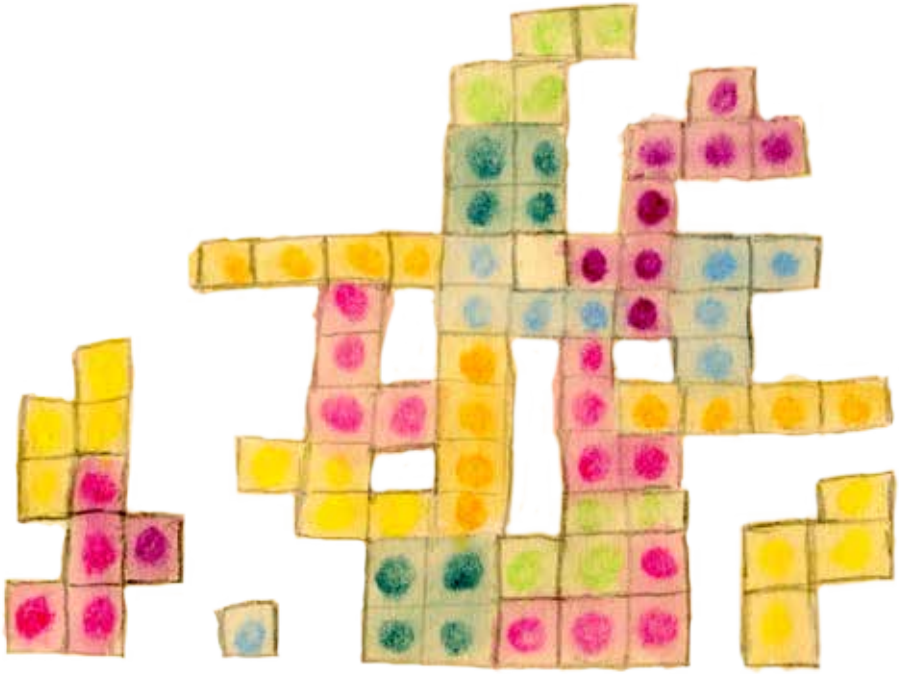
WeGov

the engine room

measuring impact on-the-go

a users' guide for monitoring tech and accountability programming



MEASURING IMPACT ON THE GO:
A USERS' GUIDE FOR MONITORING TECH
AND ACCOUNTABILITY PROGRAMMING

Content and design by the engine room.
https://theengineroom.org

This Users' Guide is the product of a collaboration
between WeGov and the engine room.

ISBN Pending.

First published, May 2014.

Table of Contents

Introduction

This is a practical guide for busy organizations with limited resources and competing priorities. It is neither comprehensive nor universally applicable, but aims to give concrete and actionable guidance, to help projects understand how monitoring can improve outcomes and avoid waste, and to set their own agendas for measurement.

The information provided in this guide becomes more detailed the further you read, and it's not necessary to use every section. If monitoring is already a priority for your project, and you know that real-time monitoring is the most appropriate approach, you can skip directly to the appropriate steps in the middle section. If measurement is an issue that your project is just beginning to explore, you might want to read the introductory pages that explain the costs and benefits of monitoring on-the-go. If you are looking for resources for specific strategies or methodologies, you might want to scroll directly to the resources section.

Here's a basic overview of how the guide is laid out:

The next four pages contain overviews of **concrete steps and rules of thumb** for developing a monitoring framework.

Immediately after this, the first two sections describe **how the guide works**, and provide an overview of the measurement problem: why it's hard, and what's unique about monitoring tech and accountability initiatives. This introduction is intended to help projects and project managers decide whether real-time monitoring will give them the information they need, and what the costs will be.

The following section is the core of the guide, and provides practical guidance for **developing a monitoring framework in 17 steps**. These steps are grouped into four phases (deciding to measure, mapping resources, developing a framework and implementation), and are designed to be applied consecutively, but can also be approached on an *a la carte* basis. Specific requirements and strategies are presented at the beginning of each of the four phases.

The guide closes with a **collection of resources** for further exploration, providing deep context and an explanation of methodologies, tools and strategies.

Overview: steps for designing a framework

Here are the basic steps to take in designing a framework for monitoring on the go. Some might not be necessary, and some can be conducted in parallel, but the sequence is generally important. Further detail, links and examples are included in the section on designing a framework.

Deciding to measure

Conversations with management and relevant staff

1 To measure or not to measure?

- Ensure proper support from management
- Identify available resources
- Discuss potential costs and benefits

2 What will the framework be for, and who needs to be involved?

- Identify key objectives for the framework, and ensure that appropriate staff are involved in developing the framework in order to meet those objectives

3 Is monitoring on-the-go the right framework?

- Consider the importance of timeliness, expertise and budgets
- Review different approaches to monitoring (ex-ante, ex-post and developmental)

Mapping resources

Coordinated parallel work by relevant focal points, or half day workshops

4 Map organizational resources

- Discuss potential costs, including staff time, hardware, software, consultancies and events

5 Map automatically generated data

- Review project tools, including social media, mailing lists, websites, digital or mobile-supported surveys or service delivery, mobile campaigns, project management tools, or security software

6 Map other data collected by the organization

- Map all data collection activities taking place within the organization
- Assign relevant focla points to chart these in a collaborative spreadsheet including individual questions or indicators
- Collaboratively identify comparable indicators across data collection efforts

7 Map external data

- Identify relevant data collected by peers, officials or international actors
- Consider risks to using this data, including sustainability

Developing the framework

In one or more workshops with all relevant staff, and ideally project stakeholders

8 Project design

- Review or develop the project's logic model or theory of change
- Identify causality and assumptions that need to be tested, and which suggest monitoring indicators

9 Understand what you'll do with it

- Review common uses for monitoring data and identify the most pressing monitoring needs for the organization
- Work with the staff responsible for these needs to develop specific use cases for monitoring data

10 Review methods

- Review commonly used monitoring methods, including participatory methods, rapid assessment methods, narrative methods, statistical methods, and others
- Balance the potential advantages of these methods, as presented in this guide, against their cost in terms of time, expertise and financial resources

11 Review data sources

- Review data sources available to the project, including usage statistics, submitted data and reports, field notes, surveys, administrative data or stakeholder communication platforms
- Match relevant sources against indicators and use cases from steps 8 and 9

12 Review indicators

- Brainstorm indicators that align with the program logic and use cases from steps 8 and 9
- Refine the indicators so that they are SMART and actionable
- Identify the different types of indicators (qual/quant, de jure/de facto, output/process), and which indicator types can be productively mixed

13 Prioritize

- Consider the methods, data sources, and indicators suggested in steps 10-12
- Agree on the three priority use cases, from those surfaced in step 9
- Identify the minimum collection of indicators to satisfy those use cases
- Identify the methods and data sources necessary to produce those indicators
- Identify the resources required to produce that monitoring data
- Determine if those resources are available, on the basis of the mapping conducted in step 4, and adjust as necessary

14 Assign timelines and roles

- Identify when monitoring data will be required by the three priority use cases
- Develop timelines that build on existing activities, that strengthen staff engagement, that create opportunities for discussion and review
- Assign specific roles and responsibilities for all necessary activities
- Ensure that timelines have enough time that review of data can realistically influence both project activities, and subsequent monitoring

Rolling it out

15 Initiate an iterative process

- Review initial monitoring data in a broad and inclusive forum that actively solicits the input of all staff and stakeholders that will use the data, or who have helped to provide the data
- Evaluate the appropriateness of indicators, methods and processes
- Learn from the review process, and design subsequent review processes with enough time and meaningful engagement to influence both project activities and subsequent monitoring

16 Analysis

- Consider challenges posed by attribution, changes over time, aggregation, and conflicting data

17 Communicating and learning

- Communicate monitoring results and their consequences openly and often
- Create spaces for meaningful feedback
- Establish clear strategies for sharing results with staff, management, donors and stakeholders

Overview: rules of thumb

Here are some basic rules of thumb for designing a monitoring framework that is efficient, actionable and meaningful.

Tech and Accountability Initiatives

Any project, large or small, that is using technology in some meaningful way to work towards improved governance, transparency or accountability.

Whether this involves using a website to raise awareness, developing an open data portal for parliament, using mobile phones to conduct surveys on service delivery, or mapping instances of political violence online, we use this phrase to refer to any project where technology supports project activities in some meaningful way.

Uses for monitoring data

- Monitoring against program goals and objectives (Is the program achieving what it intended?)
- Monitoring program outputs in key areas (What has the program delivered?)
- Monitoring short to intermediate term outcomes (What is the program beginning to achieve in key result areas?)
- Monitoring changes against a baseline (What changes have occurred over time?)
- Financial monitoring (How have funds and resources been used?)
- Monitoring management and administrative arrangements and processes (What processes have been used during program implementation?)
- Monitoring key areas such as stakeholder relationships (What are the views of stakeholders of the progress of the program, against the stated program logic?)

(adapted from Core Concepts in Developing Monitoring and Evaluation Frameworks)

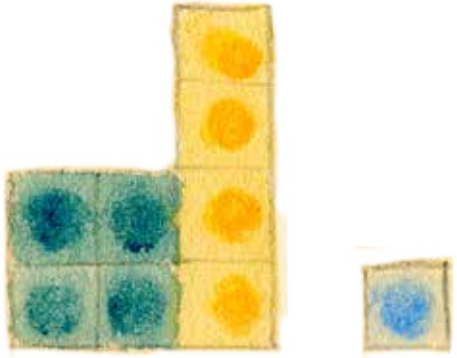
Good indicators are:

1. Mixed (there are several different types in a framework)
2. SMART (specific, measurable, achievable, relevant and time-bound)
3. Actionable (it's easy to know how to respond when they are collected and reviewed as part of a monitoring process)

Making measurement meaningful

- Whenever possible, data collection should be incorporated into other project activities. This conserves resources, but also helps to keep data collection aligned with larger project goals, increasing the chance that monitoring can be useful to project activities, and that monitoring does not get ignored or forgotten.
- Have frank conversations with people supporting the process, to ensure that there's a shared understanding of how collecting and managing data will strengthen their work.
- Share monitoring outputs widely and openly throughout the project, organization and with stakeholders, in language that is easily accessible, but also as raw data if possible. Doing so can increase the chances of getting meaningful and useful feedback.
- Create opportunities for reflection and discussion of monitoring results within organization. Something as simple as an informal lunch to discuss results can surface unexpected insights and additional resources.
- Look for ways that monitoring can strengthen project activities, for example by framing data collection as consultations to strengthen relationships with stakeholders, or by sharing data with government partners to build trust and strengthen coordination.
- Make sure that data review processes have enough time to significantly influence how the next round of data collection is conducted, and that the people reviewing the data have a real influence on how project strategies are designed and adjusted.

using this guide



What's it for?

This guide presents a series of steps that **tech and accountability initiatives** can use to develop monitoring frameworks. It will work as a roadmap for a comprehensive process that involves an entire organization, from conceptualization to implementation and learning. It will also work in bits and pieces, to refine monitoring activities, or to prioritize specific activities when time and resources are short.

