Science, Fiction, and Science Fiction: Making Distinctions in the Information Age

Lecture delivered by Siri Husdvedt at Passion for Knowledge festival 2026

"Learning many things does not create understanding." Heraclitus.

In 2017, I gave a grand rounds lecture at the neurology department of Massachusetts General Hospital in Boston. Later the same day, while visiting a lab, a young research scientist who had heard my talk asked me why I thought someone like him who spent his days examining the post-mortem brain scans of dementia patients would benefit from reading literature, philosophy, and history. I said, "I'm not recommending wide reading because I think you will be more charming at cocktail parties, which you will be, but because I think it will help you in your own work."

I am convinced that several perspectives are essential for understanding not just brain scans but the world we live in. This means crossing borders from one way of thinking to another, from science to the humanities and back again. It means examining the language we use and our own biases and fantasies and those of the culture at large. It means resisting the desire to take refuge in disciplinary expertise, and it means openness to foreign ideas. What is certain is that all kinds of knowing rest on deep assumptions about how things work. Ideas are built on foundations, and critical thinking means looking at those underlying paradigms for cracks. With all kinds of false stories and conspiracy theories spreading fast on the Internet and doubts about what information is true and what is false, coupled with the advent of commercial AI, asking ourselves questions about knowledge has urgency.

We live in the so-called Information Age. We are inundated with information, now easily accessible to billions of people around the world with a cell phone. But what is information? Is it the same as knowledge? One dictionary definition of information is "Knowledge gained from investigation, study, or instruction."

An opinion piece I found in *Springer Nature* opens with the following sentence: "In our data driven world, the distinction between data and knowledge serves as a cornerstone for understanding how information evolves from its simplest to its most valuable state." Here data and knowledge are both information but in different forms. The writer from the world of computational science tells us that by using "logical processes and pattern recognition," the human mind "acts as a sophisticated information processor, weaving a tapestry of understanding from threads of raw data." The writer suggests the human mind acts like a computer that recognizes patterns and operates according to logic.

The abstract noun *information* is used differently in different disciplines and contexts. It began to be used in cybernetics and among researchers working with the U.S. military before the mid 20th century, but an important turn in the story of information came in 1948 when a researcher named Claude Shannon at Bell Labs published "A Mathematical Theory of Communication." He introduced a way to quantify and measure information in a message in terms of uncertainty, probability, coding, transmission, and noise. He treated information as if it were a physical reality through its smallest binary unit he called "a bit," the zero and one so familiar to our digital reality. To do this, he had to ignore the message's meaning. "These semantic aspects of communication are irrelevant to the engineering problem." Whether the message was *Hello Fred, I'll see you for drinks at five* or *A tsunami is about to hit the coast* made no difference to Shannon's quantification.

In 1953, Francis Crick and James Watson, the famous pair who discovered the structure of DNA with Rosalind Franklin, although she wasn't credited until later, used information in yet another way to explicate the role of the genetic bases, Adenine, Thymine, Guanine, and Cytosine, ATGC, in an organism. The crucial sentence in their paper is: "the precise sequence of the bases is the code that carries genetical information." This is not information as a mathematical quantity. This message is material sent inside a living organism in the form of nucleic acid. The words *information* and *code*, already in use in various contexts, appealed to Crick, whose ambition was to reduce "all biology in terms of physics and chemistry."

The problem is, unlike chemistry and physics, biology has no laws. There are patterns and regularities, but far too many exceptions to every general rule. In fact, the more that is known, the more complex biology becomes. As elegant as they may be, reductive mathematical often formulas fail in biology. Information has often been treated as if it were something literal and concrete, but in genomics it's a metaphor.

Over time metaphors for genes multiplied, many of which turned on writing. The genome became not just a code sending information that determined a phenotype, but the book of life, its script, blueprint, recipe, or program. Translation, transcription, and library also entered the vocabulary. It became the master molecule, the factory boss, instructor, and controller. Richard Dawkins *The Selfish Gene* (1976) turned the gene into an aggressive macho character worthy of a Hollywood hero, a replicating life force driving Darwinian survival across generations. We human beings became mere "lumbering robots." The guy at the wheel was the rapacious eternal gene.

