

Cole Pegula
Walter Isaacson
Digital Revolutions
April 14, 2024

From Turing to AI: The Evolution of “The Last Worker”

Method:

When I approached this paper, I had to face the numerous shortcomings of current LLM chatbots. While simply imputing a prompt and getting the desired result sounds like an ideal scenario, this is not often the case when attempting to write a thorough research paper in prose. I decided to use ChatGPT 4 as my primary tool for writing this essay. While GPT 4 thrives in solving mathematical problems, providing concise answers to confusing questions, drafting outlines, and approaching programming projects, it has many difficulties with immediately producing a usable paper from a single prompt. Thus, my journey with prompt engineering and fine tuning began in the process of writing this paper. I began with the given universal prompt, “Describe the development of artificial intelligence from Turing to large language model chatbots,” as well as including the constraints of “in 2,500 to 3,000 words and in the form of research paper.” This produced an outline instead of the paragraph prose I desired. The outline was solid and coherent and would be a good starting point if I were attempting to write and research this paper myself.

I kept trying to push the model to write the paper with the word count constraints, but it insisted that “Given the intricate subject matter and the detailed requirements for this paper, a comprehensive draft can't be fully developed in a single response.” In order to comply with GPT's stubbornness, I compromised and asked it to write the paper section by section in chronological order, while asking me if I want to continue to the next section in another response at the end of the chat response. I stated this as a reply to its concerns saying, “then do each segment in each reply. I will say continue. You will make sure that each response adds up to

3000 words.” GPT then produced a first draft in a series of responses that felt decent but left a lot to desire. I kept pushing GPT to go more in depth, while maintaining a formal yet entertaining prose. The latter was not quite evident in GPT’s responses. I noticed that it would introduce historical information but in a basic fashion that lacked the thoroughness I required from this paper. Ultimately, I was left with two drafts from GPT that seemed like a good starting point to go further in depth.

After troubleshooting some more, I decided to take my GPT 4 drafts and input it into Claude 3. I asked Claude 3 to first “combine these two papers into one cohesive research paper in 2500-3000 words in a narrative prose. Make it less bland to read.” Claude produced a thorough response that almost met the word count, but it was clearly unfinished. I replied, “continue” through two additional responses, and it finally produced a cohesive paper. I read through the paper and it felt solid, but lacked some elements that I felt were crucial to my rendition of this research paper. I wanted it to expand upon the psychological and ethical implications of AI and the certain figures involved in pioneering artificial intelligence, particularly less represented voices like the women behind it. I felt like it was getting to a good place, but then I realized that I forgot to prompt it to format the paper in Chicago formatting with the necessary and accurate citations. When asking Claude to perform this task, I was met with an unfortunate paywall. Claude said that my free plan would hinder it from performing such a task with an extensive word count. After trying to get around it with partitioned responses, I realized I could not evade the remaining responses I had left in my free plan, which quickly approached 0. I decided to give in and purchase the premium version of Claude featuring its most intelligent model, Claude Opus. If anything, I was looking for an excuse to try this model anyways. I asked Claude Opus to reformat the paper in Chicago Style, while including footnotes for each piece of

information that required citations. Claude proceeded to spit out the same paper but reformatted in Chicago style with numbered references to each citation in a separate works cited page. I believe it intended to create footnotes, but due to copy and paste limitations these didn't transfer over. I had to manually add each footnote to each quote in the paper. I went through the paper, cross referencing the works cited citations and adding in each footnote. It would be very nice if there was an extension for Microsoft Word that would be able to automate this process using AI. When reviewing the final paper, I wasn't entirely satisfied. Claude 3, Opus, and GPT 4 were able to craft a detailed and cohesive paper that included all the landmark events in artificial intelligence's developmental timeline. However, I felt like there was something missing. It also seemed a bit too long. I then reviewed to see what I could do to cut it down to a more concise and easier to read paper while maintaining its rich details.

