

# Step Into the Free and Infinite Laboratory of the Mind

Science fiction is one of the best tools we have for staging inclusive, engaging conversations about science and technology policy. How can we make better use of it?

**A**liens! Killer robots! Spaceships! For a long time, I saw science fiction stories in a fairly conventional way—as a touchstone or a starting point for a conversation about a topic such as artificial intelligence (*Terminator* or *I, Robot*) or genetic engineering (*Gattaca* or *Frankenstein*). These stories act as a kind of collective shorthand, one that gets a crowd on the same page before conversation turns to the “real” challenges facing society.

That was before the future became my job. When I founded the Center for Science and the Imagination (CSI) at Arizona State University in 2012, I began to understand science fiction as part of an important feedback loop with real scientific and technological innovation. Science fiction and science are both engaged in exploring the *adjacent possible*: the worlds we might be able to reach from our present configuration of science and society. Writers and researchers scour the literature, looking for ideas that might be combined or extended in novel ways, and both groups are in the business of creating new stories or new concepts that might change the world.

That feedback loop is not merely about technological change—it’s about the full human consequences of that change. The difference between a good science fiction story and a patent application is the human beings, the characters who work through a conflict and complete some kind of narrative arc. Humans are much better at thinking about the future when we can also *feel* it, empathizing with future people who are living with the benefits and costs of particular innovations. And so good science fiction does not dream

up just the automobile, but also the traffic jam, as writers such as Isaac Asimov and Frederik Pohl have argued. It’s a dynamo we at CSI have sought to harness to explore the near-future possibilities in policy arenas like energy infrastructure, space exploration, and climate change.

Science fiction is a free, infinite laboratory of the mind that allows its audience to envision possible futures *in context*. By centering characters—people—instead of technologies, writers have to offer concrete answers to those nagging questions which can be so easy to gloss over when an invention or discovery exists only as a concept. Who will own it, use it, pay for it, maintain it? Where will it be installed or deployed? What does it look like, smell like, feel like? How does it actually work? Does it need to be plugged in? What if you drop a piece of toast into it? What else must be true in the future for this thing to exist? These questions create what I call *speculative specificity*: The craft of effective storytelling pushes authors and readers to play out second- and third-order consequences and to imagine the full context of a changed future.

## A way to “feel” the future

Well-crafted stories of possible futures invite readers to participate in those futures, to walk a mile in a character’s shoes and feel what it might be like to live in that world. They combine two key functions of the human imagination: anticipation and empathy. This is why in 2016 CSI launched Future Tense Fiction, in which we commission speculative fiction writers to create compelling

visions of possible near futures, ask artists to interpret those worlds through visual media, and task policy and scientific experts with writing nonfiction responses that consider the implications of those futures for the present. Bringing this project to *Issues* in 2024 has felt like a homecoming, because we are exploring the social implications of new knowledge and innovation through a creative lens that beautifully complements the work of this journal.

Science fiction taps into humanity's innate cultural abilities to process and manipulate narratives. As a species, we're pretty bad at statistics, higher-level systems thinking, and complex risk assessment, but we're generally good at stories. A well-told story can incorporate ambiguity, tension, nuance, and contradiction—and it can incorporate all those things in a single character. Writers can use foreground and background, allusion and inference, to depict a complex situation in a way that is broadly accessible. (I'm using "stories" to describe text here, but the same applies to film, audio, and other media.)

Stories are microcosms that embed not just the "stuff" of a world (characters, settings, objects, etc.) but the ruleset: the causal models by which that world operates. A fairy tale, a whodunit, and a romance all follow causal models that are implicitly or explicitly embedded in the text. When readers ingest these stories through the page, they recompile a version of this new world in their own heads. Ursula K. Le Guin's wonderful book *The Left Hand of Darkness*, for example, describes a world in which gender and sexuality are radically different, a departure from our familiar reality that has all sorts of broader consequences for the characters—and potentially for readers.

Once you've read a piece of science fiction, you can peek around the corner—inventing a new setting or character, say—because you can extrapolate based on the model conveyed to your imagination by the author. More importantly, you can begin to compare that fictional world to your own, stepping out of your given reality and looking back at it from a strange new horizon. Sometimes it takes voyaging to a speculative future to perceive something that's staring you in the face right here in the present. When we start to see stories as microcosms, we begin to understand how they allow us to explore different structures of power, theories of change, and perhaps even new genres of innovation.

