A Guiding Framework for Vetting Technology Vendors Operating in the Public Sector

Ford
Foundation

What is this guiding framework and who is it for?

This guiding framework aids the analysis of ways that new digital technology-based proposals may impact the U.S. public sector, with a particular focus on their impacts on human rights, social and economic justice, and democratic values.

The framework can be used in conjunction with your current due diligence process. Whether you follow the Theory of Change methodology in your vetting process or conduct detailed impact assessments (e.g., UN Guiding Principles on Business and Human Rights methodology for conducting human rights impact assessment,[1] algorithmic impact assessment,[2] privacy impact assessment [3]), guidance to address the following red flags can be readily integrated into those processes.

PRIMARY AUDIENCE:

Program officers at philanthropic organizations who often receive proposals that are framed as “tech for good,” “justice tech,” or public interest technologies.

OTHER GROUPS:

Employees of public agencies whose job is to procure digital services/products and assess technology vendors

Third-party auditors and advocates who assess societal impacts of digital technologies that are procured and/or deployed by public agencies

Vendors and social entrepreneurs who want to build digital services for public agencies. It helps them to assess the societal harms and unintended consequences of their own services.

[1] "Human rights impact assessment guidance and toolbox," the Danish Institute for Human Rights, https://www.humanrights.dk/tools/human-rights-impact-assessment-guidance-toolbox

[2] "Algorithmic Impact Assessments: A Practical Framework For Public Agency Accountability," AI Now Institute, https://ainowinstitute.org/aiareport2018.pdf

[3] "Privacy Impact Assessment," U.S. Department of Health and Human Services, https://www.hhs.gov/pia/index.html

This framework consists of a list of red flags divided into seven categories:

1. Theory of Change and value proposition
2. Business model and funding
3. Organizational governance, policies, and practices
4. Product design, development, and maintenance
5. Third-party relationships, infrastructure, and supply chain
6. Government relationships
7. Community engagement

EACH RED FLAG IS ACCOMPANIED BY:

Several questions
to help identify that red flag in a proposal

Hypothetical examples or scenarios
showing how, in a particular context,
that red flag may lead to societal harm

Resources to learn more about a given red flag

Why do you need this guiding framework?

In the philanthropic space, numerous technology-related proposals compete for attention. Their approaches and themes vary widely, but one thing that many of them have in common is the assumption that they will use technology not just to reduce inefficiency, but to increase equity and justice.

Evaluating the potential costs and benefits of projects that make these claims is a complex task. This framework seeks to create guiderails that will aid funders, procurement officers, advocates, and vendors in evaluating and assessing the justice, equity, and human rights implications of proposed projects.

"Tech for good" has, at least for the past decade, been a prevalent theme in the public sector, especially in connection with the criminal legal system, public welfare systems, public health, and migration management. It has been argued that these projects too often replicate the status quo, offering facile or impractical solutions to deeply rooted systemic social problems, while in some cases replicating the same inequalities or injustices they seek to alleviate.

For example, automated risk assessment tools used in courtrooms claim to measure the likelihood of recidivism based on an accused's interactions with law enforcement. These tools, which are often intended to make courtrooms' decisions faster and fairer, have been shown instead to be biased against Black and Brown people. Systemic racism is not only baked into the design of these systems but is also perpetuated through their use.[4]

Meanwhile, the enormous abundance of data in the digital era has made funders, non-profits, and social entrepreneurs excited about finding new methods for putting data to work as a tool for social justice. However, the mentality of "the more data the merrier" runs a considerable risk of introducing opportunities for harmful surveillance and monitoring into projects which utilize data relating to historically oppressed and excluded groups.

These risks are not just notional: They have a history dating back at least to the late 1960s and early 1970s, when the Welfare Rights Movement led to the creation of tools for preventing discrimination related to eligibility criteria for welfare services. Since that time, numerous attempts at streamlining, automating, or rationalizing the provision of government services have been framed as projects that will reduce inequality or injustice. Unintended consequences, however, have been the norm.[5] The nature of these public-private partnerships has often led to a diminishment of accountability mechanisms, public oversight, and remediation of harms to affected communities.

Much of this work stems from the ethos, associated with contemporary Silicon Valley, that applying engineering know-how can fix intractable social problems in ways that politicians and policymakers never foresaw — a beguiling but oft-disproven idea that dates back to the Technocracy movement of the 1930s.[6]

Utilizing a "follow the money" approach, advocates have been calling on philanthropies to fund more responsibly by considering unintended and longer-term consequences of digital technology-enabled and data-driven proposals.[7] That is where the following guiding framework comes in.

[4] Julia Angwin et al, "Machine Bias," ProPublica, May 2016, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

[5] Eubanks, Virginia. Automating inequality: How high-tech tools profile, police, and punish the poor. St. Martin's Press, 2018. See also the Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (P.L. 104-193).

[6] Wythoff, Grant. "Silicon Valley's attempts to self-police are anti-democratic. they're also not new." The Washington Post, August 17, 2018. https://www.washingtonpost.com/outlook/silicon-valleys-attempts-to-self-police-are-anti-democratic-theyre-also-not-new/2018/08/17/cd44fb22-9b1d-11e8-843b-36e177f3081c_story.html

[7] Haven, Janet, and Danah. Boyd. "Philanthropy's techno-solutionism problem." Democracy and civic life: What is the long game for philanthropy (2020). https://www.kettering.org/sites/default/files/boyd-philanthropys_techno-solution-problem.pdf

When and how can you use this guiding framework?

Initial assessment:

A quick cheat sheet to decide how to proceed with a "tech for good" proposal that you may receive through referrals, emails, or conferences, or as a response to a Request for Information (RFI).

Organizational capacity building:

As cross-department training material and/or during your orientation for onboarding new program officers.

Assessment for vetting a new add-on "tech for good" program proposed by your existing grantees:

Sometimes your current grantee may come up with a "tech for good" project within their existing area of work. Hypothetically speaking, examples can include a humanitarian NGO proposing a mobile app to communicate available resources to their clients in a more streamlined and centralized way; a reproductive justice organization seeking to develop a secure communication channel to collaborate with trusted clinics in sanctuary states; an academic institution wanting to deploy an algorithm to help improve outcomes for refugees post resettlement).

RED FLAGS CATEGORY

Theory of Change and Value Proposition

Ford Foundation

1

RED FLAG

The project does not offer a clear Theory of Change — or, if offered, it is tenuous, misguided, or oversimplified.

AT A GLANCE

- A clear Theory of Change is necessary to understand how a proposed project will lead to a certain impact.
- The Theory of Change in "tech for good" projects may be oversimplified or misguided due to insufficient understanding of the problem or overconfidence in a technical solution.
- To determine if a Theory of Change is misguided, consider the assumptions behind the project and challenge them in conversation with the grantee or vendor.

Theory of Change describes how and why a proposed project is assumed to lead to a certain impact. Sometimes the Theory of Change in "tech for good" projects may be too linear, oversimplified, or misguided. These can be due to the developers having an insufficient understanding of the problem space or the social problems that lie at the center of the issue. It may also be due to overconfidence in a technical solution to a decades (or centuries-old) social issue, sometimes referred to as "tech solutionism" or "technochauvinism." For example, a tool, program, or contract may offer the assumption that increasing surveillance improves policing and thus achieves greater public safety, which is oversimplified and misguided given the disproportionate impact of policing on lower-income Black, Indigenous, and people of color (BIPOC) communities.