What follows is not comprehensive, it is an introduction. Steps in the process are presented as a series of questions, considerations, examples and links. The hard work of making a monitoring framework feasible and useful will always fall on the project itself. What this guide offers, however, is an entry point for non-experts to start thinking about measurement issues and how different monitoring strategies can help their work. It is designed for people who already have too much to do, and who don't want to become experts in measurement, but want their programs to be as efficient, impactful and well funded as possible. We believe that this process will help projects to achieve better outputs, and to better manage conversations with donors about reporting and support.

What do we mean by "tech"

This guide is targeted towards projects that use some kind of technology to work towards better governance, transparency or accountability. It may also be useful for other kinds of projects, but is structured to take advantage of those circumstances, and anticipate those challenges that are most common among tech and accountability initiatives.

This guide is **not for using technology to measure** project activities and results, **but for monitoring the role of technology in projects**.[1] This is fuzzy distinction, because technology will often measure itself (more in Step 10), but tech tools for measurement are not the focus of this guide.

This guide will be most useful to people designing or managing projects, but aims to use plain language and simplify technical frameworks as much as possible, in order to facilitate discussion among teams. Links and references are provided throughout the guide to help readers find additional information and resources.

1 For guidance on using technology to conduct M&E, see the forthcoming report from the Rockefeller Foundation, authored by Linda Raftree, as well as the Clear Net Initiative. For a discussion of tools and strategies for tech tools to monitor the impact of research and information advocacy, see Nick Scott's blogpost: http://onthinktanks.org/2012/01/06/monitoring-evaluating-research-communications-digital-tools/.

What is measurement on-the-go?

This guide will focus on **strategies for monitoring the activities and consequences of tech-supported projects in real time** and on-the-go (formative monitoring).[2] This means monitoring on an on-going basis, so that projects can adapt to changes and surprises during project implementation. This is especially useful for technology-driven projects where the actual impact and consequences of project activities are not certain. While this guide makes some reference to strategies for measurement at the end of projects, the focus is on-going and regular data collection and review.

Measurement on-the-go also implies flexibility and adaptation in measurement strategies. Since time and resources tend to be scarce, this guide will emphasize strategies that take advantage of technology already being used, and integrating measurement and documentation into existing workflows and activities. It will provide projects with considerations to manage the tension between measurement strategies that are feasible and non-disruptive, and methods that are rigorous, methodologically sound, and which can produce comparable data.

Read what you need

This isn't an M&E manual or a treatise on evaluation, and there is no need for program managers to become measurement experts. This is a springboard for targeted and applied learning.

The guide begins with some background and arguments that may be helpful for initiating conversations about measurement within projects. The bulk of the guide is then devoted to a step-by-step description of how to build a framework for monitoring on-the-go. These steps are designed to be taken sequentially, and each one builds on the ones that have preceded it. You may not have time and resources to engage in a comprehensive process, however. Individual phases and steps can also applied independently in such cases. An overview of the steps, and the key questions and themes for each, is presented after this section.

2 Formative monitoring is monitoring to inform decision-making and adaptive management on an on-going basis. In the world of measurement, this is distinct from "summative" monitoring and evaluation systems, which aim to summarize what has worked and what has not, and which are primarily tools for accountability and review after a project or project phase is complete. Real-time evaluation are also relevant here. Commonly used in humanitarian projects, this method has recently been applied to budget accountability initiatives, with noteworthy results. See http://oxfamblogs.org/fp2p/watching-the-ups-and-downs-of-accountability-work-four-new-real-time-studies/ for a description

why bother?

Costs and Benefits

Measurement demands resources, there's no way around it. Choosing to measure is an allocation of time, money and expertise that could be applied to other project activities. And even when monitoring is prioritized, it may not be easy, and may not always produce clear results. So it's important to enter the world of measurement with a clear understanding of what monitoring can help a project to achieve, and what the costs will be.

Challenges to measuring the impact of investments in technology:

- There usually isn't any baseline against which to assess new improvements.
- The process of planning for IT upgrades or installation—discussing how things are done, what the bottlenecks are within an organization—can, in and of itself, free up sticking points and improve efficiency, even without the implementation of new IT systems.
- New technology can be so intimidating for staff that an organization actually becomes less efficient in the first few months following a technology upgrade—so an evaluation done too soon might result in a "false negative", reflecting inefficiencies that would be resolved with time.
- Sometimes new technologies aren't used at all (for example, when adequate training isn't available or when organizational culture presents barriers), ossifying an anti-technology mindset and causing long- term reductions in efficiency and impact. Correctly attributing improvements to proximate causes can be a challenge: at what point does the technology stop having an impact, and other factors (that may be out of the hands of the service provider) come into play?

- Evaluating the Impact of Information Technology, Allison H. Fine, INNOVATION NETWORK

There are several ways in which monitoring on-the-go can lead to more efficient processes and better project outputs.

1. **Save time and money, get better outputs.** Identifying which activities contribute to project objectives can allow projects to focus on the activities that contribute most to results. You don't have to wait until the end of a project to see these trends and make smart decisions. The point of monitoring on-the-go is to set up smart systems and adapt before it's too late.
2. **Anticipate challenges and opportunities.** Regularly collected monitoring data is likely to suggest relationships between contextual factors and projects' successes and failures. Understanding these links can help projects to use monitoring data to anticipate important external forces, whether it's market trends, social media activity,

government pronouncements or citizen reports. A well-designed system for monitoring on-the-go, potentially combining internal monitoring processes with monitoring of external processes, can flag the changes in context that mean something.

3. **Improve project design.** Real-time monitoring can support adaptation in project design along the way. Establishing pivot points in project cycles, and having robust data on what's working and what's not, can provide priceless opportunities to revisit theories of change, adjust management structures, or refine tactics. The close correspondences between a strong monitoring framework and strong program design are also important in this regard. Developing and implementing a monitoring framework can help to refine or develop a theory of change, identifying those activities that will best contribute to achieving project goals. Lastly, what is learned in monitoring one project can also feed directly into smarter design of other projects. Know what works and what doesn't, and measure in a way that gives an overview over all projects' activities.
4. **Make reporting less painful.** Donor reporting is often a challenge, pulling time and resources away from other important activities. Often this is because measurement only happens when reports need to be written, and one person will have to take the time to collect indicators and build a measurement system from scratch. Having a system in place to monitor on-the-go will put hard numbers and narrative data at the ready, already collected and set to feed into documents.
5. **Make a better case to donors.** A good monitoring framework will show what works and what doesn't, including essential, but less sexy, inputs that might not be prioritized otherwise. Hard monitoring data can make for a stronger argument about why grants should include support to boring or hard-to-fund essentials like travel, hardware and training.[3]
6. **Produce evidence for advocacy.** At the end of the day, much of the data projects collect to monitor how they are doing can also be useful for backing up campaigning and lobbying efforts. A rigorous framework for data collection and analysis can enable powerful *dual use data*.

Setting aside time & resources

The benefits of monitoring will of course vary dramatically from project to project, and will also depend on how much effort is invested to develop and implement a monitoring framework. And make no mistake, developing and implementing a monitoring system may require investments of time, energy, reputation and money. Hopefully, they will save *more* resources in the long run.

Here are some issues that will almost necessarily demand resources, and which should be considered before beginning.

3 For a history of evaluation in development as an exercise of power (from logframes to RTCs) see this framing paper Eyben (2013): http://bigpushforward.net/wp-content/uploads/2011/01/Uncovering-the-Politics-of-Evidence-and-Results-by-Rosalind-Eyben.pdf .

1. **Getting everybody on board.** Monitoring and evaluation frameworks are most effective when they have broad input, buy-in and use throughout an organization. Securing broad and early buy-in can help to secure necessary time and resources for monitoring. It can also help to ensure that monitoring frameworks measure things that actually matter, and generate information that will actually be useful. Buy-in from front-line program staff is just as important as buy-in from management. Leaving individuals or groups to struggle with monitoring in isolation can sabotage their efforts, leading them to produce systems and data that aren't useful, while the process is likely to require far more staff hours than is necessary. In some situations it can be just important to secure buy-in from funders and community members. This is especially so when they are intended to use or benefit from monitoring outputs, or when they can help to improve the quality and sustainability of monitoring processes. Getting everyone on-board will take effort, however. Staff meetings and workshops can be effective ways to collect inputs and validate frameworks, but won't completely eliminate the time cost of bringing everyone up to speed on monitoring plans.
2. **Getting the data.** This is the most obvious resource demand, but not treating monitoring as an isolated activity can help reduce that demand. Firstly, data collection for monitoring on-the-go should be aligned with ongoing project activities as closely as possible. Any information projects are already collecting (complaints about service delivery, government pronouncements, legal decisions, data on foreign direct investment), should be examined at the start, to see if it can inform a smart monitoring strategy, or what other data could be easily collected at the same time. Also consider who else is collecting data that might be relevant for monitoring your project, either as contextual data or impact data. There are a number of international actors regularly collecting comparative national data on governance and accountability,[4] and national civil society may also be active.
3. **Asking more of your community.** It will be clear to most project managers that asking for community input can have a significant relational or political cost (there are only so many times you can ask people for their time and energy), so these types of demands need to be carefully considered and prioritized. Simultaneously, any project that aims to benefit communities will want to ask those communities about how they experience project outputs or activities. To balance these competing pressures when designing data collection, projects should look to streamline monitoring data collection with existing activities, and to reinforce project brand

4 UNDP offers a comprehensive (if outdated) overview Source Guide to Global Indicators (http://gaportal.org/global-initiatives/source-guide-to-global-indicators) and the latest resources (http://gaportal.org/resources/891). A number of novel indices focusing on technology and data have also emerged in recent years, including the Open Data Index (https://index.okfn.org/) and the Web Index (http://thewebindex.org/).

and incentives with every outreach. As a point of departure, this means prioritizing whatever data can be captured incidentally (by counting social media postings or citizen reports, for example), but such data will rarely be sufficient. When projects need to directly ask communities for information, efforts should be made to build this onto other project activities, so that it constitutes an extra question in a survey or an extra box in an online form, rather than a new survey or form. When a project needs to conduct specific and independent outreach (like focus groups or independent surveys) look for ways that this can also provide something of value for the community and be sure that there are clear incentives to provide information. In some cases, the simple opportunity to air grievances may be sufficient, and including useful information or updates in outreach can also help. Sometimes added incentives such as mobile credit or social capital through group membership can also be considered (though external incentives need also to be considered for their impact on community relations as well as their impact on data validity, about which more below).

4. **Regular review** is what makes monitoring data useful on-the-go. Whether this is conducted weekly or bi-annually, by a core team or in a staff meeting, depends entirely on the context and what the data is going to be used for. When designing a monitoring framework, it will be important to understand how data will be used (program design? outreach? key lobbying opportunities? more in steps 2 & 9) and to time review accordingly. It is also important that review activities are conducted as much as possible by the individuals and teams that will actually use the information. Except for the rare instances where people using monitoring information are data scientists, this implies having a system in place that can output meaningful data with minimal analysis.

Leaving individuals or groups to struggle with monitoring in isolation can sabotage their efforts, leading them to produce systems and data that aren't useful, while the process is likely to require far more staff hours than is necessary.