I'll pause here for a reminder that the role of science fiction is not predictive, but exploratory. For every successful exercise in anticipating the future (Jules Verne's submarines in *Twenty Thousand Leagues Under the Seas*, for instance, or E. M. Forster's prophetic depiction of the internet in "The Machine Stops") there are a zillion misses. *Star Trek* is celebrated for dreaming up the communicator that prefigured early mobile phones, but the show, like nearly every other sci-fi vision of instantaneous universal

communication, mostly missed the traffic jams: social media, disinformation, spam. But that's OK, because the point of *Star Trek* was not to offer a detailed road map from the Apollo program to the warp drive—the point was to change how audiences feel about the future, to create a complex and compelling *protopia*: a future in which things are not perfect, but keep getting better. *Star Trek* modeled diversity, equity, and inclusion, empathy and ingenuity, depicting a future driven by humanity's highest scientific and social ideals. At the same time, it was a flawed articulation of those ideals grounded in its time and place. It's a vision of the future that holds up even when the gadgets it imagined seem quaint or outdated because it articulates core values and a theory of change, a *genre* of progress, that still inspires.

### Daring to imagine

I like to call this practice of bringing speculative thinking into real-world science and technology debates *policy futures*, because the same qualities that make science fiction such a great tool for engaging the public in questions of science and society make it effective for the experts as well. Increased specialization in scientific and technical fields has created a galaxy of brilliant minds who lack the *Star Trek* communicator: They don't know how to talk to one another across the inky voids between disciplines. The engineer says she's not an ethics expert. The philosopher doesn't understand the datasheets. The policy expert is stymied by the blueprints.

Specialization and siloed thinking can give everyone a hall pass to avoid considering the unintended consequences and broader ramifications of their work. And yet to create effective science and technology policy, all these people need to have a mutually intelligible conversation about possible futures that is *also* accessible and engaging to the public. The author Neal Stephenson, a longtime collaborator at CSI, once suggested that a good science fiction story can save many hours of PowerPoints and meetings because it can—literally—put everyone on the same page.

When experts and the public step into that free, infinite laboratory of the mind, they can travel to any number of possible futures with particular configurations of technological novelty, social and policy change, and reactions to or broader consequences of those new things. Better yet, once such a story-future exists, *anyone* can change the settings in the imaginary mental lab. I like to use the analogy of furniture in a room: Once a science-fictional couch has been described, everyone who reads that story can meaningfully participate in a discussion about where it should be placed.

Science fiction starts with speculative specificity, but it also awards agency to the audience, which has the

power to question or reimagine any aspect of this future. This makes it a fundamentally hopeful exercise; imagining possible futures and comparing them builds in the work of deliberation and debate. We'll never get to those better futures if we don't imagine them first. And while science fiction has given society plenty of cautionary tales and dystopian classics to choose from—yardsticks we can use to measure against the futures we *don't* want—we have very few protopian yardsticks we can use to measure our progress toward a better future for all.

There are many ways to bring science and science fiction into dialogue. Our preferred method at CSI is to host structured workshops where participants from diverse perspectives work together to collaboratively imagine a future using creative exercises and tools, collectively described as *worldbuilding*. Over the course of a day or two we ask small teams—which, depending on the project, might include engineers, social scientists, artists, and fiction writers—to identify the big ideas that will shape a particular kind of future. (For example: *It's 2050 and a*

they would never be allowed to articulate in a conference presentation or peer-reviewed article. The incentives, power structures, and territorialism that often accompany specialization can breed incremental thinking and tunnel vision when it comes to advancing real change. But imagination is a cognitive capacity available to everyone, and it can be developed like any other skill. With practice and guidance, anyone can imagine the future, and the simple exercise of envisioning a victory condition or a radical change for the better can be a powerful tonic.

The climate futures researcher Manjana Milkoreit conducted an extended study of the cognitive context of policy experts working in the United Nations international climate negotiations. She found that despite their expertise and, for some participants, decades of experience working on influencing humanity's future on this planet, her subjects had a very limited capacity to imagine possible climate futures. They generally did not have compelling stories about what “success” might look like—at best, victory might be changing projected global

## Humans are much better at thinking about the future when we can also *feel* it, empathizing with future people who are living with the benefits and costs of particular innovations.