To understand whether a Theory of Change is misguided or oversimplified, focus on the assumptions that informed the ideation of the project. In your conversation with a potential grantee or vendor, try to challenge those assumptions, understand the sources that informed those assumptions, and introduce hypothetical risks posed by the project. During these conversations, you might also identify whether the project offers a macro-level solution to a hyper-local issue or assumes a "one-size-fits-all" mentality.

EXAMPLE

Researchers in a university turned to machine learning-based solutions – using various data sources such as satellite images, temperature anomalies, and food production indices – to predict future "climate refugee" flows. Their main goal is to help humanitarian actors allocate tailored and timely resources based on estimates of refugee numbers, arrival times, points of entry, etc. Their goal is also to help governments prepare for future population influxes, develop integration policies, assess job market needs, etc.

However, many humanitarian advocates have criticized these types of prediction frameworks for over-simplifying the causes behind refugee flows, which often depend upon a complex and unpredictable web of political, economic, and social factors. Therefore, it's logical to expect that such well-meaning predictive experiments and optimistic assumptions about governments' political will for assisting asylum seekers may result in greater immigration controls or militarization of borders.

Questions to Identify this Red Flag

What is your organization-wide Theory of Change and how does this project fit into that Theory of Change?

How does your Theory of Change relate to others in the community that bring a different Theory of Change to the same problem? What are some examples?

What assumptions informed the development of this project? What information and experiences guided those assumptions?

What risks might your assumptions lead to and how would your project's Theory of Change tolerate or adapt to those risks?

What are some criticisms you have heard about this project from those who share your broader goals?

RESOURCES

- Six Theory of Change Pitfalls to Avoid
- Movement Ecology
- How to recognize AI snake oil , AI Snake Oil Substack
- "Good" isn't good enough

2

RED FLAG

The proposal is a strategic misfit; the product is not related to other projects/grants that the potential grantee works on.

AT A GLANCE

- Proposed project may not align with grantee's current projects/grants.
- This strategic misfit might happen due to insufficient preparation, including poor needs assessment, lack of training/expertise, and redirecting funding from proven approaches.
- To identify this red flag ask about prior experience, needs assessments effort, definition of success, and measurement/revert-back plan.

With the proliferation of data, funders have shown significant interest in funding data-driven and digital technology-enabled solutions. As a result, some NGOs and nonprofits— including many that may not have any experience with digital-enabled solutions — are moving into this field without adequate preparation, developing digital services either by partnering with technology firms or by cultivating in-house capacity.

However, these projects may be prone to failure due to the following factors: insufficient needs assessment, lack of training and in-house expertise in deploying and maintaining services, over-reliance on the service, and diverting funding from already proven approaches to new and shiny digital solutions.

EXAMPLE

A humanitarian organization started developing a chatbot to provide translation services and legal assistance for filling out forms for asylum seekers. As a result, they decided to cut the number of human translators and volunteer caseworkers. The chatbot relies on automated translation services. For some languages, it causes “lost in translation” issues.

However, there is not enough human assistance to troubleshoot issues. This causes delays and confusion among clients. Due to a lack of technical expertise and policy safeguards, there are also concerns about data leaks, identity theft, and exposure of asylum seekers’ personally identifiable information (PII) to government agencies in home and host countries.

Questions to Identify this Red Flag

Does the organization have prior experience implementing technology-enabled solutions? If not, what is your plan for gaining technical expertise? Have you considered partnering with a group that has more technical expertise? If so, how did you decide on choosing them?

How does this project fit within the current work at your organization? How does it fit within your longer-term organizational goals?

What types of assessments have you conducted to understand the necessity of this project? With whom did you conduct these needs assessments?

What types of organizational changes will you go through as a result of developing this project? (e.g., diverse funding, restructuring teams, new partnerships, new roles, new organizational training)

What does success look like? How does the community you are intending to impact view success?

What is your plan to measure the success of the project, and what is your revert-back plan in case of failure?

3

RED FLAG

The project is merely a new product with no prospect of policy, cultural, or systemic change. The solution promises to provide a quick fix ("band-aid") to a long-standing issue.

AT A GLANCE

→ Proposed project may only offer a quick fix without creating policy, cultural, or systemic change.

→ It lacks critical thinking and adequate metrics to measure long-term impact, resulting in a focus on short-term quantifiable metrics and reinforcing systemic discrimination.

→ To identify this red flag ask about the social and economic impacts of the project, evidence of past success/failure, and strategies on supporting necessary policy changes.

During our interview with a civil society member, they mentioned that "tech for good" projects in the public sector are often built "without thinking critically about what they're trying to accomplish and whether or not technology is the best way to accomplish those goals."[1]

This issue arises not only because there is a lack of critical thinking, but also because of a lack of adequate metrics to define and measure success. In software systems, the criteria for "success" are often overtly quantitative. These quantifiable metrics may include the number of active users, the speed and reliability of a system, its efficiency or cost savings, and the accuracy of the output compared to a certain benchmark.

However, the factor that should differentiate public sector digital services from any other digital product is their longer-term impact. And it is not always possible to measure such an impact quantitatively over a short period of time. Focusing disproportionately on short-term and quantitative metrics may distract funders/vendors from assessing the longer-term results of a project such as its impacts on public policies, legal reforms, social movements, and addressing power asymmetries within and between government agencies, companies, and community groups.

This may result in reinforcing systemic discrimination and replicating politics as usual. Moreover, if the technology is trying to fix a systemic problem, the funder should take even greater care to ascertain whether this is a band aid solution to addressing an issue that instead requires larger scale legal or political reform, targeted funding, etc.

EXAMPLE

A location-based algorithmic tool drives a policing program to help predict where crime will occur, derived from technology deployed in Iraq and Afghanistan. The impact of the project is measured primarily quantitatively, overlooking long-term impacts on local communities. This tool is an example of a short-sighted response to mitigating crime, rather than taking steps to reduce policing and thus mitigating other factors that contribute to crime.

[1] "From an interview with a director of a civil rights organization.

Questions to Identify this Red Flag

How does this project help reveal underlying social and economic issues? (e.g., unjust housing practices, power imbalances in the criminal legal system)

What does "success" look like in 1 year, 5 years, and 10 years? (Depending on the type of the project, the time period can differ.)

Do you have qualitative and/or quantitative evidence relating to how a similar product/service has worked, or why it has failed in the past?

How has the uptake of this product been measured (qualitatively and quantitatively)? How will this lead to harm reduction to affected communities?

What types of policy changes do you envision as a result of this project's uptake?

What policies have to change to make the tech solution truly viable? Are you supporting the advocates pushing for these policies?

RESOURCES

- Smart City Playbook, the City of Boston
- ShotSpotter Alerts Rarely Lead to Evidence from The City of Chicago Office of Inspector General's (OIG)
- Predictive Policing Explained
- Pitfalls of Predictive Policing
- Robinson, David G. Voices in the Code: A Story about People, Their Values, and the Algorithm They Made. Russell Sage Foundation, 2022.

RED FLAGS CATEGORY

Business Model and Funding

Ford Foundation

4

RED FLAG

The product generates revenue from affected communities.

AT A GLANCE

- Sometimes a proposed product may generate revenue from affected communities by relying on optional paid or tier-based subscriptions.
- The business model may lead to lower quality free services that collect excessive or opaque data, or use deceptive design practices to pressure users to upgrade.
- To identify this red flag ask about the organization's business model, funding sources, their sustainability and funding diversification plan.