Designing such a system at the beginning of setting up an M&E framework can be a significant, but critical, investment. Generally, systems in which teams are directly involved in picking indicators and designing data collection processes will tend to require less analysis to make output meaningful to those teams. Review processes that are lightweight and occur more regularly may also be easier to make use of, especially for large teams. For some projects, simple mechanisms might make the most sense (weekly emails with key indicators can keep large teams up to date, while also serving a team-building function, and preparing data for larger reviewing and reporting processes, while requiring very little time and energy to prepare). For others, structured interval processes make more sense, with analysis and presentation to management or all project-staff. When designing a framework, it's important not only to anticipate the costs of review systems, but also how this process can be integrated into existing workflows, to maximize efficiency.

Why tech, why now?

This guide targets voice and accountability projects using technology, but that's not because monitoring strategies are different for this type of project. They're not. In fact this type of project tends to combine some of the most challenging issues in project monitoring (more in the next section).

What *is* distinct about technology-supported projects, however, is that that **technology tends to generate metrics of its own use**. This has several implications. For projects, it means that there is often a wealth of data that may be useful, or can be made useful for monitoring. This also tends to raise the expectations of measurement experts not familiar with technology-supported programming. Many expect that this relative abundance of hard data should make M&E easy for tech-supported projects. That's not the case either. Though technology tends to produce a significant amount of data, it's not often easy to parse or make sense of on its own. Generally, making sense of the log data that technology automatically generates, and the other hard data that it can generate, requires a careful and strategic approach, which lies beyond the technical and technological familiarity of many small projects using cheap and easily accessible technologies to pursue accountability agendas. The number of visits to a project website won't likely tell you much about what you're doing right unless it's thoughtfully combined with other kinds of information.

Many expect that this relative abundance of hard data should make M&E easy for tech-supported projects.

We are, however, working in an opportune moment for the tech and accountability community to begin strengthening its capacities to measure, monitor and evaluate its own work.

Consider the following four trends:

1. There is increasingly widespread ***expert dissatisfaction*** with contemporary measurement frameworks and practices, especially for governance and accountability programming,[5] even as there is an increase in measurement among NGOs.[6]
2. International austerity has led to increased pressure on bilateral and multilateral donors, often translating into renewed calls to demonstrate "results" of international development aid. This pressure to document results has a trickle down effect, and is experienced by many small projects and recipients of grants as ***pressure to measure***.[7]
3. The ***increased use of technology*** in governance and accountability projects is accompanied by a **growing imperative to better understand** if, and under what conditions, technology can contribute to social and political change.[8]
4. The use of technology in programming tends to generate metrics of its own use. When ***tech measures itself***, tools already in use by projects can often greatly facilitate the collection of additional data.

These factors present tech and accountability projects with a unique opportunity to set the agenda for how they measure their work, and to ensure that measurement serves their own specific needs.

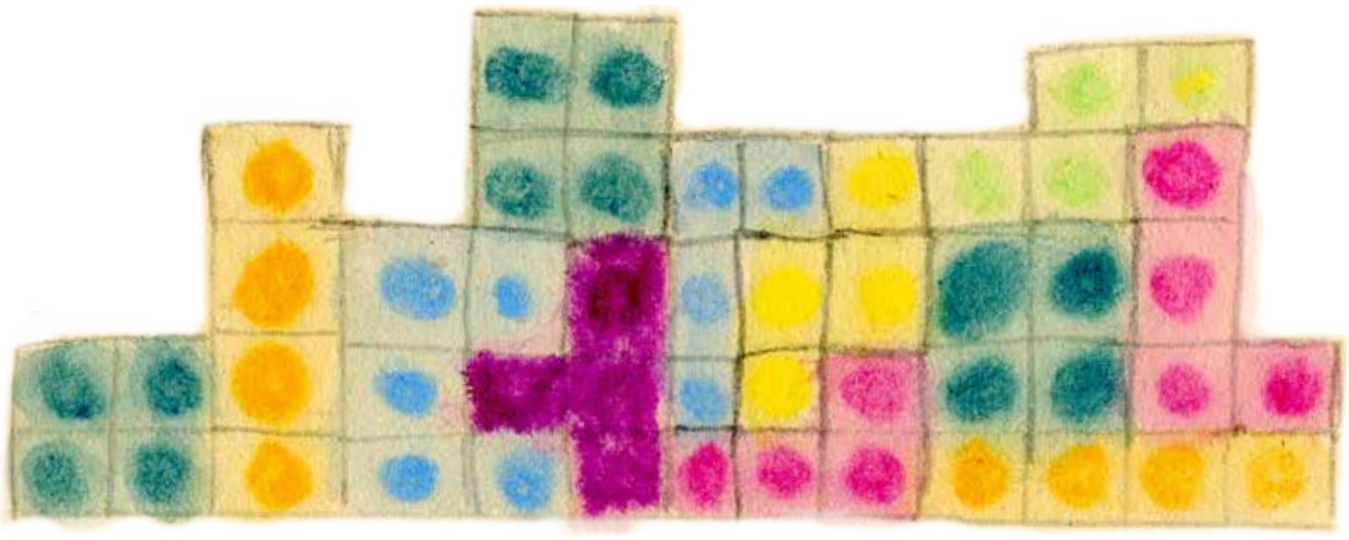
5 See, for example Starling (2010) - http://www.tandfonline.com/doi/pdf/10.1080/09614520903564215 , Shutt (2011) - http://governance.care2share.wikispaces.net/file/view/Governance+M%26E+Literature+Review+Final+6.10.11.pdf , Holland et al. (2009) http://www.oecd.org/derec/unitedkingdom/44463612.pdf , Shutt & McGee (2013) - http://www.ids.ac.uk/publication/improving-the-evaluability-of-ingo-empowerment-and-accountability-programmes, McGee & Gaventa (2013) - http://www.ids.ac.uk/publication/the-impact-and-effectiveness-of-transparency-and-accountability-initiatives .

6 Survey Morariu et al. (2012) - http://www.innonet.org/client_docs/innonet-state-of-evaluation-2012.pdf , showing the state of evaluation for nonprofit organizations. Of the 535 organizations surveyed, 90% evaluated their work. For a more general resource on assessing and learning for social change, see Irene Guijt's review of critical readings (2008) - http://bigpushforward.net/wp-content/uploads/2011/05/guijt-assessing-social-change.pdf .

7 See this Hivos blogpost (http://www.hivos.net/Hivos-Knowledge-Programme/Themes/Civil-Society-Building/News/The-Singer-not-the-Song) on donor pressure and survey Whitty (2013) - http://bigpushforward.net/wp-content/uploads/2011/01/Experiences-of-the-Results-Agenda-by-Brendan-Whitty.pdf suggesting that Results agenda is significantly affecting the way NGOs do work.

8 Organizations like MAVC (http://www.makingallvoicescount.org/) and open development technology alliance (http://odta.net/) aim to improve accountability through technology-enabled citizen engagement, but also to understand the limitations of technology (http://www.makingallvoicescount.org/blog/what-are-the-limits-of-transparency-and-technology/).

designing a framework step by step



Deciding to measure

Who's involved: Primarily management — it will also be helpful to engage focal points for learning, evaluation and reporting.

What you need: Not much. Some basic background information (this guide) and an organizational desire to do better work and make better decisions.

How to do it: In small meetings or bilateral conversations between management and relevant staff. The important thing is that time and priorities are set to allow for a meaningful conversation with management. A rubber stamp from management at this point will lead to roadblocks and wasted energies down the road.

Steps:

1. To measure or not to measure?
2. What will the framework be for and who needs to be involved?
3. Is monitoring on-the-go the right framework?

1. To measure or not to measure?

The decision to begin monitoring project activities may be prompted by external factors such as donor demands or internal pressures to improve performance or learning. Either way, it's important to ensure that the decision is informed, even if it's already been made. Ensure that appropriate decision-makers talk through the various arguments for measurement and related considerations about cost and resources. It's important that these considerations are flagged early on, to avoid delays and surprises along the way.

2. What will the framework be for and who needs to be involved?

Monitoring can serve multiple ends (identifying waste, supporting funding arguments, anticipating crises), and it's important to map these out early and in detail. Suppose a monitoring framework aims primarily to understand what components of a citizen reporting project are contributing most to the project outcome of awareness raising. By measuring this link, the project plans on adjusting the survey as it's rolled out in new geographical areas. Then it's important to have the methods person responsible for designing the survey questions, or the interface for web or mobile reporting in the conversation. If monitoring is meant to inform better donor reporting, then the person writing the reports needs to be involved. This may seem obvious, but it's often overlooked at early stages, and nothing is worse than investing time and resources to produce an indicator that can't be used.

Bill and Melinda Gates Foundation: Actionable Measurement Guidelines

Exhibit 3: what is actionable measurement? Applying actionable measurement means that teams:

1. **Consider measurement needs during strategy development and review:** Identify assumptions that should be tested and information gaps that can be filled with measurement from the beginning of strategy development and at critical stages such as strategy review.
2. **Prioritize intended audiences:** There are many potential audiences for results measurement, including foundation leadership and teams, grantees and ultimate beneficiaries, donors, national and international policymakers, and practitioners. It is important to identify and prioritize the intended audience(s) and their need for data.
3. **Do not privilege a particular evaluation design or method:** Technical decisions about evaluation design and data collection methods are driven by purpose and weighed against the feasibility of different approaches.
4. **Focus on a limited set of clearly articulated questions:** We cannot and should not measure everything. Organizing results measurement around decision points and a set of clearly articulated questions helps ensure we have the information we need, when we need it.
5. **Align results across strategy, initiatives, and grants:** Measurement can help us to confirm and adjust the alignment between our work, including the grants we make, and the objectives laid out in our initiatives and strategies.
6. **Obtain information needed to inform decisions in a timely way:** In planning results measurement, we pay particular attention to the specific information needed to make decisions and to when that information needs to be available.
7. **Allow time for reflection and the development of insight:** Data and information alone do not tell us what to do. Making properly informed decisions entails building in time to interpret and reflect upon the products of results measurement, and then applying the insight gained.

Source: https://docs.gatesfoundation.org/Documents/guide-to-actionable-measurement.pdf

3. Is monitoring on-the-go the right framework?

Ensure that key team members have a chance to discuss what monitoring information is for and when they need it. To what degree is real time information important? Is it most appropriate for that information to be collected by the project or by an external expert, and what impacts will this have on budgets? On credibility?

Ex-Ante, or Formative Evaluations (measurement on-the-go), are often applied to pilot projects, or other projects where there is a significant degree of unknown factors or causal relationships, to provide an opportunity to reconsider and adjust project strategies during implementation (*adaptive management*[1]).[2]

Ex-Post, or Summative Evaluation, takes place after projects are completed, and may be conducted by project teams, but are often conducted by external consultants. Summative evaluations generally aim for a holistic understanding of what context and process factors contributed to project successes or failures. They often target donors and researchers.

