



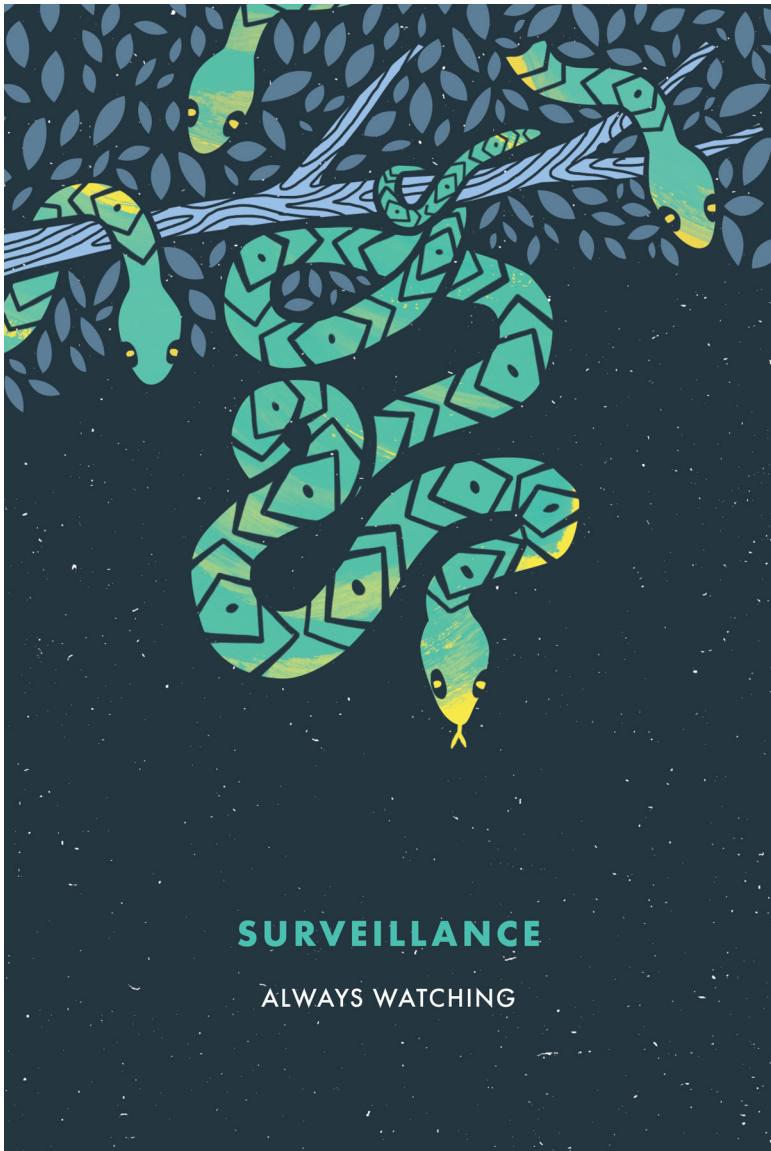
TECH RISK ZONES

As makers and builders, we're responsible for what we put into the world. That's where the eight Tech Risk Zones come in.

Split across two cards—one for understanding and one for deeper questioning—each Tech Risk Zone is designed to start thoughtful conversations around risk, responsibility, and impact. There's also a blank card you can tailor to a specific scenario for your organization.

We hope these exercises build further support for leaders like you who want to create tech that's safer, healthier, fairer, and more inclusive.

These cards are just the beginning. Visit ethicalexplorer.org for more content and to expand your experiences with the Ethical Explorer Pack.



SURVEILLANCE

ALWAYS WATCHING

FOLD



SURVEILLANCE

Privacy is in peril. From facial recognition to location tracking, technology can constantly monitor individual behavior—whether someone opts in or not.

The commercialization of surveillance means it's not just governments that can act on tracking and monitoring information—private companies and communities can, too. This information can affect everything from banks determining who gets a loan and governments deciding who is granted citizenship to police arresting individuals for political expression.

The use of surveillance tools to monitor is not a new concept—but the scale at which everyday technology, from phones to video doorbells, surveils people presents an unprecedented challenge to privacy.

