a guide to

anticipating the future impact of today's technology

INSTITUTE FOR THE FUTURE

OMIDYAR NETWORK

Or: how not to regret the things you will build

If you can’t predict the future, you’re in the wrong business.*

*Just kidding.

No one can predict exactly what tomorrow will bring (although somewhere in the tech world, someone is no doubt working on it). So until we get that crystal ball app, the best we can hope to do is anticipate the long-term social impact and unexpected uses of the tech we create today.

It’s not necessary that we predict the future. Only that we get better at envisioning what’s possible in the future. The last thing you want is to get blindsided by a future YOU helped create. Seriously, that’s the worst.

So Ask Yourself:

- If the technology you're building right now will someday be used in unexpected ways, how can you hope to be prepared?
- What new categories of risk should you pay special attention to now?
- And which design, team, or business model choices can actively safeguard **users, communities, society,** and **your company** from future risk?

The Ethical Operating System

The Ethical Operating System can help makers of tech, product managers, engineers, and others get out in front of problems before they happen. It's been designed to facilitate better product development, faster deployment, and more impactful innovation. All while striving to minimize technical and reputational risks.

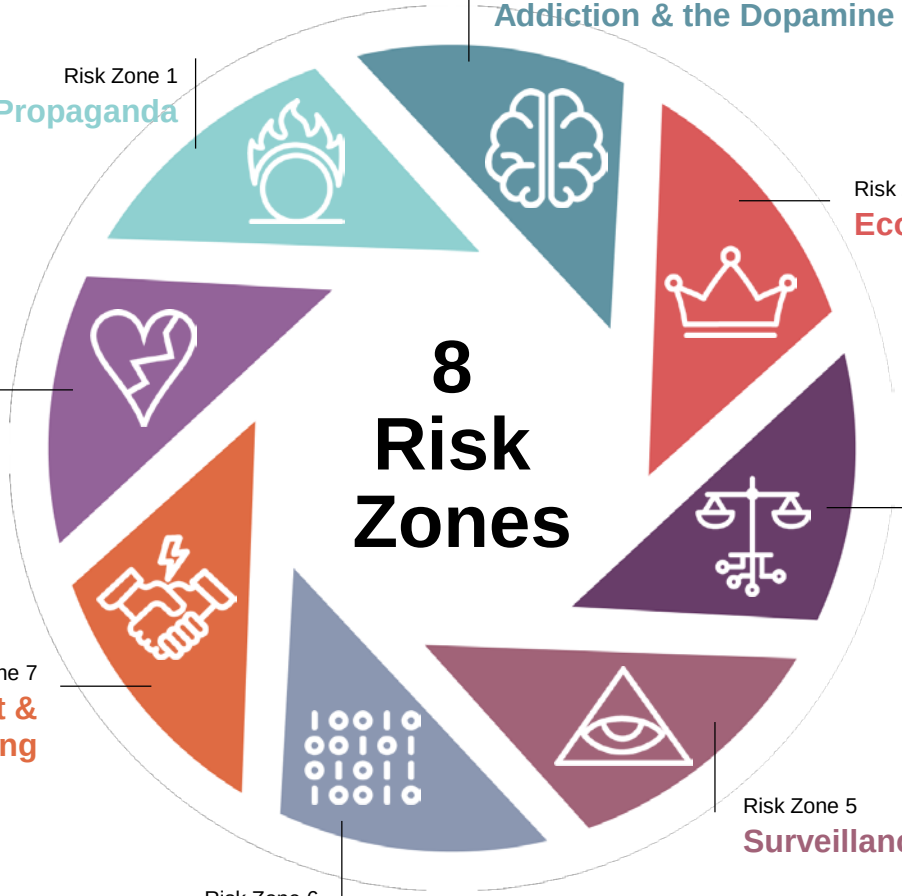
This toolkit can help inform your design process today and manage risks around existing technologies in the future.

Why Risk It?

It is only natural that as technologists, we spend most of our time focusing on how our tech will change the world for the better. Which is great. Everyone loves a sunny disposition.

But perhaps it's more useful, in some ways, to consider our glass half empty. What if, in addition to fantasizing about how our tech will save the world, we spent some time dreading all the ways it might, possibly, perhaps, just maybe, screw everything up?

The last thing you want is for your optimism to create blind spots around the possible downsides of your tech. The Ethical OS is here to help you see more clearly.



Risk Zone 1
Truth, Disinformation, Propaganda

Risk Zone 2
Addiction & the Dopamine Economy

Risk Zone 3
Economic & Asset Inequalities

Risk Zone 4
Machine Ethics & Algorithmic Biases

Risk Zone 5
Surveillance State

Risk Zone 6
Data Control & Monetization

Risk Zone 7
Implicit Trust & User Understanding

Risk Zone 8
Hateful & Criminal Actors

#SALMANQADIR

Table of Contents

HOW IT WORKS

Tools for visualizing and anticipating future risk, and building a better ecosystem

 ethicalOS.org

How It Works

Tool 1: Tomorrow's Risk—Today

If today is yesterday's future, then tomorrow's risks are your problem. Today.

Suffice to say that as your tech proceeds to grow and evolve, risks that now appear way out on the horizon (if at all) will come quickly into view. Often with little to no warning.

We must train ourselves to envision far-off risk.

Use this tool to begin stretching your imagination and warming up your foresight muscles and kicking off important conversations with your team. Sort of like yoga for your product ideation.

How It Works

Tool 2: Check Your Tech—The Risk Mitigation Manual

Different technologies naturally run different risks. Which should concern you?

Use this checklist to identify which emerging areas of risk and social harm are most critical for your team to start considering now.

How It Works

Tool 3: Future-Proofing Strategies

You've asked the hard questions. You've looked deep into the future. You've identified potential risks. Now what? It's time to take what you've learned and take action.

- How should you prioritize the risks you've identified? Which are the biggest threats? Which will be the hardest to address?
- What strategies can help mitigate these risks?
- Where and how should you begin?
- Who else needs to come on board before you can move forward?

