“We become builders by building and we become harpists by playing the harp,” he wrote. “Similarly, then, we become just by doing just actions, temperate by doing temperate actions, brave by doing brave actions.”

Would you have a great empire? Rule over yourself.

Freedom, as Eisenhower famously said, is actually only the “opportunity for self-discipline.”

When you love the work, you don’t cheat it or the demands it asks of you. You respect even the most trivial aspects of the pursuit

You have to do your best while you still have a chance. Life is short. You never know when the game, when your body, will be taken away from you.

You say you love what you do. Where’s your proof? What kind of streak do you have to show for it?

“Writers all devise ways to approach that place where they expect to make the contact,” she’d later reflect, “where they become the conduit, or where they engage in this mysterious process.

Has anyone ever drunk or eaten their way to happiness? No. An early grave? Misery? Regret? You bet.

the habit itself is less important than what we’re really quitting, which is dependency. What the Buddhists call tanha. The thirst. The craving.

The cost is not just personal but shared by us all, in symphonies never written, feats never accomplished, in good never done, the potential of an ordinary day never fulfilled.

Be regular and orderly in your life, so that you may be violent and original in your work.

The modern conveniences we can trace to his lab then, owe far more to his body than his brain, to the compounding power of consistency rather than sheer brilliance. It wasn’t about inspiration. It was about getting to work.

We don’t rise to the occasion, we fall to the level of our training.

Always and forever, the reward is the work. It is a joy itself. It is torture and also heaven—sweaty, wonderful salvation.

We think we can make up for it with brilliance or creativity, but what we really need is commitment. What we need is a willingness to put our body where the problem is, throwing ourselves completely into solving it, to show that

We think we can make up for it with brilliance or creativity, but what we really need is commitment. What we need is a willingness to put our body where the problem is, throwing ourselves completely into solving it, to show that we are not for turning, that we will not be deterred.

Edison tested six thousand filaments, one by one, in his laboratory,

Edison tested six thousand filaments, one by one, in his laboratory, before he found the one that brought us light.

We know that between every stimulus and its response, every piece of information and our decision, there is space. It is a brief space, to be sure, but one with room enough to insert our philosophy.

It’s just a fact. The muses never bless the unfocused. And even if they did, how would they notice?

Toni Morrison didn’t get up before dawn for some “me time.” The mornings weren’t for catching up on the news or folding laundry. She had a short window and she used it to write—seizing the day while others weren’t yet stirring.

“It took about four hundred years for the Lord Jesus Christ to have his message accepted. Up to that time he would be considered a ‘failure.’

No leader, no matter how good they are, can hope to avoid criticism.

the second we stop trying to get better is the moment we start gradually getting worse.

Ben Franklin, many generations later, would put forth an even better rule: “Search others for their virtues, thyself for thy vices.”

The leader shows up first and leaves last. The leader works hardest. The leader puts others before themselves. The leader takes the hit. Everything else is just semantics and titles.

Coming up with the idea for a book is a creative pursuit. Actually creating the book is a work of excruciating manual labor, sitting in a chair, grinding out each consecutive sentence—a process not measured in hours or days, but months and years. It’s a marathon of endurance, cognitive and physical.

a talented leader identifies the one or two critical issues in the situation—the pivot points that can multiply the effectiveness of effort—and then focuses and concentrates action and resources on them.

The core of strategy work is always the same: discovering the critical factors in a situation and designing a way of coordinating and focusing actions to deal with those factors.

discovering the critical factors in a situation and designing a way of coordinating and focusing actions to deal with those factors.

A good strategy includes a set of coherent actions. They are not “implementation” details; they are the punch in the strategy. A strategy that fails to define a variety of plausible and feasible immediate actions is missing a critical component.

A good strategy has an essential logical structure that I call the kernel. The kernel of a strategy contains three elements: a diagnosis, a guiding policy, and coherent action.

