

Responsible Data Program

Proposal for Funding 2015/2016

the engine room

<https://theengineroom.org>

September 2014

Summary.....	2
Context.....	4
Framing the challenge	4
Power dynamics and incentives	5
The emerging community	6
What we've learned so far	8
Work to Date	10
Grants	10
Activities	11
Proposed Activities.....	11
Awareness Raising and Advocacy.....	11
Data Support Clinics	12
Resources and Research	13
What does change look like?.....	14
Project Outputs.....	15
Preliminary Outcomes.....	15
Intermediate Outcomes.....	16
Long term objectives	17
People and Partners	18
The engine room team	18
Collaboration	18
Budget	19

Summary

Activists and advocates are using data and technology to address a variety of obstacles and with varying degrees of success. The new opportunities of data also imply new risks, however, and failing to use data and technology responsibly can result in real harm to projects, relationships and beneficiaries. This includes potential harms that have been traditionally addressed within the context of digital security and information security, but also challenges more deeply embedded in information strategies, relationships with local communities and the ways in which information can be adapted and repurposed outside of project activities. As more groups turn towards novel tools to accomplish their aims, there is an increasing awareness about the responsible data risks posed by new technologies, but little understanding about how to mitigate those risks.

Our responsible data work addresses the challenges faced by activists and organizations using data and technology in advocacy. This is related to, but very distinct from the societal questions about ethics, data and security posed by big data and new media. The engine room works directly with advocates and the challenges they face, supporting groups who recognize these challenges, and raising awareness among projects that are facing responsible data challenges but don't know it.

Over the course of 2013 and 2014, the engine room has engaged in a broad range of activities to promote more responsible use of data in advocacy. With funding from Open Society Foundations and considerable support from partners and organizational allies, we have:

- Designed, organized, and led 5 events addressing challenges faced by specific sectors and data uses
- Introduced responsible data topics and ran sessions at more than 25 events
- Provided concrete strategic support to approximately 10 advocacy projects, addressing dynamic responsible data challenges for approximately 10 projects
- Consolidated conceptual and practical frameworks for understanding what responsible data means
- Convened 19 organizations as formal partners in the Responsible Data Forum, for a discussion of how to scale advocacy towards widespread changes in responsible data policy and practice for advocacy and advocacy support communities
- Provided substantive input to academic research initiatives led by Stanford, Tufts, Harvard and others

Using this work as a springboard for more responsive support and further momentum-building, the engine room plans to scale up the Responsible Data Program to address some of the considerable demand we see. We will also be coordinating with other organizations who are tackling responsible data challenges from different angles and with different expertise.

Responsible Data Program Activities will include:

- **Advocacy and Awareness Raising**

Events are critical for building responsible data skills and awareness. Engine room staff have been active in both organizing issue-focused events, and injecting responsible data discourse into others' events. We have seen that this work has large reach and can lead to behavior change in norm-setting groups, turning them into high-impact advocates for responsible data. Live events also give us an opportunity to better understand the challenges organizations are facing, and provide on-the-spot clinic support for at-risk projects.

We will continue and increase our advocacy efforts at events (our own and others), targeting both people working on the front-lines of data-driven advocacy, and groups and organizations that are in position to reach and influence larger numbers of those end users. We will nurture strategic relationships with specific individuals and organizations that can have a scaling and amplification effect for responsible data advocacy within their respective communities, sectors and organizations, including trainers, funders, and key individuals in large, agenda-setting organizations. Identifying and supporting a loose network of responsible data ambassadors will be a key part of this work moving forward. This stream of work will also include producing and disseminating both harm stories, including case studies with illustrative pitfalls, as well as success stories.

- **Data Support Clinics**

Promoting greater awareness (and encouraging organizations to be more responsible) is only part of the process. Providing accessible, relevant, and project-ready resources and support is what makes change possible. In some cases, this is possible in groups, at events or remotely. Often, however, the deep contextual challenges and complexities of responsible data dilemmas necessitate focused support with organizations. These interventions can be the most meaningful for resolving specific challenges and minimizing specific risks. They can also provide the most insight into what types of general resources are the most useful, and can also produce some of the most powerful responsible data ambassadors. The engine room will provide responsive, applied support to directly improve advocacy projects, and identify generalizable approaches and resources for others. Recipients of direct support will be identified through engine room networks and our partner networks. In some cases this will benefit from the engine room's other support activities and its mechanisms for identifying other types of partners. The engine room is already receiving a significant amount of referrals from partners and peers, however, and expects that this will increase with effective advocacy, surfacing a number of individuals and organizations in need of responsible data support. Partner selection will be based on a number of factors, including: the level of risk a partner has, our ability to support them, the likelihood that support outcomes can be used to support other groups, their ability to meet the demands that support will require, whether they are in a position to set norms, and whether they have a demonstrated history of effective work on which support can build.

- **Reusable Approaches and Solutions**

With the help of partners and participants at Responsible Data Forums, the engine room has identified several issue areas and problem archetypes that need attention and support. We have also been working hard to prototype tools and strategies to rapidly and easily mitigate responsible data risks in

applied advocacy. This work will continue, but we are now in a position to strengthen these processes and begin testing resources, in events and through pilot work. The Responsible Data Program will, when appropriate, facilitate or lead the development of responsible data approaches, solutions, and resources towards a functional prototype, and assist advocacy initiatives in piloting these prototypes. We will communicate openly about these activities to feed into awareness raising and to build a community of resource providers and archetypal projects.