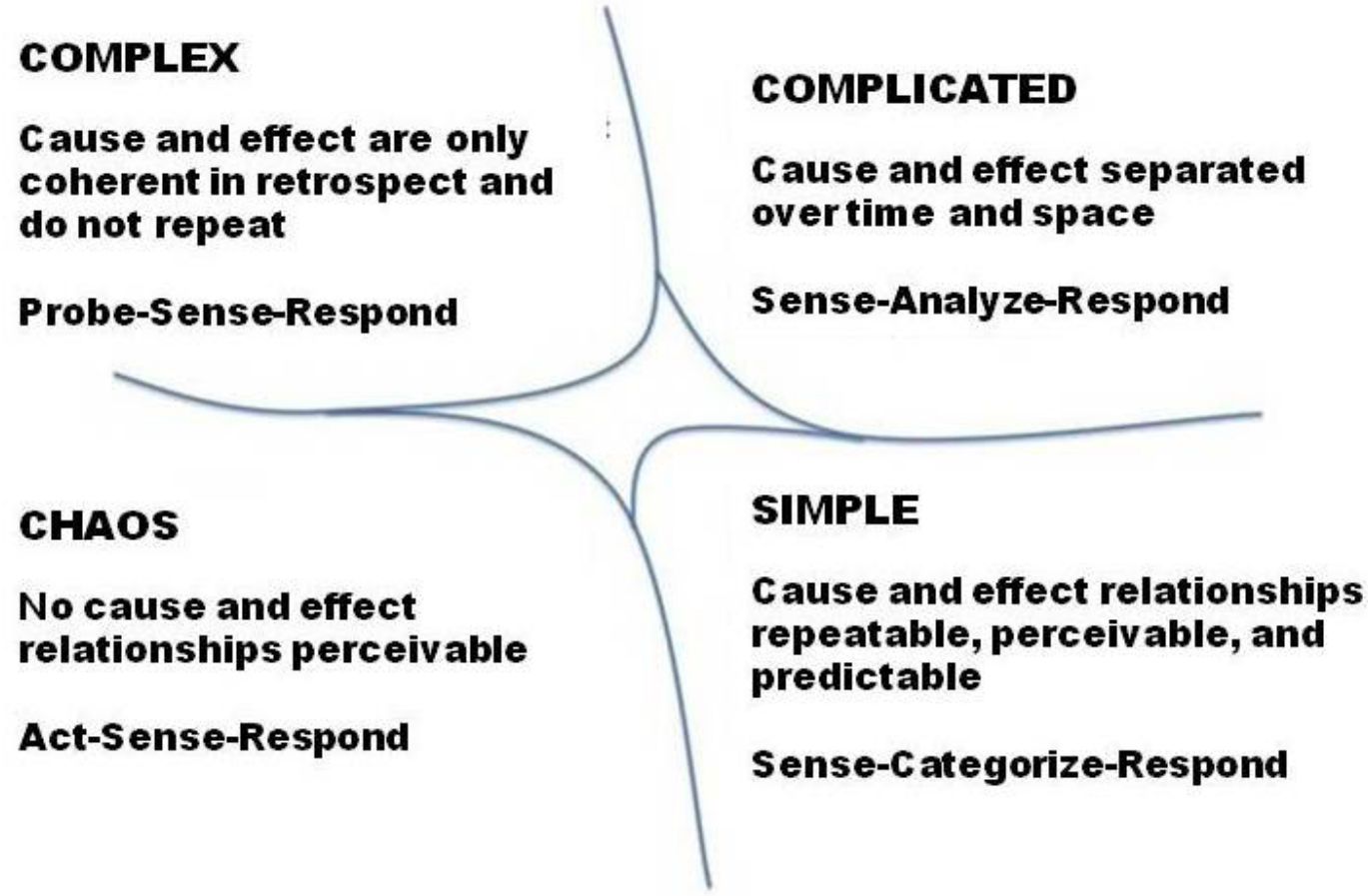
Developmental evaluation is a method that combines the adaptive and iterative aspects of measurement on-the-go, with the assessment dynamics of an external evaluation. It is designed to track and support program innovation in situations of high complexity.[3]

1 Adaptive management (https://en.wikipedia.org/wiki/Adaptive_management) is an iterative process of decision making with the aim to lessen uncertainty via system monitoring.

2 There is, however, increasing recognition that these dynamics are equally useful for any project operating in contexts dominated by uncertain or complicated causal relationships. For an excellent discusson of how "real-time evaluation" methods can be usefully applied to transparency and accountbaility initatives, see Duncan Greens review of four evaluations supported by the International Budget Partnership, at http://oxfamblogs.org/fp2p/watching-the-ups-and-downs-of-accountability-work-four-new-real-time-studies/.

3 A Developmental Evaluation Primer Gamble (2008) - http://tamarackcommunity.ca/downloads/vc/Developmental_Evaluation_Primer.pdf - suggests using a developmental evaluation when working in situations of high complexity or when working on early stage social innovations.

It may not always be easy to determine what kind of framework is most appropriate, and if measurement is the best solution in the first place. One tool for answering that big question is the Cynefin Framework, a decision making tool that identifies appropriate methods of investigation, according to the kinds of contexts projects operate in.[9]



This representation of the Cynefin framework, though taken from an entirely different context, captures the key areas and their correlated responses. Adapted by Philip J. Palin at http://www.hlswatch.com/2010/07/29/tara-the-bodhisattva-of-risk-management/ .

9 For more information on the Cynefin Framework and its five domains, Complex, Complicated, Chaotic, Simple and Disorder, see Wikipedia's description of the framework (https://en.wikipedia.org/wiki/Cynefin#Description_of_the_framework).

Mapping resources

Who's involved: Management, project management, resource managers and budget managers as appropriate, focal points for different data collection activities, people with deep technical understanding of the technology the project is using.

What you need: Management support, in order to motivate multiple focal points to actively supply detailed information.

How to do it: This phase can be completed in a half day workshop if it's possible to convene all the relevant people. Steps 4, 5 and 7 may be conducted in parallel and in groups during the workshop. Step 6 will itself require several groups, each led by focal points for different data collection activities in the organization. Active participation of these focal points is essential if this step is to be taken in workshop form, as is strong facilitation. This workshop can also into the design workshop (steps 8-14), if appropriate.

This mapping work can also be conducted as separate independent processes, but this will require strong and active coordination, which can be demanding. The coordinator will also need to collate all the mapped information into a useable format for the design workshop, which can be quite demanding (this is especially true for the spreadsheet on organizational data collection).

Regardless of whether this phase is conducted independently or in a workshop form, all mapping outputs need to be collected and presented in clear formats before beginning the design phase (Step 8).

Steps:

4. Map organizational resources
5. Map automatically generated data
6. Map other data collection in the organization
7. Map external data

This is the time to take a quick overview of all the resources and information that will support a lean and useful monitoring framework. This includes gaining an overview of the staff and financial resources available in the organization and project. Most importantly, it involves understanding what data is already (or can easily be made) available.

Mapping what data is available or already being collected is an important early step for saving time and resources, and to make sure that the monitoring framework is as efficient as it can be. It can also be an opportunity to synchronize data collection efforts across the organization, and identify opportunities for collaboration or data sharing with peers.[10] That said, it can sometimes be difficult to get information about what is happening in different projects and activities across an organization. Keep an eye on how demanding these mapping activities will be, and ensure that you have strong buy-in ahead of time, to ensure that focal points have the time and mandate to collect the information you need. If it turns out that this process is too demanding, and is threatening to slow down the entire exercise, be prepared to skip the problem steps or cut them short.

There are different approaches for mapping different types of data, but for each of them, the key aim is to identify the following:

- Data that speaks directly to project outputs or the project's causal model
- Data that has been collected long enough to provide the project's *baseline*
- Data collection processes that can provide additional information

4. Map organizational resources

As mentioned in the section on costs and benefits, setting up a measurement framework implies a number of costs:

- Staff time spent planning, as well as collecting and reviewing information
- Hardware or software for collecting, managing and analyzing information
- Consultancies and technical support
- Events, such as consultations or focus groups

It's too early to map specific expenditures at this stage. But it can be smart to flag potential costs, and map what resources are available to meet them.

10 Often, surveys, consultations and project evaluations are planned in silos and with limited resources, producing similar but incompatible indicators (f.e. number of times a service was used in the last six months, and whether a service was used in the last month). Harmonizing these indicators can allow for more powerful diagnostics across the organization, and make data management a lot easier. Streamlining data collection processes is primarily about increasing efficiency and minimizing costs, and implies finding opportunities in one data collection process, to collect information for another.

Questions to Ask

1. Is there budget allocated to measurement?
2. Do staff have capacities and tools to manage large amounts of data should that be desirable?
3. Does the organization have partners or peers with expertise or tools that might be accessible? (If cost is an inhibiting factor for conversations at this stage, it might be worth reviewing some of the cost mitigation strategies laid out in the section on Costs and Benefits.)

5. Map automatically generated data

One of the unique aspects of monitoring technology-supported projects is that technology can automatically generate metrics of its own use. This can go far beyond tracking the number of "likes" or "shares" in online campaigns, to provide detailed and nuanced running information about who is using the technology and how. What this entails depends entirely on what technology the project is using. A few starting points are listed below, but this mapping exercise should be sure to include people who know how the technology being used works, at a fairly technical level. This might be project staff, but it might also involve including an external expert or consultant.

Below is a list of common technology tools and platforms, and some of the useful information they might be producing. Use this list as an entry point to map what relevant data is being automatically generated, or could be.

Social Media for campaigning

Metadata provides a tremendous amount of information about the people on social networks that are following or engaging in campaigns. Not all of this is obvious, but each tweet, for example, contains information about the person tweeting, their profile, their location, language users, etc.[11] Similar data exists for other social networks, though methods and ease of extracting it vary.[12] Network analysis can also provide useful information about the relationships between people following or participating in a campaign and might be able to surface changes in how outreach is performing over time.[13]

11 For a full review of possible metadata, see Slaw's "The Anatomy of a Tweet: Metadata on Twitter" (http://www.slaw.ca/2011/11/17/the-anatomy-of-a-tweet-metadata-on-twitter/).

12 Popular and accessible platforms for analyzing this kind of content including Splunk, which offers special licences for non-profits (http://www.splunk.com/view/SP-CAAAG6E).

13 The Robert Wood Johnson Foundation has also produced a useful overview of tools to monitor social media activity, see Social Media Monitoring Tools.(http://www.rwjf.org/content/dam/files/rwjf-web-files/Resources/1/SMMonitoring_2013.pdf)

- Facebook: Likes and group memberships are easy starting points. Comments can also be useful qualitative data when submitted to natural language processing or other processes. Facebook data outside of a profile or group that you manage is notoriously difficult to access, but for tracking and automatically collecting information posted to you, there are some useful built in tools, including Graph.[14]
- Twitter: Important data points include @ replies (to track the characteristics of engagement), retweets (to assess the characteristics and the growth of networks engaging with your issue) and hashtag references (to understand how your issue is being referenced outside of your immediate networks, in conjunction with what other issues, and by whom).[15]
- Other social media. The dynamics and data available on other social media platforms varies widely. Some useful dynamics to consider include changes in the size and characteristics of communities, how users interact with your issue area in terms of both frequency and quality.

"Not your grandfather's document review, key informant interviews or probability surveys."

