

NOVEMBER 2, 2017

Falling Walls: The Past, Present and Future of Artificial Intelligence

This is more than just another industrial revolution—it is something that transcends humankind and even biology

BY JURGEN SCHMIDHUBER

Computing ▾

This article was published in Scientific American's former blog network and reflects the views of the author, not necessarily those of Scientific American

Editor's Note: The [Falling Walls Conference](#) is an annual, global gathering of forward thinking individuals from 80 countries organized by the Falling Walls Foundation. Each year, on November 9—the anniversary of the fall of the Berlin Wall—20 of the world's leading scientists are invited to Berlin to present their current breakthrough research. The aim of the conference is to address two questions: Which will be the next walls to fall? And how will that change our lives? The author of the following essay is speaking at this year's Falling Walls gathering.

As a boy, I wanted to maximize my impact on the world, so I decided I would build a self-improving AI that could learn to become much smarter than I am. That would allow me to retire and let AIs solve all of the problems that I could not solve myself—and also colonize the universe in a way infeasible for humans, expanding the realm of intelligence.

So I studied mathematics and computers. My very ambitious 1987 diploma thesis described the first concrete research on meta-learning programs, which not only learn to solve a few problems but also learn to improve their

own learning algorithms, restricted only by the limits of computability, to achieve super-intelligence through recursive self-improvement.

I am still working on this, but now many more people are interested. Why? Because the methods we've created on the way to this goal are now permeating the modern world—available to half of humankind, used billions of times per day.

As of August 2017, the five most valuable public companies in existence are Apple, Google, Microsoft, Facebook and Amazon. All of them are heavily using the deep-learning neural networks developed in my labs in Germany and Switzerland since the early 1990s—in particular, the Long Short-Term Memory network, or LSTM, described in several papers with my colleagues Sepp Hochreiter, Felix Gers, Alex Graves and other brilliant students and postdocs funded by European taxpayers. In the beginning, such an LSTM is stupid. It knows nothing. But it can learn through experience. It is a bit inspired by the human cortex, each of whose more than 15 billion neurons are connected to 10,000 other neurons on average. Input neurons feed the rest with data (sound, vision, pain). Output neurons trigger muscles. Thinking neurons are hidden in between. All learn by changing the connection strengths defining how strongly neurons influence each other.

Things are similar for our LSTM, an artificial recurrent neural network (RNN), which outperforms previous methods in numerous applications. LSTM learns to control robots, analyze images, summarize documents, recognize videos and handwriting, run chat bots, predict diseases and click rates and stock markets, compose music, and much more. LSTM has become a basis of much of what's now called deep learning, especially for sequential data (note that most real-world data is sequential).

In 2015, LSTM greatly improved Google's speech recognition, now on over two billion Android phones. LSTM is also at the core of the new, much

better Google Translate service used since 2016. LSTM is also in Apple's QuickType and Siri on almost 1 billion iPhones. LSTM also creates the spoken answers of Amazon's Alexa.

As of 2016, almost 30 percent of the awesome computational power for inference in all those Google data centers was used for LSTM. As of 2017, Facebook is using LSTM for a whopping 4.5 billion translations each day—more than 50,000 per second. You are probably using LSTM all the time. But other deep learning algorithms of ours are also now available to billions of users.

We called our RNN-based approaches “general deep learning,” to contrast them with traditional deep learning in the multilayer feed-forward neural networks (FNNs) pioneered by Ivakhnenko & Lapa (1965) more than half a century ago in the Ukraine (which back then was part of the USSR). Unlike FNNs, RNNs such as LSTM have general purpose, parallel-sequential computational architectures. RNNs are to the more limited FNNs as general-purpose computers are to mere calculators.

By the early 1990s, our (initially unsupervised) deep RNNs could learn to solve many previously unlearnable tasks. But this was just the beginning. Every five years computers are getting roughly 10 times faster per dollar. This trend is older than Moore's Law; it has held since Konrad Zuse built the first working program-controlled computer over the period 1935–1941, which could perform roughly one elementary operation per second. Today, 75 years later, computing is about a million billion times cheaper. LSTM has greatly profited from this acceleration.

Today's largest LSTMs have a billion connections or so. Extrapolating this trend, in 25 years we should have rather cheap, human-cortex-sized LSTMs with more than 100,000 billion electronic connections, which are much faster than biological connections. A few decades later, we may have cheap

computers with the raw computational power of all of the planet's 10 billion human brains together, which collectively probably cannot execute more than 10^{30} meaningful elementary operations per second. And Bremermann's physical limit (1982) for 1 kilogram of computational substrate is still over 10^{20} times bigger than that. The trend above won't approach this limit before the next century, which is still "soon" though—a century is just 1 percent of the 10,000 years human civilization has existed.

LSTM by itself, however, is a supervised method and therefore not sufficient for a true AI that learns without a teacher to solve all kinds of problems in initially unknown environments. That's why for three decades I have been publishing on more general AIs.

A particular focus of mine since 1990 has been on unsupervised AIs that exhibit what I have called "artificial curiosity" and creativity. They invent their own goals and experiments to figure out how the world works, and what can be done in it. Such AIs may use LSTM as a submodule that learns to predict consequences of actions. They do not slavishly imitate human teachers, but derive rewards from continually creating and solving their own, new, previously unsolvable problems, a bit like playing kids, to become more and more general problem solvers in the process (buzzword: PowerPlay, 2011). We have already built simple "artificial scientists" based on this.