Context

Framing the challenge

The use of data and new media has become prominent in contemporary advocacy strategies. As access to mobile and digital technologies becomes more widespread, less expensive and more user friendly, we've seen growing excitement surrounding the idea of these tools and their potential to enhance civic action, political participation and accountability. Small organizations and activist groups campaigning in-country, established human rights groups and multinational development organizations all appear to be adopting novel tools and data-driven strategies with staggering enthusiasm. It is not clear, however, that this increased uptake of tools and strategies is accompanied either by critical thinking about the relationship between digital tools, stakeholders and impact, or the potential risks surrounding the use of data and ICTs.

And the risks are significant. We don't know much about them, because when things go wrong, there are few incentives for projects to share their experiences. But anecdotally, we know that there are a number of ethical, privacy and security harms that can result from well-intentioned data-driven advocacy. Some recent examples:

- A campaign organization working in a highly oppressive context saves the names and personal information of people reporting abuse of public service on a desktop, and keeps backups on a USB. Though similar information is protected in a safe when collected in paper formats, the organization does not have strong protocols for passwords and access to data. One night the organization's offices are burglarized and a USB with these names goes missing. That week several of the people who had reported corporate abuse are arrested.
- A project working to improve public health services, including HIV treatment centers, develops an online map based on individual reports, and uses patient data to map treatments sites. Failing to understand the local context, the map inadvertently "exposes" individuals who are receiving HIV treatment in a country where HIV is associated with homosexuality and homosexuality is strongly taboo.
- A project advocating for the rights of a marginalized community conducts surveys with that community to document their satisfaction and needs as related to public service delivery. Survey results are used in advocacy and the raw data is released under an open license in keeping with the project's aspiration towards transparency. Shortly thereafter, the municipal government launches an infrastructure project that clearly excludes the marginalized community, apparently on the basis of GIS data released by the project.

- An international organization advocating for the rights of a local community determines that mobile survey technology, thanks to its savings of cost and time, would allow for much more regular data collection from that community. They secure funding to increase data collection using mobiles, also by supplying community members with hardware to conduct surveys. The community experiences the increase in data collection, but does not see an increase in advocacy impact (services and accountability do not improve over time). To the contrary, the allocation of novel technologies to specific individuals exacerbates social tensions within the community, and some community members express a feeling of exploitation, provoking general resentment towards the project. The relationship between the project and the community deteriorates, worsening data quality and limiting the project's scope of activity.

These anecdotes represent different kinds of harm. Some are more obvious and damaging than others. Similar stories have circulated in different contexts and with different types of tools. What they have in common is that projects were not able to anticipate the dynamics and exchanges implied by digital information. Whether this is a failure to translate traditional security practices into a digital context (as with the USB drive), or to understand how technology can impact local social relationships (as in the mobile surveys), it's challenging and dangerous despite the fact that most projects are smart, motivated, thoughtful and well intentioned. It's simply that with new technologies come new risks, and most of us aren't familiar with them yet.

Power dynamics and incentives

This lack of familiarity is compounded by a set of perverse incentives in the advocacy and advocacy support communities. The hype cycle of digital activism and ICT4D is widely recognized, as is the trend by which innovative and early adoption of technology takes place among small organizations "at the fringes", while large and established organizations with more significant bureaucracies and network positions tend to be late adopters.

Ironically, these late adopters tend also to be "agenda setters" within the community, and have a significant role in determining which projects and strategies feature in conference programs, publications and funding strategies. As a result, their use of technology draws much attention and becomes associated with the supposed success of a larger organization, without a nuanced discussion of how these tools fit into the organization's strategy. In addition, the resources and priorities of those setting the agenda are often vastly different of those whose work they are influencing. This reinforces the prominence of tech-driven projects (they are simply more visible, because they're online) and the conception of technology as a "game changer", while de-incentivizing discussions of harm and failure. It's not uncommon to hear about projects that are developed solely on the belief that a tech focus will secure funding, or projects being funded simply because they have a significant technology component. That such projects are often designed, implemented and funded by individuals with limited technological familiarity is problematic for their potential impact. Equally problematic is the lack of capacities to recognize and mitigate responsible data challenges. As with any kind of strategy, failure to carefully assess the contextual impacts of tech and data in the project design phase greatly increases the chances that something can go wrong.

Perceived obstacles to efficiency can also provide important disincentives. Among front line advocacy and social justice initiatives, resources are often scarce and workloads heavy. It's not uncommon to find five people doing the job of ten, and the most critical program components underfunded. In such contexts, additional ethical safeguards or privacy checklists can be seen as an impediment to "getting the important work done". Even in large organizations accustomed to significant procedural obstacles, ethical safeguards can be seen as additional hurdles that do not add clear value. This is at least in part due to the lack of awareness about potential harms, but also the esoteric language of privacy and ethics in which these discussions tend to be framed. Few things can make an activist's eyes glaze over as quickly as suggesting a Privacy Impact Assessment, and "it's just too complicated" is a common refrain among project implementation staff considering responsible data challenges for the first time. These challenges are especially acute in high-risk contexts such as humanitarian emergencies, when procedural delays can cost lives.