Automatic time stamps are added to Skype chats, sms messages, email and web postings, and this is an invaluable tool for evaluators when trying to triangulate information. On the other hand, there are significant challenges of analyzing so many small and often disparate threads of conversations. Reinforced and adaptive models of program logic and improved techniques for intelligent analysis of these data, including multi-media data will become increasingly important for evaluators of this generation.

(from Ushahidi Haiti evaluation: https://sites.google.com/site/haitiushahidieval/news/finalreportindependentevaluationoftheushahidihaitiproject)

Mailing Lists

Analytics functionality is available on most mailing list platforms, and can be added to others, to provide information about who is opening what and when. Understanding what parts of emails attract the most attention for which types of readers can be useful information to improve mailings.

14 For an introduction to some of these, see B2C's Facebook's Hidden Market Research Tools (http://www.business2community.com/facebook/facebooks-hidden-market-research-tools-0659436#!F9XUD).

15 For a preliminary introduction to tools for collecting and analyzing tweets, see Karr (2013) - http://www.marketingtechblog.com/hashtag-research-tools/ . For an Excel-based method that does not even require a twitter account, see Conducting a Hashtag Network Search of Twitter (https://blogs.k-state.edu/it-news/2013/04/06/the-nodexl-series-conducting-a-hashtag-search-of-twitter-part-5/). See also the tutorial How to Get Data from Twitter from the Meta Activism Project (http://www.meta-activism.org/2014/05/how-to-get-data-from-twitter/).

Interactive websites and citizen reporting websites

Users and content providers will provide information both automatically and actively. This can be used to judge the quality of that engagement and how it changes.

- User metadata associated with IP addresses can provide information about users' locations. Time stamps on contributions can reveal patterns and factors that encourage engagement. Metadata detailing users' software and hardware specifications can also highlight access to information and infrastructure, and serve as proxies for socioeconomic class.
- Comments will often contain the same information described above, and can also be subjected to natural language processing, to understand sentiments and contents for different types of users.
- Short surveys and evaluation questions can be built into interactive websites, so that users provide information for monitoring as part of the engagement process. This can mean including a question in an online form, or asking an additional question when users post content.

Tech-supported data collection (enumerated surveys and consultations)

Data collection efforts will of course produce useful information for measuring their own efficiency. Simple metrics on the number of completed questionnaires will be useful for this, and substantive results (identifying members of target groups, for example) can also inform monitoring efforts. In addition consider:

- Metadata on how information is collected (frequency, geographic information, and hardware/operating systems used by respondents), can all reveal how well data collection is reaching specific target groups.
- The cost of reporting and contributing (SMS fees versus fees at cybercafes or free online submission) can also be tracked over time, to show levels of engagement.

Mobile campaigning, education and engagement

Numbers of subscriptions are the most obvious point for tracking activities such as mass SMS communications, SMS surveys and **interactive voice response**. Subscriptions may also include additional information about subscribers, that aligns with project goals (socioeconomic groups, political affiliation). This information can then be tracked according to engagement patterns, including:

- response rates and call in rates,
- disengagement points in surveys and question patterns, and
- multiple uses/responses.

Digital or mobile service delivery

This is the area in which automated metrics are perhaps most obviously useful for monitoring frameworks. Platforms and projects will vary dramatically, so it's important to look closely at what data is being generated or can be generated by specific platforms, but in general, it can be useful to begin with the following:

- response rates and response times,
- geographic data about requests for service, and
- success rates for requests.

Project management software or hardware

The administrative use of technology for accountability projects may play a key background role in those projects, whether for coordinating project management through online platforms like Basecamp,[16] managing contacts through a dedicated Client Relationship Management platform (CRM), or establishing infrastructure for conducting teleconferences and virtual meetings between remote teams. To understand the use patters for such tools, and the potential for measuring their contributions to project outputs, it may be worth reviewing log times that show:

- uptime for phones and devices,
- location of field staff (relative to costs that would have been incurred without the technology), and
- staff response times to HQ requests.

Security or circumvention software

This type of technology may be a key focus of project activities (for example, in promoting access to information within repressive environments) or a secondary mechanism designed to protect and enhance project activities. Either way, such software will provide information on how it is used, including:

- logs,
- surveys, and
- response times and rates for intermittent check-ins over secure channels.

16 See Basecamp website (https://basecamp.com).

6. Map other data collection in the organization

Challenges to mapping data collection

- Projects might not have clear and accessible documentation of their data collection processes
- Focal points might not be able to dedicate the time necessary to map what data they collect
- A harmonization exercise (and even the mapping exercise suggested below) can quickly devolve into a protracted discussion about appropriate indicators, stalling progress.

Benefits to mapping data collection

- Can reveal data collection redundancies and wasted resources
- May reveal data that can be used for monitoring
- The mapping exercise itself can help teams within an organization to share experience and strategies
- Allows for harmonizing indicators across data collection activities, which in turn can enable powerful diagnostics across organizational activities, and make data management easier.

This step is about getting an overview of what other data is being actively collected by the project, and by other activities across the organization. Most organizations don't have this, as data collection (surveys for advocacy, evidence on human rights abuses, mapping of community needs and priorities) gets planned and executed according to specific needs and under pressure. Mapping out these processes and what data is collected can pose a number of challenges, but also a number of benefits.

One the next pages, there is a detailed description of how to go about this mapping exercise. Review this carefully and determine whether it's realistic in your organization. If it is, that's great, you'll be able to produce something useful for the organization as a whole, and likely identify resources for your framework. If not, consider conducting a more limited version of this exercise. If you need to skip it altogether, keep an eye out during other mapping and development exercises, for relevant pieces of data that are being conducted by other projects.

1. Together with management or other team members who should have an overview, make a list of every data collection effort in the organization. This means every ongoing or regular survey, questionnaire, consultation or interview.
2. Enter these as column headers in a spreadsheet, and identify focal points for each one.

Focal person:	Maya	Christopher	Tin	Maya	Alix
Name of process:	Public Survey	Baseline Survey	Video	Gender Project	Citizen

3. Have the focal points enter the information they collect as rows in the spreadsheet. For questionnaires, this means one row per question, for less structured formats (like video testimonies), this means a row for each piece of information that is coded or extracted in some way.

Focal person:	Maya	Christopher	Tin	Maya	Alix
Name of process:	Public Survey	Baseline Survey	Video	Gender Project	Citizen
Number of people in household (exact number)	x				
Household income (exact number)	x				
Household income (range)		x			
How serious of a problem is sexual harrassment in your workplace? (10 pt scale)		x			
Geolocation (City)		x			
Geolocation (lat/long)			x		
Have you faced sexual harrassment at your workplace? (Yes/No)				x	
Geolocation (Address)				x	
Have you been harrassed at your place of work? (Yes/No)					x
Harassment incidents (aggregate number)					x
Geolocation (lat/long)					x

4. You can have rows filled out simultaneously in an online spreadsheet[17] or sequentially. This will take time, and the process will depend on how available contact points are to enter the information. Either way, **rows should be combined as they are added.** If a survey is entered into a spreadsheet and has a question very similar to a question that has already been entered for a community consultation ("How corrupt do you believe the judicial system to be?" and "In your experience, is it necessary to pay a bribe in order to access legal services?") these rows should be combined, and the differences between the two surveys should be expressed where the row and column meet.[18]

Focal person:	Maya	Christopher	Tin	Maya	Alix
Name of process:	Public Survey	Baseline Survey	Video	Gender Project	Citizen
Number of people in household (exact number)	x				
Household income (exact number)	x				
Household income (range)		x			
How serious of a problem is sexual harrassment in your workplace? (10 pt scale)		x			
Geolocation (City)		x			
Geolocation (lat/long)			x		
Have you faced sexual harrassment at your workplace? (Yes/No)				x	
Geolocation (Address)				x	
Have you been harrassed at your place of work? (Yes/No)					x
Harassment incidents (aggregate number)					x
Geolocation (lat/long)					x

Focal person:	Maya	Christopher	Tin	Maya	Alix
NAME OF MACRO-AREA	Public Survey	Baseline Survey	Video	Gender Project	Citizen Reporting
Number of people in household	Exact number				
Household income	Exact number	Range			
Sexual harrassment in the workplace		How serious of a problem is sexual harrassment in your workplace? (10 pt scale)		Have you faced sexual harrassment at your workplace? (Yes/No)	Have you been harrassed at your place of work? (Yes/No)
Geolocation		City	Lat/long	Address	Lat/long
Harassment incidents					Aggregate number

17 Google spreadsheets (https://support.google.com/drive/) will do this well, as will Zoho docs. (https://www.zoho.com/docs/). Pirate spreadsheets (http://piratepad.net/front-page/) will do so without compromising your control over the data, but with less functionality.

18 Note that this can get messy. If, regarding the example above, another survey asks people why they do not access legal services, and one of several multiple choice responses could be that they feel the system is corrupt. It's important to include this indicator in the same row as the other two, but it will also required entering the question in other row(s) to capture all of the potential responses.

When this spreadsheet is complete, it will give a good overview of all the ongoing data collection processes available to the project. The design team should ask the following questions:

- Could any of this data directly inform project monitoring? Does any of it directly speak to project outputs or the project's causal model?
- Has any of this data been collected long enough to provide a baseline for the project (ie: measuring how things were before the project started, in order to demonstrate change)?
- Can any of these data collection processes be used to collect additional information for project monitoring (by adding or reformulating single questions)?

Note that this mapping can be completed in a workshop environment (potentially even during the workshop for developing the framework). To do so, it will be necessary to have all of the data collection initiatives (the columns) identified ahead of time, and all of the focal points participating. Each focal point should then lead its own group, and groups should work at the same time, using the same collaborative document (for example, an online spreadsheet that allows for synchronous editing).

Harmonizing Indicators and Data Collection

This spreadsheet exercise will quickly illustrate where data collection efforts across an organization are collecting similar information, and where there are opportunities to harmonize indicators. A spreadsheet that is longer than it is wide will indicate that efforts are generally collecting different types of information. A spreadsheet that is as wide or wider than it is long will indicate that several initiatives are collecting complementary information.

Harmonizing these indicators implies choosing one indicator and using it for all similar information types. Streamlining data collection processes implies both using standardized methods and procedures (to avoid continuously reinventing the wheel), as well as consolidating data collection in teams and processes in order to conserve resources.

All of these are ways to increase efficiency and minimize costs at the organizational level. For the process of developing a monitoring framework, this exercise is about finding opportunities in one data collection process, to collect information for another.

7. Map external data

In some instances, data collected by third parties can be just as useful as data automatically produced by the technology you're trying to measure, or data collected by other projects in the organization. Mapping this probably won't be as lengthy or involved a process as the other two types of data; a quick brainstorm with the design group should suffice. But getting a quick overview may well surface some opportunities and prevent resource waste.

Have the group reflect on the same three types of data (data that directly speaks to project outcomes or causalities, data that could inform baselines and change over time, and processes that might be used for additional data collection). Consider the following questions:

- What data is being collected by peers? Is any of this directly relevant to the project outcomes or causality? Could this data potentially be accessed or shared on a regular basis so that the project would not need to collect it?
- What data is being collected by government institutions or officials? Would any national statistical data be useful? Are there any open data initiatives to be reviewed?
- What data is being collected by other international actors? Is the project's country included in international indices on governance or technology? Do large IGOs like the World Bank or large NGOs like Transparency International run regular surveys?

