

Christof Teuscher

ECE 410/510: Hardware for AI and ML

Welcome and Introduction

Portland State University
Department of Electrical and Computer Engineering (ECE)
www.teuscher-lab.com
teuscher@pdx.edu

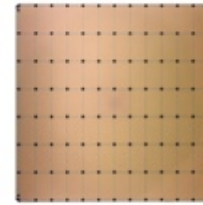


Christof Teuscher • teuscher@pdx.edu

www.teuscher-lab.com/teaching

Cerebras Wafer-Scale Engine

Fabrication process
5nm
Silicon area
46,225mm²
Transistors
4 Trillion
AI-optimized cores
900,000
Memory (on-chip)
44GB
Memory bandwidth
21PB/s
Fabric bandwidth
214Pb/s



Cerebras WSE-3
4 Trillion Transistors
46,225 mm² Silicon

**Electronic chips:
the most complex
human-made
systems ever
created**

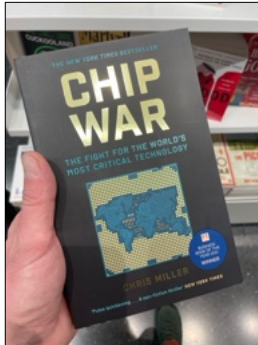
Largest GPU
80 Billion Transistors
814 mm² Silicon



The human brain consists of 100 billion neurons and over 100 trillion synaptic connections.

Christof Teuscher • teuscher@pdx.edu

www.teuscher-lab.com/teaching



Christof Teuscher • teuscher@pdx.edu

www.teuscher-lab.com/teaching

Commodore C64



1982



Radio-frequency
identification (RFID) tag

Same amount of memory!

Christof Teuscher • teuscher@pdx.edu

www.teuscher-lab.com/teaching

**The way we build, program, and use computers has
fundamentally changed**

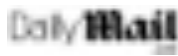


MOS 6510, 1 MHz, 64 kB RAM, ~9000 transistors

Christof Teuscher • teuscher@pdx.edu

www.teuscher-lab.com/teaching





The perfect piece of toast: Scientists test 2,000 slices and find 216 seconds is the optimum time

By GUY MARRAS, SCIENCE REPORTER
Updated on 26/07/2017, 22 July 2017

136
View comments

Scientists today revealed the mathematical formula for a perfect slice of toast, showing that it is best cooked for exactly 216 seconds.

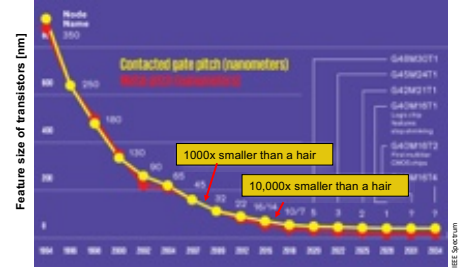
A team of researchers carried out a study which found the optimum thickness is 16mm and the ideal amount of butter is 0.44 grams per square inch.

The recommended cooking time gives the slice a golden-brown colour and the optimum balance of external crunch and internal softness.



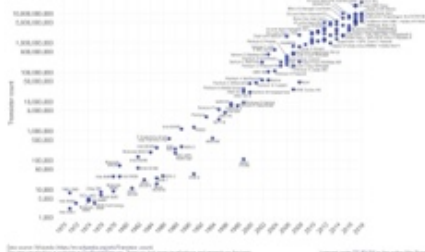
Perfect: The ideal slice of toast is cooked for just over three minutes, according to scientists

Smaller smaller smaller smaller smaller smaller smaller smaller smaller Miniaturization: the cornerstone of progress in computer science

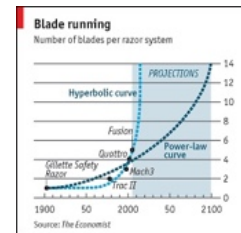


Moore's Law - The number of transistors on integrated circuit chips (1971-2018)

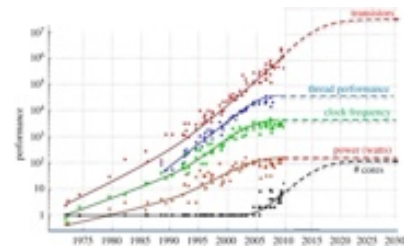
Moore's Law describes the exponential increase in the number of transistors on integrated circuit chips over time. This is a key factor in the exponential growth of computing power.



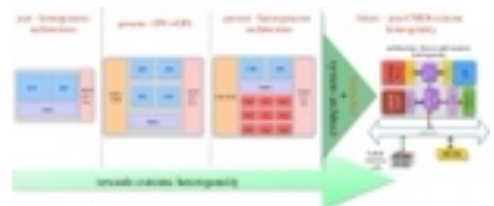
The Cutting Edge: A Moore's Law for Razor Blades?



