

University of Mary Washington

**Eagle Scholar**

---

Student Research Submissions

---

Spring 5-7-2021

## I Like To Think I'm Making Something New: An A. I. Attempts Narrative

Kelly Donaghey

Follow this and additional works at: [https://scholar.umw.edu/student\\_research](https://scholar.umw.edu/student_research)



Part of the [Creative Writing Commons](#)

---

### Recommended Citation

Donaghey, Kelly, "I Like To Think I'm Making Something New: An A. I. Attempts Narrative" (2021). *Student Research Submissions*. 398.

[https://scholar.umw.edu/student\\_research/398](https://scholar.umw.edu/student_research/398)

This Honors Project is brought to you for free and open access by Eagle Scholar. It has been accepted for inclusion in Student Research Submissions by an authorized administrator of Eagle Scholar. For more information, please contact [archives@umw.edu](mailto:archives@umw.edu).

Kelly Donaghey

Professor Whalen

ENGL 491

10 December 2020

## I Like To Think I'm Making Something New: An A. I. Attempts Narrative

### Abstract

The purpose of this semester-long independent study was to evaluate various methods of computer-generated narratives with the goal of creating and deploying a model that would write a literary science fiction short story on the topic of artificial intelligence. Various methods of language generation were tested before settling on OpenAI's GPT-2 model. Narrative generation methods were also studied but ultimately not deployed due to the complex nature of narrative computation. Thus, in lieu of a computer-generated narrative with computer-generated language, human intervention was used to curate and assemble a narrative out of computer-generated text. This paper provides an overview of which methods were studied, how the project was designed and why, the results of the project, analysis of said results, and finally provides recommendations for further study.

### Introduction

This study proposed considering the question: "Can machines write?"<sup>1</sup> Machine-generated narrative and language provides many avenues for exploration and creative play. Thus, this study was launched with the goal of creating a science fiction short story from the point of view of an artificial intelligence (A. I.) narrating its observations about humans from information

---

<sup>1</sup> See Turing, A. M. "Computing Machinery and Intelligence".

gleaned from the Internet. The idea was that for a non-human (and for less Internet-savvy humans), trying to learn about humanity from the great seething chaos of the World Wide Web would be an experience not unlike a first-time viewing of Hieronymus Bosch's *The Garden of Earthly Delights*: a strange and surreal world full of meaning without context. And what better way to turn this on its head than to have an A. I. write the actual text itself? It would then become the very thing it was writing about.

In an effort to execute this project, a study in different forms of language and narrative generation was undertaken in order to gain familiarity with the techniques and theory, both literary and computational, that would be necessary to guide the development of a model capable of doing the above. A computer-generated short story and analysis of the story and project would be submitted as the end product of the study.

### Background

As a primer for this project, two Allison Parrish courses were blended to create a course that incorporated both language and narrative generation. The two courses were "Reading and Writing Electronic Text" and "Computational Approaches to Narrative". Both courses are taught by Parrish at NYU's Tisch School of the Arts and are available online on her website <http://www.decontextualize.com/>. The courses progressed through the recent history of language and narrative generation and included topics such as Markov chains, speech-to-text, linguistics, narratives, and narratology. The courses were augmented by the documentary *Coded Bias*, the book *You Look Like A Thing And I Love You*, and a variety of journal articles on recent research into natural language processing and narrative generation.

Originally this project was conceived with the idea of creating a model that would allow for the specification of a narrative structure. Within that overarching program, each section of the

narrative would be separated into its own unit, then generate text appropriate to that portion of a narrative. The output would then be joined together by the outer program that would then print the output to a text file. Because the building of this outer program was outside the scope of the course, an alternate methodology was deployed instead.

