MY WEIRD PROMPTS

Podcast Transcript

The 70-Year Overnight Success: How AI Finally Arrived

Published January 20, 2026 • Runtime: 26:02

https://myweirdprompts.com/episode/history-of-ai-evolution/

EPISODE SYNOPSIS

While the world was stunned by the sudden arrival of generative AI in late 2022, the technology was actually the result of a grueling seventy-year marathon. In this episode, Herman Poppleberry and Corn peel back the layers of AI history, from the optimistic beginnings of the 1956 Dartmouth Workshop to the dark periods known as "AI Winters." They explore why early symbolic logic failed to capture the messiness of the real world and how a small group of dedicated researchers—the "Canadian Mafia"—kept the dream of neural networks alive when no one else would. The duo breaks down the "three pillars" that finally allowed AI to reach its tipping point: sophisticated algorithms, the massive data of the internet, and the unexpected computing power provided by video game hardware. From the "Attention Is All You Need" paper to the emergent behaviors of modern LLMs, this episode provides a comprehensive look at the persistence and breakthroughs that turned a fringe academic curiosity into the defining technology of the 21st century.

DANIEL'S PROMPT

Daniel

I'd like to discuss the evolution of AI and how it has transitioned from an obscure academic field into the mainstream. While it feels like AI exploded overnight with the arrival of the Transformer architecture and ChatGPT, it has a much longer history. I want to explore what the early days of AI were like before it became a practical, everyday tool and how it moved from the margins of technology to dominating headlines today.

TRANSCRIPT

Corn

You know Herman, I was looking at some old tech magazines from the late nineties the other day, and it is wild how we used to talk about the future. Everything was about the information superhighway and pocket organizers. But there was this one tiny column in the back of a two thousand one issue that mentioned something called neural networks as a fringe academic curiosity. It really puts things in perspective when you look at where we are today in early two thousand twenty-six.

Herman

Herman Poppleberry here, and Corn, you are hitting on exactly what our housemate Daniel was asking about in his prompt this week. He was reflecting on how it feels like A-I just fell out of the sky a few years ago. One day we are using Clippy to write a letter, and the next, we have these massive models that can write code, compose symphonies, and simulate entire worlds. But as Daniel pointed out, and as you just hinted at with those old magazines, this was anything but an overnight success. It was a slow, sometimes painful, seventy-year grind that only recently hit an inflection point.

Corn

It is that classic quote, right? Every overnight success is ten years in the making. But with A-I, it is more like seventy years. It is fascinating because most people's memory of A-I starts with ChatGPT in late two thousand twenty-two. Maybe they remember AlphaGo beating Lee Sedol in two thousand sixteen. But before that, A-I was almost a dirty word in some circles. It was the thing that promised the world and never delivered.

Herman

Oh, absolutely. If you were a researcher in the nineteen eighties or nineties, you often had to hide the fact that you were working on artificial intelligence just to get funding. They called them the A-I Winters. There were two big ones. The first happened in the mid-seventies after the initial hype of the nineteen fifties and sixties died down. People thought we would have human-level intelligence in a decade. When it did not happen, the government and private investors just pulled the plug.

Corn

I think that is a really important point to start with. Why did they think it was going to be so easy back then? I mean, we are talking about the Dartmouth Workshop in nineteen fifty-six, which is generally seen as the birth of the field. What was their approach, and why did it hit such a massive wall?

Herman

Well, the early pioneers, like John McCarthy and Marvin Minsky, were focused on what we now call symbolic A-I, or Good Old Fashioned A-I. The idea was that intelligence is basically just logic. If you could write enough if-then rules, you could simulate a human mind. It was very top-down. If you want a computer to know what a chair is, you write a thousand rules describing the legs, the seat, the back, and the function.

Corn

Right, but the problem is the real world is messy. There are three-legged chairs, beanbag chairs, and chairs that look like art pieces. You cannot possibly write enough rules to cover every edge case. We talked about this a bit back in episode ninety-six when we were discussing the evolution of barcodes and how machines struggle with visual patterns. It is that same fundamental issue of trying to define the world through rigid logic.

Herman

Exactly! That is called the combinatorial explosion. As the world gets more complex, the number of rules you need grows exponentially until the computer just chokes. That is why the first A-I Winter happened. The symbolic approach was great for playing chess or solving math theorems, but it could not handle a simple conversation or recognize a face.

Corn

So, if the symbolic approach failed, what was happening in the background? Because the neural networks we use today were actually being discussed even back then, right?