Lastly, it's worth noting that while there has been a significant interest recently in issues such as the "ethics of data", these conversations tend to be highly conceptual, addressing abstract, rather than practical challenges. While this is better than no conversation at all, and will hopefully develop concepts and frameworks with a practical application, it is important to acknowledge the lack of engagement with front line advocates who face responsible data challenges on a daily basis. There does not seem to be a common language or forum for people discussing these issues in universities and people facing them in applied advocacy. This will make it difficult to deliver solutions and strategies to the people who need them most. These abstract conversations, that will inevitably shape how we understand the ethics of data in a global information economy, are also likely to be much poorer for failing to incorporate advocacy's programmatic realities.

These dynamics are distinct, but closely interrelated. It's safe to assume that they will correct themselves in the long term, as data-driven advocacy becomes less novel and experiences are shared. In the near term, however, advocacy initiatives large and small will continue to turn to data and technology for their potential and promise, often with unrealistic expectations about its costs and benefits. To help them to maximize their impact and to avoid doing harm, it is critically important to broadly promote awareness, strategies and tools for meeting and mitigating responsible data challenges.

The emerging community

The engine room has been working explicitly towards this end since early 2013. We've done so iteratively, trying different approaches to understand how responsible data challenges manifest themselves, and what kinds of responses make sense. In doing so, we've bumped into many individuals and organizations that are eager to understand and address these challenges from a variety of perspectives, disciplines and biases.

Early conversations in the advocacy and advocacy support communities suggest that across institutional and sectoral divides, there are small groups of people keenly aware of responsible data issues and eager to see them addressed.

Interviews with **the donor community** suggested that among funding institutions, these proponents were a decided minority, working within institutional and cultural parameters that were largely not attuned to responsible data issues. Though many of these individuals are already actively discussing these issues, they tend to do so largely within the framework of digital security and data integrity, which de facto excludes many important issues and favors technical solutions over organizational/programmatic responses. Several respondents to these interviews cited institutional and procedural obstacles to proactively helping advocates to mitigate responsible data risks.

Advocates on the other hand, seem to be keenly aware of these issues when framed in terminology that speaks directly to the political and social context in which they work. For advocates publishing land data, concerns that data release will lead to international land speculation or fuel local conflicts are extremely acute. Concerns about “data re-use” or “the ethical consequences of open licensing” are less so. A consequence of this context-specific understanding is that few people from the advocacy community are speaking amongst themselves about these issues, and there is little knowledge about either risks or responses in comparable advocacy contexts. Conversations and support to advocacy groups also indicated that responsible data rhetoric was most compelling when it focused on efficiency and avoiding harm to others (activists are often more cavalier with their own privacy and safety), but that they found it extremely difficult to operationalize responsible data strategies. Even when advocates recognize risks and challenges, they simply don’t know where to start.

There was perhaps greatest awareness and engagement among **advocacy support groups** that provide training and strategic support to projects. Like the donor community, individuals from these groups tend to conceptualize responsible data in terms of digital security (which provides a familiar conceptual framework for considering risk) but struggled to consider these challenges and appropriate responses at a more strategic level of project design. Individuals from this group were also less likely to discuss these issues amongst themselves, unless in the context of an existing initiative (usually a digital security initiative).

Early investigations into resources in the **research community** (such as the Ethical Review Board mechanism) also uncovered significant interest among academics. This community is struggling to understand how a traditional approach to ethics and human subjects should be applied to research on, and using, digital and mobile technologies. Simultaneously, several researchers with advocacy-relevant issue focus are studying the use of new technologies by advocacy groups, or using new technologies to study advocacy groups. These conversations have to-date been muddled, including many perspectives and mixing methodological and “ethical” debates. And while there are a great number of them, they appear to be largely disconnected. Similar workshops and symposiums have been held monthly throughout 2014. Though the resources in this community (intellectual, methodological, strategic) are tremendous, it is not yet clear if these discussions can be made directly relevant to applied advocacy.

We also acknowledge the communities on the periphery of advocacy work, especially **service providers**. There are many organizations and practitioners who are working for social change by providing direct services (medical, psychosocial support, etc.) to vulnerable populations and survivors of

abuses. These efforts often collect and store information, and their activities overlap significantly with advocacy efforts, though they do not themselves conduct advocacy. With these actors, we intend to explore the challenges they face and the lessons they have learned in the responsible use of data, and identify ways we can support their work.

Across these different groups, there is an increasing awareness that data and technology-driven advocacy can cause harm, and that some kind of response or exploration is warranted. During the course of interviews, direct support and convening events, the engine room has found that the rhetoric of responsible data is very useful for framing and facilitating conversations across these groups. That the benefits of data and technology are accompanied by responsibilities, and that people using data have some kind of due diligence obligation, seems to be an idea that everyone can get behind. We have found it a very useful frame for bringing together groups that have previously not been able to discuss these issues: scrappy activists and corporate social responsibility lawyers, tenured professors and crypto fundamentalists.