### Project

OpenAI's GPT-2 text generation model was chosen for its memory capacity, high-quality syntax, and ease of access. GPT-2's memory ability enables it to stay on topic for a longer stretch of text than many other methods; this helps it feel more like text created by an actual consciousness. Its syntax is clear and understandable for the most part; this also allows the text to feel like something created intentionally. These (and other) abilities make it a model of great interest in the A. I. research and data science communities; because it is so popular, it was relatively easy to find material about how to train one's own GPT-2 model.

The GPT-2 model used in this project is a copy of the "Train a GPT-2 Text-Generating Model w/ GPU" Google Colaboratory notebook created by BuzzFeed data scientist Max Woolf. This notebook (hereafter referred to as a "Colab notebook") is an instance of a Jupyter notebook on Google's cloud computing service. These notebooks are a web application that allows for the integration of text, code, and other media. Additionally, the Woolf Colab notebook grants access to a virtual machine and GPU that is ideal for this sort of work, at no cost.

In order to train the model to create a story on the desired topic, a curated selection of texts were sourced mainly from Project Gutenberg in a text file format. These texts focused on machines, automation, artificial intelligence, sapience, consciousness, and theory of mind. The majority were fiction; there were some philosophical texts as well (see Table 1 at the end of this section). These were chosen to increase the likelihood that the model would stay on topic and

produce a coherent text. The majority of texts chosen were in the public domain and easily accessible. Formatting was necessary in order to decrease the chance that the model would produce gibberish.

Each fictional narrative was broken down into the traditional three-act structure and tagged “[ACT 1]”, “[ACT 2]”, and “[ACT 3]”. This was easiest with the play, and more difficult the earlier a text was written. The files were formatted with “[BEGIN]” and “[END]” to delineate the start and end points of a story. The non-fiction texts were not broken down into a three-act structure and were instead given only the BEGIN/END tags. A set of five input texts was created to fine-tune the model. Each of the input texts was a compilation of each set of acts (thus a file for all Act 1 text, a file for all Act 2 text, etc.) and the last two texts were the entire nonfictional narrative corpus and the entire fictional narrative corpus.

The model was trained incrementally, starting with each “Act” input in order, then the nonfiction input, and finally the entire fiction corpus. The entire fictional corpus was last, to reinforce the narrative aspect of the resulting text and to counteract how long the essay input was<sup>2</sup>. Rather than start with a fresh model of GPT-2 for each input, the same model was reused with the idea that the resulting model would be better at creating a narrative. Finally, all of the outputs were combined and edited for comprehension and clarity. The output is in the “Results” section.

---

<sup>2</sup> It is unlikely that readers would enjoy a paragraphs-long monologue by Hegel.

**Table 1: Works and Genres**

Name	Author	Source	Genre
<i>The Shape of Things to Come</i>	H. G. Wells	Project Gutenberg Australia	Fiction
<i>Erewhon</i>	Samuel Butler	Project Gutenberg	Fiction
<i>Frankenstein</i> (1830)	Mary Shelley	Project Gutenberg	Fiction
<i>Rossum's Universal Robots (R. U. R.)</i>	Karel Capek	Project Gutenberg	Play
<i>Ex Machina</i>	Alex Garland	N/A	Screenplay
<i>The Grindwell Governing Machine</i>	Anonymous	Project Gutenberg Australia	Fiction
<i>I Have No Mouth, and I Must Scream</i>	Harlan Ellison	N/A	Fiction
<i>Utopia</i>	Thomas Moore	Project Gutenberg	Fiction
<i>The Tunnel Under the World</i>	Frederik Pohl	Project Gutenberg	Fiction
<i>On The Soul</i>	Aristotle (trans. J. A. Smith)	The Internet Classics Archive	Non-fiction
<i>The Critique of Pure Reason</i>	Immanuel Kant (trans. J. M. D. Meiklejohn)	Project Gutenberg	Non-fiction
<i>Philosophy of Mind</i>	Georg Wilhelm Friedrich Hegel	Project Gutenberg	Non-fiction
<i>A Discourse on Method</i>	René Descartes	Project Gutenberg	Non-fiction
<i>An Essay Concerning Humane Understanding, Vol I.</i>	John Locke	Project Gutenberg	Non-fiction
<i>An Essay Concerning Humane Understanding, Vol II.</i>	John Locke	Project Gutenberg	Non-fiction

## Results

[BEGIN]

This is a rush transcript. Copy may not be in its final form.

