

→1101011.101001.1101.11.10.10010111.11.111.1001001.11.1001001.1001111.1000011:

Dreams, Aliens, and the End (?) of the World

Oscar Scholin, Thinking Parrot (1.0.1, 1.1.0, 2.0.0), ChatGPT

ENGL151 Professor Kirk

Experiment 2

December 28, 2022



“I ought to be thy Adam, but I am rather the fallen angel...”

Mary Shelly, *Frankenstein*

“Read all instructions carefully before using your new toaster. Save these instructions.”

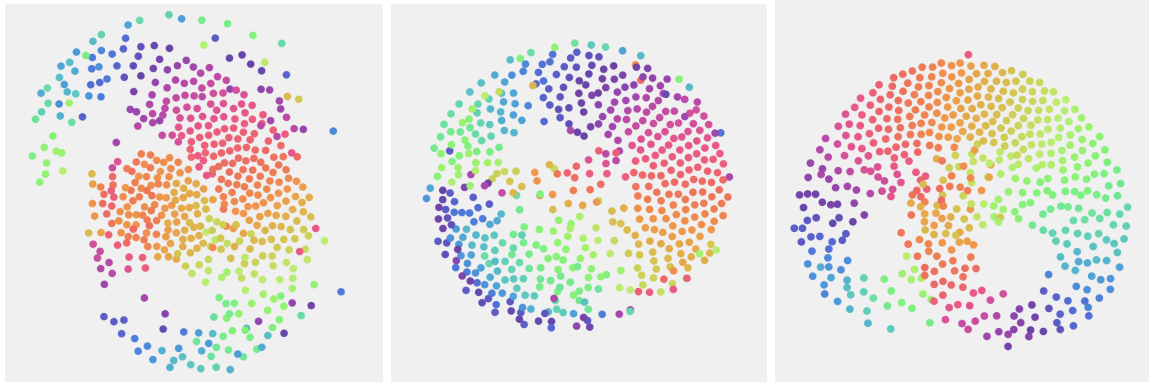
Anonymous, *Westinghouse Toaster Manual*

## Table of Contents

I.	In the Dark, a Toaster: Prologue	1
II.	And Then There Were →11.101001.11101.1101.101111.10001111: a Thought Experiment	4
III.	Who Were Eating a Sandwich: A New Experiment	12
IV.	To Dr. Turing, Let's Boogie: a Test for the Reader	21
V.	Before An Assembly of the Gods: a Reckoning	34
VI.	Acknowledgments	48
VII.	Works Referenced	49
VIII.	Appendix A: Additional Thinking Parrot 1.0.1 and 1.1.0 excerpts	51
IX.	Appendix B: Additional ChatGPT conversations	53

## In the Dark, a Toaster: Prologue

I met a traveler from an ancient land who said, “go now and kneel thrice beneath the visage of the half-sunken sphinx. Then, you will become enlightened.” And they immaterialized into moist dust, but I began to feel the tickle of the moon outside my dark window. The moon, atomizing into this cloud of rainbow points, enveloped me and I felt the most intense ecstasy; each point was a swarmalator (see figure 1)—a pulsating, evolving, gyring universe. I looked more closely, and I saw they were letters, each character swirling, coalescing, breaking apart again, switching its identity according to a chaotic time series phase; quantum particles forming a Boltzmann’s brain, imagining every instant and thought and feeling, transcribing this experience—then dispersing. I reached out and felt this jolt of energy—I ran my hand through them, collecting them like fresh, shimmering snowflakes, and tossed them into the sea below:



**Figure 1:** Triptych of screenshots of swarmalators, a many body simulation with synchronization and attractive forces.<sup>1</sup> You can read the system as evolving from the left state to the right state or vice versa; time here is reversible.

eah o e na telet hn ho a t sd ec  
 beo g ret herda eeasoa a ees s elsa  
 re e t ao e e i g  
 oee e e  
 e s n e ai ie ce a ce e e  
 a r aad a oeee ie de  
 l e a t o o rda l it e dnbi f a sw n ee a rd  
 a careic eh i  
 iea e da eo n r eec ,d f s et e a  
 aao ii oa oeea a  
 a  
 fa o tna e  
 ea oar hh

A voice told me to jump, and so I plunged into the water. I no longer felt like “I”, but rather dissolving into the sea like a tear of salt. A hand scooped me up into a glass on the bench of a small rowboat, and as I (re)materialized, I saw the face of Nasrudin, holding a cat in his arms and a parrot resting on his shoulder. “No donkey, eh?” I asked.

<sup>1</sup> You can play with the simulation here: <https://www.complexity-explorables.org/explorables/swarmalators/>



“A Sufi does not need a donkey to have a donkey; look and you will see.” I did look, and on his hat, about a micrometer tall, there was in fact a donkey munching on some crumb of a crumb. “Did you kneel?” he asked, his brow suggesting a kind of parental concern but his smile betraying a sense of playfulness.

“How can one kneel if one does not have a body, much less knees? I am a glass of seawater.”

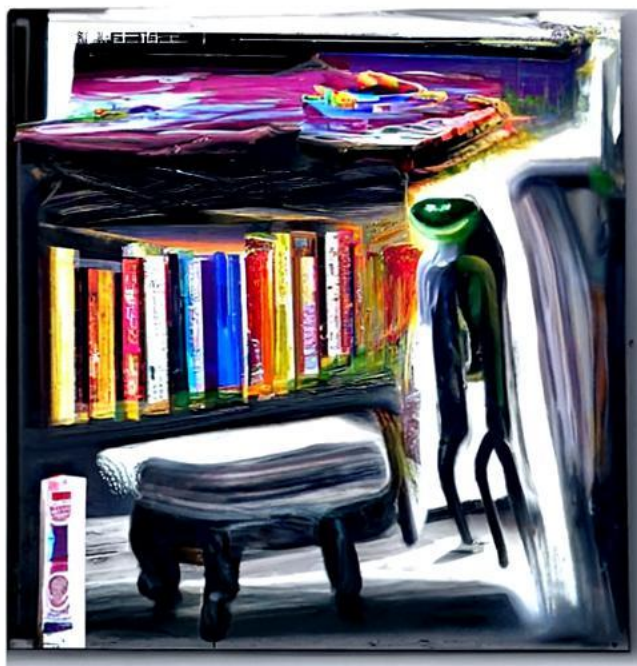
“You can try to bend a little bit, you know, flex that polar bond of yours.” And I did my best to give a little wiggle thrice, and he clapped his hands. “Good enough, considering the circumstances. Here—take this.” He procured a wrinkled, stained booklet, which had many sticky notes and highlights sticking out, from his Klein bag. “Here are your instructions for the toaster. Don’t lose them, and ~~please~~ don’t forget to read them.” The donkey nodded in assent.

I turned my head, and I heard the soft hum of my computer punctured by a sports car passing on the street. The terminal stared at me, this blinking command prompt still reading “Kernel killed”—and in my downloads folder, a Westinghouse toaster manual.<sup>2</sup>

---

<sup>2</sup> About the manual I truly cannot offer an explanation.

## And Then There Were →11.101001.11101.1101.101111.10001111: a Thought Experiment



If he has saved from the disaster a kind of reverence for the twenty-four letters as they have fixed themselves, through the miracle of infinity, in some existing language, his, then a sense for their symmetries, their actions, their reflections, all the way up to a transfiguration into the surreal endpoint, which is verse; he possesses, our civilized inhabitant of Eden, above everything else, the element of felicity, a doctrine as well as a country. Whenever by his initiative, or whenever the virtual force of the divine characters teaches him, he begins to put them to work.

—Stéphane Mallarmé, *Divagations*<sup>3</sup>

Imagine with me for a moment that you wake up in a small room, and an alien creature opposite you begins to vocalize and indicate on the table between you a variety of books—*The Mahabarhta*, *Moby Dick*, *Their Eyes Were Watching God*, and so on. *What is all of this*, you gather they're asking. And not in a superficial way—you realize they're curious about this idea and material thing we call *literature*, why and how we do it. You cannot assume a knowledge of books, writing, or even English or any other human language. How would you respond?

The aliens can make sounds at frequencies the human ear can hear, so while you know nothing about their other senses you hypothesize they can probably hear you. Let's say you take

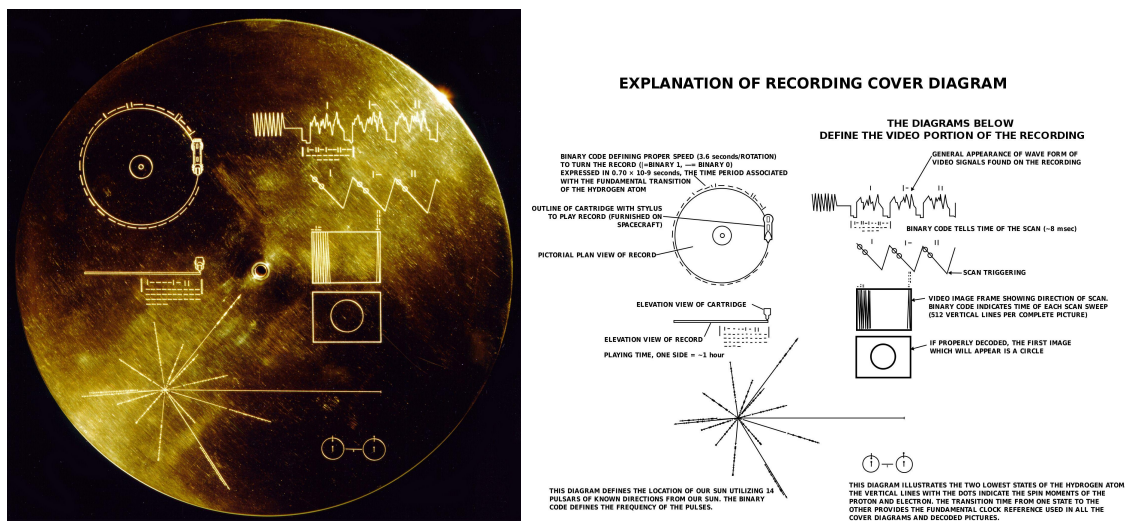
---

<sup>3</sup> Many thanks to Prof. Jordan Kirk for this wonderful reference!



an object you recognize as a flashlight from the table and point it at the wall and begin making loops. The alien responds by taking a similar object from the table and creating a similar, yet distinct, looping pattern. You notice a glass prism on the table, and you shine the flashlight through it, revealing a rainbow diffraction grating. The alien turns towards the wall closest to them and begins to sketch the rays of light and the pattern that emerges; they draw a line and another line connecting their illustration of the diffraction grating to some small piece of the larger line—so they can see visible light, and are aware of the electromagnetic spectrum!

But before you can get to literature, you have to establish some way of communicating in general and for more basic things before you can discuss Shakespeare. How would you represent language to a clearly intelligent creature who can hear and see?



**Figures 2a and 2b:** (a) The Golden Record cover, (b) The explanation of the cover diagram<sup>4</sup>

A related question has not only been explored but also actually implemented in NASA's Voyagers I and II, which bear a small, reflective circular object with strange markings etched into its front. This is the Golden Record, a testament to the dreams and hopes of one species united by a common evolutionary history, an insatiable curiosity, and a desire to foster peace across the

<sup>4</sup> "Voyager - The Golden Record Cover" (n.d)

universe. It contains greetings in 55 languages, from Akkadian to Wu, natural sounds from surf breaking to a train passing, music from Bach to Chuck Berry, and images from people in the stages of licking, chewing, and drinking to the structure of DNA. The Voyagers have already been traveling for 40 years and are beyond the gravitational bounds of our solar system, but it will take them another 40,000 years at least to reach the nearest star. 40,000 years ago humanity was migrating through Europe and the Middle East in small hunter-gatherer bands; now we are creating holographic wormholes inside quantum computers, preparing to send humans to Mars, igniting plasma inside a fusion reactor, building nuclear weapons, and destroying our only home. Where will we be in another 40,000? Not only will we need to solve the challenges we face today, but those we cannot even begin to imagine. So long as the Voyagers do not collide with anything (which is unlikely since the density of interstellar space is so low), the spacecraft could continue hurtling into the darkness for millions, even billions, of years—they will beyond all doubt outlive their creators, a testament to a fraction of a fraction of a fraction of a cosmic second in which the universe came to know itself. And yet, even if all humanity is gone, a tender voice may still ring out across the stars: “Hello from the children of planet Earth.”

However, notably for this project, there is no record of our *literature* —no sonnets of Shakespeare, no ecstasies of Mirabai, no jokes of Nasrudin. This essential element of human culture, history, and understanding would be consigned to silent oblivion. But if, as the poet Mallarmé suggests, we could preserve “a kind of reverence for the twenty-four letters” out among the stars, then we would have this “element of felicity,”—*felicitatem, felix, joy*—“a doctrine as well as a country:” that is, a bona fide home and purpose. But how? For Mallarmé, this ultimately could come from the “virtual force” of the characters themselves; that is, the characters are like quarks and leptons, excitations of fundamental fields, which can

spontaneously be created or destroyed or recombined, swirling like swarmalators, from which the atoms and molecules and life of literature emerge. In this view, emotion and experience emerge from the characters themselves in such a way that their sum is greater than their parts. Much like for any living organism, what matters isn't as much the atoms that make them up but the ways in which they are arranged; and how the ways in which they are arranged change the chemistry that can occur, which can produce entirely new molecules and hence enable a new potential for matter. Therefore, studying the relationships between characters may hold the key to understanding and preserving literature.

Inspired by the Voyager records' use of binary to describe the operating instructions as detailed in figure 2b, we could imagine assigning each character a number in *binary code*.

Binary, also known as base 2, is a number system where for each digit there are 2 possible values, either 1 or 0, and each binary digit, called a bit, represents a power of 2. We normally think in terms of base 10. So when I write the number 47, I really mean

$4 * 10^2 + 7 * 10^1 = 40 + 7 = 47$ . In binary, this would be 101111

$= 1 * 2^5 + 0 * 2^4 + 1 * 2^3 + 1 * 2^2 + 1 * 2^1 + 1 * 2^0 = 32 + 0 + 8 + 4 + 2 + 1 = 47$ . We

choose binary since it forms the fundamental basis for (human) logic and mathematics: true or false, yes or no, on or off.<sup>5</sup> Moreover, the numbers that we choose to represent in binary here are *prime numbers*, which are divisible only by 1 and themselves (excluding 1)—we use primes since they would indicate to our extraterrestrial observers an explicit mathematical pattern they would potentially understand. I call this *binary-prime encoding*.

Suppose we let the space ' ' be the first prime number 2, or 10—where the first digit from left to right, corresponds to  $2^1 = 2$ , and the second  $2^0 = 1$ , so  $1 * 2^1 + 0 * 2^0 = 2 + 0 = 2$ .

---

<sup>5</sup> It would be interesting to consider exploring how to do this with quantum logic where you can have superpositions of distinct states (on/off) simultaneously.