The most basic idea of strategy is the application of strength against weakness. Or, if you prefer, strength applied to the most promising opportunity

The most basic idea of strategy is the application of strength against weakness. Or, if you prefer, strength applied to the most promising opportunity A good strategy doesn’t just draw on existing

the actual discovery of power in a situation—the creation or revelation of a decisive asymmetry. How someone can see what others have not, or what they have ignored, and thereby discover a pivotal objective and create an advantage, lies at the very edge of our understanding, something glimpsed only out of the corner of our minds.

use your relative advantages to impose out-of-proportion costs on the opposition and complicate his problem of competing with you.

strategy is a way through a difficulty, an approach to overcoming an obstacle, a response to a challenge. If the challenge is not defined, it is difficult or impossible to assess the quality of the strategy.

Being a general manager, CEO, president, or other top-level leader means having more power and being less constrained. Effective senior leaders don’t chase arbitrary goals. Rather, they decide which general goals should be pursued. And they design the subgoals that various pieces of the organization work toward. Indeed, the cutting edge of any strategy is the set of strategic objectives (subgoals) it lays out. One of the challenges of being a leader is mastering this shift from having others define your goals to being the architect of the organization’s purposes and objectives.

A great deal of strategy work is trying to figure out what is going on. Not just deciding what to do, but the more fundamental problem of comprehending the situation.

The coordination of action provides the most basic source of leverage or advantage available in strategy.

A “threshold effect” exists when there is a critical level of effort necessary to affect the system. Levels of effort below this threshold have little payoff. When there are threshold effects, it is prudent to limit objectives to those that can be affected by the resources at the strategist’s disposal.

An important duty of any leader is to absorb a large part of that complexity and ambiguity, passing on to the organization a simpler problem—one that is solvable.

Good strategy is design, and design is about fitting various pieces together so they work as a coherent whole.

When faced with a question or problem to which there is no obvious answer, it is human nature to welcome the first seemingly reasonable answer that pops into mind, as if it were a life preserver in a choppy sea. The discipline of analysis is to not stop there, but to test that first insight against the evidence.

advantage is rooted in differences—in the asymmetries among rivals. In real rivalry, there are an uncountable number of asymmetries. It is the leader’s job to identify which asymmetries are critical—which can be turned into important advantages.

It is hard to show your skill as a sailor when there is no wind. Similarly, it is in moments of industry transition that skills at strategy are most valuable.

If you can peer into the fog of change and see 10 percent more clearly than others see, then you may gain an edge.

strategy is, in the language of science, a hypothesis,

A new strategy is, in the language of science, a hypothesis, and its implementation is an experiment. As results appear, good leaders learn more about what does and doesn’t work and adjust their strategies accordingly.

Given that we are working on the edge, asking for a strategy that is guaranteed to work is like asking a scientist for a hypothesis that is guaranteed to be true—it is a dumb request.

The heart of the scientific method is that the worth of a hypothesis is determined by empirical data drawn from the physical world, not by the author’s popularity, caste, or wealth. That is the radical revolution wrought by the Enlightenment.

When one has an initial insight into what to do about a challenging situation, it never occurs in the form of a full-blown strategy. Rather, the lightning of insight strikes in one of the three elements of the kernel.

Trying to destroy your own ideas is not easy or pleasant. It takes mental toughness to pick apart one’s own insights.

Even after the antibiotic was discovered, humans could not articulate precisely why it worked. The AI did not just process data more quickly than humanly possible; it also detected aspects of reality humans have not detected, or perhaps cannot detect. A

Even after the antibiotic was discovered, humans could not articulate precisely why it worked. The AI did not just process data more quickly than humanly possible; it also detected aspects of reality humans have not detected, or perhaps cannot detect.

It is reasonable to expect that over time, AI will progress at least as fast as computing power has, yielding a millionfold increase in fifteen to twenty years.

is reasonable to expect that over time, AI will progress at least as fast as computing power has, yielding a millionfold increase in fifteen to twenty years.

FOR AS LONG as history has been recorded, security has been the minimum objective of an organized society. Cultures have differed in their values, and political units have differed in their interests and aspirations, but no society that could not defend itself—either alone or in alignment with other societies—has endured.

Cyber arms-control negotiators (which do not yet exist) will need to solve the paradox that discussion of a cyber weapon’s capability may be one and the same with its forfeiture (permitting the adversary to patch a vulnerability) or its proliferation (permitting the adversary to copy the code or method of intrusion).

In 1450, Johannes Gutenberg, a goldsmith in the German city of Mainz, used borrowed money to fund the creation of an experimental printing press. His effort barely succeeded—his business floundered, and his creditors sued—but by 1455, the Gutenberg Bible, Europe’s first printed book, appeared.