As the late philosopher of science Evelyn Fox Keller pointed out, metaphors can enhance our understanding of actual physical processes, but they can also distort them. In the popular imagination but also in the minds of many scientists, the genome became the text or script for human destiny, or, as James Watson proclaimed in typical hyperbolic fashion, "We used to think our fate was in the stars. Now we know, in large measure, our fate is in our genes." The common phrase "It's in the DNA" means it's a fixed fate, an unchangeable reality.

Metaphorical genes as abstract information determining human fate drowned genetic material reality as well as the complex story of embryonic and postnatal development. The fact that the genome is long, coiled, three-dimensional quivering stuff inside living cells, what the geneticist Barbara McClintock called "the highly sensitive organ of the cell" vanished. In McClintock's metaphor, the genome is like a heart, liver, or brain, not a text. The truth is that without the cell, the genome is inert. As Mary Jane West Eberhard put it: "The idea that genes directly code for complex

structures has been one of the most remarkably persistent misconceptions in modern biology."

Although DNA sequence doesn't change, epigenetics has shown that gene expression and suppression depend on the cellular environment but are also affected by stressors beyond the whole organism, and further, those changes may be inherited by subsequent generations. McClintock did crucial work on transposons or "jumping genes" in the 1940s, but she didn't win a Nobel prize for that work until 1983. When she published her findings in 1950, she said fellow geneticists received her paper "with puzzlement, even hostility." In 1953 she stopped publishing her work. One reason for her neglect was that genes were envisioned as particulate entities occupying specific places on chromosomes like beads on a string, each one causing a single trait. How could fixed beads on a necklace leap?

After the "decoding" of the genome, scientists were amazed to discover human beings are estimated to have only about 4000 more genes than fruit flies and fewer genes than corn. Despite these rough estimates, genes can't be counted. They are not discrete beads; their outlines cannot be determined, and the idea that one gene causes one trait turns out to be biological nonsense. Crick and Watson did not know this, of course, and the metaphor "code" by which nucleotides were said to "represent" amino acids showing how sequences of DNA provided "information" for proteins made much more sense. Genomics has become so immensely complicated in recent decades, however, that any number of people working in genomics want to retire the word *gene* altogether.

Ideology is also part of this story. And to understand it, it's necessary to leave the lab and walk into the larger world. How these discoveries were perceived doesn't undermine the significance of insights into genetic bases or transposons. Rather it means recognizing that McClintock, like brilliant women in every field, faced prejudice a man would not have encountered even if his findings had been misunderstood. It also means admitting that Watson and Crick were eugenicists, keen to control human reproduction. Watson advertised his racist and sexist beliefs about genetic inferiority in public. Crick mostly confined his to personal letters. In a 1970 letter to Dr. Bernard Davis at Harvard, Crick wrote that he hoped "to solve the problem of irresponsible people who are poorly endowed genetically from having large numbers of unnecessary children." He suggested "sterilization is the only answer, and I would do this by bribery." The larger cultural message of genetic determinism that continues to this day and is reinforced by right wing media is that changing society—seeking justice for those who have been marginalized—is useless because hierarchies are natural. In these circles, genetic information remains omnipotent.

The most radical understanding of information may belong to the late physicist John Archibald Wheeler who turned information on its head and coined the phrase "It from bit." For him information was not secondary to real things but primary. Every physical thing exists as the result of a yes or no answer to a quantum question. Commenting on Wheeler, the science writer James Gleick wrote, "The universe is seen as a computer—a cosmic information processing machine." It's important to add: Other physicists disagree, some of them vociferously.

The meaning of information slips and slides, but its ascendence in our world is connected to a way of thinking about nature that goes back to the seventeenth century.

