

Chapter4

- Until recently, the history of technology could be encapsulated in a single phrase: humanity's quest to manipulate atoms. From fire to electricity, stone tools to machine tools, hydrocarbons to medicines, the journey is a vast, unfolding process in which our species has slowly extended its control over atoms
- At root, the primary driver of all of these new technologies is material – the ever-growing manipulation of their atomic elements
- Then, starting in the mid-twentieth century, technology began to operate on higher level of abstraction
- In the decades after WW2, scientists, technologists, and entrepreneurs funded the fields of computer science and genetics, and a host of companies associated with both
- Eventually, the technologies matured and gave us everything from smartphones to genetically modified rice. But there were limits to what we could do.
- Those limits are now being breached. We are approaching an inflection point with the arrival of these higher-order technologies, the most profound in history.
- Industry research output and patents soared.
- In 1987 there were just ninety academic papers published at Neural Information processing systems, at what became the field's leading conference.
- By the 2020s there were almost 2 thousand
- In the last six years there was a six fold increase in the number of papers published on deep learning alone, tenfold if you widen the view to machine learning as a whole.
- AI really isn't emerging anymore. It's in products, services, and devices you use every day. Across all areas of life, a raft of application rely on techniques that a decade ago were impossible
- AI is already here. But it's far from done

- A big part of what makes human intelligent is that we look at the past to predict what might happen in the future.
- In this sense intelligence can be understood as the ability to generate a range of plausible scenarios about how the world around you may unfold and then base sensible actions on those predictions.

Language Learning Models

- LLMs take advantage of the fact that language data comes in a sequential order.
- Each unit of information is in some way related to data earlier in a series
- The model reads very large numbers of sentences, learns an abstract representation of the information contained within them, and then, based on this, generates a prediction about what should come next.
- The challenge lies in designing an algorithm that “knows where to look” for signals in a given sentence.
- What are the keywords, the most salient elements of a sentence, and how do they relate to one another? In AI this notion is commonly referred to as “attention”.
- When a LLM ingests a sentence, it constructs what can be thought of as an “attention map”.
- It first organizes commonly occurring groups of letters or punctuation into “tokens”, something like syllables, but really just chunks of frequently occurring letters making it easier for the model to process the information.
- It’s worth noting that humans do this with words of course, but the model doesn’t use our vocabulary.
- Instead, it creates a new vocabulary of common tokens that helps it spot patterns across billions of billions of documents
- In the attention map, every token bears some relationship to every token before it, and for a given input sentence the strength of this relationship describes something about the importance of that token in the sentence.
- In effect, the LLM learns which words to pay attention to.

- These systems are called transformers
- Much of AI's progress during the mid-2010s was powered by the effectiveness of "supervised" deep learning. Here AI models learn from carefully hand labeled data
- Today's LLMs are trained on trillions of words.
- Transistors are getting so small they are hitting physical limits: at this size electrons start to interfere with one another, messing up the process of computation.
- While this is true, it misses the fact that in AI training we can keep connecting larger and larger arrays of chips, daisy-chaining them into massively parallel supercomputers. There is therefore no doubt that the size of the large AI training jobs will continue to scale exponentially
- Researchers meanwhile see more and more evidence for "the scaling hypothesis" which predicts that the main driver of performance is, quite simply, to go big and keep going bigger.
- Keep growing these models with more data, more parameters, more computation, and they will keep improving – potentially all the way to human-level intelligence and beyond
- AI researchers are racing to reduce costs and drive up performance so that these models can be used in all sorts of production settings.
- In the last four years, the costs and time needed to train advanced language models have collapsed.
- AI systems now help engineers generate production ready code
- In 2022, OpenAI and Microsoft unveiled a new tool called Copilot, which quickly became ubiquitous among coders.
 - One analysis suggest it makes engineers 55% faster at completing coding tasks, almost like having a second brain on hand
 - Manu coders now increasingly outsource much of their mundane work, focusing instead on knotty and creating problems
- In the words of an eminent computer scientist "It seems totally obvious to me that of course all programs in the future will ultimately be written by AIs, with humans relegated to, at best, a supervisory role".

- Anyone with an internet connection and a credit card will soon be able to deploy these capabilities-an infinite stream of output on tap
- In the words of John McCarthy, who coined the term “artificial intelligence”: “As soon as it works, no one calls it AI anymore”
- AI is -as those of us building it like to joke – “what computers can’t do”. Once they can, it’s just software
- Critics like NYU professor Gary Marcus believe deep learning’s limitations are evident, that despite the buzz of generative AI the field is “hitting a wall”, that it doesn’t present any path to key milestones like being capable of learning concepts or demonstrating real understanding
- The eminent professor of complexity Melanie Mitchell rightly points out that present-day AI systems have many limitations: they can’t transfer knowledge from one domain to another, provide quality explanations of their decision-making process, and so on.
- Today we have narrow or weak AI: limited and specific versions
- GPT-4 can spit out virtuoso texts, but it can’t turn around tomorrow and drive a car, as other AI programs do.
- Existing AI systems still operate in relatively narrow lanes
- Over the last decade, intellectual and political elites in tech circles became absorbed by the idea that a recursively self-improving AI would lead to an “intelligence explosion” known as the Singularity
- The first stage of AI was about classification and prediction – it was capable, but only within clearly defined limits and preset tasks
- AIs represent the next stage of AI’s revolution
- A system that not only could recognize and generate novel images, audio, and language appropriate to a given context, but also would be interactive-operating in real time, with real users.
- Conscious super intelligence? Who knows? But highly capable learning systems, AIs, that can pass some version of the modern Turing test? Make no mistake: they are on their way, are already here in embryonic form.
- There will be thousands of these models, and they will be used by the majority of the world’s population.