.....
**How will we
protect privacy?**



SURVEILLANCE

WHERE WE STAND

Can someone use our product or service to track or identify other users?



Is this how our tech is intended to be used?

Do we have privacy features that limit the collection or sharing of personally identifiable information?



What could governments or third-party organizations do with the information we collect?

FOLD



How will we protect privacy?

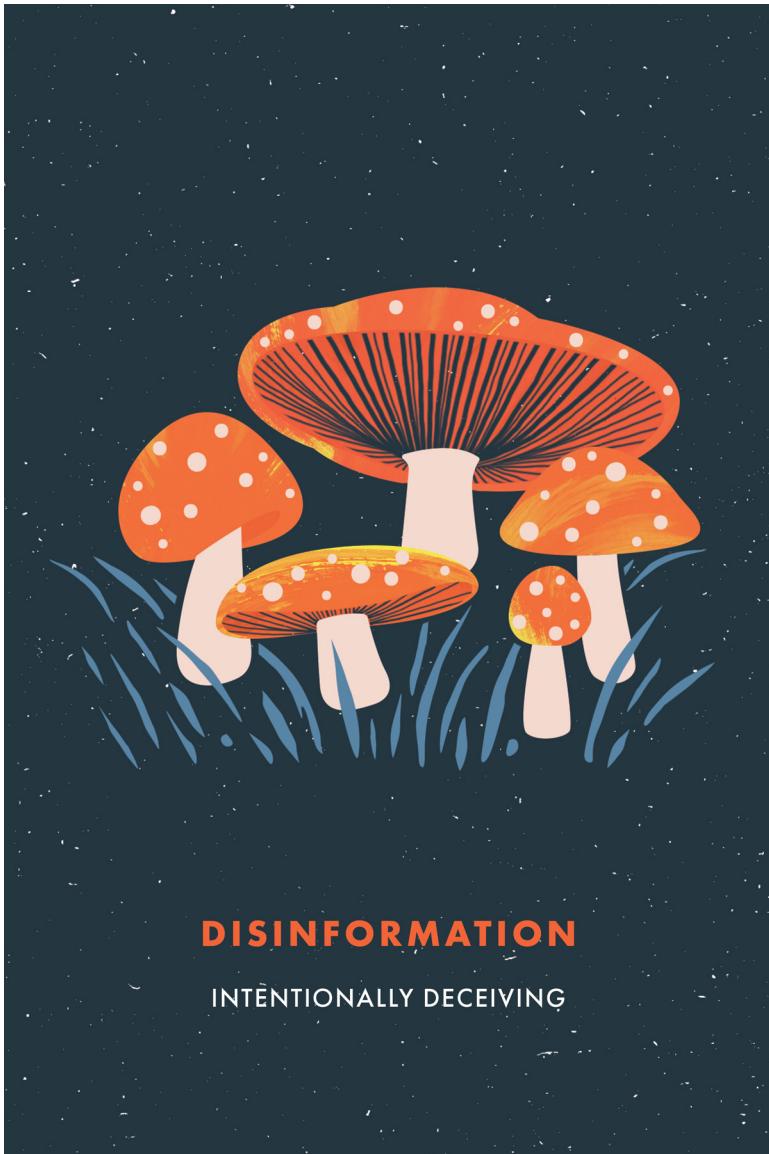
ANTICIPATING RISK

How might our technology be used to discriminate, oppress, or target specific groups?

How would we respond to law enforcement or government requests for user information?

LEADING THE WAY

What policies can we implement to ensure we balance a great user experience with protecting privacy?



DISINFORMATION

INTENTIONALLY DECEIVING

FOLD



DISINFORMATION

Shared facts are under attack and dangerous disinformation is on the rise. Bots pose as real people. Deepfakes use altered speech or images to convince people of things that never happened. And some political ads contain outright falsehoods.

The speed and breadth at which disinformation spreads is alarming. One study showed that false news spreads up to six times faster and is 70% more likely to be shared than facts. Disinformation can also have real impact during times of crises, like health pandemics.

There have always been individuals and groups compelled to subvert the truth, but the digital age provides them unprecedented efficiency and scale.