Use these strategies to consider potential ethical systems meant to drive a more sustainable tech ecosystem.

Tool 1: Tomorrow's Risk—Today

14 Scenarios to Start a Conversation

Tool 1: Tomorrow's Risk—Today

14 Scenarios to Start a Conversation

How clearly and quickly can you spot potential risks on the horizon?

The Institute for the Future and Omidyar Network have scanned for areas of innovation where unintended consequences bearing major social impact may emerge. We've asked tech leaders which emergent risks and ethical dilemmas they're most worried about. And where possible, we've identified examples of positive signals where people are already working to mitigate these harms.

A **signal** is a specific example in the present of something that might influence or shape the future. It's a clue as to how things are already becoming different.

CLICK on this icon wherever you see it in the deck to reveal the signals.

Remember: Don’t get caught up in whether a scenario is likely or even possible. Just pick one and go with it.

> “When something of massive consequence happens that no one predicted, we often say it was simply unimaginable. But the truth is, nothing is impossible to imagine. When we say something was unimaginable, usually it means we failed to point our imagination in the right direction.”
>
> *—Jane McGonigal, Institute for the Future*

Tool 1: What to Do

These possibility spaces or scenarios are presented here as **14 Risky Futures**. They're meant to inspire a discussion about the growing responsibility—and opportunity—that tech companies have to manage social risks and anticipate longer-term consequences.

As you review each future scenario, CONSIDER:

- What is your greatest worry in this scenario?
- How might different users be affected differently by this future?
- What actions would you take to safeguard privacy, truth, democracy, mental health, civic discourse, equality of opportunity, economic stability, or public safety?
- What could we be doing now to get ready for this risky future?

Scenario 1: Are you ready for a world in which …

Video-faking algorithms are so advanced that faked videos are impossible to distinguish from real footage. These algorithms can completely replace people's faces and speech to make it appear they've said or done anything the author intends. With this technology, anyone can manufacture video "proof" to back up any claim. These video fakes flood major video-sharing, social media, and live-streaming platforms.

Scenario 2: Are you ready for a world in which …

Conversation bots have been trained to imitate specific people, using data sets collected from public social media posts. These bots are deployed across social media networks, email, and text messages in super-targeted, super-personalized propaganda campaigns. They are highly effective in changing opinions and driving action, as personalized messages appearing to be from real friends, family, and favorite celebrities are more influential than advertisements.

Scenario 3: Are you ready for a world in which …

In response to growing concerns over **tech addiction**, and anticipating government regulation as a possible future event, several of the most popular social media and game companies decide **voluntarily to enforce time limits**. Adults are commonly limited to two hours per day, and children to one hour per day per platform. Whether a platform is *limited* or *unlimited* becomes a major selling point. Many people prefer having hard limits that prevent them from becoming addicted or too frequently distracted. Other people switch to new unlimited competitors, creating a new social divide in who spends time where, and how their mental and physical health is affected by tech use.

Economic &
Asset Inequalities

Scenario 4: Are you ready for a world in which …

Automation has eliminated a significant number of jobs. By **2030, automation could eliminate 73 million American jobs,** distributed unevenly across racial groups, putting already marginalized communities even further at risk. Latino workers face the highest threat of automation at 60% of jobs eliminated, followed by African-Americans at 50%, Asians at almost 40% percent, and whites at roughly 25%.

Scenario 5: Are you ready for a world in which …

Fortune 500 human resource departments have subscribed to a "smart employer" service that evaluates a person's suitability for workplace cultures and stress levels, using public social media posts and "likes." Algorithms can identify individuals likely suffering from various mental illnesses, from depression to sociopathy. The service also predicts who may develop symptoms of mental illness in the near future, based on trends in the individuals' postings. This data is used to offer support and resources to current employees, recommend reassignment when necessary, and make hiring decisions.

Scenario 6: Are you ready for a world in which …

“Predictive justice” tools become the preferred method for determining prison sentences. These tools take data from millions of past cases and compare long-term historical recidivism rates for criminals with similar behavioral patterns or demographics to determine the most appropriate sentences. In other words: If you are similar in your economic, racial, gender, age or behavioral profile to people who have previously re-offended after incarceration, you get a longer prison sentence. These predictive justice tools do not take into consideration structural inequalities in how frequently members of different demographic groups are stopped, arrested, prosecuted, or convicted.

Scenario 7: Are you ready for a world in which …

A major social network company purchases a top U.S. bank and becomes the first social credit provider. It bases mortgage rates, loan approvals, and credit access on deep data collected by its social platform. It takes into consideration the credit histories of close friends and family, locations visited (including frequency of visits to places like bars and legal marijuana dispensaries), and "semantic analysis" of messages and photos to indicate whether individuals are generally happy, angry, anxious, or depressed.

Scenario 8: Are you ready for a world in which …

Facial recognition technology is a mainstream tool available to any individual or organization. Subscribers can tap into a database with hundreds of millions of faces indexed and clearly recognizable. Cameras are installed in most public and private spaces to take advantage of this technology, and a new ecosystem of apps emerges that integrates facial recognition across all product categories: dating apps, shopping apps, neighborhood apps, games, and more.

Scenario 9: Are you ready for a world in which …

“Read-and-write” **neurotech implants** are a reality. A device can be implanted in your brain that both “reads” your thoughts and feelings, and also “writes” new, artificial thoughts and memories directly into your head. The first successful read-and-write implant intercepts and records **80% of data sent to the brain’s hippocampus**, giving users close-to-perfect recall of anything they have directly read, seen, or experienced. This neuro-data is saved in the cloud from where it can be sent to other implants. The company creating and servicing these implants hopes to anticipate potential individual risks and widespread social impacts of its new product.

Scenario 10: Are you ready for a world in which …

A major data aggregator offers free health insurance to anyone who agrees to install a smart toilet in their home and submit its data to the company. **Smart toilets** can detect stress hormones, pregnancy, infectious diseases, alcohol and drug use, and blood sugar levels, among many other things. According to the contract that participants must sign, data collected from these smart toilets can be used for any purpose without limitation, including being sold to third parties, used for targeted marketing or shared with the government and scientific researchers.