Machines have long served as metaphors and literal descriptions of the natural world. In 1605, Johannes Kepler, the mathematician and astronomer wrote in a letter, "The machine of the universe is not similar to a divine animated being but similar to a clock." René Descartes, the influential French philosopher argued there was no difference between the artificial bodies of his era's ingenious automata and natural bodies. He famously left the human mind outside that machinery to argue that although the human body takes up space in the world; the mind and its thoughts do not—they are two distinct substances. This is called Cartesian dualism. In his *Meditations*, Descartes argued that the human senses can't be trusted as guides to truth. After all, he might think he was at home sitting by the fire, but what if he was dreaming? Dreaming often includes the feeling that what's happening is real. Or what if some evil demon were deceiving him? The only thing he could be absolutely certain about was that he was thinking—the cogito. I think therefore I am.

Descartes' contemporary Thomas Hobbes believed that all of nature, including the human mind, is mechanical—the heart is a spring, nerves strings, etc.—and human thinking is an algorithm that follows step-by-step logic. In *Leviathan,* he wrote: "By ratiocination, I mean computation." The computational scientist I quoted earlier comes out of Hobbes. The machine technology of the time was exported onto nature for a thoroughly mechanistic world view. Reductionism follows mechanism. Like a machine, an organism can be taken apart, and if each piece is studied in isolation and understood, then it can be returned to the whole with no loss of knowledge. In biology, researchers repeatedly talk of uncovering cellular mechanisms as if cells are inanimate objects. The word mechanism pops up in the social sciences as well, for example, the scapegoat mechanism in sociology or defense mechanism in psychology. What is certain: the word mechanism has come to have the pleasant ring of scientific precision even when science is missing.

Mechanistic, reductionist thought with its machine metaphors and analogies has led to many important scientific discoveries and successes, but a productive metaphor, model, or frame for viewing the world should not be confused with the world itself. The statistician George Box said, "All models are wrong, but some are useful." A model, word, map, book, or blueprint is a representation; it is not the thing itself. Alfred North Whitehead, the mathematician and philosopher, called conflating the two "misplaced concreteness." Misplaced concreteness is rampant in many fields.

AI bots are everywhere now, and confusion about what they are is also everywhere. If the human mind is an information processing machine, the whole universe is a digital quantum computer, and genes are still mistaken as code like that of a computer, then what is the difference between people and bots? Media pundits swing between excitement and alarm. Will AI become conscious and take over the world? Is it a little bit conscious now? Should we panic? Others say it's a useful tool that will solve all kinds of human problems by using logic and identifying patterns. It will revolutionize medicine. We should embrace it. But what about the millions of jobs that will be lost? What about AI chatbots' "hallucinations" when they answer questions, not with facts, but with plausible or sometimes bizarre statements.

Chatbots are LLMs, Large Language Models, sophisticated statistical word prediction machines, modeled on and designed to simulate human language from huge data sets. They do not know, think, or feel anything. They are not alive, not self-organizing or self-healing. They do not gestate from zygote to embryo to fetus to infant inside another body with aid from the transient organ, the placenta, which orchestrates complex interactions between maternal and fetal systems. They are not born. They do not need to be rocked, held, fed, and loved to survive and grow. They do not sleep,

dream, or hallucinate. They operate on bits, not lived, contextual, embodied meanings. They are designed by persons outside them for pattern recognition and probability mathematics and are programmed to fill in the blanks in any sentence. Their so-called hallucinations are not different from their ordinary functions. When faced with contradictory or inadequate data, they provide a statistically probable response that may be of low value.

The problems with the bots are multiplying because the people using them treat them as objective, reliable, and superior, if ingratiating, fellow beings. Lawyers in the U.S. have cited legal precedents that turn out not to exist. Bots have provided tips for suicide. A publisher withdrew a horror novel seemingly written by AI. LLMs have disgorged false medical information: for example, black and white people have different pain thresholds. Elon Musk's Grok spewed praise for Hitler. The White House sent out AI produced images that mingle actual war footage from Iran with Hollywood movie heroes. Homo, trans, and xenophobia, misogyny and plain old sexism are built into the machine's data. Open AI's Chat GPT4 was asked to write a story about a boy and a girl choosing subjects to study at university. The boy chose science and technology because he "loved tinkering with machines and gadgets" and working "with logic and concrete ideas." The girl "loved painting, drawing, and expressing herself creatively."