I told her that one of the most important things Apple does is trust itself. During my time at Chiat, we didn’t test a single ad. Not for print, TV, billboards, the web, retail, or anything.

Apple’s approach embraced the idea that it’s okay to make a mistake, that it’s better to shoot for the stars and fall short on occasion than to burden itself with processes that drain the creativity from its ads.

Steve Jobs looked at pretty much everything with the idea of cutting it down to its essence, whether it was a new product or a new ad.

Apple got it right by empowering a small group of smart people and creating a schedule that didn’t allow the process to stagnate.

You can tell a lot about someone by the people he or she admires, and that would be the philosophy of our campaign. By celebrating the lives of those who inspired Apple, Apple would be telling the world exactly what kind of company it is—without using any more words than “Think different.” We wouldn’t be churning out ads. Instead, we’d be crafting posters that paid tribute to Apple’s heroes.

The Think different ads were a vivid reminder that a single iconic image can be the most powerful form of communication.

To me, marketing is about values. This is a very complicated world, it’s a very noisy world, and we’re not going to get a chance to get people to remember much about us. No company is. And so we have to be really clear on what we want them to know about us.

Product naming is the ultimate exercise in Simplicity. It requires one to capture in a single word, possibly two, the essence of a product or company—or in some cases create a personality for it.

Human beings are naturally programmed to identify products by single words. Ask anything more of them and you’re bound to be disappointed. People will say, “I’ll look it up on my iPhone” but never “I’ll look it up on my Apple iPhone.”

Steve Jobs was so determined to use this name for his world-changing device, he chose to announce it without first securing the legal rights. As a born marketer, Steve was willing to do risky things (and obviously expensive things) when he saw great marketing potential in a name.

Steve would just share his point of view about why the product existed, how it worked and what set it apart. When necessary, he’d jump to the whiteboard to illustrate his point. In effect, Steve acted as his own slide show. There was no more efficient way to hear about the philosophy behind a product than to hear it from the CEO himself.

The informality of Steve’s conference room style was based on the fact that he considered most meetings to be brainstorming sessions. Even if you brought finished work to the meeting to share, it would be discussed and dissected before you left the room.

It would have been offensive if Steve had asked us to work harder than he did. But he didn’t. He only asked us to work exactly as hard as he did. I have no idea how he did it, but in the years following his return to Apple he would be CEO of Apple three days a week, CEO of Pixar two days a week, and CEO of both on weekends. Yet there were still times when he’d be unavailable because he’d promised his family he’d take them to a movie.

Human-speak is a hallmark of Simplicity. It’s the recognition that the best way to connect with people is to put things in human terms and use the words that people use in everyday conversation.

Steve Jobs was one of the most forward-thinking people on this planet, yet he was refreshingly old-fashioned when it came to the use of analytics. He demanded all the information he could get, and he would digest every bit of it—but he took it all in context. He never lost sight of the fact that at the end of the day, technology is about people: what stirs their imaginations, what keeps them satisfied, and what makes them smile. He would never sacrifice that kind of connection in favor of a decision that somehow got Apple a few more clicks on its website.

When you first start off trying to solve a problem, the first solutions you come up with are very complex, and most people stop there. But if you keep going, and live with the problem and peel more layers of the onion off, you can oftentimes arrive at some very elegant and simple solutions

start thinking in outcomes rather than outputs. That means rather than defining your success by the code that you ship (your output), you define success as the value that code creates for your customers and for your business (the outcomes). Rather than measuring value in features and bells and whistles, we measure success in impact—the impact we have had on our customers’ lives and the impact we have had on the sustainability and growth of our business.

definition of continuous discovery: At a minimum, weekly touchpoints with customers By the team building the product Where they conduct small research activities In pursuit of a desired outcome

I had an hour to solve a problem, I’d spend 55 minutes thinking about the problem and 5 minutes thinking about solutions.”