.....
**How will we
promote truth?**



DISINFORMATION

WHERE WE STAND

Does our tech have the ability to influence people's worldviews?



Do we have systems in place to prevent the dissemination of falsehoods?

Are people using our tech to subvert or attack facts?



Are we responsible if disinformation is spread using our tech?

FOLD

How will we promote truth?

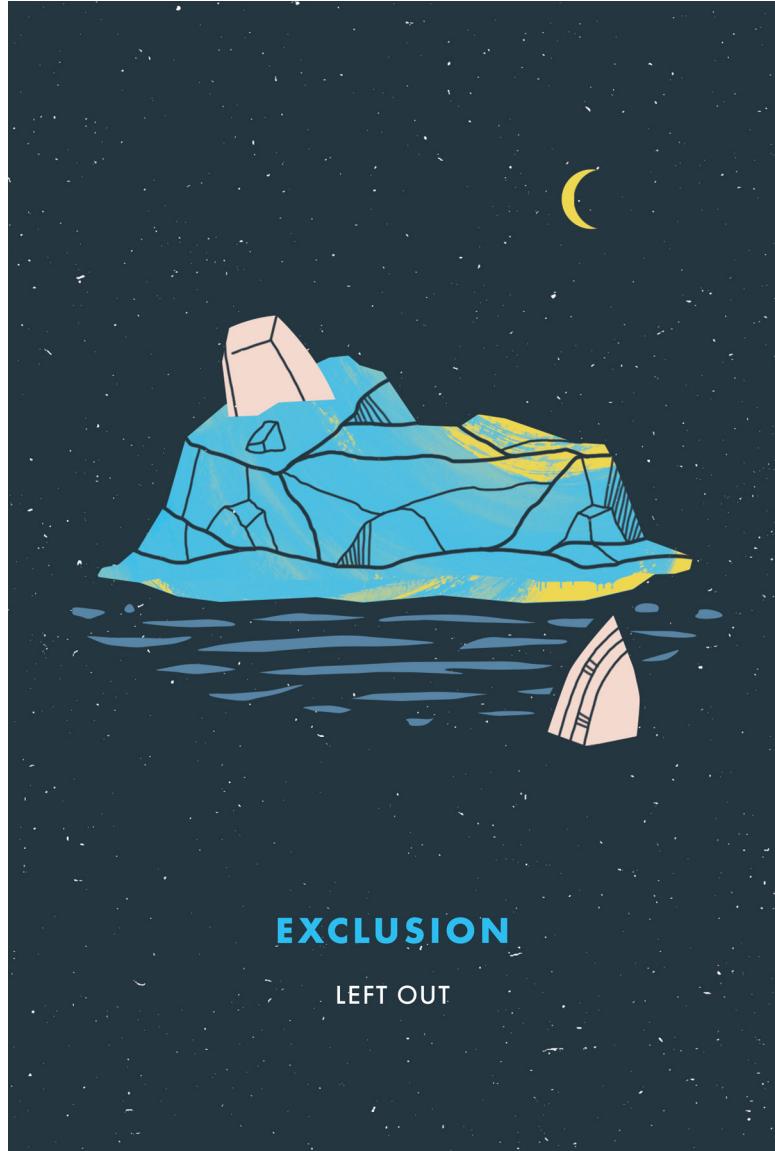
ANTICIPATING RISK

What could become the equivalent of false news, bots, or deepfakes on our platform?

How might our tech be co-opted to undermine trust in our societal institutions?

LEADING THE WAY

What might fostering an environment for facts to flourish look like in our tech?



FOLD



EXCLUSION

Technology products and services touch many aspects of our society, yet only a limited few influence and create them.

Technology's benefits are increasingly a privilege for the wealthy and digitally literate. Meanwhile, historically marginalized populations suffer higher consequences from a lack of data privacy and literacy, like overrepresentation in the criminal justice system and decreased access to tools that impact social or financial mobility.

When technology is designed *for* a group of people, and not *with* them, inequities inevitably emerge.

.....

**How will we
enable equity?**



EXCLUSION

WHERE WE STAND

How did we decide who our target audience should be?



How can we benefit from more diversity in our audience?