Note that there is a risk inherent in using third party data for monitoring. What if they suddenly decide to change their indicators, or stop collecting data altogether? When third party data sources are mapped, take a moment to note:

- How essential is the data? Is it key to monitoring the success of a project, or is complementary, a "freebie" or a "bonus"?
- How likely is it that the data will continue to be produced?
- How difficult would it be for the project to generate the same data if it suddenly became unavailable? (Civil society managed consultations with a key stakeholder group the project already engages will be much more feasible to replicate than household surveys conducted by the World Bank).

Developing the Framework

Who's involved: Management, people who will be collecting the data, and everyone who will be using the data. Ideally, this also involves some form of input from everyone involved in the project, including communities and local counterparts.

What you need: Buy-in and support from management, information from the mapping exercises, broad participation from the people who will be using the monitoring data.

How to do it: Ideally this work will take place in a single workshop (1-2 days), where all relevant staff are participating for the whole time. Breaking it up into separate meetings is not impossible, but will risk losing momentum and missing opportunities. Note that these steps are presented in the order they should be conducted, and if steps do need to be separated into different meetings, try to keep 7-8, 9-11 and 13-14 together. Note also that it may be useful to re-order steps 9-11 (to review methods, data sources and indicators), or to run them in parallel groups. It's important, however, to address steps 9-11 ***after*** reviewing project design and use cases, and ***before*** prioritizing.

Steps:

8. Project Design
9. Understand what you'll do with it
10. Review methods
11. Review data sources
12. Review Indicators
13. Prioritize
14. Assign timelines and roles

8. Project Design

One of the most fundamental truisms of project measurement, is that it's not all that different from project design. A well designed project will have identified a causal chain that links activities to desired outputs and outcomes, as well as assumptions and risks along the way. These causal links and risks are often the best place to measure project activities, and will very naturally suggest indicators and data sources. If the project already has a theory of change or other kind of logic model, now is the time to take it out and dust it off. What are the key points at which project activities are supposed to lead to responses? What are the key assumptions? Discuss this with the project team, and see if it suggests any indicators for measurement.

If your project does not have a logic model or theory of change, it's never too late to build one. Doing so isn't always easy, but for projects already underway, it can be an excellent mechanism for surfacing hidden lessons, ensuring a common vision among team members, spotting waste and finding surprise opportunities. A well structured and facilitated half-day meeting with all staff can go a long way.[19]

19 For a general introduction to the Theory of Change see the Theory of Change tool (https://www.theoryofchange.org/what-is-theory-of-change/). For hands-on instruction on how to develop a ToC collaboratively, see the Aspen Institute's Practical Guide to Theories of Change and Community Development Exercises (http://www.aspeninstitute.org/sites/default/files/content/docs/rcc/rcccommbuildersapproach.pdf). See also the International Budget Partnership's Super Duper Impact Planning Guide (http://internationalbudget.org/wp-content/uploads/Super-Duper-Impact-Planning-Guide.pdf)

9. Understand what you'll do with it

Hopefully this was discussed in the first step when deciding to measure, so there should already be some clear ideas on the table about whether measurement is primarily a tool to improve social impact, increase project efficiency, facilitate reporting, inform policy and management decisions, support advocacy or manage donor relationships.

Now is the time to take a closer look at who is going to use the information towards those ends - and how. To start this conversation, it may be useful to review the following types of measurement objectives together with staff:

- Monitoring against program goals and objectives (Is the program achieving what it intended?)
- Monitoring program outputs in key areas (What has the program delivered?)
- Monitoring short to intermediate term outcomes (What is the program beginning to achieve in key result areas?)
- Monitoring changes against a baseline (What changes have occurred over time?)
- Financial monitoring (How have funds and resources been used?)
- Monitoring management and administrative arrangements and processes (What processes have been used during program implementation?)
- Monitoring key areas such as stakeholder relationships (What are the views of stakeholders of the progress of the program, against the stated program logic?)

(adapted from Core Concepts in Developing Monitoring and Evaluation Frameworks)

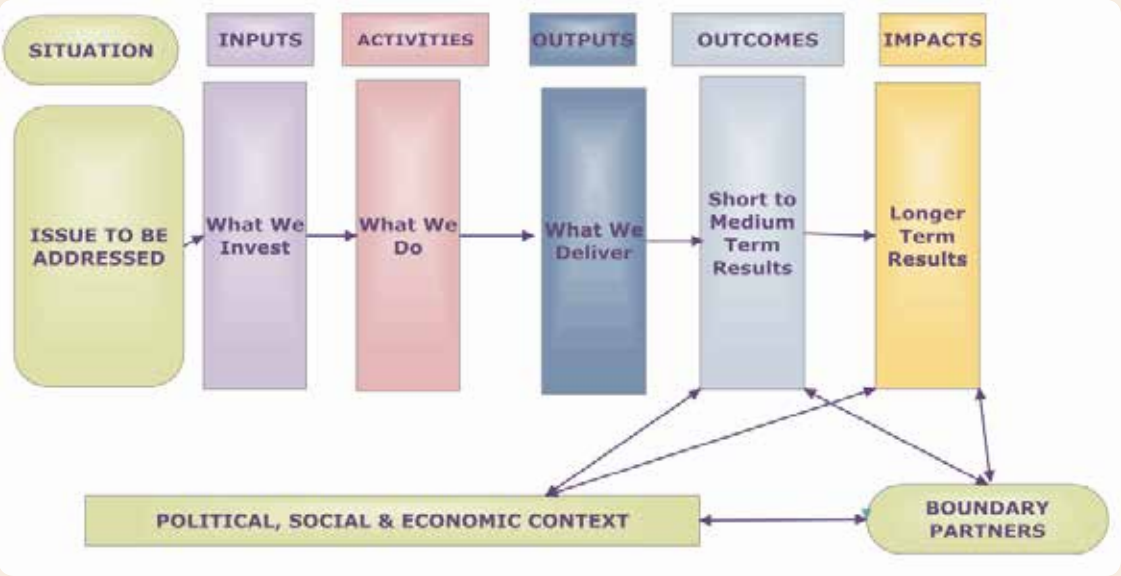
This list is neither required nor complete. In fact, the most useful and interesting monitoring results might involve mixing and combining these.

Use this list as a starting point to begin developing use cases (short descriptions of how information will be used).[20] To do this, each of the staff representing the different ways in which monitoring data might be used (project manager or project board for strategic adaptation, development director for reporting, communications focal point for advocacy) present their needs for monitoring information in a way that is specific.

This means moving beyond what people want to know and the abstract level of "what kinds of people are using our SMS reporting system?", to specific questions like "are women using our SMS reporting system less than men, and if so, are they reporting on different issues, and why?".

20 Borrowed from software development, "use case" refers to a specific description of how a tool or piece of information will be used, including details about the person using it, their motivations for doing so, and what they require to do so effectively.

"A program logic model identifies the expected outcomes and impacts arising from program activities and outputs, and presents a results chain or indication of how the activities and outputs are expected to lead to the achievement of the intended outcomes and impacts."



- Core Concepts in Monitoring and Evaluation Frameworks.

Note that the above model suggests a quite clean and linear model of causation. Such models may not often apply to transparency and accountability initiatives, which often operate in contexts of high uncertainty, where the relationships between contextual and project factors are often difficult to disentangle. For this type of work, logic models and theory of change diagrams need to reflect these complicated relationships.[1]

1 For a discussion of how complex political contexts impact project design (and logic models by extension), see the Transparency and Accountability Initiative's draft think piece on working politically, at http://www.transparency-initiative.org/wp-content/uploads/2014/03/TP.-Thinking-and-Workin-Politically.-Draft.pdf. For suggestions on how to go about developing a logic model that fits with your project, see http://betterevaluation.org/blog/drawing_logic_models.

This last formulation has enough detail that it can be broken down into several different indicators (percentage of users by gender, a categorization of reporting types broken down by users' gender, and potential reasons for not using the service, which would need to be developed on the spot). Mapping out these specific types of information for different types of use cases is a good way to identify which kinds of information are most useful across the organization or project, and for meeting different types of needs. It's also a good way to identify when similar pieces of information need to be presented differently for different uses. For example, simple disaggregated use statistics may be sufficient for the person writing donor reports, while the person responsible for designing the reporting app or writing the script for an ***interactive voice response*** may need to see this expressed as a function of literacy, or disaggregated by times of day to understand it against the background of social factors that limit access to media.

10. Review methods

There are a number of established methodologies for monitoring and evaluation, each with its advantages and disadvantages, its specific use cases, and contexts for which it is particularly well fitted. Most methods are also elaborate and involved enough to support expert careers, but it's safe to say that most project staff will not have the time or desire to become experts.

The tables below give a quick overview, grouping commonly used methods into loose categories of rapid assessments, participatory methods, narrative analysis and formal methods. These distinctions are not strict, and the brief descriptions they contain are intended to provide project staff with an introduction and links for further information. By understanding basic differences between these methods, project staff can make more informed decisions about how much time and resources are worth investing in learning the methods, and at what point external expertise may or may not be necessary. To consider these and other methods in greater detail, it may also be worth reviewing the Better Evaluation Website, the World Bank's handbook, *Monitoring and Evaluation: Some Tools, Methods and Approaches*, and Learning for Action's *Advocacy Evaluation Mini-Toolkit: Tips and Tools for Busy Organizations*.[21]

When reviewing methods, it's important to think about how they might be combined. It's also important to note that all methods can, and should, be adapted to individual project contexts. This includes customizing when and how often information is collected.

21 See Better Evaluations's guide to approaches (http://betterevaluation.org/approaches), The World Bank (2004) - http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2006/02/15/000012009_20060215093620/Rendered/PDF/246140UPDATED01s1methods1approaches.pdf, and LFA Group (2013) - http://www.lfagroup.com/wp/wp-content/uploads/2013/04/Advocacy-Evaluation-Mini-Toolkit.pdf, respectively.