Extrapolating from this work, I think that within not so many years we'll have an AI that incrementally learns to become as smart as a little animal—curiously and creatively and continually learning to plan and reason and decompose a wide variety of problems into quickly solvable (or already solved) sub-problems. Soon after we develop monkey-level AI we may have human-level AI, with truly limitless applications.

And it won't stop there. Many curious AIs that invent their own goals will quickly improve themselves, restricted only by the fundamental limits of computability and physics. What will they do? Space is hostile to humans but friendly to appropriately designed robots, and offers many more resources than our thin film of biosphere, which receives less than a billionth of the sun's light. While some AIs will remain fascinated with life, at least as long as they don't fully understand it, most will be more interested in the incredible new opportunities for robots and software life out there in space. Through innumerable self-replicating robot factories in the asteroid belt and beyond they will transform the solar system and then within a few hundred thousand years the entire galaxy and within billions of years the rest of the reachable universe, held back only by the light-speed limit. (AIs or parts thereof are likely to travel by radio from transmitters to receivers—although putting these in place will take considerable time.)

This will be very different from the scenarios described in the science fiction novels of the 20th century, which also featured galactic empires and smart AIs. Most of the novels' plots were very human-centric and thus unrealistic. For example, to make large distances in the galaxy compatible with short human life spans, sci-fi authors invented physically impossible technologies such as warp drives. The expanding AI sphere, however, won't have any problems with physics' speed limit. Since the universe will continue to exist for many times its current 13.8-billion year age, there will probably be enough time to reach all of it.

Many sci-fi novels featured single AIs dominating everything. It is more realistic to expect an incredibly diverse variety of AIs trying to optimize all kinds of partially conflicting (and quickly evolving) utility functions, many of them generated automatically (my lab had already evolved utility functions in the millennium that just ended), where each AI is continually trying to survive and adapt to rapidly changing niches in AI ecologies driven

by intense competition and collaboration that lie beyond our current imagination.

Some humans may hope to become immortal parts of these ecologies through brain scans and "mind uploads" into virtual realities or robots, a physically plausible idea discussed in fiction since the 1960s. However, to compete in rapidly evolving AI ecologies, uploaded human minds will eventually have to change beyond recognition, becoming something very different in the process.

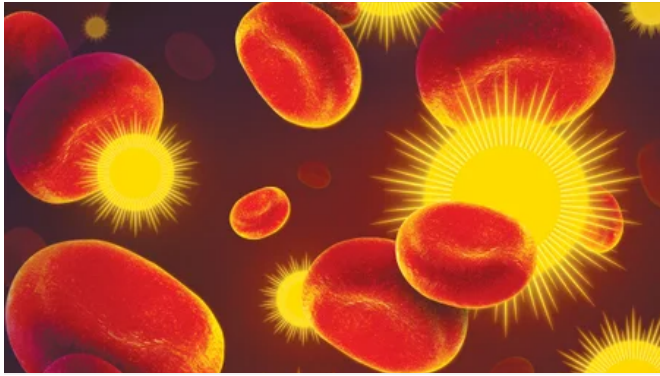
So humans won't play a significant role in the spreading of intelligence across the cosmos. But that's OK. Don't think of humans as the crown of creation. Instead view human civilization as part of a much grander scheme, an important step (but not the last one) on the path of the universe towards higher complexity. Now it seems ready to take its next step, a step comparable to the invention of life itself over 3.5 billion years ago.

This is more than just another industrial revolution. This is something new that transcends humankind and even biology. It is a privilege to witness its beginnings, and contribute something to it.

The Falling Walls Conference is supported by the German Federal Ministry of Education and Research, the Helmholtz Association, the Robert Bosch Stiftung and the Berlin Senate. It receives support and advice from a wide variety of international top-class universities and research institutions as well as foundations, corporations, noted individuals and non-governmental organizations.

[RIGHTS & PERMISSIONS](#)

Popular Stories



HEALTH | MAY 20, 2025

Can Sunlight Cure Disease?

Sunshine may hold healing rays for a variety of autoimmune diseases such as multiple sclerosis. Scientists are turning this surprising discovery into treatments

ROWAN JACOBSEN



BIOLOGY | MAY 20, 2025

Mitochondria Are More Than Powerhouses—They're the Motherboard of the Cell

When these energy-giving organelles thrive, so do we

MARTIN PICARD



EVOLUTIONARY BIOLOGY | MAY 20, 2025

DNA Studies Uncover Unexpected Evolutionary Changes in Modern Humans

Mounting evidence from genome studies indicates that, contrary to received wisdom, our species has undergone profound biological adaptation in its recent evolutionary past

KERMIT PATTISON



QUANTUM COMPUTING | MAY 21, 2025

Single Atom Acts as a Quantum Computer and Simulates Molecules

A quantum computer has used a single atom to model the complex dynamics of organic molecules interacting with light

DAVIDE CASTELVECCHI, NATURE MAGAZINE



CLIMATE CHANGE | MAY 20, 2025

Inside the Bold Geoengineering Work to Refreeze the Arctic's Disappearing Ice

Researchers are trying to rebuild sea ice above the Arctic Circle so it can reflect the sun's warming rays, slowing climate change

ALEC LUHN



ANIMALS | OCTOBER 4, 2023

Cats Are Perfect. An Evolutionary Biologist Explains Why

Cats have attained evolutionary perfection

KATE WONG

Subscribe to *Scientific American* to learn and share the most exciting discoveries, innovations and ideas shaping our world today.

[Subscription Plans](#)

[Give a Gift Subscription](#)