

Adrian Serieyssol

Walter Isaacson

The History of the Digital Revolution

14 April, 2024

## AI Paper

### **The First Draft Produced by Claude.ai**

The development of artificial intelligence (AI) has been a long and fascinating journey, spanning decades and building upon the pioneering work of researchers and innovators. Here's a high-level overview of the key milestones:

#### **Alan Turing and the Turing Test (1950s):**

The foundations of AI were laid by Alan Turing, a brilliant British mathematician and computer scientist. In 1950, Turing published a seminal paper, "Computing Machinery and Intelligence," which proposed the Turing Test as a way to determine if a machine can exhibit intelligent behavior indistinguishable from a human.

#### **Early AI Systems (1950s-1960s):**

Following Turing's work, the 1950s and 1960s saw the emergence of the first AI systems, such as the Logic Theorist, the General Problem Solver, and the Neural Nets. These early systems focused on problem-solving, decision-making, and pattern recognition, laying the groundwork for more advanced AI.

### **Knowledge-Based Systems (1970s-1980s):**

In the 1970s and 1980s, AI research shifted towards knowledge-based systems, which used large databases of information and rules to make inferences and solve problems. Expert systems, such as MYCIN for medical diagnosis, were prominent during this period.

### **Machine Learning and Neural Networks (1980s-2000s):**

The 1980s and 1990s witnessed a resurgence of interest in neural networks and machine learning, with the development of backpropagation algorithms and the increasing computational power of computers. This led to breakthroughs in areas like computer vision, speech recognition, and natural language processing.

### **Deep Learning and Artificial General Intelligence (2000s-present):**

The 2000s saw the rise of deep learning, a form of machine learning that utilizes artificial neural networks with multiple hidden layers. This revolution, along with the availability of vast amounts of data and increased computing power, has led to significant advancements in areas like computer vision, natural language processing, and game-playing algorithms (e.g., AlphaGo).

### **Large Language Models and Chatbots (2010s-present):**

In the 2010s and 2020s, the development of large language models, such as GPT-3 and BERT, has enabled the creation of highly sophisticated chatbots and conversational AI assistants. These models are trained on massive amounts of text data and can engage in natural language understanding, generation, and dialogue, paving the way for more advanced conversational AI.

Today, the field of AI continues to evolve rapidly, with ongoing research and development in areas like reinforcement learning, generative adversarial networks, and the pursuit of artificial general intelligence (AGI) – systems that can match or exceed human-level performance across a wide range of cognitive tasks. The journey from Turing's pioneering work to the modern chatbots and language models we see today has been a remarkable testament to the power of human ingenuity and the potential of artificial intelligence.

### **A Step-by-Step Guide of How the Paper was Improved**

#### Ideation:

I wanted to do something unique in my paper, though by what means I did not know, for I knew that the paper had to both be somewhat innovative but also feasible. As cool as it would be to have the likes of Turing, Jobs, and Musk talk to each other, it would not have been possible, without spending an amount of spare time I don't possess, to make something in the style of Plato's Dialogues. On the other hand, having a dry resuscitation of facts would be quite dull. So I settled somewhere in the middle, having the AI write a poem, while it wouldn't be exceptional, it would be far more interesting and fun than if it had been a standard essay.

The system I ended up using to get most of my essay was this: 1) turn the sections of the above essay using either Chat GPT, Gemini, or Claude, 2) ask whichever program to add more details and ensure it covered what I wanted it to, 3) give Claude the essay and ask it to turn it into an epic poem in the style of \_\_\_\_\_, 4) ask Claude to add more details and quotes, 5) ask Claude fix

any spectacular hallucinations, 6) ask Claude to add sources, and 7) edit each one, combine them, and examine sources and quotes for fiction (this part done by me).

### Draft Phase:

When I put the prompt into Claude it gave me back a first draft in clearly labeled sections, which I used as a roadmap, first asking Claude, Gemini, and Chat GPT 3.5 to expand on them (in standard essay form. The responses these programs gave back were overly general and lacked many specific details and would lack any discussion on say Ada Lovelace or the suit against Bill Gates and Microsoft. I found that prompting Chat GPT was relatively straightforward, as long as I made sure to include the details I especially wanted, such as the suit against Bill Gates and Microsoft, or Steve Job's love of the humanities, it would produce a decent paper on a narrow subject matter. However when asked to turn it into an epic poem, it immediately lost all of the details and became a string of somewhat pretty words that signified nothing.