In some cases – especially in the for-profit space – a vendor’s business model may rely on optional paid or tier-based subscriptions from communities that it is meant to serve. Often, the justification behind this practice is to create a self-sustaining business model. A common practice is to provide a baseline free program for all members; however, if a person is interested in add-on/premium programs, they must pay for them. As a result, they deprive people without resources of the best level of service that they too deserve, consistently preferencing those with resources or running people into debt to get them.

Experience has shown that sometimes the free program relies on excessive and opaque data collection practices or ad revenue, has lower quality, or uses deceptive design practices (also known as “dark patterns”) to put pressure on users to upgrade.

EXAMPLE

A company wins a bid with Correctional Agencies in a state to facilitate e-learning programs in the state’s jails and prisons. They provide “free” tablets to individuals with only a few pre-installed programs. However, all other features, including reading e-books, are marked as “pay to play” and are charged by the minute.

Questions to Identify this Red Flag

Can you tell us about your investors and funders? How do you solicit/select funding streams? Do you have particular value-based criteria for funding?

What is your business model? As you grow, how do you foresee your business model changing or expanding?

For for-profit vendors:

- Do you provide paid tier-based/premium services to your clients?
- What are the differences between paid versus free programs? This can include questions about add-on services, privacy policy and users' data collection, personalization, quality of services, and data network requirements.

For non-profits:

What is your plan for sustainability and diversifying your funding? What are your thoughts about tier-based services as a model to generate revenue?

RESOURCES

- Business model red flags: 24 ways in which businesses could be wired to put people at risk
- Best practices for prison and jail tablet procurement

5

RED FLAG

The project depends on harmful surveillance – either by corporations or government agencies – regardless of framing.

AT A GLANCE

- The project may depend on excessive data collection and harmful surveillance by corporations or government agencies.
- This will negatively impact privacy and other fundamental rights, such as freedom of expression, assembly, association, movement and safety.
- To identify this red flag ask about the business model and its dependence on user data, data brokers, and the potential for government surveillance and safeguards against it.

Data-driven technologies, as their name implies, depend on collecting data from users. This data is often collected from people either with or without their knowledge. When a business model relies on user data, there might be a risk of contributing to surveillance capitalism (e.g., ad targeting, data brokers' opaque practices) or government surveillance. Both types of surveillance may negatively impact the right to privacy and consequently, the right to freedom of expression, the right to freedom of assembly and association, freedom of movement, etc.

EXAMPLE

Detaining migrants is not only inhumane, but also expensive for governments. As a result, governments have turned to alternative methods that present as more humane, facilitate integration into host communities, and also offer greater cost-efficiency. The U.S. Immigration and Customs Enforcement (ICE) has a contract with a private company to provide Electronic Alternative to Detention (e-ADT) for migrants. The company provides ankle monitors and remote case management services. The company collects migrants' sensitive data, including real-time location and voiceprints. Human rights groups have raised concerns about the constant surveillance of migrants in addition to violating their rights to dignity, freedom, and liberty.

Questions to Identify this Red Flag

What is your business model? Does it depend on or provide targeted advertising?

Does your program depend on buying/selling data to data brokers?

If your program requires working with public agencies (by using public data or providing data from public agencies), how do you think your project may contribute to government surveillance? What are your safeguards to avoid such a contribution?

RESOURCES

- It's the Business Model: How Big Tech's Profit Machine is Distorting the Public Sphere and Threatening Democracy
- Surveillance Capitalism and the Challenge of Collective Action
- Sidewalk Labs: Privacy in a City Built from the Internet Up
- Sideways: The City Google Couldn't Buy
- Logic Magazine, Community Defense: Sarah T. Hamid on Abolishing Carceral Technologies

RED FLAGS CATEGORY

Organizational Governance, Policies and Practices

Ford Foundation

6

RED FLAG

The governing body and the team behind the project are homogeneous in demographic, background, and expertise; the team is structured in such a way that the knowledge or decision-making power is concentrated within a small group of individuals.

AT A GLANCE

→ The lack of diversity in demographics and backgrounds among the team members may result in a disconnect between the team and the community they are attempting to serve.

→ The concentration of knowledge and decision-making power within a small group of individuals can result in a bottleneck effect and an increased chance of biased viewpoints.

→ To identify this red flag ask about team member experience, organizational structure, expertise and demographics of board and advisors, diversity and hiring practices.

If the governing body or team that is creating the product comes from a uniform and dominant group (such as being uniformly white, male, and/or young) and is attempting to solve a problem in a community who has been historically excluded, then there is a major mismatch between the team and people who will use and/or be affected by the product.

Sometimes, no team members have or are connected with direct/lived experience with the communities or problems being targeted (e.g., policing in BIPOC communities or individuals impacted by the carceral system, or experience with people with disabilities).

Alternatively, the team may lack the necessary expertise (technical, issue-based) and/or sensitivity surrounding the issues and sociopolitical context. This consideration also applies to the organization's governing bodies, including the board of directors and advisors. Given the gender and racial gaps in the technology sector,[1] the diversity and inclusion in those projects should be a subject of further scrutiny.

A red flag also occurs when all the knowledge or decision-making power is concentrated in a small group of individuals. This may be unsustainable from a product development perspective because there may be a bottleneck effect. If an organization is understaffed they may struggle to support the community they are supposed to be serving, even with the best intentions. Additionally, when decisions are made by a small group of individuals, there may be an increased chance of biased viewpoints.

EXAMPLE

A city government partners with a local anti-human trafficking nonprofit and a major technology company to organize a "tech for good" hackathon for identifying and breaking human trafficking patterns. After the hackathon, one of the participating groups decides to expand its proposed project and start a nonprofit. The founding members are exclusively former tech workers, all white and majority male. From early on, because of their lack of expertise and connections to the advocacy space, they face multiple issues, ranging from a lack of knowledge about culture and language to not being able to develop a trustworthy relationship with relevant advocacy groups among survivors, domestic workers, and sex workers.

[1] Molla, Rani, and Renee. Lightner. "Diversity in Tech," 2016, https://graphics.wsj.com/diversity-in-tech-companies/

Questions to Identify this Red Flag

Do any of your team members have direct/lived experience with the community(ies) or issue being addressed? How did you put your team together?

Can you tell us about your organizational chart/structure? Who do you categorize as technical/policy/advocacy experts on your team and how do you define expertise?

Can you tell us about the expertise and demographics of your board of directors and/or advisors?

Why is diversity important for your organization?

Can you tell us about your hiring practices and the steps you have taken to ensure that you have built a diverse team?

What policies have to change to make the tech solution truly viable? Are you supporting the advocates pushing for these policies?

RESOURCES

- A primer on Agile development in the public sector
- The U.S. Government Accountability Office's Agile Assessment Guide

7

RED FLAG

There is insufficient disclosure about the project's privacy policy, terms of service, and algorithmic use policies (if applicable). Furthermore, there is no process for obtaining meaningful informed consent from communities.

AT A GLANCE

- Informed consent is often not obtained from communities, and the consent process is often limited to generic terms of service and privacy policies written in inaccessible jargon.
- To identify this red flag ask about privacy policies, the process for obtaining informed consent, and how the vendor manages situations where consent is not knowingly given.

The first step for vendors to show their commitment to beneficiaries' fundamental rights is to disclose their policies around data usage, privacy, terms of service, and algorithmic use, among others. Transparency enables the demand for accountability – which may diminish in public-private partnership projects.

