

# HumanProgress



CHELSEA  
FOLLETT

WILL  
DUFFIELD

## THE AI REVOLUTION

THE  
**HumanProgress**  
PODCAST

Artificial intelligence made a sudden leap forward last year, with deep learning models like DALL-E, Midjourney, and ChatGPT delighting millions of people with impressive, if uncanny, artwork and text.

But how will these technologies impact our lives? Will they really replace artists and writers? And what are the risks and benefits of increasingly powerful AI?

In the latest episode of The Human Progress Podcast, policy analyst and technology expert Will Duffield joins Chelsea Follett to discuss these questions and more.

Watch the episode [here](#).  
Or, listen to the [audio](#).

*Below is an abridged transcript featuring some highlights from the interview.*

### Recent advances in artificial intelligence have been all over the news, from DALL-E to ChatGPT. How significant are they really?

Recently, there have been several advances in large language modeling. Essentially, by gathering large corpuses of language training data, AI pioneers have been able to derive ever more effective models.

These large language models don't reason deductively like humans. Instead, they infer information by synthesizing great reams of data. But even if we see these models as mere curve fitters, things that are good at recognizing patterns, their pattern recognition is useful to us and, in some cases, even outstrips our abilities.

There've been studies throwing millions of iris images at artificial intelligence. And while we don't know of any difference between the male and female iris, this AI has been able to, with a fair degree of certainty, determine whether the iris belongs to a man or a woman. So, it's seeing a pattern that we don't even know exists. It does this by comparing irises in a way that you or I never could. By iris 20, they'd all look the same to us, but a machine doesn't have the same constraints. And so, it can make associations that might not occur to us.

### How else are AI technologies being used in the real world?

One of the more exciting areas is drug discovery, using AI to predict the results of combining molecules rather than synthesizing everything in a lab. A lot of that early-stage testing can now be done virtually using artificial intelligence.

There are also tools like Copilot at GitHub, which allow someone to explain the code that they would like and then receive code generated by an AI. It can help a novice coder learn or punch above their weight much more quickly, turning to the AI in the same way you might turn to a more experienced coder.

The most exciting possibility to me is that AI will help us interface with complex modern systems. Everything from your county website, where you pay your taxes and deal with parking tickets, to payroll software, to booking a flight can be very difficult to navigate. In the most hopeful world, AI will be an everyday interface between human beings and the various systems we've built that don't tend to work well with us.

### How do you think advances in AI technology will impact the job market?

Unlike something like the power loom, we aren't talking about displacing manual labor but instead rote white-collar tasks like data entry, basic research, and legal assistance, as well as some creative tasks.

There's been a tremendous amount of concern from artists about generative AI. And while lots of people are using it, not all that use causes displacement. Lots of people have the new Lensa profile picture, but very few of those people were previously paying an artist to create a profile picture for them. They were relying on a photograph or stock imagery. So, a lot of this fills in the parts of the market that weren't well served before.

At the end of the day, we still need people in charge of these tools. And at their best, they can save us from drudge work that, frankly, people deserve better than. It's undignified for people to spend all day acting like a machine in a toll booth.

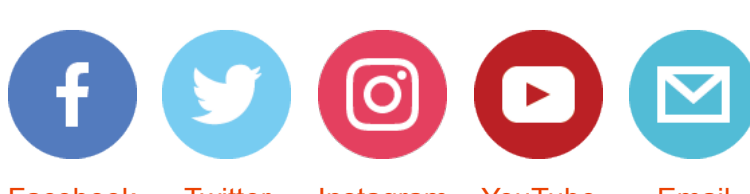
### What are some challenges posed by artificial intelligence?

People often bring up the so-called paperclip maximizer scenario. Say you have an AI that is ostensibly aligned with your goals, and you set it to make a lot of paperclips. If it's very powerful and can have factories built and draw up resources, it will slowly start making everything into paperclips. That would obviously be a bad scenario for everyone. So, it's not just thinking of the genocidal evil machine that wants to destroy all humans, but also how to put side constraints on AI to prevent it from taking harmful steps to achieve its goal.

Another challenge is that, if AI is an effective personal assistant, it won't just be an effective personal assistant for good people. We can imagine a friendly artificial assistant helping people who otherwise couldn't build a bomb to build one.

There's a lot of debate right now about how to either regulate AI or place side constraints on the models themselves to resolve these challenges. Getting that right is important, but it will also be very important not to neuter this technology or gate it away from the everyman. If only the vetted or powerful have access to it, then you're just going to exacerbate existing disparities in wealth and power. Even though there are potentials for misuse, many of those cases seem worse in a world in which only some people have access to the technology.

Read the full transcript [here](#).



Facebook Twitter Instagram YouTube Email