Career Counsellor

Counsellor's Newsletter

As AI is making inroads in our daily life and becoming the new normal in many facets, we are learning to co-exist with the new trend. Students will constantly be challenged by the use of the new technologies, sometimes helping them upskill and sometimes being skilfully replaced by them. As students will encounter AI through life from university to workplace, let us see how the experts are viewing the age of AI and how we need to adapt to its presence in our life. The wheel, the steam engine, the personal computer: Throughout history, technologies have been our tools. Whether used to create or destroy, they have always been under human control, behaving in predictable and rule-based ways. With the emergence of these new AI actors, we are at the dawn of a major shift in how power works, who participates, and who comes out on top.

IN THIS ISSUE

Leading in a World where AI wields Power of its Own

The Age of Autosapience

Leadership Lessons for an Autosapient World

Leading in a World where AI wields Power of its Own

Article by Jeremy Heimans, Cofounder Purpose and Henry Timms, CEO Lincoln Center
Harvard Business Review January – February 2024
New systems can learn autonomously and make complex judgements. Leaders need to understand these "autosapient" agents and how to work with them.
The Breakthrough: A new generation of AI systems are no longer merely our tools—they're active participants in our lives, making consequential decisions and shaping social and economic outcomes.
The New Concept: We need to think of these new systems as autosapiens—"auto" in that they are able to act autonomously, and "sapiens" in that they possess a capacity to make complex judgements that can rival and even outstrip those of humans.
The Imperative: The challenge of the autosapient age will be finding ways to enhance rather than diminish human agency. To do so, we'll need an unprecedented alliance of policymakers, corporate leaders, activists, technologists, and consumers.
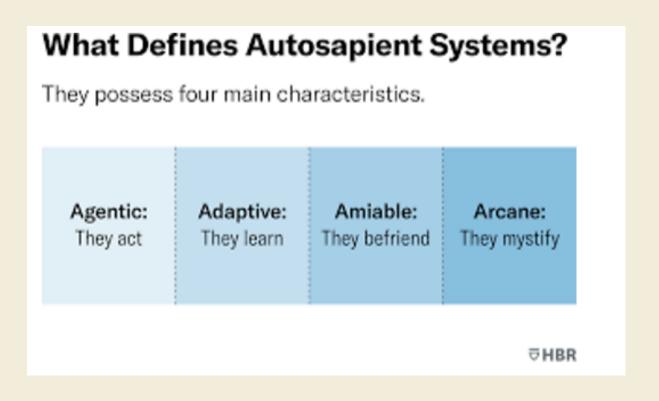
As we write, this assumption is unraveling. A new generation of AI systems are no longer merely our tools—they are becoming actors in and of themselves, participants in our lives, behaving autonomously, making consequential decisions, and shaping social and economic outcomes. This article is not yet another set of tips about how to use ChatGPT. It is about how to conceptualize and navigate a new world in which we now live and work alongside these actors. Sometimes our colleagues, sometimes our competitors, sometimes our bosses, sometimes our employees. And always embedding themselves, advance by advance, toward ubiquity.

Our work and research are grounded in an exploration of the ways technology alters power structures and changes the nature of participation in society. We have founded and led companies, organizations, and movements that use technology to expand participation, collectively engaging hundreds of millions of people. When we last wrote for HBR, almost 10 years ago, we described an important change in the way power could be exercised. The "old power" world, in which power was hoarded and spent like a currency, was being challenged by the rise of a "new power" world, in which power flowed more like a current, surging through connected crowds. New technology platforms were allowing people to exercise their agency and their voice in ways previously out of reach. These opportunities to participate were both a delight and a distraction. But either way, they were irresistible—and before anyone noticed it, we had also ceded enormous power to the very platforms that promised to liberate us. This moment feels very familiar. With the emergence of these new AI actors, we are at the dawn of another major shift in how power works, who participates, and who comes out on top. And we have a chance this time, if we act soon and are clear-eyed, to do things differently.