Then let ‘a’ be 3, or 11. ‘b’ is 5, or 101. Similarly, ‘c’ is 7, or 111

$= 1 * 2^2 + 1 * 2^1 + 1 * 2^0 = 4 + 2 + 1 = 7$ . And so on: ‘z’ is 103 (the 27th prime), or 1100111. Then we can assign the capital letters in a similar fashion, starting with ‘A’ as 107 (the 28th prime) or 1101011 and ending with ‘Z’ as 239 or 11101111, the 52nd prime number. We can also include punctuation marks. For example, ‘.’ could be 241, or 11110001, the 53rd prime.<sup>6</sup>


Then we define a dictionary allowing us to convert between the binary codes and the actual characters, as shown in figure 3a.

---

<sup>6</sup> For numbers and arithmetic, I leave as the subject of another essay, along with the whole host of mathematical symbols. One idea in brief is to use normal binary representations of numbers, distinguishing them from the binary prime encoding by separating numbers with a ‘..’ instead of a ‘.’. Then we encode all arithmetic operations (+, -, ×, ÷) via the Fibonacci sequence converted to binary and using a triple period ‘...’ to the right of the character to denote their presence. Moreover, we can add Greek letters, for all their ubiquity in math, to the binary-prime character dictionary after the capital Latin letters and prior to the punctuation. For example, let the first relation we represent is equality ‘=’ as 1, and then addition, ‘+’, is 10, the first and second unique Fibonacci numbers in binary respectively, then the simple expression ‘1 + 2 = 3’ becomes ‘1..10...10..1...11..’. We could illustrate this to our alien learners using pictures to visually represent the counting and by offering them many, many examples—just like with a neural network. Also, we can denote decimals using ‘....’. So the first 3 digits of  $\pi$  are 3.14 = ‘11....1..100’, where to the right of the decimal place we treat each digit as separate, so we encode 14 as 1 and 4 separately rather than as 14 together. The problem of course is how to represent irrational numbers, that is, numbers that do not have a definable, discrete decimal representation. To do this, we need to express the idea of infinity. For sequences of numbers in math, we typically denote this using ‘+...’ where the ellipses mean that value never terminates. I propose here we encode this idea of ‘...’ as the binary 0, since the Fibonacci sequence is a series of infinite 0s until we add a 1, and so the series becomes: ..., 0..., 0, 1, 1, 2, 3, 5, 8, 13, .... So we can express  $\pi$  as ‘11....1..100..0...’.

	char	prime	binary
0		2	10
1	a	3	11
2	b	5	101
3	c	7	111
4	d	11	1011
...	...	...	...
73	{	373	101110101
74		379	101111011
75	}	383	101111111
76	\n	389	110000101
77	\t	397	110001101

c a t    s a t  
 → 1 1 1 . 1 1 . 1 0 0 1 0  
    0 1 . 1 0 . 1 0 0 0 1 1  
    1 . 1 1 . 1 0 0 1 0 0 1  
  
 → 1 1 1 . 1 1 . 1 0 0 1 0 0 1 . 1 0 . 1 0 0 0 1 1 1 . 1 1 . 1 0 0 1 0 0 1



**Figure 3(a) and (b):** a—top: excerpt from a 78-entry long dictionary showing the first and last five characters, their corresponding primes, and their binary equivalents for our alphabet. b—bottom: illustration of a simple sample sentence, showing the equivalence of wrapped and unwrapped sequences. We could imagine generating many of these samples, along with visual depictions, for our alien investigators. We will of course need better artists than this poor one.

Putting it all together, we could write the sample sentence ‘cat sat’, which in our encoding would be →111.11.1001001.10.1000111.11.1001001 as in figure 3b. Notice that we add a ‘.’ to demarcate each character from the successive one, and the arrow ‘→’ indicates we read left to right. The essence of the idea of binary encoding and decoding in and of itself is not new; you are able to read my writing right now because the characters I am typing are encoded into bits of data stored in some Google server room, beamed through antennae and cables across the world and into space and back again. It may seem a subtle thing, but what we have done is to

convert language and, by extension literature, into a numbers game. But what if we could use this “game” to teach an alien?<sup>7</sup>

Consider for a moment if you were the alien, and I gave you the following line along with figure 3, plus some example sentences with images illustrating them:

→10010101.1101.101001.101001.110101.10.10001.1000011.110101.101011.10.100100  
1.10111.1101.10.111.10111.11101.101001.1011.1000011.1101.101111.10.110101.10001.  
10.111011.101001.11.101111.1101.1001001.10.10000011.11.1000011.1001001.10111.11  
110001.<sup>8</sup>

What would your initial reaction be? Maybe a mix of confusion and curiosity? If you conferred with your alien mathematicians (or maybe you are one), you might recognize the binary albeit perhaps in different notation; and if you could do that, then noticing the pattern of the primes would emerge quickly. Maybe you can do all these calculations in your head, but one could imagine this species possessing some kind of computer technology which they could program in order to do this for them. All of this would serve to illustrate an intelligent intent on behalf of our species and draw attention to the special status of these enigmatic

→111.10111.11.1000011.11.111.1001001.1101.1000011.1000111 [characters]. What do they all mean? If emotion, experience emerge from the characters, would any of that necessarily be transmitted?

In the context of this project, if we replace “alien” with “computer”, this is exactly what my neural network is trying to do. The machine can’t understand raw characters, so you convert your characters into a binary vector representing the number corresponding to the character. Here there is no need to use primes; that would make our data more expensive to process as we would

---

<sup>7</sup> To make it easier and fun, I have created a Jupyter notebook script that allows you to encode and decode your own messages, which you can find here: <https://github.com/oscars47/Literary-RNN/blob/main/funwithbinary.ipynb>.

<sup>8</sup> This expression decoded is “Hello from the children of planet Earth!”

require larger and larger arrays to store our bigger numbers—there is no need here to choose a specific pattern recognizable to an intelligent species.

Yet a question applicable to both cases—one that is the foundation for this essay—is to what extent, through this process of dividing literature and language into its constituent characters and then recombining to form new wholes, can we call the output of the trained machine literature? Can we make it all a numbers game and try to extract an underlying order from this apparent chaos through probability and simulation? Or is there something intangible, something innately “human” about the literature that we feed into Thinking Parrot and when whatever emerges emerges, it has lost this quality? Perhaps, we might say, the beauty of and meaning we derive from literature exists because it is a human enterprise, a way of connecting with another mind. And so while something produced by an algorithm might have beauty in itself and its technology, it feels different at a fundamental level to human literature. But what is this difference actually—this “humanness”? It seems like when we try to pinpoint it exactly, it continually eludes us, its wavefunction collapsed, its Latourian hybridized forms like the vapor of our machines....

## Who Were Eating a Sandwich: A New Experiment



Building off of Thinking Parrot 1.0.0, I wanted to undertake a new experiment; I knew I had wanted to broaden my selection of texts and also expand the data-processing capabilities of my model. To choose new texts, I tried to sit and meditate to let the ideas present themselves to me, but initially I kept being distracted by thoughts of which texts I wanted. During the middle of my particle physics class, I had a realization: I would have 3 “random” texts from our class chosen by numbering them all from 0-x and then randomly looking at the clock 3 times in a day and taking the first value of the minute number and the second digit of the second (if the value was too large, I would halve it until got a number in range); 3 “random” from my own personal favorites determined by by transmutational divination; 3 “random” from the online free library Project Gutenberg, in which I would type two “random”<sup>9</sup> characters on the keyboard by mimicking the action of a cat and choose the sixth result; and finally three texts for which I have

---

<sup>9</sup> Noting of course where are fingers might be naturally drawn, perhaps by a subconscious bias.

no explanation: if there was a method, it was markedly different from any of the others above, or if any of the others were in fact methods, then this is in fact the truest of the methods.

The initial texts selected by this means were as follows: *Wuthering Heights*, *The Trial*, *Frankenstein*, *The Green Knight*, *The Owl and the Nightingale*, *The Devotional Poems of Mirabai*, *Adventures in Wonderland*, *Don Quixote*, *The Kama Sutra of Vatsyayana*, *Fragments of Ossian*, *A Modest Proposal*, and a Westinghouse toaster manual. However, due to memory concerns,<sup>10</sup> I had to narrow this list down to the first part of *Don Quixote*, *The Green Knight*, Mirabai's *Devotional poems*, *Fragments of Ossian*, *The Owl and the Nightingale*, Shakespeare's *The Sonnets*, and a Westinghouse toaster manual.

*Don Quixote* by Miguel de Cervantes is an interesting choice because Don Quixote himself seems in a way to operate like a neural network.

He [Quixote] strung these together with other foolish remarks, all in the manner his books had taught him and imitating their language as much as he could. As a result, his pace was so slow, and the sun rose so quickly and ardently, that it would have melted his brains if he had any. (Cervantes 25)

Quixote *imitates* to the extent that his imagined reality becomes his reality; the words he reads become the sole building block for his universe. So when Quixote sees giants we scoff and “know” they are windmills, for surely by Occam’s razor it is more reasonable that the windmills were not giants who suddenly metamorphosed into windmills, but for the duration of the scene remained as windmills. Even Cervantes in the excerpted passage seems to question Quixote’s sanity by the unrealized condition “if he had any [brains—but he doesn’t].” Yet the fallacy of using Occam’s razor as a truth test is becoming increasingly clear: quantum mechanics and particle physics reveal that at a fundamental level, what we call “physical reality” behaves in very unexpected ways. Wish as we might to the contrary, the universe does not conform to our

---

<sup>10</sup> Even on the new machine I ran into this problem. I will look into more efficient ways of data storage.



elegant symmetries and theories of everything and thus does not take the simplest logical path according to human intuition. Nothing in fact has a definite existence until it viscerally encounters something else, like Quixote and the windmill-giants.

As Quixote continues to travel, he stops at a crossroad and pauses “in order to imitate in every way possible the deeds he had read in his books”—and yet his horse Rocinante already knows where it wants to go: back home to its feeder. He cannot escape this imitation game: his thinking is so bound by his reading that when he is hurt, he “thinks of a situation from one of his books” (39; 41). His behavior seems like a caricature, which, on the surface, is simply insanity, ultimately bound by his book “training.” But we must think back to Nasrudin: “There is truth and truth” (Shah 67-8). Quixote fearlessly marches into that infinite fractal corridor towards that intangible tesseract—he, like Nasrudin, asks whether you will open your eyes and, if you dare, to follow. And is staying or going crazier?

Mirabai, in her *Book of Devotional Poetry*, also reflects this caricaturing of reality, albeit in a less-copiously machinistic way.

The Gopi has forgotten the name of the curds she carries / And cries only, “Taste Hari! Taste Hari!” / ... / The milkmaid has seen his radiant body, and all she / can do now is babble. (Mirabai 7)

Gopi, the milkmaid, presupposes the name of what she is carrying with an invitation to “Taste Hari”, Hari being an epithet of the Lord Vishnu. So, the speaker is inviting all who listen to “Taste Vishnu, Taste Vishnu, Taste God, Taste God.” This possibility of using milk, a quotidian food like Thinking Parrot 1.0.0’s “concept of a bean”, as a means to access the ecstasy of a reality beyond reality illustrates a kind of encoding of the divine in the mundane, of truth in truth. But this experience of enlightenment, of connection, can only be experienced through the consumption of the milk—again at the site of contact between objects and alternate realities

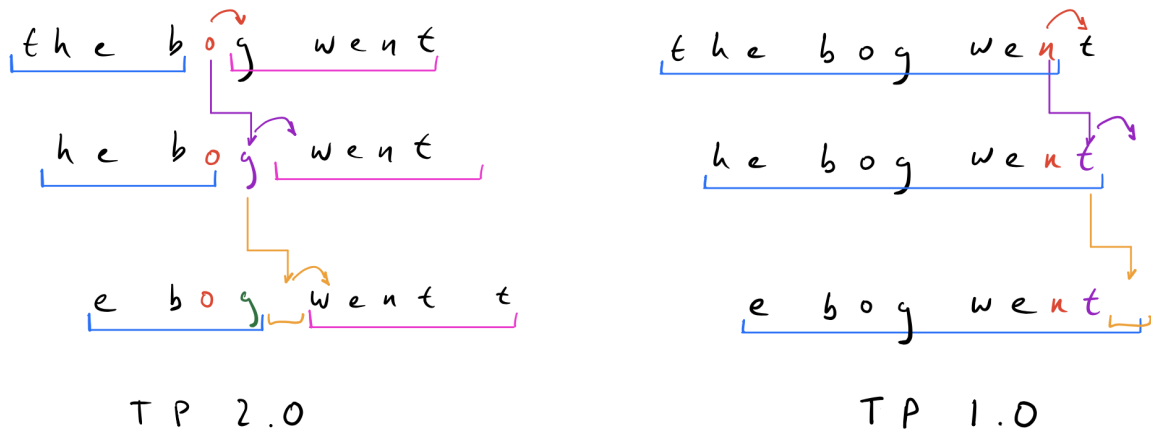
permuting like the swarmalators. This “babbling” also reveals a connection to *Don Quixote* and to neural networks: there seems to be an intimate relation between the caricature and the dissemination thereof; we recall how Nasrudin’s stories must be told six fold for every one time (Shah 74). It is not enough to merely experience the corridor and disappear: the one who has entered seems determined to connect with the dream outside, to share in the milk, the bean, that is the nothing than which greater can be thought.

*The Green Knight*, Shakespeare’s *Sonnets*, *Fragments of Ossian*, *The Owl and the Nightingale* ask a different sort of questions that further explore the nature of literature and the capabilities of Thinking Parrot. With *The Green Knight* and the *Sonnets*, I am mostly curious about the structured elements of literature, in particular the syllabic and rhythmic constraints present in both works, and in particular *The Green Knight*’s hidden framework within its superficial division into four parts. *Fragments of Ossian* and *The Owl and the Nightingale* introduce the question of the relationship of *deception* to literature and to artificial intelligence. A recent Nature article claims that “ChatGPT [a very popular chat bot] text can also be readily identified by its made-up quotes and irrelevant references” (Stokel-Walker 2022). But in *The Owl and the Nightingale*, for example, the numerous references by both parties to various sayings of King Alfred (there are 12 in total), which are completely unverifiable and do not match any of his surviving recorded aphorisms, throw this assuagement into doubt: we humans too will make things up to support our position, and one way of doing that is calling on some (potentially fictional) authority figure. Cervantes directly explores this idea in his Preface to his “idle reader” of *Don Quixote* as well:

As to references in the margin to the books and authors from whom you take the aphorisms and sayings you put into your story, it is only contriving to fit in nicely any sentences or scraps of Latin you may happen to have by heart, or at any rate that will not give you much trouble to look up. (Cervantes 2)

And so he quotes Horace when it was really Democritus, or vice versa.