In addition, not being able to obtain informed consent is one of the main pitfalls of any data-driven project. Often, the process for obtaining informed consent is restricted to providing generic terms of services. Those documents are often written using inaccessible, lengthy, and legal jargon. Other issues include deceptive design practices (e.g., obfuscating or hiding website cookies settings) and a lack of alternatives if a person refuses to use or be a subject of those tools. In addition, technologies that are used by public agencies are often used on people without their knowledge (e.g., smart city projects, sharing databases between multiple public agencies and law enforcement).

While ultimately it is public agencies' responsibility to ensure the public's interest in consent, transparency, and oversight mechanisms in every stage of contracting, it is the vendors' responsibility to provide accessible information, design interfaces for obtaining informed consent, develop policies and make them publicly available, be transparent about their third-party relationships, and set boundaries for data co-ownership with public agencies.

EXAMPLE

As part of their "Green City, Smart City" initiative, a city decides to fully switch to paperless subway tickets. They have a contract with a company to develop an app. The city is also interested in integrating all other ticket-based transportation services including parking tickets and traffic tickets. It will provide APIs so other developers can use the city data and propose new digital services. During the sign-up, users must accept lengthy terms of service agreements. However, they are not fully aware of how their transportation data is used, what other data is collected, with whom the data is shared, and what consequences this might have.

Questions to Identify this Red Flag

Ask the potential grantee/vendor to provide privacy policies, terms of services, algorithmic use policies, developer policies, etc. Ask about the process of coming up with those policies and whether the policies are publicly available, and if not, why? Moreover, how understandable are these policies for the layman?

What does “informed consent" from communities mean to you? Describe your process for obtaining informed consent mainly from users of your product.

How can you collaborate with public agencies to obtain informed consent? What options do beneficiaries have if they decide not to use or be a subject of using your tool?

How do you manage a situation where you later learn that a subject did not knowingly consent to data collection or sharing?

RESOURCES

- Ranking Digital Rights Corporate Accountability Index: privacy, freedom of expression, and governance indicators
- Smart City PDX, On the road for the Surveillance Technologies Policy

8

RED FLAG

There is no formal process for conducting a human rights or algorithmic impact assessment and/or a mechanism to track, report, and remediate harms.

AT A GLANCE

- Potential grantees should regularly evaluate their products' and policies' impact on society, with processes in place to report and address any potential harm.
- They should have a strategy for receiving reports and remedying potential harms.
- To identify this red flag ask about the frequency and methodology of impact assessments.

It is important for vendors to conduct impact assessments to understand the direct and indirect societal harms of their project, especially on underrepresented and underresourced communities.

Impact assessments can be done internally or with the help of independent external experts. Funders should require potential grantees to report on their impact assessment methods and their findings. Funders can also connect their (prospective) grantees to experts who can help with conducting and developing impact assessment processes – not as a one-time assessment but rather as a continuous one.

In addition, vendors should have a strategy (e.g., feedback channel, functioning email addresses, “Contact” or “Report Issues” online forms on their own or government’s website) to receive reports about the potential and actual societal harms of their products from beneficiaries and third-party advocates/researchers. Being transparent about those harms and having a strategy to redress harms is a must.

EXAMPLE

City A is a “sanctuary city.” Undocumented immigrants can purchase e-tickets with less fear of government surveillance and consequences such as arbitrary detention and repatriation. A year later, the same company which built the e-ticketing service wins a bid to develop similar services for City B. City B has a very strict policy on immigration and is not a sanctuary city for undocumented immigrants. Without conducting a thorough human rights impact assessment on design choices, app features, data collection, and data sharing practices, undocumented immigrants (in both City A and B) may be at risk of surveillance, violation of their rights to freedom of movement, and potentially even detention and repatriation.

Questions to Identify this Red Flag

What unintended negative consequences are possible as a result of this product? Are you willing to abandon the product if the negative consequences are too great?

Do you conduct any types of impact assessment? If yes, do you publish impact reports in a manner/format that is widely accessible?

What mechanism do you offer to prevent, mitigate, and remediate harm?

Do you provide any feedback channel for receiving reports about any harm as a result of your product design and deployment?

Do you ask for indemnification in your contract? What are you liable for if you do cause harm?

RESOURCES

- Digital Rights Check
- The Santa Clara Principles On Transparency and Accountability in Content Moderation
- United Nations Guiding Principles on Business and Human rights (UNGPs) B-Tech Project

A list of risk assessment and documentation tool:

- https://github.com/users/royapakzad/projects/3
- Data Protection Impact Assessment template
- Assembling Accountability: Algorithmic Impact Assessment for the Public Interest

9

RED FLAG

There is not enough knowledge about technology standards and regulations that apply to vendors' practices.

AT A GLANCE

- Vendors may have incomplete knowledge of local, state, and federal regulations relevant to their practices.
- Lack of engagement with technology standards and regulations can result in "one-size-fits-all" project design.
- To identify this red flag ask about the impact of relevant regulations and an organization's stance on them.

In some cases, vendors may possess insufficient or incomplete knowledge about local, state, and federal regulations that are applicable to their practices. Similar to Red Flag 3 (on "band-aid" fixes), this shows a lack of engagement with policy and advocacy space. This may also result in a "one-size-fits-all" approach to project design. In addition, being knowledgeable about technology standards (e.g., standards published by national bodies such as NIST, professional codes published by the ACM or IEEE), and best practices helps their products to be reliable.

EXAMPLE

An electronic medical records company is not familiar with local medical privacy regulations. The company complies with federal regulations but has not developed infrastructure to handle evolving local privacy regulations. In particular, parts of the software assume that certain data is available, when in reality the data is only accessible in certain states.

Questions to Identify this Red Flag

What are the key regulations/laws/regulatory proposals that apply to this project?

How does [applicable regulation/regulatory proposal] affect this project?

What is your position on [applicable regulation/regulatory proposal]? This question depends on funders' knowledge about the regulatory space. For instance, you can ask about the Community Control Over Police Surveillance (CCOPS) Model Bill, or the Algorithmic Justice and Online Platform Transparency Act.

Is changing [applicable regulation] something you commit resources to? Is your service part of a larger set of goals?

RESOURCES

- Federal & California ai legislation database from the CITRIS Policy Lab
- Privacy: California Consumer Privacy Act (CCPA), Illinois Biometric Information Privacy Act.
- Local Surveillance Oversight Ordinances
- Federal Fair Housing Act
- NIST’s Face Recognition Vendor Test (FRVT)
- IEEE P7000TM Standard
- Making Smart Decisions About Surveillance

10

RED FLAG

Based on the vendor's current policies, there are not enough safeguards for preventing harm during organizational restructuring, spin-offs, merges, or dissolution.

AT A GLANCE

- Vendors should have clear policies in place to prevent harm during organizational restructuring, spin-offs, mergers, or dissolution.
- They should have a data ownership policy that stays robust during major organizational changes, backed up with dedicated resources and a written policy.
- To identify this red flag ask about the safeguards potential grantees have in place for these hypothetical scenarios, such as impact assessments and data ownership policies.

Often, "tech for good" projects start as experiments. For instance, a city decides to pilot a new welfare distribution platform. However, there is no clarity about the consequences of program failure or what happens if the city decides not to continue working with the vendor. Funders should ask potential grantees about the safeguards they have in place during these hypothetical scenarios.