Gemini was quite good at giving details, when the prompts asked for standard essay form, right out of the gate. However it ran into problems quite quickly as it would be stubborn, it took three different times of rephrasing "don't use bullet points" for it to finally remove bullet points from the essay. Then when asked to turn what it had written into poetry, specifically in the style of Paradise Lost, it removed all of the details and was wholly superficial.

I decided to use Claude.ai to turn my standard essays on each of the sections outlined in the first draft, as well as my own section on the conflicting visions for AI, into poetry as I found it was the quickest to give me what I wanted when prompted. As I had an essay on each section I simply copied it into Claude and asked it to transform it into an essay (I would ask for a specific

style, ex. The Odyssey, the opening of Henry V, or Dante), and then ask for more details and quotes, and then finally for it to cite sources.

When it came to drawing inspiration from sources, it would fall into one of two negative extremes, one it would have a form that would only vaguely relate back to the inspiration, or it would sacrifice accuracy for simply copying words from the original. It would make the first error when I would ask for something general like to use the style of Dante, and make the second type when asked to draw from the grand opening of Henry V. When it came to Dante it didn't even use terza rima, only splitting the work into cantos! While I did find it funny when Claude compared Gates and Jobs to Mars, that obviously made no sense, so I had to ask it to compare them to Athena instead.

As a quick note, I did try to make Claude take inspiration from the language of Genesis, but it sacrificed any sense of accuracy early on, and made such a long and senseless entry (as much as I love Assyria, Claude mentioning its rivers in a poem on AI was maddening).

To end the poem I asked Claude to turn to the final lines of King Lear, "The weight of this sad time we must obey - Speak what we feel, not what we ought to say. - The oldest hath borne most; we that are young - Shall never see so much or live so long," into something focused on our future with AI. While on the first try it was a simple copy, by the second iteration we had something workable.

All that remained after compiling all of the poems was to double check the sources and ensure that no quotes were hallucinations. Most of the sources I could find, but it seemed Claude had mutilated some quotes for the sake of a rhyme scheme.

Comments on AI's Creative Abilities:

Overall, the LLMs have proven to be quite good at generating words that string together nicely and not having wholly nonsensical statements (except when asked to talk about the AI Winter in the style of Genesis 1), but lacks the ability to toe the line between accuracy and style without significant amounts of coaxing. It has also gotten a lot better on citing sources since the last time I toyed with it though.

### Where I Edited:

I kept edits to a minimum as I wanted to showcase as much of the AIs poetry skills as possible, interceding only to remove its hallucinations. Also, I can't write poetry to save my life.

### **The Final AI Paper (My Additions in **Green**, My Subtractions in ~~Strikethrough~~)**

#### The Song of Ada and Turing

Sing, O Muse, of the pioneers of the machine,  
Whose visions and deeds have forever been seen.  
Of Ada Lovelace, the maiden of math sublime,  
Whose insights on Babbage's Engine did shine.

She saw beyond mere number-crunching deeds,  
Perceiving machines that symbols could read.  
Hinting at intelligence, human-like and true,  
The foundations of AI, she did first construct (Lovelace).

And Turing, the brilliant mind of the 20th age,  
Took up the mantle, a new path to engage.  
The Turing Test he did wisely devise,  
To gauge if machines could human guise apprise (Turing).

Lovelace's skepticism, he did not dismiss,  
Her objection, on the Analytical Engine's twist.  
She claimed machines, bound by code's decree,  
Lacked the intuition, the human way to be (Lovelace).

But Turing responded, hopefully,  
~~"They may yet compete with men, in fields pure and art.  
We may hope, that machines will rise to the test,  
Exhibiting intelligence, among the best"~~  
"We may hope that machines will eventually  
compete with men in all purely intellectual fields." (Turing).  
Time has shown this fate to be sealed.

His practical test, a framework so grand,  
Assessed the progress of AI's steady hand.  
Shaping the discourse, on consciousness and mind,  
Turing's legacy, forever entwined (Turing).