*city like Phoenix has successfully transitioned to renewable energy; this new infrastructure is highly centralized and sited in urban areas.*) From there participants drill down into timelines, important people, places, technologies, and points of friction. They might flesh out key characters and plot points for a story set in this future world. Typically, we ask everyone who participates to contribute something within a few weeks of the workshop: a short story, an essay grounded in technical expertise, or a piece of art, all grappling with the same speculative future questions. The CSI team edits and compiles these contributions into a book or other output exploring multiple speculative futures on the same general theme. These books and media products go on to spur new conversations about what a positive future might look like, and how society can get there.

The speculative part of speculative fiction (I'm using this term interchangeably with science fiction, loosely defined as “useful stories about possible futures”) brings another benefit to the policy world: It's made up! Make-believe, pretend, fictional. There are many experts who have participated in collaborative imagination projects at CSI who will, under the aegis of speculative leaps into the future, dare to imagine the kind of positive, transformative change

temperature rise or carbon parts-per-million from one number to another. If those who are best informed about the real context and possibilities of the human response to the climate crisis cannot imagine what our future should actually look and feel like, how can the rest of us do it?

### **An early warning system for the near future**

This brings me to the most important aspect of science fiction for science and technology policy. A lot of commercial science fiction explores the distant future because it is easier to posit exciting new technologies and radical change (and it's harder to be proven wrong). In the field of artificial intelligence, for example, our cultural imaginaries of the technology are still dominated by a familiar clutch of old nightmares, especially the autonomous killer robots (*Terminator, Ex Machina*) and the godlike supermachines (*The Matrix, Her*). Perhaps these threats loom on the horizon, but our social imaginary about AI has lost sight of the road immediately ahead.

Humanity needs to get better at telling compelling stories about near-future change, because the accelerating pace of technological innovation is rapidly delivering

## RULES FOR SPECULATORS

**Embrace the Adjacent Possible:** Bring together diverse perspectives and bodies of knowledge, including the creative arts; humanities; social sciences; and local, Indigenous, and community knowledge; as well as technical and natural science expertise.

**Speculative Specificity:** Use the tools of worldbuilding and speculative fiction to create concrete, richly detailed, nuanced, complex future worlds and populate them with compelling characters who have real problems to solve.

**Collaborative Imagination:** Imagination is a team sport. Policy futures can only be effective when it engages broad audiences and cuts across disciplines, ideologies, and other social fault lines. A good story will travel far.

**History and the Future Are the Same Place:** We rely on imagination to access both the past and the future, and the stories we tell about history offer many lessons for how the world could be (again).

**Two Futures (or Three) Is Better Than One:** Building imaginative capacity requires exploring the full possibility space of multiple futures, especially hopeful futures. Avoid imaginative path dependency by practicing exercises of radical imagination and empathy.

**Have Fun:** Seriously. To change the future, we need to change the stories we tell about the future. But we also need to change how we feel about the future, making it a site of hope and joy rather than anxiety and fear.

futures that we don't have the words to talk about, much less regulate. We have no good metaphors to grapple with the latest iterations of tools from companies like OpenAI, which are already transforming whole economic sectors (including mine). Designers give these systems human names such as Claude and Siri, or eerily human voices, like the latest iterations of ChatGPT; but to pretend that they're actually similar to people is to make a dangerous category error. It might make more sense to understand a generative large language model as like a forest or a flock of starlings: a large, complex system that the user can only partially perceive.

Government and industry continue to invest power and agency in computational systems whose inner workings and possible effects we all struggle to describe. This is a problem, because while AI is sifting through job candidates, evaluating loan applications, and making decisions in health care, logistics, and finance that could affect huge numbers of people, we cling to simplistic metaphors like the "digital assistant," apocalyptic language about "existential threats," or dismissive anachronisms like "glorified autocomplete." We lack the words to describe what these tools are actually doing right now and in the near future, or how to articulate what equitable, just, and inclusive futures with AI could look like.

Several years ago, some colleagues and I conducted a taxonomy study to understand how AI is depicted in contemporary science fiction. We were hoping to learn whether the speculative side of our cultural imaginary on AI has better answers for what kinds of stories we should be telling, as policymakers try to understand and regulate these rapidly evolving tools. The answer? Nobody knows what AI is—even in fiction, where the author can make all the decisions about how the world works. The AI systems we studied *in fiction* were ambiguous in terms of their agency, their boundaries or extent of operations, the question of who owned or controlled them, etc. Little wonder that we collectively have such a poor grasp of what AI means in the real world, much less how to manage it.