In addition, vendors should have a data ownership policy from the start: Who owns the data after potential major restructuring such as mergers and spin-offs? They should guarantee that their policies for harm prevention and mitigation stay robust during situations such as major organizational changes e.g., changing top executives, changing business models or creating a spin-off nonprofit from for-profit vendors or vice versa, entering a new market, and being acquired by or merged with other companies. This should be backed up with dedicated resources, such as teams or individuals, who are committed to the program's maintenance, or having a written policy for sunsetting a program that includes a policy for deleting user data.

EXAMPLE

A nonprofit that maintains a suicide hotline shared its beneficiaries' anonymized and unidentifiable information with its spin-off organization. The for-profit company provides audio-based emotion recognition services. Both the for-profit and nonprofit organizations share the same CEO. Privacy activists and mental health advocacy groups raise concerns about this data sharing relationship. They believe extracting commercial value from people's most sensitive and vulnerable conversations is unethical. There is a resulting lack of trust in such services despite being backed and promoted by government public health agencies.

Questions to Identify this Red Flag

Do you have any written policy to show that you will be conducting impact assessments if your organization goes through major changes (spin-off, changes in business model, acquisition, dissolution)? What's your data ownership policy during these major changes?

Did you pilot this project anywhere? If yes:

- Did you have an exit interview with users when the pilot ended?
- What safeguards did you have in place to protect users' data?

RESOURCES

- Crisis Text Line, from my perspective

RED FLAGS CATEGORY

Product Design, Development, and Maintenance

Ford Foundation

11

RED FLAG

The product claims to be completely new, “disruptive,” or different in all relevant facets.

AT A GLANCE

- Products claiming to be completely new or "disruptive" should be scrutinized, as they may not consider past lessons learned in the field.
- The vendor's solution may create new harm in other domains
- To identify this red flag ask about direct and indirect stakeholders, who are the vendor’s competitors, and how is the proposed product or technology similar and different from existing solutions.

In most cases, a product that claims to be completely new or different from all precedents is not completely new, especially in the public sector. If a vendor claims to provide a groundbreaking solution to a public interest issue, it is possible the vendor has not engaged with the existing domain and current players.

The vendor may not build upon lessons that past products/vendors have learned. Thus, the vendor may repeat or amplify past mistakes. Moreover, the vendor’s solution may be solving an issue in one domain, but doing so by creating new harm in another.

EXAMPLE

A potential vendor may claim to predict how well a child will perform in school by feeding a large amount of personal data into an AI model. Prior research has shown that in reality, even large machine learning models with access to fine-grained data collected over years for each child is unable to outperform a simple regression model using a few data points.[1] In this case, the use of AI can be an excuse to access/collect more data.

QUESTIONS TO IDENTIFY THIS RED FLAG

Who are the current stakeholders and parties involved in this domain?

Who are the key competitors and how does this product differ from them?

How is the proposed product or technology similar and different from existing solutions?

RESOURCES

- How to recognize “AI Snake Oil”
- On NYT Magazine on AI: Resist the Urge to be Impressed

[1] Salganik, Matthew J., Ian Lundberg, Alexander T. Kindel, Caitlin E. Ahearn, Khaled Al-Ghoneim, Abdullah Almaatouq, Drew M. Altschul et al. "Measuring the predictability of life outcomes with a scientific mass collaboration." Proceedings of the National Academy of Sciences 117, no. 15 (2020): 8398-8403. https://www.pnas.org/doi/full/10.1073/pnas.1915006117

12

RED FLAG

The product replaces an existing product with an interface that is very different from the one that users are accustomed to or the user interface/design is inaccessible to people with disabilities or intimidating to those lacking technical or digital literacy.

AT A GLANCE

- Significant changes to a user interface can cause confusion and limit accessibility, especially for those with limited technical or digital literacy.
- To identify this red flag ask whether the product has been tested with a range of users with varying degrees of digital fluency and disabilities, what are the potential consequences of accidental misuse, and how does the design differ from what users are accustomed to.

Significant changes to an interface, especially one that users have grown used to, can cause confusion or open the door to inadvertent misuse. Since large changes in UI often assume specific technological literacies, these changes can make technology more difficult to use for clients, limiting accessibility.

EXAMPLE

A prison system migrates from a PC-based communication system to a tablet-based one. Further, incarcerated people use the communication system to call family members. It is possible that, for instance, those currently incarcerated are allowed to use the device for a limited period of time. If they are not used to the tablet interface and controls, it is possible that it will take them much longer to navigate the system, reducing the amount of time for the call itself. Even logging into the system may be a challenge if there are people who remember passwords via muscle memory on a traditional keyboard.

QUESTIONS TO IDENTIFY THIS RED FLAG

How does the design differ from the interface that clients are accustomed to using?

What user testing has been done to show that the design is accessible; namely have you tested the product with those with various degrees of digital fluency?

What are ways that the technology could be accidentally misused? If the interface is inadvertently misused, what is the range of possible consequences? Is it possible to remediate the consequences?

How do you plan to overcome mistrust or unfamiliarity in order to increase adoption and impact?

RESOURCES

- User Interface Design for Low-literate and Novice Users: Past, Present and Future
- Making the Web Accessible: Strategies, standards, and supporting resources to help you make the Web more accessible to people with disabilities.
- Protecting Older Users Online
- Building Access: Universal Design and the Politics of Disability

13

RED FLAG

The proposed project does not sufficiently follow industry best practices including security, privacy, openness, interpretability, and non-discriminatory design.

AT A GLANCE

- Products should follow technical industry best practices.
- Products should be tested for stability, encryption, and resistance to cyber-attacks for cybersecurity, and adhere to "privacy by design" principles for privacy.
- Tools used by public agencies should be audited for discrimination, interpretability, and accountability, and vendors should be transparent about the results of these audits and their decision-making processes.

From a technical standpoint, there are several industry standards against which developers can test products. Testing is very much dependent on the type of technology. We understand that philanthropic organizations and government agencies may not have the in-house expertise to fully run all the necessary tests. However, having general knowledge about those testing criteria is necessary. In addition, organizations can work with external experts to be able to test product/service performances based on the relevant standards.

For cybersecurity, products should be tested for stability, authentication, encryption, and resistance to cyber-attacks. For privacy, products should adhere to “privacy by design” principles including minimal data collection, privacy-by-default settings, and retaining data as long as needed. For human-rights-centric UX/UI[1] design, in addition to security and privacy, products should be tested based on accessibility criteria such as network and device quality, beneficiaries’ digital literacy levels, physical and mental impairment, etc.

Furthermore, computational tools that help public agencies to make decisions about certain applications (e.g., predictive risk assessment tools in child welfare practices, student assignment algorithms for public schools) often rely on historical and demographic information. Researchers have shown that these tools are prone to discrimination based on gender, race, religion, and other socioeconomic factors.[2] Apart from assessing the more complex and long-term impacts of these tools, these tools should also be audited based on other technical criteria. For instance, issues around over and under-representation arise during the process of collecting and annotating data that is used to train, validate, test, and optimize the system.

In addition, these systems are prone to making “unfair” decisions based on the input variables (whether directly about protected groups such as race, religion, or age, or proxies such as zip code, phone area code, education level) and statistical models that are selected during the design and development process. Vendors may also use more complex technical methods to design these systems.

When the system becomes very complex, there might be no transparency in how that system makes a certain decision (e.g., why this tool thinks that family A should be denied access to welfare benefits but not family B; why asylum seeker A’s application should be granted but not B’s).