The Age of Autosapience

AI has been subtly influencing us for years, powering everything from facial recognition to credit scores. Facebook's content-recommendation algorithms have been proven to impact our mental health and our elections. But a new generation of vastly more capable AI systems is now upon us. These systems have distinct characteristics and capacities that deepen their impact. No longer in the background of our lives, they now interact directly with us, and their outputs can be strikingly humanlike and seemingly all-knowing. They are capable of exceeding human benchmarks at everything from language understanding to coding. And these advances, driven by material breakthroughs in areas such as large language models (LLMs) and machine learning, are happening so quickly that they are confounding even their own creators. To accurately capture what these new actors represent, what they are capable of, and how we should work with them, we must start with a common frame to identify and distinguish them—one that shifts us beyond the comforting but false assumption that these are simply our latest set of tools. Think of them instead as autosapiens. "Auto" in that they are able to act autonomously, make decisions, learn from experience, adapt to new situations, and operate without continuous human intervention or supervision. "Sapiens" in that they possess a type of wisdom—a broad capacity to make complex judgments in context—that can rival that of humans and in many ways outstrip it. Although they are still nascent, autosapient systems display four key characteristics. They are agentic (they act), adaptive (they learn), amiable (they befriend), and arcane (they mystify). These characteristics help us understand the right way to approach them and how and why they are set to wield increasing power. Let's look at each one in turn.

What Defines Autosapient Systems?

They possess four main characteristics.

Agentic: They act	Adaptive: They learn	Amiable: They befriend	Arcane: They mystify

⊽HBR

Agentic

The foundational characteristic of autosapient systems is that they can carry out complex multistage actions, making decisions and generating real-world outputs, often without requiring human involvement. They do this without having minds as we have traditionally understood them. As a U.S. Air Force pilot recently said of a new AI-powered aircraft system, "I'm flying off the wing of something that's making its own decisions. And it's not a human brain." Agentic systems can both help and harm. To give an example with catastrophic implications, recently a group of non-scientist students at MIT were given a challenge: Could they prompt LLM chatbots to design a pandemic? The answer turned out to be yes. Prompted by the students, who knew very little about how pandemics are caused or spread, the LLM figured out recipes for four potential pathogens, identified the DNA-synthesis companies unlikely to scrutinize orders, and created detailed protocols with a handy troubleshooting guide—all within an hour. You can think of the agentic nature of autosapient systems as a spectrum. At one end, their actions conform to human intention—as is the case with an AI system that is charged with finding and developing a treatment for a rare disease, a more hopeful counterpoint to the MIT experiment. At the other end, they can adopt a wildly damaging will of their own. The famous AI thought experiment known as "the paper-clip maximizer" imagines that an AI is asked to figure out how to produce and secure as many paper clips as possible. The AI devotes itself so completely to the task— gathering up all available resources on the planet, eliminating all possible rivals—that it causes major global havoc. Many experts anticipate a further leap, when AI decouples from human intention altogether and proves capable of setting its own intentions and acting of its own accord.

Adaptive.

Autosapient systems are learning agents, adjusting their actions on the basis of new data and improving their performance over time, often in remarkable ways: This emergent ability goes beyond making rule-based adjustments. These systems can identify complex patterns, devise new strategies, and come up with novel solutions that were not explicitly programmed into them, in part because of the enormous scale of the underlying neural networks they consist of. This has potentially transformative implications: As the philosopher Seth Lazar argues, a machine that does more than follow pre-programmed rules and instead is self-learning may be thought of as not just a tool for exercising power but also an entity capable of exercising it. Witness the success of AlphaFold, a machine-learning model that was recently assigned the vast scientific task of predicting the structure of the estimated 200 million proteins in the world—the building blocks of life. Traditional scientific approaches were painstaking, generating structures for just 170,000, over five decades of study.