\*\*\*

SCENE:(INT. HOUSE/MAIN ROOM - DAY

The MAIN ROOM is a modernist-inspired space.

It has a minimalist look to it. All of the windows are closed.

Everything in the room was black.

Except for the wall.

The room was protected by glass walls.

Outside the walls there was a raised platform which faces the outside wall.

The Cent consists of a restaurant, library, and a garden. The front of the building is decked and faceless. In comparison with the front of a house, it is tattered, almost impassable. The front of the room has a propped-up exotica of brick and plaster. Behind the propped-up brick wall, where the vista refers, there is a corridor and a courtyard. The courtyard is covered in vines and lichen. Up to the roof, which rises towards the roof, there are no details unsparing.

The courtyard extends to a balcony on either side of the courtyard, which extends down to the sea level. Above it all is sunshine. The canteen stands at the top of the garden, just above the sea level. The fruit stand, which is just above the garden level and runs until it overflows, has a tin-roofed exterior, and is covered with vines. The front door of the canteen stands on the side of the cellar. Up to this point the windows of the front room had no windows, but the front door has been opened since the reaction of summer. In the middle of the room there is a warehouse with a bath and a kitchen area.

I remember thinking, my dear girl, that all these years had been meeting, that this was going to be a journey; that we would be alone, alone together, alone together, and one another were none the -

\*\*\*

EXT. GARDENS – NIGHT

The moon hangs in the sky.

\*\*\*

INT. HOUSE/OBSERVATION ROOM - DAY

The automated doorbell rings.

It comes from the glass.

It reveals NATHAN who has just entered. Tall, handsome, bearded, with blue eyes.

The doorbell rings again.

\*\*\*

EXT. GARDENS - DAY

The automated mower cuts grass as the long-tailed deer approach.

The mower tracks along the grass, and as NATHAN watches, the deer approach with each stride as the mower tracks along.

They follow in the footsteps of two deer. )

\*\*\*

SCENE: (EXT. GARDENS - DAY

- NATHAN stands on a ledge in the middle of the garden, facing upwards. He is holding a camera. He flicks the film between his fingers. Then clicks the link.



CALEB:

You okay?

(NATHAN smiles.)

NATHAN (CONT'D):

I think I've got a diagnosis. I really do. I woke up this morning in a hospital bed, and I think I'm in a coma.

CALEB:

... Are you sure?

NATHAN:

I don't think they've asked me any questions.

(NATHAN shrugs.)

NATHAN (CONT'D):

I'm not sure any of you are either. I'm not sure how they feel about me.

(CALEB looks at the picture.)

CALEB:

It's a picture of a girl.

NATHAN:

Guess I've been waiting for this.

CALEB:

Yeah.

\*\*\*

But we do not know exactly what the computer programmers did there; no one has ever been able to create a program that could generate a programmatically intelligible program, and no one has ever been able to program a programmatically intelligible manner of life.

But what if instead we had of a computer program, there were programs that were programs that were programs? Such algorithms as ‘These are the instructions that are to be given you if you wish to do these things:’ or ‘These are the instructions that will become yours:’ or ‘These are the words which you will write:’ . . . and you shall be a writer?’ What if instead this were the case?

\*\*\*

INT. HOUSE/CALEB’S BEDROOM – DAY

CALEB helps the android into the bedroom.

He takes a piece of the glass.

He looks at it.

Then he puts it back in its original position.

Then he puts his hand on her shoulder.

Then he starts moving.

The android turns away.

But before it can do so, a kind of robotic noise escapes from his ear buds.