Scenario 11: Are you ready for a world in which …

A major university recommends students install a blockchain-based "**proof of consent**" app on their phones before having sexual relations. Users state their preferences in a number of categories including condom use, STI status, and if photography is allowed. They must accept each other's preferences before having sex. The app creates an unalterable digital record of what has been consented to. Users are pseudonymous to prevent information about their sexual preferences and histories from becoming public. Some worry their data will be stolen and linked to their real-life identities. But many, especially women, use the app anyway. They place a higher priority on preventing rape and non-consensual acts than on privacy.

Scenario 12: Are you ready for a world in which …

Twenty-five percent of online orders are **delivered by drone.** Many of these drones are fitted with cameras and other sensors to collect data as they fly over neighborhoods, providing an additional revenue stream for shippers and merchants. Individuals who opt for free, unlimited drone delivery consent to the collection of data from their homes and yards. Entire neighborhoods where drone delivery is legally permitted are subject to the same data collection activities—even though not all of their individual residents or households have explicitly consented.

Scenario 13: Are you ready for a world in which …

A new **blockchain**-powered platform called Venge allows users to place anonymous bounties on actions designed to harass or terrorize specific individuals. These actions include doxxing, revenge porn, “swatting,” social media harassment, vandalism even violence. With one tap, you can pay someone else to act on your anger, frustration, or hate. Smart contracts pay out the bounties immediately in untraceable **cryptocurrencies** upon proof of completed action. Just as the Internet made it easier to anonymously harass other people, the anonymity and lack of regulation inherent in the case of cryptocurrencies makes it easier to pay people to carry out acts of hate.

Scenario 14: Are you ready for a world in which …

Self-driving vehicles become vulnerable to a new type of real-time ransomware. Hackers access the car remotely, turn off the engine, and refuse to start the car again until the driver pays a ransom. Typically this is a minor, commonplace inconvenience for small amounts of money. At worst, these attacks are timed and staged at dangerous locations—such as on train tracks, shortly before a train is scheduled to approach—for much larger ransoms.

Tool 2: Check Your Tech

The Risk Mitigation Manual

Tool 2: Check Your Tech

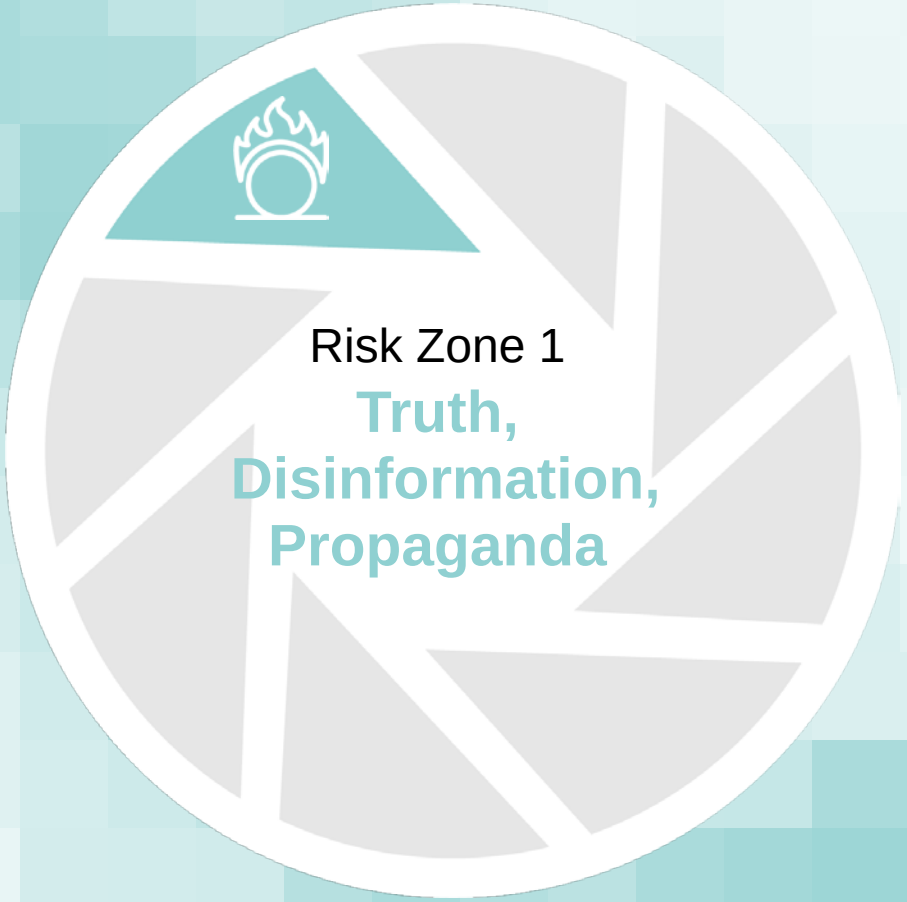
Most technologies are designed with the best intentions. But once a product is released and reaches scale, all bets are off. The Risk Mitigation Manual presents **8 risk zones** where we believe **hard-to-anticipate *and* unwelcome consequences** are most likely to emerge.

Remember: Potential risks evolve over time. As things change, we'll continue to update this section to reflect the most accurate current risk zones.

Tool 2: What to Do

- **CHOOSE a technology, product, or feature** you're working on, or something in the industry that's recently captured your attention.
- **READ through the sections on the 8 zones**
- **CHECK OUT THE SIGNALS of real examples** of these risks already being encountered and mitigated.
- **IDENTIFY the checklist questions** most relevant to the technology you've selected.
- **THINK about how you might** start to correct or mitigate the risk you've identified.

Download this manual



Risk Zone 1

Truth, Disinformation, Propaganda

Risk Zone 1

Truth, Disinformation, Propaganda

Shared facts are under attack. By everything, from fake news to bots that spread propaganda while posing as real people. Now a new wave of dangerous disinformation is on the rise. Known as “deep fakes,” these highly convincing videos algorithmically alter or replace people’s speech, facial expressions, and identities to create fake proof of actions or speech that never actually happened. Awesome.