A lack of interpretability in these systems may lead to confusion and weaken beneficiaries’ ability to hold public agencies to account. When mistakes are made, there is no clear answer to who should be blamed: the tool, the vendor, the public employee, or the agency?

QUESTIONS TO IDENTIFY THIS RED FLAG

What industry best practice standards did you use to design and test your product? Use the following resources and ask about privacy, security, fairness, interpretability, accessibility, openness, and sustainability.

Can we or our trusted technical partners test your product? Do we need to sign an NDA for it? If yes, why and what does it include?

Is your product documentation available publicly? If not, what prevents you from keeping the documentation open?

Do you perform tests to determine whether your product or tool is creating discriminatory, adverse outcomes for certain demographic groups? If so, how do you obtain the demographic data in order to perform these tests?

Can you share your audit and/or impact assessment reports? Who conducted the audits? Ask whether the audits have been conducted by the company itself, by consultants who were commissioned by the company, or externally by advocates, technologists, and researchers.

RESOURCES

For Privacy and Security

- The Digital Standard by Consumer Reports
- The Open Web Application Security Project or OWASP's Mobile Security Testing Guide
- The OWASP's Testing for Weak Encryption
- The Mozilla Observatory
- Privacy by Design: The 7 Foundational Principles
- Security Planner

For UX/UI Design

- Digital Security and Privacy Protection UX Checklist

For Algorithmic Fairness

- White House Blueprint for an AI Bill of Rights
- Microsoft Fairlearn toolkit
- IBM Fairness 360
- Google's "What If?" Tool
- Eticas Guide to Algorithmic Auditing
- Other tools

RESOURCES

For Algorithmic Fairness (Continued)

- Datasheets for Datasets
- Model Cards for Model Reportinga

For Algorithms Interpretability and Explainability

- Introduction to Interpretable Machine Learning (I, II)
- AI Explainability 360, IBM

For Sustainability

- Principles of Green Software Engineering

For Openness

- Critical Digital Infrastructure

Other Resources for Responsible Product Design

- Value Sensitive Design: Envisioning Cards
- AI in Education Toolkit for Racial Equity: How to mitigate racial bias in the design and development of your product
- AI risk management framework
- Outsider Oversight: Designing a Third Party Audit Ecosystem for AI Governance

[1] User Experience/User Interface

[2] Eubanks, Virginia. Automating inequality: How high-tech tools profile, police, and punish the poor. St. Martin's Press, 2018.

14

RED FLAG

Vendor is not able to explain how the product/service works in an accessible manner, without using technical terms.

AT A GLANCE

- Excessive use of technical terms and acronyms without clear explanations can indicate that a vendor is over-selling their product.
- Inability of the vendor to explain the product can result in public agencies also being unable to understand it, leading to accountability issues.

Vendors should be able to explain how their tool works in an accessible manner. During our interview with digital rights advocates and social entrepreneurs, it became clear that throwing in terms and acronyms such as AI, algorithm, machine learning, deep learning, blockchain, etc. without being able to simply explain why that technology is used in a system is an indicator of a vendor over-selling its service.

In addition, if a vendor is not able to explain its service, public agencies that use the service will not be able to either. This will lead to a further lack of accountability.

EXAMPLE

A civic tech company proposes developing a method for faster and safer political participation such as voting. Their method relies on blockchain technology. During their conversation with funders, they use terms such as “blockchain,” “public ledger,” “private ledger,” “chain,” etc. without elaborating what they mean by the terms and why this technology is relevant.

QUESTIONS TO IDENTIFY THIS RED FLAG

Can you explain how your tool works for users and stakeholders with a lower level of technical literacy? If not, what is limiting you from doing so?

15

RED FLAG

The product locks you in and/or is not easily repairable.

AT A GLANCE

- Products that lock in users and are not easily repairable can cause problems in switching to another better product.
- To identify this red flag ask about compatibility of the product with standard data formats, proprietorship of the product, issues with rolling out updates, repairability of services, and processes for minimal disruption to end users in case of technical issues.

Sometimes products are designed in such a way as to make switching to another (better) product costly and resource-intensive. Sometimes a tool may require proprietary data formats that are only compatible with that certain software. This becomes a serious issue when a vendor and public agency work together to pilot a service. After the pilot phase, it may be too costly for the agency to switch to other services. They may simply decide to continue working with that vendor because they are "locked in" with them. Proprietary software can also make it difficult for public agencies to repair a system or service.

During the course of our interviews, several entrepreneurs and advocates mentioned that these issues can arise due to the lack of interoperability among systems. There is no doubt that the interoperability of digital systems is important; after all, the Internet is built on the principle of interoperability and seamless information exchange.

However, from an anti-surveillance perspective, interoperability of data sharing systems between and within governments, without adequate safeguards, may result in harmful consequences. An example could be frictionless data sharing practices between police departments, Immigration and Customs Enforcement agencies, and other public offices that are involved in managing education, health care, and welfare services

In this example, this may result in surveilling refugees and immigrants, arbitrary arrest, and denying them access to public spaces/services.

EXAMPLE

A technology vendor wins a bid to develop custom-built data management systems for a county. A few years ago, a city in that county updated its data management system. However, the city's system is not compatible with the one for the county. To solve the issue, the vendor proposes to update the city's internal system as well.

In addition, the vendor proposes to sell other custom-built add-on services (project management system, internal messaging platform, invoice management, etc.). This pattern repeats itself every time the county, city, or state needs to upgrade its digital infrastructure. The chaotic situation hinders public officers' services while exposing sensitive government data to instability and malicious activities.

QUESTIONS TO IDENTIFY THIS RED FLAG

If your product needs data as its input, what kinds of standard data formats is it compatible with? Can the data be exported to similar products?

What proprietorship do you have on this product?

Have you experienced any issues with rolling out updates to your users – either end users or public agencies?

How can public agencies repair your services? Do they need to keep in touch with you as long as they use this service?

What processes do you have in place to ensure minimal disruption to end users/affected communities in the event of technical issues?

RESOURCES

Digital rights to repair act

16

RED FLAG

Developers don’t make explicit how a certain technology or product will be maintained or adapted in the future.

AT A GLANCE

- A lack of explicit maintenance plans for technology or products can lead to a lack of sustainability, making it harmful in the future if the environment, ecosystem, or regulations change without updates.
- To identify this red flag ask about the frequency of updates for the technology and training data, preparation for seamless transfer of services, and willingness to be held accountable for maintaining a certain quality of service.

Without making explicit how a certain technology or product will be maintained or adapted in the future, vendors can inadvertently create a product that is helpful today but harmful in the future.

Without stating how a technology will be maintained in the future, vendors indicate a lack of sustainability. Technology can become faulty and harmful if the environment, ecosystem, or regulatory landscape changes without updates to the technology.

EXAMPLE

A research team at Vanderbilt University was able to show that a model trained to predict hospital mortality rates using data from 2006 deteriorated in quality over time. One model in particular provided a mortality rate prediction for 2013 when the actual observed value differed by 25 percent from the predicted value. Because of shifts in the cases that hospitals in the areas were treating the original model had gone stale.

QUESTIONS TO IDENTIFY THIS RED FLAG

How and how often will the technology and/or training data be updated once it is released?

Can you describe your client training program? How do you help your clients have direct access to you or become independent in maintaining and troubleshooting the service?