Herman

They were! It is honestly a bit tragic. Frank Rosenblatt came up with the Perceptron in nineteen fifty-eight. It was a very basic version of a single neuron in a neural network. He was so confident that he told the New York Times that the Navy would soon have a machine that could walk, talk, and see. But then, Minsky and Papert wrote a book in nineteen sixty-nine that mathematically proved a single-layer Perceptron could not even solve a simple X-O-R logic gate problem. That book basically killed neural network research for over a decade.

Corn

Wow, so one book essentially froze an entire branch of science? That is incredible. But then we get to the eighties, and things start to thaw out a bit. I remember reading about expert systems. Was that just symbolic A-I making a comeback?

Herman

Pretty much. Expert systems were the big thing in the eighties. Companies spent millions on these massive rule-based systems to help with things like medical diagnosis or oil exploration. And they worked, up to a point. But they were brittle. If you gave them a piece of information that was slightly outside their rule set, they would give you a completely nonsensical answer. This led to the second A-I Winter in the late eighties, triggered by the collapse of the specialized Lisp machine market and the failure of massive projects like Japan's Fifth Generation Computer Systems.

Corn

It is interesting how the history of A-I is basically a series of hype cycles followed by crashes. It makes me wonder about our current moment in early twenty-six. Are we in another hype cycle, or is this time actually different because the underlying technology shifted?

Herman

That is the big question, isn't it? But to understand why today feels different, we have to look at what happened during that second winter. While the world was ignoring A-I, a small group of researchers, often called the Canadian Mafia, kept working on neural networks. People like Geoffrey Hinton, Yoshua Bengio, and Yann LeCun. They were supported by the Canadian Institute for Advanced Research, or C-I-F-A-R, which was one of the few places willing to fund what everyone else thought was a dead end.

Corn

The Canadian Mafia. I love that. It sounds like a group of polite but very determined scientists. And it clearly paid off—I mean, Geoffrey Hinton shared the Nobel Prize in Physics in twenty-twenty-four with John Hopfield for their work on neural networks and machine learning. So, they were working on backpropagation and multi-layer networks while everyone else was focused on the internet and dot-coms. What was the missing ingredient back then?

Herman

It was the three pillars: algorithms, data, and compute. They had some of the algorithms, like backpropagation, which allows a network to learn from its mistakes by adjusting the weights of its connections. But they did not have the data, and they definitely did not have the compute. In the nineties, if you wanted to train a network to recognize a cat, you had to manually feed it thousands of photos, and a top-of-the-line computer would take weeks to process them.

Corn

And then the internet happened. Suddenly, we have billions of photos, articles, and videos being uploaded every day. We inadvertently created the world's largest training set. But even with all that data, you still need the horsepower to crunch it.

Herman

Right, and that horsepower came from an unlikely place: video games. This is one of my favorite parts of the story. Researchers realized that Graphics Processing Units, or G-P-Us, which were designed to render pretty pictures in games like Quake and Doom, were actually perfect for the type of math needed for neural networks. Specifically, matrix multiplication. A C-P-U is like a very smart professor who can do one hard problem at a time. A G-P-U is like a thousand high schoolers who can each do one simple multiplication problem simultaneously. For A-I, you need the thousand high schoolers.

Corn

That is a great analogy. It is funny to think that the reason we have modern A-I is partially because teenagers wanted more realistic explosions in their games. So, when does the explosion Daniel mentioned actually start? Is there a specific moment where the academic world realized the game had changed?

Herman

There is a very specific moment! It is two thousand twelve. The ImageNet competition. For years, researchers had been trying to get computers to identify objects in photos, and the progress was slow, maybe improving by a percent or two each year. Then, a team from the University of Toronto, led by Hinton, entered a neural network called AlexNet. They absolutely obliterated the competition. Their error rate was fifteen percent lower than the next best team. That was the moment the industry realized that deep learning, which is just neural networks with many layers, was the future.

Corn

I remember that being a big deal in the tech news. But even then, it felt very specialized. It was about computer vision. It was not something the average person was interacting with daily, unless they were using Google Translate or something. How do we get from recognizing a picture of a cat to a model that can explain quantum physics to a five-year-old?

Herman

That is where we get into the realm of Natural Language Processing, or N-L-P. For a long time, N-L-P was stuck in the same rut as computer vision. We were using things called Recurrent Neural Networks and L-S-T-Ms, or Long Short-Term Memory networks. They were good, but they had a problem with memory. If you gave them a long paragraph, they would forget the beginning by the time they got to the end. They processed words one by one, in order.

Corn

Which is how humans read, but I imagine it is very inefficient for a machine that wants to understand the context of an entire document at once.