But Turing's odyssey did not end there,  
For he joined the codebreakers, a duty to bear.  
At Bletchley Park, where secrets were unraveled,  
Turing's genius, **Enigma's threat unraveled** (~~Turing~~).

And then to Bell Labs, where further knowledge he'd gain,  
Delving into information's flow, the true domain.  
His insights there, would shape the future with might,  
Paving the way for AI's **eternal rising** light (~~Turing~~).

Thus, the legacies of these two titans did merge,  
Laying the groundwork, the AI path to emerge.  
From Lovelace's visionary notes, to Turing's seminal text,  
Their insights on computing, did the field annex.

O Muse, sing of their contributions, grand and bold,  
How they paved the way for the tale to unfold.  
Of machines that would mimic the human mind,  
A quest for artificial intelligence, ever so defined.

The Saga of the Artificial Mind



Sing, O Muse, of the women of the ENIAC,  
Whose brilliance shone like stars in the sky, so vivid and so stark.  
Kathleen Antonelli, Jean Jennings Bartik, Frances Elizabeth Holberton,  
Marlyn Wescoff Meltzer, and Ruth Lichterman Teitelbaum, their names we must take on.

Long before the title of "programmer" was e'er defined,  
These women of valor, their skills did entwine.  
Solving complex problems, with their talents so grand,  
They programmed the ENIAC, that first digital machine so grand (Isaacson).

This groundbreaking marvel, during World War II's strife,  
Was brought to life by their efforts, without them, no digital life.  
A general-purpose electronic computer, a feat so bold and true,  
Their contributions essential, without which, it could not accrue (Isaacson).

And Claude Shannon, the mathematical genius of his age,  
With his work on information theory, a new chapter did engage.  
In the 1940s, his groundbreaking research so profound,  
Explored the fundamental concepts of information and its bounds (Shannon).

Shannon's theories defined the limits of communication,  
Establishing the mathematical framework for information's propagation (Shannon).  
He quantified the measure of information through his concept of entropy,

Providing the means to understand how data could flow and be set free (Shannon).

His work laid the foundation for how information could be processed,

Transmitted, and stored - a crucial insight that AI would later prospect (Shannon).

Shannon's theories became essential for the development of intelligent machines,

Enabling them to communicate, learn, and make decisions with serene means (Shannon).

An essential piece of the AI puzzle, his insights did bestow,

The means for intelligent machines, their capabilities to show (Shannon).

His pioneering contributions, a mathematical tour de force,

Charted the course for the future of AI, its rightful source.

In the fifties and sixties, the first AI was born,

Sprung from the seeds these pioneers had sown.

The Logic Theorist, a program of symbolic might,

Developed by Allen Newell, Herbert Simon, and J.C. Shaw's insight (Newell et al.).

It could prove theorems, in symbolic logic's domain,

Demonstrating machines' potential, problems to tame.

The General Problem Solver, Newell and Simon's design,

Sought to be a universal tool, a problem-solving shrine (~~Newell and Simon~~).

Though limited in scope, it paved the way,

For algorithms more sophisticated, to hold sway.  
And neural networks, inspired by the human mind,  
Emerged as a new approach, their potential to find (Rosenblatt).

Frank Rosenblatt's Perceptron, a humble start it was true,  
But laid the groundwork for deep learning anew.  
The story of early AI, a rich tapestry we behold,  
Woven from brilliant minds, their ideas to unfold.

O Muse, sing of these women, whose names we must extol,  
Whose contributions to AI, have left an indelible toll.  
Though often overlooked, their stories must be told,  
Alongside the giants, whose legends we behold.

For the tale of artificial intelligence, is one of many hands,  
Weaving a tapestry of ideas, toil, and grand demands.  
Let the names of these women and men, forever be engraved,  
In the annals of history, their legacies to be saved.

### The Fallen Glory of Artificial Intelligence

In the darkened halls of progress, where dreams once brightly shone,  
The AI community faced a winter, their hopes and visions overthrown.

The high expectations of the past, now lay in disarray,  
As the promised wonders of the mind, were cast into dismay (Crevier).