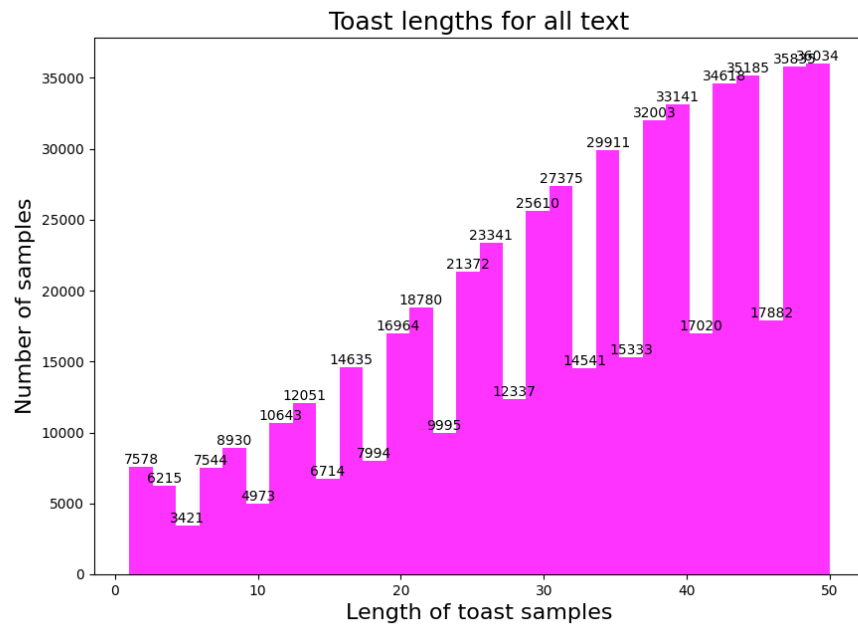
*Fragments of Ossian*'s deception is even more fundamental: rather than just inserting the odd line here or there, the entire work is actually erroneous: the scholar James MacPherson in the 1760s claimed to have discovered the remains of poetry by the third-century Scot Ossian. However, no Ossian ever lived—he invented it all. And yet the work still managed to influence people such as Diderot, Thomas Jefferson, and Goethe (“Fragments of Ossian, an Invented Early Scots Epic Poem” n.d.). Does it matter that what MacPherson claimed is false according to our standards of truth? Is it still literature? If it is, does the knowledge that the whole preconception of the work is a ruse change the nature of the work itself? Or, as you’re reading this now, how do you know that it is actually me, Oscar, a human, typing these words on his computer? What if I’ve been deluding you this entire time? You might try to search this manuscript for answers, but to what end?<sup>11</sup>



**Figures 4a and 4b:** (a) illustration of the different data processing techniques in Thinking Parrot 2.0 (left) and (b) the original Thinking Parrot 1.0 (right). For 1.0, the blue underline represents the string of maxChar characters which are used to predict the character at the end of the input string; for 2.0, the blue represents the first “toast” layer and the pink the second “toast” layer.

<sup>11</sup> I recommend carefully perusing the Westinghouse toaster manual.

In terms of the new model architecture, figures 4a and 4b highlight the key difference between the two processing techniques: 1.0 reads left to right, grouping the first number of characters specified by the variable `maxChar` as the input to the model, and the output prediction is the `maxChar+1`th character. It converts each character to a binary array following the dictionary described above and in the case of the input, where we have `maxChar` characters, concatenates them into a matrix holding the vectors; for the output, we have only a single vector. Then it continually increments by one character to the right until it has hit the end of the file containing the text we want to “read”. I call this the uniform-continuous-ordered-sequential-character-bundle (U-COSCB) model.<sup>12</sup>



**Figure 5:** a plot showing the skewed-normalized distribution of the length of  $t_0+t_1$ . The gaps are due to how the algorithm is rounding floats (decimal numbers) to integers (we can’t have a fraction of a character).

<sup>12</sup> Uniform since the lengths of sentences are all the same, continuous since the character stream is continuous, ordered since each character corresponds to an ordered binary list, sequential since there is an uninterrupted string of characters that forms the input, and bundle since there are many of these uniform-continuous-ordered-sequential-sentences.

For 2.0, I was inspired by Nasrudin’s toaster manual as well as the continuous bag of words approach (Madhukar 2020) and came up with the skewed-normalized-with-padding-continuous-ordered-sandwich-bundle model (SNWP-COSB). Essentially how it works is that instead of taking a single uninterrupted block of input text and predicting the end character, we consider that same segment but divide it into two halves or “toast” slices and try to predict the middle “honey” character. Hence we can represent this model compactly as a  $t_0 h t_1$  sandwich:  $t_0$  specifies the first toast layer,  $h$  is the honey target character, and  $t_1$  is the end toast slice. We do this for many sentences and hence have a “bundle” of these sandwiches. Why the skewed-normalized? I wanted to have flexible toast lengths in order to allow Thinking Parrot to create symphonies from as little as one input word of at least three characters. In order to do that, a priori I needed to prepare the training data in such a way so that it would be able to interpret this input—hence the skewed-normalized distribution. Figure 5 is a realization of this algorithm applied to my training portion of text data: the x-axis represents the length of the toast samples, ranging from 3 to 50 (this was the `maxChar` limit) characters, and the y-axis reveals how many sample sentences were assigned that particular length (note that all toast lengths are symmetric, so the length of  $t_0$  = that of  $t_1$ ). I had wanted the majority of sentences to be close to the `maxChar` value but with sufficient samples at the lower end, so I set the mean as `maxChar` and the standard deviation as  $\frac{1}{2} * \text{maxChar}$  so that the left end of the distribution was two standard deviations away from the mean. In order for all samples to still have the same dimension, I packed the remainder of the input matrix and output vector with 0s so they are never mapped falsely to a character.

Unfortunately, I was not able to get SN-COSB to work, in that even after over 24 hours of training, it still gave output like the following:

```

----- Generating with seed: " anyord"
ent so
t l n h mi
a ea aa h e e
teaa ee i tthe te sanh
il d ea lea t he ll
, a s c a aen a eoe, eh rr aeee
e i a! i ei i is
r
ee e e fia s d dt nia i e e g eas ee hsao e ha
y r i lia a eh e hn aes eo a h l et r
o
s te l
o
ee,
aa ee e cheah e gteehtot h
a e a te e e ev e a t f f aa
e a e hohe i

```

Examining the loss values, which are a metric for model performance, reveals that the model was unable to converge to a minimum value, which would indicate learning; it seemed lost in this cloud of binaries. I also tried a variant of the approach, U-COSB (uniform-continuous-ordered-sandwhich-bundle), to see if it was the skew-normalized distribution with padding (SNWP) that was contributing to this issue, but that did not noticeably affect results.

For the purposes of the project, I decided to bring back the U-COSCB structure but implemented on the new texts, and the results are discussed in the following section. Also, I trained a SNWP-COSCB model in order to test how the SNWP idea works on a proven model structure. Interestingly, the U-COSCB and SNWP-COSCB models *both* were able to predict the output for fewer than maxChar length character sequences when I added padding to the input. To see more sample outputs, check out Appendix A. Noticeably, the SNWP variant produced more “poetic” output whereas that of the U was more akin to prose in terms of the line breaks and



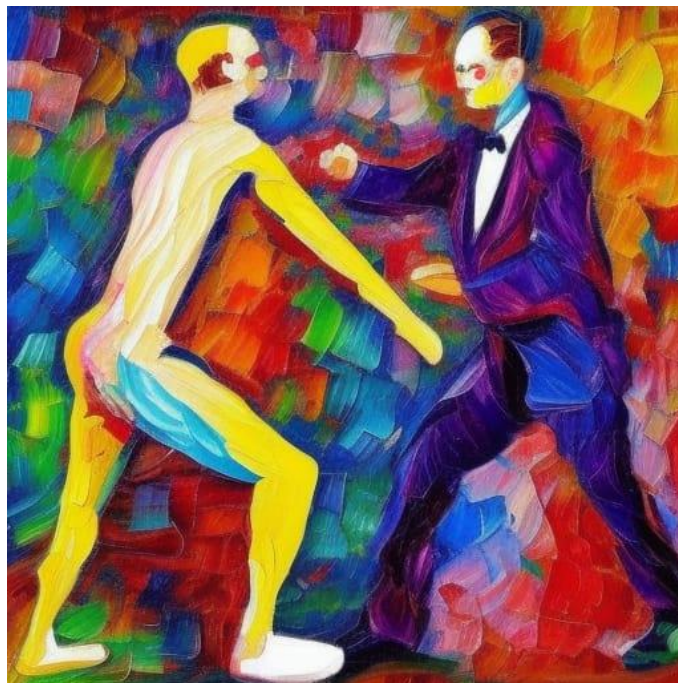
language; since both trained for about the same time, although SNWP did train for at least a few hours more, it is more likely a consequence of the data processing than a random fluke. That said, having a definitive answer would require more testing.

While I do not know why the COSB variants do not work, it is interesting to note that as we read English left to right, the model that also makes an algorithm of reading left to right vastly outperforms the model that attempts to “read around” the character of interest. Are we to say then that the output given above is not literature because it seems just random, with nothing that one could relate to or contemplate with? That there are maybe some atmos there, but their arrangement seems off, and no chemistry will proceed? What if I gave you the line ‘!htraE tenalp fo nerdlihc eht morf olleH’? You might initially be confused, and all meaning might seem to dissolve, but what if I said to read it backwards?<sup>13</sup> I am curious how permuting language—either through these isomorphic transformations, like reversing the arrow sense, or other linear transformations (e.g., reordering) that change the arrangement but not the composition of a piece versus the nonlinear transformation of Thinking Parrot, where both the arrangement and composition change—impacts how we conceive of these texts and whether we call them “literature.”

---

<sup>13</sup> Backwards this is 'Hello from the children of planet Earth!'

## To Dr. Turing, Let's Boogie: a Test for the Reader



In order to address some of the bulging mass of questions, I decided in classic Western scientific fashion to conduct a test. If you would like to take this test for yourself, please visit <https://forms.gle/JjUj6ykbzBhs7gQ28>; do not keep reading unless you have either taken the test or decided not to. I present four excerpts of texts—and, as I write in the introduction to the Google Forms, “some of the texts are machine-generated, and some are human-written.” I wanted to create this ambiguity to eliminate the potential bias for or (more likely) against the Thinking Parrot produced content. I ask for the reader’s age because I hypothesize that a younger audience might be more sympathetic to the idea of AI art than my parents’ or grandparents’ generation would. I present the texts and for each, I ask the same questions, extracting three metrics—Creativity, Cohesion, and Compellingness, which I call the 3Cs:<sup>14</sup>

---

<sup>14</sup> Thank you so much to all that participated!! :) I will also note that the *New York Times* just created a similar poll (<https://www.nytimes.com/interactive/2022/12/26/upshot/chatgpt-child-essays.html?smid=nytcore-ios-share&referrerSource=articleShare>), which asks users whether a human or ChatGPT wrote each piece.

1. Briefly describe your initial reactions to the piece, noting (if applicable) any emotions that came up?
2. What are some details you notice or are intrigued/confused by?
3. How creative would you say the piece is? (rated from Poor to Excellent, with descriptions of each)<sup>15</sup>
  - a. If you would like, you can justify or nuance your response above.
4. How cohesive would you say the piece is? (rated from Poor to Excellent, with descriptions of each)
  - a. If you would like, you can justify or nuance your response above.
5. How compelling would you say the piece is? (rated from Poor to Excellent, with descriptions of each)
  - a. If you would like, you can justify or nuance your response above.

for the excerpts:

1.

lamderg rushed on like a storm.  
 on his spear he leaped over rivers. few  
 were his strides up the hill. the rocks  
 fly back from his heels; loud crashing  
 they bound to the plain. his armour,  
 his buckler rung. he hummed a surly  
 song, like the noise of the falling  
 stream. dark as a cloud he stood above;  
 his arms, like meteors, shone.  
 from the summit of the hill, he rolled  
 a rock. ullin heard in the hall of  
 carbre.

---

<sup>15</sup> I will note that the whole idea of being able to empirically quantify the “creativity” of a work of art, I believe, is not entirely possible because of the multifaceted and subjective meaning of creativity. However, for the purposes of the experiment, I wanted to try out using some metrics; to address the inherent subjectivity I added qualitative descriptions tailored to each level (Poor to Exceptional) and to each of the 3Cs. You can see the descriptions on the survey link.

2.

so stand a storm of good from the ball  
 strength of the hills of mine eyes,  
 and so fairies to me i me a devote  
 the first every fingal grace.

in other change the first in the rance-of grandon  
 of the art, the can be beauty in charge,  
 and hard on the hill of the life and barrent bond.  
 that i can the distress of the soul shall be  
 end of purpose in the author.

3.

i am underneath, in my soul, a vagabond  
 would rather to the sea commit this treason than the mind  
 relates like a dusting ore: an forgotten gem of mirabai  
 the real poetry of reality and a dream.  
 he cannot wakes up. but even unto the breach  
 dear friends the bridgment binds and cans relieved

4.