Are we missing any perspectives that would provide important inputs on our product?



How can these perspectives help us better meet our goals?

FOLD



How will we enable equity?

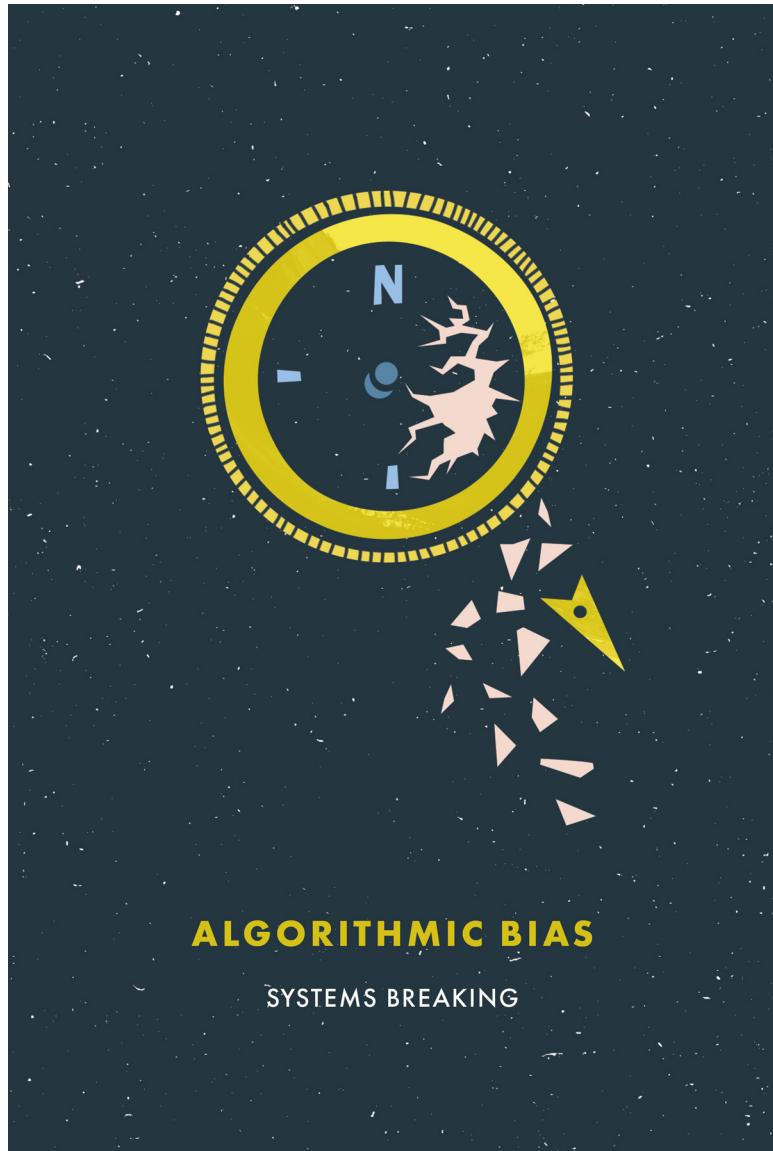
ANTICIPATING RISK

How would society be impacted if marginalized groups couldn't use our product?

How could our technology potentially be misused to harm or exclude certain populations?

LEADING THE WAY

How might people benefit if our design and decision-making included people from historically marginalized populations?



FOLD



ALGORITHMIC BIAS

Objective AI is an illusion. As humans, we all have unconscious biases that impact algorithms, potentially causing or amplifying harm through predictive policing, hiring decisions, and more.

Researchers found racial bias in a medical algorithm used across the US to determine which patients require further care. Although the algorithm didn't explicitly include racial data, using information impacted by societal biases created a system consistently favoring white patients over sicker black patients.

Simply diversifying datasets will not fix these issues. Addressing biases requires further understanding of society's inequalities and accounting for them in tech.

.....
**How will we
promote fairness?**



ALGORITHMIC BIAS

WHERE WE STAND

Can our tech unintentionally reinforce or amplify existing biases?



How can we create or update our tech to avoid harm and promote neutrality?

Do we have systems in place to limit algorithmic biases?