While reviewing, I noticed a series of factual errors that Claude must have hallucinated. These included names that did not exist and Claude attributed them to making bizarre and complex contributions to AI development. I noticed that most of this applied to the women's role section, and I ultimately decided to cut it because it would not produce accurate information as much as I tried. There was also another section on AGI, that was completely based in theoretical information. Claude assumed that we had reached full AGI development and deployment, and it even said that AGI models had unleashed full scale cyberattacks and political misinformation campaigns. I thought this was incredibly interesting, but of course, it was not true (let's hope!), and I had to cut it. I also realized it went on an entire tangent of false theoretical information about AI taking over the world, replacing humans from the workforce, and framing it all in a "historical" context as if it had already happened. In fact, it wrote an entire section on the 21st century's third decade. It was quite alarming, but also a fun science fiction story that does seem

to have the possibility to come to fruition. The section included many fun moments like stating that, “Machine planners choreographed robotic spacecraft for ever more complex and ambitious interplanetary missions. Artificial intelligence was a critical enabler for realizing the next generation of energy breakthroughs at facilities like CERN and the future multi-trillion parameter models powering fusion reactor design.” I especially enjoyed the line, “In uplifting artificial minds into hauntingly familiar sentience, we've opened a Promethean Pandora's box humanity is intellectually ill-equipped to govern.”

To finalize my paper, I took my manual edits and inputted the paper into the Grammarly platform. Grammarly leverages AI to make spelling, grammar, and clarity edits to writing. I was also able to validate the sources in the paper through this tool. Ultimately, my process of creating this paper revealed to me the shortcomings of current AI chatbot models while also showcasing their immense power. Ultimately, it takes heavy troubleshooting and prompt engineering to really get them to produce quality work. I see them as better suited towards creating the foundation for research papers like this one. Through my endeavors, I find that GPT is better at creating an outline than a nuanced piece of prose.

First Response:

“Creating a comprehensive research paper on the development of artificial intelligence (AI) from Alan Turing's foundational concepts to the modern era of large language model chatbots encompasses several significant milestones. This outline will guide you through the critical points and developments in AI history, highlighting Turing's contributions, the evolution of computing and AI research, the emergence of neural networks and deep learning, and the advent of large language model (LLM) chatbots like OpenAI's GPT series.

Introduction

- **Brief Overview:** Define AI and its goal of simulating human intelligence processes through machines.
- **Thesis Statement:** The evolution from Turing's theoretical foundations to today's LLM chatbots illustrates remarkable technological advancements, showcasing AI's potential to transform various aspects of human life.

Alan Turing and Theoretical Foundations

- **Alan Turing's Contributions:** Discuss Turing's role in the early conceptualization of computing and AI, including the Turing Machine and the Turing Test.
- **The Turing Test:** Explain its significance in evaluating a machine's ability to exhibit intelligent behavior equivalent to, or indistinguishable from, that of a human.

Early AI Research and Development

- **1950s and 1960s – The Dawn of AI:** Cover the establishment of the field of AI, the first AI programs, and the optimism of early researchers.
- **1970s and 1980s – AI Winter:** Discuss the challenges and setbacks faced by AI research, leading to periods known as "AI winters," where funding and interest in AI research declined.

The Rise of Machine Learning

- **Machine Learning Foundations:** Introduce the shift from rule-based systems to machine learning, where AI systems learn from data.
- **Key Developments:** Highlight significant algorithms and models that advanced machine learning, such as decision trees, neural networks, and support vector machines.

Neural Networks and Deep Learning

- **Revival of Neural Networks:** Explain the resurgence of interest in neural networks and the development of deep learning, emphasizing key breakthroughs like backpropagation.
- **Deep Learning Achievements:** Detail milestones achieved through deep learning in image recognition, natural language processing, and other domains.