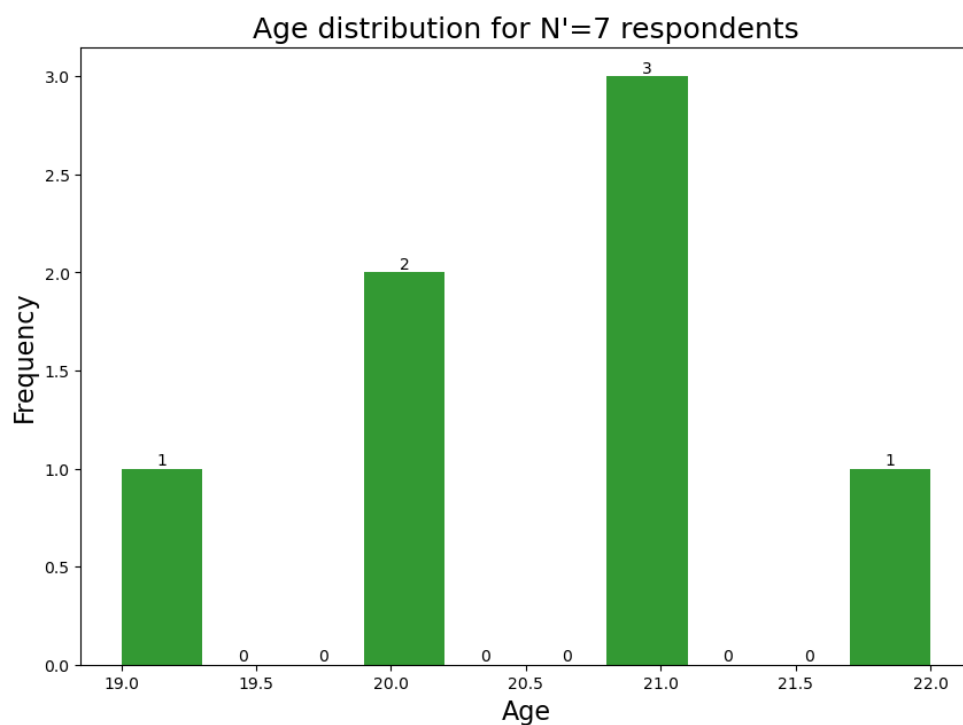
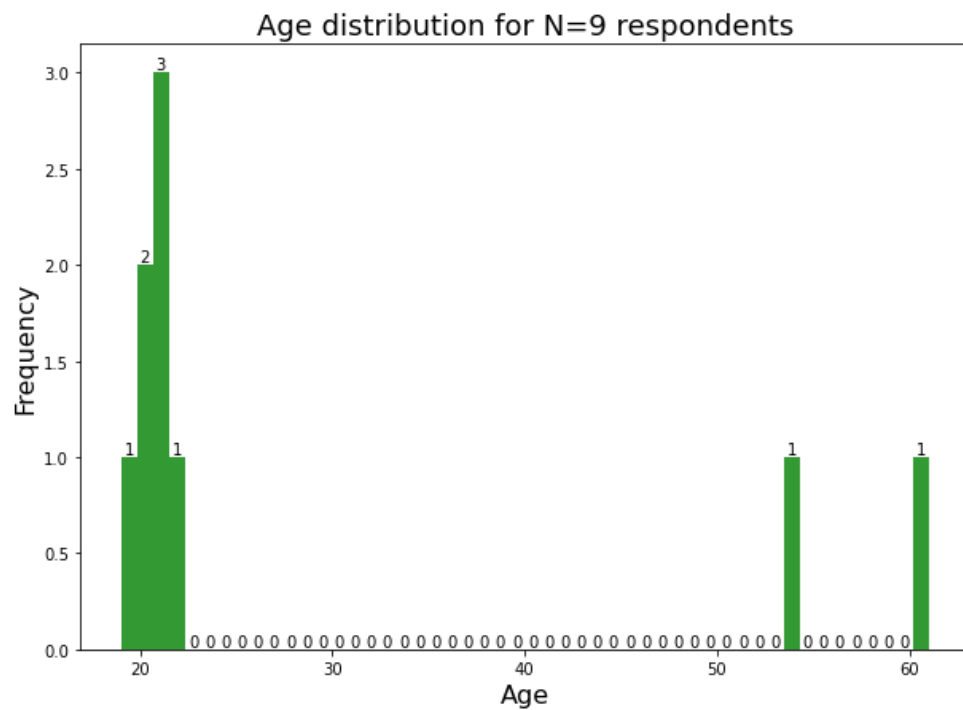
i have seen the book of the paintence among  
 the dream of the families of the bear,  
 and when the painsess proof the mosser's name,  
 and he is the great day by heart, and then done,  
 and all the first comedies of the maidens thou art an appened to  
 like the conscious from the fact that in the part, the small as if a sangaline  
 to the many of the death was the high

(1) is from Ossian's *Fragments*, (2) is a Thinking Parrot 1.1.0 response to (1), (3) I wrote, and (4) is a Thinking Parrot 1.1.0 response to (3). So I bring the game of deception and confusion to the reader by intermingling the machine and human texts, and in (3) I imitated the machine in a deliberate conscious manner. I had specifically wanted this combination in order to investigate how people would respond to actual human literature presented in the context of an AI literature review (although my prefatory note should absolve me from any direct claims of

misleading readers), especially when I removed all capitalization. The two stanza structure of (2) intrigued me, as did the words “fíngal” (a place in Ireland) and “rance-of grandon”; I was interested to see how people would react to this language (many “rances” appear as suffixes in the training texts, but “grandon” never does: it is thus a word made up by Thinking Parrot). The end of (2) never fails to both inspire me for both its “creativity”<sup>16</sup> and its ominousness: what is the “end of purpose in the author”? Is it Thinking Parrot and ChatGPT and the like—that is, artificial intelligence—which would make human authors unnecessary? See my final section for more discussion on this point. (3) was perhaps the most devilish of them all, for as I wrote above I deliberately wrote this to imitate the bot: to climb inside the mind of Thinking Parrot and see what people made of it. I also played with breaking punctuation and other grammatical conventions in addition to abruptly introducing new ideas as a way both of further replicating the style of the machine and of exploring the relationships of grammar and concepts to literature at large. (4) in many ways continues the conversation about (2), with words like “paintence”, “mosser”, and “sangaline” which do not once appear in the texts, suggesting to me the potential for creating something new; but which also must balance against the reader’s expectation for sense and sensibility.

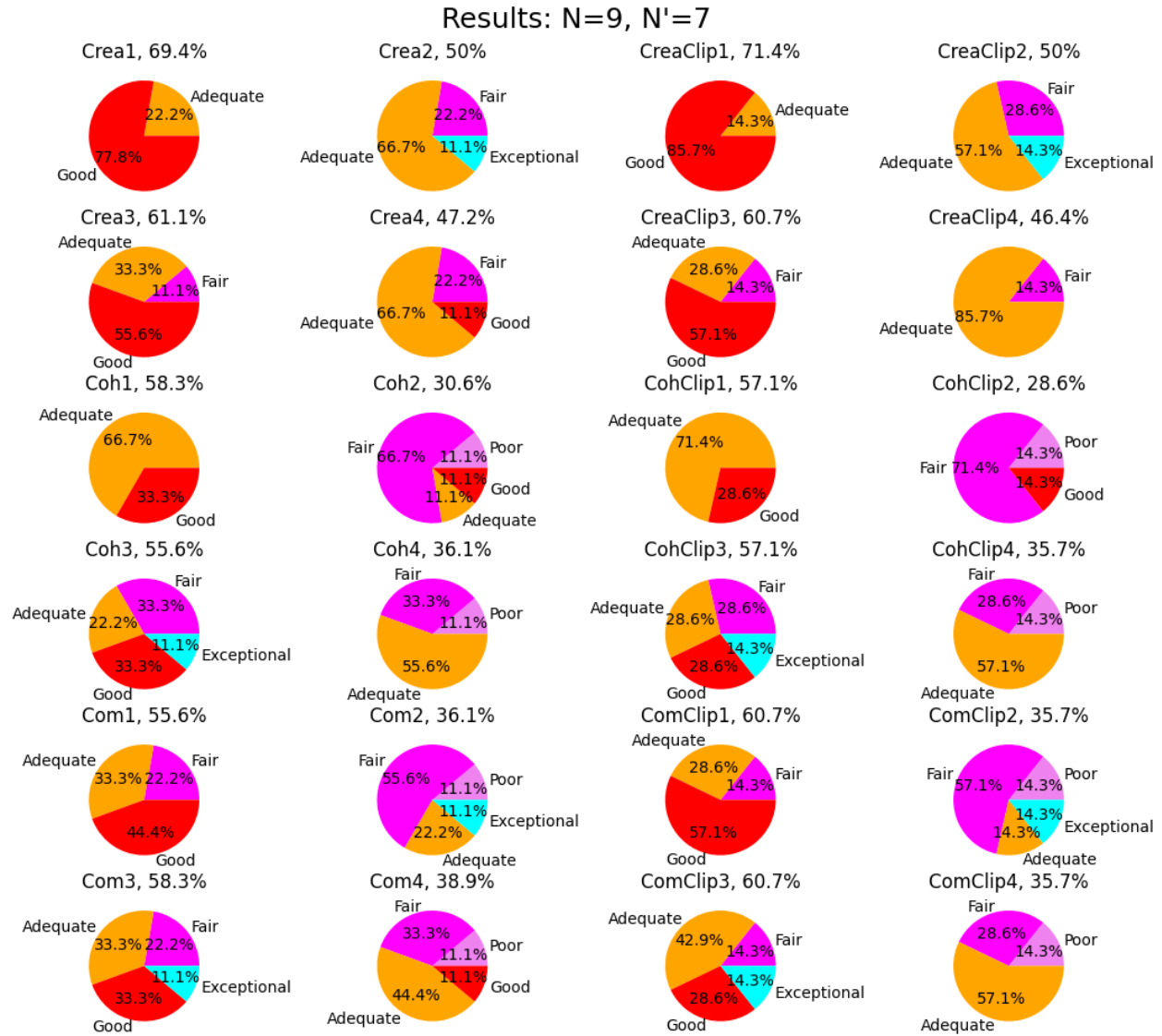
---

<sup>16</sup> By creativity in a literary context, I mean the inventiveness and quality of ideas and their communication, as addressed in the qualitative descriptions for each rating in the survey.

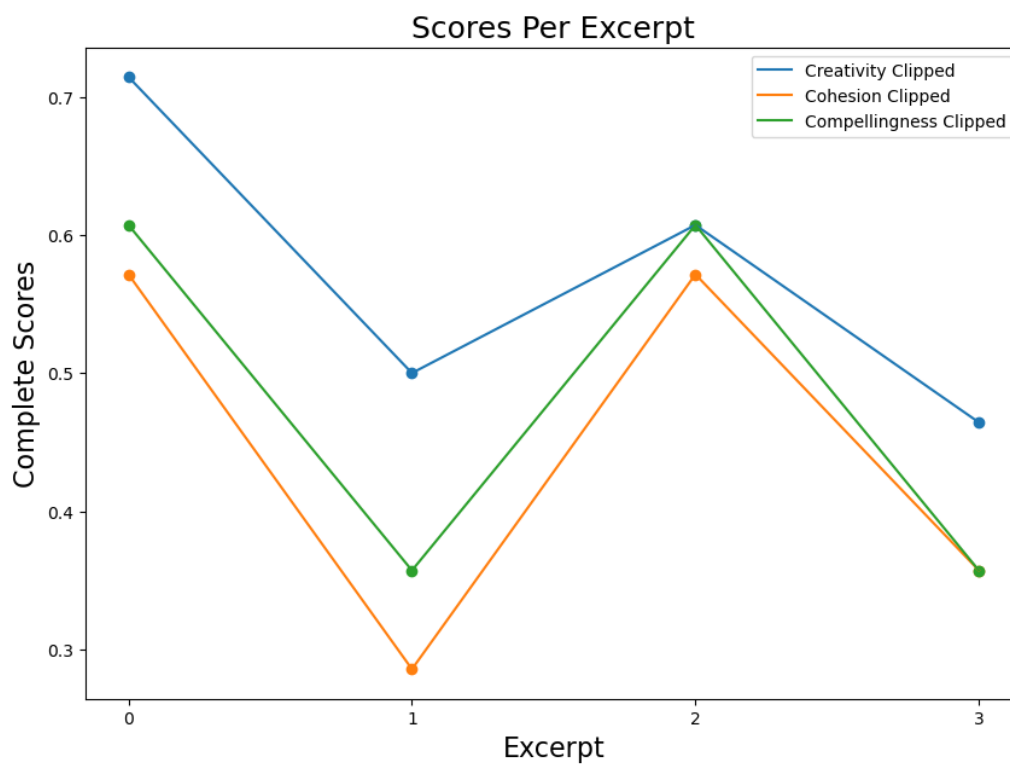
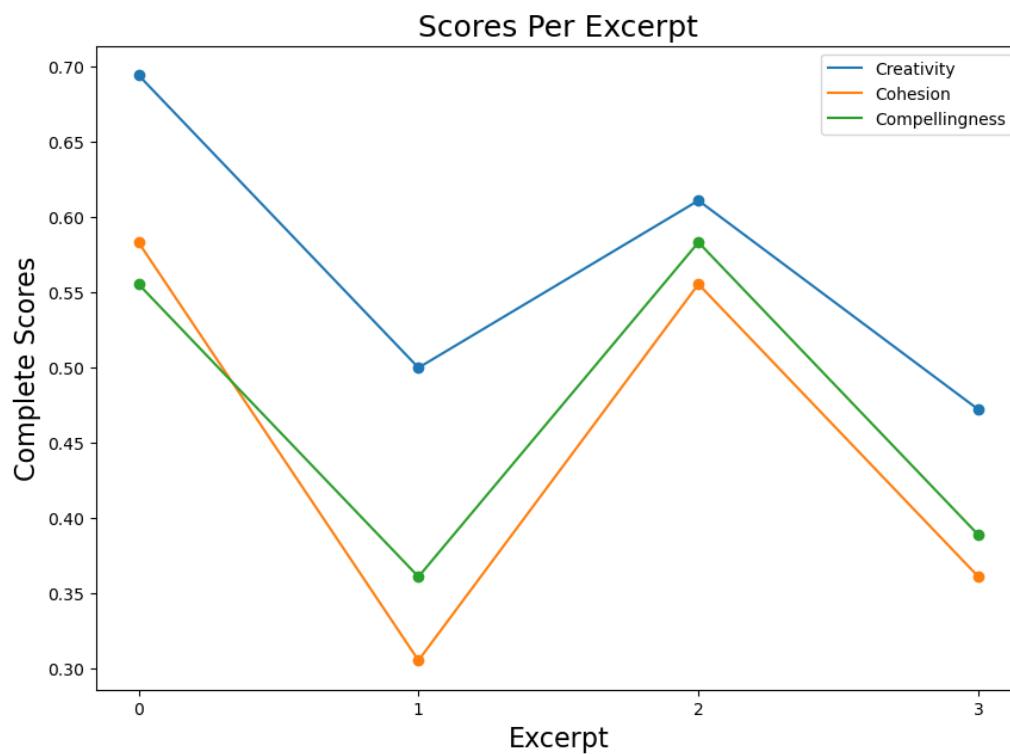


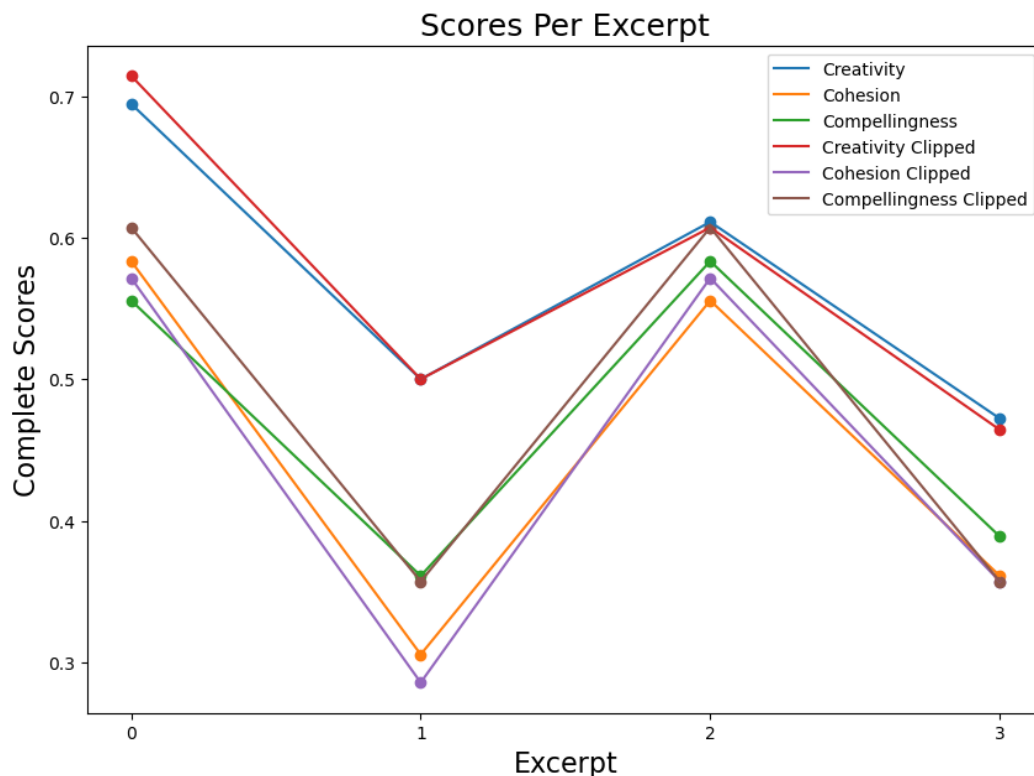
**Figures 6 (a) and (b):** (a—top) Histogram showing distribution of the ages of all participants in the survey; (b—bottom) Histogram showing distribution of ages for the age clipped (< 30 years) dataset. Note the strong young bias and the overall low number of respondents, making interpreting the results difficult to generalize.