RAPID ASSESSMENT

Name	Description	Demands
Mini-Survey	A structured questionnaire with a limited number of close ended questions that is administered to 50-75 people. Selection of respondents may be random or 'purposive' (interviewing stakeholders at locations such as a clinic for a health care survey).	■ human resources ■ basic methodological expertise (qual & quant)
Focus group discussions	A facilitated discussion in a small group (approx 8-12) of carefully selected participants. Participants might be beneficiaries or program staff, but should have similar backgrounds. Discussions are semi-structured and recorded in detail.	■ minimal ■ human resources: workshop expenses
Key Informant Interview	A series of Open-ended questions posed to individuals selected for their knowledge and experience in a topic of interest. Interviews are qualitative, in-depth, and semi-structured. They rely on interview guides that list topics or questions. Although these interviews are not appropriate if quantitative data are needed, this method can deliver information quickly and at low cost.	■ access to experts ■ basic expertise in qualitative methods ■ interview guides
Direct Observation	An approach that records information about processes, activities, discussions or impressions observed at a program site, and recorded using a detailed observation form. Direct Observation is susceptible to observer bias among other issues, and requires some methodological rigor to maintain credibility.	■ statistical rigor to combat observer biasprepared observation forms

For further guidance and resources on mini-surveys, see The World Bank (2004) - http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2006/02/15/000012009_20060215093620/Rendered/PDF/246140UPDATED01s1methods1approaches.pdf

For further guidance and resources on focus groups, see Better Evaluation's overview - http://betterevaluation.org/evaluation-options/FocusGroups

For further guidance and resources on key informant interview, see Better Evaluation's overview - http://betterevaluation.org/resources/guides/interviews/conducting_keyinformant_int

For further guidance and resources on direct observation, see USAID Center for Development Information and Evaluation (1996) - http://dmeforpeace.org/sites/default/files/USAID_TIPS-Using DirectObservation Techniques.pdf

PARTICIPATORY METHODS

Name	Description	Demands
Beneficiary Assessment	An investigation of perceptions among a systematic sample of beneficiaries and other stakeholders, to ensure that their concerns are heard and incorporated into project and policy formulation. The general purpose is to 'give voice' to poor and other hard-to-reach beneficiaries, highlighting constraints to beneficiary participation, and obtain feedback on interventions.	■ time ■ basic expertise with sampling
Community group interviews	A facilitated discussion (or series of them) open to all community members and in which set questions are discussed. The interviewer follows a carefully prepared questionnaire.	■ community engagement ■ skilled facilitation
Most Significant Change	Collects and analyses personal accounts of change, includes processes for learning about what changes are most valued by individuals and groups. Accordingly, MSC does not provide information about the usual experience but about the best experiences. This method requires significant time and is especially useful for understanding how and why change occurs.	■ time ■ human resources

For further guidance and resources on beneficiary assessment, see The World Bank's overview - http://go.worldbank.org/3AIUQJ5WP0

For more information about Participatory Evaluation and rapid appraisal methods see USAID Center for Development Information and Evaluation (1996) - http://dmeforpeace.org/sites/default/files/USAID_TIPS-Using DirectObservation Techniques.pdf

For further guidance and resources on most significant change, see Better Evaluation's overview - http://betterevaluation.org/plan/approach/most_significant_change

NARRATIVE ANALYSIS

Name	Description	Demands
Contribution Analysis	Assesses evidence that an intervention has contributed to observed outcomes and impacts. This method is most useful when there is a clear causality to be tested. This method relies primarily on desk research and analysis, and does not require systematic data collection. This approach may also be applied through a stakeholder workshop, which will require strong facilitation skills. Unlike most methods listed here, episode studies will be conducted after project implementation.	■ a clear theory of change ■ desk research and analytical capacity ■ strong facilitation if conducted in a workshop
Collaborative Outcomes Reporting	Collaborative Outcomes Reporting builds on contribution analysis, adding expert review and community review of the assembled evidence and conclusions. Like contribution analysis, it is useful in contexts where the project is not experimental and after a theory of change has already been developed, where it can help confirm or revise a theory of change	■ a clear theory of change ■ desk research and analytical capacity ■ engagement with communities and experts
Episode Studies (tracer studies)	A case study that focuses on a clear policy change and tracks back to assess what impact research had among the variety of issues that led to the policy change. They could be focusing on a single episode or comparative episodes. It differs from other case study approaches, which usually take an initiative as the starting point and look forward. Unlike most methods listed here, episode studies will be conducted after project implementation.	■ stakeholder workshop ■ skilled facilitators
Process Tracing	Process tracing is a methodology composed of four tests, to assess causality in specific cases, and is a widely accepted model for evaluating causality. It is implemented by applying observations about a case to four specific tests, as a means of evaluating alternative explanations for outcomes.	

Name	Description	Demands
Cost Benefit and Cost-effectiveness Analysis	Tools for assessing whether or not the costs of an activity can be justified by the outcomes and impacts. Cost-benefit analysis measures both inputs and outputs in monetary terms. Cost-effectiveness analysis estimates inputs in monetary terms and outcomes in non-monetary quantitative terms (such as improvements in student reading scores). Though these methods can be implemented at any time during a project cycle, cost effectiveness analysis is most useful prior to implementation.	■ methodological expertise ■ financial resources ■ human resources

For further guidance and resources on episode studies, see Better Evaluation's overview - http://betterevaluation.org/evaluation-options/episode_studies

For further guidance and resources on contribution analysis, see Better Evaluation's overview - http://betterevaluation.org/plan/approach/contribution_analysis - and their review of Collaborative Outcomes Reporting - http://betterevaluation.org/plan/approach/cort

See Collier (2010) - http://dmeforpeace.omnidev3.com/sites/default/files/Collier_Process%20Tracing.pdf - and Collier (2011) - http://www.ukcds.org.uk/sites/default/files/uploads/Understanding-Process-Tracing.pdf - for further resources

See Better Evaluation's overview of Cost-Effectiveness Analysis - http://betterevaluation.org/evaluation-options/CostEffectivenessAnalysis - and overview of Cost-Benefit Analysis - http://betterevaluation.org/evaluation-options/CostBenefitAnalysis - for further guidance and resources

FORMAL METHODS

Name	Description	Demands
Randomised Controlled Trials	RCTs are an experimental method, especially well suited to identifying causality when there is no clear or broadly accepted causal hypothesis. This method randomly assigns a control group that is not impacted by a project, as well as an experimental group that is. Both groups are then monitored, and the results compared to understand the impact of a project. The method is fairly demanding and requires significant technical expertise. Because of its experimental nature, this method needs to be adopted prior to project implementation.	■ statistical expertise ■ financial resources ■ time

Name	Description	Demands
Formal Surveys	Formal surveys can be used to collect standardized information from a carefully selected sample of people or households. Surveys often collect comparable information for a relatively large number of people in particular target groups. Examples include Household Surveys, which randomly select respondents by selecting houses for specific neighborhoods, or Exit Surveys, which interview people that have just engaged with a service. Formal surveys will require representative samples and methodological rigor, so will tend to be demanding in terms of expertise, human resources and financial cost. In addition, processing and analysis of data can prove to be a significant bottleneck for organizations with limited statistical capacity.	■ statistical expertise ■ human resources ■ financial cost ■ time
Panel Data	Expert panels can be composed for projects or contexts that are particularly complex, or when monitoring findings may be contentious and there is a need for credibility. Panel members are selected on the basis of their expertise, standing and ability to analyze complex phenomenon. Panels are regularly asked the same questions (either individually or in plenary, either live or virtually) and their answers are tracked over time.	■ engagement with experts ■ time ■ skilled facilitation if conducted in live groups

For further guidance and resources on RCTs, see Better Evaluation's overview - http://betterevaluation.org/plan/approach/rct

For further information, including additional reading, see the World Bank publication, Monitoring and Evaluation: Some Tools, Methods and Approaches, The World Bank (2004) - http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2006/02/15/000012009_20060215093620/Rendered/PDF/246140UPDATED01s1methods1approaches.pdf

For further guidance and resources on panel data, see Better Evaluation's overview - http://betterevaluation.org/evaluation-options/expert_panel

Mixing methods

None of these methods are a magic bullet, and there is no broad agreement about any of them being more generally appropriate than any other. There is broad agreement, however, that **mixed methods are almost always advisable**.[22] Generally, the term mixed methods refers to using a combination

22 For background on the advantages of using mixed methods to evaluation progress towards political objectives, see Roch and Kelly's Advocating a mixed methods approach to the evaluation of politics (http://www.dlprog.org/news-events/advocating-a-mixed-methods-approach-to-the-evaluation-of-politics.php).

of quantitative and qualitative data; that is, both numbers and narratives. The idea is that together, different kinds of measurement will do a better job of answering nuanced questions about complex phenomena. As phrased by the Hewlett Foundation:

> This process of triangulation allows one method to complement the weaknesses of another. For example, randomized experiments can determine whether a certain outcome can be attributed to an intervention. But complementary qualitative methods are also needed to answer questions about how and why an intervention did or didn't work.[23]

Table 1 **Evaluation designs and basis for causality claims**[5]

Evaluation design approaches	Basis for claims of causality
Experimental Randomised Controlled Trials (RCTs), quasi-experiments, natural experiments	Counterfactuals; the co-presence of cause and effect
Statistical including descriptive statistical modelling and longitudinal studies	Correlation between cause and effect or between variables
Theory based Using a theory of change (ToC), process tracing, contribution analysis and impact pathways	Identification or confirmation of causal processes or chains and supporting factors and mechanisms at work in the context. May include some statistical analysis
Case based Interpretive – naturalistic, grounded theory or ethnography	Comparison across and within combinations of causal factors
Participatory Participatory evaluation for empowerment.	Validation by participants that their actions and experienced effects are caused by the programme.
Agency-led Learning by doing, collaborative action research	Involves adoption, customisation and commitment to a goal
Synthesis Meta analysis, narrative or realistic-based synthesis (Pawson 2002)	Accumulation and aggregation within a number of perspectives above

(Adapted from Stern *et al.* 2012)

When considering how to combine methods and triangulate for insights, it's also important to think about the way different types of methods vary in their assumptions and their credibility. Relying solely on quantitative data will often provide only modest insights about causality or complex relationships. Over-reliance on participatory qualitative data can expose measurement data to criticism that it is "only stories." In striking a balance between different methodologies and data types, it can be useful to review the below table, produced by the Centre for Development Impact.[24]

On the basis of this, and keeping in mind the use cases and available resources, brainstorm potential methodologies for the monitoring framework.

23 Hewlett Foundation's interal working paper, Evaluation Principles and Practices Twersky et al (2012). For further discussion of complementary qualitative methods, see also blogposts and discussions at the Better Evaluation website (http://betterevaluation.org/blog/mixed_methods_part1).

24 From Working Paper 01 by Shutt & McGee (2013) - http://bigpushforward.net/wp-content/uploads/2011/01/CDIPracticePaper01finalR.pdf.

11. Review data sources

Building on the discussions so far, this step aims to identify possible sources for monitoring data. Quickly review the mappings of already available data from technology, other projects and internal sources. Then discuss where the project would go for additional data to monitor project activities.

Some possible sources include:

- Usage statistics for project technology
- Qualitative and quantitative data actively submitted as part of the project (through reporting platforms, organized consultation or other)
- Field notes and interviews conducted during project work
- Exit surveys or other surveys conducted when accessing services
- Interviews or surveys with peripheral stakeholders, including government or corporate representatives
- Administrative data
- Community communication systems, both online and offline

Use this list to begin a conversation about appropriate data sources. Keeping in mind the use cases and available resources, brainstorm potential data sources for the monitoring framework. For each, determine what kinds of questions they would be able to inform.

12. Review Indicators

Indicators are the nuts and bolts of monitoring and evaluation. They are the pieces of information that get collected and made sense of, and their selection and formulation is incredibly important for determining what a monitoring framework can and can't do.

Good indicators are:

1. Mixed (there are several different types in a framework)
2. SMART (specific, measurable, achievable, relevant and time-bound)
3. Actionable (it's easy to know how to respond when they are collected and reviewed as part of a monitoring process)

In simple terms, an indicator is a piece of information, formulated in a way to allow measurement and comparison, across time and contexts. Good indicators are also easy to use, understand and respond to, and can withstand statistical analysis. As a point of departure, this step should aim to brainstorm a collection of monitoring indicators that satisfy the following three criteria: they should be mixed, SMART, and actionable.