- It will take us to the point where anyone can have an ACI in their pocket
- The future of AI is, at least in one sense, fairly easy to predict.
- Over the next five years, vast resources will continue to be invested
- Some of the smartest people on the planet are working on these problems
- Orders of magnitude more computation will train the top models
- AI is far deeper and more powerful than just another technology.
- The risk isn't in overhyping it, it's rather in missing the magnitude of the coming wave
- It's not just a tool or platform but a transformative meta-technology, the technology behind the technology and everything else, itself a maker of tools and platforms, not just a system but a generator of systems of any and all kinds

Chapter 6

- AIs are products of bit and code, existing within simulations and servers.
- Robots are their bridge, their interface with the real world
- If AI represents the automation of information, robotics is the automation of the material, the physical instantiations of AI, a step change in what it is possible to do
- Mastery of bits comes full circle, directly reconfiguring atoms, rewriting the bonds not just of what can be thought or said or calculated but what can be built in the most tangible physical sense
- The remarkable thing about the coming wave is that this kind of blunt atomic manipulation is nothing compared with what's on the horizon
- In 2019, Google announced that it had reached "quantum supremacy".
- Researchers had built a quantum computer, one using the peculiar properties of the subatomic world
- Google's machine used an understanding of quantum mechanics to complete a calculation in seconds that would have taken a conventional computer ten thousand years
- It had just 53 "qubits", or quantum bits, the core units of quantum computing

- To store equivalent information on a classical computer, you would need 72 billion GB of memory
- Its key attraction is that each additional qubit doubles a machine's total computing power.
- Start adding qubits and it gets exponentially more powerful
- If you wanted to write the crudest possible equation for our world it would be something like this:

$$(\text{Life} + \text{Intelligence}) \times \text{Energy} = \text{Modern Civilization}$$

- Increase any or all of those inputs (let alone supercharge their marginal cost toward zero) and you have a step change in the nature of society
- Energy – expensive and dirty as often is – is at present a limiter on technology's rate of progress
- As the elements of AI, advanced biotechnology, quantum computing, and robotics combine in new ways, prepare for breakthroughs like advanced nanotechnology, a concept that takes the ever-growing precision of technology to its logical conclusion
- The ultimate vision of nanotechnology is one where atoms become controllable building blocks, capable of automatically assembling almost anything
- Nanomachines would work at speeds far beyond anything at our scale, delivering extraordinary outputs: an atomic-scale nanometer, e.g., could rotate 48 billion times a minute.
- At its core, the coming wave is a story of the proliferation of power
- If the last wave reduced the costs of broadcasting information, this one reduces the costs of acting on it

Chapter 7

- New technologies create previously unthinkable vulnerabilities and pressure points against seemingly dominant powers
- They are developing fast, a kind of hyper-evolution, iterating, improving, and branching into new areas at incredible speed

- They are often omni-use, i.e., they can be used for many different purposes.
- They increasingly have a degree of autonomy beyond any previous technology
- Prospects for asymmetric impact are growing all around, and also in the positive sense – single systems can deliver huge benefits as well
- The reverse of asymmetric action is also true
- The very scale and interconnectedness of the coming wave create new systematic vulnerabilities: one point of failure can quickly cascade around the world
- The less localized a technology, the less easily it can be contained-and vice versa
- If you want to contain technology, you might hope it develops at a manageable pace, giving society time and space to understand and adapt to it
- AI already helps find new materials and chemical compounds. For example, scientists have used neural networks to produce new configurations of Lithium, with implications for battery technology
- AI has helped design and build a car using 3-D printers
- In some cases the final outcome looks bizarrely different from anything designed by a human, resembling the undulating and efficient forms found in nature
- Configurations of wiring and ducting are organically melded into the chassis for optimal use of space
- Parts are too complex to build using conventional tooling and have to be 3D printed
- One of the promising areas of AI is automated drug discovery
- AI techniques can search through the vast space of possible molecules but helpful treatments
- In 2020 an AI system sifted through 100 million molecules to create the first machine learning derived antibiotic – called halicin – which can potentially fight tuberculosis

- There's a flip side. Researchers looking for these helpful compounds raised an awkward question
 - What if you redirected the discovery process
 - What if, instead of looking for cures, you looked for killers
 - They ran a test, asking their molecule generating AI to find poisons.
 - In six hours it identified more than 40 thousand molecules with toxicity comparable to the most dangerous chemical weapons, like Novichok.
- It turns out that in drug discovery, one of the areas where AI will undoubtedly make the clearest possible difference, the opportunities are very much "dual use"
- Dual-use technologies are those with both civilian and military applications
- A more appropriate term for the technologies of the coming wave is "omni-use", a concept that grasps at the sheer level of generality, the extreme versatility on display
- Omni-use technologies like steam or electricity have wider societal effects and spillovers than narrow technologies
- Omni-use technologies are more valuable than narrow ones
- Over time, technology tends toward generality. What this means is that weaponizable or harmful uses of the coming wave will be possible regardless of whether this was intended
- Many technologies and systems are becoming so complex that they are beyond the capacity of any one individual to truly understand: quantum computing and other technologies operate toward the limits of what can be known
- A paradox of the coming wave is that its technologies are largely beyond our ability to comprehend at a granular level yet still within our ability to create and use.
- In AI, the neural networks moving towards autonomy are, at present, not explainable
- Engineers can't peer beneath the hood and easily explain what caused something to happen