We are also convinced that responsible data rhetoric has been useful for mapping and facilitating the emergence of a community that, if nurtured and provided with support to address the challenges it finds most important, can grow and thrive. We see this community developing now in a number of ad hoc and disparate events, joined by the idea that there is a responsibility that accompanies using data, and that using technology should do no harm. From its kick off in March 2014 to Budapest Forum in September the same year, the Responsible Data Forum has accrued 18 institutional partners. We believe this is the tip of the iceberg, and that this work will continue to generate interest and momentum, presenting opportunities for greater advocacy and support over time.

What we've learned so far

Our research, support and discussions so far have led us to some preliminary conclusions about responsible data:

Lesson #1: Framing: a conceptual framework is useful, but not for everyone

We believe that it makes sense to think about disparate issues like participatory ethics and protecting anonymity together, because in an applied sense, they involve the same kinds of assessments and processes for advocacy initiatives. For example, we have found that for analysis and comparison, it makes sense to consider four overarching issues in which responsible data challenges emerge (Identification, Consent, Agency and Data Re-use) and which we can usefully differentiate between harm to individuals, to communities and to advocacy efforts themselves. These distinctions (more on our conceptual framework at <https://www.theengineroom.org/responsible-data-a-conceptual-framework/>) are useful for mapping out the types of risks and applicability of responses across diverse contexts and groups. This can be an important tool for helping funders assess their portfolios and internal capacities, or to help training organizations review their curriculum.

These categories aren't at all useful for individuals or groups facing specific challenges, however. The advocacy initiative that discovers its SMS reports are being surveilled and aggregated by the government's security forces isn't concerned with these categories, nor is the project that is worried

about how their research on slum dwellers might misrepresent gender dynamics and have an adverse effect on municipal budget allocation.

When considering how to support front line advocates, we need to think in terms that resonate. This means understanding the types of **decision points** that recur across most projects, where there are opportunities in program cycles and hectic workdays to assess and respond to risks. Anticipating the pressures and limited resources with which most advocacy initiatives operate is essential and needs to be built into tools and support strategies. This approach needs to draw on the actual experiences of real advocacy initiatives, which are not yet sufficiently understood in any operational sense within the support community.

Lesson #2: Evidence: we need better and more balanced real life examples

Among donors and the support community, there is a palpable desire for case studies and harm stories. The dissatisfaction with generalities, anecdotes and alarmism is profound and widespread. This is driven largely by an assumption that having specific and detailed actual harm stories will be a useful advocacy tool, especially with leadership of large organizations, and will help to put responsible data on the agenda. In addition to supporting advocacy, reliable evidence on how responsible data challenges are experienced in the field, if comprehensive and structured, would also provide invaluable information for fine-tuning support methods.

Previous efforts to collect such evidence have been hampered by a lack of resources and the reluctance of groups and individuals to share attributable stories. We believe that in pursuing this evidence, it will be important to have a clear research method, work within trusted advocacy networks, and to collect a large number of cases, according to a common data structure and with a limited degree of detail.

It will also be important to balance harm stories with success stories. While the “scare tactic” is intuitively appealing for many seeking greater awareness, experiences from the realm of digital security suggest that this approach must be followed immediately by solutions and support if it is to change behavior. Examples of successfully mitigating risks are equally important.

Lesson #3: Networks: we need to scale outreach and support

The need for responsible data discussion and support is tremendous and nearly universal. A single project or collaboration could never hope to resolve it. We believe that in order to foster change in the long term, advocacy needs to also strategically target individuals within key networks and organizations. By identifying and supporting these individuals, we believe they can play a critical role in developing capacities and promoting the responsible data agenda within their communities. We believe that this is inevitably the way that change occurs (albeit slowly) in organizations like the UNDP and in national civil society contexts. We believe that a consciously considered strategy to groom responsible data ambassadors can help to facilitate that change.

Lesson #4: Support: is there anything it can't do?

The incredible variation across advocacy contexts means that some important realities simply cannot be captured by comparative research or superficial engagement. Deep engagement and support to

initiatives is the only way to truly understand the micro-level pressures, power dynamics and other factors which determine how and if an advocacy initiative can identify and mitigate responsible data challenges. It is important that broad and comparative research on responsible data challenges be used in conjunction with (and reality checked by) actual response and support to specific responsible data challenges. This support should be well documented.

Simultaneously, despite the great need for advocacy and awareness raising around responsible data challenges, in some cases, direct support - through advice or workshops that address an organization's particular challenges, or through concrete tools such as checklists or draft policies - may be the most effective advocacy mechanisms. Being able to reference success stories about mitigating responsible data risks for small organizations shows that it is possible, and will help to combat the "it's just too complicated" malaise. Providing and applying concrete tools (such as checklists and draft policies), even when imperfect, can open the door to responsible data discussions, and provide organizations large and small with an opportunity to review and consider their practices.

Lesson #5: Language and ownership: planting seeds and letting them grow

We have learned that targeted language is key for raising and implementing a responsible data agenda. This means using appropriate terminology with different actors, but also using the language of responsibility and harm avoidance as a means to gather disparate sectors, disciplines and perspectives.