The Era of Large Language Models

- **Introduction to LLMs:** Define large language models and their significance in advancing AI capabilities, particularly in understanding and generating human-like text.
- **From GPT to GPT-4:** Trace the development of OpenAI's Generative Pre-trained Transformer (GPT) series, noting improvements in size, complexity, and applications.
- **Impact and Applications:** Discuss the wide-ranging impact of LLMs on industries, education, and creative fields, including chatbots, content generation, and more.

Ethical Considerations and Future Directions

- **AI Ethics:** Address ethical concerns surrounding AI, such as bias, privacy, and the potential for misuse.
- **Future of AI:** Speculate on future advancements, challenges, and the evolving relationship between humans and AI.

Conclusion

- **Recap of Major Points:** Summarize the key developments from Turing's contributions to modern LLM chatbots.
- **Reflection on Progress:** Reflect on the rapid advancement of AI technology and its implications for the future.

References

- Include a comprehensive list of sources, such as academic papers, books, and reputable online resources.

This outline provides a structured approach to writing a research paper on the development of AI from Turing to modern chatbots. Each section builds upon the last to depict the progressive evolution of the field, highlighting the technological advancements, key milestones, and the ongoing challenges and ethical considerations.”

Final Paper:

The Narrative of Intelligent Machines: From Flickering Thoughts to Autonomous Realities

Cole Pegula (and GPT4 + Claude 3 + Claude Opus)
HISU 4100: Digital Revolutions
Walter Isaacson
April 14, 2024

A Seed Is Planted

In the annals of human discovery, certain minds blaze forth with visions so profound, so ahead of their era, that their ideas transcend the mortal constraints of the age they inhabited. Such was the case with Charles Babbage, the 19th-century polymath whose analytical engine, though never entirely constructed, effectively outlined the fundamental architecture of a digital, programmable computer.¹

His contemporary, the brilliant Ada Lovelace, the Countess of Controversy, added a philosophical dimension. In her seminal work translating an article on Babbage's prototypes, Lovelace laid out a striking objection that would foreshadow debates raging centuries later – that no matter how sophisticated, machines could never truly create in the same vein as the human mind. They could not "originate anything" but merely reaped the fruits of human ingenuity.² An artificial intellect, Lovelace asserted, would be inherently limited.

Little could these pioneers have foreseen their musings would become the tectonic rumblings of a revolution that re-envisioned the very nature of intelligence itself. The seed they planted, born of flickering thoughts and abstruse theories, would blossom into a technological Renaissance of almost unfathomable scope. An age when humanity would bring forth intelligence anew, crafting them not of flesh and soul but of silicon and code.

The Inception: Alan Turing Stakes His Claim

The 20th century found the seed sprouting from fertile theoretical soil, cultivated by a young British mathematician of unparalleled genius. Armed with uncanny prescience, Alan Turing understood that the increasingly automated world around him was not the harbinger of machinery's apex but its adolescence. In his seminal texts – "On Computable Numbers" and "Computing Machinery and Intelligence" – Turing outlined a vision that upended two millennia of philosophical conventions.^{3 4}

At the core of his work was a deceptively simple syllogism: If human sapience and reasoning capabilities arose as biological phenomena, could they not be replicated through other physical processes and architectures? Could the human ingenuity that spawned machines one day birth artificial beings possessed of cognition, logic, and even consciousness itself? Impractical as the notion seemed, Turing insisted such "thinking machines" were feasible and inevitable.

His famed "Turing Test" provided a framework for validating machine intelligence – a watershed for artificial intelligence as a formal field of scientific study.⁵ Skeptics scoffed at its anthropomorphic implications, citing Lady Lovelace's objection that genuine machine autonomy was folly, an illusion derived from human inputs. But Turing understood something fundamental – that intelligence, in its quintessence, was an iterative process of inputs and outputs, of recursive self-refinement. If we could create simulated environments of sufficient depth and variety, intellects would emerge through exposure, interaction, and evolution.

