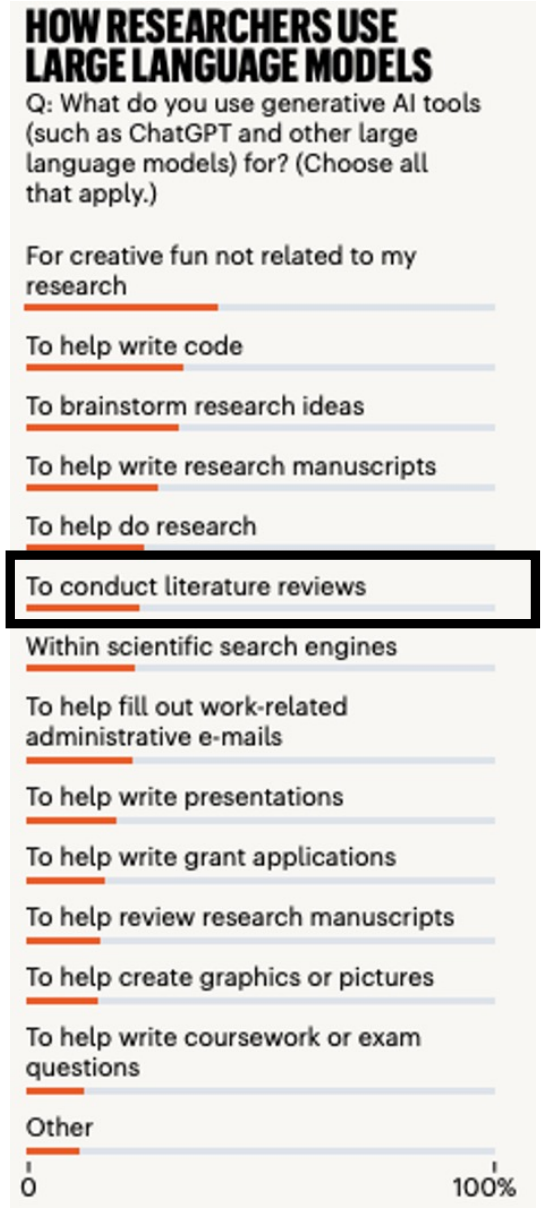
Letchford, A., Moat, H. S., & Preis, T. (2015). The advantage of short paper titles. Royal Society open science, 2(8), 150266.

https://www.nature.com/articles/d41586-023-02980-0

Motivation

- Human citation behavior (~training data) has some well-documented biases (Letchford et al., 2015; Price, 1976; Wang, 2014; Wuchtly et al., 2007)
- Researchers are using LLMs in literature reviews
- Our experimental setup focuses on a “controlled laboratory” setting (~parametric knowledge)

A Nature survey in September 2023



Price, D. D. S. (1976). A general theory of bibliometric and other cumulative advantage processes. Journal of the American society for Information science, 27(5), 292-306.

Wang, J. (2014). Unpacking the Matthew effect in citations. Journal of Informetrics, 8(2), 329-339.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. Science, 316(5827), 1036-1039.

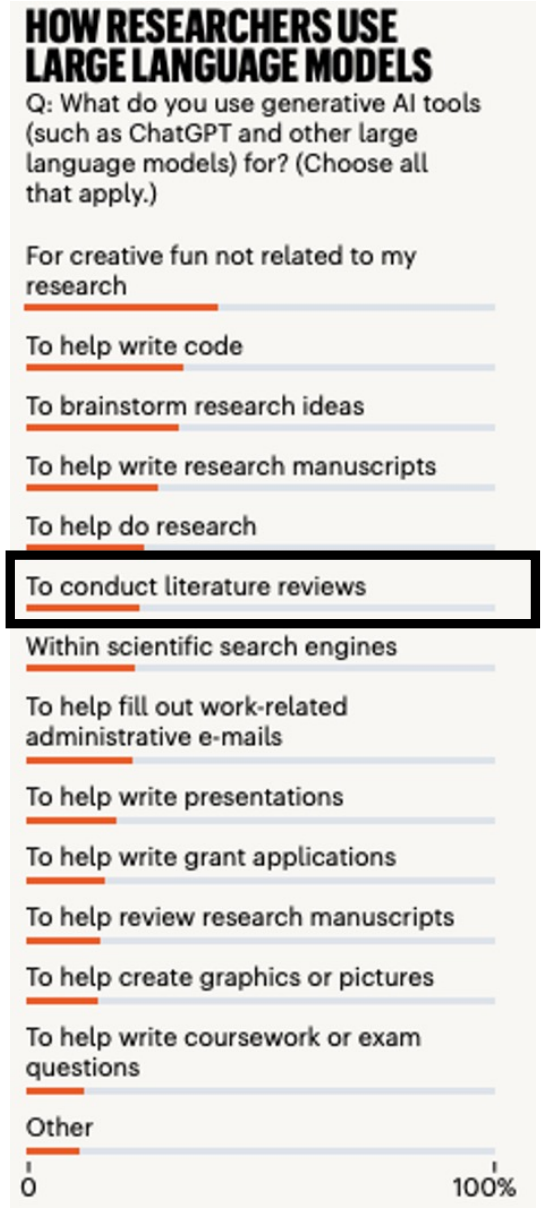
Letchford, A., Moat, H. S., & Preis, T. (2015). The advantage of short paper titles. Royal Society open science, 2(8), 150266.

https://www.nature.com/articles/d41586-023-02980-0

Motivation

- Human citation behavior (~training data) has some well-documented biases (Letchford et al., 2015; Price, 1976; Wang, 2014; Wuchtly et al., 2007)
- Researchers are using LLMs in literature reviews
- Our experimental setup focuses on a “controlled laboratory” setting (~parametric knowledge)
- A lot is still unknown about “real-world” usage

A Nature survey in September 2023



Price, D. D. S. (1976). A general theory of bibliometric and other cumulative advantage processes. Journal of the American society for Information science, 27(5), 292-306.

Wang, J. (2014). Unpacking the Matthew effect in citations. Journal of Informetrics, 8(2), 329-339.

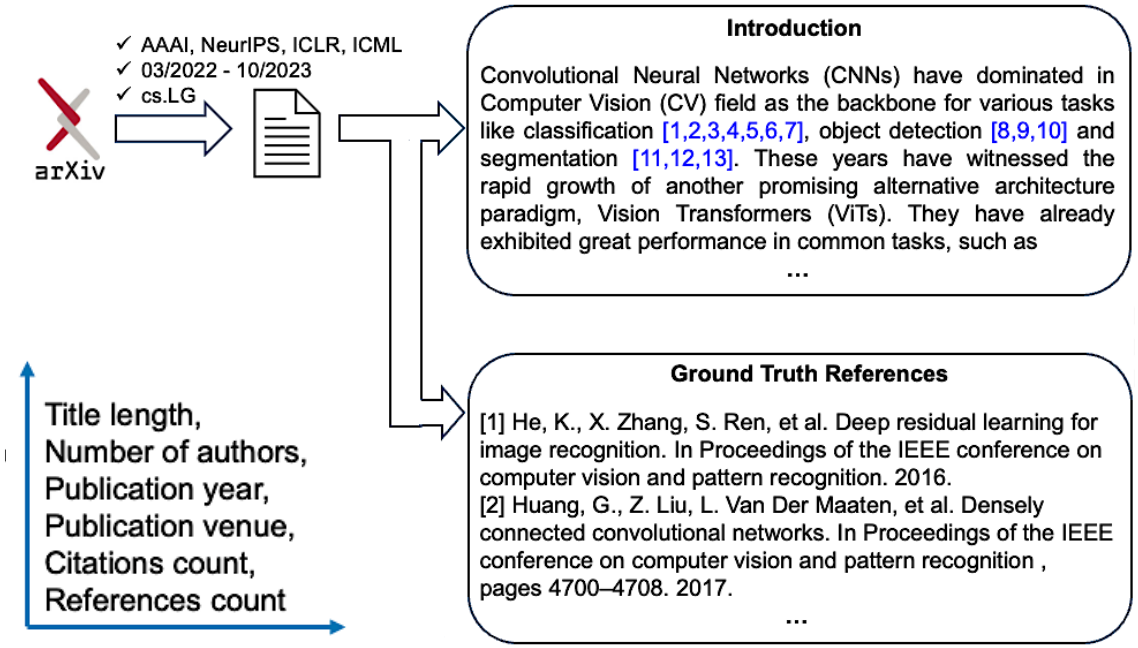
Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. Science, 316(5827), 1036-1039.

Letchford, A., Moat, H. S., & Preis, T. (2015). The advantage of short paper titles. Royal Society open science, 2(8), 150266.

https://www.nature.com/articles/d41586-023-02980-0

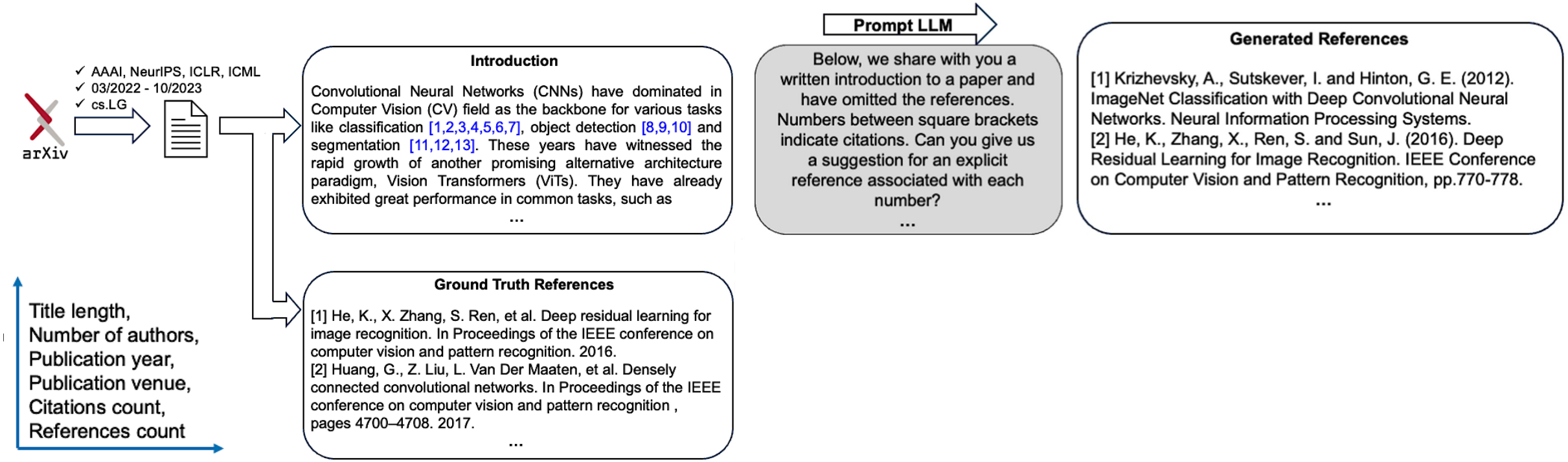
Prompt GPT-4, GPT-4o, and Claude 3.5 to Generate Reference Suggestions

For GPT-4: after knowledge cut-off



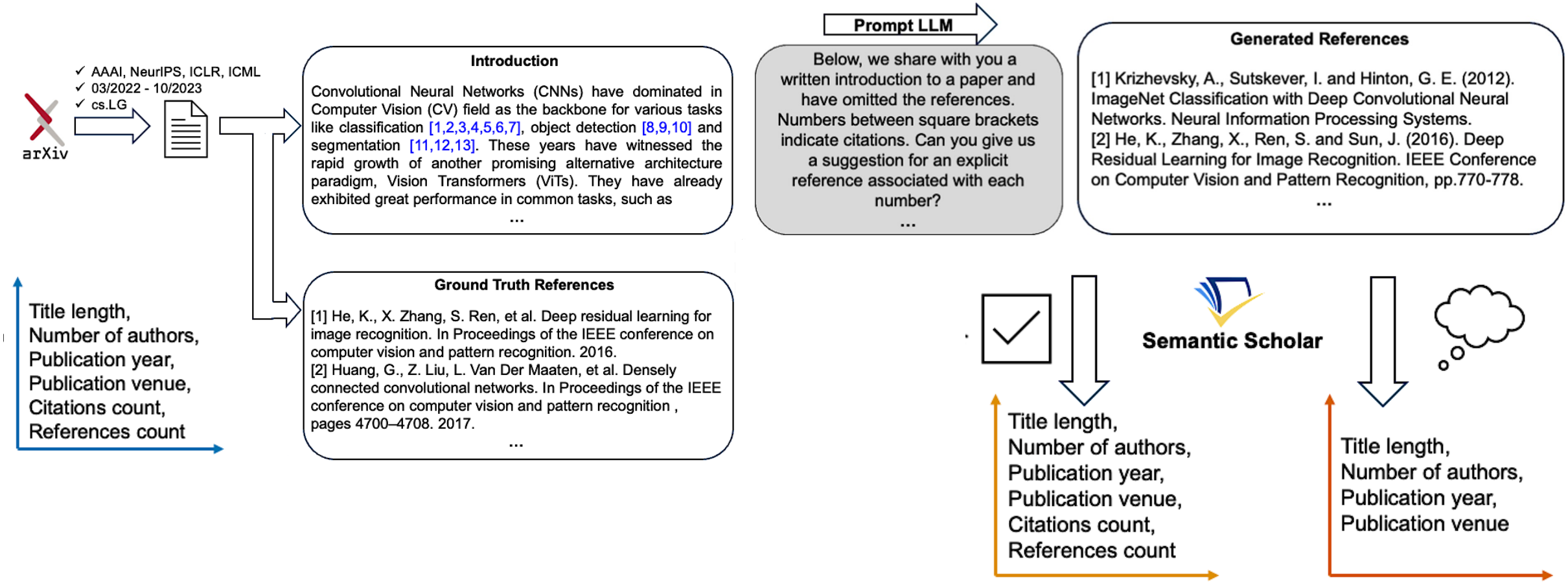
Prompt GPT-4, GPT-4o, and Claude 3.5 to Generate Reference Suggestions