We believe that such an approach is most effective if it is not branded. We have been careful not to over-associate the engine room's brand with Responsible Data Forums in 2014, and believe this is part of the reason that we have seen such an uptick in the use of the term "responsible data"¹. We see this as positive, since it is a rhetorical frame that encourages a sense of obligation and a focus on solutions. To encourage increased use of this rhetorical frame and engagement with a wider group of actors, we will be exploring the potential of other groups to drive the responsible data agenda, organizing their own forums, producing their own tools and leading their own discussions, as determined by the needs and priorities as they experience them. We expect this will require significant network engagement, followed by capacity development and light touch input and support from us, primarily in the form of sharing what has worked and has not worked during the Responsible Data Forum's short career.

Work to Date

Grants

So far we have received the following grants:

- \$48,830 from Open Society Foundations to implement the Responsible Data Forum.
- \$39,130 from HIVOS to develop a Responsible Data Toolkit for development programming

1 This rhetorical shift is modest but significant, in trainers' online communications and project framings, researchers project descriptions and conference titles. Without claiming attribution, we believe that our work has contributed to a general environment in which rhetorical changes like these and others are more likely.

- \$100,000 in core funding from Oak Foundation, which has enabled a significant amount of self-funded responsible data work

Activities

- Research on Donors and Responsible Data (self-funded): conducted interviews with 25 representatives of funders on how they support responsible data practices among grantees, in the grant making relationship and in their internal data management. Report forthcoming.
- Responsible data support and consultancies for organizations facing specific problems: to date we have supported 6 organizations through regular meetings or focused clinics to help them identify and mitigate specific responsible challenges on an ad hoc and self-funded basis. We have also been contracted by 3 organizations to provide similar support on a consultancy basis.
- Responsible Data Forums: We have organized and executed 4 Responsible Data Forums, including a launch event and events on responsible hosting, development programming, and accessing private sector data. Two additional Forums are scheduled (Responsible Data Sprint in Budapest, Sept 30-Oct 1, and Consent and Crowdsourcing in Nairobi, Oct 28), and two additional Forums are being planned with partners.
- Online conversations and gatherings. We have organized and co-organized several online conversations and hangouts to discuss particular responsible data issues, both publicly and with closed groups.
- Tools, outputs and promotional materials: Responsible Data Forums have generated over a dozen tool prototypes with enough support among partners that we can reasonably expect them to be completed and piloted. We have prepared a GitHub (<https://github.com/the-engine-room/>) site for sharing these, and some have already been applied by organizations. We have also produced a number of promotional materials, including the www.shouldmywebsitehavessl.com.

Proposed Activities

Awareness Raising and Advocacy

The engine room will advocate, debate, and act on responsible data issues in a variety of sectors in which responsible data challenges are most pressing.

The engine room will focus advocacy efforts on broad convenings such as Stanford's recent Ethics of Data conference, communities with ripe responsible data challenges such as the human rights documentation community, highly trafficked intersections of data technology and advocacy such as the open data and civic hacking spheres, and support communities particularly positioned to reach many organizations such as international issue networks, funders, or large, multilateral development institutions.

The engine room will:

- **Identify and attend events to promote responsible data practices**

We will continue to organize and facilitate event sessions, blog issue-specific responsible data challenges and disseminate posts around the events, live tweet with a responsible data angle, post up small tables to triage responsible data problems of participants, actively participate in others' sessions to promote responsible data practices, find responsible data evangelists to support, encourage attendance at responsible data forums, and design new forums with enthusiastic and sector-appropriate partners.

- **Produce compelling and informative materials about responsible data**

We will continue to develop analytical and informational pieces on responsible data for a variety of audiences, particularly targeting advocacy support groups, funders and advocacy project designers. We will draft blog posts, articles, and other analytical reports to promote responsible data practices by bringing harm stories, challenges, and success stories to the forefront of discussion about data and technology use in advocacy.

- **Disseminate and seed responsible data discussion and information**

Convenings and events provide clear opportunities to share information about responsible data, but events can be ephemeral. We will continue to collect and share information that we (and others) have produced about responsible data challenges and successes in mailing lists, on blogs, in partnerships with other organizations, and in targeted outreach work.

- **Promote strategic planning in support communities for responsible data approaches**

Support communities obviously have a large effect on practices and possibilities for responsible data. As such, we will work closely with partners that provide support, such as Open Knowledge's School of Data, HURIDOCs or DataKind, to ensure that they have space and a strategic partner to carefully consider responsible data challenges so that they can provide holistic support to organizations they work with.

Awareness raising and advocacy work is a key part of the responsible data program. It is a way to listen to the needs of different communities, encourage better practices, develop common approaches, and recruit ambassadors into the responsible data community.

We anticipate that these activities will result in an increase of demand for direct support that is tailored to particular, complex project problems, and a need for development and maintenance of the responsible data community. To address these demands, the Responsible Data Program will devote staff time and energy to community management to support these processes, and to Data Support Clinics devoted to directly supporting advocacy projects and communities as need arises.