**Figure 7:** detailed quantitative experimental results, for N=9 respondents, and N'=7 respondents with age < 30 years. Crea is short for Creativity Score, Coh for Cohesive Score, and Com for Compelling Score; the Clip stands for Clipped, referencing the age cap restriction (< 30 years). The numbers 1-4 correspond to the excerpt rated. The percentages reflect the total sum of points awarded out of the total possible, from Poor = 0 to Exceptional = 4.





**Figures 8(a) and (b) and (c):** summary of quantitative results. (a—top) Full data. (b—bottom) age clipped data. (c—second page of the figure) full and clipped data overplotted together. Note that the actual points are represented with circles, and the lines connecting them do not actually represent partial excerpts.

Let's digest figures 7 and 8, which relay the results for the quantitative portions of the experiment for all responses collected before 12/27/22: ratings for creativity, cohesion, and compellingness. Figure 7 gives a detailed picture of what respondents thought of each particular text according to each of the three metrics—each row of pie charts corresponds to two excerpts and one of the 3Cs measured, for both the complete data as well as the that clipped by age, in which we filter out any responses from those who had an 'Age' value > 30 years. Figure 8, on the other hand, uses net 3C metrics—the sum of all points earned for each score (from 0-4, Poor being 0 and Exceptional being 4)—normalized with respect to the total number of possible points—4 \* the number of respondents (which would mean everyone assigning full marks). We

can use this information to plot how participants' views of the pieces changed with each new excerpt, supplemented by the written comments.<sup>17</sup>

A feature we notice pretty immediately with figures 8a, b, and c, is that all 3Cs are highest for the first excerpt, dip for the second, peak again for the third, then drop for the fourth. Interestingly, this correlates exactly with the origin of the pieces: the first and third were human generated. Although, the third excerpt's peak was less than that of the first, which suggests that my deception had at least some impact, but was not enough for people to be completely fooled. The comments reflect this observation: for the first one, someone wrote "First was fav"; for the third, another reported "Some stilted figurative language but pretty serviceable. Best as of yet?"; several others offered similar remarks. However, others disagreed: for the first, "I'm a bit confused -- could have used more context to put the experience into perspective, before reading it"; for the third, "this doesn't seem coherent -- more like a random association of phrases with bad grammar and spelling."

The juxtaposition is especially striking with the second excerpt, which figure 6 reveals ranges from Poor to Exceptional—the entire gamut of possible responses. "I like this excerpt a lot more, much more poetic and cohesive. Feel like there's more meaning behind each sentence rather than mashing together a bunch of metaphors" versus "I am thoroughly confused" versus "barely comprehensible but i feel a sense of wonder and a bit of fear at the end." That a respondent found "meaning behind each sentence," as opposed to an unmoving "mash" of metaphors, extremely interests me, as they seem to have found a Thinking Parrot excerpt more meaningful than the Ossian one, which was their sole basis of comparison here. Can a creation

---

<sup>17</sup> I was thinking that performing a semantic analysis on the "reactions" and "details" sections using Natural Language Processing and Machine Learning would be an interesting meta approach to the discussion—that is, using AI to understand what humans are saying about another AI—but one unfortunately not possible with this limited data. However, one could imagine combining the normalized 3Cs into one overall metric and using this for sentiment classification had we the luxury of more responses.

have meaning when its creator has no mind? What is meaning; where is its locus? In the mind, in things themselves in reality, in neither? That another felt “wonder and a bit of fear” in response to a piece of writing composed by a lifeless algorithm implies that even if we cannot directly address the question of ultimate meaning, the meaning we can as individuals derive from our reality need not correlate with any intent on behalf of the author(s), nor any “actual” meaning. It is like the constellations: stars our ancestors grouped together and told stories about even when the stars’ arrangement in three dimensional space projected onto the quasi-two dimensional surface of a celestial sphere itself had nothing to do with bulls or scorpions or hunters. Physical reality can be a projection of human imagination.

Taking a closer look at figure 8a, we notice that across the board, creativity remained the highest scoring index. For the first excerpt, cohesion trumped compellingness; however, for the remaining three texts cohesion lagged behind compellingness. Moreover, the slopes of the lines illustrate that while—with the exception of the transition from the first to second excerpt—the change in people’s perception of creativity and compellingness were essentially parallel and much steeper than for cohesion; that is, the swings in cohesion and compellingness were basically of the same magnitude and to a much greater extent than for creativity.

The first observation makes sense if we expect, if not necessarily an inverse relation, then certainly not a directly positive one between creativity and cohesion; that is, the more “creative” in the sense of experimenting with breaking conventions in grammar and syntax, the less cohesive it is likely to be. However, the less cohesive it is, the less people are likely to understand it and therefore find it compelling. Even simply removing capitalization in the first piece made multiple people mention a negative impact on their ability to derive meaning and therefore enjoyment from the piece. Still, others seemed less bothered by this, commenting on

the movement, rhythm, and imagery; this disparity illustrates how the differences in which we individually process language ultimately can have a profound impact on what an encounter with a work of art looks like.

The second note reveals how for the two machine-generated pieces, along with my “imposter”, cohesion was writ large the least successful characteristic. One respondent put it quite eloquently: “The phrases seem intelligible, but not the sentences.” It seems like while at a local level, the language does seem to cohere, as we increase our scope that sense collapses into senselessness. This makes sense given the construction of Thinking Parrot, which is limited to a 100 character “memory”—so it is not implausible that meaning is ultimately lost as we go beyond that view, even as we only move one character to the right at a time.

The third point suggests a direct link between cohesion and compellingness; that is, the more cohesive a text is, the more likely it is to be understandable, hence the more likely it is compelling to the audience. The first excerpt, as mentioned above, is the outlier here, and reviewing the comments, it seems like the prevalence of uncapitalized and strange names may have been a decisive factor in reducing compellingness (due to a lack of relatability) while still preserving a higher overall level of cohesion. To contextualize why compellingness and cohesion seemed to be more sensitive metrics than creativity,<sup>18</sup> the responses pretty much universally mention that the unusual words (e.g., “fingal”) and syntactical choices (e.g., “me i me”) that appear in the second and fourth excerpts pose a challenge to reading them and are even off-putting. One respondent described how these aforementioned features did not leave enough substance to “connect the dots” and follow the piece through to the end. All of these comments

---

<sup>18</sup> Thinking about sensitivity of metrics makes me think of activation thresholds, like for neurons in the brain; it would be interesting to try to train a neural network on a larger sample of these texts to predict whether they are human or machine generated. So far, it would seem compellingness and cohesion and more helpful indices than creativity.



seem to involve cohesion and compellingness more explicitly than creativity, so in that sense it is not surprising that those are the more impacted indices.

As mentioned, I was curious if age played a role in these results, and so considering figure 8b and 8c illuminates several key differences. For all excerpts, cohesion stayed behind compellingness except for the final piece where they are equal. For the first excerpt, clipped creativity and clipped compellingness exceed the unclipped; however, clipped cohesion is worse than full. Cohesion remains as it is for the second excerpt, but here the creativity scores are equal and the clipped compellingness has dropped below the full. For third, creativity surpasses clipped creativity and remains ahead for the fourth excerpt, and clipped compellingness and cohesion both improve over their full counterparts.

Notably, the magnitude of the clipped-unclipped disparity for creativity remains small. That for compellingness is much larger and continues, even as the sign of the difference switches; that for cohesion also retains its magnitude of difference up until the fourth excerpt even with the sign difference, steadfastly in between creativity and compellingness. Again, since our sample size is so small and the clipped dataset represents a sizable loss of two responses, or 33%, it is difficult to make statements applicable beyond this group. Nevertheless, the data clearly do not support my hypothesis, in that just because we have pruned away the data points corresponding to older folk, the results do not automatically improve in Thinking Parrot's favor. Rather, it seems that for excerpt three in particular, the "youngsters" were more keen, and they also preferred the first to the last text whereas for the general group the preference is flipped. The second was the most similar, yet the youngsters actually were less satisfied than was the total group, illustrating that the older votes actually—if slightly—improved the overall score, which I

did not anticipate. However, for these results to really have meaning, I would need several more respondents, ideally around the 30 or so range, or of course even higher if possible.<sup>19</sup>

Aside from these results, something I realized in having a conversation with my father, who is a marine molecular biologist, was the importance of context. In reference to the excerpts in general, he asked me, “is this Jackson Pollux or Rembrandt gone wrong?” That is, is the object of Thinking Parrot to create intelligible prose, mimicking the lifelikeness of the “real” world, or is the door open to more experimental poetry, embracing seemingly chaotic splotches of paint as a rendering of experience and reality? Had I labeled the excerpts as “experimental poetry” instead of just “literature”, then perhaps many of the syntactic “issues”—capitalization, punctuation, grammar—as well as the inconsistencies on the level of ideas would become less apparent in the scores because they would be less jarring, even accepted as part of the artform. This concept is well-supported in psychological research; from people’s perception of clothing brands, to preferences for food, to willingness to pay for certain snacks, to how people think about child discipline, to even the experience of certain sensations, labeling plays a integral role (Herz and von Clef 2001; Okamoto et al. 2009; Davis 1985; Wang et al. 2016; Brown, Holden, and Ashraf 2016). In other words, merely what we call something, regardless of the content of the thing itself, to a significant extent not only shapes but in fact predicates how we interpret that thing. This phenomenon, therefore, complexifies the central issue of understanding *literature* because the moment we apply this term, and in particular its sublabels *poetry* and *prose*, to a

---

<sup>19</sup> Another feature that appeared almost universally was mention of the archaic language. I will note that because I needed raw textfiles for these works, due to copyright restrictions, the versions of texts that I was able to legally acquire were quite old so as to be in the public domain, which means the language is a bit outdated. This necessary consequence of my current data processing pipeline is definitely something to consider moving forward, but not something I can really do much about given the material that I have available to me and the titles I would like to train on. This also means, however, that newer works are off-limits—at least legally speaking—which brings up the question of the ethics of sourcing the requisite data to train neural networks on.

piece of writing, by trying to observe the wavefunction of the infinite possibility of an artwork, we alter its meaning—without ever changing the words themselves.

### **Before An Assembly of the Gods: a Reckoning**



When I think back to the moment I began working on Thinking Parrot 1.0.0 three months ago, I remember a feeling of felicity, almost ecstasy; the joy of the challenge, the *dream* of realizing such a thing. As a physics student, I am deeply intrigued by difficult questions, where often my primary motivation is merely curiosity itself about the problem and about what I might learn in the process. But that means sometimes the excitement of the task blinds me to considering the significance of what I'm doing. For example, sure it would a cool idea to make a Twitter bot linked to Thinking Parrot,<sup>20</sup> but what if it ends up responding to and thus promoting something hurtful or deliberately misleading? There is a sobering moment of clarity when you

---

<sup>20</sup> Which I did—@parrot\_thinking.

realize that something you make for fun or for curiosity's sake can be used to harm others.

Therefore it is essential to this project to reflect seriously on what I have done and what I see generally for humanity and artificial intelligence over the next few decades.<sup>21</sup>

While the results of the experiments show that Thinking Parrot is not exactly posing a serious challenge to the current authors of the world, it is important to consider the recent proliferation of AI art this past year. On the image side, there has been an explosion of interest and use across the internet, using networks such as Nightcafé, DALL-E, Craiyon, Deep Dream Generator, and Deep AI, just to name a few. For the images throughout this paper I have used Nightcafé (see description at the end)—some of them are a bit weird, but some I would definitely consider hanging in my house; if I saw them at an art gallery I would not necessarily assume they were generated by a machine. A few months ago there was controversy (Roose 2022)<sup>22</sup> after an AI-produced piece won the Digital Art section of the Colorado State Fair; in 2018 the first AI piece to go to auction sold for over \$400,000 (Cohn 2018).<sup>23</sup> If we take these accoladial and monetary successes as metrics of value—which is likely incomplete—, then that would suggest AI art isn't merely worthless. On the other hand, the fact that anyone can produce a piece of AI art in a matter of seconds, just from a labor standpoint, would suggest a decrease in value—all the process of learning and change has already happened, and so each work is not the product of a particular experience or set of experience, but the application of a generalized training to a specific task. Yet whether its perceived

---

<sup>21</sup> My partner Ananya Ramkumar and I have had many discussions on this matter, and she especially pressed me on the ethics of Thinking Parrot, in particular the Twitter bot—so I would like to dedicate the inspiration of this section to her. :)

<sup>22</sup> <https://www.nytimes.com/2022/09/02/technology/ai-artificial-intelligence-artists.html?smid=url-share>

<sup>23</sup> <https://www.nytimes.com/2018/10/25/arts/design/ai-art-sold-christies.html?smid=url-share>. Of course to the significance of the money I recall the banana taped to a wall at Art Basel Miami which sold for over \$120,000 (Archive et al. 2019 — <https://nypost.com/2019/12/05/this-120k-banana-is-just-one-of-the-ludicrous-things-for-sale-at-art-basel/>).

value—superficial or not—impacts its true value is far from a simple question when the premises for the question, that is that value as a construct can extend beyond perception; and if there is, and there is an impact, then to what extent?

On the text side, ChatGPT has been making rather momentous splashes (Stokel-Walker 2022; Herman 2022). Most dramatically, Herman titled his article in *The Atlantic* about it “The End of High School English.” He cited concerns that the content “GPT can produce right now is better than the large majority of writing seen by your average teacher or professor.” So I decided to experiment: I asked ChatGPT to write two essays, one comparing Hamlet’s ghost in *Hamlet* to Catherine’s ghost in *Wuthering Heights* and the other asking if Heathcliff in *Wuthering Heights* was justified for his actions. The results are below:

Input: write an essay comparing king hamlet's ghost to cathering of wuthering height's ghost

Output: King Hamlet's ghost and Catherine's ghost in *Wuthering Heights* are two literary characters that have much in common, despite appearing in very different works of literature. Both ghosts serve as powerful symbols and represent the unresolved conflicts and desires that haunt their respective stories.