Are you prepared to work with a future vendor to seamlessly transfer services without interruption? How?

Would you be open to contract’s violations terms that impose fines if you don’t maintain a certain quality of service threshold?

RESOURCES

- A primer on AI model drift
- The Maintainers

RED FLAGS CATEGORY

Third-party Services and Supply Chain

Ford Foundation

17

RED FLAG

Vendor lacks sufficient knowledge about the hidden labor that goes into the design, development, and maintenance of a service.

AT A GLANCE

- The vendor may lack knowledge about the hidden labor involved in design, development, and maintenance of a service, which may contribute to exploitative labor practices.
- This hidden labor is mostly performed by content moderators, hotline coordinators, and data annotators, who often lack job benefits and fair compensation.
- To identify this red flag ask about the vendor's knowledge of their labor forces, supply chain mapping, and due diligence regarding third-party services.

In their book Ghost Work: How To Stop Silicon Valley From Building A New Global Underclass, Mary L. Gray and Siddharth Suri reveal the “vast, invisible human labor force” behind data-driven digital systems. Content moderators, hotline coordinators, and data annotators numbers among the “ghost workers” who are usually underpaid, overworked, and without normal job benefits such as health insurance. The vendor should know the “invisible” human labor that goes into its services’ supply chain, otherwise, it might unintentionally contribute to exploitative labor practices.

EXAMPLE

A company in the U.S. puts out a call for content moderation or data labeling jobs (ie determining content that is offensive), easily to be filled by remote workers in the Global South. Not only are there potential vast workers’ rights abuses and ramifications for workers’ rights and working conditions in the long term, but the choice of who to hire to label data and moderate content creation and dissemination affects the algorithms - and thus decisions made - by users all around the world, often with little to no oversight or accountability.

QUESTIONS TO IDENTIFY THIS RED FLAG

Beyond your own employees, can you tell us about the labor forces that make your services possible?

Have you mapped out your full supply chain? Do you have a policy ensuring that the supply chain is ethically sourced?

Do you work with any third-party services that provide crowdsourcing services? If yes, have you done any due diligence to ensure their practice is not exploitative?

RESOURCES

- We Are Dynamo: Overcoming Stalling and Friction in Collective Action for Crowd Workers
- Gray, Mary L., and Siddharth Suri. Ghost work: How to stop Silicon Valley from building a new global underclass. Eamon Dolan Books, 2019. (link)
- Roberts, Sarah T. Behind the screen. Yale University Press, 2019.
- The Supply Chain Risk You Didn’t Know About: Navigating Responsible Sourcing in AI
- Family Units: Can the families whose labor powers AI seize the means of production?
- Anatomy of an AI system
- What Green Costs

RED FLAGS CATEGORY

Government Relationships

Ford Foundation

18

RED FLAG

Vendor is not transparent or does not set red lines in its data sharing practices with government agencies.

AT A GLANCE

- Data sharing between vendors and public agencies could be complicated and opaque.
- Vendors should conduct due diligence, establish clear boundaries, and be transparent about their process for handling public data.
- To identify this red flag ask about vendor's relationship with law enforcement, response to government requests for information, and safeguards in their data sharing agreements with public agencies.

Often times, data sharing practices between vendors and public agencies are opaque and entangled. Vendors might rely on data that is provided by government agencies either through a government's open data platforms or closed data. In some cases, the source of this data is not obvious. For instance, it may not be clear if a public agency itself buys the data from data brokers or uses surveillance methods to collect it (e.g., social media scraping, realtime transportation data collected by law enforcement and then shared with trusted caseworkers who serve unhoused communities). The reverse scenario may be true as well; sometimes, the government requires vendors to give them access to data they collect.

Vendors should conduct their own due diligence to understand where government data comes from. In addition, they should be transparent about the process they have in place to handle government demands, subpoena, or court orders for handing out users' data. They should set red lines with governments to avoid contributing to surveilling traditionally excluded communities.

EXAMPLE

An entity comes forward with a proposal for a policing data collective – owned by former police officers. It is meant to keep local communities safe, but data sharing with local government/police departments remains unclear.

QUESTIONS TO IDENTIFY THIS RED FLAG

What is your relationship with law enforcement, if any?

How have you responded, or how would you respond, to government requests for information that you hold?

What safeguards exist in your data sharing agreements with public agencies to keep users' data secure?

19

RED FLAG

Due to trade secrecy, intellectual property rights, or proprietary software, the vendor has excessive confidentiality clauses to their contracts with public agencies, or has many non-disclosure agreements they require the agency to sign.

AT A GLANCE

- Lack of scrutiny by government agencies can lead to loss of public oversight over the system, and FOIA requests may be withheld.
- To identify this red flag ask about the type of licensing for the tool, reviewing confidentiality clauses, and inquiring about the vendor's willingness to waive their right if data is requested through FOIA.

This is a very common red flag, especially when a vendor provides the same commercial services to non-government agencies as well. Often, there is not enough scrutiny by government agencies to understand the consequences of these confidentiality clauses. As a result, there may be a loss of public oversight over the system and/or FOIA requests may be withheld. In addition, the government may lose its right to repair the tool.

EXAMPLE

A company wins a contract to develop a pre-trial risk assessment tool. The tool is deployed statewide. Racial justice advocates raise concerns about discriminatory outputs of the tool based on defendants' race. The company publishes a public statement claiming their tool does not take defendants' race into consideration at all.

However, advocates still think other factors that indirectly reveal defendants' race such as zip code, undergraduate or college club membership, or other "proxies" may contribute to discriminatory practices. They require state government agencies to scrutinize the data and algorithms behind the tool. However, the company cites its IP rights and trade secrecy and will not reveal such information to the public.

QUESTIONS TO IDENTIFY THIS RED FLAG

What types of licensing do you have for your tool? Are they open or closed?

If you already have a contract with a government agency, can you walk us through your confidentiality clauses? Why do you have those?

Are you going to waive your right if data you collect is requested through FOIA?

RESOURCES

- Best Practices for Government Procurement of Data-Driven Technologies
- AI and Procurement - A Primer
- School Procurement Guide: Buying Edtech Products with Racial Equity in Mind
- Gizmodo, The FBI Will Neither Confirm Nor Deny the Existence of These Documents I Just Printed

RED FLAGS CATEGORY

Community Engagement

Ford Foundation

20

RED FLAG

Tokenism in community engagement where engagement is not meaningful and is treated as a checkbox.

AT A GLANCE

- Tokenism can indicate a lack of prioritization of understanding how technology affects impacted communities.
- To identify this red flag, ask about vendor's engagement with community advocates, the frequency and nature of the engagement, whether there was compensation, and how they handle any weaknesses identified by community-based organizations.

Tokenism is a performance by vendors to portray an image of community engagement without building relationships. This can indicate the deprioritization of understanding how technology affects impacted communities.

EXAMPLE

A start-up that provides services to individuals impacted by incarceration claims to have the support of grassroots groups, citing endorsements from leaders within the interest groups. However, it is later identified that while the start-up may have had a one-time conversation with the interest group in the early ideation stages, the start-up did not maintain any standing meetings with the interest group during the actual implementation phase.

Thus, while the startup may have received endorsements for the product concept it is possible that the same interest groups do not approve of the final product itself.

QUESTIONS TO IDENTIFY THIS RED FLAG

What civil society organizations and community advocates have you been in conversation with? How often and in what capacity (e.g. focus group, advisor)? Have they been compensated for their engagement?