Data Support Clinics

The engine room will identify 10 projects that are facing concrete responsible data challenges, and will support them to confront those challenges. Direct support to projects means providing strategic consultation on responsible data issues, working with groups to weigh considerations, supporting concrete project adjustments, and matchmaking projects with resources and expertise to take these adjustments further should more specialized support be needed.

These support clinics will move along a common pipeline, but will be implemented in different ways depending on circumstance and need. The pipeline will include:

- **Initial fact-finding and consultation**
The engine room will lead structured discussions and exploration of the projects' unique responsible data challenges.
- **Direct support in developing strategies to address responsible data challenges**
Using the consultative fact-finding process, the engine room will make concrete suggestions and help the project design a strategy to address problems without undermining project goals.
- **Matchmaking partners with appropriate resources**
In many instances, the strategy to address responsible data challenges will require specialized expertise - be it technical, methodological, or other types of expertise - and the engine room will support the partner in finding and acquiring those resources.
- **Documentation of challenges and response that can be adapted and reused**
Recognizing that there will be patterns and common challenges, the engine room will document direct support to seed generalizable resources that can be applied to other projects facing similar challenges.

Methods for implementing this stream of work can include in-person workshops between the projects receiving support, the engine room, and appropriate experts. Projects that receive this direct support will also have the opportunity to engage with their peers who are facing similar challenges or working towards similar goals. This peer-to-peer exchange (online and/or in-person events) will be facilitated by the engine room and will feed into advocacy activities outlined above. Given the costs of an in-person convening, and the growing demand for responsible data support, these activities can also be managed remotely when appropriate.

Data support clinics will strengthen our understanding of how responsible data challenge manifest themselves and will directly inform the development of resources and research design.

Resources and Research

Over the next two years, the engine room will lead the development of at least 15 responsible data resources. We will also lead the production of four in-depth analytical reports on the responsible data challenges in specific issue areas and/or sectors.

Resources will take the shape of in-depth how-tos, checklists, strategic frameworks, and other tools. These resources will be designed to promote responsible data practices within organizations, to consolidate outputs from forums and direct support clinics, and to facilitate the responsible deployment of specific advocacy tasks.

Resources will be developed responsively and the engine room will manage the process through the following steps:

- Collaboratively identify resources that can make a difference in responsible data practices within advocacy organizations

- Develop clear scenarios for how the resource might be used, and by whom, to identify a feasible scope for the resource
- Identify appropriate format, depth, and dissemination strategy
- Produce and disseminate resource

Research areas will be chosen based on initial scoping processes at events and convenings and the interest of issue-area partners who can collaborate with the engine room on the frame of research questions and the grounding of responsible data challenges. Research outputs will be used to activate interest in responsible data issues in particular sectors, and ground future support in real-world problems that are surfaced during the research process.

Research will be developed responsively and the engine room will manage the process through the following steps:

- Identify issue area or sector to explore further²
- Work with partners to identify research questions and illustrative examples within the sector
- Collect further evidence of both challenges and successes in the sector
- Produce analysis of responsible data landscape in the sector
- Use the research produced in advocacy and outreach in that sector to promote targeted approaches to pressing challenges

What does change look like?

Responsible data needs are urgent, widespread and not yet well understood. We're in early days and don't presume to offer a complete solution. However, we are capable of influencing the awareness and discussion that is emerging around these issues, and helping to frame them in terms of responsibility and agency will have important consequences for people working at the front lines of data-driven advocacy. We also believe that our focus on these front liners will inject a healthy degree of reality into an otherwise overly abstract discussion, hopefully shortening the time before that discussion can be applied to real life responsible data challenges. Another critical aspect of this approach is that it works to impact widespread awareness and discussion by providing direct support and meeting specific needs. In this way, the Responsible Data Program seeks to support change at both the macro and micro levels.

Below we outline the project outputs and objectives towards which we will be working, identifying appropriate indicators for success, and exploring the causality and assumptions that link them.

2 Issue areas might include human rights or land advocacy, sectors might include humanitarian interventions or open government. These focus areas will be defined by the expressed needs within a specific group or discussion.

Project Outputs

The Responsible Data Project will produce outputs continuously under each of the three activity streams. Advocacy and awareness raising activities will promote responsible data issues with new audiences and increase engagement with individuals and organizations already participating in the responsible data discussion. Clinics and direct support will meet the specific needs of individual organizations. Research and Resources will produce thematic reports, as well as tools and documented strategies for use by advocacy initiatives. Together, these outputs represent the raw materials that we will be using to reach a tipping point, after which responsible data challenges will become an inevitable and integral part of advocacy projects with a data and technology component.

Indicators

- One third of Responsible Data Forum partners increase their engagement, potentially by leading conversations, developing resources or organizing responsible data activities independent of engine room collaboration
- 2000 individuals reached directly and in-person through responsible data advocacy and awareness raising
- 10 data clinics conducted to provide direct support to advocacy initiatives facing specific responsible data challenges
- 4 research-based reports published on responsible issues or domain areas
- Of the 15 resources produced by the program (such as checklists or draft policies) over the two years, 4 will be directly piloted with advocacy organizations

Preliminary Outcomes

We believe that the above activities and outputs will contribute directly to a number of preliminary outcomes. Specifically, we believe that compelling advocacy and awareness-raising, together with the effective dissemination of resources and research will strengthen the degree to which practical responsible data **challenges are discussed and understood** in global discussions about advocacy and that this will lead to a more informed discussion around ethics, privacy and security issues surrounding data-driven advocacy. Generally, we hope that this will result in **more regular framing in terms of responsibilities and avoiding harm**, and that global discussions will be more relevant to, and cognizant of front-line advocacy realities.