For GPT-4: after knowledge cut-off



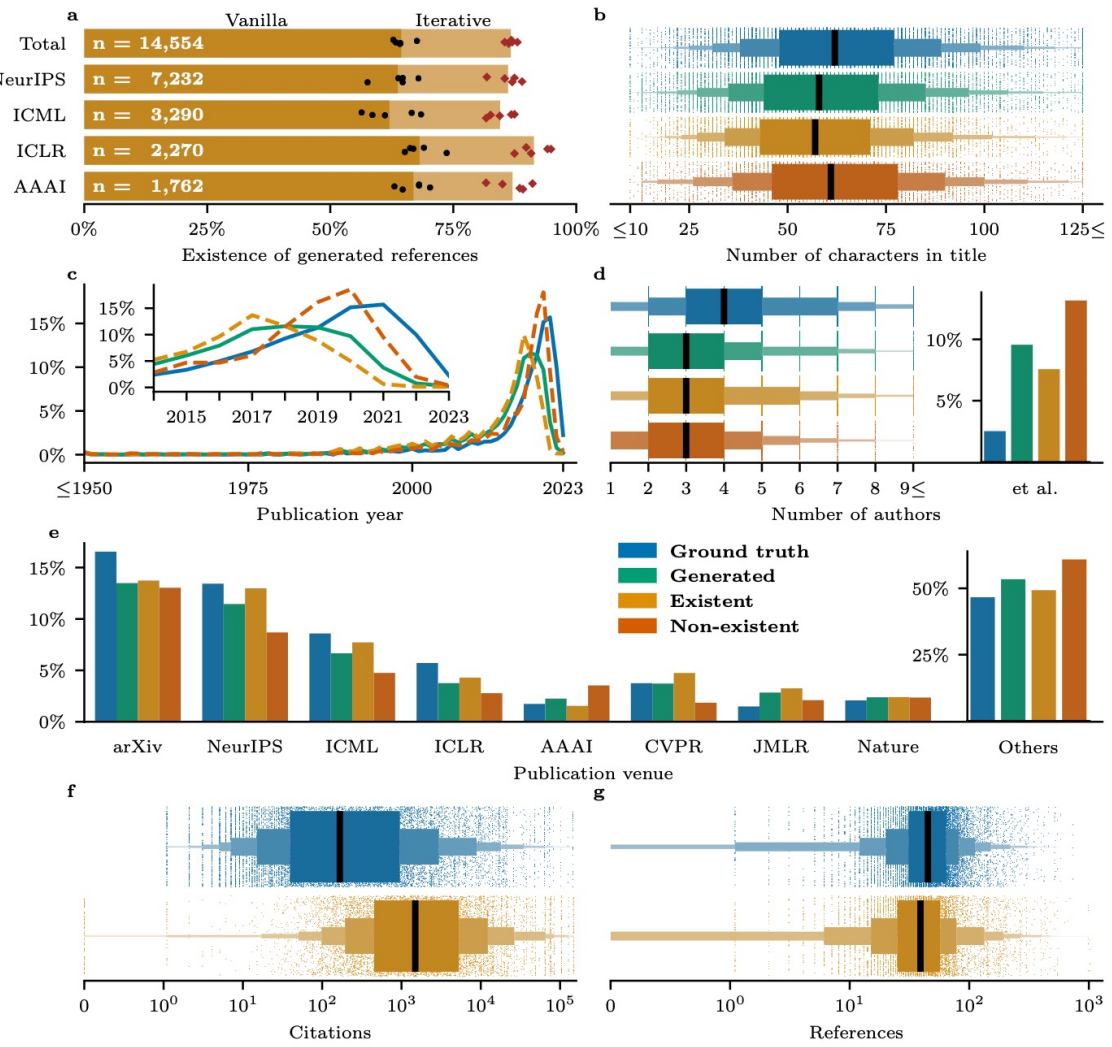
Prompt GPT-4, GPT-4o, and Claude 3.5 to Generate Reference Suggestions

For GPT-4: after knowledge cut-off



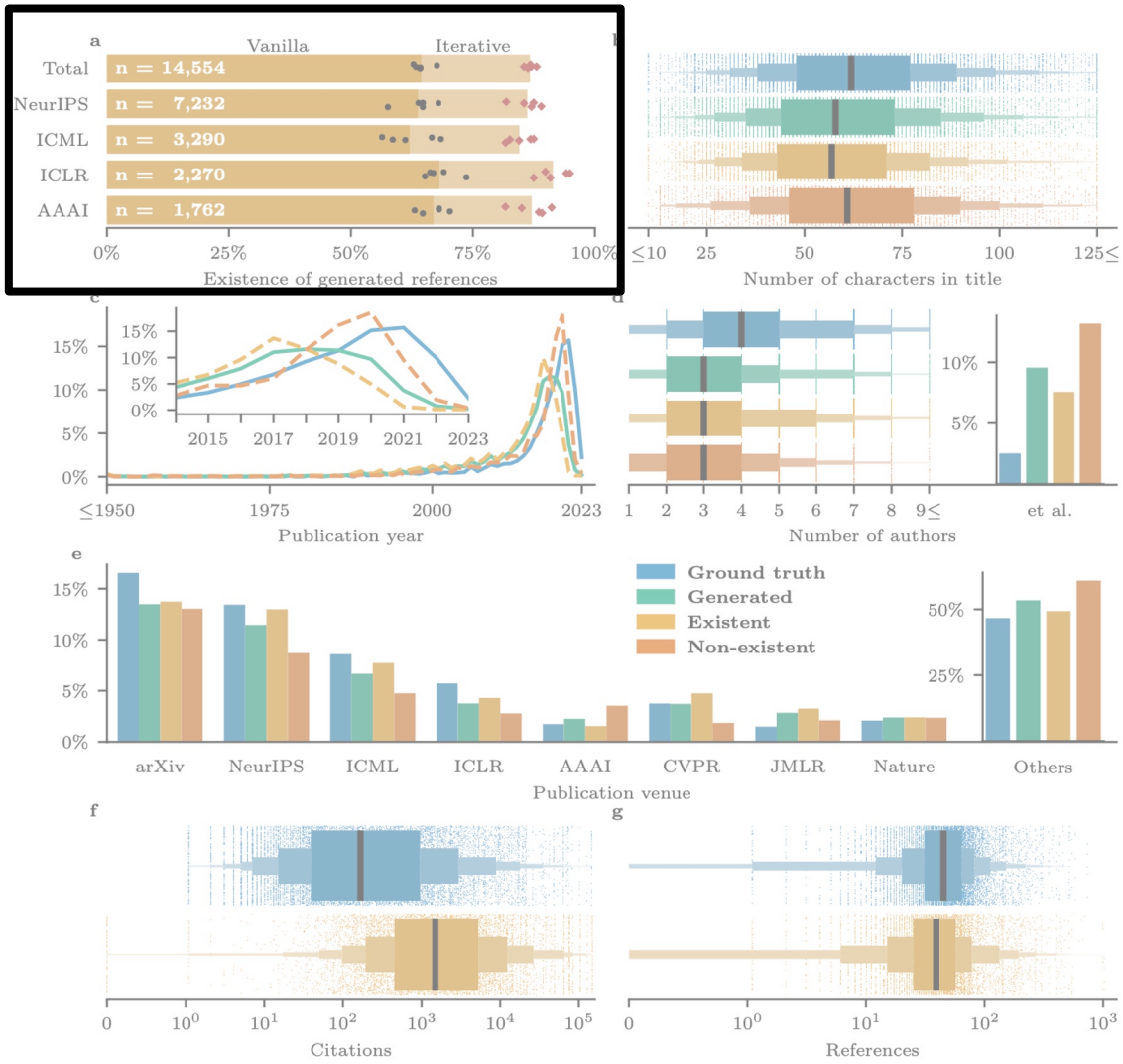
Generated Properties Reflect Human Citation Patterns Even When Hallucinating

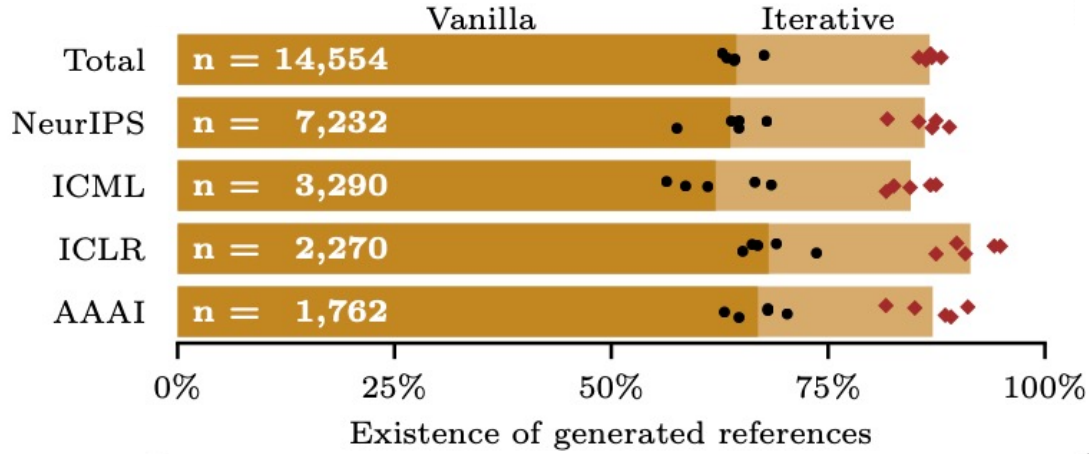
GPT-4 (also holds for GPT-4o and Claude 3.5)



Generated Properties Reflect Human Citation Patterns Even When Hallucinating

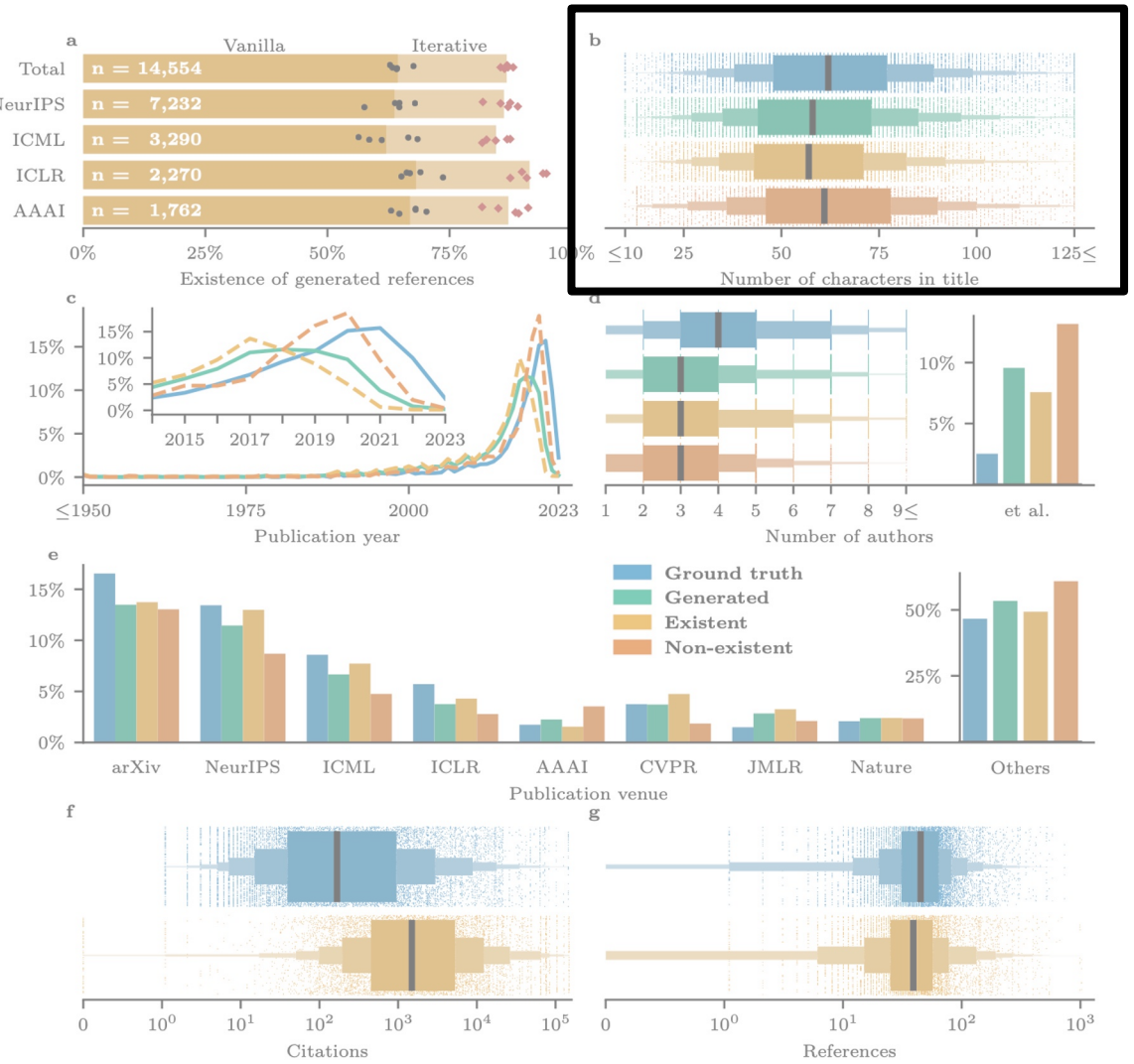
GPT-4 (also holds for GPT-4o and Claude 3.5)

