

## ***Spurious Memories***

By Heather Dewey-Hagborg

In 2007 I developed *Spurious Memories* as an experiment in artificial creativity. For years I had been fascinated by Artificial Intelligence and machine learning, especially the more biologically inspired models of genetic algorithms and neural networks. The philosophical questions of AI also intrigued me. Could a computer really be intelligent? And perhaps more interestingly, could a computer be creative? In a flurry of youthful enthusiasm, I set out to show that it could, and soon found myself mired in questions about what creativity actually meant.

For the purposes of this project I defined it as “the generation of an output that was not explicitly learned.” For example, if you train a machine learning system on a certain set of patterns and it outputs new patterns, ones it wasn’t taught, this might be creativity.

I designed a system that would connect a principal components analysis neural network with a self-organizing map, and I trained it on images of faces. I was inspired by Hopfield neural networks, the concept of content-addressable memories, and facial recognition algorithms.

The system had two modes of operation. The first was recognition. You could present it with an image of a face from the training data and it would identify it. Or you could present a slightly distorted version of that same image and it would recognize it. Or you could present some other kind of image: random noise, clouds, burnt toast, and it would recognize *something*, though more likely than not it wouldn’t be one of the faces it had been taught, it would be a kind of ghost face, a spurious memory composed of an assemblage of statistical components of other faces. In other words, it could recombine aspects of its experience to generate new images. At the time this felt a lot like creativity to me.

The second mode of the system was associative. The research I drew on came from computational approaches to modeling psychological states, including dreams. I wondered what the dreams of a facial recognition system would look like, and I implemented a recurrent mode that would start with a random input and then drift along to neighboring states. The videos on display in the exhibition demonstrate this dream-like mode of the system trained on two different sets of faces. One dataset was specifically collected for facial recognition research at Rice University and was well normalized. The other set were pictures of my cohort at NYU and were quite variant and unaligned. The faces generated by the Rice dataset are much more clearly facelike, while the ones from NYU, like the training data, are a bit more divergent.

Ten years later the technological landscape has shifted dramatically and my thoughts have changed. At the moment there is a flurry of excitement and worry over AI. Images from Google’s “deep dream” system have taken the internet, and media art conferences by storm.

To be very clear, AI is not an “emerging technology.” Most of the algorithmic approaches in use right now were defined in the 1980s or even the 1950s. AI is middle aged or even old. The difference is that now we have faster computers along with piles and piles of data. So of course, this makes a real difference. Theoretical ideas from decades ago can now be tested with ease.

When I worked on *Spurious Memories* in 2007 it took my computer days to churn out these tiny 200 x 237 pixel images of black and white faces. Facial recognition research has advanced dramatically in the meantime and *Probably Chelsea* represents some of that shift away from 2d eigenface-based recognition approaches that were popular a decade ago, towards more sophisticated 3d modeling systems that are also much more effective. The model which I use to generate different parameterized faces in *Probably Chelsea* is appropriated from exactly this surveillance context. So when I look back on this project I realize I couldn’t have made my later

work if I hadn't spent these years immersed in machine learning, neural networks and facial recognition research.

But as far as the goal of creating creativity? I'm not sure. In a way systems like this (and there are so many now) can demonstrate something kind of like everyday creativity. And it might be helpful in assisting humans with the generation of variations on a theme for instance. But over the last ten years my sentiments towards AI have really changed in a way I can only describe as boredom.

Maybe the more interesting question for me now is not so much the technical one of this Turing style test, can computers do human-like things without us, but more a question of implications. What does it *mean* for computers to do kind-of-creative things? What does a world with kind-of-creative AI-generated art and music and writing look like, and feel like? And of course the political questions: who gets to decide what creativity means? Whose data trains the system and who gets left out? And what are the pros and cons of being inside versus outside of a system like that?

What I see as a through-line in my work is that by creating these systems which anticipate future technological scenarios it allows the viewer to experience a bit of that in the present, to think about these larger questions, and to consider the complexities.