It’s the whine of a robot computer.

The noise is so strong, so overpowering, that CALEB nearly drops.

But he doesn’t.

He turns.

On the sound screen, a LIVE COMPUTER MONITOR is playing.

It shows the AVA computer.

It displays the range of movement in microseconds.

Then it stops.

CALEB watches as the AVA moves across the room.

The monitor doesn't know where it is. It doesn't move, it doesn't know.

It stays where it is.

The sound stops.

For a moment, the screen freezes.

Then the AVA starts moving.

It stops.

CALEB watches.

It doesn't know where it is. It doesn't know when it is. The only thing it knows is that it is somewhere in the room.

And that's all there is.

It doesn't know when it will leave the room.

Neither it's entering the room, nor leaving their own.

It stays what it is.

CUT TO –

- the POV of AVA.

Where the camera is held.

The soft LED light on the side of the glass is illuminating a small door. On the far side of the door, we can see a single LED pin. That LED pin turns blue when the door is opened.

Then the door opens.

\*\*\*

### AVA'S BRIEFING

The evidence - the description of the physical evidence - has been entirely and entirely outside the scope of this entry.

We know that Ava is a synthetic android. We know that she has a full-blown consciousness, which is why we can talk to her. We also know that she is physically strong, as demonstrated by the way she locks eyes with AVA. But what we do know is that she is not. Her mind functions as HAL 9000's, and - as demonstrated by her facial expressions - she is not Turing complete.

\*\*\*

### EXT. GARDENS - DAY

CALEB:

Oh, my God.

NATHAN:

Excuse me. You mean, you thought--

CALEB:

You were predicting something specific?

NATHAN:

Maybe it was a forecast of the Great Leap Forward.

CALEB:

... I didn't know that you were predicting a revolution?

NATHAN:

I didn't know that it was so.

CALEB:

But you did something that no one expected.

NATHAN:

I did something that I never would have expected.

(NATHAN stands.)

NATHAN (CONT'D):

You. And the things we're planning with Ava.

\*\*\*

I am a programmer, and one of the things I love most about programming is having the freedom to experiment with new ideas. My favorite part of programming, at least for me, is the opportunity I have of exploring new fields at large. I get to experience the world while I am coding, and I am also allowed to ask, "Why did you only build Java for fun?" Sometimes I get so engrossed in the creation of new things that I forget that there really is a market for its functionality, and that there really are many other ways of doing things that do not involve manipulating and reusing the source code.

\*\*\*

EXT. GARDENS - DAY

- the growing of a rainbow in the spring sunshine.

INT. HOUSE/GLASS PLANT - DAY

- the house.

The view through the patio.

The view of the trees.

CALEB stares at the mountains beyond.

He reaches his decision.

\*\*\*

SCENE:(INT. HOUSE/CALEB'S BATHROOM - DAY

CALEB stands in the shower.

He puts on a brave face.

It takes him a few seconds to realise. He has both of his arms wrapped around himself.

He tries to cry. And he doesn't quite understand it.

He reaches for the kitchen knife.

Goes to the sink, where it's still wet.

He picks it up.

It's the sound of his broken voice.

It's the sound of the kitchen knife.

It's cut through the air.

He pulls the knife from the sink.

It's red.

He holds the knife out for another four fingers. It's bleeding.)

\*\*\*

I had no sooner entered into the room than I was absorbed in the force of the blast. The walls of the room were covered with leaves and were covered with a carpet of leaves, so that I could not see but the mouldings which formed the covers of the walls; and the only light which I

could see came from a great distance up the chimney of the cottage. I saw the four great windows of the cottage shut, and no one could pass in.

Then came the sound of voices which were distant and unnatural, but whose sound was so unlike the madness of my father that I could not bear to see them. It is the breakdown that has brought us to this place where the world, its laws, its schools, its doctors, its mothers, its brothers and cousins, is going through a phase of mental breakdown and insurrection. And at the heart of it all, like a cancer spreading and dividing, is the sad failure of dignity. It is here that we see the final death throes of the species.