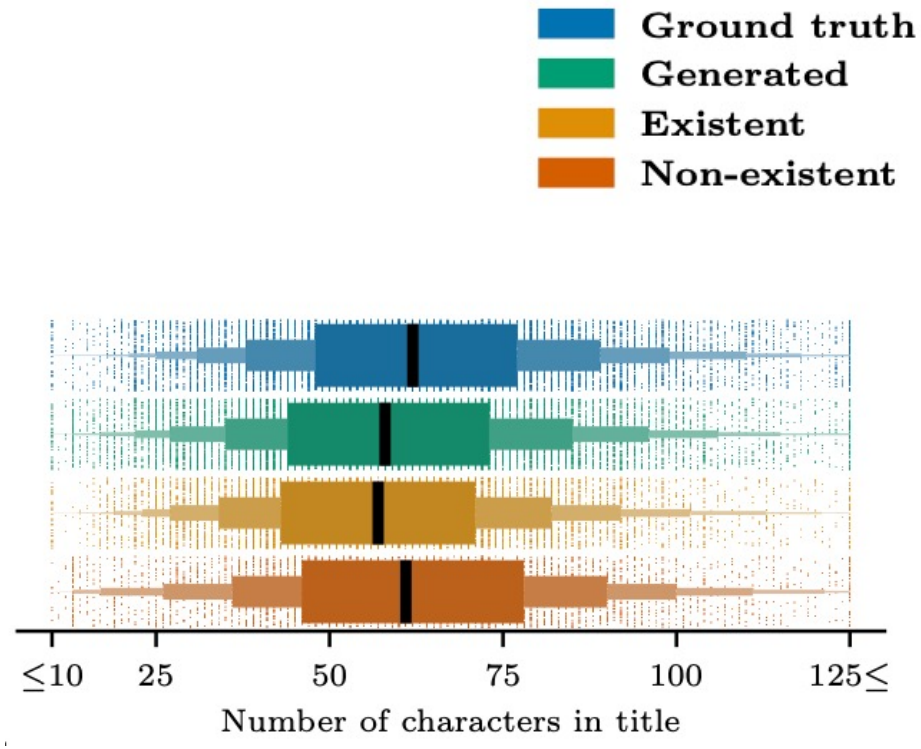




Generated Properties Reflect Human Citation Patterns Even When Hallucinating

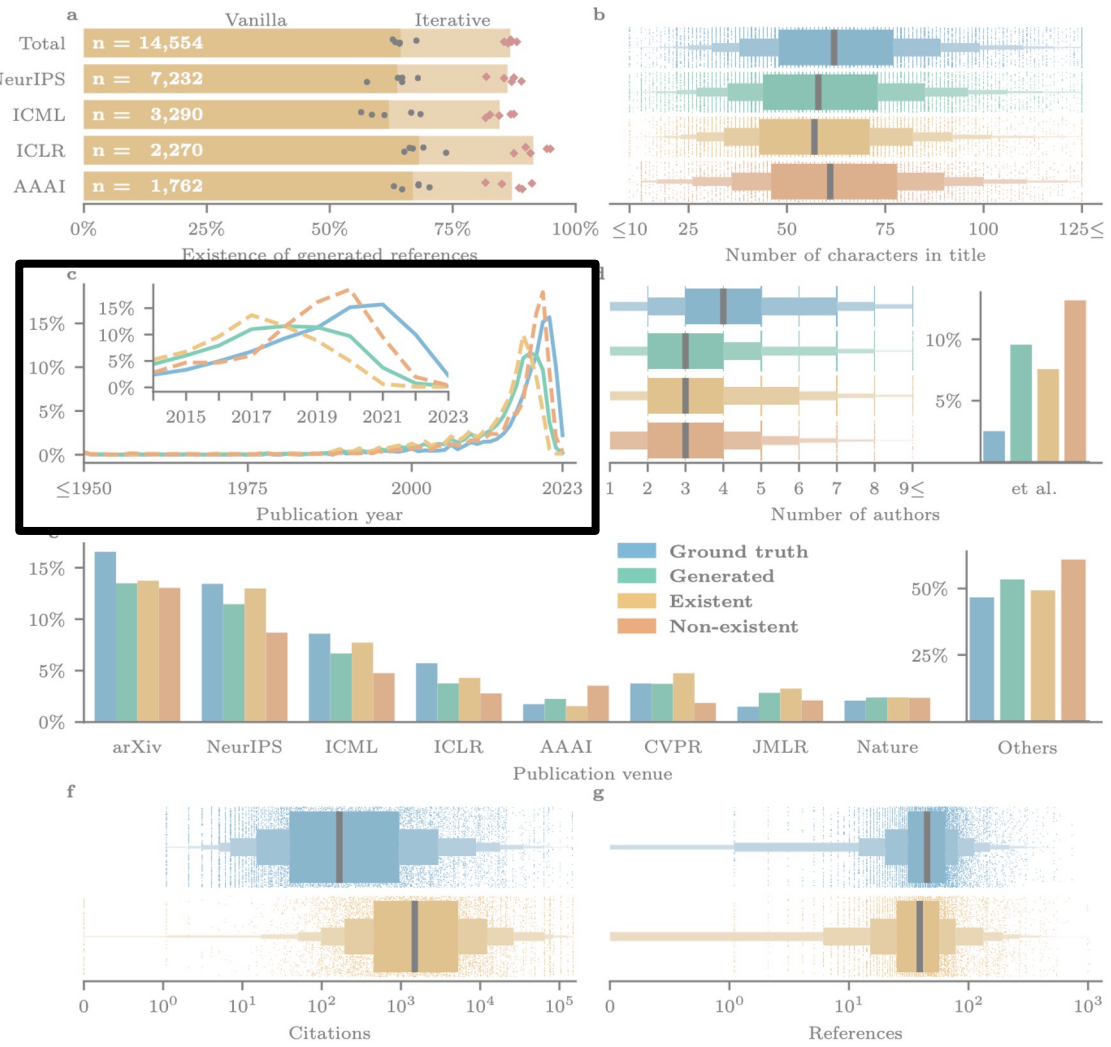
GPT-4 (also holds for GPT-4o and Claude 3.5)

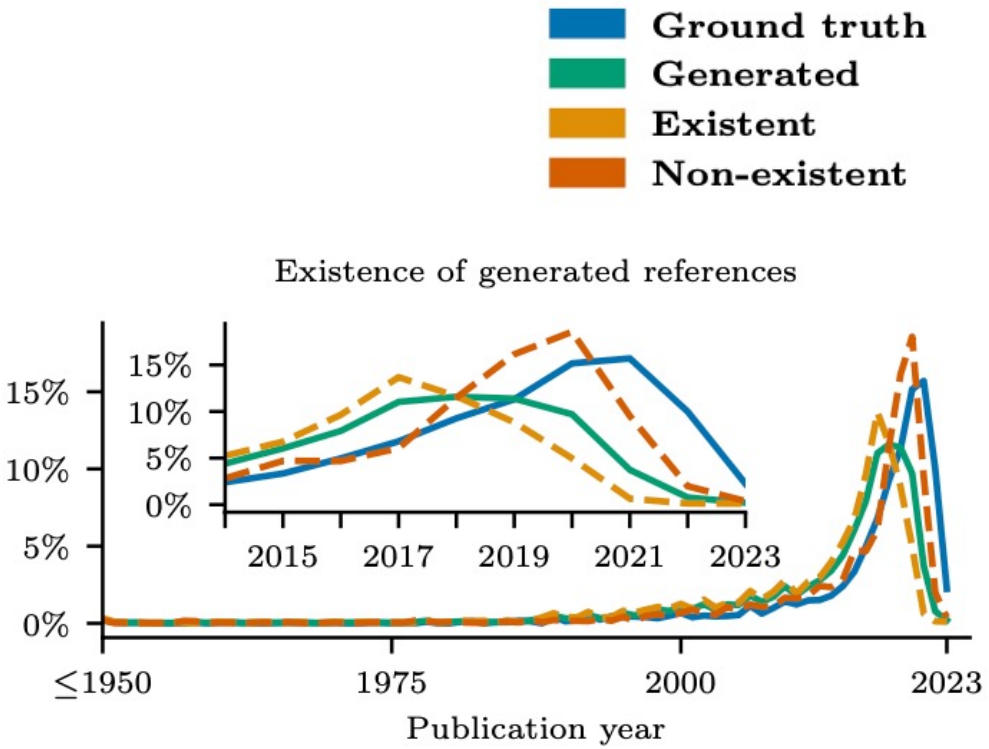




Generated Properties Reflect Human Citation Patterns Even When Hallucinating

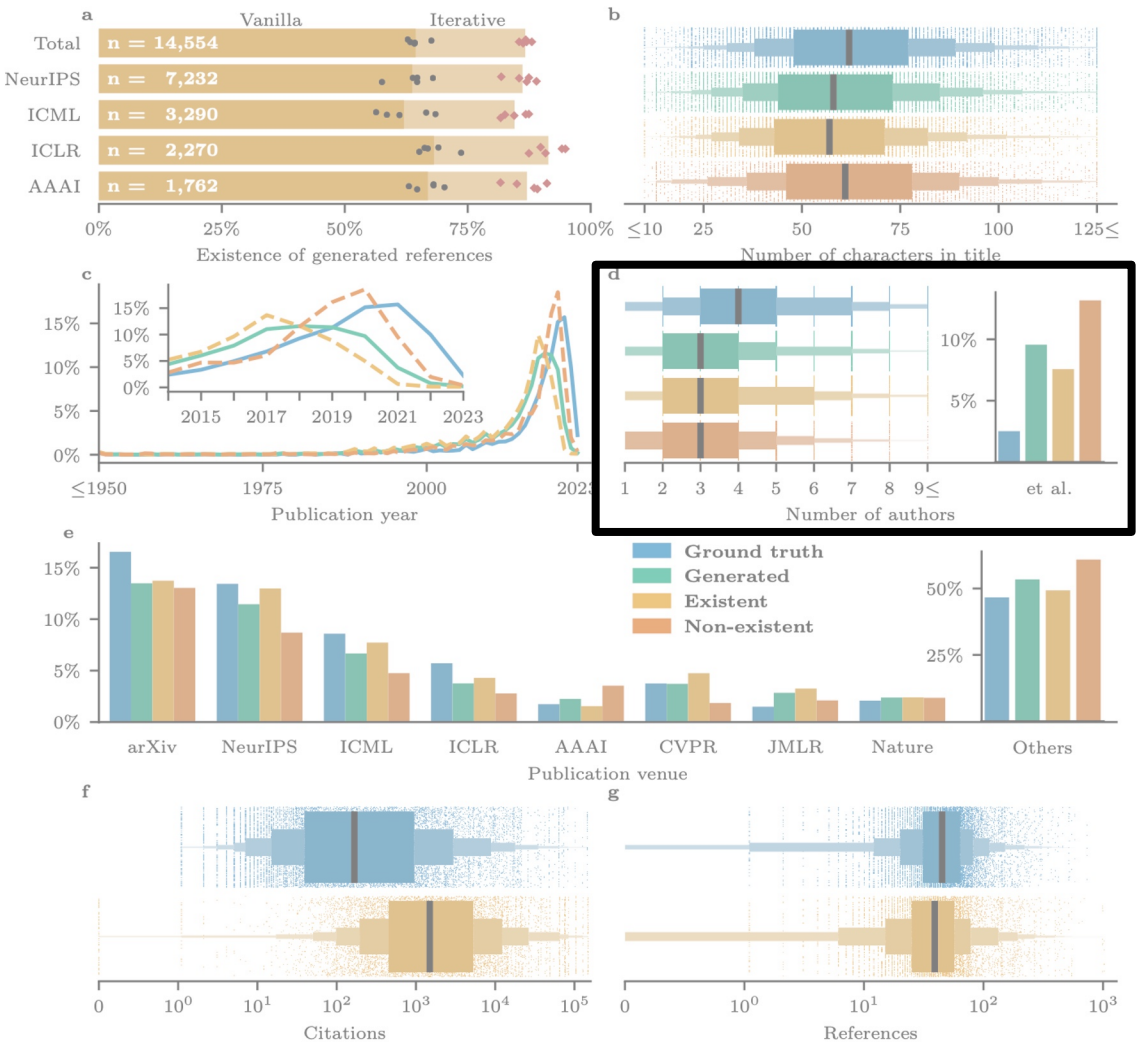
GPT-4 (also holds for GPT-4o and Claude 3.5)