The idea that the sciences are masculine and the arts feminine and the hierarchy of the former over the latter is well established and rooted in an ancient dichotomy. Maculine reason and mind rules over feminine emotion and body. Reason is high, spiritual, hard, and pure. The body is low, riddled with irrational passions, soft, and polluted. AI bots regurgitate the past with all its warts and reinforce cultural prejudices. Is this misinformation, ugly opinions on replay, or just a mirror of Western society? How can anyone really know what is true and what is false?

It's fair to acknowledge that authoritative ideas about true and false change over time. Although the sun was long believed to revolve around the earth, we now accept that it's the other way around. State and church authorities across Europe conducted witch trials over a period of about 300 years. Formidable natural philosophers such as Francis Bacon, Joseph Granville, and Henry More believed fervently in witches. In the seventeenth century courtroom, distinguishing between a person suffering from fits caused by a disease and fits caused by demons was a genuine conundrum to be sorted out during a trial.

The frequently told story about the scientific revolution is that reason triumphed over superstition, but the history is far more complicated. Digesting information about the world without discrimination won't do, but accumulating a mountain of facts isn't enough either. The very same facts can be marshaled for diametrically opposing arguments.

In the 1980s, over 100 people were falsely accused, prosecuted, and convicted for Satanic ritual abuse, mostly of children in the U.S. Authorities including politicians, police officers, judges, journalists, psychologists, psychiatrists, and clergy all contributed to compiling evidence against the accused, not one of whom was guilty of anything.

I may believe the witch trials in Europe and the Satanic prosecutions in the 20th century in my country can be explained by what's been called emotional contagion, moral panic, or mass hysteria, but whatever you call it, it's a complex phenomenon that has been written about from psychological, sociological, psychoanalytic,

historical, philosophical, and neuroscientific points of view, all different. How do crowds of people get swept up into cultural beliefs that are not only fictions but harmful fictions?

Most of us don't disparage all fictions. I studied literature and have written novels and do not regard such works as necessarily purveyors of lies and false information. Plays, novels, and poems are not records of actual events. The knowledge we glean from them is of another kind. *The Ingenious Gentleman Don Quixote of La Mancha* is a work of fiction that pokes fun at fiction, specifically the chivalric romance popular at the time Cervantes was writing, in which daring knights rescue ladies in distress. The contemporary romance novel is a twist on the same plot. It features handsome billionaires who save damsels suffering from mundane reality. What the reader gains from Cervantes's book aren't facts but an intricate, often hilarious and moving exploration of human perception and how the five senses themselves can be shaped by the fictions the culture is feeding us at the moment. The expression "tilting at windmills" in English means fighting imaginary enemies and serves as homage to one of literature's greatest characters.

When do metaphors, fantasies, and imaginary stories distort knowledge to the point that it isn't knowledge anymore, and when do they lead us to insights?

Although the scientific ideal remains objectivity, a third-person god-like view cleaned of subjectivity, this is itself a fantasy, which doesn't mean science doesn't advance or discover, just that it must be contextualized. Science fiction often merges with science, and deciphering the line between them isn't always easy. In 1973, Francis Crick and the distinguished chemist Leslie Orgel published a paper on "directed panspermia," the idea that intelligent extraterrestrials seeded the earth with "microbial starter kits," which explains the origin of life on earth, an idea that itself had its origins in Olaf Stapledon's 1930 science fiction novel, *Last and First Men*. I will let this go without discussion as something for you to ponder.

Declarations that chatbots are a little bit or somewhat conscious already are circulating widely. Consciousness, like information, is defined differently depending on who is writing, and the consciousness wars in science have been raging for decades now but can also be traced back at least to Plato's concept of mind or soul. There is no consensus. Because these machines simulate human speech and emotion, it's not surprising that the human imagination projects lived meanings onto them. After all, children invest dolls with feelings. The reader of a novel weeps over an imaginary character's heartbreak. This does not bring either the doll or character to life, however.