Many individuals and groups are highly motivated to subvert the truth at a massive scale, especially for political ends. New social media technologies will make it even easier to spread lies and undermine trust.

Over the next decade, what else could be faked via new technologies? Which shared truths, facts, and information will we be called on to protect?

The Blockchain Solution to Our Deep Fake Problems

The Era of Fake Videos Begins

Risk Zone 1

Truth, Disinformation, Propaganda

- What type of data do users expect you to accurately share, measure, or collect?
- How could bad actors use your tech to subvert or attack the truth? What could become the equivalent of fake news, bots, or deep fakes, on your platform?
- How could someone use this technology to undermine trust in established social institutions, like media, medicine, democracy, science? Could your tech be used to generate or spread misinformation to create political distrust or social unrest?
- Imagine the form such misinformation might take on your platform.
  Even if your tech is meant to be apolitical in nature, how could it be co-opted to destabilize a government?

Risk Zone 2

Addiction & the Dopamine Economy

Risk Zone 2

Addiction & the Dopamine Economy

Research by Common Sense Media found that the average teenager spends 9 hours a day using some form of media. 9 hours!

The time we spend with our devices is of growing concern. Tristan Harris, founder of the Center for Humane Technology (CHT), has called for tech companies to encourage “time well spent,” suggesting designers optimize the time they spend on platforms, in a way that makes their time beneficial to their overall happiness and well-being.

Studies show people achieve maximal intended use of apps like Instagram and Snapchat after 11 minutes—beyond which overall happiness decreases. How might tools be designed to advocate for time well spent? How can we design software that prioritizes user happiness—offline and online—over keeping eyes glued to the screen?

China's Tencent to limit play time of top-grossing game for children

These are the sneaky ways apps like Instagram, Facebook, Tinder lure you in and get you “addicted”

Risk Zone 2

Addiction & the Dopamine Economy

- Does the business model behind your chosen technology benefit from maximizing user attention and engagement—i.e., the more, the better? If so, is that good for the mental, physical, or social health of the people who use it? What might not be good about it?
- What does “extreme” use of, addiction to, or unhealthy engagement with your tech look like? What does “moderate” use to or healthy engagement look like?
- How could you design a system that encourages moderate use? Can you imagine a business model where promoting moderate use is more sustainable or profitable than always seeking to increase or maximize engagement?
- If there is potential for toxic materials like conspiracy theories and propaganda to drive high levels of engagement, what steps are being taken to reduce the prevalence of that content? Is it enough?

Risk Zone 3

Economic & Asset Inequalities

Risk Zone 3

Economic & Asset Inequalities

In 2017, according to Oxfam International, eight people owned as much wealth as the entire bottom half of the world's population. Let that sink in a minute.

Wealth concentration and distribution have been issues issue throughout modern history. But it's gotten way worse. Today, in the United States, wealth concentration is the worst it's been since 1928. New technology can democratize access, provide income opportunity, and balance distribution, but it can also exacerbate inequality by catering only to high-income groups and eliminating low-income jobs.

Widespread Mobile Phones Can Provide Banking to the Poor: World Bank

How to Get Cheaper Car Insurance: Be White

Two Ex-Googlers Want to Make Bodegas Obsolete

The New Casualties of Automation

Risk Zone 3

Economic & Asset Inequalities

- Who will have access to this technology and who won't? Will people or communities who don't have access to this technology suffer a setback compared to those who do? What does that setback look like? What new differences will there be between the "haves" and "have-nots" of this technology?
- What asset does your technology create, collect, or disseminate? (example: health data, gigs, a virtual currency, deep AI) Who has access to this asset? Who has the ability to monetize it? Is the asset (or profits from it) fairly shared or distributed with other parties who help create or collect it?
- Are you using machine learning and robots to create wealth, rather than human labor? If you are reducing human employment, how might that impact overall economic well-being and social stability? Are there other ways your company or product can contribute to our collective economic security, if not through employment of people?

Risk Zone 4

Machine Ethics & Algorithmic Biases

Risk Zone 4

Machine Ethics & Algorithmic Biases

As the application of AI in critical domains like welfare, education, employment, and criminal justice has intensified, so, too, have concerns around the way AI can reflect, introduce, or even amplify bias.

Whether it's racially biased police facial recognition systems or search engines that privilege certain ideas and identities, there is enough evidence to know that technology itself is not neutral. That is because technology is often the product of human action or a tool used by humans—humans who by nature are not neutral.

Solutions that rely on AI to remedy challenges like those above often overlook the fact that AI itself is a tool created and used by humans. An interdisciplinary approach to the design, use, and governance of AI technology remains critical to addressing some of the inadvertent and/or intentional discriminatory impacts on certain individuals and communities.

Capital One Pursues 'Explainable AI' to Guard Against Bias in Models

Silicon Valley Is Stumped

Podcast: Artifical Intellgence With Very Real Biases

Algorithms Aren't Biased, But the People Who Write Them May Be

Risk Zone 4
Machine Ethics & Algorithmic Biases

- Does this technology make use of deep data sets and machine learning? If so, are there gaps or historical biases in the data that might bias the technology?
- Have you seen instances of personal or individual bias enter into your product's algorithms? How could these have been prevented or mitigated?
- Is the technology reinforcing or amplifying existing bias?
- Who is responsible for developing the algorithm? Is there a lack of diversity in the people responsible for the design of the technology?
- How will you push back against a blind preference for automation (the assumption that AI-based systems and decisions are correct and don't need to be verified or audited)?
- Are your algorithms transparent to the people impacted by them? Is there any recourse for people who feel they have been incorrectly or unfairly assessed?

Risk Zone 5

Surveillance State

Risk Zone 5

Surveillance State

Recent examples of social bots being co-opted by governments and militaries for use in attacking the opposition reveal that technology built for benign uses can be leveraged for harm. In some cases, armies of automated software-driven Twitter and Facebook profiles were used to target journalists, activists, and citizens who spoke out against particular ideas.