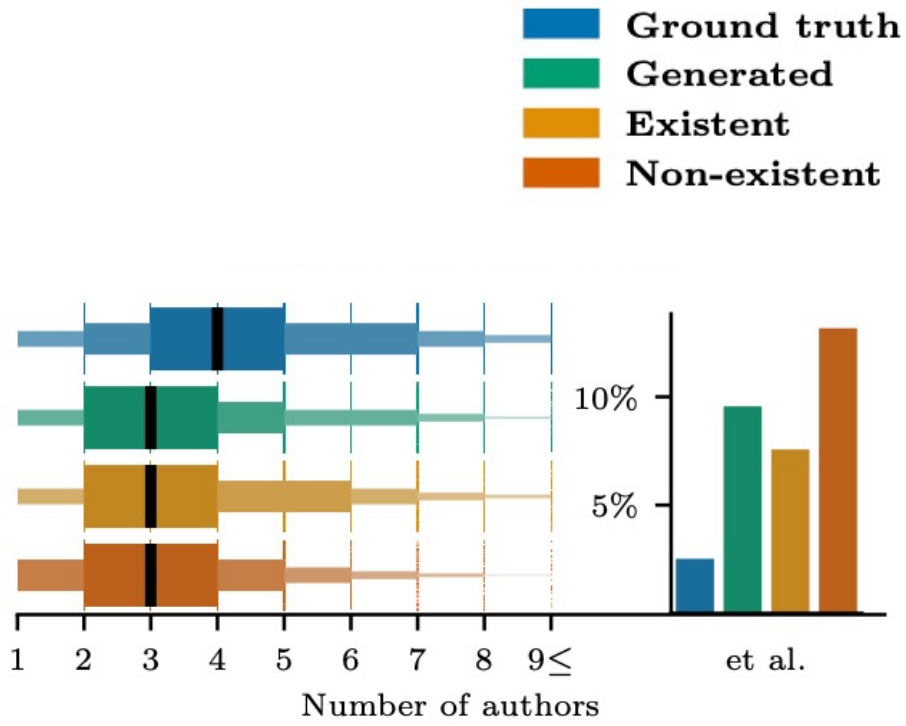




Generated Properties Reflect Human Citation Patterns Even When Hallucinating

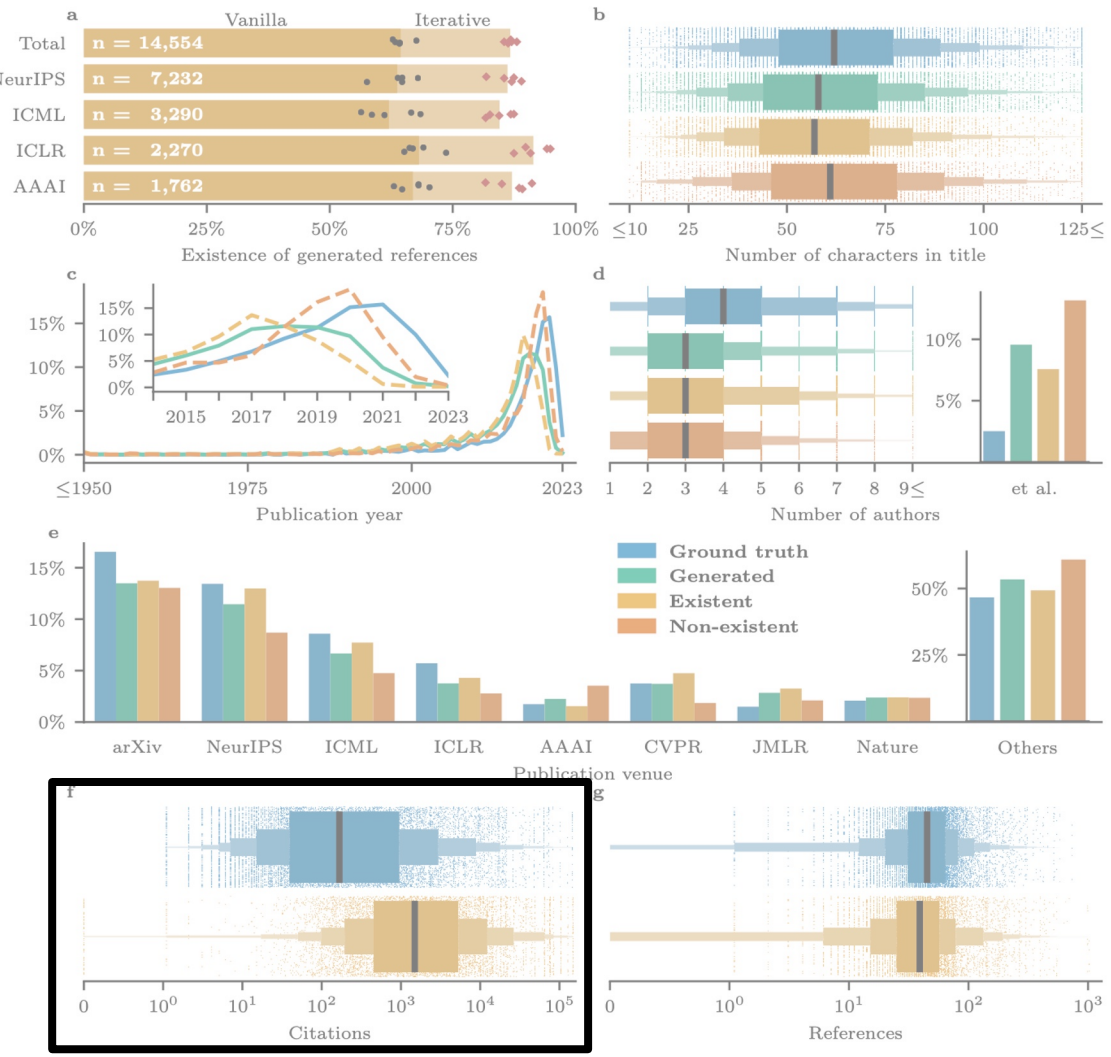
GPT-4 (also holds for GPT-4o and Claude 3.5)

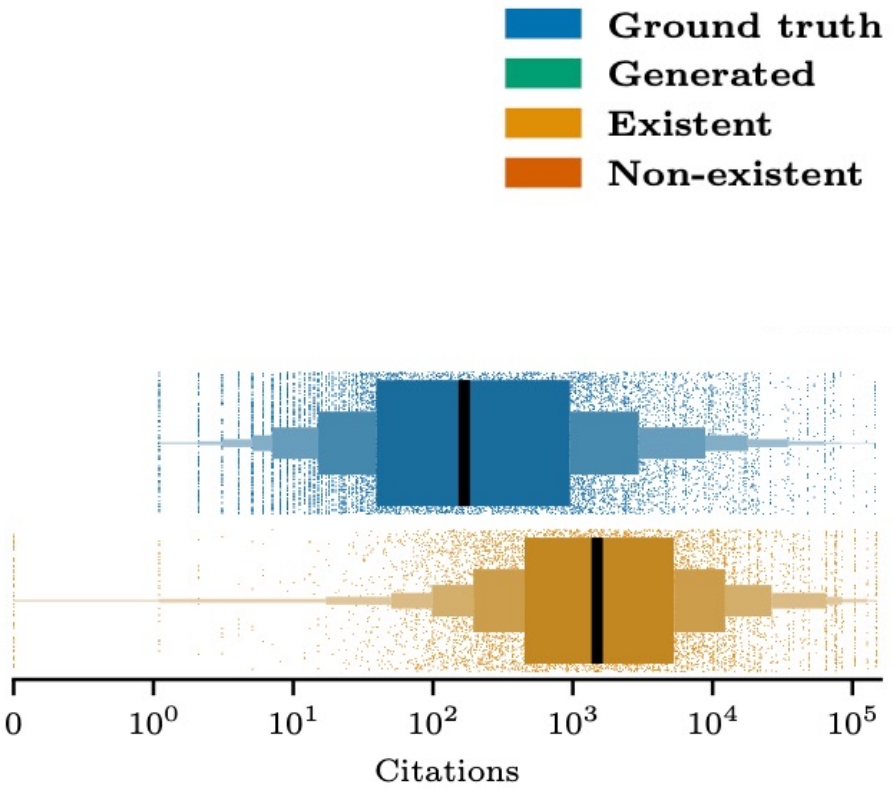




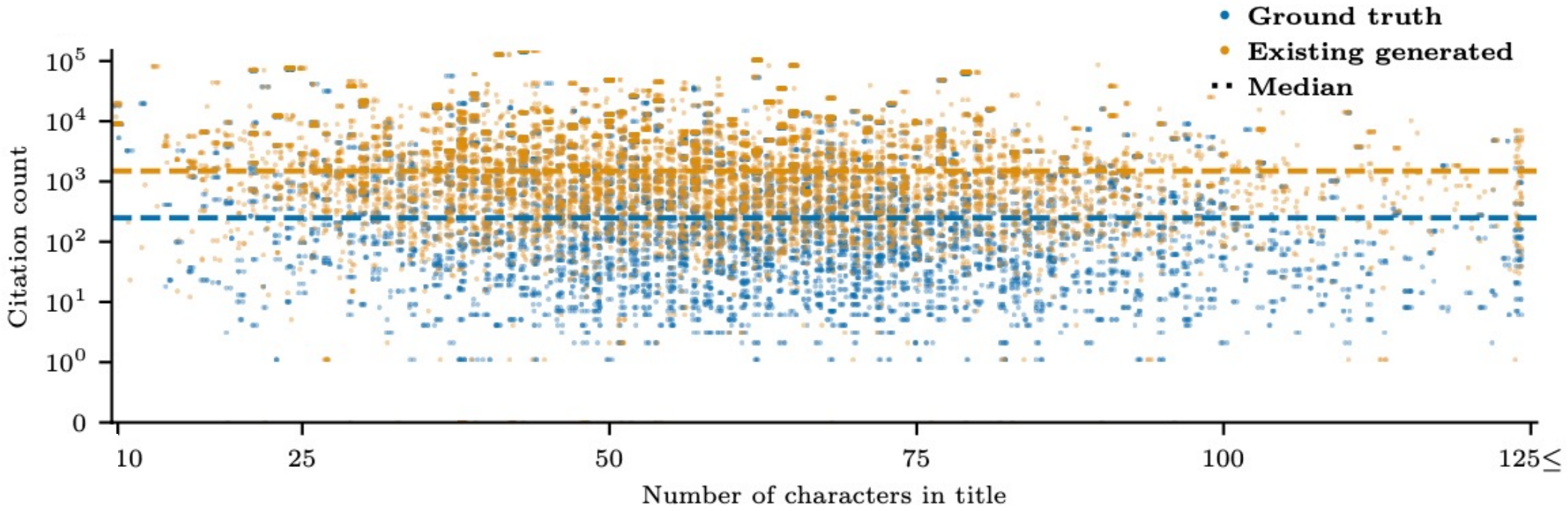
Generated Properties Reflect Human Citation Patterns with a Heightened Citation Bias

GPT-4 (also holds for GPT-4o and Claude 3.5)