We also believe that advocacy, direct support and research and resource outputs will significantly raise awareness among advocacy initiatives, producing an **increased demand for support**, both expressed to the responsible data program and allies, but also directly to funders and other training and support organizations.

Lastly, we believe that direct support and resource production will **strengthen responsible data practices** of those advocacy initiatives we reach. Though it will be difficult to measure harms that are averted, we believe that advocates' own assessments of the support they have received (whether it added value, whether they feel more comfortable in how they handle data) will be an appropriate measure of this value added.

Assumptions

- Relevant stakeholders at the global, regional and national registers continue to be receptive to the rhetoric of responsible data
- The engine room will be able to identify and engage with appropriate initiatives facing responsible data challenges, and that those initiatives will want, and be in a position to receive help in addressing those challenges
- A better understanding of how responsible data challenges manifest themselves will also suggest appropriate responses, which are actionable for advocacy initiatives and support organizations

Indicators

- Increased use of “responsible data” rhetoric in articles, blog posts and policy documents relevant to data-driven advocacy (list of online references, including blog posts and training curricula produced by prominent members of the advocacy support community /or international networks, funders and multilateral development organizations)
- Actors and discussions currently engaged in abstract debates about ethics and data begin to reference that actual challenges faced by advocates and data users (list of online references)
- Increase in the number of direct requests for responsible data support received by the engine room from advocacy initiatives and support organizations
- Recipients of direct support report stronger capacities to develop responsible data strategies in program design and implementation, to identify and assess risks before harm occurs, and to effectively respond to challenges.

Intermediate Outcomes

As awareness and familiarity with responsible data harms become more common, we believe that this will impact the ways in which advocacy support organizations operate. We hope to see an increase in explicit responsible data support to advocacy initiatives by other actors. We believe and hope that an increase in general awareness, availability of tools, and subtle changes in institutional cultures will slowly begin to influence the mandates and practices of organizations that have traditionally provided digital security support. As this happens, we hope to see these groups demonstrating that when appropriate, they are capable of providing support for more strategic and amorphous problems surrounding data and technology use, addressing issues that have to do with information security and risk assessments more broadly construed, including methodological questions and ethical considerations of representation, consent and agency.

We also believe that increased awareness and successful support to advocacy initiatives will result in increased capacities among advocates to provide support to others, and we anticipate early signs of peer-to-peer support networks for responsible data.

Assumptions

- Digital security support organizations such as Internews and Tactical Technology Collective will be able to successfully integrate more general responsible data support and strategy into end-user digital security tool trainings
- Training and capacity development in providing responsible data support will be possible
- Support organizations will be able to secure funding and staff expertise for support that does not focus on specific tools or digital security, but takes a more critical and three-dimensional approach to the responsible use of data across a project cycle, including the impact of data sharing or publication

Indicators

- Training and support groups traditionally providing digital security support to advocacy initiatives begin providing more holistic responsible data support, addressing strategic consideration across the project cycle, including the impacts of data strategies on local stakeholders or counterparts and the consequences of data sharing or publication
- Funding organizations traditionally providing digital security support to advocacy initiatives begin providing support for addressing responsible data challenges
- Responsible data harms are more frequently reported by advocacy initiatives to peers and support organizations
- Peer-to-peer responsible data support between advocacy initiatives is more common
- Data support for projects more holistically includes responsible data support

Long term objectives

Eventually, we hope that the above activities and outcomes will create an environment for data-driven advocacy in which the principle of “do no harm” is both a presumptive norm and immediately actionable. This, in turn, we hope will lead to the more responsible use of data by advocacy initiatives, and less incidents of unanticipated harm caused by the use of data and technology in advocacy. The lack of information about how and how often such harms currently occur makes it impossible to determine a change in the frequency of harm over time. However, we believe that perceptions of relevant stakeholders are a sound proxy measure for this.