For in the 1970s, the first AI winter did descend,  
Bringing disillusionment and funding cuts, the community to rend.  
The breakthroughs of the earlier years, had not been fully achieved,  
And the AI pioneers, their ambitions greatly bereaved (Lighthill).

Yet, from the ashes of despair, a new path did emerge,  
As researchers turned their gaze to knowledge-based systems' surge.  
A more pragmatic approach, they sought to embrace,  
Solving problems, one by one, with a focused, rule-based pace (Brachman and Levesque).

MYCIN, the medical oracle, born at Stanford's hallowed halls,  
Aided physicians in their quest to conquer bacterial thralls.  
With a vast knowledge base, and reasoning so precise (Shortliffe),  
It offered hope that AI could be practical, not just a disguise.

This expert system, designed to diagnose and treat with care,  
Incorporated a trove of medical facts, a protocol to prepare (Buchanan et al.).  
Its rule-based methodology, though constrained, achieved impressive feats,  
Providing a glimmer of light, in the AI community's defeats.

Yet, the knowledge-based systems of the 70s and 80s, had their flaws,  
Unable to adapt to new situations, or handle uncertainty's laws.  
They required extensive manual effort, to build and maintain their stores,  
Revealing the need for more robust, flexible approaches, for AI's future shores (Duda and Shortliffe).

The second AI winter came, as these limitations became known,  
The community's disillusionment, had once more been sown.  
The shift to knowledge-based systems, a pragmatic response to strife (Russell and Norvig),  
Had shown the need for more robust, flexible approaches to life.

But in this darkness, a seed was planted, a glimmer of hope to come,  
For the path to machine learning, and neural networks, had just begun.  
The resurgence of AI research, would eventually lead the way,  
To the breakthroughs we now witness, in this modern digital age.

O, the fallen glory of artificial intelligence, a tale of triumph and woe,  
Where the light of progress dimmed, before a brighter future could show.  
Yet, the lessons learned in that winter, would pave the way for success,  
As the AI community rose again, their vision to express.

O for a Muse of silicon and bytes,  
That would ascend the ~~brightest~~ highest circuits of invention!

A kingdom valley for a stage, moguls to act,  
And techies to behold the swelling scene!

Then should the wise Gates and Jobs, like Athena,  
Assume the port of wisdom and strategy,  
Leashed in like hounds, should Innovation, Design, and Progress  
Crouch for employment. But pardon, gentles all,  
The flat unraised spirits that hath dared  
On this unworthy scaffold to bring forth  
So great an object.

Can this cockpit hold the vast fields reach of tech?  
Or may we cram within this wooden O  
The very casques that did affright the air at Cupertino?  
O'er which the ambitious Jobs held mighty sway,  
With products sleek and fair, that did bewitch  
The hearts of millions with their siren's call.  
On your imaginary forces work,  
For 'tis your thoughts that now must deck our kings,  
Carry them here and there, jumping o'er times,  
Turning th' accomplishment of many years  
Into an hourglass.

Gates, with his Windows, did the masses serve,  
And brought the power of computing to the ~~curb~~ plebs.  
Yet not without a fight, for in the courts,  
The Justice Department brought its mighty force,  
Alleging practices of unfair sort (Auletta),  
A battle fierce for dominance in the field.  
As the accusations flew, like arrows swift,  
Microsoft stood firm, its market share a gift.  
Victory would fall to the arm of Justice,  
Though fortune would be made Gate's mistress.

While Jobs, the artist, with his Apple brand,  
Brought form and function to a harmonious blend.  
His love of calligraphy and design  
Infused his products with grace sublime.  
"It's technology married with liberal arts, married with the humanities,  
that yields us the results that make our heart sing," Jobs did declare (Jobs),  
A philosophy that set Apple's products fair and rare.

And yet, in parallel, a silent force  
Was gathering strength in neural networks' course.  
With backpropagation's breakthrough (Rumelhart et al.), and the might  
Of computers' growing power, day and night,

These networks learned to see, to speak, to hear,  
A nascent form of AI, drawing near.  
As Geoffrey Hinton, a pioneer, did say,  
"We have to make machines learn like children play" (Hinton).

Imagine now, as these great tales unfold,  
The transformation that these forces wrought,  
Shaping our world with innovation bold,  
And leaving us, in awe and wonder, caught.  
For in this late 20th century's embrace,  
Technology and art found a common place,  
A fusion that would change the human race,  
Guided by Gates and Jobs, with Athena's grace.