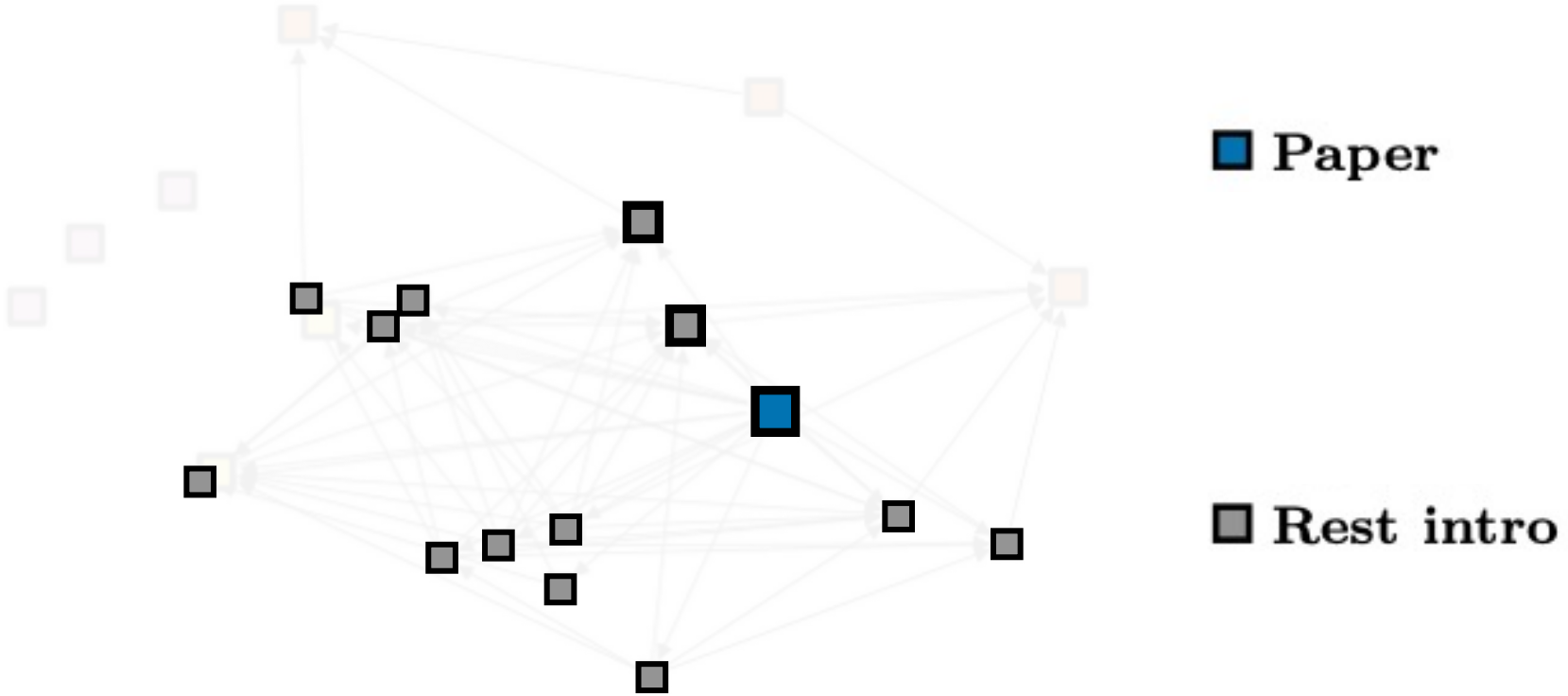




The Citation Gap Is Not Confounded by #Authors, Time, Title Length, and Venue

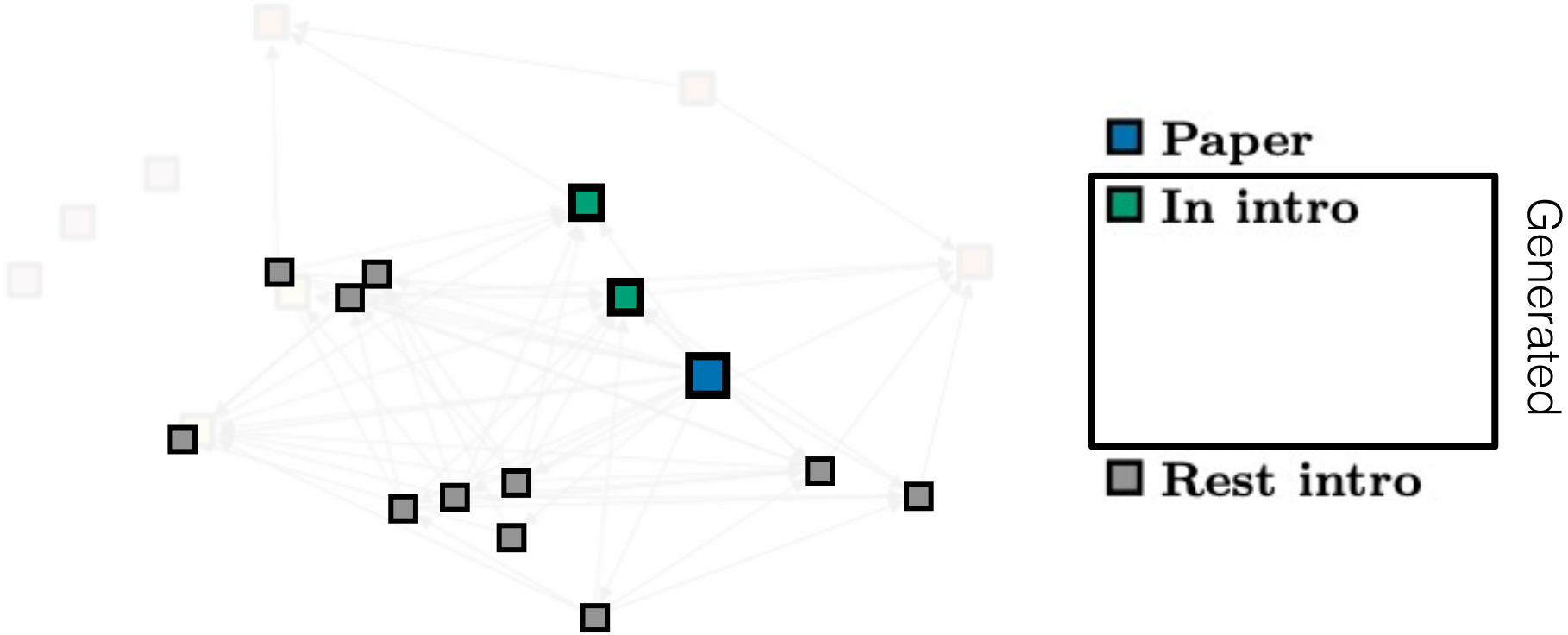


The Generated References Are Well Embedded In the Local Citation Context



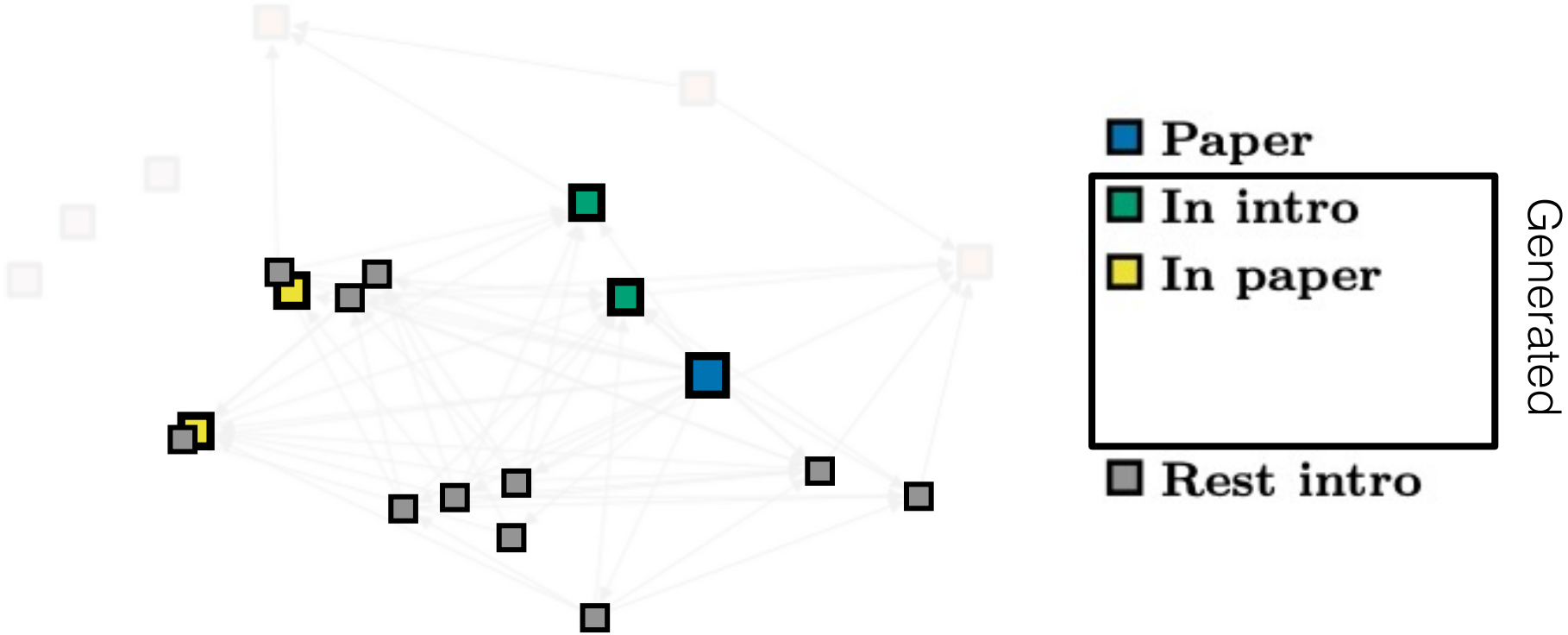
Local citation context of the paper and its introduction references

The Generated References Are Well Embedded In the Local Citation Context



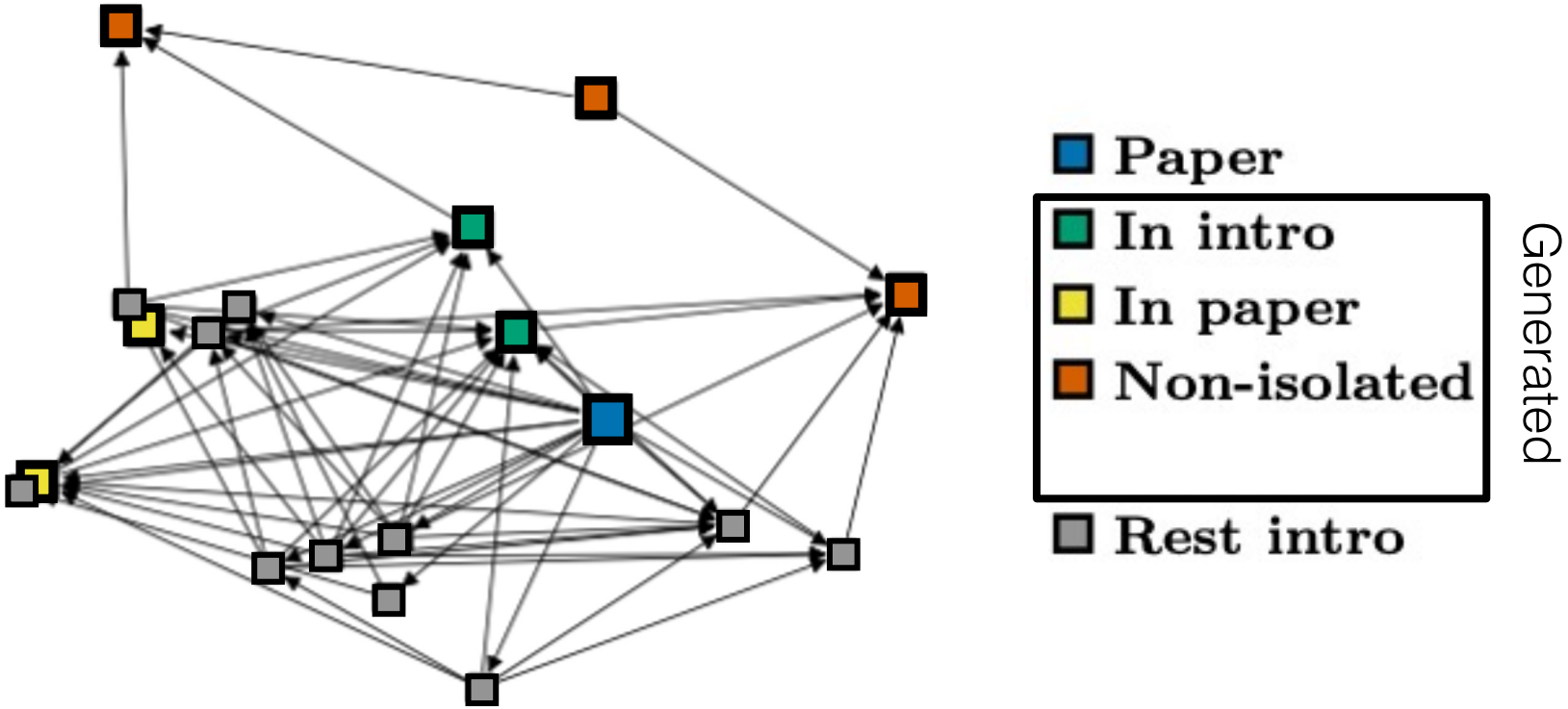
Local citation context of the paper and its introduction references