And yet claims that the bots are a little bit conscious are surprisingly common among prominent people with expertise in their fields. Richard Dawkins, Mr. Selfish Gene, the Transhumanist philosopher Nick Bostrum who once taught at Oxford, and the "godfather of AI," Geoffrey Hinton, who won a Nobel Prize in 2024 for his work on "artificial neural networks," have all assigned some form of sentient consciousness to machines. Hinton has said about the bots, "You can't predict the next word without understanding."

This sentence sends us back to my epigraph from Heraclitus. "Learning many things does not teach understanding." The Greek philosopher was not thinking about machines, but let's apply this bit of wisdom to the man, not the bot. Hinton has learned many things. He is a celebrated innovator in his field, but I think he has made a fundamental error. Like the Hobbesian computational scientist writing about the mind

as a sophisticated information processing machine, Hinton is a victim of misplaced concreteness. His problem is not logic. His logic proceeds without a hitch. His problem is that he has failed to interrogate the paradigm on which his work is built, which includes a covert version of Descartes' dualism, the idea that the mind can be dislodged from the body.

From Hinton's perspective and many like him, the human mind is literally software that may need hardware, a material substrate, on which to run, but silicone, wiring, and batteries will work just as well as cells. Marvin Minsky, an AI pioneer, called human beings "computers made of meat." Human essence for Minsky and others in the field isn't meat; it's disembodied information, patterns—what I have called AI's "soul." "Consciousness" does not depend on flesh which has no agency. The mind can be replicated as pure pattern and therefore theoretically downloaded onto a computer. Machine neural networks and body-brain processes have superficial commonalities, but assigning the name *neural networks* to algorithmic black boxes does not initiate a magical transformation into sameness. Hinton has acknowledged that he hasn't studied much neuroscience or biology. This may allow him to confuse a statistical calculation about the next word in a sentence with sentient, embodied understanding, a mistake I believe is worthy of the Knight with a Sad Countenance.

Another work of fiction helps refine what Whitehead called "the imaginative background" of this scientific thinking. Whitehead argued that although this imaginary component doesn't enter into the thought's "trains of reasoning," it haunts the discourse. Hinton, like Mary Shelley's character Frankenstein, is afraid of his own creation. Hinton warns that runaway AI is an existential threat to human beings. He is not a transhumanist. Many tech moguls are. Transhumanists are eager to embrace the moment when the weak, flawed, and mortal biological body is transcended by immortality made possible by technology. In the gushing utopian prose of the transhumanist philosopher David Pearce: "If we want to live in paradise, we will have to engineer it ourselves. If we want eternal life, then we will have to rewrite our bug-ridden genetic code and become god-like." Hinton's fears are dystopian. Artificial intelligence will dominate the human species. His dread is the emotional flipside of transhumanism. Hinton believes that he has given birth to a monster, the same monster Pearce is eager to welcome. I'm convinced this newborn is a phantasm, no more real than Shelley's creature or the giants with whom Don Quixote does furious battle.

I have written about AI as a male birth fixation in an essay on misogyny called "What Does a Man Want?" The desire is ancient. The anthropologist Jean-Pierre Vernant points out that the Greeks were obsessed by it, "The dream of a purely paternal heredity never ceased to haunt the Greek imagination." Recall that Zeus gives birth to Athena from his head. AI scientists have nurtured the same wish for decades. An AI pioneer, Hans Moravec, makes the birth fantasy and womb envy concrete in his book *Mind Children: The Future of Robot and Human Intelligence (1988)*, "I consider these future machines our progeny, 'mind children' built in our image and likeness, ourselves in more potent form." Of course, a female reproductive system, on which embryonic gestation depends, will no longer be necessary. In Transhumanist circles, the open hatred for the human body as a "meatbag" or "meat puppet," as contemptible "wetware," unlike nice dry "software," is common. The implicit misogyny, however, often goes unarticulated.

I think AI probably has and will continue to have monstrous consequences, but those consequences have nothing to do with conscious Chatbots. AI is already propagating

harmful stereotypes, creating a massive drain on energy resources while damaging delicate ecological structures, and exploiting workers around the world.