Source: The Economist, March 16, 2006



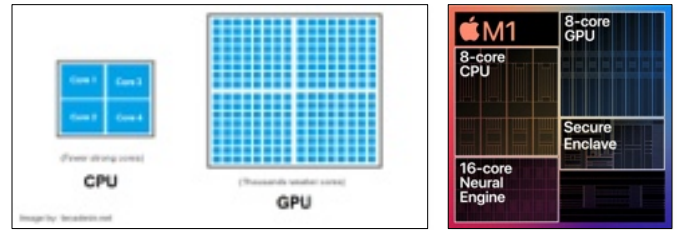
<https://royalsocietypublishing.org/doi/10.1098/rsta.2019.0061>



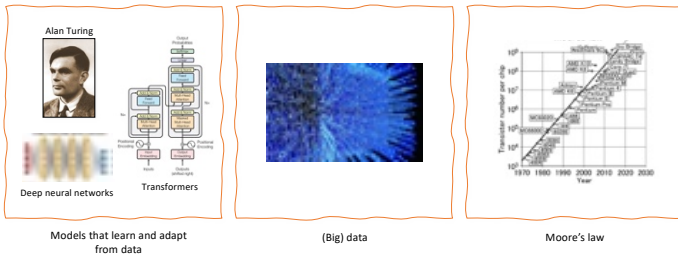
<https://royalsocietypublishing.org/doi/10.1098/rsta.2019.0061>



<https://royalsocietypublishing.org/doi/10.1098/rsta.2019.0061>



How did we get here?

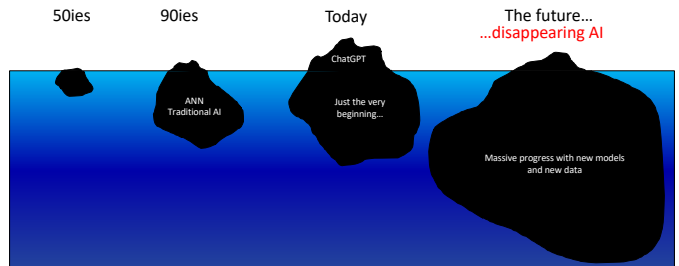


Models that learn and adapt from data

(Big) data

Moore's law

Where is this going?



Hype Cycle for Artificial Intelligence, 2023



AI is no longer hype—it's transforming how we work, innovate, and invest.

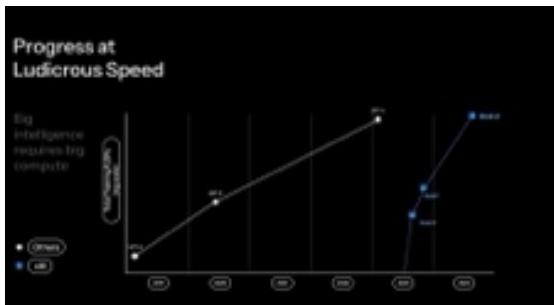
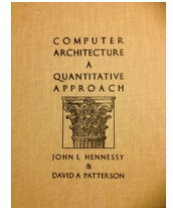
How does AI as a revolutionary technology compare to the steam engine, telegraph, photograph, telephone, automobile, airplane, radio, television, computer, the internet, mobile phone, and the personal computer?



Traditional computer engineering and architecture is dead

AI/ML moves fast

- “In AI, something breaks, ships, or goes viral every six hours.”
- Timescale: ~weeks
- For comparison: Hennessy & Patterson, first edition, 1990
- I have a beautiful 2nd edition from 1996 that I bought for my computer architecture class.



What challenges do computer engineers face?

- Deal with unseen complexities (server-scale AI machines, chiplets, trillions of transistors,...)
- Application-specific architectures for entirely new workloads
- Power barriers
- Shorter time to market
- Drastically increased need to keep up with state-of-the-art technological developments across the entire compute stack.
- Ethical issues
- ...

Synopsys lays out strategy for AI 'agents' to design computer chips

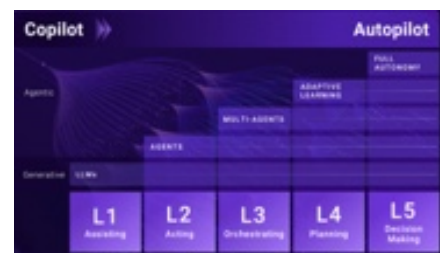


<https://www.reuters.com/technology/artificial-intelligence/synopsys-lays-out-strategy-ai-agents-design-computer-chips-2025-03-19/>

“[...] shift from designing single chips to AI server systems with hundreds or even thousands of chips in them while aiming to release a new server each year”

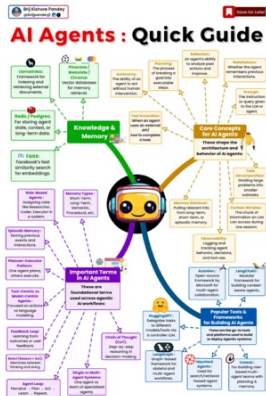
“These are very complex and difficult to design,” Ghazi said of new AI computers. “The pressure engineers are feeling today is not only complexity, it is **complexity** and the **pace by which** they need to deliver these products, as well as the cost.”

AgentEngineer: an “agent” that a human engineer can give instructions to.