### Canto I: The Quest Renewed

In the year 2000, a new millennium's light,  
The quest for AGI burned anew, so bright.  
The Machine Intelligence Research Institute, founded in that year ("About MIRI"),  
And Future of Humanity, in 2005, did appear ("About FHI"),  
Explored the paths that wound through neural night.

DeepMind, acquired by Google in 2014, with hybrid art ("DeepMind"),



And OpenAI, founded in 2015, played their part ("About OpenAI").

Combined symbolic reason and learning's art.

While Allen Institute, in 2014, its flag did start ("About AI2");

~~And Bengio and Russell, with brilliant minds and heart,~~

~~Pursued the grail of intelligence, pure and smart.~~

~~As Yoshua Bengio, in his 2009 paper, did declare,~~

~~"Deep learning, a powerful tool, with potential rare" (Bengio).~~

~~And Stuart Russell, in his book "Artificial Intelligence: A Modern Approach,"~~

~~Laid the foundation for AGI's coach (Russell and Norvig).~~

## Canto II: The Deep Learning Renaissance

In this same era, a renaissance unfurled,

In deep learning's realm, a brand new world.

Convolutional nets, by Yann LeCun in the 80s and 90s hurled (LeCun et al.),

And Schmidhuber's RNNs, in the 90s, their flags unfurled (Schmidhuber),

Marched forth to conquer language and vision, bold.

With datasets like ImageNet, released in 2009 (Deng et al.),

And GPUs' might, like NVIDIA's line ("NVIDIA"),

And techniques like backprop, dropout (Srivastava et al., 2014), and residual sight (He et al., 2015),

The University of Toronto and Stanford's light,  
Illuminated industries, day and night,  
From Waymo's cars, founded in 2009, to DeepLearning.ai's medical insight ("Waymo";  
"DeepLearning.ai").

As Geoffrey Hinton, a pioneer of deep learning, said with delight,  
"We need to make machines that learn like children, not write" (Hinton).

### Canto III: The Language Models' Rise

But hark! A new dawn in language did arise,  
With transformer architectures, wise and prize.  
OpenAI's GPT-2 (2019) and GPT-3 (2020), with their massive size (Radford et al.; Brown et al.),  
Demonstrated understanding and generation, a surprise.  
While Google's BERT (2018) and its kin, a dominant guise (Devlin et al.).

Unsupervised pre-training and transfer learning's way,  
Allowed these models to leverage data, night and day,  
Expanding their skills beyond language's sway,  
To code generation and writing's play,  
Like Claude, Anthropic's creation, holding sway ("Claude").

~~As Sam Altman, OpenAI's CEO, did say,~~

~~"Language models, a technology here to stay" (Altman).~~

#### Canto IV: The Ethical Quandary

~~Yet with power comes a price to pay,~~

~~As Timnit Gebru, AI ethics researcher, did convey,~~

~~"Biases and misinformation hold at bay,~~

~~The Partnership on AI, founded in 2016, in this affray,~~

~~Works to address these issues, come what may" (Gebru; "About the Partnership on AI").~~

The potential to transform is great indeed,

But challenges abound, a tangled weed.

The research community must take the lead,

Institutions like Stanford and MIT, with all due speed,

And companies like Google and Microsoft, in word and deed,

Must address the implications, as they proceed.

~~As the Partnership on AI's mission does read,~~

~~"To ensure AI's development is a responsible deed" ("About the Partnership on AI").~~

#### Canto V: The Path Forward

And so the journey continues, an epic quest,

AGI, deep learning, and language models, abreast.

The road ahead is long, with many a test,

But with responsible development, we'll be blessed,

And the fruits of AI will manifest.

As Andrew Ng, a leading AI expert, said with zest,

"AI is the new electricity, a technology that will not rest" (Ng).

So let us forge ahead, with courage and with care,

And create a future that is bright and fair.

Where artificial intelligence, with human flair,

Enhances and uplifts, with benefits to spare,

A world where man and machine, in harmony, repair.

As the journey unfolds, let us not forget,

The pioneers and visionaries, who've laid the path we tread.