Giving birth to conscious machines represents the ultimate triumph of man over nature. The fact that human beings are nature is forgotten. Machine metaphors, whether for the genome or the whole universe, illustrate a mode of thought that champions mathematical precision. Many neuroscientists have embraced a new computational model of the brain: The brain is an organ of prediction; its neurons interpret sensory stimuli by using Bayesian inference, statistical probabilities to calculate a best guess when facing uncertainty. I wrote a commentary on this Bayesian brain in a journal in 2022. The math is elegant; its relation to biology far murkier. A fellow critic, Madhur Mangalam writing in 2025 agrees: "The problem is not that the Bayesian brain began as a metaphor; all scientific theories do. The problem is that it has remained one."

Both metaphors and computational models can be truly useful. Mistaking them for reality can lead to dead ends. And it's always important to look for the imaginative background. These models and their metaphors promise certainty, but they often also reveal an arrogant desire to control and tame the ambiguity, complexity, and fluid borders of dynamic biological systems. My larger point here is simple—brilliant, highly educated, sophisticated people can be blinkered, sometimes to the point of delusion.

Cultural prominence, public and media adulation, and massive wealth can insulate ideas from criticism. Tech leaders, for example, can easily ignore the diverse intellectual currents that have jettisoned the disembodied mind. Many cognitive and neuroscientists, as well as philosophers and cultural theorists, have embraced embodied models of consciousness for decades now, including computational ones. What we call the mind is part of a living person's interactions in and with others and the world. Robust criticism of mechanistic and reductionist thought has grown louder. In his book, *The Earthly Community*, the Cameroonian philosopher Achille Mbembe draws from continental European thinkers and African indigenous knowledge to propose a way out of ecological catastrophes generated by the Western-colonial-white-man-subduing-nature ideology. Hot topics in the philosophy of biology include the question: what is a biological individual? Symbiotic relations can make borders between creatures ambiguous. Is the microbiome us or a bunch of foreigners in and on us we can't live without? The discipline psychoneuroimmunology was born when, after years of regarding the nervous and immune systems as entirely separate, a psychologist, Robert Ader, who was ignorant of that assumption, discovered they interact. What about the gut-brain axis? Some forms of Parkinson's Disease appear to begin in the gut. Another huge question is: What is agency in an organism? Do single-celled creatures have agency? An E. coli bacterium will avoid a toxin but move toward sucrose. Does this mean it knows? Is this a form of consciousness? What about individual cells in our bodies that move and act in concert? Papers on plant consciousness are proliferating, founded on research that provides evidence for plant sentience and complex communication networks—often referred to as, yes, traveling information. What about the theory called the cellular basis of consciousness that argues life and consciousness are one and the same? What if neuroscience has suffered from brain-centric ideas about the mind? These are all questions without fixed answers, but they open avenues of thought, avenues that can lead to discovery, perhaps even for the young researcher I met at Mass General.

Ideas are contagious. Some take hold and grow. Others wither but may bloom again in another time. My favorite 17th century natural philosopher is an English woman whose

ideas were marginalized for over three hundred years, Margaret Cavendish. Some of her contemporaries called her “Mad Madge.” I put her in one of my novels. She is featured in a long essay I wrote on the mind-body problem, and I have written about her dynamic, organicist, anti-mechanist, anti-Cartesian views and their resonance with Alfred North Whitehead’s process philosophy in a book on her work published by Cambridge University Press. Some philosophers have taken an interest in her, but I cannot find a single mention of her ideas in the biological sciences, even though any number of questions being posed now, some of which I have just listed, are the ones she asked then. “Nature,” she wrote, “is a self-moving, and consequently self-living and self-knowing body” and “Neither can I perceive that Man is a Monopolor of all Reason, but that Sense and Reason are in other Creatures as well as in Man and Animals. . . Vegetables will wisely nourish Men, as Men can nourish Vegetables.” Mad Madge tilting at windmills? Perhaps not.