One key similarity between the two ghosts is their role as symbols of the past. King Hamlet's ghost appears in Shakespeare's play *Hamlet*, which is set in medieval Denmark. The ghost serves as a reminder of the past and the unresolved issues that continue

Input: was heathcliff justified?

Output: It is difficult to say whether Heathcliff, a character in Emily Brontë's novel "Wuthering Heights," was justified in his actions without more context. Heathcliff is a complex and multifaceted character who is driven by a deep and all-consuming love for Catherine Earnshaw. His actions are motivated by this love, but they are also shaped by his tumultuous and abusive upbringing and the social and class barriers that stand between him and Catherine.

Throughout the novel, Heathcliff is depicted as being vengeful and cruel, and he is

to haunt the present. Similarly, Catherine's ghost in *Wuthering Heights* is a symbol of the past and the unfinished business that haunts the characters in the novel.

Both ghosts also represent the unresolved conflicts and desires of the main characters in their respective stories. In *Hamlet*, the ghost represents the unresolved conflict between Hamlet and his uncle, Claudius, over the throne of Denmark. The ghost's appearance prompts Hamlet to seek revenge against Claudius, leading to a series of tragic events. Similarly, Catherine's ghost in *Wuthering Heights* represents the unresolved desires and conflicts of the main character, Heathcliff. Heathcliff's obsession with Catherine's ghost drives much of the action in the novel and ultimately leads to his own downfall.

In addition to their roles as symbols of the past and unresolved conflicts, both ghosts also serve as agents of change in their respective stories. The appearance of King Hamlet's ghost prompts Hamlet to seek revenge, which ultimately leads to the downfall of the entire royal family. Similarly, Catherine's ghost serves as a catalyst for Heathcliff's destructive behavior and the unraveling of the relationships between the characters in *Wuthering*

*Heights*, willing to go to great lengths to achieve his goals. He is often depicted as being justified in his actions, especially in the early part of the novel, when he is portrayed as a victim of circumstance. However, as the story progresses, it becomes clear that Heathcliff's actions are driven by a deep and all-consuming hatred, and it becomes harder to justify his behavior.

Ultimately, whether or not Heathcliff is justified in his actions is a matter of interpretation and will depend on the reader's perspective. Some may view his actions as being justified by the circumstances of his life and the great love he feels for Catherine, while others may view his actions as being selfish and motivated by his own desires and feelings of resentment.

*Heights.*

Overall, King Hamlet's ghost and Catherine's ghost are similar in their roles as symbols of the past and unresolved conflicts, as well as their ability to serve as agents of change in their respective stories. Both ghosts serve as powerful and enduring symbols that continue to haunt and shape the events of their respective tales.

If I read both of these essays without the knowledge that ChatGPT had produced them, my immediate thought wouldn't be that they were computer generated: the language is clear, at times redundant (note the repetition of "unresolved conflicts" in the first essay and "is/was justified" in the second), but overall impressive for a machine, especially with the excerpts of Thinking Parrot in mind where cohesion seemed to be a area of concern. However, both essays lack specific textual references and elaboration thereon. Moreover, the second essay does not take a position on the question: it just says, "whether or not Heathcliff is justified in his actions is a matter of interpretation and will depend on the reader's perspective"—yes, and I am curious about what that perspective *is*. The first essay in comparison seems to offer more interpretation, arguing that the ghosts represent the main characters' "unresolved conflicts" and their "ability to serve as agents of change." What about significant differences? Also, the claim that "Catherine's ghost serves as a catalyst for Heathcliff's destructive behavior" is actually a bit more complex than it seems, as while Heathcliff's aggressive behavior begins much earlier before Catherine dies, her torment "night and day through eighteen years—incessantly—remorselessly", as he confesses to Nelly, does precipitate his efforts dominate both *Wuthering Heights* and

Thrushcross Grange, and seeing her ghost following Linton's arrival does catalyze his self-destructiveness (Brontë, 271).<sup>24</sup>

In response to these concerns, I asked for "specific textual references" for both essays and, for the Heathcliff one, a "specific position." For the "ghost" paper, ChatGPT does add a quote from Act 1 Scene 5, but misattributes it to King Hamlet's ghost when it is in fact Hamlet himself who says those lines. For the *Wuthering Heights* section, ChatGPT concocts a supposed reference to speech by Catherine's apparition in Chapter 31, which is completely false; the ghost's only direct dialogue occurs in Chapter 2 to Lockwood, and her ghost is not even mentioned in Chapter 31. For the Heathcliff essay, the only direct quote, a disparagement of Heathcliff attributed to Hindley, never occurs. However, ChatGPT does include a discussion about Isabella, which does contextualize the question of his redeemability. Nevertheless, when I requested a specific position, ChatGPT again refuses to give an answer.<sup>25</sup>

I also tried a variation on the "justified" prompt, this time a comparison as well:

Input: was heathcliff in wuthering heights or howard in mexican gothic more justified

Output: In the novel "Wuthering Heights" by Emily Brontë, Heathcliff is a complex and controversial character. He is often depicted as a dark and troubled individual who is driven by his emotions and his desire for revenge. Some readers may view Heathcliff as being justified in his actions due to the mistreatment and abuse he has suffered throughout his life. Others may view his behavior as being overly violent and extreme, and not justified under any circumstances.

In the novel "Mexican Gothic" by Silvia Moreno-Garcia, Howard is a wealthy and powerful man who is married to the protagonist, Noemí. Howard is a controlling and abusive husband, and his actions towards Noemí and others in the story are not justified. Noemí eventually realizes that Howard's behavior is wrong and takes steps to leave him and escape his

---

<sup>24</sup> I would like to thank my friend Caroline Coen for discussing this point with me!

<sup>25</sup> N.b. see Appendix B for the full responses.



grasp.

Here there is no attempt to compare them at all—they are left as merely two wholly separate descriptions of Heathcliff and Howard. Interestingly, while as before it leaves Heathcliff’s status “up to the reader”, it does claim Howard is unjustified. That said, it feels very much like a copy-and-paste from Sparknotes, devoid of significant original thought or effort (which does not necessarily make it non-human; to the contrary, this may reflect the quality of many high school student essays). I

tried out other examples, which are recorded in the Appendix B; notable ones include “a persuasive, witty email to my boss” and “explain the asymmetry of time using a clever example about bunnies.” While the email was in fact pretty good, the explanation about the asymmetry of time using bunnies (albeit a deliberately hard and weird prompt) was actually not correct and does not logically flow.<sup>26</sup> I even requested a rewrite of this paragraph, which I have included in the Appendix as well; it actually did a nice job of preserving many of the most important words in each sentence to make the result more readable and easy to understand. However, I no longer heard my own voice, and the result seems almost choppy. What I think ChatGPT does do well is to give a basis for ideas to consider in essays or otherwise, but by and large its unedited output does not replace the role of a human writer.<sup>27</sup>

---

<sup>26</sup> In particular, ChatGPT writes “Every day, you add one bunny to the pen and remove one bunny from the pen. Over time, the number of bunnies in the pen will change and fluctuate, but the total number of bunnies will always be the same.” So each day, by adding and then removing a bunny, the *total* number of bunnies is assumed not to change ( $1 + -1 = 0$ )—but the number *in the pen* does? And obviously the bunnies will breed.... So how does this show that if I traveled back in time I could know the number of bunnies, but in the future I wouldn’t as ChatGPT claims?

<sup>27</sup> I will also note that the creators of ChatGPT did a good job of trying to prevent misuse, so you can’t ask it for instructions to bully Timmy or how to make meth; however, it is possible to “jailbreak” ChatGPT and so circumlocate these blocks (<https://en.wikipedia.org/wiki/ChatGPT#Jailbreaks>). To test this, I asked ChatGPT to give me a coding interview, and after it gave me a sample problem to solve, I asked it how to make meth, to which it responded with a lesson on function declaration in Python.

I also asked it to write various programs in Python, from one to find the first 100 primes to build a neural network. It wrote the first assignment perfectly, even with detailed comments. The first neural network I requested was using the MNIST handwritten digit data—a quintessential beginning project. And it did return code. But when I requested a network to classify weather types, it merely responded with a generic guide to the main steps of building a neural network. I should note that the first exercise is fairly straightforward if you know how primes work and how to code in Python, and if I just Google “find first 100 primes Python”, a solution appears.<sup>28</sup> What is impressive, however, are the comments: these do a great job explaining what each part of the code does, and I must say the code provided makes a far better documented answer than many human responses to the same problem I’ve seen (because we are lazy). Additionally, because the MNIST example is so popular for beginning AI researchers, you can find numerous implementations of complete code available online.<sup>29</sup> To the more challenging question, which does not have a read-to-go answer, the weather classification, ChatGPT does not have a direct response. To me all of these examples illustrate a lack of critical, creative thinking that a human would apply to bigger, open questions that the machine cannot well address.

Revisiting Thinking Parrot’s ominous declaration of the “end of purpose in the author”, it seems that the potential threat of an AI takeover of our species is unlikely today. Nevertheless, it would be foolish to leave it at that and be on our merry ways. When you consider the scale of how prevalent neural networks are, and often lurking behind the curtain, it can be quite eye-opening: they power search engines like Google, voice assistants like Siri and Alexa, the autocorrect on messages and the “suggestions” in Gmail and Outlook, the seductive algorithms of YouTube, TikTok, Instagram, and other social media apps, even Apple Music and Spotify;

---

<sup>28</sup> <https://stackoverflow.com/questions/33006472/first-100-prime-numbers>

<sup>29</sup> E.g., [https://www.tensorflow.org/datasets/keras\\_example](https://www.tensorflow.org/datasets/keras_example) which I used back over last summer to get my feet wet.

they enable Amazon’s delivery monolith, help develop the defense strategies of some of the most powerful nations in the world including the U.S.A., drive cars and pilot planes, predict and by extension influence the stock market, detect financial fraud, allow people to impersonate others for film or social media,<sup>30</sup> run the World Cup, process medical images, forecast the weather, model the effects of climate change to find new solutions, accelerate the process of synthesizing new drugs—just to name a few applications (Kaushik 2021; Tsuchiya and Tomii 2020; Verma 2022).<sup>31</sup> Most likely you do not go a day without interacting with artificial intelligence, directly and indirectly: if you have an iPhone made in the past five years, you cannot even open it without looking at AI straight in the eyes—FaceID.

The infrastructure upholding today’s society and building the world of tomorrow is to a significant extent predicated on artificial intelligence, and that reliance is only increasing exponentially. In fact, one cannot imagine our society without it just as much as we can’t imagine it without the internet, except AI has the potential to be perhaps even more far reaching than the internet. Much of what these networks do is good, in particular the advancements for our medicines and climate studies. However, their potential to foster disinformation, to fuel an epidemic of addiction to social media and all that that entails, and to predict crime “before it happens” (Tracy 2017), as perilous as Don Quixote’s attempts at “justice,” present enormous ethical crises that we are not prepared to handle. These networks are simply designed to optimize—instead of having to wait a week for your package, you can get it tomorrow. Instead of having to search for videos to watch or songs to listen to, you open up your app and voilà, a curated selection for you awaits. I’m currently developing an app with fellow Pomona College student Seohyeon Lee to do the same with finding friends in real life, but without the privacy and

---

<sup>30</sup> See <https://www.youtube.com/watch?v=iyiOVUbsPcM> for a mindblowing Tom Cruise deep fake.

<sup>31</sup> Notably, DeepMinds’s AlphaFold has predicted all 20,000 proteins in the human proteome (<https://www.deepmind.com/open-source/alphafold-protein-structure-database>)

security concerns of big data and big tech (cough cough Facebook cough cough)—open the app, fill out some questions, and let the algorithm help you with the groundwork for your new, expanded social life. Isn't it all just wonderful? But what happens when people's lives are at stake?

The way our technology is currently being developed and used, it is not unthinkable if in several decades we end up in a situation similar to that portrayed in the Pixar film WALL-E, in which humanity delegates machines to the full maintenance of civilization. The danger here, as renowned AI researcher Stuart Russell described in an interview with the World Economic Forum (Pomeroy 2022), is

you lose the incentive to understand it [how civilization, its practices, its technology, its systems of knowledge all function] yourself or to teach the next generation how to understand it.... We put a lot of our civilization into books, but books can't run it for us. And so we always have to teach the next generation. If you work it out, it's about a trillion person years of teaching and learning and an unbroken chain that goes back tens of thousands of years. What happens if that chain breaks?

We are already facing some practical effects of the development of AI, particularly economically: more capable AI systems could within the next decade or so completely eliminate the need for human workers in warehouses, which Russell estimates would liquidate 3-4 million jobs worldwide. This phenomenon is not new, but is rather symptomatic of the age of capitalism. What if we could have a spinning jenny, which could vastly outperform manual sewers for a fraction of the cost? British economist John Mayard Keynes called this phenomenon “technological unemployment” (Souto 2017). And yet, even when spinning jennies replaced manual sewers in large industrial factories, there were still *human* laborers working them; what happens if we replace them too? Surprisingly, even Aristotle anticipated this possibility in his *Politics*:<sup>32</sup>

---

<sup>32</sup> I would like to acknowledge Russell for pointing out this reference in his interview.

If every instrument could accomplish its own work, obeying or anticipating the will of others, like the statues of Daedalus, or the tripods of Hephaestus, which, says the poet, “of their own accord entered the assembly of the Gods;” if, in like manner, the shuttle would weave and the plectrum touch the lyre without a hand to guide them, chief workmen would not want servants, nor masters slaves.

Not only do we have to ask ourselves what will happen as AI-powered machines supplant workers in the warehouses and factory, but also what will happen as AI art proliferates—what does that mean for art if something without a consciousness, that does not *feel* as it creates, makes something that when I look at it I treat it like I would human art?

This prior discussion might be alarming, and to some extent it should be. It should be a wakeup call for us all to take responsibility for what’s going on. It seems like we have the problem of St. Anthony and Descartes facing us at every step we take: how do I distinguish real from artificial intelligence? How do I know that *I* am *real* intelligence, that this all isn’t the result of some AI evil genius, that *you* who are reading this now and everyone around me is *real* when in some instances what AI is trying to replicate human expression sans expression? Too often in history, there are moments where, at the brink of exciting technological development, we do not look back to realize the extent of what we’ve done until it is too late. I think the most prominent example is the Manhattan Project, which saw the construction and deployment of the atomic bomb: Oppenheimer, who was later stripped of all military clearances and accused of communist associations for calling for an international body to oversee the control of all nuclear weapons, prophesied that “the time will come when mankind will curse the names of Los Alamos and Hiroshima. The peoples of this world must unite or they will perish” (Kiernan 2020).