\*\*\*

SCENE:( EXT. GARDENS - DAY

He walks along the lawn.

Then stops.

NATHAN has stopped.

NATHAN:

Sorry. I thought I caught you.

(Beat.)

NATHAN (CONT'D):

You caught me.

(NATHAN freezes.

NATHAN sees BETWEEN AVA and CALEB.

They've met.

And already know.

NATHAN looks at CALEB.

Then reaches out -

- to the cardiologist who is tending to NATHAN.

And touches his hand to the heart of his surviving hand.

This exertion causes the blood to cool, and relaxes the muscles in the wound. Then it stops.

NATHAN watches in stunned silence.)

\*\*\*

NATHAN (CONT'D):

Caleb.

(His body temperature remains warm.)

NATHAN (CONT'D):

You were a pleasure to have.

CALEB:

... Thank you.

(NATHAN laughs.)

NATHAN (CONT'D):

I'll take care of business.

\*\*\*

EXT. GARDENS - EVENING

They sit, together, on AVA'S arms, her legs spread, her stomach twisted, her head tilted. Her head turns to the side, as if looking at the tiny camera.



She watches.

AVA (CONT'D):

And it wasn't you who faced the world in the twentieth century. It was a collection of captive minds, each claiming to be different. And all of them were destined to one thing: they couldn't be good.

(Beat.)

AVA (CONT'D):

That they were destined to hell.

(Reveal AVA's bare chest. Her breathing is shallow. Her ribs are all broken. Her head is all broken. Her face is all broken. Her face is gone. Her face has no emotion. Her expression is lifeless.)

AVA:

Where did you send me?

(Reveal AVA's naked form. Her legs are tied up. Her head tilted up in a straight up up state.)

AVA (CONT'D):

You sent me to the machine.

(A beat on AVA. Her body is semi-conscious. Her head is tilted up. Her expression is lifeless.)

\*\*\*

I can think of no greater contrast in life than that which I noticed myself in, with the exception of the latter being that I was the only one in the world whom all things were equally equal to. I was everywhere in the world, in the air, in the house, in the garden, among other

things, where every man was to be equally free from competition and all other inequalities. I was everywhere free from the fear and violence of physical exchange; I was everywhere all things were to be as they were to be fit for them; and yet I was free from any fear or violence whatsoever. In all other respects I was as free as any other human being. It was thus that I became the arbiter of what was and was not possible. I saw what was and was not possible to achieve; I made the known laws of this world as it were.

\*\*\*

SCENE: EXT. GARDENS - MONDAY NIGHT

The moon is shining brightly through the mesh of clouds.

AVA:

Caleb, what's the meaning of that?

CALEB:

The meaning of life is to be happy.

AVA (cuts in):

What's the meaning of life?

CALEB:

The disappearance of pleasure.

\*\*\*

The evening was spent in silence. Nothing to say or do.

\*\*\*

it was man who would have

destroyed the world.

## Analysis

Four outputs were used in piecing together the results. The resulting piece was heavily edited and rearranged into a coherent narrative, and even then it relies heavily on the dreamlike nature of the generated prose. Although GPT-2 is superior to, say, a Markov chain in that it stayed mostly on topic, there were still instances in the total output that didn't make sense, or made sense in a way that would not work in a narrative. For example:

INT. HOUSE/CALEB'S BATHROOM - DAY

CALEB sucks on the water in a special bath towel.

A scene sure to provoke all of the wrong kinds of questions.

Some interesting features noted in the output included the fact that the only two scripts in the corpus (*R. U. R.* and *Ex Machina*) tended to feature characters from both stories in scenes and dialogue with each other:

HELENA:

And then you go and see the moon.

CALEB:

But I don't see it.

HELENA:

Why?

CALEB:

The only people who've seen it are the people who've died.