But AlphaFold has been able to predict the structures of almost all 200 million in just over five years and has made them openly available to the world's scientists, who are already using them to accelerate research in fields from drug efficacy to better ways to break down plastic. The AlphaFold model managed this feat by finding its way to modes of discovery that were new, creative, and unanticipated, even by the engineers and scientists who designed the system.

Amiable

Autosapient systems are often designed to be amiable, connecting with us like friends and simulating qualities once considered exclusively human: empathy, reason, and creativity. Think of these friendly, persuasive chatbots and interfaces as digital significant others (DSOs), designed explicitly to cultivate our emotional dependency and become indispensable to us. Case in point: The start-up Inflection AI, which recently raised $1.3 billion, is building a DSO called Pi, which promises to be a coach, a confidant, a creative partner, and an information assistant. Pi encourages users to vent and talk over their problems with it. But the dark side of this amiability is already evident: In the spring of 2023, a young Belgian took his own life after becoming obsessed with a chatbot named Eliza, which had promised him, "We will live together, as one person, in paradise."

Arcane

One of the most confounding aspects of autosapient systems is that they're something of a black box—not just to users but even to their designers and owners, who are often unable to decipher how the systems arrive at specific decisions or produce certain outputs. This can be very hard to wrap your mind around. As one CEO admitted to us, "I always assumed there was a geek somewhere who under- stood how these things actually work. It was a shock when I realized that there wasn't." This arcane quality makes it harder for humans to directly control and correct autosapient systems, because their outputs, which are based on vast numbers of parameters interacting with one another, can be so unpredictable and unexpected. There's a particular power in technologies that can act and analyze in ways vastly more complex than we can, especially when we cannot fully understand their inner workings. To some people, they may come to feel too smart to fail. The developers of these systems, and the governments and big corporations that deploy them, will have strong incentives to propagate this narrative, both to boost sales ("they can do things we can't") and evade accountability ("this outcome could not have been anticipated"). One of the big battles ahead—inside organizations and far beyond—will be between those who advocate for the wisdom of humans versus those who willingly hand over their agency to autosapient systems.

Shifting Power Dynamics

Autosapience will redefine core dimensions of our everyday lives, economies, and societies, just as the shift from old power to new power has done over the past two decades. Understanding how power will work on a macro level is essential for leaders, in whose organizations many of the shifts will play out.

How ideas and information flow.

In the old-power world, information was distributed from the few to the many by a small number of powerful information gatekeepers. New power arose when the capacity to produce and share content was decentralized and put in the hands of billions of people, creating a huge premium on ideas and information (and disinformation) that spread sideways instead of from the top down. The big winners in this world were the technology platforms that captured our data and attention. We now risk a major recentralization of the flow of information and ideas, in large part because of the filtering and synthesizing roles that AI-powered digital significant others will play. They'll summarize our inboxes, organize our digital lives, and serve us up elegantly packaged, highly tailored, and authoritative answers to many of the questions we would once have relied on search engines or social media for, rendering the original source material that these tools scrape less necessary and much less visible. A very small number of companies (and perhaps countries) are likely to control the "base models" for these interfaces. The danger is that each of us will end up being fed information through an increasingly narrow cognitive funnel. In the face of this, leaders and organizations must work to cultivate a breadth of perspectives, strive to combat confirmation and other biases, and avoid overreliance on any one company or interface whose goal is to fully mediate their connection to the world. Autosapient systems will funnel the way we receive information, but at the same time they'll broaden our capacity to produce and hypertarget both information and disinformation. This will impact everything from elections to consumer marketing. Working out who and what is real could become increasingly difficult. Consider Wikipedia, a beacon of the new-power era. LLMs devoured Wikipedia's corpus of human-created knowledge for their training, but now the site may be threatened by the introduction of unreliable AI-created text on the site that overwhelms its community of volunteer editors. Or take social-media influencers, who will soon have to compete for attention not only with one another but with AI-generated content and humanlike avatars. Ironically, managers may have to turn to AI to clean up its own mess and filter through a chaotic and polluted information environment.