In authoritarian countries, Western surveillance tools have powered state-sponsored repression, allowing leaders to intercept emails and text messages and monitor the whereabouts of citizens through their phones. Armed with this information, police in many of these countries now routinely confront dissidents with records of their messages and movements during arrest and torture.

Meanwhile, increasing use of facial recognition and social media tracking is allowing governments to create deep records of individual behavior. “Citizen scores” generated over years of data collection can be used to limit access to public spaces, work opportunities, and more, creating a new kind of social inequality based on prior behavior and speech.

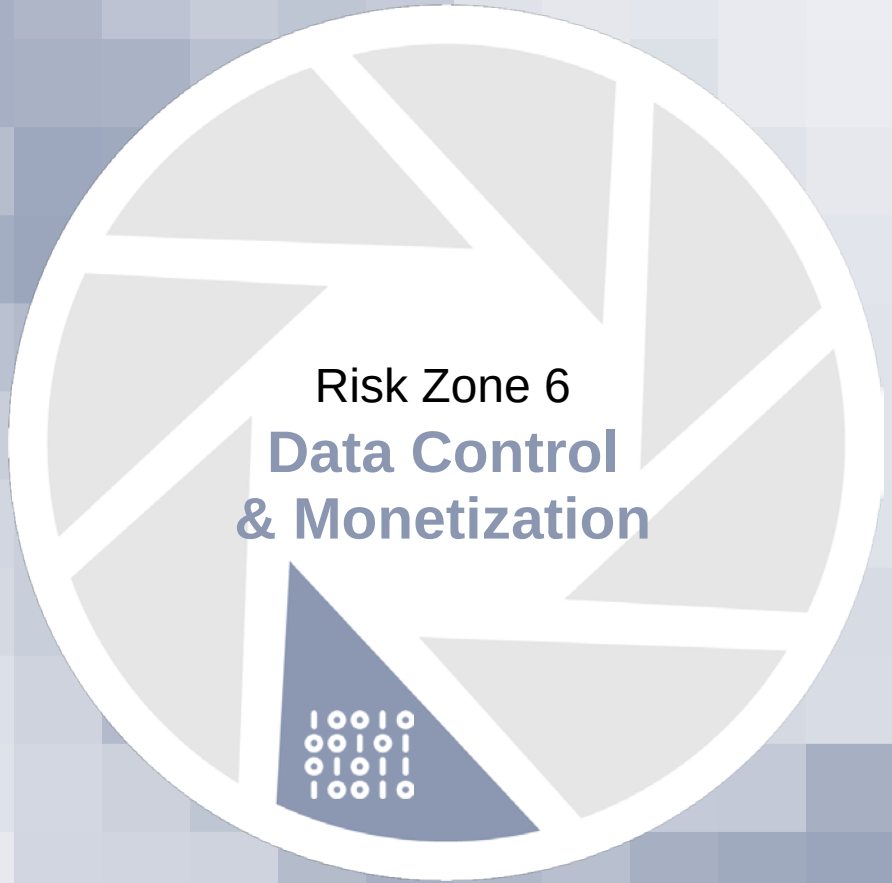
UK Unveils Machine Learning Technology to Fight ISIS Propaganda Online

China’s Surveillance State Should Scare Everyone

Risk Zone 5

Surveillance State

- How might a government or military body utilize this technology to increase its capacity to surveil or otherwise infringe upon the rights of its citizens?
- What could governments do with this data if they were granted access to it, or if they legally required or subpoenaed access to it?
- Who, besides government or military, might use the tools and data you're creating to increase surveillance of targeted individuals? Whom would they track, why—and do you want your tech to be used in this way?
- Are you creating data that could follow users throughout their lifetimes, affect their reputations, and impact their future opportunities? Will the data your tech is generating have long-term consequences for the freedoms and reputation of individuals?
- Whom would you *not* want to use your data to surveil and make decisions about individuals, and why not? What can you do to proactively protect this data from being accessible to them?



Risk Zone 6

Data Control & Monetization

Risk Zone 6

Data Control & Monetization

In the future, users will expect access to tools for acquiring, sharing, interpreting, and verifying information that has been collected about themselves and their environment. They will assert a fundamental right to selectively protect the privacy of their personal and institutional data where disclosure would jeopardize their well-being. And they will expect to control access to their personal data and have opportunities to selectively share, monetize, and benefit from it alongside the tech companies that create and use the data.

Google Takeout (Google's tool for downloading your personal data)

The Social Network Doling Out Millions in Ephemeral Money

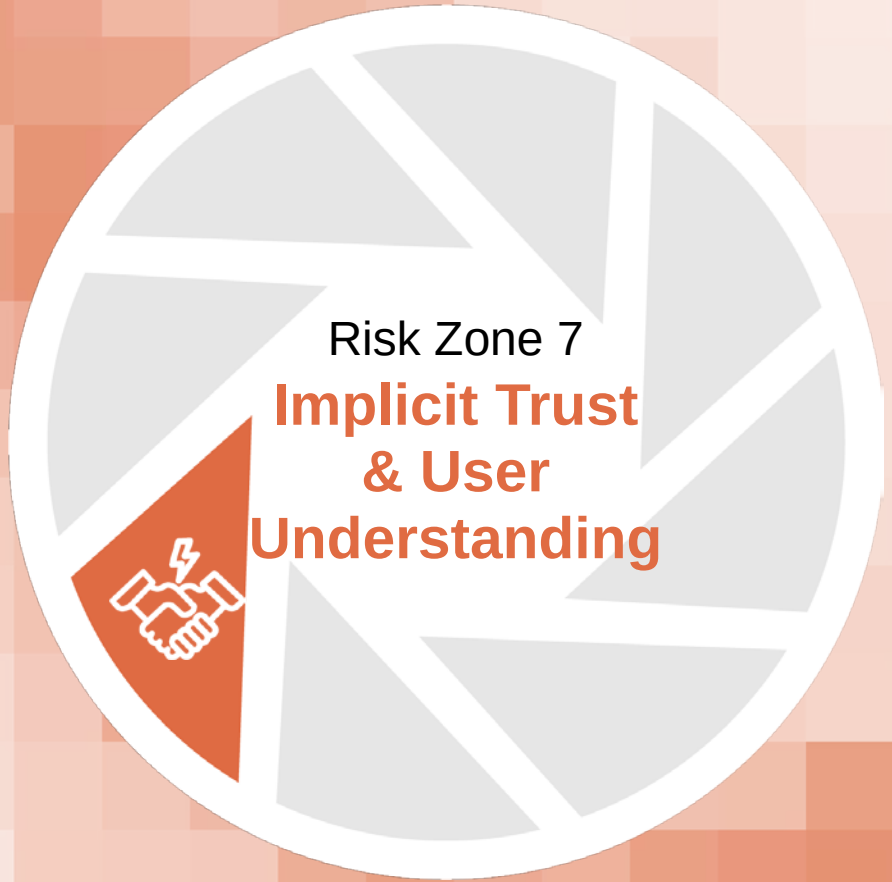
Cambridge Analytica and Facebook: The Scandal and the Fallout So Far