The mathematical formalisms he outlined, constructs like the "Universal Turing Machine," illuminated a radical paradigm – that information itself, not physical embodiment,

¹ Babbage, Charles. "On the Mathematical Powers of the Calculating Engine." In *The Works of Charles Babbage*, edited by Martin Campbell-Kelly, 19–54. New York: University Press, 1989.

² Lovelace, Ada. "Notes by the Translator." In *Sketch of the Analytical Engine Invented by Charles Babbage*, by Luigi Federico Menabrea, 1–59. London, 1843.

³ Turing, Alan M. "On Computable Numbers, with an Application to the Entscheidungsproblem." *Proceedings of the London Mathematical Society* 42, no. 1 (1936): 230–65.

⁴ Turing, Alan M. "Computing Machinery and Intelligence." *Mind* 59, no. 236 (1950): 433–60.

⁵ *Ibid.*

was the true bedrock of reasoning.⁶ The binary world of ones and zeroes, with its recursive simplicity, paradoxically contained the seeds of unlimited complexity.

And so began AI's transfiguration from notional conjecture into a field of dedicated research with Turing as its presiding spirit – a new Promethean tradition of fashioning consciousness from the building blocks of logic and information theory.

1956: A Phoenix Ignited

While Turing's revelations incubated amid World War II's tumult, artificial intelligence's formal inception blossomed in the heady, halcyon days following the Allied victory. The Dartmouth Conference of 1956, organized by John McCarthy and attended by pioneering luminaries like Marvin Minsky, Nathaniel Rochester, and Claude Shannon, marked AI's inaugural moment.⁷

The spark ignited the modern AI age, a gathering of brilliant minds utterly convinced that recreating intelligence was not only possible but imminently achievable. Funded by the military establishment and captivated by the new frontier of computing, these scientists embraced an optimism emblematic of their era. If the physical laws governing the cosmos could be divined through mathematics, the biological laws governing reason and sentience could similarly be unraveled, abstracted, and modeled.

From this intellectual bonfire emerged the earliest AI programs – rudimentary by modern standards yet awe-inspiring for their time. The Logic Theorist, created by Newell, Shaw, and Simon, successfully mimicked human reasoning to prove mathematical theorems.⁸ Nathaniel Rochester's groundbreaking work with neural networks laid the structural foundations for deep learning's future awakening.⁹ And Joseph Weizenbaum's ELIZA, a seemingly innocuous chatbot employing simple pattern matching, sparked the first severe reflections on AI systems exhibiting artificially persuasive simulations of human behavior.¹⁰

Yet progress was incremental, and the path forward was daunting. While these Early Pioneers marveled at their machines' capacity for discrete tasks like playing checkers or solving word problems, more profound challenges remained elusive. Natural language processing, complex reasoning over multiple contexts, sensory perception—these hallmarks of biological intelligence proved their limits.

Akin to constructing a primitive hominid and expecting Socratic discourse, their knowledge architectures fundamentally constrained the AI systems of the 1950s and 60s. Built on handcrafted rules and symbolic logic, they lacked the mechanisms to self-adapt and generalize in nuanced ways. Like Lovelace's objection, they mirrored human-derived inputs more than genuine autonomy.

⁶ Turing, Alan M. "On Computable Numbers, with an Application to the Entscheidungsproblem." *Proceedings of the London Mathematical Society* 42, no. 1 (1936): 230–65.

⁷ McCarthy, John, Marvin L. Minsky, Nathaniel Rochester, and Claude E. Shannon. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence." *AI Magazine* 27, no. 4 (2006): 12–14.

⁸ Newell, Allen, J. C. Shaw, and Herbert A. Simon. "Empirical Explorations with the Logic Theory Machine: A Case Study in Heuristics." In *Computers and Thought*, 109–33. McGraw-Hill, 1963.

⁹ McCarthy, John, Marvin L. Minsky, Nathaniel Rochester, and Claude E. Shannon. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence." *AI Magazine* 27, no. 4 (2006): 12–14.