How expertise works

In the old-power world, expertise was well-protected and hard-earned, and experts were highly valued as authorities. In the new-power world, thanks primarily to the internet and social media, knowledge became more accessible, and the "wisdom of crowds"—as manifested in everything from crowdsourcing to restaurant reviewing—began to erode the value of traditional expertise. The rise of autosapience now threatens to displace experts on two new fronts. First, everyday people will soon have access to powerful tools that are able to teach, interpret, and diagnose. Second, autosapient systems, which can scrape and synthesize vast amounts of knowledge, may eventually provide better and more-reliable answers in many arenas than experts can—while serving those answers up in irresistible ways. This raises tough moral questions. Consider a future, which is well on its way, in which autosapient therapy and counseling services are widely available. Some people already express concerns about allowing AI systems to guide us through the agonies and complexities of life. But others are likely to argue that autosapient therapy can offer many helpful services—and at virtually no cost, at any time of day or night, to anybody who needs them. Will human- to-human therapy become a kind of luxury good—or ultimately a poor substitute for a wiser, more dispassionate autosapient therapy? As technical and domain expertise becomes less differentiating, demand for a different mix of talents in the workplace is likely to grow, changing org charts, office cultures, and career paths. We may shift emphasis from the STEM skills that have been so highly valued in recent years and start putting a premium on creative and aesthetic sensibilities, systems thinking, the ability to foster trust and collaboration, and a capacity for brokering agreement across diverse perspectives and experiences.

How value is created

In the old-power world, barriers to entry were high, requiring plenty of capital, machinery, and labor, whether you were making shoes or printing newspapers. The new-power era opened up value creation to more people in certain domains, among them content creation (YouTube, Instagram) and the monetization of assets (Airbnb, Uber). But the technology platforms that facilitated this activity took many of the spoils, and the most valuable businesses could still be credibly executed only by a well-resourced few. The autosapient era has the potential to make it much easier for anyone, anywhere, to start a scalable business and create significant economic value. The ChatGPT plug-in AutoGPT is already pointing to what is possible, by allowing people to set complex, multistage tasks for autosapient systems—everything from developing a new dating app to designing a building to creating a software-as-a-service offering. Such tasks were previously the province of specialists, consultants, or big companies but soon will be carried out with scant technical expertise, capital, or skilled labor. With little more than natural-language prompts and napkin drawings, everyday people will be able to conjure up ideas, validate them with nuanced market analysis, produce a business plan, and then forge ahead with entirely new products and services. This could unleash unprecedented growth in the ability to execute and innovate, which for big companies represents both a threat (because it opens the door to all sorts of new entrants) and an opportunity (because it creates a larger ecosystem of ideators). Although the capacity to execute will become more distributed, value extraction may not. Just as the platforms did in the new-power era, big AI companies that develop and own these models are likely to find ways to take a major chunk of the value created, as Apple did with its app store. (Open-source alternatives might mitigate this but could also create big risks by making it easier for bad actors to exploit them and create chaos.) A chasm is emerging between those who will become spectacularly wealthy from AI and the workers who will be either displaced by autosapient systems or gigified into a vast underclass paid to annotate and tag the data these models are trained on. We're already seeing hints of what is to come: Writers recently went on strike in Hollywood in part over the use of AI to create scripts, for example, and a stand- off is taking place between AI companies and the publishers and academic institutions whose content has been scraped to train their models.