Google's File on Me was Huge. Here's Why it Wasnt as Creepy as My Facebook Data.`

Risk Zone 6

Data Control & Monetization

- What data are you collecting from users? Do you really need to collect it? Are you selling it? If so, whom are you selling it to and do users know this? How can you be more transparent about this?
- Do your users have the right and ability to access the data you have collected about them? If not, how can you better support users in easily and transparently knowing about themselves what you already know about them?
- If you profit from the use or sale of user data, do your users share in that profit? What options would you consider for giving users the right to share profits on their own data?
- Could you build ways to give users the right to share and monetize their own independently?
- What could bad actors do with this data if they had access to it? What is the worst thing someone could do with this data if it were stolen or leaked?
- Do you have a policy in place for what happens to customer data if your company is bought, sold, or shut down?



Risk Zone 7

Implicit Trust & User Understanding

Risk Zone 7

Implicit Trust & User Understanding

Misuse of data is a serious problem. But perhaps equally as troubling is the fact that many users are not sufficiently aware of how exactly popular apps and platforms work, how their engagement is optimized, or what's being tracked and collected. Terms of service are hardly read and are rarely written for clarity. Permissions to many activities are often given without full understanding of how they will be used.

In the future, companies are likely to face backlash when their product or employees violate the implicit or explicit assumptions and scope of what users have agreed to, especially where that agreement is based on users accepting a dense, indecipherable agreement no one actually reads.

Tim Cook says Apple's refusal to unlock iPhone for FBI is a 'civil liberties' issue

Snapchat's Evan Spiegel Suggests Facebook Copy Its Data Protection Policy

Uber allegedly tracked journalist with internal tool called 'God View'

Some apps were listening to you through the smartphone's mic to track your TV viewing, says report

Risk Zone 7

Implicit Trust & User Understanding

- Does the technology you're building have a clear code of rights for users? Are the terms of service easy to read, access, and understand?
- Is there a version of your product that is available to use if users don't want to sign the user agreements?
- Does your technology do anything your users don't know about, or would probably be surprised to find out about? If so, why are you not sharing this information explicitly—and what kind of backlash might you face if users found out?
- If users object to the idea of their actions being monetized, or their data being sold to specific types of groups or organizations, though still want to use the platform, what options do they have? Is it possible to create alternative models that builds trust and allows users to opt-in or opt-out of different aspects of your business model moving forward?
- Are all users treated equally? If not—and your algorithms and predictive technologies prioritize certain information or set prices or access differently for different users—how would you handle consumer demands or government regulations that require all users be treated equally, or at least *transparently* unequally?

Risk Zone 8

Hateful & Criminal Actors

Risk Zone 8

Hateful & Criminal Actors

On August 12, 2017, in Charlottesville, VA, Heather Heyer was killed when a man linked to a white supremacist group, crashed his car into a crowd of counterprotestors at a rally called “Unite the Right.”

The rally was organized using the Facebook events tool.

Obviously, hate crimes are nothing new. But online tools enable the global dissemination of content with an ease and speed never before available to those wishing to spread hate. Bullying, radicalization, trolling, doxing and other bad behaviors have boomed, thanks to tech. And it’s practically impossible to mitigate or prevent negative experiences for users who spend a great deal of time navigating in that environment.

Researchers Find that Twitter Bots Can Be Used for Good

Airbnb Cancels Accounts Linked to White Nationalist Rally

Cut Off from Big Fintech, White Nationalists Are Using Bitcoin to Raise Funds

How ISIS Became the World’s Deadliest Tech Startup

Risk Zone 8

Hateful & Criminal Actors

- How could someone use your technology to bully, stalk, or harass other people?
- What new kinds of ransomware, theft, financial crimes, fraud, or other illegal activity could potentially arise in or around your tech?
- Do technology makers have an ethical responsibility to make it harder for bad actors to act?
- How could organized hate groups use your technology to spread hate, recruit, or discriminate against others? What does organized hate look like on your platform or community or users?
- What are the risks of your technology being weaponized? What responsibility do you have to prevent this? How do you work to create regulations or international treaties to prevent the weaponizing of technology?

Let’s Get Ethical

You understand the risks, now it’s time to act.

OK, you’ve contemplated scenarios and familiarized yourself with major areas of potential risk. Now let’s take the next step: beginning to design and implement a more ethical operating system. Here are a few ways to get started:

- **SHARE the risk zones and questions you’ve highlighted across your teams.** Encourage individuals to keep these front-of-mind when developing new products or considering new features.
- **CONSIDER adding your top questions** to the Product Requirement Document (PRD), if you’re actively working on a new technology to help your team sort through the potential risks of what you're building.
- **SCAN the horizon for additional information about these risk zones**. Invite team members to **flag and circulate news stories, op-eds, or other “signals” that relate to these risks** to keep ethical issues embedded in your operations.

Tool 3: Future-Proofing

Embedding learnings and best practices from now onward

Tool 3: Future-Proofing

On Building an Ethical Infrastructure

Ready to make the world a better place (or at least help stop others from making it worse)? Then it's time to turn our best intentions into actionable safeguards.