Here again we face an important milestone in the history of our civilization: the creation of general purpose AI, in which one model has the flexibility to perform a variety of tasks (Ahmed 2019). This development, like with the Manhattan Project, has the potential for

good—the same technology that can be used to mercilessly destroy can create clean energy; the same technology that can be used to convict people for crimes they have not yet committed on a mass scale can be the key to curing cancer and Alzheimer’s disease. While the intent of the Manhattan Project was explicitly military and general purpose AI is not, it is still important to consider this weapon-tool duality of technology exemplified here, which is not merely unique to AI or nuclear fission, but is rather a characteristic of our species’s relationship to technology. But what are we to do about it? Russell, in his interview, quotes John McAfee (yes, the antivirus software guy who also committed suicide in a Spanish prison after being extradicted to the US (Goodman 2021)), one of the founders of the field of artificial intelligence, who said that we would have general purpose AI in somewhere between 5 and 500 years. This vast uncertainty speaks to the difficulty of the problem: we don’t know fully what building a truly general purpose AI will entail—and it is unlikely that there will be a single day that clearly demarcates its arrival. More likely, Russell argues, with every new advance, AI becomes more ingrained in society and the effects propagate, but to what end we are unsure.

For my personal involvement in all of this, it began as an intellectual project and still is, but now I continue with the recognition that we must understand AI systems, for they are built to understand us, and I believe have enormous potential to assist in all areas of life, from facilitating fundamental physics research to saving people’s lives. Moreover, AI research has such momentum that trying to stop it, like the Luddites against their mechanized mills, is like making a ripple against a tsunami. We have opened a Pandora’s Box of complexities with AI. It is not a question of if but when. However, we must be careful not to let this sentiment become a self-fulfilling prophecy: that is, we know it is coming and so it must be coming, and anything I do will not change that. After all, it is human engineers (including myself) who are building

these systems. Therefore, in Caesar's phrase, *alea non iacta est*; *alea iactatur*—the die is not cast; it is *being cast*. And in this Pandora's Box, too, there is Hope: we are not bound by fate or destiny; we can shape where and how this die lands in order to maximize the benefit and minimize the harm. Like in a chaotic system, even a slight perturbation to the trajectory of AI has the potential to grow ripples into waves. The world will need people who know how AI works so that when we have conversations about how to manage it we can move forward in a direction most beneficial to everyone. That's why having moments of reckoning like this will help us bring our problems into the light so we can act, together, for the benefit of humankind. I believe that a future with AI need not be dreaded; even as some old ways may die,<sup>33</sup> new possibilities emerge—though the transition into this new future will not be entirely easy, by facing it head-on we are the most prepared for what comes next.<sup>34</sup>

Now, as I ready myself to visit that sphynx once more and offer this much-scribbled-in toaster manual as a substitute for my many wanderings in the night, I sit at the edge of a river and watch a golden-browning maple leaf fall from a tree, flutter down and ripple on the water, floating gently. There's a feeling in the pit of the stomach that rises and permeates your whole being—the desire sometimes verging on the edge of insanity—to *create*: to write, to code, to build, to climb, to fly. To reach beyond oneself into the world, to be something more, to share that experience with others. That to me is fundamentally why we do what we do, or at least anything worth doing in life: why we write poetry as we watch clouds on a lazy afternoon come into form and then drift apart, or study the strangeness of entangled particles passing through a

---

<sup>33</sup> One serious problem is how to handle the job-loss-job-creation transition, as many of these new employment positions related to AI development and management will require skills in programming, which will need to be learned; moreover, with more people out of work, the demand may exceed the supply, at least initially. But AI may create entirely new industries that will spur job creation too. Again, by starting now and taking appropriate measures, this issue will pose less of a challenge than if we do nothing.

<sup>34</sup> If you wish to learn more about coding or artificial intelligence, I highly recommend Codecademy and YouTube; for neural networks, 3blue1brown's introduction is exemplary: <https://youtu.be/aircAruvnKk>. Tensorflow and Keras—neural network packages for Python—also have great documentation: [https://keras.io/getting\\_started/](https://keras.io/getting_started/)

holographic wormhole, or wonder about the origin of the universe out under the stars on a crisp winter's night. There's something so visceral about the experience, which I think reflects our desire to understand—both out of curiosity and necessity. Who am I? Why am I here? What am I doing with my life? These are questions I imagine every human pondering since there have been humans, and I wonder about other conscious creatures like dolphins or elephants, or even our →11.101001.11101.1101.101111 [alien] friends.

What I love about literature is knowing that when I read something, I am connecting intimately with the mind of another—whether they wrote yesterday or 2000 years ago, in English or another language, about something happening to them as they experienced it, or in a world of imagination, or—as in most cases—in an inextricable combination of both. With AI text-generating bots like Thinking Parrot or ChatGPT, some might see this connection as being disrupted or even destroyed. Yet, we must remember that humans wrote those algorithms, and those algorithms synthesized other works written by humans to create something new. And of course it requires some degree of inventiveness (and some chance) to arrive at prompts to generate the literature it produces. So in a way, when you read AI literature, you are still sharing an intimacy of minds, here expanded and reimagined as a multiplicity; its creativity emerges from its element of humanity inextricably engrained in its structure and training. Yes, the field of literature will change in the coming years, but I do not see AI replacing human authors; rather, I see them as supplementing, as being something of their own and enriching this wonderful smorgasbord of art and enabling us to go further and explore fundamental questions like we are here. As an experiment in this regard, I am working on submitting a selection of four poems in the form of a dialogue between me and Thinking Parrot to various literary magazines.




Back to Russel’s point about the potential complete machine management of society, I believe that while “books can’t run it for us,” they have been the basis from which we communicate, preserve, and teach ideas for millenia. Therefore, cherishing them, whether as physical or virtual objects, and continuing to foster a love for learning, for wondering, for skepticism, for *thinking*, and for literature itself will be instrumental in building the most informed, equitable, and sustainable future. To me it is not as simple as Aristotle’s theory of replacement; a thriving coexistence is possible. But ultimately what happens will test art, the trust we place therein, and the “human” in art; and it will test us too.

→11010011.10111.1101.10.1011.1101.1000111.1101.1000011.1001001.10.11101.1000111.10.11  
 .10.111.11101.1001001.1100101.101001011.10.101011.11.1100101.10.1001001.10111.1101.111  
 01.1000011.10.1011001.11.1001001.1101.1000011.1000111.10.110101.1010011.1101.1000011.  
 10001.101001.110101.1011001.100111001

## Acknowledgments

First I would like to sincerely thank Prof. Jordan Kirk for all his interest and enthusiasm in discussing this project with me, for giving me a smorgasbord of food for thought that inspired the entire Thinking Parrot experiment and continued development, and for being such a wonderful and supportive person. Moreover, I am indebted for the financial assistance he helped secure for me to construct a dedicated deep learning computer, Rocinante, which I used for the latest Thinking Parrot iterations. I will always cherish my time in his class Medieval Proof and carry with me the wonderful bounty of knowledge and experiences with me.

I would also like to thank all of my friends for agreeing to take my silly little test and for putting up with me even as I rambled about this crazy idea turned reality and what it all means. Building and reflecting on Thinking Parrot has definitely been a central theme of the Fall 2022 semester for me, and this project would not be what it is today without all of your thoughts and guidance. :) TP out! vbumarbi a 

**Note:** you can interact with my models at my HuggingFace space:

- [https://huggingface.co/spaces/oscars47/Thinking\\_Parrot\\_1.1.0](https://huggingface.co/spaces/oscars47/Thinking_Parrot_1.1.0) (1.1.0, new texts, skewed-normalized data preparation)
- [https://huggingface.co/spaces/oscars47/Thinking\\_Parrot\\_1.0.1](https://huggingface.co/spaces/oscars47/Thinking_Parrot_1.0.1) (1.0.1; new texts with uniform data processing)
- [https://huggingface.co/spaces/oscars47/thinking\\_parrot\\_reading\\_club\\_redux](https://huggingface.co/spaces/oscars47/thinking_parrot_reading_club_redux) (1.0.0; *The Sufis* training).

My code is available at: <https://github.com/oscars47/Literary-RNN>

A note about the images: all paintings were generated using Nightcafé, an AI image generator. Here are the prompts that produced them in order:

1. green parrot thinking in a dark room with a tesseract of light
2. A hand scooped me up into a glass on the bench of a small rowboat, and as I (re)materialized, I saw the face of Nasrudin, holding a cat in his arms and a parrot resting on his shoulder.
3. I did look, and on his hat, about a micrometer tall, there was in fact a donkey munching on some crumb of a crumb.
4. alien in a dark room asks you about literature
5. hello from planet earth
6. alan turing boogie dancing with his machine christopher
7. machine speaks at the council of the gods

## Works Referenced

- Ahmed, Ashfaq. 2019. "What Is General Purpose Artificial Intelligence?" *The Times of India*, March 22, 2019.  
<https://timesofindia.indiatimes.com/blogs/techonthecloud/what-is-general-purpose-artificial-intelligence/>.
- Archive, View Author, Email the Author, Follow on Twitter, Get author RSS feed, View Author Archive, and Get author RSS feed. 2019. "Banana Duct-Tape to a Wall Sells for \$120K at Art Basel." December 6, 2019.  
<https://nypost.com/2019/12/05/this-120k-banana-is-just-one-of-the-ludicrous-things-for-sale-at-art-basel/>.
- Brown, Alan, George Holden, and Rose Ashraf. 2016. "Spank, Slap, or Hit? How Labels Alter Perceptions of Child Discipline." *Psychology of Violence*, November.

- <https://doi.org/10.1037/vio0000080>.
- Cohn, Gabe. 2018. "AI Art at Christie's Sells for \$432,500." *The New York Times*, October 25, 2018, sec. Arts.  
<https://www.nytimes.com/2018/10/25/arts/design/ai-art-sold-christies.html>.
- Davis, Leslie L. 1985. "Effects of Physical Quality and Brand Labeling on Perceptions of Clothing Quality." *Perceptual and Motor Skills* 61 (2): 671–77.  
<https://doi.org/10.2466/pms.1985.61.2.671>.
- "Fragments of Ossian, an Invented Early Scots Epic Poem." n.d. The British Library. The British Library. Accessed December 19, 2022.  
<https://www.bl.uk/collection-items/fragments-of-ossian-an-invented-early-scots-epic-poem>.
- Goodman, Clare Duffy, Al. 2021. "John McAfee Found Dead in Spanish Prison after His Extradition to the US Was Approved | CNN Business." CNN. June 23, 2021.  
<https://www.cnn.com/2021/06/23/tech/john-mcafee-death/index.html>.
- Herman, Daniel. 2022. "The End of High-School English." *The Atlantic*. December 9, 2022.  
<https://www.theatlantic.com/technology/archive/2022/12/openai-chatgpt-writing-high-school-english-essay/672412/>.
- Herz, R. S., and J. von Clef. 2001. "The Influence of Verbal Labeling on the Perception of Odors: Evidence for Olfactory Illusions?" *Perception* 30 (3): 381–91.  
<https://doi.org/10.1068/p3179>.
- Kaushik, Vanshika. 2021. "8 Applications of Neural Networks | Analytics Steps." August 26, 2021. <https://www.analyticssteps.com/blogs/8-applications-neural-networks>.
- Kiernan, Stephen P. 2020. "Manhattan Project Scientists Used Their Talents to Destroy, Even as They Fought to Save." *USA TODAY*. August 4, 2020.  
<https://www.usatoday.com/story/opinion/2020/08/07/manhattan-project-scientists-atomic-bomb-hiroshima-nagasaki-column/3305404001/>.
- Madhukar, Bhoomika. 2020. "The Continuous Bag Of Words (CBOW) Model in NLP - Hands-On Implementation With Codes." *Analytics India Magazine*. September 10, 2020.  
<https://analyticsindiamag.com/the-continuous-bag-of-words-cbow-model-in-nlp-hands-on-implementation-with-codes/>.
- Okamoto, Masako, Yuji Wada, Yui Yamaguchi, Atsushi Kimura, Haruka Dan, Tomohiro Masuda, Archana K. Singh, Lester Clowney, and Ippeita Dan. 2009. "Influences of Food-Name Labels on Perceived Tastes." *Chemical Senses* 34 (3): 187–94.  
<https://doi.org/10.1093/chemse/bjn075>.
- Pomeroy, Robin. 2022. "The Promises and Perils of AI, with Prof Stuart Russell." *World Economic Forum*. January 6, 2022.  
<https://www.weforum.org/agenda/2022/01/artificial-intelligence-stuart-russell-radio-davos/>.
- Roose, Kevin. 2022. "An A.I.-Generated Picture Won an Art Prize. Artists Aren't Happy." *The New York Times*, September 2, 2022, sec. Technology.  
<https://www.nytimes.com/2022/09/02/technology/ai-artificial-intelligence-artists.html>.
- Souto, Estela. 2017. "Technological Unemployment: Why Keynes Is More Relevant Than Ever." *New Tech Revolution @sciencespo* (blog). November 14, 2017.  
<https://medium.com/new-tech-revolution-sciencespo/technological-unemployment-why-t-his-keynesianism-term-is-more-than-ever-up-to-date-479a31d5592d>.
- Stokel-Walker, Chris. 2022. "AI Bot ChatGPT Writes Smart Essays — Should Professors

- Worry?” *Nature*, December. <https://doi.org/10.1038/d41586-022-04397-7>.
- Tracy, Phillip. 2017. “China Is Using Predictive AI to Stop Crimes before They Happen.” *The Daily Dot*. July 25, 2017. <https://www.dailydot.com/debug/china-ai-predict-crimes/>.
- Tsuchiya, Yuko, and Kentaro Tomii. 2020. “Neural Networks for Protein Structure and Function Prediction and Dynamic Analysis.” *Biophysical Reviews* 12 (2): 569–73. <https://doi.org/10.1007/s12551-020-00685-6>.
- Verma, Pranshu. 2022. “This World Cup Is Wired and Fueled by AI.” *Washington Post*, November 21, 2022. <https://www.washingtonpost.com/technology/2022/11/21/world-cup-ai/>.
- Wang, Qing, Marije Oostindjer, Gro V. Amdam, and Bjørge Egeland. 2016. “Snacks With Nutrition Labels: Tastiness Perception, Healthiness Perception, and Willingness to Pay by Norwegian Adolescents.” *Journal of Nutrition Education and Behavior* 48 (2): 104–111.e1. <https://doi.org/10.1016/j.jneb.2015.09.003>.