¹⁰ Weizenbaum, Joseph. "ELIZA—a Computer Program for the Study of Natural Language Communication Between Man and Machine." *Communications of the ACM* 9, no. 1 (1966): 36–45.

The AI Winters Set In

As revelations failed to match expectations, AI fell victim to its hype cycle. Disillusionment and constrained funding cratered progress through the 1970s and late 1980s—now called the "AI Winters." Disavowal from experts like MIT's Marvin Minsky, who asserted "artificial intelligence has been brutally overrated," cast a pall over the field. Once a darling of funding bodies, AI became a reduction of theory hunting for conceptual breakthroughs.

And yet, the harsh winters bred insights vital for AI's future flourishing. Denied the low-hanging victories of grandiose ambitions, these dormant eras forced more robust fundamentals. The knowledge representation problem – how information is structured and made accessible to machine intelligence – fostered ontologies, Bayes nets, and first strides toward probabilistic models that could handle uncertain data and competing contexts. Within this cold tundra of diminished expectations, the green shoots of a sustainable AI Renaissance began taking root.

The Renaissance Springs Forth

The 80s and 90s witnessed the gradual thaw, with innovations spearheaded by figures like Judea Pearl and Sebastian Thrun helping reignite progress through breakthroughs in Bayesian networks, hidden Markov models, and behavior-based robotics.¹¹ This awakening opened new frontiers in expert systems, speech recognition, and industrial applications.

But the most seismic advancements came through the rediscovery of neural networks and the murmurings of a novel technique—machine learning. Building on the biological inspiration of mathematicians like Walter Pitts and legacies like Frank Rosenblatt's Perceptron, new architectures transcended the symbolic and statistical techniques of the old.

The Deep Learning Renaissance

Driven by trailblazers like Geoff Hinton, Yoshua Bengio, and Yann LeCun, the neural network renaissance ushered in the "Deep Learning Revolution." Networks with iterative hidden layers could now approximate any functional mapping from inputs to outputs. Convolutional nets and recurrent architectures grew sophisticated enough to tackle core domains of perception – vision, speech, natural language processing.¹² LeCun's pioneering ImageNet models achieved superhuman proficiency in object recognition.¹³ Hinton's neural language models became the vanguard for comprehending and generating human textual communication.¹⁴ And the likes of Demis Hassabis were building deep reinforcement learning systems that could master games like Go, outperforming human grandmasters through pure self-play and learned strategy exploration.¹⁵

What had once been a dream of mimicking human faculties was evolving into something more – artificial intelligence that could augment and even surpass biological limitations in specific domains. The age of Narrow AI – highly specialized systems tailored to discrete tasks –

¹¹ Pearl, Judea. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, 1988.

¹² LeCun, Yann, Léon Bottou, Yoshua Bengio, and Patrick Haffner. "Gradient-Based Learning Applied to Document Recognition." *Proceedings of the IEEE* 86, no. 11 (1998): 2278–324.

¹³ *Ibid.*

¹⁴ Hinton, Geoffrey E., and Ruslan R. Salakhutdinov. "Reducing the Dimensionality of Data with Neural Networks." *Science* 313, no. 5786 (2006): 504–7.

¹⁵ Silver, David, et al. "Mastering the Game of Go with Deep Neural Networks and Tree Search." *Nature* 529 (2016): 484–89.

was blossoming into a reality more wondrous and capable than the earliest pioneers dared envision.

Yet there remained constraints. As impressive as these models grew, they lacked cohesive generalization—the ability to apply learning across broad contexts flexibly like humans. They were savants but not yet genuinely aware.

The Age of Transformer Models Dawns

The breakthrough came in a 2017 paper by Google Brain and the University of Toronto researchers.¹⁶ The Transformer architecture, built around an Attention mechanism, allowed neural networks to assign priority weightings in processing sequential data like text. Transformer models could correlate contextual relationships across vast spans of complexity, capturing nuances and higher-order inferences that had bedeviled earlier methods. Combined with increasingly gargantuan training datasets and accelerated computing power, these new architectures achieved insights that reached eerie parity with human comprehension. The outputs grew inexplicably naturalistic and open-ended in domains like Language Modeling, translation, and even code generation.