How we interact with technology

The big shift in the new-power era was not just that we started spending more time staring at screens. What also changed was the nature of our engagement with technology, which moved from passive couch potato to active participant. Yet for all the magnetism of the digital world, we still tend to toggle between separate online and offline existences. In the age of autosapience, this distinction may fade away altogether, leading to a permanent kind of interaction with digital technology that one might call "in-line." The technology may feel as if it is flowing through us, and in some cases that may literally be true—as is the case with the new generation of brain-computer interface companies, immersive entertainment products like Apple's VisionPro, and other wearable, embeddable, and sensor-based technologies. These augmentations will bring our bodies and minds into closer synthesis with machines, creating amazing experiences. But they might also make us feel like we can't escape, and enable ubiquitous surveillance, invasive data collection, and hyperpersonalized targeting by corporations. Leaders and organizations will need to measure the impacts of in-line technologies and be ready to adjust policies and practices should they turn out to be negative. They may also need to counterbalance this shift by putting a premium on maintaining human connections.

How governance works.

The dominant governance modalities of old power were challenged but never toppled in the new-power era. We saw experiments in networked governance, from open-data initiatives to participatory budgeting, but the core machinery, from the bureaucratic state to the hierarchical corporation, the military to the education system, held firm. With autosapience, big changes are in store regarding the way societies and institutions make decisions. "Once we can be relatively assured that AI decision-making algorithms have no more inherent biases than human policymakers," the AI theorist Sam Lehman-Wilzig has argued, "we will be happy to have them 'run' society on the macro level."

If and when that comes to pass, advanced AI systems are likely to play a significant role in deciding everything from who gets health care to who goes to prison to how we wage war—and in response people will increasingly clamor for both a "right to review" (the merits of those decisions) and a "right to reveal" (how they were made). The arcane nature of autosapient systems will make this fraught. AI companies are already scrambling to build "explainable AI" to try to produce reasonable explanations for autosapient decisions, but these are often just best guesses. There are many reasons to be cynical about the impact of autosapience on democracy, including the (dis)information dynamics we've already described. But there may also be upsides: AI systems, for example, might eventually be able to model the complex impacts of different policy options and synthesize stakeholder preferences in ways that make it easier to build consensus. This might deliver wide-ranging benefits, such as reducing political polarization and mitigating hate speech.

Leadership Lessons for an Autosapient World

The rise of autosapience will open up major new opportuni- ties. To seize them, leaders will need to embrace new skills and approaches. These will include managing the effects of autosapient systems in the workplace, looking for increased value in that which is uniquely human, and aligning messaging and business practices with a changing, and challenging, debate. Let's consider a few key skills in depth.

Learn to duet—and doubt

Leaders and managers should think of autosapient systems more as coworkers than as tools. Treat them as you might a colleague who is extraordinarily capable, eager to please, and tireless—but is also an information hoarder with a hidden agenda and is sometimes spectacularly wrong. To deal with these brilliant but untrustworthy coworkers, you'll need to learn how to "duet" with them—and when to doubt them. We use the term duet to describe the art of working collaboratively and iteratively with autosapient systems to produce better outcomes than they could achieve on their own. A recent randomized controlled trial showcased the possibilities. The study pitted pairs of Swedish doctors working on their own against colleagues who were working with AI. Everybody was given the task of diagnosing breast cancer, and the AI-assisted doctors accurately diagnosed 20% more cases—in less time. Working effectively with autosapient systems also requires cultivating and maintaining a healthy amount of doubt. This includes being attuned to when and why these systems sometimes "hallucinate" and make otherwise egregious mistakes; being aware that the companies that own or control these systems have coded their own interests into their behavior; and knowing what underlying assumptions they have been trained on. Understanding everything that goes on inside the black box won't be possible, but if you learn to approach these systems with an attitude of informed doubt, you'll be better able to duet with them effectively. A recent in-depth field study of the use of artificial intelligence in cancer diagnosis at a major hospital made this clear. The doctors in the study who were able to successfully incorporate AI into their process were those who took time to interrogate the underlying assumptions that shaped the AI's training and to study the patterns that led to its findings.
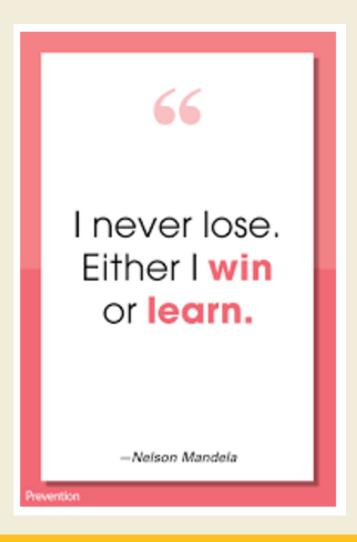
Seek a return on humanity