Herman

Exactly. And then came two thousand seventeen. A group of researchers at Google published a paper with what might be the best title in the history of computer science: Attention Is All You Need. They introduced the Transformer architecture. Instead of reading words in order, the Transformer uses a mechanism called self-attention to look at every word in a sentence simultaneously. It figures out which words are most relevant to each other, regardless of how far apart they are.

Corn

So, if I say The bank was closed because the river overflowed, the Transformer knows that bank refers to the edge of the river, not a financial institution, because it sees the word river at the same time.

Herman

Precisely. It can parallelize the processing, which means we could suddenly train models on much, much larger datasets. We went from training on books to training on the entire public internet. And because it was so efficient, we could make the models bigger. More parameters, more layers, more intelligence.

Corn

This is where the scale comes in. We talked about this back in episode two hundred when we looked at the modern A-I tech stack. It seems like once we had the Transformer, the path to G-P-T-three and G-P-T-four was basically just a matter of scaling up the compute and the data. But it still feels like there was a jump in capability that surprised everyone, even the researchers.

Herman

It did! These are called emergent behaviors. When you scale these models to a certain point, they suddenly start being able to do things they were not specifically trained for. They were not trained to write Python code; they were trained to predict the next word in a sequence. But it turns out that to predict the next word in a Python script, you have to actually understand the logic of Python.

Corn

That is the part that blows my mind. It is like if you memorized every word in the library, and suddenly you realized you could speak five languages and solve calculus problems just because you've seen the patterns so many times. But let's bring this back to Daniel's question about the transition to the mainstream. Why did it take from two thousand seventeen to two thousand twenty-two for the general public to feel the impact?

Herman

It was the interface. We had these models, but they were hard to use. You needed to be a developer, you needed to know how to prompt them via an A-P-I, and they were still a bit wild and unpredictable. OpenAI's genius with ChatGPT was not just the model, which was a fine-tuned version of G-P-T-three-point-five, it was the chat interface. They made it feel like you were talking to a person. They used something called Reinforcement Learning from Human Feedback, or R-L-H-F, to align the model's responses with what humans actually find helpful and safe.

Corn

It is like they took this raw, powerful engine and finally put a steering wheel and a dashboard on it. And suddenly, my mom is using it to plan a vacation, and students are using it to help with their homework. It moved from a lab to the kitchen table.

Herman

And that is the transition Daniel was talking about. But even since then, things have moved again. We have transitioned from the era of fast chat to the era of reasoning. In late twenty-four and throughout twenty-five, we saw emerging models like OpenAI's o-one and o-three, and Google's Gemini two-point-zero, which actually pause and think before they answer. They use something called inference-scaling to work through complex problems step-by-step. It is what researchers call System Two thinking.

Corn

So we have gone from a statistical engine that guesses the next word to a system that actually plans its answer. That is a huge shift. Does the fact that it feels like it exploded overnight make it more dangerous because we have not had time to adapt our culture and laws?

Corn

That is a deep one. I think there is a real risk of future shock. When technology moves faster than our ability to understand its implications, we make mistakes. We saw this with social media. We did not realize how it would affect mental health or democracy until it was already woven into the fabric of society. With A-I, the stakes are even higher because it touches everything—work, education, truth itself. We are still catching up to the fact that a video can be perfectly faked or that an essay might not have a human author.

Herman

I agree. And it is not just the social impact; it is the technical one. Because it happened so fast, we are still figuring out how these models actually work under the hood. We know the math, we know the architecture, but we do not always know why a model makes a specific decision. It is the black box problem. As we discussed in episode two hundred fifty-one regarding privacy and backdoors, if we do not understand the internal logic, how can we truly trust it with critical infrastructure?

Corn

It is a bit like we've tamed a wild animal, but we're not quite sure if it's actually tame or just waiting for the right moment to do something unexpected. But let's pivot to the practical side of this. For someone listening who is not a computer scientist, what is the takeaway from this long history? Why does it matter that it started in nineteen fifty-six?

Herman

I think it matters because it gives us a sense of perspective. It reminds us that we are not at the end of the story; we are probably just at the end of the first chapter of the practical era. If it took seventy years to get here, imagine where we will be in another twenty. It also helps to bust the myth that A-I is magic. It is not magic; it is an incredible feat of engineering and math that relies on very human things: our data, our feedback, and our curiosity.