What else must we consider when building and refining algorithms?

FOLD



How will we promote fairness?

ANTICIPATING RISK

How could we monitor whether our product's algorithm benefits or hurts individuals, communities, and societies?

What might we do if we discovered our users have been profiled or discriminated against?

LEADING THE WAY

How can we encourage a diverse and global community of experts on AI to counteract biases?



ADDICTION

SINKING DEEPER

FOLD



ADDICTION

Digital notifications are irresistible by design.

According to a US survey, 72% of teens and 48% of parents feel the need to immediately respond to notifications like texts and messages on social networks. Within a business model where time on an app equals shareholder revenue, these dopamine-driven feedback loops are intentionally created to capture and sustain our attention.

But constant engagement doesn't feel good and can lead to tangible societal harms. For example, increased smartphone use correlates to higher symptoms of depression, loneliness, distorted social reality, and even suicide in teenagers.

.....
How will we promote healthier behaviors?



ADDICTION

WHERE WE STAND

Is our business model based
on engagement?



What might happen if our algorithms
and other features promoted
moderate use?

What drives high levels
of engagement in our tech?



Is this good for the mental, physical,
or social health of those who use it?

FOLD

How will we promote healthier behaviors?

ANTICIPATING RISK

What might extreme use
or unhealthy engagement look
like with our tech?

How might high engagement levels
negatively change people's habits
or our collective social norms?

LEADING THE WAY

What are different engagement metrics
we might prioritize for user wellbeing
online and offline?



DATA CONTROL

INSATIABLE APPETITE

FOLD



DATA CONTROL

The devices and apps that help simplify people's lives can also collect vast amounts of personal data, from web browser and purchase history to keystrokes and clicks.

Too often, consumers have little to no choice in how their valuable data is used—or even in claiming their rights over it—which allows some companies to buy, sell, or benefit from this data unfairly. Data usage agreements don't inspire much confidence in how user data is handled either.

In a system with little negotiation or informed consent, data for services is no longer an empowered or even voluntary exchange.

.....
**How will we
enable transparency?**



DATA CONTROL

WHERE WE STAND

What data are we collecting from users?



Do we really *need* to collect every piece of this data?

Do users have the option to choose what data they share with us or use our platform without being tracked?



What is our policy and workflow if a user wants their data to be removed from our systems?

FOLD



How will we enable transparency?

ANTICIPATING RISK

What could we do to make it harder for hackers to access or combine the user data we store?

Do we collect any data that puts us at greater risk of legal liability? If yes, how are we planning to address that?

LEADING THE WAY

Are there ways we could limit personal data collection or sharing that don't hurt our product's functionality?



FOLD



BAD ACTORS

Digital tools enable the spread of ideas at unprecedented speed and scale. While this generally benefits global communication and democratizes knowledge, it also empowers bad actors disseminating malicious content.

Bullying and trolling, as well as radicalization and exploitation, have also boomed. In Germany, one study showed that towns with heavier social media usage saw more anti-refugee attacks. And encrypted services, while vital for privacy and security, make it hard to identify those leading child and sex trafficking operations.

Bad actors using online platforms to hurt others creates an enduring struggle: balancing free expression and privacy with protecting the public from harm.

.....
**How will we
promote civility?**



BAD ACTORS

WHERE WE STAND

Can our feature or product be used to spam, harass, target, or exploit other people?



What safeguards do we have in place to minimize bad behavior?

Can our users block or mute bad actors—and request they be banned?



What other tools should we consider to help all of our users feel safe?

FOLD



How will we promote civility?