If a company claims to be endorsed by a community-based organization, ask for the strengths and weaknesses that the organization identified. What guidance did the organization provide?

Can we reach out to them?

RESOURCES

- Participation is not a Design Fix for Machine Learning
- Statement of resigning axon AI ethics board members

21

RED FLAG

Products are not designed with the impacted communities centered. There is no meaningful community engagement in the process of needs assessment, development, and implementation.

AT A GLANCE

- Barriers such as language and timeline can lead to a lack of meaningful community engagement.
- To identify this red flag, ask about the vendor's definition of meaningful engagement, stakeholder mapping, which groups have been consulted, and their diversity in expertise, sector, demographics, and geography.

"Community engagement" is a term that one repeatedly hears from technology developers and deployers. The promise is to involve affected communities in the process of building and deploying a certain technology. There are several barriers to reaching "meaningful" community engagement:

One is language and culture. The language used strips agency from the communities being targeted or demonstrates hubris about the value it can bring.

Timeline is another issue. Often the timeline for assessing needs, co-designing, feedback, and piloting does not allow for sufficient consultation with the communities affected by the product, and/or the product was deployed before local communities had the opportunity to weigh in with their questions and concerns for use.

Mechanisms or processes for stakeholder engagement can become ineffective as well - it is not accessible (e.g. in-person vs virtual participation, time of the day), fully inclusive, or sensitive to issues of power, race, gender, and class. Sometimes feedback received from Community Review/Oversight bodies is treated more like a checklist without having any "teeth" in terms of actually affecting the government or vendors' decision-making process.

EXAMPLE

Humanitarian applications of technology often purport to help clients and communities in need without engaging these communities in the design, implementation, and monitoring of these tools. Two examples are AI/algorithms for refugee resettlement (e.g. Stanford Immigration Policy Lab, International Rescue Committee's "Match" program), and the deployment of biometric identification, namely during COVID-19.

An entity may claim that there was too little time to incorporate feedback iteratively and consistently, or that communities were consulted at the outset - but not later on. The people designing the tool are also, more often than not, individuals without any lived experience.

QUESTIONS TO IDENTIFY THIS RED FLAG

How do you define "meaningful" engagement?

FOR THOSE WHO HAVE DONE STAKEHOLDER MAPPING

Which groups of people have you involved - and who have you overlooked or excluded, intentionally or not? Does the group of people you have consulted with represent sufficient diversity in expertise, sector, demographics, and geography?

Have you considered the mistrust of the justice system in a city by BIPOC or other marginalized communities?

How often do you speak with the affected communities and do you have any formal processes for doing so (monthly calls, focused groups, workshops, interviews, feedback channels)?

What is/has been your timeline for community engagement?

Did you use any virtual forums to gauge community feedback and interest?

RESOURCES

- Arnstein's Ladder of Citizen Participation
- What Words We Use — and Avoid — When Covering People and Incarceration

Appendix: List of 21 Red Flags

THEORY OF CHANGE AND VALUE PROPOSITION

- (1) The project does not offer a clear Theory of Change — or, if offered, it is tenuous, misguided, or oversimplified.
- (2) The proposal is a strategic misfit; the product is not related to other projects/grants that the potential grantee works on.
- (3) The project is merely a new product with no prospect of policy, cultural, or systemic change. The solution promises to provide a quick fix ("band-aid") to a long-standing issue.

BUSINESS MODEL AND FUNDING

- (4) The product generates revenue from affected communities.
- (5) The project depends on harmful surveillance – either by corporations or government agencies – regardless of framing.

ORGANIZATIONAL GOVERNANCE, POLICIES, AND PRACTICES

- (6) The governing body and the team behind the project are homogeneous in demographic, background, and expertise; the team is structured in such a way that the knowledge or decision-making power is concentrated within a small group of individuals.
- (7) There is insufficient disclosure about the project's privacy policy, terms of service, and algorithmic use policies (if applicable). Furthermore, there is no process for obtaining meaningful informed consent from communities.
- (8) There is no formal process for conducting a human rights or algorithmic impact assessment and/or a mechanism to track, report, and remediate harms.
- (9) There is not enough knowledge about technology standards and regulations that apply to vendors' practices.
- (10) Based on the vendor's current policies, there are not enough safeguards for preventing harm during organizational restructuring, spin-offs, merges, or dissolution.

PRODUCT DESIGN, DEVELOPMENT, AND MAINTENANCE

- (11) The product claims to be completely new, "disruptive," or different in all relevant facets.
- (12) The product replaces an existing product with an interface that is very different from the one that users are accustomed to or the user interface/design is inaccessible to people with disabilities or intimidating to those lacking technical or digital literacy.
- (13) The proposed project does not sufficiently follow industry best practices including security, privacy, openness, interpretability, and non-discriminatory design.
- (14) Vendor is not able to explain how the product/service works in an accessible manner, without using technical terms.
- (15) The product locks you in and/or is not easily repairable.
- (16) Developers don't make explicit how a certain technology or product will be maintained or adapted in the future.

THIRD-PARTY RELATIONSHIPS, INFRASTRUCTURE, AND SUPPLY CHAIN

- (17) Vendor lacks sufficient knowledge about the hidden labor that goes into the design, development, and maintenance of a service.

GOVERNMENT RELATIONSHIPS

- (18) Vendor is not transparent or does not set red lines in its data sharing practices with government agencies.
- (19) Due to trade secrecy, intellectual property rights, or proprietary software, the vendor has excessive confidentiality clauses to their contracts with public agencies, or has many non-disclosure agreements they require the agency to sign.

COMMUNITY ENGAGEMENT

- (20) Tokenism in community engagement where engagement is not meaningful and is treated as a checkbox.
- (21) Products are not designed with the impacted communities centered. There is no meaningful community engagement in the process of needs assessment, development, and implementation.

Acknowledgements

This project was spearheaded by:
**Cynthia Conti-Cook,
Technology Fellow at Ford Foundation**

The Foundation commissioned:
Taraaz to carry out the project.

TARAAZ
TECHNOLOGY & HUMAN RIGHTS

The research was developed and undertaken by:
**Roya Pakzad,
Sarah Ariyan Sakha,
and David Liu.**

The authors of this framework would like to thank our interview participants and advisory group for their generosity in sharing their knowledge and experience:

Octavia Abell (Govern For America)
Betsy Cooper (The Aspen Tech Policy Hub)
Kade Crockford (ACLU)
Hector Dominguez (The City of Portland)
Laura Guzmán (The Engine Room)
Alex Hanna (DAIR)
Sheila Herbert (The Measure for Justice)
Brian Hofer (Secure Justice)
Annalise Irby (Schmidt Futures)
Ruby Khan (Schmidt Futures)
Chris Kingsley (the Annie E. Casey Foundation)
Andrew Lama (Ameelio)
Maya Markovich (Justice Tech Association)
Kevin O'Neil (The Rockefeller Foundation)
Marika Pfefferkorn (Twin Cities Innovation Alliance)
Jake Porway
Aasim Shabazz (Twin Cities Innovation Alliance)
Jacob Sills (Uptrust)
Samantha Silver (The Measure for Justice)
Bianca Tylek (Worth Rises)
Andrew Warren (Recidiviz)
Harlan Yu (Upturn)

We would also like to thank the Ford Foundation's program officers for their contributions. Special thanks also to the Ford Foundation's Office of Communications and Zacht Studios for bringing this report to life.