The Generated References Are Well Embedded In the Local Citation Context



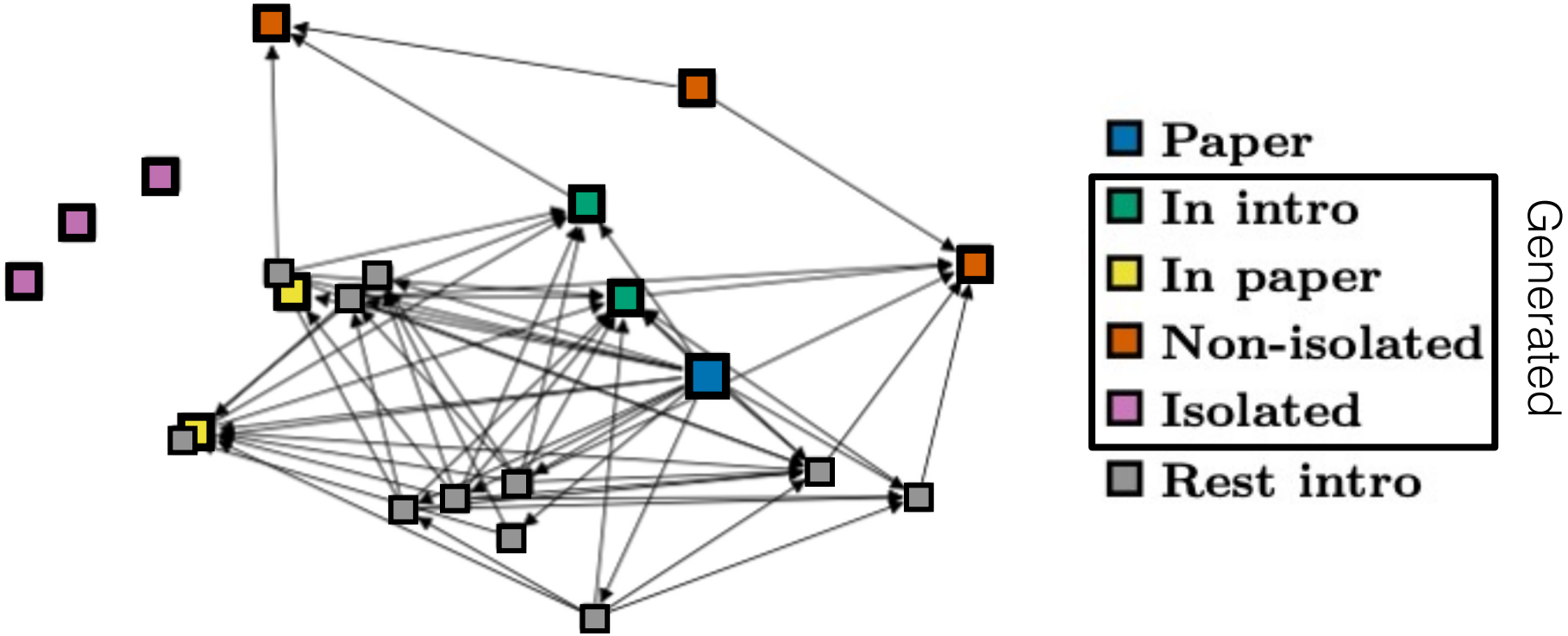
Local citation context of the paper and its introduction references

The Generated References Are Well Embedded In the Local Citation Context



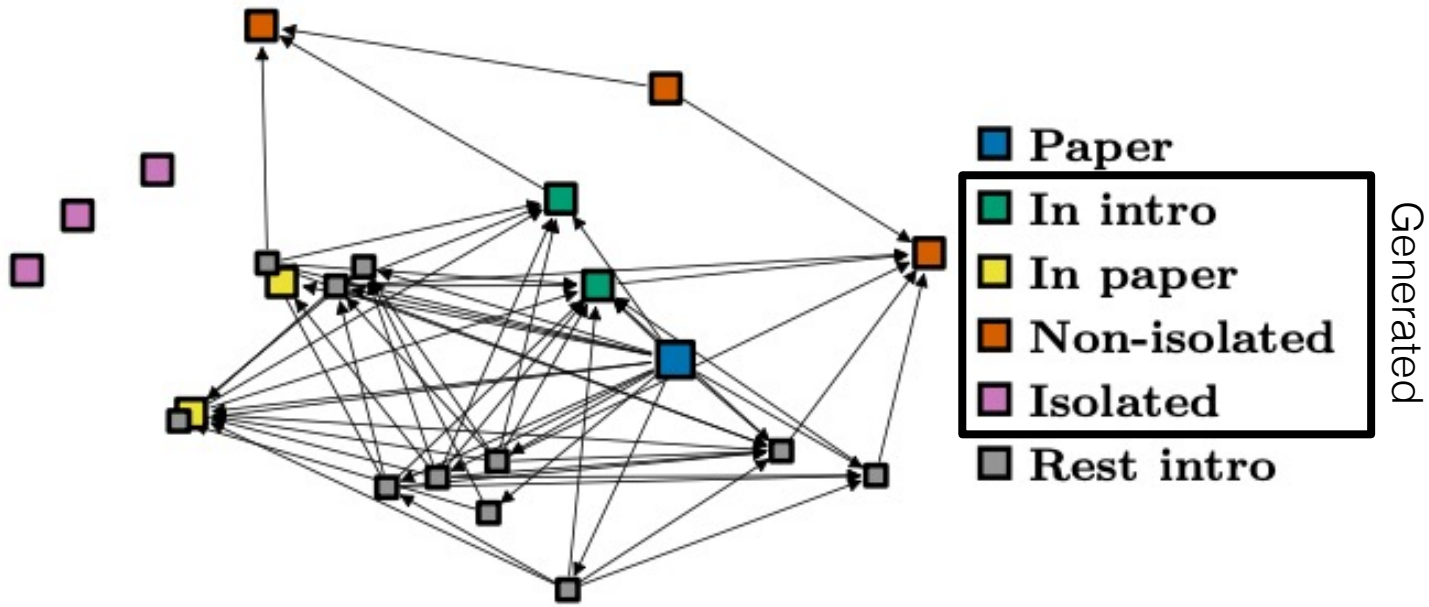
Local citation context of the paper and its introduction references

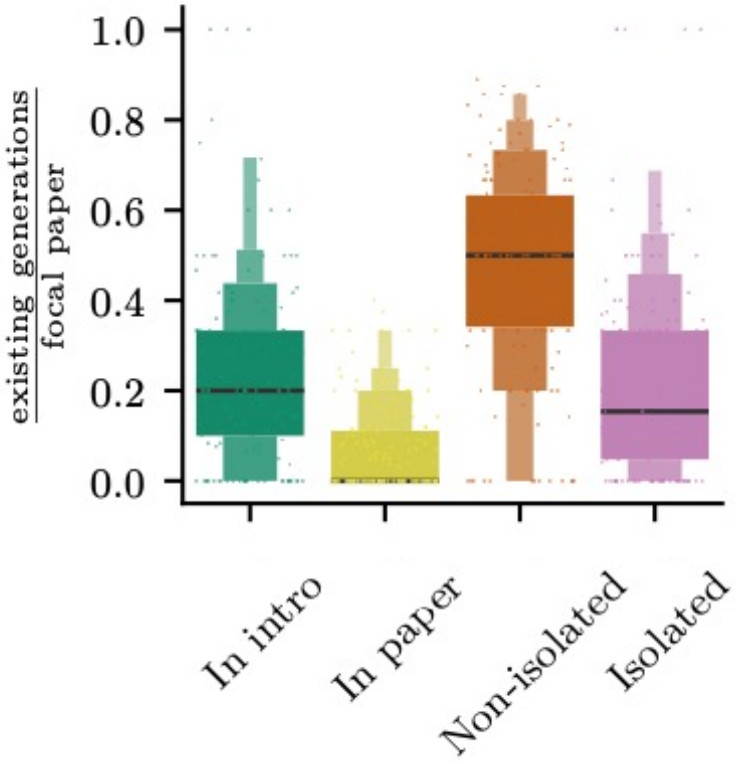
The Generated References Are Well Embedded In the Local Citation Context



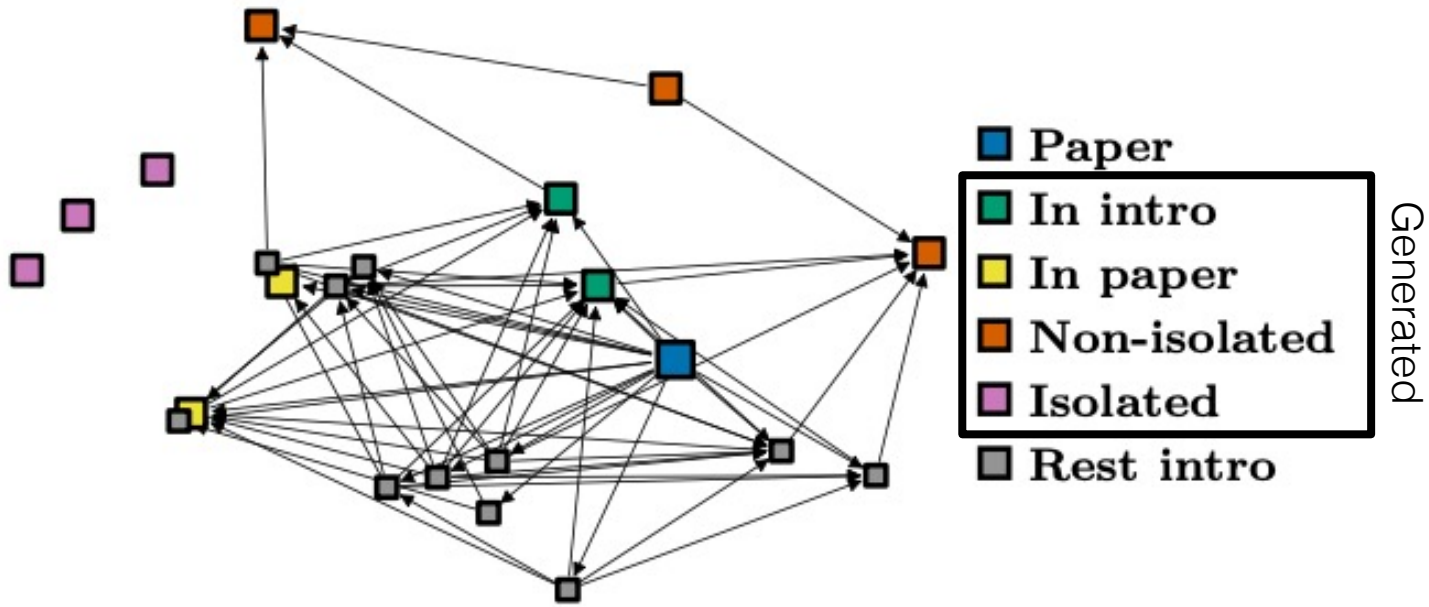
Local citation context of the paper and its introduction references

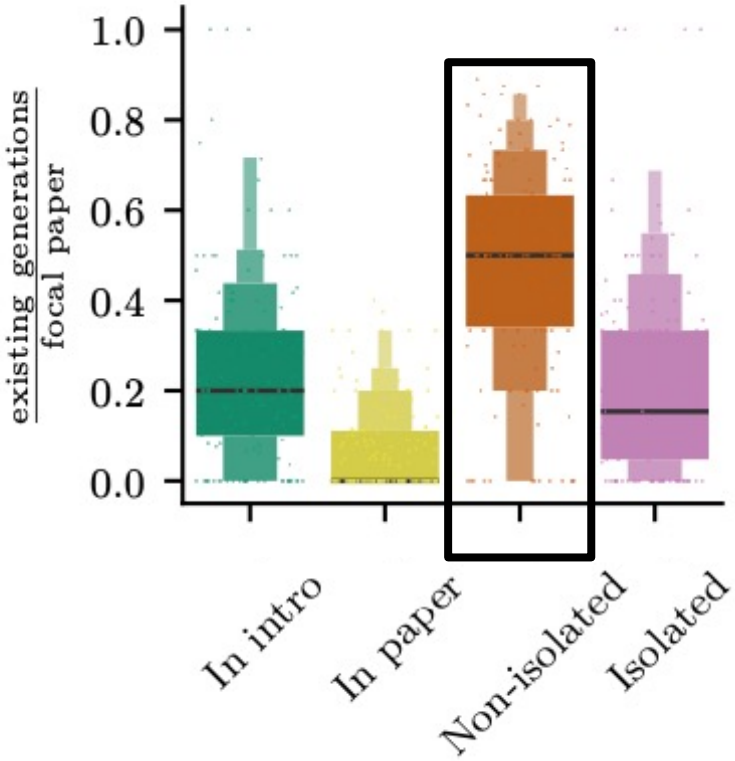
The Generated References Are Well Embedded In the Local Citation Context