Corn

Right. It is a mirror of us, in a way. It is trained on our collective knowledge. One of my takeaways is that we should be looking for the next Transformer. There is a tendency to think that Large Language Models are the final form of A-I, but history shows us that every dominant paradigm eventually hits a wall and gets replaced by something better. Right now, in early twenty-six, everyone is talking about Physical A-I—putting these brains into robotic bodies that can actually do laundry or cook a meal.

Herman

I love that you brought that up! Physical A-I and autonomous agents are the huge areas of research right now. It is the idea that we need to move beyond the screen. If we can give a neural network the ability to interact with the physical world and follow hard logic rules when it needs to, we might finally solve the remaining hallucination problems.

Corn

Exactly. It would be like giving the creative, intuitive part of the brain a more rigorous, logical partner to check its work. Which, funny enough, sounds a lot like our dynamic here, Herman. You dive deep into the research and the data, and I try to poke at the implications and the logic of it all.

Herman

Guilty as charged! And that is why I find this field so exciting. It is not just about the code; it is about how we think and how we learn. Another practical takeaway for listeners is to realize that the A-I you see in the headlines is just the tip of the iceberg. There is so much happening in specialized fields like protein folding for medicine or climate modeling that does not get the same buzz as a chatbot but might actually have a bigger impact on our lives in the long run.

Corn

That is a great point. We focus on the things we can talk to, but the things that are working silently in the background are often more transformative. It is like the difference between a flashy new car and the invention of the internal combustion engine. One is what you see, the other is what actually changes the world.

Herman

So, looking at where we are now, in early twenty-six, what do you think is the biggest misconception people still have about A-I, given this long history?

Corn

I think the biggest misconception is that A-I is thinking the way we do. Because it is so good at language, we naturally anthropomorphize it. We think it has intentions or feelings. But when you look at the history, from the Perceptron to the Transformer, you see it is really about pattern recognition at a scale we can barely comprehend. It is a statistical engine, not a sentient being. And confusing the two can lead to some really bad decisions, whether it is in how we regulate it or how much we trust it.

Herman

That is spot on. It is a tool, an incredibly sophisticated one, but a tool nonetheless. It is like a super-powered version of those old expert systems, but instead of us writing the rules, the machine found the rules in our data. But it does not know what it is saying. It just knows that in this context, these words are the most likely to follow those words.

Corn

It makes me think about Daniel's point about it becoming an everyday tool. We are moving from the wow phase to the utility phase. It is becoming like electricity or the internet. You do not think about the history of the power grid when you flip a light switch; you just expect the light to come on. We are reaching that point with A-I where it is just... there. It is in our emails, our maps, our medical records.

Herman

And that is when it gets really interesting, because that is when the second-order effects start to kick in. When everyone has access to a world-class tutor or a legal assistant in their pocket, how does that change the economy? How does it change the value of a college degree? These are the questions we are going to be tackling for the next decade.

Corn

It is a bit like what we discussed in episode one fifty-one about internet speeds. Once the infrastructure is there, the way we use it changes completely. We move from just having the internet to living on it. We are starting to live with A-I.

Herman

We really are. And I think we should take a moment to appreciate the sheer human effort that went into this. All those researchers who worked through the A-I Winters when no one cared, who were told their ideas were dead ends. People like Hinton, who spent thirty years being the odd man out in computer science. Their persistence is the reason we are having this conversation today.

Corn

It is a good reminder to stay curious and not to dismiss things just because they are currently out of fashion. The weird idea of today could be the mainstream tool of tomorrow. That is basically the mission statement of this podcast, right?

Herman

Exactly! Exploring those weird prompts and seeing where they lead. And speaking of curiosity, if you have been enjoying our deep dives into these topics, we would really appreciate it if you could leave us a review on your podcast app or over on Spotify. It genuinely helps other curious minds find the show.

Corn

It really does. And if you want to get in touch or see the show notes for this episode, you can always find us at myweirdprompts.com. We have the full archive there, including all those early episodes Daniel was mentioning earlier.

Herman

Yeah, we have come a long way since those first hundred episodes. It has been quite a journey, much like the history of A-I itself.

Corn

Well, I think we have covered a lot of ground today. From the Dartmouth workshop to the Transformer, and why Daniel's feeling of an overnight explosion is both right and wrong. It is a fascinating story of human persistence and the power of scaling simple ideas.

Herman

It really is. Thanks for the great discussion, Corn. And thanks to Daniel for the prompt. It is always fun to look back at the roots of the tech that is shaping our world.

Corn

Definitely. Alright everyone, thanks for listening to My Weird Prompts. We will be back next week with another deep dive into whatever is on Daniel's mind.

Herman

Or yours, if you send us a message! Until next time.

Corn

See ya.