“If I had an hour to solve a problem, I’d spend 55 minutes thinking about the problem and 5 minutes

two of the most important steps for reaching our desired outcome are first, how we map out and structure the opportunity space, and second, how we select

Researchers have found that teams that set specific, challenging goals outperform teams who don’t. Challenging goals create focus, inspire effort and persistence, and help to surface relevant organizational knowledge. However, there are some caveats. The team has to believe that they can achieve the goal, and they need to be committed to the goal, further supporting the idea that teams need to be involved in defining their own outcomes. Teams also need continuous feedback on their progress toward their goal, supporting the argument that goals should be measurable.

“Confidence is a feeling, which reflects the coherence of the information and the cognitive ease of processing it. It is wise to take admissions of uncertainty seriously, but declarations of high confidence mainly tell you that an individual has constructed a coherent story in his mind, not necessarily that the story is true.” — Daniel Kahneman,

The purpose of customer interviewing is not to ask your customers what you should build. Instead, the purpose of an interview is to discover and explore opportunities. Remember, opportunities are customer needs, pain points, and desires. They are opportunities to intervene in your customers’ lives in a positive way.

If you want to build a successful product, you need to understand your customers’ actual behavior—their reality—not the story they tell themselves.

Our primary research question in any interview should be: What needs, pain points, and desires matter most to this customer?

A good way to do this is to ask, “If you had that feature, what would that do for you?” For example, if an interviewee says, “I wish I could just say the name of the movie I’m searching for,” that’s a feature request. If you ask, “What would that do for you?” they might respond, “I don’t want to have to type out a long movie title.” That’s the underlying need. The benefit of capturing the need and not just the solution is that the need opens up more of the solution space. We could add voice search to address this need, but we also could auto-complete movie titles as they type.

John Dewey, an American educational philosopher, encourages us to “carry on systematic and protracted inquiry.” Rather than jumping to the first need that we might address, Dewey argues, good thinking requires that we explore our options—that we carry out a systematic search for longer than we feel comfortable. We should compare and contrast the impact of addressing one opportunity against the impact of addressing another opportunity. We want to be deliberate and systematic in our search for the highest-impact opportunity.

When a customer expresses emotion in an interview, it’s usually a strong signal that an opportunity is lurking nearby. However, don’t capture the feeling itself as the opportunity. Instead, look for the cause of the feeling. When we capture opportunities like “I’m frustrated” or “I’m overwhelmed,” we limit how we can help. We can’t fix feelings. But if we capture the cause of those feelings—“I hate typing in my password every time I purchase a show” or “I’m way behind on this show”—we can often identify solutions that address the underlying cause.

Product strategy happens in the opportunity space. Strategy emerges from the decisions we make about which outcomes to pursue, customers to serve, and opportunities to address.

When you are generating assumptions, always phrase your assumptions such that you need them to be true: “Customers will remember their passwords.” For many assumptions, you’ll find that this positive framing will make them easier to test.

A good mentor won’t hand you the answers, but they will try to help you see your problem from a new perspective.

Do, fail, learn. The rest will follow.

The critical thing is to have a goal. To strive for something big and hard and important to you. Then every step you take toward that goal, even if it’s a stumble, moves you forward.

If you’re not solving a real problem, you can’t start a revolution.

Steve Jobs once said of management consulting, “You do get a broad cut at companies but it’s very thin. It’s like a picture of a banana: you might get a very accurate picture but it’s only two dimensions, and without the experience of actually doing it you never get three dimensional. So you might have a lot of pictures on your walls, you can show it off to your friends—I’ve worked in bananas, I’ve worked in peaches, I’ve worked in grapes—but you never really taste it.”

“I can’t make you the smartest or the brightest, but it’s doable to be the most knowledgeable. It’s possible to gather more information than somebody else.”

The key is persistence and being helpful. Not just asking for something, but offering something. You always have something to offer if you’re curious and engaged. You can always trade and barter good ideas; you can always be kind and find a way to help.

Being exacting and expecting great work is not micromanagement. Your job is to make sure the team produces high-quality work. It only turns into micromanagement when you dictate the step-by-step process by which they create that work rather than focusing on the output.

You should have a weekly crib sheet that helps you keep your priorities and the questions you need to ask top of mind.

To get people to join you, to truly become a team, to fill them with the same energy and drive that’s bubbling within you, you need to tell them the why.

If you’re testing the core of your product, if the basic functionality can flex and change depending on the whims of an A/B test, then there is no core. There’s a hole where your product vision should be and you’re just shoveling data into the void.