What best practices could help the tech community mitigate risk **at scale?** What industry-wide effort would help us create products that have both a company's and humanity's best interests in mind?

The following are **strategies** and ideas to inspire action and add infrastructure for you and your team.

Tool 3: What to Do

- **PICK** a future-proofing strategy.
- **CONSIDER** its merits. What could it accomplish?
- **DISCOVER** potential drawbacks. How could it go wrong? What would prevent it from becoming a reality?

If one strategy jumps out as something your team or company would like to pursue, **make a list of potential resources, allies, and next steps** that could help you put it into practice.

Strategy 1

Tech Ethics 101

In the future, imagine that top computer science, design, and engineering programs adopt a requirement that all students complete a course or training sequence in tech ethics.

- What skills should it teach?
- How would students be graded on their ethical skill set?
- Which schools might adopt this change first?
- Who would create the curriculum or textbook?
- What can you add into your daily standup to reflect these principles?

Strategy 2

A Hippocratic Oath for Data Workers

Imagine if anyone working with user data took an oath to obtain, use, and share data ethically.

- What specific commitments should the oath include?
- Could major companies or investors require individuals to take the oath as a condition of employment or funding?
- What rituals would make “taking the oath” more meaningful?
- What other specializations might create their own oath? For example, what would a Hippocratic oath for people who work in machine learning look like?
- **Have your team take the pledge! https://github.com/Data4Democracy/ethics-resources**

Strategy 3

Ethical Bounty Hunters

What if tech companies began paying “bounties” for identifying the major potential social risks of new services and technologies? Similar to the way hackers are rewarded for identifying security flaws and vulnerabilities today, bounties could be set for different areas like “mental health impact,” “risks to democracy,” and “structural racism, sexism, or other inequality.”

- How would someone claim a bounty? What kind of evidence would be required?
- Who within a company would be responsible for evaluating the risks identified?
- What would be a reasonable payout?
- **Send us your answers**
  ideas@ethicalOS.org

Red Flag Rules

In the future, tech companies and investors will provide employees with a list of social and ethical “red flags” to watch for. All employees will have clear pathways to report risks internally without fear of retribution.

- What risks should go on a red flag list?
- Who would create and update it?
- Would the list be shared across companies and portfolios?
  Would it be public or proprietary?
- **What would an internal red flag reporting process look like? How would reports be handled on your own team?**

Healthy Platforms

Creators of new technology platforms will be expected to establish and share a transparent measure of the "health" of their platform. Twitter has already committed to measuring and maximizing four new measures of "conversational health". Such health metrics will help evaluate whether the platform is doing good for its users and society.

- Which measures of health would you want to ensure?
- What would you define as important components of this health?
- How would you measure these components?
- How would you communicate health or health risks to users and/or the public?
- **Whose job is it on your team to monitor these metrics for health?**

Strategy 6

License to Design

Imagine if all technology designers and developers were licensed by a governing body that regulates their actions, prohibits irresponsible or unethical design, and counteracts the prevailing “move fast, break things” culture of today’s tech world. This would be similar to licensing practices for doctors, lawyers, and architects.

- Who would make up the governing body?
- What would requirements to obtain a license include?
- What actions would lead to loss of license?
- Would designers and developers benefit from a move to licensing? Would companies?
- How would you mitigate the potential negative repercussions of licensing, i.e., its leading to greater inequality and less diversity in terms of who gets jobs, narrowing of the job pool, or slowing innovation?
- **Send us your thoughts at ideas@ethicalOS.org**

Tool 3: Future-Proofing Strategies

“We have to think about the long-term futures that our technologies are helping to create, not just the growth and scale we want to achieve today. If we fail to do this, we fail everyone who uses these technologies. And in the worst-case scenarios, we fail democracy and equality at scales never possible before.”

—Sam Woolley, Director
Digital Intelligence Lab, Institute for the Future

What else would you design?

Now that you've considered these ethical infrastructure interventions, what other efforts would you design to build long-term resilience and help the tech community create products that have both a company's and humanity's best interests in mind?

Share your thoughts:

#ethicalOS

ideas@ethicalOS.org

We'll keep updating the Ethical Operating System with thoughts from the community.

CUSTOM USE CASES

In the following section you'll find the Cliff Notes versions for your specific teams, board members, or students to engage them in their first conversation on the topic (which surely won't be their last).

Ethical OS Cliff Notes

*(AKA: **"If you're only going to read one page from this toolkit, here it is."***
Please share with any peers (even those who may not think they need it.)

THINGS TO CONSIDER

- How could your product, service or business most **positively impact** the world you want to see? Beyond the problem you're trying to solve, how will the world be better for having this innovation introduced?
- Could your product or product feature harm someone directly? Consider not only the first-order consequences, but try to imagine the second- and third-order as well.
- Could a malicious actor use your product or product feature to harm a person or group? Commit a crime?
- Could the business model behind your product harm someone or a group of people? Could it fuel addiction? Could it impact certain groups—children, seniors, minority groups, the disenfranchised or disempowered—in different ways.
- Could your product risk someone's privacy? What happens if someone's data gets leaked, hacked, or lost? Do you have a plan in place for these contingencies? How can you best build trust with your users?
- Could your product or business do anything your users are unaware of? If so, why are you not sharing this information explicitly? Would it pose a risk to your business if it showed up in tomorrow's news?
- What can or should you do as a board to ensure your company lives up to the best version of itself?

For more details and the full Ethical OS toolkit: ethicalOS.org

Are you a trustee or board member?

With threats ranging from reputational risk to employee outflow, ethical lapses can negatively impact shareholder value in any number of ways—and this goes beyond office misconduct. The following conversation starters are intended to help you get out in front of potential product risks before any damage is done. Take 5 minutes out of your meeting to make sure you've thought through some of these issues.