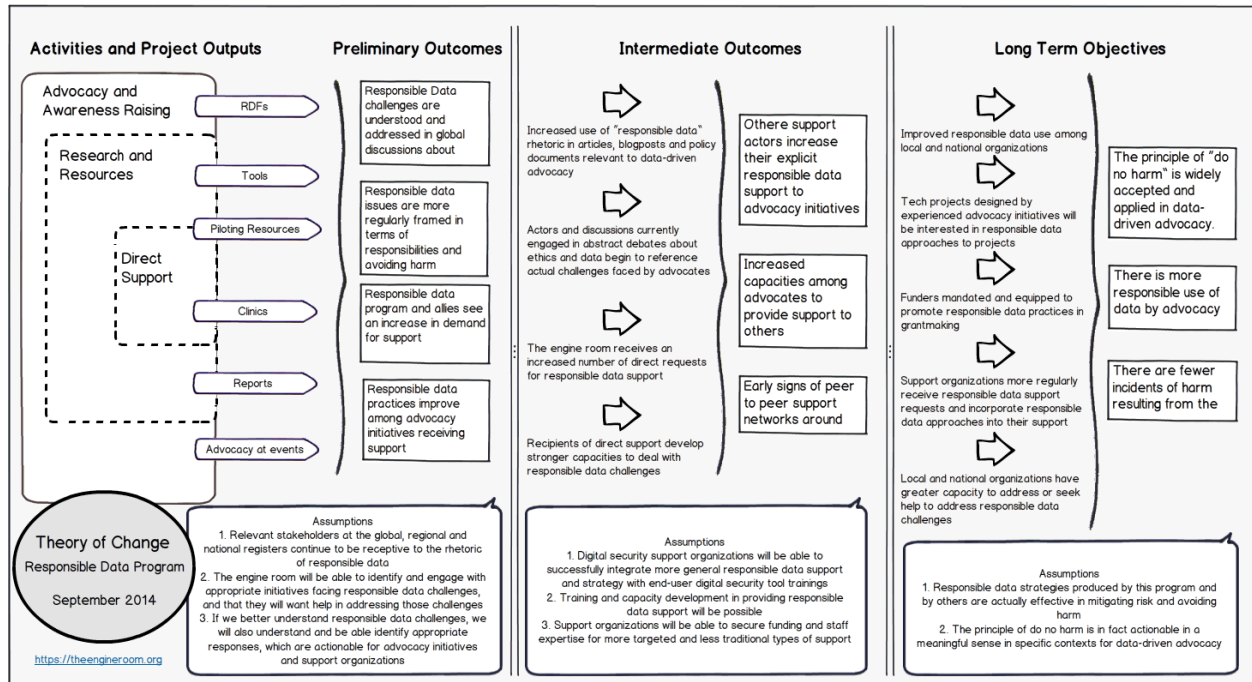
Assumptions

- Responsible data strategies produced by this program and by others are actually effective in mitigating risk and avoiding harm
- The principle of do no harm is in fact actionable in a meaningful sense in specific contexts for data-driven advocacy

Indicators

- Improved responsible data use among local and national organizations
- Support organizations get more requests for support on using data responsibly
- Tech projects designed by experienced advocacy initiatives will be interested in responsible data approaches to projects

- Funders will be keenly interested in and equipped to promote responsible data practices in proposal submission processes
- Support organizations will incorporate responsible data approaches into their support methodologies
- Local and national organizations will have greater capacity to address or seek help to address responsible data challenges



People and Partners

The engine room team

Over the next two years, the engine room will be building a responsible data team within the organization. This team will include **program management** responsible for supporting the strategic development of program (15%), an overall **project manager** (100%), a **community manager** responsible for supporting advocacy and awareness raising (25%), a **project deputy** responsible for supporting research and resource development and event management (20%), and **editorial support**. Some of these positions will be filled by engine room staff and peers who have been working with us for the past 18 months to map and wrestle the responsible data landscape. Others will be new hires. All team members will work together across the project's activity streams, and this work will be guided by a program director, who is currently managing the engine room's responsible data work.

Collaboration

Collaboration has been a core component of responsible data work done to date. Understanding how responsible data challenges manifest themselves in different contexts has required us to work across a variety of institutional and sectoral domains, and partners have been key to making this possible.

Sometimes this collaboration takes shape around loose conversations and a common interest, but more structured collaboration has also been key.

The Responsible Data Forum has established formal partnerships with 14 thought leaders in this field.³ Several of these partners have been active publicly and behind the scenes, working actively to advance our thinking, identify opportunities to leverage resources and to identify the areas in which responsible data advocacy is most needed. Several partners are also well positioned to scale responsible data advocacy within their respective networks and sectors.

In its next phase, the Responsible Data Program will be working to develop these relationships and to establish new ones. Partnerships will serve a variety of different functions--reaching specific audiences, understanding specific sectors, ensuring that program work speaks to the needs of specific constituencies--but will generally be premised on a shared commitment to promote responsible data practices in advocacy.

In addition to these core partners, we expect to engage with a number of actors who are eager to work on responsible data issues more independently. Where appropriate, we plan to provide them with strategies and experience to help them do so. In some instances this might involve light touch support to help them organize responsible data forums, develop strategies or provide support independently. As with the Responsible Data Forum, this work will emphasize the concept and rhetoric of responsible data and the role of the responsible data community, while de-emphasizing the engine room's role and organizational brand.

Budget

A detailed budget is attached, totaling \$356,500 to cover staff and activity expenses over a 2- year period. We are seeking support from multiple donors, and believe that a strong mix of donors will enhance the impact of the program.

³ RDF partners include: [Amnesty International](#), [Aspiration](#), [Berkman Centre for Internet and Society](#), [Central European University](#), [Data & Society Research Institute](#), [DataKind](#), [Greenhost](#), [iHub Research](#), [HIVOS](#), [HURIDOCS](#), [Kurante](#), [MIT Center for Civic Media](#), [Open Knowledge](#), [Privacy International](#), [UN Global Pulse](#), and [Ushahidi](#).