Storytelling is how you get people to take a leap of faith to do something new. It’s what all our big choices ultimately come down to—believing a story we tell ourselves or that someone else tells us. Creating a believable narrative that everyone can latch on to is critical to moving forward and making hard choices. It’s all that marketing comes down to. It’s the heart of sales.

But when you’re creating a new product, regardless of whether it’s made of atoms or electrons, for businesses or consumers, the actual thing you’re building is only one tiny part of a vast, intangible, overlooked user journey that starts long before a customer ever gets their hands on your product and ends long after. So don’t just make a prototype of your product and think you’re done. Prototype as much of the full customer experience as possible. Make the intangible tangible so you can’t overlook the less showy but incredibly important parts of the journey. You should be able to map out and visualize exactly how a customer discovers, considers, installs, uses, fixes, and even returns your product. It all matters.

“Don’t tell me what’s so special about this object. Tell me what’s different about the customer journey.”

You should be prototyping your marketing long before you have anything to market.

Every product should have a story, a narrative that explains why it needs to exist and how it will solve your customer’s problems. A good product story has three elements: » It appeals to people’s rational and emotional sides. » It takes complicated concepts and makes them simple. » It reminds people of the problem that’s being solved—it focuses on the “why.”

That “why” is the most critical part of product development—it has to come first. Once you have a strong answer for why your product is needed, then you can focus on how it works. Just don’t forget that anyone encountering your product for the first time won’t have the context you have. You can’t just hit customers on the head with the “what” before you tell them the “why.”

Steve didn’t just read a script for the presentation. He’d been telling a version of that same story every single day for months and months during development—to us, to his friends, his family. He was constantly working on it, refining it. Every time he’d get a puzzled look or a request for clarification from his unwitting early audience, he’d sand it down, tweak it slightly, until it was perfectly polished.

A good story is an act of empathy. It recognizes the needs of its audience. And it blends facts

A good story is an act of empathy. It recognizes the needs of its audience. And it blends facts and feelings so the customer gets enough of both.

That’s why analogies can be such a useful tool in storytelling. They create a shorthand for complicated concepts—a bridge directly to a common experience. That’s another thing I learned from Steve Jobs. He’d always say that analogies give customers superpowers. A great analogy allows a customer to instantly grasp a difficult feature and then describe that feature to others. That’s why “1,000 songs in your pocket” was so powerful.

The tools you need to make those decisions are below, organized by order of importance: 1. Vision: Know what you want to make, why you’re making it, who it’s for, and why people will buy it. You’ll need a strong leader or a small group to ensure the vision is delivered intact. 2. Customer insights: This is what you’ve learned through customer or market research or simply by thinking like your customer: what they like, what they dislike, what problems they experience on a regular basis, and what solutions they’ll respond to. 3. Data: For any really new product, reliable data will be limited or nonexistent. That doesn’t mean you shouldn’t make a reasonable attempt to gather objective information—the scope of the opportunity, the way people use current solutions, etc. But this information will never be definitive. It won’t make your decisions for you.

You should also keep in mind that you’re not just making V1 or V2 of your product—you’re building out the first or second version of your team and processes.

You may not come back from it. You may have run out of money, lost the team or your credibility. But the only way to move forward is to do an honest accounting of the past. Learn your lessons—especially the hard ones. Then try again. Back to the drawing board. V1.

write a press release. But don’t write it when you’re done. Write it when you start. I began doing this at Apple and eventually realized other leaders had figured it out, too (looking at you, Bezos). It’s an incredibly useful tool to narrow down what really matters.

We forced as many constraints on ourselves as possible: not too much time, not too much money, and not too many people on the team. That last point is important.

I was exactly the kind of founder investors like. Four failed startups and years of professional disappointment had paved the way for a decade of success. I was forty years old, knew exactly how hard this was going to be and which mistakes not to make again. I’d worked on hardware and software at tiny and enormous companies. I had contacts, credibility, and enough experience to know what I didn’t know.

According to the book Super Founder, by Ali Tamaseb, around 60 percent of the founders of billion-dollar startups started another company before their wild success and many lost a ton of money. Just 42 percent of them had a previous exit of $10 million or more, so the majority “failed” by the standards of venture capital. But they came out on the other side with a basic mental model of a startup. They understood the operational details and what it might look like if that tiny startup became successful. That’s it. That’s the magical key to success.