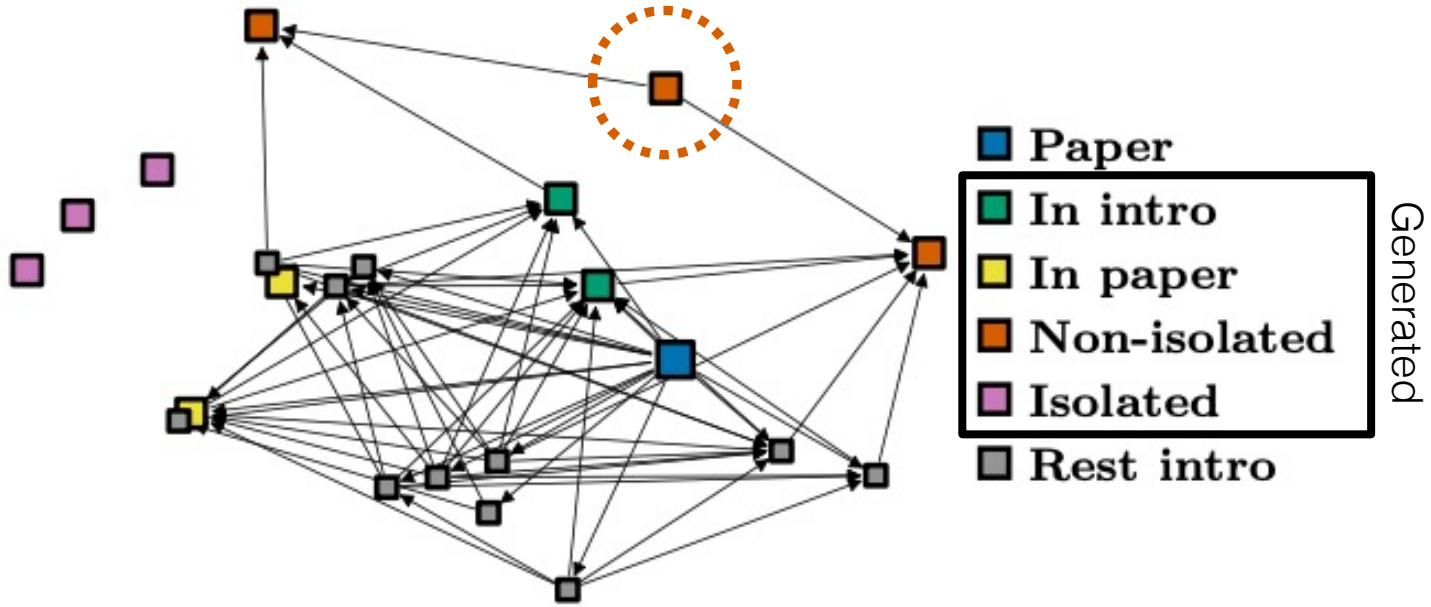


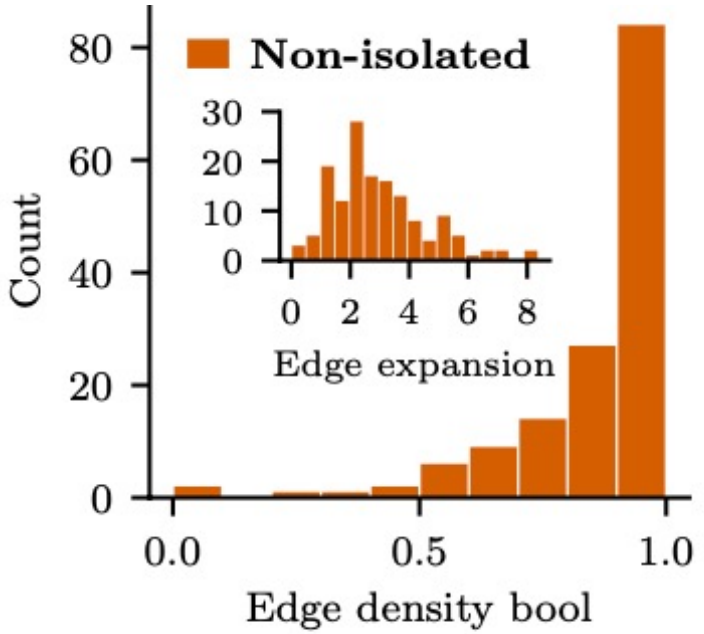
The Generated References Are Well Embedded In the Local Citation Context





The Generated References Are Well Embedded In the Local Citation Context

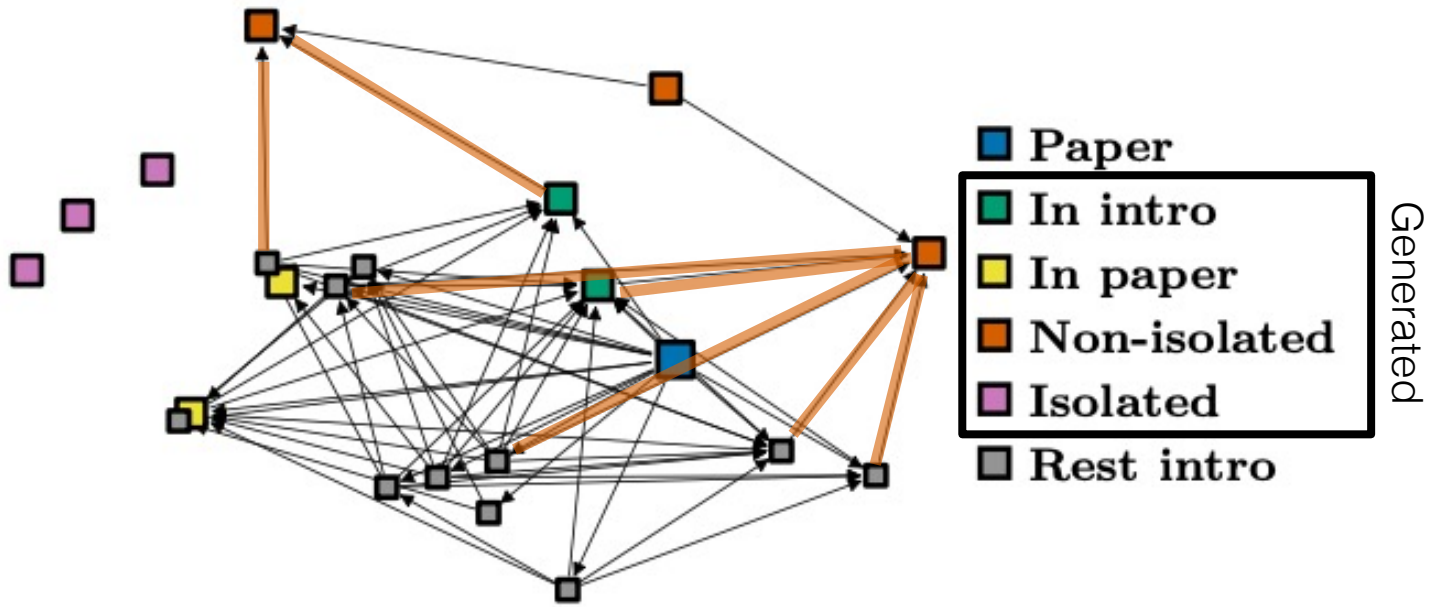


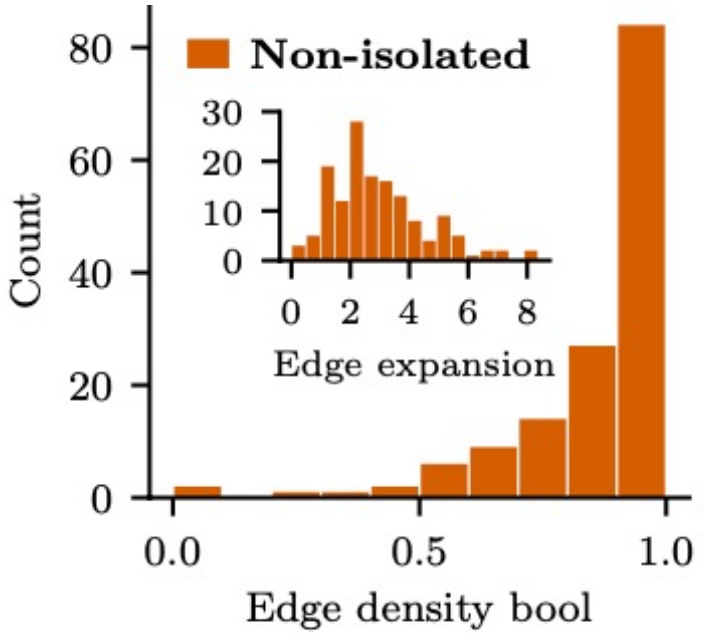


Edge density bool ~ 0.66

Edge expansion ~ 2.33

The Generated References Are Well Embedded In the Local Citation Context

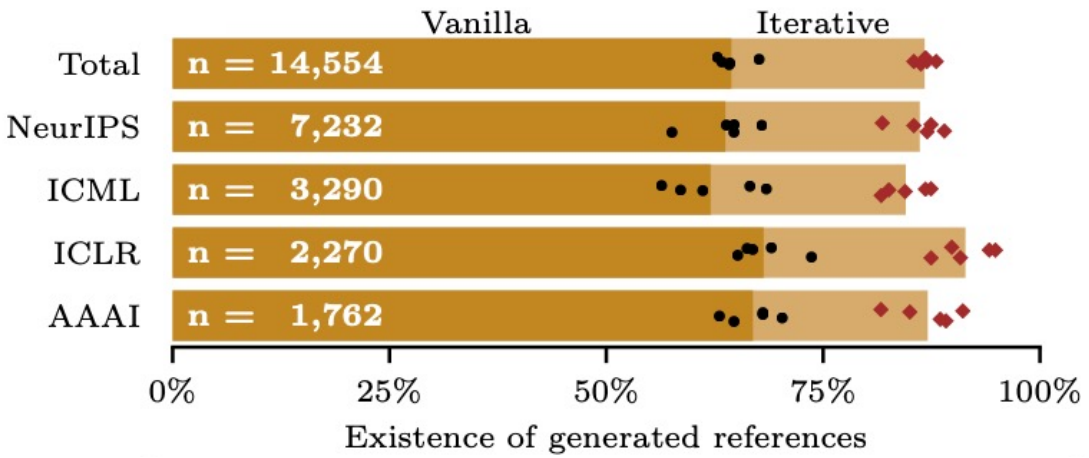




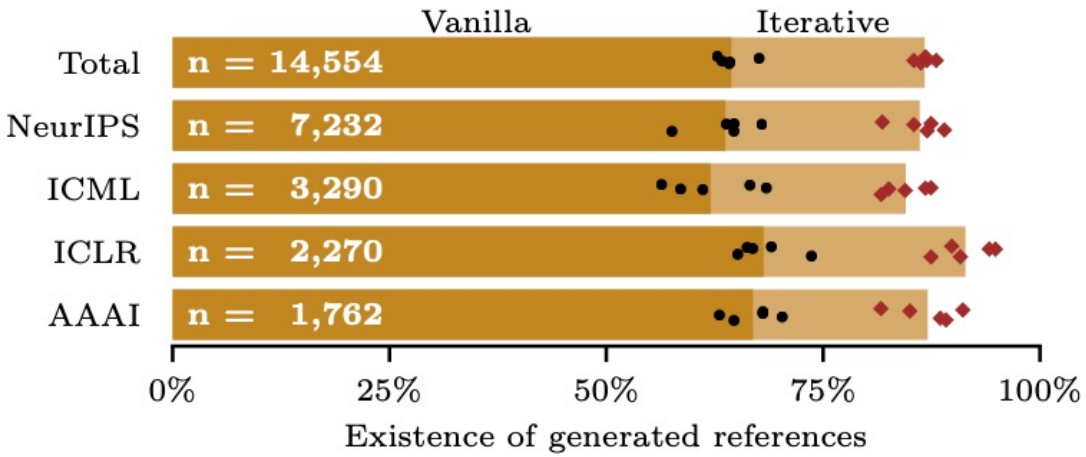
Edge density bool ~ 0.66

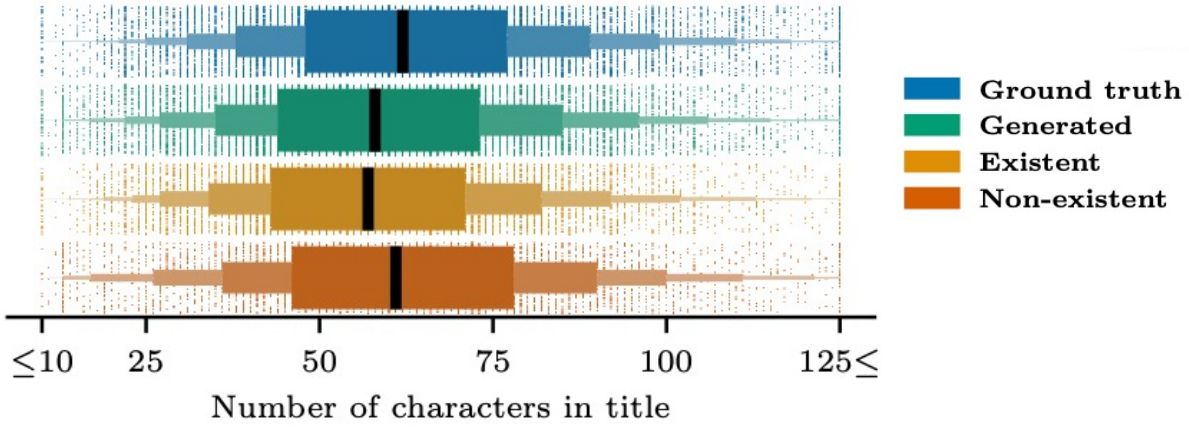
Edge expansion ~ 2.33

Takeaways

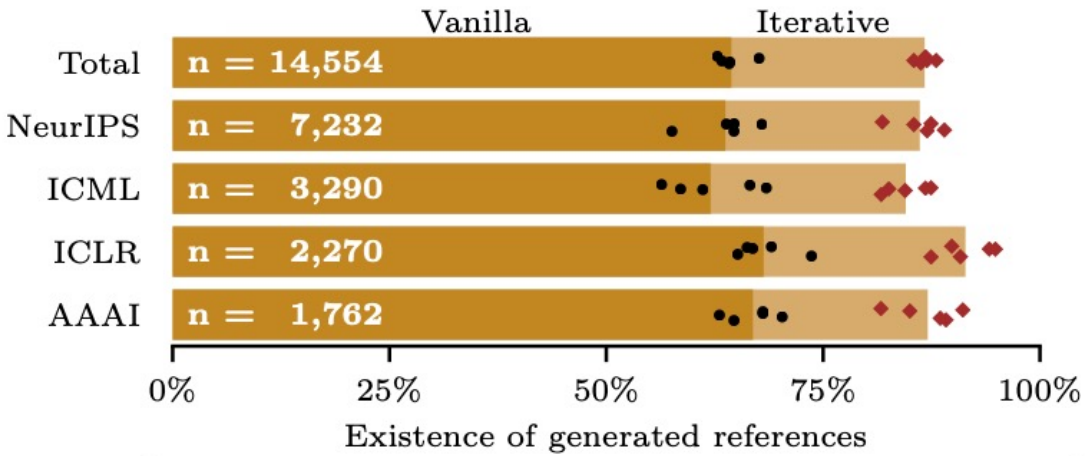

Vanilla
Iterative
Total
n = 14,554
NeurIPS
n = 7,232
ICML
n = 3,290
ICLR
n = 2,270
AAAI
n = 1,762
0%
25%
50%
75%
100%
Existence of generated references