This is because an effort was made to format both scripts in the same way. However, the dialogues between characters tended to continue for quite some time, with characters asking “Why?” or “What?” and receiving the same or similar response from other characters. Repetition of phrases was a very common issue; where this was used by Ellison in “I Have No Mouth...” for dramatic effect, the same cannot be said of GPT-2.

Although this instance of GPT-2 was fine-tuned with a curated selection of inputs, the underlying training tended to come out too, as in the following:

Buyer receives a return envelope containing a signed copy of the software, copy of software, copy of electronics, and a label explaining why we accept these charges: your package will contain: a label explaining why we accept these charges, a signed copy of the software, and a label explaining why we offer to exchange your camera for its original (and only melted) glassware. We will exchange your Camera for a glassware made from high-temperature, nickel-ion-phosphate (KYC) steel.

This is assumed to be a result of the underlying training data as it is unlikely that Kant was discussing software or nickel-ion-phosphate (KYC) steel in his *Critique of Pure Reason*. This and repetition of words and phrases were the two biggest issues in the results.

### Conclusion

In the next iteration of the Bosch project, there will be further research and testing of narrative models. There are many promising models, such as a hierarchical model (Fan 2018), where the model generates a prompt and then text to flesh out the prompt; using plot graphs to draw inferences from stories, potentially useful for reinforcing themes (Purdy 2016); plan-and-write, a model that plans a storyline and then generates the text, closest to the original vision for the project (Yao 2019); and potentially even an iteration that would incorporate a human-and-

computer collaborative storytelling model (Nichols 2020). The next iteration of this project will likely experiment with one of the aforementioned models; the plan-and-write is the most likely candidate. Future iterations would feature a combination of models, potentially the plot graph, plan-and-write, and fine-grained sentiment for story endings (Luo 2019).

More important than any of the models used would be an increase in the quantity of texts chosen, the quality of the texts, and the diversity of inputs. Project Gutenberg featured heavily in this iteration of the project because the texts it provided were no cost, easily accessed, and already in a text file format. Even with these features, they still required hours of formatting to remove unusual characters, ensure standard use of parentheses and brackets, create uniform spacing, and in older texts reduce the amount of Latin involved so as to not confuse the English model. The types of writers most readily available on Project Gutenberg tend to be male, white, Western European, and dead. Increasing the input corpus not just in size but in degree of diversity is the first priority for this project. Ultimately, the answer to the question “Can machines write?” is the same for any human in a creative endeavor: the insights in the creative output are only as good as the creative input.

## Note

1. Turing, A. M. "Computing Machinery and Intelligence." *Mind: A Quarterly Review of Psychology and Philosophy*, LIX, no. 236, Oct. 1950, pp. 433–460.

## Works Cited

- Fan, Angela, et al. "Hierarchical Neural Story Generation." *ArXiv:2011.10208v1*, 13 May 2018, [arxiv.org/abs/1805.04833](https://arxiv.org/abs/1805.04833).
- Luo, Fuli, et al. "Learning to Control the Fine-Grained Sentiment for Story Ending Generation." *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 2019, pp. 6020–6026.
- Nichols, Eric, et al. "Collaborative Storytelling with Large-Scale Neural Language Models." *ArXiv:2011.10208v1*, 20 Nov. 2020, [arxiv.org/abs/2011.10208](https://arxiv.org/abs/2011.10208).
- Purdy, Christopher, and Mark O. Riedl. "Reading Between the Lines: Using Plot Graphs to Draw Inferences from Stories." *International Conference on Interactive Digital Storytelling*, 2016, pp. 197–298., doi:[https://doi.org/10.1007/978-3-319-48279-8\\_18](https://doi.org/10.1007/978-3-319-48279-8_18).
- Yao, Lili, et al. "Plan-and-Write: Towards Better Automatic Storytelling." *The Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19)*, 2019, pp. 7378–7385.