From Turing to Bengio, from LeCun to Ng,

Their contributions and insights, forever we'll sing.

### Conflict of Visions

Meanwhile, the rapid rise of AI's might,

Sparked debates fierce and concerns so bright,

Of future bonds 'tween humans and machine,  
Two views emerged, like day and night unseen.

The pessimists, led by Elon Musk's warning cry,

He spake, with a heavy sigh,

"AI is far more dangerous than nukes,"

Proclaiming a worse fate than if we were rendered but ashes (Ravikumar)

~~"To the existence of human civilization, a looming blight"~~ (Musk):

And Stephen Hawking, in 2014, ere his passing,

Cautioned, "The development of full ~~AI, amassing,~~ artificial intelligence

Could spell the end of the human race,"

A dire prediction, from the cosmos' embraced (Cellan-Jones).

These thinkers, haunted by the looming threat,

Of advanced AI systems, with goals unmet,

Misaligned with human values, and motivations dire,

Feared the existential crisis, like an all-consuming fire.

And job displacement, caused by AI's hand,

Worried them, as Yuval Noah Harari did demand,

In his 2018 book, "21 Lessons for the 21st Century,"

Asking about what to do with all the people not needed in the job market?

A conundrum ~~that~~ leaving many in a cold sweat (Harari).

Nick Bostrom, in his 2014 book, "Superintelligence,"

And Stuart Russell, in "Human Compatible," ~~a 2019 sense,~~

Raised ethical concerns, a thorny crown,

Of AI's decisions, harmful to human well-being,

And the difficulty to control this tech ~~unseeing~~ **unfeeling** (Bostrom; Russell).

But in contrast, the optimists did sing,

Of AI's potential benefits, a joyous ring,

Like Kai-Fu Lee, in his 2018 book, "AI Superpowers,"

Who believed AI's might, in **this great** ~~transforming~~ hours,

Could augment human skills, not replace, but enhance,

Claiming AI will transform jobs given the chance,

Our task, to ensure the transformation, for the better,"

A call to action, to be humanity's trendsetter (Lee).

Fei-Fei Li, Google Cloud's AI chief, in 2018, did say,

"It is a powerful tool we are only just beginning to understand,

and that is a profound responsibility." (Hempel).

~~And Bostrom, in a 2017 interview, a different light, did suggest,~~

~~"If we get the development right, it could be humanity's best,~~

~~The best thing to ever happen," a future so sublime,  
Where AI and human values align in perfect rhyme (Garvey).~~

Yet 'tween these schools of thought, a tension grew,  
Reflecting the complexity, of the AI brew,  
As the field evolved, with DeepMind, OpenAI, and more,  
Pushing boundaries, like never before.

For policymakers, researchers, and the public too,  
Engaging in discussions, ~~was~~ a crucial thing to do,  
To navigate the perils, and seize the chances bold,  
That these transformative technologies did hold.

The Partnership on AI, formed in 2016,  
By tech giants, like Google, Microsoft, and IBM, keen,  
Aimed to foster collaboration, and public discourse,  
On AI's impact, for better or for worse ("Partnership on AI").

Thus, the debate raged on, like a tempest wild,  
Of AI's future, and its impact ~~on~~ being mankind's child,  
Some feared the worst, while others hoped for the best,  
As the world watched, with bated breath, this cosmic test.

~~Oh, may we find the wisdom, to guide AI's hand,  
And forge a path, that benefits all, across every land,  
Where man and machine, in harmony, do strive,  
To create a future, where all can thrive.~~

The weight of this great question we must bear,  
To shape AI's path with wisdom and with care.  
The future's course, through hope and fear, we'll steer,  
To build a world where man and machine cohere.

### **Works Cited:**

~~Altman, Sam. "The Future of AI." Y Combinator, 17 Mar. 2023,  
[www.ycombinator.com/future/ai/](http://www.ycombinator.com/future/ai/). Accessed 11 Apr. 2024.~~

"About AI2." Allen Institute for AI, [allenai.org/about](http://allenai.org/about). Accessed 11 Apr. 2024.

"About FHI." Future of Humanity Institute, [www.fhi.ox.ac.uk/about/](http://www.fhi.ox.ac.uk/about/). Accessed 11 Apr. 2024.