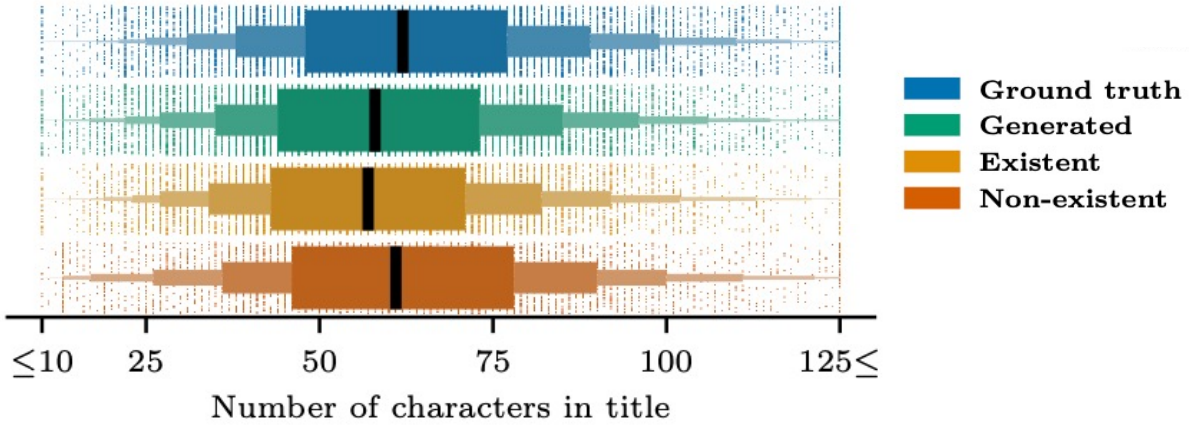
Takeaways

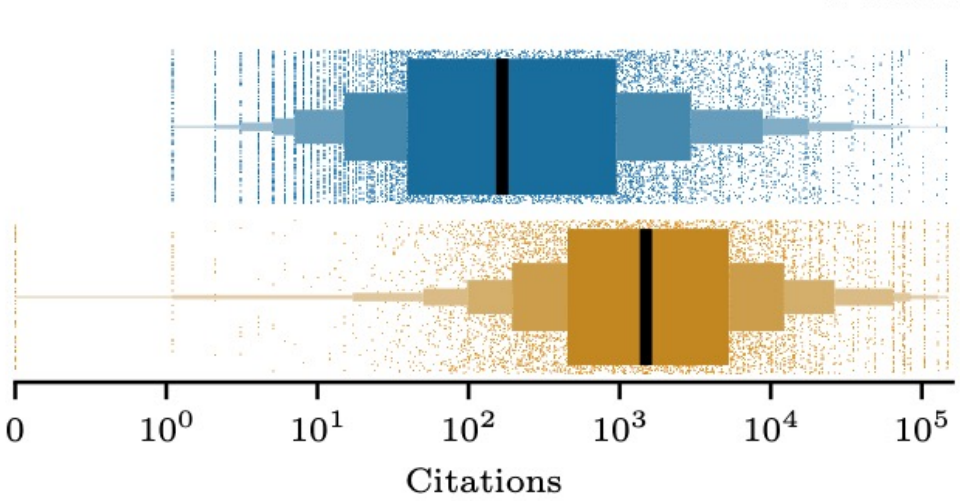

Vanilla
Iterative
Total
n = 14,554
NeurIPS
n = 7,232
ICML
n = 3,290
ICLR
n = 2,270
AAAI
n = 1,762
0%
25%
50%
75%
100%
Existence of generated references


Ground truth
Generated
Existent
Non-existent
≤10
25
50
75
100
125≤
Number of characters in title

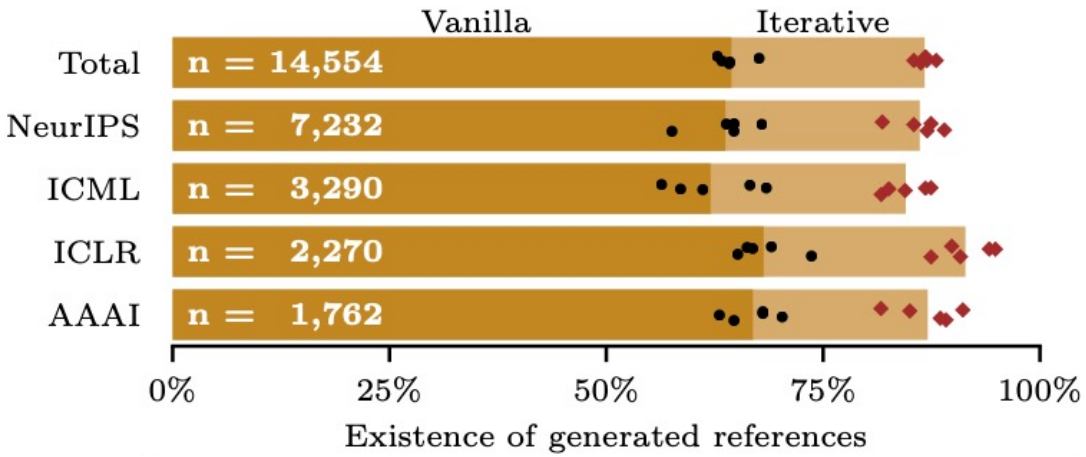
Takeaways

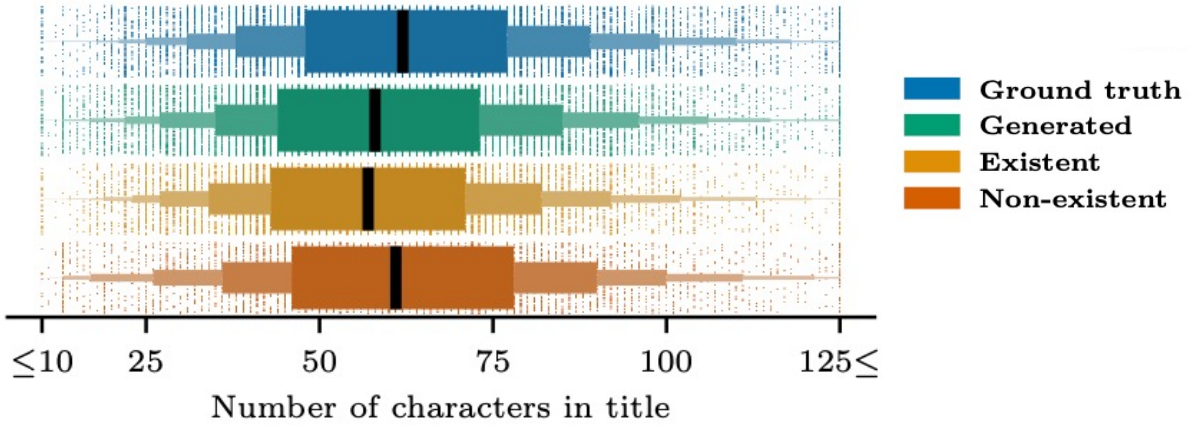

Vanilla
Iterative
Total
n = 14,554
NeurIPS
n = 7,232
ICML
n = 3,290
ICLR
n = 2,270
AAAI
n = 1,762
0%
25%
50%
75%
100%
Existence of generated references

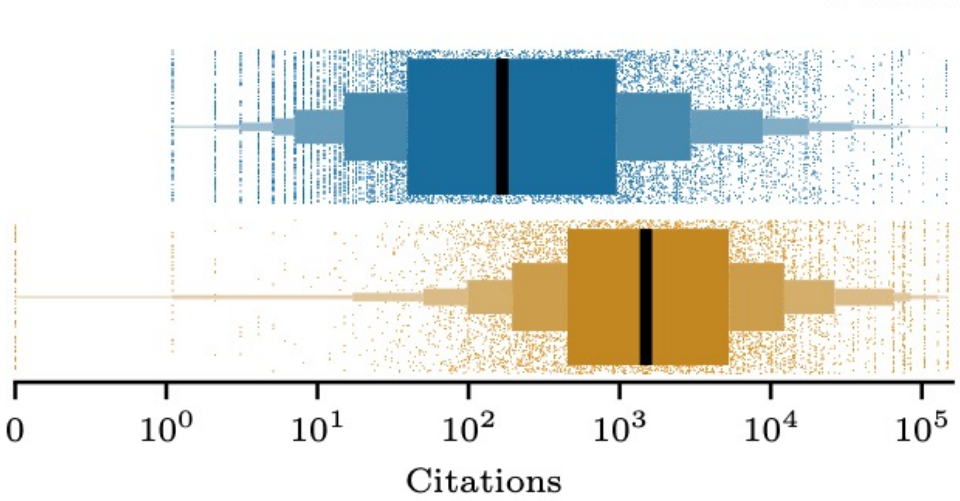

Ground truth
Generated
Existent
Non-existent
≤10
25
50
75
100
125≤
Number of characters in title

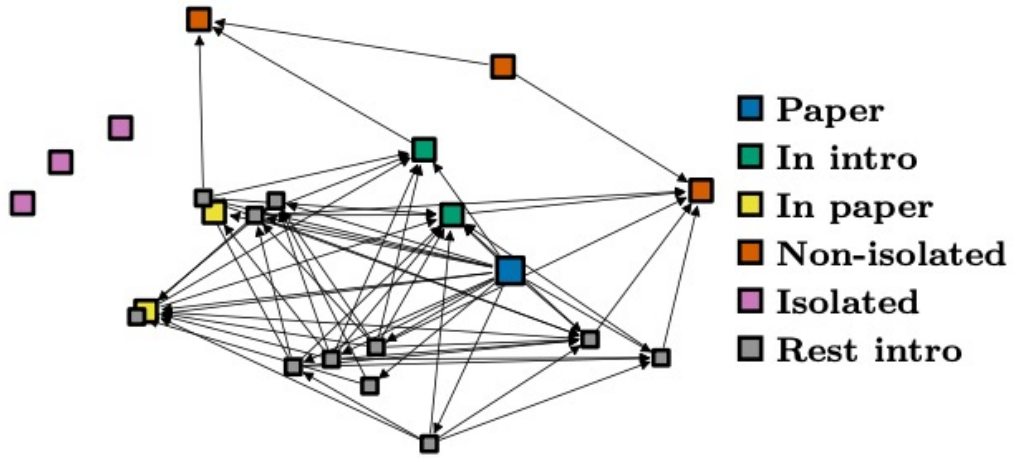

0
10^0
10^1
10^2
10^3
10^4
10^5
Citations

Takeaways


Vanilla
Iterative
Total
n = 14,554
NeurIPS
n = 7,232
ICML
n = 3,290
ICLR
n = 2,270
AAAI
n = 1,762
0%
25%
50%
75%
100%
Existence of generated references


Ground truth
Generated
Existent
Non-existent
≤10
25
50
75
100
125≤
Number of characters in title


0
10^0
10^1
10^2
10^3
10^4
10^5
Citations


Paper
In intro
In paper
Non-isolated
Isolated
Rest intro

Large Language Models Reflect Human Citation Patterns with a Heightened Citation Bias

Main paper

Follow-up work

Across scientific domains (preprint)

Graphs & embeddings (ICLR tiny paper)


Andres Algaba


Carmen Mazijn


Vincent Holst


Floriano Tori


Sylvia Wenmackers


Vincent Ginis

https://github.com/AndresAlgaba/LLM_citation_patterns

https://zenodo.org/records/11299894

VUB DATA ANALYTICS LABORATORY RESEARCH GROUP

KU LEUVEN Centre for Logic and Philosophy of Science (CLPS)

Harvard John A. Paulson School of Engineering and Applied Sciences