Forming that team and shepherding it through its many transitions is always the hardest and most rewarding part of building anything.

What you’re building never matters as much as who you’re building it with.

solve it yourself, especially if solving that problem is core to the future of your business. If it’s a critical function, your team needs to build

you shouldn’t outsource a problem before you try to solve it yourself, especially if solving that problem is core to the future of your business. If it’s a critical function, your team needs to build the muscle to understand the process and do it themselves.

Steve Jobs called it “staying a beginner.” He was constantly telling us to look at what we were making with fresh eyes.

Steve Jobs often said, “The best marketing is just telling the truth.”

The job is to give a shit. To care. About everything.

So don’t worry about picking your battles. Don’t rack your brain trying to decide which parts of your company need your attention and which don’t. They all do. You can prioritize, but nothing ever comes off the list. Avoiding or ignoring any part of your company only comes back to haunt you sooner or later.

Steve took a lot of risks, made bad decisions, launched products that didn’t work—the original Apple III, the Motorola ROKR iTunes phone, the Power Mac G4 Cube, the list goes on. But if you aren’t failing, you aren’t trying hard enough. He learned from the screwups, was constantly improving, and his good ideas, his successes, totally wiped away his failures. He was constantly pushing the company to learn and try new things.

Bad CEOs come to board meetings and expect the board to help them make decisions. Good CEOs walk in with a presentation of where the company was, where it is now, and where it’s headed this quarter and in the years to come. They tell the board what’s working but they’re also transparent about what isn’t and how they’re addressing it. They present a fully formed plan that the board can question, object to, or try to modify. Things might get a little heated, a little bumpy, but in the end everyone walks out of the meeting understanding and accepting the CEO’s vision and the company’s path forward.

Most people and companies need a near-death experience before they can really change.

In the end, there are two things that matter: products and people. What you build and who you build it with. The things you make—the ideas you chase and the ideas that chase you—will ultimately define your career. And the people you chase them with may define your life.

Our normal habit in life is to develop a quick belief about a situation and then seek out information that bolsters our belief.

Throughout Walton’s career, he kept his eyes out for good ideas. He once said that “most everything I’ve done I’ve copied from someone else.”

“The use of analogies is one of the main mechanisms for driving research forward.” And the key to using analogies successfully, he said, was the ability to extract the “crucial features of the current problem.” This required the scientist to think of the problem from a more abstract, general perspective, and then “search for other problems that have been solved.”

they found that for every favorable article written in a major publication about the CEO, the acquisition premium paid went up by 4.8%. That’s a $4.8 million boost on a $100 million acquisition! Because of one flattering article! And a second article would inflate it by another $4.8 million.

“What would have to be true?” question has become the most important ingredient of his strategy work,

when you need trustworthy information, go find an expert—someone more experienced than you. Just keep them talking about the past and the present, not the future.

“Go and see what’s happening,” he told one. “See the end product of what we are doing. Talk to people; get the wind in your nose.”

When people share the worst decisions they’ve made in life, they are often recalling choices made in the grip of visceral emotion: anger, lust, anxiety, greed.

Our calendars are the ultimate scoreboard for our priorities.

An extensive body of research confirms that procedural justice is critical in explaining how people feel about a decision. It’s not just the outcome that matters; it’s the process.

Short-run emotion, as we’ve seen, makes the status quo seductive. But when researchers ask the elderly what they regret about their lives, they don’t often regret something they did; they regret things they didn’t do. They regret not seizing opportunities.

Short-run emotion, as we’ve seen, makes the status quo seductive. But when researchers ask the elderly what they regret about their lives, they don’t often regret something they did; they regret things they didn’t do. They regret not seizing opportunities. They regret hesitating. They regret being indecisive.

Neural networks became the first major game changer in the field of NLP by making it feasible to perform complex natural language tasks, something that had previously been possible only in theory.

Generative modeling is a branch of statistical modeling. It is a method for mathematically approximating the world.

GPT-3 is pre-trained on a corpus of text from five datasets: Common Crawl, WebText2, Books1, Books2, and Wikipedia:

To vet for quality, the authors scraped all outbound links from Reddit that received at least three karma (an indicator for whether other users found the link interesting, educational, or just funny).