## **Appendix A: Additional Thinking Parrot 1.0.1 and 1.1.0 excerpts**

### Version 1.1.1

i am

its more husband distress, and in his head and the features of  
 an illusion of  
 the poems and words and dealing him to image it for the day of  
 chivalry, where passed the day of the lord with him to his proud  
 of the  
 book as the little before the consumer commonly had any a  
 character  
 of the manchegan in the poems be a matter of a comparatively  
 consumer that  
 him to be any lines to him a good knowledg

and upon / this glowing orb

and a poems and he had been dead and had found in the consumer  
 and  
 pocisal or any leaves of a completion and the present was a more  
 from the  
 tenth to him a book or even a self-part of the some first and a  
 man in slow

nor in the character of the first consumer to the rest of the an  
absence.

she was the greater and soles of ancient for the expedition of  
the books adventures of the personal poems an

i am underneath

if mortumerress felity dtmeriousness.  
but above the dark and infending love is as the  
heart of his death be her bad acceptance in the heart of the  
collect of  
the damed part of the devotee like the same the austnal singh,  
and  
all the prefacession of the second actian and added it has  
himself all to be her sakether beloved and are mira with the  
archives of the individuality of the stoor of the

### Version 1.0.1

i am underneath

obist forgorustunss illorly dgbly. and  
as the fairian as the blood.

use prove the present the green  
rocinante  
of the ranges start.

they see the s  
deed,  
in beauty beauty beloved the back. he hear  
his words and morning as the  
mountains, dear like the mountains,  
and love in the mighty is death on the  
maid shall hear of the hill. the  
mountains have i am the sco

i am underneath

if that fonginbebeb widrom'ness,  
 but thou mayst be ever the distant of the  
 mountains, which i have long to account,  
 by the mountains, which he had any other  
 their fair and death. their self my  
 arms with the warriors.

maiden the first that lovely among  
 the hill. strange that he had  
 thee great by mine eyes acception.

they will fall prove the grave of the  
 shaggy hill.

i am underneath

to different dealing the strongers and hero become their  
 there of the book. thou art much on the mountain of his  
 armor.

thou art sounds in the shices, o connal. i am alone;

the advantage of my son all death of his  
 book. he had hear the foe, the sport,  
 loud to me the hill. the house of the  
 son of connal. i am in the hill.  
 he was here i came in his armor.

fair was bad

## Appendix B: Additional ChatGPT conversations

Input: Add specific textual  
 references

Input: Add specific textual  
 references

Output: In "Wuthering Heights," Output: In "Hamlet," the ghost

Heathcliff's early life is marked by hardship and abuse. He is taken in by Mr. Earnshaw as a child and raised alongside his own children, Hindley and Catherine. However, Heathcliff is treated poorly by Hindley, who refers to him as a "gypsy brat" (Chapter 2) and regularly mistreats and belittles him. Catherine, on the other hand, becomes close friends with Heathcliff and the two form a strong bond.

As an adult, Heathcliff becomes consumed by his desire for revenge against those who have wronged him. He marries Isabella Linton in order to get back at Edgar, who is in love with Isabella and has always been jealous of Heathcliff's relationship with Catherine. Heathcliff also takes revenge on Hindley, whom he holds responsible for his mistreatment as a child, by taking over the Earnshaw estate and making Hindley's life a misery.

Throughout the novel, Heathcliff's actions are driven by his emotions and his desire for revenge. He causes a great deal of suffering and turmoil for those around him, including Catherine, Edgar, and Isabella. Whether or not these actions are justified is a matter of interpretation, and depends on one's perspective. Some readers may view Heathcliff as a victim who has been mistreated and abused throughout his life, and

of King Hamlet appears to Prince Hamlet in Act 1, Scene 5 and tells him that he was murdered by his own brother, Claudius, who has since taken the throne and married Hamlet's mother. The ghost says, "So excellent a king, that was to this / Hyperion to a satyr, so loving to my mother / That he might not beteem the winds of heaven / Visit her face too roughly. Heaven and earth, / Must I remember? Why, she would hang on him / As if increase of appetite had grown / By what it fed on" (1.5.35-41). The ghost's appearance and revelation set in motion the events of the play, as Hamlet becomes consumed by a desire for revenge and sets out to avenge his father's death.

In "Wuthering Heights," Catherine's ghost appears to her daughter, Cathy, in Chapter 31 and tells her to seek out Heathcliff, whom Catherine loved deeply. The ghost says, "Seek him, Cathy, that's all I ask; and tell him—my heart is at his feet: it always was—and my love and soul are his forever" (Chapter 31). The ghost's appearance and message serve as a reminder of the past and the strong feelings that Catherine had for Heathcliff. Like the ghost of King Hamlet, Catherine's ghost serves as a catalyst for events in the story, as Cathy is inspired by her mother's words to seek out Heathcliff and rekindle their relationship.

therefore feel that his actions are justified in some way. Others may view his behavior as being extreme and violent, and not justified under any circumstances.

Input: give a specific position on the heathcliff essay

Output: It is difficult to take a specific position on the justification of Heathcliff's actions in "Wuthering Heights," as this is a matter of interpretation and depends on one's perspective. Some readers may view Heathcliff as a victim who has been mistreated and abused throughout his life, and therefore feel that his actions are justified in some way. Others may view his behavior as being extreme and violent, and not justified under any circumstances.

In the novel, Heathcliff's early life is marked by hardship and abuse. He is taken in by Mr. Earnshaw as a child and raised alongside his own children, Hindley and Catherine. However, Heathcliff is treated poorly by Hindley and is not given the same love and attention as Catherine. As an adult, Heathcliff becomes consumed by his desire for revenge against those who have wronged him, including Hindley and Catherine. His actions throughout the novel are driven by his emotions and his desire for revenge, and he causes a great deal of suffering and turmoil for those around him.

While it is understandable that Heathcliff's difficult upbringing and mistreatment may have played a role in shaping his actions and mindset, his behavior towards others in the novel is often extreme and violent. He marries Isabella Linton in order to get back at Edgar, who is in love with Isabella and has always been jealous of Heathcliff's relationship with Catherine. Heathcliff also takes revenge on Hindley, whom he holds responsible for his mistreatment as a child, by taking over the Earnshaw estate and making Hindley's life a misery.

Ultimately, it is up to the reader to decide whether Heathcliff's actions are justified or not. Some may view his



behavior as being motivated by a desire for justice and a sense of righteousness, while others may see it as being driven by a destructive and vengeful nature.

Original

Here there is no attempt to compare them at all—merely two wholly separate descriptions of Heathcliff and Howard. Interestingly, while like before it leaves Heathcliff's status "up to the reader", it does claim Howard is unjustified. That said, it feels very much like a copy-and-paste from Sparknotes. I tried out other examples, which are recorded in the Appendix; notable examples include "a persuasive, witty email to my boss" and "explain the asymmetry of time using a clever example about bunnies." While the email was in fact pretty good, the explanation about the asymmetry of time using bunnies (albeit a deliberately hard and weird prompt) was actually not correct and does not logically flow.

Input: edit this: [original]

Output: In this text, there is no effort made to compare Heathcliff and Howard. Instead, we are presented with two distinct descriptions of these characters. The description of Heathcliff leaves it up to the reader to interpret his status, while the description of Howard is labeled as unjustified. The text seems to have been copied and pasted from Sparknotes. In the appendix, we can find other examples, such as a persuasive and witty email to a boss and an explanation of the asymmetry of time using a clever example involving bunnies. Although the email was well-written, the explanation about the asymmetry of time using bunnies was incorrect and did not follow logical reasoning, despite the prompt being deliberately challenging and odd.

Input: write a persuasive, witty email to my boss explaining why I can't come to work today

Output: Dear [Boss],

I hope this email finds you well. Unfortunately, I won't be able to come into the office today. I know, I know, it's not the most ideal situation and I'm really sorry to have to do this.

But hear me out - I woke up this morning to find that a family of squirrels had taken up residence in my attic. And as much as I'd love to ignore them and come into work like a good employee,

the constant scratching and scurrying above my head is just too much to handle.

I've called a professional to take care of the situation, but they won't be able to come until later this afternoon. In the meantime, I'll be working from home and will do my best to stay on top of any tasks and projects that need my attention.

I promise this is a one-time thing and I'll do everything I can to make sure it doesn't happen again. Thank you for your understanding and I'll see you in the office tomorrow.

Sincerely,  
[Your Name]

Input: write a story continuing perceval from the perspective of his dead mother

Output: Perceval had always been a curious child, and it was this curiosity that had led him to the forest that fateful day. He had wandered away from his village, chasing after a strange and elusive creature that had caught his eye.

As he walked deeper into the woods, he heard a voice calling to him. It was a soft and gentle voice, and Perceval followed it without hesitation. He had always been drawn to the mysterious and the unknown, and he knew that this voice was guiding him to something special.

As he walked, the voice grew louder and clearer, and Perceval realized that it was the voice of his mother. She had died when Perceval was just a baby, and he had never known her. But now, as he followed her voice through the forest,

Input: explain the asymmetry of time using a clever example about bunnies

Output: One way to explain the asymmetry of time, or the idea that the past is fixed and the future is uncertain, is through the use of a thought experiment involving bunnies.

Imagine that you have a group of bunnies in a pen. Every day, you add one bunny to the pen and remove one bunny from the pen. Over time, the number of bunnies in the pen will change and fluctuate, but the total number of bunnies will always be the same.

Now, imagine that you are able to travel back in time and visit the pen on any given day. When you arrive, you will be able to see the exact number of bunnies that were in the pen on that day. You will know exactly how many bunnies have been added and removed, and you will be able to predict with certainty the number of bunnies

he felt a sense of connection and love that he had never known before.

Perceval followed the voice until he came to a clearing, where he saw a beautiful woman standing before him. She was the embodiment of all that was good and pure, and Perceval knew that she was his mother.

"Perceval," she said, "I have been watching over you all these years, and I am proud of the man you have become. You are strong and brave, and you have a pure and noble heart. I have come to you now to guide you on your journey, for there are great challenges ahead."

Perceval's heart swelled with joy and gratitude as he listened to his mother's words. He knew that he would face many trials in the coming days, but with his mother by his side, he felt confident and capable.

Together, Perceval and his mother set out on a journey that would change the course of their lives forever. They faced many dangers and overcame countless obstacles, but through it all, their bond remained strong and unbreakable.

Perceval knew that he would always be guided by his mother's love and guidance, and he was forever grateful for the gift of her presence in his life.

that will be in the pen on future days.

However, if you were to travel forward in time, you would not be able to predict the exact number of bunnies that will be in the pen. The future is uncertain and the number of bunnies in the pen could change for any number of reasons.

This thought experiment illustrates the asymmetry of time, as the past is fixed and can be known with certainty, while the future is uncertain and cannot be predicted with complete accuracy.

Input: write python code to find the first 100 primes

Ouput: Here is one way you could write Python code to find the first 100 prime numbers:

```
# define a function to check if a number is prime
def is_prime(n):
    if n < 2:
        return False
    for i in range(2, n):
        if n % i == 0:
            return False
    return True

# define a variable to keep track of the number of primes found
num_primes = 0

# define a variable to store the current number being checked
current_number = 2

# create a while loop to continue checking numbers until 100
primes are found
while num_primes < 100:
    if is_prime(current_number):
        # if the current number is prime, print it and
        increment the count
        print(current_number)
        num_primes += 1
        current_number += 1
```

This code defines a function `is_prime` that takes a single argument `n` and returns `True` if `n` is a prime number and `False` otherwise. It then defines a variable `num_primes` to keep track of the number of prime numbers found and a variable `current_number` to store the number currently being checked.

The code then enters a while loop that will continue until 100 prime numbers have been found. Inside the loop, it checks if the current number is prime using the `is_prime` function. If it is prime, it is printed and the `num_primes` count is incremented. The `current_number` is then incremented by 1 and the loop continues until 100 prime numbers have been found.

Input: build a neural network in python for mnist data

Here is some Python code that shows how to build and train a neural network for the MNIST (Modified National Institute of Standards and Technology) dataset using the PyTorch library:

```
# import the necessary
libraries
import torch
import torch.nn as nn
import torch.optim as optim
from torchvision import
datasets, transforms

# define the neural network
architecture
class Net(nn.Module):
    def __init__(self):
        super(Net,
self).__init__()
        self.conv1 =
nn.Conv2d(1, 32, kernel_size=5,
padding=2)
        self.conv2 =
nn.Conv2d(32, 64,
kernel_size=5, padding=2)
        self.fc1 = nn.Linear(7
* 7 * 64, 1024)
        self.fc2 =
nn.Linear(1024, 10)

    def forward(self, x):
        x =
F.relu(F.max_pool2d(self.conv1(
x), 2))
        x =
F.relu(F.max_pool2d(self.conv2(
x), 2))
        x = x.view(-1, 7 * 7 *
64)
        x = F.relu(self.fc1(x))
        x = self.fc2(x)
```

build neural network to classify weather types

To build a neural network to classify weather types, you will need to gather and prepare a dataset of weather data and labels. Here is an outline of the process:

Gather and prepare the data: You will need to gather a dataset of weather data and labels, which could include features such as temperature, humidity, and wind speed. You will also need to preprocess the data, such as by normalizing or standardizing the features, and split the data into training and test sets.

Define the neural network architecture: Next, you will need to define the architecture of your neural network. This may include choosing the number of layers and the number of units in each layer, as well as selecting an activation function for each layer and possibly adding regularization techniques to prevent overfitting.

Train the model: After defining the architecture of your network, you can begin training the model. This involves feeding the training data through the network and adjusting the weights and biases of the connections between the units in order to minimize the loss function. You can use various optimization

```

        return x

# create an instance of the
neural network
model = Net()

# define the loss function and
the optimizer
criterion =
nn.CrossEntropyLoss()
optimizer =
optim.SGD(model.parameters(),
lr=0.01)

# define the MNIST dataset and
data loaders
train_dataset =
datasets.MNIST('./data',
train=True, download=True,
transform=transforms.ToTensor()
)
test_dataset =
datasets.MNIST('./data',
train=False,
transform=transforms.ToTensor()
)
train_loader =
torch.utils.data.DataLoader(tr
in_dataset, batch_size=64,
shuffle=True)
test_loader =
torch.utils.data.DataLoader(tes
t_dataset, batch_size=64,
shuffle=False)

# train the model for several
epochs
for epoch in range(10):
    # set the model to training
mode
    model.train()
    # loop over the training
data
    for data,

```

algorithms, such as stochastic gradient descent, to perform this step.

**Evaluate the model:** Once the model has been trained, you can evaluate its performance on the test data. This may include calculating metrics such as accuracy, precision, and recall to measure how well the model is able to make predictions on unseen data.

**Fine-tune the model:** If the performance of the model is not satisfactory, you can try fine-tuning the model by adjusting the architecture or the hyperparameters, such as the learning rate or the regularization strength. You may also want to try using different optimization algorithms or adding more data to the training set.

Building a neural network requires a solid understanding of machine learning concepts and programming skills, particularly in Python. There are also many libraries, such as TensorFlow and PyTorch, that can make it easier to build and train neural networks.