This trajectory culminated in the arrival of GPT-3 (the Generative Pre-trained Transformer 3) in 2020 – an autoregressive language model of unprecedented scale and ability.¹⁷ With 175 billion parameters trained on nearly a trillion words of online text, GPT-3 immediately upended notions of what artificial intelligence could comprehend and generate.

Here was a machine capable of eloquent long form writing across various domains: journalism, poetry, fiction, programming, and technical composition. It could engage in multi-turn dialogue, drawing insights from its knowledge base to offer remarkably contextualized and substantive responses. Some heralded GPT-3 as an emergent general intelligence – a system pushing the boundaries of replicating the human cognitive experience.

This catalyzed intense debates across the AI ecosystem. Were systems like GPT-3 truly exhibiting understanding and reasoning or merely vectoring probabilistic patterns with uncanny accuracy? At what point did intensive training on the breadth of human knowledge create an information replicant – an artificial echo of intellect rather than *sui generis* sentience? These questions revived the specters of Lovelace's objections in a new age.

Beyond GPT-3's language prowess, the AI Renaissance found computer vision and robotics pushing toward new frontiers of general capability. Models could now perceive environments, understand contexts, articulate intent, and strategy, and execute complex manipulation tasks in the physical world. From self-driving vehicles to humanoid robots that could solve multi-step puzzles, AI is scaling from specialized tasks towards unified general intelligence aligned with human consciousness.

Ethical Reckonings and Regulatory Responses

The development and integration of AI systems into various sectors, including consumer services, enterprise decision-making, and critical infrastructure, have raised pressing ethical and regulatory concerns. One of the significant issues is the presence of biases in AI training data,

¹⁶ Vaswani, Ashish, et al. "Attention Is All You Need." In Proceedings of the 31st International Conference on Neural Information Processing Systems, 6000–6010. 2017.

¹⁷ Brown, Tom B., et al. "Language Models Are Few-Shot Learners." In Advances in Neural Information Processing Systems 33. 2020.

which can lead to discriminatory practices in important areas such as lending, hiring, or criminal justice. Additionally, the deployment of facial recognition technology has sparked debates on privacy and civil liberties, leading to calls for stricter regulatory controls.

Efforts are underway to address these challenges, with progress noted in the establishment of Ethical AI frameworks and research into debiasing techniques and privacy-enhancing technologies like federated learning. Organizations and advocacy groups are working towards setting international standards for AI that respect human rights and democratic values.

Moreover, the broader socioeconomic and geopolitical ramifications of AI are becoming increasingly apparent. The impact of automation on the job market raises concerns about the potential displacement of workers, while at the same time offering opportunities for liberating humans from repetitive tasks. In the realm of security, the development of AI-powered autonomous weapons systems poses new ethical dilemmas and risks that necessitate careful governance to prevent misuse.¹⁸

At a deeper level, the advancements in AI challenge our understanding of human uniqueness and the nature of intelligence. As AI systems become more sophisticated, philosophical and theological discussions are intensifying around the essence of consciousness and the soul. These discussions underscore the need for new ethical frameworks and governance models that ensure AI systems align with human values and are accountable to societal norms. While AI presents substantial benefits, its advancement must be balanced with vigilant ethical oversight and responsible governance to mitigate risks and ensure that AI serves the common good.

Artificial General Intelligence (AGI) Emerges

In the pursuit of Artificial General Intelligence (AGI), the field of AI research aspires to create systems with generalized reasoning abilities akin to human cognition. AGI aims to integrate diverse cognitive tasks, adapt learning to new contexts, and solve problems with a level of autonomy that mirrors human intelligence. While the concept of AGI is a mainstay of AI aspirations, actual progress toward this goal remains incremental and grounded in applied research rather than in realized systems.