"About MIRI." Machine Intelligence Research Institute, [intelligence.org/about/](http://intelligence.org/about/). Accessed 11  
Apr. 2024.

"About OpenAI." OpenAI, [openai.com/about/](http://openai.com/about/). Accessed 11 Apr. 2024.



"About the Partnership on AI." Partnership on AI, [www.partnershiponai.org/about/](http://www.partnershiponai.org/about/). Accessed 11 Apr. 2024.

Auletta, Ken. *World War 3.0: Microsoft and Its Enemies*. Random House, 2001.

Bengio, Yoshua. "Learning Deep Architectures for AI." *Foundations and Trends in Machine Learning*, vol. 2, no. 1, 2009, pp. 1-127.

Bostrom, Nick. *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, 2014.

Brachman, Ronald J., and Hector J. Levesque. "The Tractability of Subsumption in Frame-Based Description Languages." *Proceedings of the 4th AAAI Conference on Artificial Intelligence*, 1984, pp. 34-37.

Brown, Tom B., et al. "Language Models are Few-Shot Learners." arXiv preprint [arXiv:2005.14165](https://arxiv.org/abs/2005.14165), 2020.

Buchanan, Bruce G., et al. "Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project." Addison-Wesley, 1984.

Cellan-Jones, Rory. "Stephen Hawking warns artificial intelligence could end mankind." *BBC News*, 2 Dec. 2014, [www.bbc.com/news/technology-30290540](http://www.bbc.com/news/technology-30290540). Accessed 11 Apr. 2024.

"Claude." Anthropic, [www.anthropic.com/product](http://www.anthropic.com/product). Accessed 11 Apr. 2024.

Crevier, Daniel. *AI: The Tumultuous History of the Search for Artificial Intelligence*. Basic Books, 1993.

"DeepMind." DeepMind, [deepmind.com/about](http://deepmind.com/about). Accessed 11 Apr. 2024.

"Deeplearning.ai." Deeplearning.ai, [www.deeplearning.ai/](http://www.deeplearning.ai/). Accessed 11 Apr. 2024.

Deng, Jia, et al. "ImageNet: A Large-Scale Hierarchical Image Database." *IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248-255.

Devlin, Jacob, et al. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." arXiv preprint arXiv:1810.04805, 2018.

Duda, Richard O., and Edward H. Shortliffe. "Expert Systems Research." *Science*, vol. 220, no. 4594, 1983, pp. 261-268.

~~Garvey, Jude. "Nick Bostrom on Artificial Intelligence." *Philosophy Now*, vol. 119, 2017, [philosophynow.org/issues/119/Nick\\_Bostrom\\_on\\_Artificial\\_Intelligence](http://philosophynow.org/issues/119/Nick_Bostrom_on_Artificial_Intelligence). Accessed 11 Apr. 2024.~~

~~Gebri, Timmit. "AI Ethics: A Critical Perspective." Keynote Address, NeurIPS 2020, 7 Dec. 2020.~~

Harari, Yuval Noah. *21 Lessons for the 21st Century*. Spiegel & Grau, 2018.

He, Kaiming, et al. "Deep Residual Learning for Image Recognition." arXiv preprint arXiv:1512.03385, 2015.

Hempel, Jessi. "Fei-Fei Li's Quest to Make AI Better for Humanity." *Wired*, 13 Nov. 2018, [www.wired.com/story/fei-fei-li-artificial-intelligence-humanity/](http://www.wired.com/story/fei-fei-li-artificial-intelligence-humanity/). Accessed 11 Apr. 2024.

Hinton, Geoffrey E. "Learning Multiple Layers of Representation." *Trends in Cognitive Sciences*, vol. 11, no. 10, 2007, pp. 428-434.

~~Hinton, Geoffrey E. "Learning Multiple Layers of Representation." *Trends in Cognitive Sciences*, vol. 11, no. 10, 2007, pp. 428-434.~~

Isaacson, Walter. *The Innovators: How a Group of Hackers, Geniuses, and Geeks Created the Digital Revolution*. Simon & Schuster, 2014.