I offer this not as a lament about our lack of imagination in this arena, or the unsurprising news that humanity's technological reach has once again exceeded our cultural grasp. The point is that we need more policy futures: research-based, technically plausible, hopeful stories about scientific and technological change. The policy and regulatory apparatus for science and technology is understandably oriented to thinking about what could go wrong. But to engage the public's imagination, policy futures also need to embrace hope and the possibility—or better, the design aspiration—for positive change.

A second insight: Policy and technology decisions are deeply informed by the adjacent possible futures that science fiction offers up. Dissecting how society has imagined emerging technologies can reveal not only our fears, but all the things we're failing to talk about—all the potential traffic jams we might have missed. Doing some science on the fiction can reveal new ways to consider policy and shape the larger conversation about how we should adapt to change.

I knew we were on to something with Future Tense Fiction when we published our first story by Paolo Bacigalupi and the response essay, by legal scholar Ryan Calo, noted that Bacigalupi's story "revealed an important connection in robotics law that had never before occurred to me." More policy (and policy-adjacent) experts need to engage in the speculative, generalist work of imagining a complete future. This is not something anyone has to do on their own; in fact, we've learned from over a decade of work at CSI that the results are much better when groups collaborate to imagine futures they can mutually agree on and contribute to. Imaginative capacity can be expanded, both on an individual and an organizational level. The most important outputs of our workshops and projects are not the stories we share, but the people who have been affected and see their work differently, like Ryan Calo, or readers who are inspired by one of our stories to change their own futures.

### Moving beyond incrementalism

Building this mental laboratory into the everyday practice of policy work and scientific research will accomplish two basic goals. First, it will improve the anticipatory capacities of those involved by helping them to identify and address the unintended consequences of new ideas and potential roadblocks to positive change. Second, it will invite broader public engagement into debates that shape society at large, but are often overlooked in the current model of expert-led, incrementalist policy work.

What might this approach look like? One example is the story of N Square, a funder, network, and skunkworks initiative focused on reducing the risks of nuclear weapons and working toward nuclear disarmament. Years ago, I helped facilitate a futures workshop they convened with a number of policy experts in the field. These participants found it extremely difficult—almost impossible—to imagine a world without nuclear weapons. For several, there was simply no future with a plausible path to effective global governance of all nuclear technologies and the achievement of "global zero." It was striking to see how hard it can be to imagine a transformational positive change, even, or perhaps especially, for people dedicated to working toward that change.

After several years of collaboration with futurists and speculative fiction writers (at CSI but also a number of other places, including the Center for Complexity at the Rhode

Island School of Design), N Square decided to reform itself around Horizon 2045, a project on understanding interconnections between global risks and fundamentally reimagining human and planetary security. One of their key outputs so far has been Far Futures, wherein writers, artists, musicians, and policy experts were commissioned to imagine a world in which nuclear weapons have been removed from the human story. Horizon 2045 plans to extend Far Futures to create protopian visions for climate, public health, AI, and democracy.

I have talked about the mental laboratory of speculative fiction as free, or almost free. But it does cost something, beyond coffee and donuts or the commissions CSI pays to writers and others for their time and creative work. (Such payments are, I should note, pocket change compared to the billions of dollars invested annually in science and technology research, often funded with only the flimsiest examination of potential societal consequences.) The real price lies in questioning assumptions about what progress really means—first for participants in the room who have to examine their own preconceptions, but also for the broader policy communities, stakeholders, and members of the public who might start asking the same questions after experiencing a very different speculative future. Are social systems for innovation driving transformational change (and if so, for the benefit of all or for only some?), or are they only endorsing modest, incremental shifts that perpetuate existing structures of power and investment? A future where many more people feel invited and empowered to imagine their own futures, or a future defined by the tiny sliver of humanity who are currently privileged to realize their stories about what tomorrow should look like?

Letting more people—ideally everyone—answer these questions might cost those who benefit from the status quo. Empowering people in such a way may push them to have uncomfortable conversations about power and agency, or to realize that the biggest changes humanity needs to make to achieve a better future are social and ethical, not technological. But that price is well worth paying—in fact, I think the survival of our species depends on it. Building imaginative capacity is an essential collective project if we hope to successfully navigate the tidal waves of change that have already been unleashed, like AI, climate chaos, and the complex global economy. Engaging in grounded, structured acts of speculation should become an essential part of every policy process, because doing so allows all to imagine and debate not just how the world could be, but how it should be.

*Ed Finn is the founding director of the Center for Science and the Imagination at Arizona State University, where he is an associate professor.*