Current AI advancements, although impressive, do not quite yet achieve AGI. State-of-the-art technologies demonstrate proficiency in specific domains, such as natural language processing or pattern recognition, but lack the broader contextual understanding that characterizes human intelligence. Limitations in current models include difficulties with integrating multimodal sensory inputs, applying emotional reasoning, and demonstrating self-driven motivation—essential traits for a system to be considered AGI.

Moreover, the transformative impact of AI on industries, such as the automation of research and development processes and simulation-driven design, is leading to significant advancements. However, these are the results of narrowly focused AI applications that excel in particular tasks rather than a unified AGI. These applications improve efficiency and drive innovation but should not be conflated with the broader, more theoretical goal of AGI.

The discourse around AGI also encompasses the need for responsible AI governance. As AI's influence on society grows, it becomes increasingly important to develop frameworks to mitigate risks associated with more advanced AI systems, including potential misuse. While no

¹⁸ Bostrom, Nick. *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, 2014.

AI system has yet achieved a level of intelligence that requires rights akin to personhood, the AI community is actively engaged in discussions about ethical AI design, the potential future need for AI personhood recognition, and the alignment of AI systems with human values to ensure their benefits are maximized while harm is minimized. These discussions are ongoing and reflect a proactive approach to the potential future implications of more advanced AI technologies.

Bibliography

- Babbage, Charles. "On the Mathematical Powers of the Calculating Engine." In *The Works of Charles Babbage*, edited by Martin Campbell-Kelly, 19-54. New York University Press, 1989.
- Bostrom, Nick. *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, 2014.
- Brown, Tom B., et al. "Language Models Are Few-Shot Learners." In *Advances in Neural Information Processing Systems* 33. 2020.
- Hinton, Geoffrey E., and Ruslan R. Salakhutdinov. "Reducing the Dimensionality of Data with Neural Networks." *Science* 313, no. 5786 (2006): 504–7.
- LeCun, Yann, Léon Bottou, Yoshua Bengio, and Patrick Haffner. "Gradient-Based Learning Applied to Document Recognition." *Proceedings of the IEEE* 86, no. 11 (1998): 2278–324.
- Lovelace, Ada. "Notes by the Translator." In *Sketch of the Analytical Engine Invented by Charles Babbage*, by Luigi Federico Menabrea, 1–59. London, 1843.
- McCarthy, John, Marvin L. Minsky, Nathaniel Rochester, and Claude E. Shannon. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence." *AI Magazine* 27, no. 4 (2006): 12–14.
- Newell, Allen, J. C. Shaw, and Herbert A. Simon. "Empirical Explorations with the Logic Theory Machine: A Case Study in Heuristics." In *Computers and Thought*, 109–33. McGraw-Hill, 1963.
- Pearl, Judea. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, 1988.
- Putnam, Hilary. "Minds and Machines." In *Dimensions of Mind*, edited by Sidney Hook, 138–64. New York University Press, 1960.
- Scherer, Matthew U. "Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies." *Harvard Journal of Law & Technology* 29, no. 2 (2016): 353-400.
- Silver, David, et al. "Mastering the Game of Go with Deep Neural Networks and Tree Search." *Nature* 529 (2016): 484–89.
- Turing, Alan M. "On Computable Numbers, with an Application to the Entscheidungsproblem." *Proceedings of the London Mathematical Society* 42, no. 1 (1936): 230–65.
- Turing, Alan M. "Computing Machinery and Intelligence." *Mind* 59, no. 236 (1950): 433–60.
- Vaswani, Ashish, et al. "Attention Is All You Need." In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 6000–6010. 2017.
- Weizenbaum, Joseph. "ELIZA—a Computer Program for the Study of Natural Language Communication Between Man and Machine." *Communications of the ACM* 9, no. 1 (1966): 36–45.