This corpus includes nearly a trillion words altogether.

An attention mechanism is a technique that mimics cognitive attention: it looks at an input sequence, piece by piece and, on the basis of probabilities, decides at each step which other parts of the sequence are important.

An attention mechanism is a technique that mimics cognitive attention: it looks at an input sequence, piece by piece and, on the basis of probabilities, decides at each step which other parts of the sequence are important. For example, look at the sentence “The cat sat on the mat once it ate the mouse.” Does “it” in this sentence refer to “the cat” or “the mat”? The transformer model can strongly connect “it” with “the cat.” That’s attention.

Releasing GPT-3 via an API was a revolutionary move. Until 2020, the powerful AI models developed by leading research labs were available to only a select few—researchers and engineers working on those projects. The OpenAI API gives users all over the world unprecedented access to the world’s most powerful language model via a simple sign-in. (OpenAI’s business rationale for this move is to create a new paradigm it calls “model-as-a-service” where developers can pay per API call;

“imagine GPT-3 as a friend or a colleague that you’re asking to do something. How would you describe the task that you want them to do? And then, see how GPT-3 responds. And if it doesn’t respond in the way that you want, iterate on your instructions.”

The secret to writing good prompts is understanding what GPT-3 knows about the world. As Awan points out, “It has only seen text. That means you shouldn’t expect that it knows about the physical world, even though it obviously does. It could describe the Mona Lisa, [could] tell you [about] the significance, the importance, the history [of] it probably, but it’s never seen [the painting] because it’s only trained on text.”

Your job is to get the model to use the information it already has to generate useful results. In the game of charades, the performer gives the other players just enough information to figure out the secret word. Similarly, with GPT-3, we give the model just enough context (in the form of a training prompt) to figure out patterns and perform the given task.

You can think of GPT-3 processing the input in the same way as the human brain. When somebody asks us any question without proper context we tend to give random responses. This happens because without any proper direction or context, it’s difficult to get to the precise response. The same is true of GPT-3; its universe of training data is so big that it’s difficult to navigate to a correct response without any external context or direction.

Embeddings are the core of any machine learning model and allow you to capture semantics from the text by converting it into high-dimensional vectors. Currently, developers tend to use open source models to create embeddings for their data that can be used for a variety of tasks like recommendation, topic modeling, semantic search, etc.

Customizing GPT-3 seems to yield better results than what can be achieved with prompt design, because during this process you can provide more examples.

By customizing GPT-3, Keeper Tax experienced an increase in accuracy from 85% to 93%. And it continuously improves thanks to adding 500 new training examples to its model once a week, which is leading to about a 1% accuracy improvement per week.

Text classification is the process of categorizing text into organized groups. By using NLP, text classification can automatically analyze text and then assign a set of predefined tags or categories based on its context.

“you pregenerate a bunch of options and then can use search to find the right option to respond back with,”

“Keep a very close eye on what you are feeding to the model. Be careful with the punctuation, grammar, and the wording of the prompt. I guarantee you’ll have a much better experience with the model output.”

GPT-3 has no way to verify the truth, logic, or meaning of any of the millions of lines of text it produces on a daily basis.

The steps to building a data learning effect with intelligent machines are (1) capturing a critical mass of data, (2) developing capabilities to process that data into information, and (3) feeding that information into a computer that runs calculations over data to learn something new. Here’s the formula: data learning effects = economies of scale to data + data processing capabilities + data network effects.

Products with next-level data network effects often get data from customers that provide data as part of using the product, sometimes called feedback data, but—crucially—that data goes into a model that compounds the value of this data by transforming it into a prediction.

80 percent of the time dedicated to building AI is spent preparing data and the other 20 percent is spent creating the models.

The ongoing challenge is to make sure that the models don’t break or lose touch with reality. AI-First products are powerful because they constantly adapt, evolve, and spawn new data, but this constant change makes them hard to manage, like a multidimensional balancing act of the following factors.

Design products to constantly collect feedback data; for example, ask customers if the prediction was right, to correct it, or to just accept it. Feedback data is proprietary because it comes from something customers observed in the real world only after acting on the prediction made by the model. Competitors won’t have those observations.