



Citations

Cite a key from the .bib: the Transformer replaces recurrence with pure attention (Vaswani et al. 2023).

Suppress the author with [-@key]: Vaswani et al. show (2023) that attention alone suffices.

Add a page with [@key, p. 42]: the attention formula sits in (Vaswani et al. 2023, 4).

Give several sources at once: the concept has multiple roots (Vaswani et al. 2023, 1; 2023, 2).

Citation notes (citekey)

Link a note that carries a citekey: tracing JITs specialize dynamic-language code at runtime (Gal et al. 2009).

A minimal note with citekey, author, title and year also works (Rumelhart et al. 1986).

A link with a heading/block suffix still resolves the citekey, not the suffix as text: (Gal et al. 2009).

Bibliography

Gal, Andreas, Brendan Eich, Mike Shaver, et al. 2009. "Trace-Based Just-in-Time Type Specialization for Dynamic Languages." *PLDI '09: Proceedings of the 2009 ACM SIGPLAN Conference on Programming Language Design and Implementation*, 465–78. <https://doi.org/10.1145/1542476.1542528>.

Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams. 1986. *Learning Representations by Back-Propagating Errors*.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, et al. 2023. *Attention Is All You Need*. arXiv:1706.03762. arXiv. <https://doi.org/10.48550/arXiv.1706.03762>.