Evolution from generative to agentic AI. Source: Synopsys/SNUG

<https://semiengineering.com/first-time-silicon-success-plummets>



From LLMs to Agentic Systems

Christof Teuscher • teuscher@pdx.edu

www.teuscher-lab.com/teaching

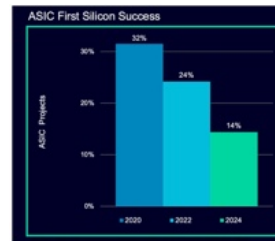
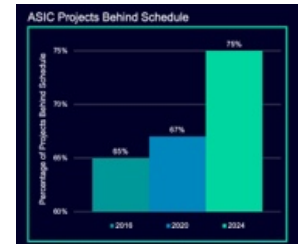


Fig. 1: Number of designs that are functionally correct and manufacturable is declining. Source: Siemens EDA/Wisdom Research Group 2024 Functional Verification Study/DVCon

<https://semiengineering.com/first-time-silicon-success-plummets>



<https://www.nytimes.com/2025/03/14/technology/why-im-feeling-the-agi.html>

Christof Teuscher • teuscher@pdx.edu

www.teuscher-lab.com/teaching

Man files complaint after ChatGPT said he killed his children



<https://www.bbc.com/news/articles/c0kydkr516o>



<https://www.nytimes.com/2025/03/29/opinion/ai-tech-innovation.html>

- LLMs prevent you from thinking and writing.
- You can't just learn how to craft a prompt for an A.I. chatbot without first having the experience, exposure and, yes, education to know what the heck you are doing.
- Learning is a messy, nonlinear human development process that resists efficiency. A.I. cannot replace it.

Christof Teuscher • teuscher@pdx.edu

www.teuscher-lab.com/teaching

Experimental Evidence of the Effects of Large Language Models versus Web Search on Depth of Learning

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5104064

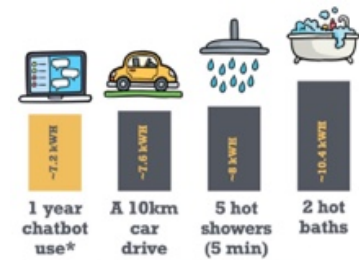
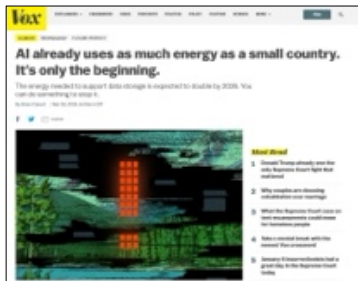
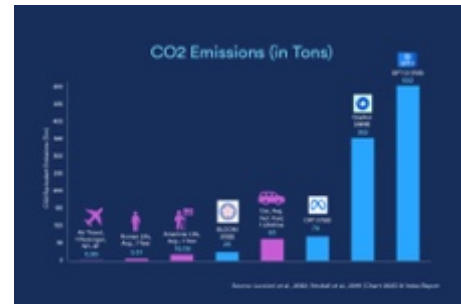
Abstract
The effects of using large language models (LLMs) versus traditional web search on depth of learning are explored. Results from four online and laboratory experiments (N = 6,072) tested support for the predictions that when individuals learn to deal with complex tasks, they tend to develop shallow knowledge that when they have to deal with complex tasks, they often use the same information in the same way. This shallow knowledge is characterized by a lack of understanding of the underlying principles of the task and a reliance on surface-level information. The results suggest that LLMs may lead to shallower learning compared to traditional web search, which requires learners to process and synthesize information themselves. However, in some cases, LLMs may lead to deeper learning, particularly when the task is complex and requires a high level of understanding. The results also suggest that LLMs may lead to a more efficient learning process, as they can provide learners with a large amount of information in a short amount of time. The results have implications for the design of learning environments and the use of LLMs in education.

- Using LLMs leads to developing shallower knowledge compared to standard web search.
- "Processed" knowledge delivered to the learner, even if perfectly summarized, will often lead to shallower learning.
- Active learning is the best learning.

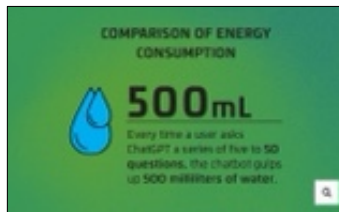
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5104064



<https://www.nytimes.com/2025/03/25/business/economy/white-collar-layoffs.html>



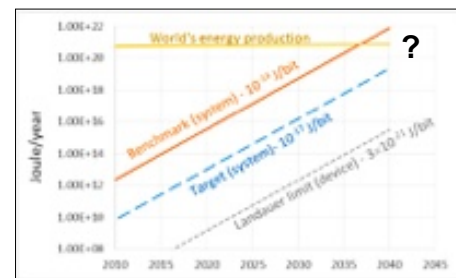
<https://engineeringprompts.substack.com/p/ai-energy-use>



<https://wired.me/science/energy/ai-vs-bitcoin-mining-energy>



<https://www.oregonlive.com>



Estimated total energy expenditure for computing, directly related to the number of raw bit transitions. Source: SIA/SRC.

Progress in AI/ML is currently largely dependent on advances in hardware

Thus this course.
Thus why this is a timely topic of interest to any computer engineer.

Co-design