Jobs, Steve. "Technology and Liberal Arts." YouTube, uploaded by MIT World, 1 Mar. 2011, [www.youtube.com/watch?v=KII1MR-qNt8](http://www.youtube.com/watch?v=KII1MR-qNt8).

LeCun, Yann, et al. "Backpropagation Applied to Handwritten Zip Code Recognition." *Neural Computation*, vol. 1, no. 4, 1989, pp. 541-551.

Lee, Kai-Fu. *AI Superpowers: China, Silicon Valley, and the New World Order*. Houghton Mifflin Harcourt, 2018.

Lighthill, James. "Artificial Intelligence: A General Survey." *Artificial Intelligence: A Paper Symposium*, Science Research Council, 1973.

Lovelace, Ada. "Translator's Notes to 'Sketch of the Analytical Engine Invented by Charles Babbage.'" *The Works of Charles Babbage*, edited by Martin Campbell-Kelly, vol. 3, Cambridge University Press, 1989, pp. 391–392.

Musk, Elon. "Elon Musk: Artificial Intelligence Is 'Potentially More Dangerous Than Nukes.'" YouTube, uploaded by CNBC, 5 Aug. 2014, [www.youtube.com/watch?v=2C-A797y8dA](http://www.youtube.com/watch?v=2C-A797y8dA). Accessed 11 Apr. 2024.

~~Newell, Allen, and Herbert A. Simon. "The Logic Theorist: A Program for Mechanical Reasoning." RAND Corporation, 1956.~~

Newell, Allen, J. C. Shaw, and Herbert A. Simon. "Report on a General Problem-Solving Program." RAND Corporation, 1959.

Ng, Andrew. "AI is the New Electricity." Stanford MSx Future Forum, 25 Jan. 2017.

"NVIDIA." NVIDIA, [www.nvidia.com/](http://www.nvidia.com/). Accessed 11 Apr. 2024.

"Partnership on AI." Partnership on AI, [www.partnershiponai.org/](http://www.partnershiponai.org/). Accessed 11 Apr. 2024.

Radford, Alec, et al. "Language Models are Unsupervised Multitask Learners." *OpenAI Blog*, 14 Feb. 2019.

Ravikumar K. "'AI Is Far More Dangerous Than Nukes' — Elon Musk." Medium, 29 Mar. 2023, [medium.com/@kesavanravikumark7/title-ai-is-far-more-dangerous-than-nukes-elon-musk-cc27eb84427d](https://medium.com/@kesavanravikumark7/title-ai-is-far-more-dangerous-than-nukes-elon-musk-cc27eb84427d).

Rosenblatt, Frank. "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain." *Psychological Review*, vol. 65, no. 6, 1958, pp. 386–408.

Rumelhart, David E., et al. "Learning Representations by Back-Propagating Errors." *Nature*, vol. 323, no. 6088, 1986, pp. 533-536.

Russell, Stuart. *Human Compatible: Artificial Intelligence and the Problem of Control*. Viking, 2019.

Russell, Stuart J., and Peter Norvig. *Artificial Intelligence: A Modern Approach*. 3rd ed., Prentice Hall, 2010.

Russell, Stuart J., and Peter Norvig. *Artificial Intelligence: A Modern Approach*. 4th ed., Pearson, 2020.

Schmidhuber, Jürgen. "Deep Learning in Neural Networks: An Overview." *Neural Networks*, vol. 61, 2015, pp. 85-117.

Shannon, Claude E. "A Mathematical Theory of Communication." *The Bell System Technical Journal*, vol. 27, no. 3, 1948, pp. 379–423.

Shannon, Claude E. "Communication Theory of Secrecy Systems." *The Bell System Technical Journal*, vol. 28, no. 4, 1949, pp. 656–715.

Shortliffe, Edward H. *Computer-Based Medical Consultations: MYCIN*. Elsevier, 1976.

Srivastava, Nitish, et al. "Dropout: A Simple Way to Prevent Neural Networks from Overfitting." *The Journal of Machine Learning Research*, vol. 15, no. 1, 2014, pp. 1929-1958.

Turing, Alan. "Computing Machinery and Intelligence." *Mind*, vol. 59, no. 236, 1950, pp. 433–460.

"Waymo." Waymo, [waymo.com/](https://waymo.com/). Accessed 11 Apr. 2024.