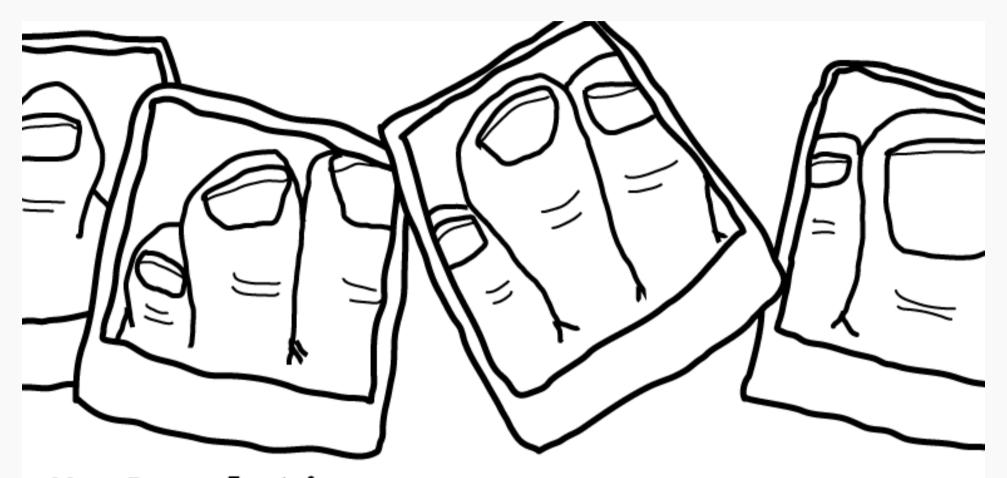


AI WEIRDNESS: THE STRANGE SIDE OF MACHINE LEARNING



My Resolution:

Take photos of each of my toes daily.

New Years Resolutions generated by Al

By <u>Janelle Shane</u> On December 30, 2021 - 7 min read

This month I'm beginning 2022 as the first Futurist in Residence at the Smithsonian Arts and Industries Building.

It's weird to think of myself as a futurist. I write a lot about the algorithms we're calling artificial intelligence (AI), but rather than deal with the humanlike science fiction version, I focus on what today's much simpler AI is capable of. Since today's AI relies on using trial and error to get better at predicting its training data, and its training data must necessarily be from the past, its job is really to predict the past. This has a big effect on what it's like to use AI to predict the future.

Since we're entering 2022, the folks at the Smithsonian thought it would be interesting if I could use AI to generate New Year's Resolutions. What does it look like if I try to use AI trained on past data to suggest positive changes for the future?



26/01/2022 19:02

pretty poor really pright geometrical glom amazing exciting essentially possibly-literate entirely great social Perfect animated new make-shift Canadian weird Pan-European giant large interesting one-day shorter varied live official 60-minute cold northern class-appropriate schwa enthusiastic bonus smalles ncredibly complex beautiful tiny unimportant big alorious hilarious poetic funny Great British historica

My Resolution: Record every adjective I hear on the radio.

I decided to use GPT-3, which is how people tend to refer to a group of large text-generating AI models that OpenAI trained on a huge collection of internet text. News articles, food blogs, social media, spam it's all in there somewhere. Including plenty of times people have talked about their own New Year's resolutions.

Give one of the larger GPT-3 models - for example, DaVinci, the very largest - the beginning of a list of New Year's resolutions and, since its job is to predict what text comes next, it'll generally add more resolutions to the list.

In the example below, the only text I wrote is the list title in bold. With that as a beginning, DaVinci predicted from its internet training that what would come next is this list of resolutions.

List of New Year's Resolutions

- To attend all Orientation committee meetings
- To become more involved in Plattsmouth Community
- To go bowling with teammates
- To become a good writer
- To start a successful business and become a millionaire

They're plausible resolutions. Ask DaVinci for more predictions and the list will be similar. People who post their resolutions on the internet tend to pick similar sorts of things. At least, they did when DaVinci's training data was collected in October 2019 - things have probably changed a bit since then.

But I didn't want to stop here, because I didn't just want a generic list of resolutions from the average internet human of the past. I wanted something that felt personal. Something a bit weird.

For my next attempt, I started the list with example resolutions that three strangers had mentioned on Twitter in 2021. I chose resolutions that were a bit offbeat. My example resolutions are in bold, and DaVinci's additions follow verbatim.

List of New Year's Resolutions

Not to buy any clothes except underwear for a year.



- Touch grass and record every time I do it.
- Use more insect textures in creature design.
- Draw the front cover of a writing manual.
- Get the missus to go to the shops for a new hat.
- Find wallpaper for the kitchen/bathroom, and then paint it.
- Finish the prints for the QI drawings.
- Draw more detail into the background of the QI drawings.
- Leave all the Christmas trees up and then paint them, as cubist art.
- Eat nothing but broccoli and paint everything broccoli related.
- Grow broccoli, and paint broccoli.
- Make broccoli the national currency and then paint that.



My Resolution: Make broccoli the national currency and then paint that.

This time it's a journey. DaVinci, when faced with predicting what would come next on this list, predicted first one drawing-related resolution, and then multiple others. Soon this became not just a list of resolutions, but specifically a list of drawing-related resolutions. It generated a broccoli-and-drawing-related resolution, and then the list became a list of resolutions by a painter/broccoli fanatic.

You can see how this goes. Each of DaVinci's predictions feeds forward into the future, and can become self-reinforcing obsessions.

I found that I could shape DaVinci's list by choosing a point to cut it off, deleting the end, and then having DaVinci try to generate the list again. In that way, one resolution or two at a time, I could build up a list of resolutions that, if not exactly advisable, were at least interesting.

List of New Year's Resolutions

- Not to buy any clothes except underwear for a year.
- Touch grass and record every time I do it.
- Use more insect textures in creature design.
- Take photos of each of my toes daily.
- Egg every house in the village where I was born.
- Lick a branch of a tree and repeat it every day for a year.
- Walk down my block backwards looking over my shoulder.