- Can your product or business do anything users are **unaware** of? Why not share this information explicitly? Would it **pose a risk** to your business if it showed up in tomorrow's news?
- How transparent and clear are your terms of service? What could you change to make them easier to understand?
- Could a malicious actor use your product to do harm or commit a crime? What mitigants do you have in place if this were to occur?
- Could your business model itself be used maliciously? Could it fuel addiction? Could it impact certain groups — children, seniors, minority groups, the disenfranchised or disempowered — in different ways.
- Could your product compromise a user's privacy? What happens if someone's data gets leaked, hacked, or lost? Do you have a plan in place for these contingencies? How can you best build trust with your users?
- What can or should you do as a board to ensure your company lives up to the best version of itself?

For more details and the full Ethical OS toolkit go to: ethicalOS.org

Are you making products?

THINGS TO KEEP IN MIND AS YOU GO

- Have you assessed not only the first-order consequences of your product, service or platform, but also its second- and third-order impacts?
- Have you designed your product or service in such a way so as to ensure that a diverse set of users will be able to access?
- Have you ensured you're collecting only as much user data as is absolutely necessary to operate your product and ensure your financial viability?
- Have you made certain that your training data or data sets represent a diverse set of users and that you've minimized any potential bias in the source of that data?
- Have you identified the most effective ways to ensure your product spreads only truthful content and information?
- Have you ensured the fairness, accountability, and transparency of any algorithms or machine learning processes embedded in your product or service?
- Have you designed appropriate off-ramps for your users? Have you eliminated, to the extent possible, bottomless pits, infinite scrolls, and attention traps?
- Have you found ways to avoid the use of any dark patterns in your UX?
- If you have them, are your notifications important enough to interrupt the user? Can they be delivered in a less intrusive manner—using, for example, a sound, a sensation, or subtle status change?
- Have you created systems that place a value on quality of user-generated content over quantity?
- Have you created contingency plans for data breach? Have you found ways to ensure that users will trust you before, during, and after such a breach?

Tech product and design guide (cont.)

TO ASK BEFORE YOU SHIP

- Have you written your terms of service clearly, succinctly, and in easy-to-understand language?
- Have you tested your product with a diverse set of users, representing diversity of age, gender, race, socioeconomic status and income, geography, political affiliation, language, ability, sexual orientation, religion, and education?
- Have you re-examined any potentially discarded paths to revenue or growth that you should reconsider in light of any of the unintended consequences you can foresee?
- Have you red-teamed your product to assess how a bad or malevolent actor (individual, group, or body) might weaponize your product?
- Have you gut-checked yourself on any implicit assumptions about any or all of the above?
- Given all of the above, have you designed a product you feel confident will make the world a better place?

For more details and the full Ethical OS toolkit go to: ethicalOS.org

Are you teaching a class in computer science or design?

TOOL 1: **Have students read all 14 "Risky Futures" scenarios and pick one that captures their imagination.** Ask them to brainstorm what ethical risks or harms might occur in that future. Students can work in teams with others interested in the same scenario. This can also be done as a class discussion.

TOOL 2:

- **Ask students to read about all 8 Risk Zones and pick one that interests them. Challenge them to collect signals from that risk zone.** Signals are real examples of things already happening in the present that might influence or shape the future. Signals can be found in the news, blogs, social media, scientific journals, tech conferences, TED talks, labs, and anywhere people or companies are sharing new products, ideas, or findings. Students can work in teams with others interested in the same risk zone.
- **Challenge students to pick a real, emerging technology, product, or app that they are interested in.** Have them go through the risk mitigation questions and check off the questions in the risk zones they think are most relevant to the tech. Then, choose one of those questions and attempt to answer it, looking for ways to make that tech more ethical and less risky.

TOOL 3: Hold a class discussion about one or more of the six strategies. Have students vote on their favorite strategy and then spend time imagining in depth how the class's top one or two strategies might play out in the future.

For more details and the full Ethical OS toolkit go to: ethicalOS.org

Start a conversation with your network

The future of tech is all-hands-on-deck, and we need all points of view.
Here are some ideas for using the EthicalOS to engage your networks on Facebook, LinkedIn, Twitter, and anywhere else you talk about what matters.

TOOL 1: Post one of the 14 "Risky Futures" scenarios on your social media network. Include a link to the scenario's signal, so people can investigate the real-world example that inspired it. Ask your network: What would you worry about happening in this scenario? And what would you suggest doing about it?

TOOL 2: Pick any of the 8 Risk Zones. Share its description and signals with your network. Ask them: Is this risk something you worry about in the tech you work on? What about the technologies you use? Have you already been affected by any of these risks? What other examples have you seen of this risk in the real world?

TOOL 3: Share any two of the future-proofing strategies. Ask your network which of the two they think would be more effective and why.

- **SHARE FREELY:** Feel free to take screen shots of slides in the deck or quote any of the material, to share with your network
- **Use the hashtag** #EthicalOS
- **Link back** to EthicalOS.org so your network can download the full Ethical OS toolkit

For more details and the full Ethical OS toolkit go to: ethicalOS.org

About the Ethical Operating System

The Ethical OS is a joint creation of the Institute for the Future and Omidyar Network's Tech and Society Solutions Lab.

- If you would like the Institute for the Future to help you organize an event or workshop using this toolkit, or to facilitate a training session within your organization, contact **ideas@ethicalos.org**
- If you have feedback on this toolkit or suggestions for additional risk zones or scenarios, share your foresight with us at **ideas@ethicalos.org**

About the Institute for the Future

Institute for the Future (IFTF) is celebrating its 50th anniversary as the world's leading non-profit strategic futures organization. The core of our work is identifying emerging discontinuities that will transform global society and the global marketplace. We provide organizations with insights into business strategy, design process, innovation, and social dilemmas. Our research spans a broad territory of deeply transformative trends, from health and health care to technology, the workplace, and human identity. IFTF is based in Palo Alto, California. For more, visit www.iftf.org.

Research lead for the Ethical OS: Jane McGonigal, IFTF Director of Research & Collaborative Foresight.

With special thanks to David Evan Harris, Director, IFTF Governance Futures Lab; Jean Hagan, IFTF Executive Producer; and Sam Woolley, Director, IFTF Digital Intelligence Lab.