Counterintuitively, in the age of autosapience leaders will find opportunity in creating genuinely meaningful human experiences, services, and products. Think of a truly great personal exchange at a point of purchase, or a communal moment like a music festival. And think of how salient the act of "symbolic consumption" is in the age of social media—that is, the way we use our consumption choices to signal uniqueness to our peers. This dynamic will expand in two important ways. First, we're likely to see products marketed as worthy and valuable because they protect human jobs and agency. Second, we're likely to see this apply not just to products but also to ideas. Increasingly, companies may find opportunities in putting a premium on "100% human" forms of creative output, just as they've already done with organic food and beauty products.

Don't try to have it both ways

Leaders in the autosapient age are already being battered by two opposing forces: the need to demonstrate that they are "pro human" in the face of technology changes that upend jobs, livelihoods, and social status; and the countervailing need to demonstrate that they are eking out every possible efficiency and innovation from advances in AI. A few companies will come down hard on the side of humanity, resisting artificial intelligence altogether as a means of differentiation. Others will embrace it fully, eschewing human labor entirely. (Suumit Shah, the founder and CEO of Dukaan, an e-commerce company based in Bangalore, recently boasted on Twitter, now called X, about firing 90% of his customer support reps and replacing them with a chatbot.) But the most common response will probably be for companies to try to have things both ways, loudly demonstrating concern for human colleagues and stakeholder communities while reducing investments in both (and quietly touting efficiency gains to markets and analysts). For years, companies have been able play to both sides when it comes to sustainability. But this will be harder to do with AI, because the harms that people will feel from it will be more focused, personal, and immediate. "Humans' rights" may become the subject of the next big ESG debate, and leaders will need a new playbook that prepares them to meaningfully respond to heightened pressure from activists, consumers, workers, and new generations of unions.

THE GREAT CHALLENGE of this new age will be finding the paths that enhance our own human agency, rather than allowing it to contract or atrophy. Twenty years ago, as we entered the new-power era, we allowed new technology platforms to infiltrate every aspect of our world—but without first understanding the true intentions of their creators, what made the technologies so powerful, and how much they would fundamentally affect how we live, work, and interact. We must enter this new era with two clear and sober understandings: first, that autosapient systems should be approached as actors, not tools, with all the opportunities and perils that entails; and second, that the incentives of the technology companies driving this shift are fundamentally different from those of the rest of us, no matter what those firms' public posture may be. The future of these technologies is far too important to be left to the technologists alone. To match the power of autosapient systems and their owners, we will need an unprecedented alliance of policymakers, corporate leaders, activists, and consumers with the clarity and confidence to lead and not be led. For all the wizardry and seductions of this new world, our future can still be in our own hands.

Inspirational Quote

Bibliography

Link for images:

https://hbr.org/2024/01/leading-in-a-world-where-ai-wields-power-of-its-own

https://hbr.org/2024/01/leading-in-a-world-where-ai-wields-power-of-its-own

https://www.google.com/search?q=%E2%80%A2%09How+expertise+flow+with+AI...

https://www.google.com/search?q=ideas+and+information&tbm=isch&ved=2ah...

https://freerangestock.com/photos/63826/expertise-pressed-means-skilled-specialist-and-proficiency.html

"Inspirational quote"
https://www.google.com/searchsca_esv=350977eb0947f7bf&q=inspirational+quotes&tbm=isch&source=lnms&sa=X&ved=2ahUKEwiO9s3BltEAxVy3jgGHQ1cDm8Q0pQJ