



Friday, August 23, 2024

## Today's top trending papers in Computer Science

650,786 papers ranked by PageRank\*. +417 new papers added in the last 1 hour. [Read more.](#)

Filter and sort

### Chip-Chat: Challenges and Opportunities in Conversational Hardware Design

PageRank: 35,621 Growth: +289% Citations: 62

Blocklove, Jason | Garg, Siddharth | Karri, Ramesh | Pearce, Hammond

May 22, 2023 – This paper discusses the challenges and opportunities in using artificial intelligence (AI) and conversational language models for hardware design. The authors present a case study where a...

### TPU v4: An Optically Reconfigurable Supercomputer for Machine Learning with Hardware Support for Embeddings

PageRank: 30,168 Growth: +178% Citations: 97

Jouppi, Norman P. | Kurian, George | Li, Sheng | Ma, Peter | Nagarajan, Rahul | Nai, Lifeng | Patil, Nishant | Subramanian,...

Apr 3, 2023 – TPU v4 is a new supercomputer developed by Google for machine learning models. It utilizes optical circuit switches to improve performance and efficiency, and includes SparseCores that...

### Retrospective: Flipping Bits in Memory Without Accessing Them: An Experimental Study of DRAM Disturbance Errors

PageRank: 51,524 Growth: +164% Citations: 46

Mutlu, Onur

Jun 28, 2023 – The ISCA 2014 paper introduced the RowHammer vulnerability in DRAM chips, demonstrating that it is possible to induce bitflips in real systems by repeatedly accessing a DRAM row...

### Sustainable AI: Environmental Implications, Challenges and Opportunities

PageRank: 10,076 Growth: +160% Citations: 148

Wu, Carole-Jean | Raghavendra, Ramya | Gupta, Udit | Acun, Bilge | Ardalani, Newsha | Maeng, Kiwan | Chang, Gloria |...

Oct 30, 2021 – This paper examines the environmental impact of AI and proposes ways to reduce its carbon footprint through hardware–software design and optimization. It also highlights the challenges an...

### RTLLM: An Open-Source Benchmark for Design RTL Generation with Large Language Model

PageRank: 67,351 Growth: +145% Citations: 39

Lu, Yao | Liu, Shang | Zhang, Qijun | Xie, Zhiyao

Aug 10, 2023 – This paper introduces an open–source benchmark called RTLLM for generating design RTL using natural language instructions. It also proposes a prompt engineering technique called self-...

### ChatEDA: A Large Language Model Powered Autonomous Agent for EDA

PageRank: 68,506 Growth: +144% Citations: 38

He, Zhuolun | Wu, Haoyuan | Zhang, Xinyun | Yao, Xufeng | Zheng, Su | Zheng, Haisheng | Yu, Bei

**Aug 20, 2023** – This research paper introduces ChatEDA, an autonomous agent for Electronic Design Automation (EDA) that utilizes a large language model called AutoMage. ChatEDA streamlines the design...

### ChipGPT: How far are we from natural language hardware design

PageRank: **68,575** Growth: **+121%** Citations: **31**

Chang, Kaiyan | Wang, Ying | Ren, Haimeng | Wang, Mengdi | Liang, Shengwen | Han, Yinhe | Li, Huawei | Li, Xiaowei

**May 23, 2023** – This work explores the potential of using large language models (LLMs) like ChatGPT to assist hardware engineers in generating hardware logic designs from natural language specifications. Th...

### Chasing Carbon: The Elusive Environmental Footprint of Computing

PageRank: **11,360** Growth: **+113%** Citations: **60**

Gupta, Udit | Kim, Young Geun | Lee, Sylvia | Tse, Jordan | Lee, Hsien-Hsin S. | Wei, Gu-Yeon | Brooks, David | Wu, Carole...

**Oct 28, 2020** – This paper highlights the environmental impact of computing and quantifies the carbon emissions associated with computer systems. It reveals that while operational energy consumption is...

### Gemmini: Enabling Systematic Deep-Learning Architecture Evaluation via Full-Stack Integration

PageRank: **33,772** Growth: **+100%** Citations: **37**

Genc, Hasan | Kim, Seah | Amid, Alon | Haj-Ali, Ameer | Iyer, Vighnesh | Prakash, Pranav | Zhao, Jerry | Grubb, Daniel | Liew...

**Nov 22, 2019** – Gemmini is an open-source DNN accelerator generator that considers the cross-stack, system-level effects in real-world environments, allowing for the evaluation of deep-learning architectur...

### GPT4AIGChip: Towards Next-Generation AI Accelerator Design Automation via Large Language Models

PageRank: **94,717** Growth: **+96%** Citations: **27**

Fu, Yonggan | Zhang, Yongan | Yu, Zhongzhi | Li, Sixu | Ye, Zhifan | Li, Chaojian | Wan, Cheng | Lin, Yingyan

**Sep 19, 2023** – This paper explores the use of large language models (LLMs) to automate the design of AI accelerators, aiming to democratize the process and make it more accessible to non-experts. The author...

### RTLcoder: Outperforming GPT-3.5 in Design RTL Generation with Our Open-Source Dataset and Lightweight Solution

PageRank: **103,107** Growth: **+90%** Citations: **27**

Liu, Shang | Fang, Wenji | Lu, Yao | Zhang, Qijun | Zhang, Hongce | Xie, Zhiyao

**Dec 13, 2023** – This study introduces a new open-source language model for generating RTL code, which outperforms commercial models like GPT-3.5. The model is efficient, with a small parameter count and t...

### Splitwise: Efficient generative LLM inference using phase splitting

PageRank: **98,562** Growth: **+88%** Citations: **43**

Patel, Pratyush | Choukse, Esha | Zhang, Chaojie | Shah, Aashaka | Goiri, Íñigo | Maleki, Saeed | Bianchini, Ricardo

**Nov 30, 2023** – The paper discusses the challenge of efficient inference in large language models (LLMs) and proposes a technique called Splitwise, which splits the two main phases of LLM inference onto...