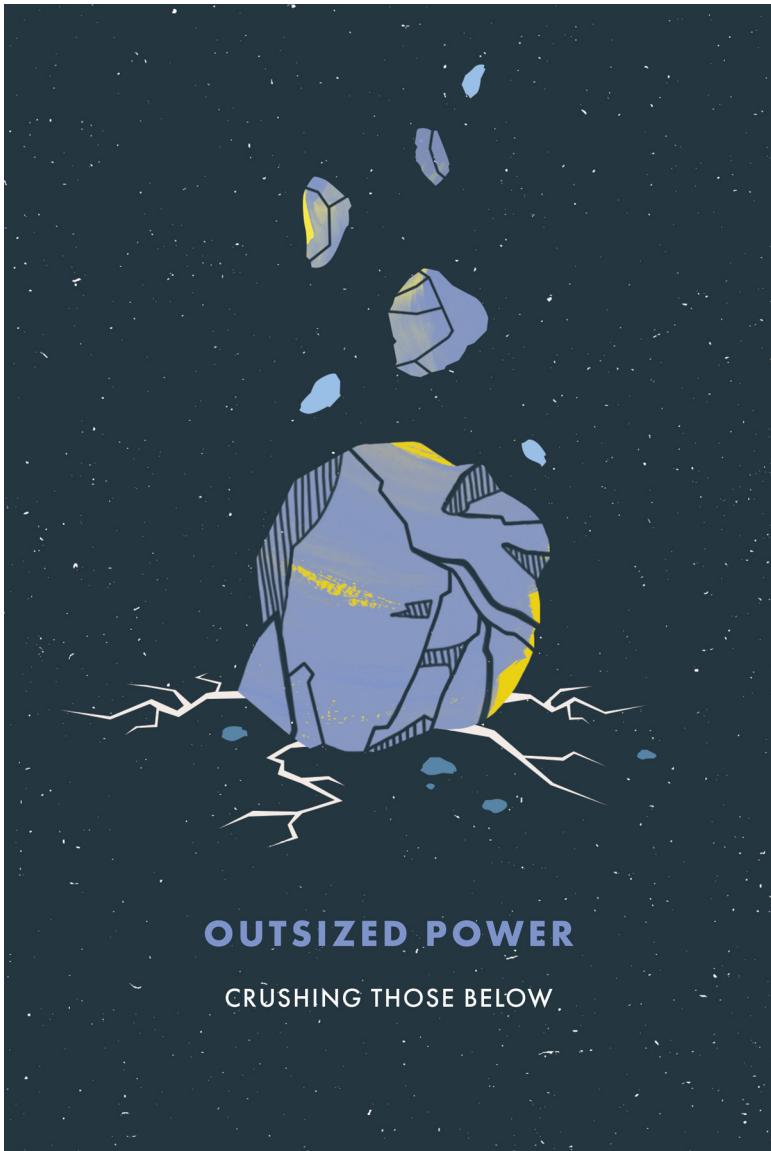
ANTICIPATING RISK

What crimes could arise in or around our tech?

Without violating privacy, how could we detect and address patterns of behavior that indicate our tool is being used to organize hate or harassment?

LEADING THE WAY

How can we prevent bad actors of all kinds from acting in the first place?



OUTSIZED POWER

CRUSHING THOSE BELOW.

FOLD



OUTSIZED POWER

Too much power in the hands of just a few creates problems. Tech has accelerated this issue by helping companies achieve market dominance at greater speed and scale.

It's natural for all companies—big *and* small—to try to grow their market share. But concentrated industries make it easy for companies to take shortcuts—or worse, abuse their power. Whether it's charging users higher prices after acquiring competitors or squeezing suppliers into unfavorable terms, anti-competitive behavior can have profound consequences.

The ability for consumers, suppliers, and employees to pursue alternatives if they're dissatisfied with their current experience is a key tenet of our economy.

.....
**How will we
promote choice?**



OUTSIZED POWER

WHERE WE STAND

Is there a dominant player in our industry? Is it us?



What impact does this have on our organization and the broader industry?

Do our growth targets result in us compromising our values or harming customers, suppliers, or employees?



Are there ways we can achieve growth that benefits us *and* our broader stakeholders?

FOLD



How will we promote choice?

ANTICIPATING RISK

How could we preserve our company's values—and ensure our tech isn't misused—if we're acquired?

Do we have a path to profitability that does not include an acquisition by a dominant platform?

LEADING THE WAY

Is it possible for us to ensure our product is interoperable with the broader market, not just tech giants?

CREATE YOUR OWN

RISK ZONE TITLE

DESCRIPTION

How will we
promote _____?

HUMAN VALUE

FOLD

RISK ZONE TITLE

QUESTIONS