Mixed indicators:

As with mixed methods, a good monitoring will be composed of different types of indicators, and there are a few different ways to distinguish between indicators. Here are a few:

- **Quantitative** indicators are expressed as numbers (67% of users were satisfied with a service), while
- **qualitative** indicators are expressed as text (participants in a consultation described wide satisfaction with a service).

- **De jure** indicators reflect what the law is (it is illegal for politicians to accept undisclosed gifts),
- **process** indicators reflect what is being done to address an issue (politicians are [not] being prosecuted for not disclosing gifts), and
- **de facto** indicators reflect what is actually happening (politicians are [not] disclosing gifts).

- **Input** indicators reflect resources allocated to affect change (a budget is established, a program created),
- **output** (or process) indicators reflect activities undertaken (a workshop was held),
- **outcome** indicators reflect how outputs were used or had an impact (workshop participants acquired new skills), and
- **impact** indicators reflect meaningful, long term consequences (these skills were used in their organizations and institutions to implement meaningful policies).

SMART indicators:

SMART refers to a set of criteria for indicators, understood as targets or indicators of success, to be useful and manageable. Though often debated and refined, they are also widely regarded as good practice, and the below version provides a useful checklist for quality controlling indicators.

- Specific – Double the percentage of local youth that have access to independent media before the next parliamentary election, in three years time
- Measurable – It should be possible to poll a representative sample of local youth.
- Achievable – Not 100% of local youth, unless that really is feasible
- Relevant – This indicator would be relevant to a project aiming to increase political engagement among youth, but less relevant to a project seeking to increase youth enrollment in education.
- Time-bound – specifies when the result(s) can be achieved.

Actionable indicators:

Actionable indicators are indicators that can be easily acted upon. Consider the following survey questions:

- Are you able to access the mobile health clinic as often as you would like to? (yes/no)
- Do any of the following factors prevent you from accessing the mobile health clinic as often as you would like to? (language, timing, social stigma, distance, etc)
- How often do you access the mobile health clinic last month? How often would you like to be able to access the mobile health clinic in a month?
    - (if language): What language would you prefer to receive the service in?
    - (if timing): What times of the month/week/day would you be able to access the service? (open question).

Each of these questions provides a dramatically different indicator or set of indicators, and are also dramatically different in their actionability. The first may warn a project that there is a problem with accessibility of services, but offers no information on what to do about that problem. The second offers information about the nature of the problem, but needs guessing or additional data collection in order to respond. The third set of questions provides actionable data that can be used to correct project challenges.

You should have already spent some time mapping out specific uses and use cases for monitoring data, which will help identifying and selecting actionable indicators.

Based on these three criteria for indicators (mixed, SMART and actionable), and keeping in mind potential data sources, brainstorm potential indicators for the monitoring framework. For each, determine what kinds of questions they would be able to inform.

Choose "One Metric that Matters"

"It focuses the entire company. Avinash Kaushik has a name for trying to report too many things: data puking. Nobody likes puke. Use [the one metric that matters] as a way of focusing your entire company."

From "Lean Analytics Use Data to Build a Better Startup" - http://www.amazon.com/Lean-Analytics-Better-Startup-Faster/dp/1449335675

13. Prioritize

By now, the project should have a significant amount of information at hand. Having mapped resources, needs, use cases, data sources, methods and indicators, it may seem as if everything needs to be measured all the time. That's *not* the case. This step is perhaps the most difficult and the most important, because it involves doing away with everything that isn't absolutely necessary, and working towards a monitoring framework that can be implemented without placing a strain on the project.

It's been suggested that every organization or project should be able to identify the one metric that matters, and that doing so can help organizations to manage chaos, to set clear priorities, to focus teams, and to foster cultures of experimentation.[25]

A single metric might be excessive, but it is definitely a good idea to streamline data collection and avoid measurement for measurement's sake.[26]

Doing so will simply require placing all options on the table and having the design team argue for the most important metrics. To facilitate this conversation, try **working backwards** from the informational needs, to what would actually be necessary to meet them, using these four questions in order:

1. On the basis of all the mapped data and resources, and noting what is likely feasible for the organization in terms of methodologies and data sources, what are the **3 highest priority use cases** for monitoring data?
2. What is the **minimum collection of indicators** (that are mixed, SMART and actionable) that will meet the needs of those use cases?
3. What **methods and data sources** can realistically produce those indicators?
4. What **resources** would be required (financial, staff, expertise) would be required to collect those indicators, using those data sources and methodologies?

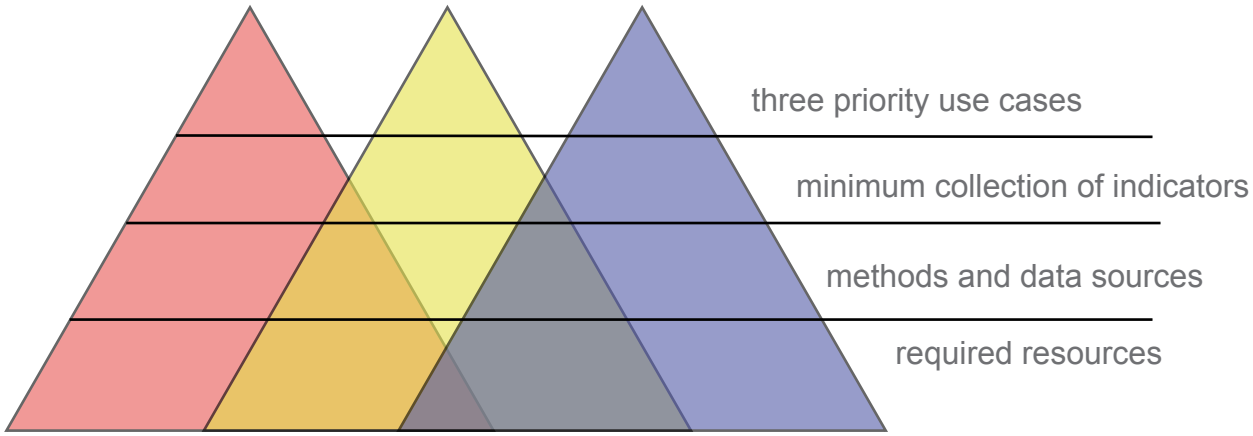
Map the answers to these questions as overlapping pyramids, with the first question on top, requiring more detail as you move down to practical questions about resources, and tracking the data flow with lines up to the top.

When you've finished drawing out these pyramids, make a point of highlighting indicators and methods that support more than one use case. Note that getting to this point may take some time, and when you're finished you may find that you have proposed an impossible framework, which requires much greater resources that the project is able to mobilize. That's ok. Do it again, working from the bottom up:

25 Beth Kanter reviewed Lean Analytics Use Data to Build a Better Startup Faster in her blogpost What is your organization's "One Metric That Matters"? (http://www.bethkanter.org/performance-management/).

26 For discussions of the "obsessive measurement disorder" some see sweeping the non-profit field, see Rogers (2010) - http://devpolicy.org/usaid-obsessive-measurement-disorder20101122/ and Pratt (2012) - http://www.intrac.org/blog.php/22/obsessive-compulsive-measurement-disorder .

Remove the methods and data sources that are too demanding and try to find alternatives that will still produce the indicators you've prioritized. If you can't find feasible methods and data sources to populate the desired indicators, you may have to switch out those indicators, suggesting new ones that will still inform your priority use cases. Keep replacing what you can at the bottom of each pyramid, working upwards, until you have mapped direct and feasible relationships between the resources you have available, and the questions that need to be answered for your three priority use cases.



What to do if it's hard:

It's ok if you're ***not*** able to devise a framework that supports all three priority use cases. Even if you are only able to support one of the resources you have available, take comfort in knowing that you are developing framework that can be realistically sustained over time. It's a framework that is not likely to not buckle under administrative pressure or be abandoned when time is scarce. More importantly, it will meet actual project needs, instead of simply producing data for data's sake, or to fill donor reports. This is a strong practice that be built on.

What to do if it's easy:

If you find that you are able to easily meet all three priority use cases with methodologies and data sources that are realistically sustainable for the project, resist the temptation to begin adding additional use cases and widening the base of the pyramid. Unless the other use cases are truly imperative, it might be wiser to pilot a manageable monitoring framework first, and add additional use cases and data collection processes later, once a resilient process is established. It may even be wise to start only with one or two of the use cases, and work out the kinks in that before adding additional information.

Questions to Ask

1. What do you really want to know?
2. How can we craft questions that are as specific as possible in terms of identifying what we want to know? (For example, don't ask "Is our strategy working?" Instead ask questions like: "Has this set of advocacy activities helped us to get new policymaker champions on board?")
3. Are some activities and benchmarks more important than others – are there certain goals that must be accomplished for many other goals to be reached? (It can be tempting to want to evaluate everything – resist the temptation, or evaluation will become overwhelming.)
4. Is there a particular strategy that could benefit from continuous feedback for the purposes of ongoing improvement? (There may be strategies that you already know are working well – you get less "bang for the buck" by concentrating your evaluation resources on those strategies.)
5. Is there a strategy in which the advocates are particularly curious about whether they are effective?
6. Conduct a thought experiment: what will you do with this information once you get it? Will it be truly helpful?

Adapted from Advocacy Evaluation Mini-Toolkit: Tips and Tools for Busy Organizations (LFA Group 2013 - http://www.lfagroup.com/wp/wp-content/uploads/2013/04/Advocacy-Evaluation-Mini-Toolkit.pdf).

14. Assign timelines and roles

Having drawn out monitoring pyramids, you've conceptualized a direct connection between your project's capacities and the production of useful monitoring information. You've come a long way towards designing a feasible and actionable monitoring framework. What's left is to plot this onto timelines and processes.

To begin, start by asking ***when*** you need the data. This should be quite clear from the use cases at the top of the pyramids. Is the information for project boards that meet annually? For weekly team meetings? To respond to specific things that are going to happen at some point, but you don't know when?

Work backwards from this question about when you need the data, to determine how often the data will be collected. Using the different data collection methods in the pyramids, write out a timeline for data collection and data review processes. When doing so, keep the following points in mind:

- Whenever possible, **data collection should be incorporated into other project activities**. This conserves resources, but also helps to keep data collection aligned with larger project goals, increasing the chance that monitoring can be useful to project activities, and that monitoring does not get ignored or forgotten.

- **Have frank conversations with people supporting the process**, to ensure that there's a shared understanding of how collecting and managing data will strengthen their work.[27]
- **Share monitoring outputs widely and openly** throughout the project, organization and with stakeholders, in language that is easily accessible, but also as raw data if possible. Doing so can increase the chances of getting meaningful and useful feedback.
- **Create opportunities for reflection and discussion of monitoring results** within organization. Something as simple as an informal lunch to discuss results can surface unexpected insights and additional resources.
- **Look for ways that monitoring can strengthen project activities**, for example by framing data collection as consultations to strengthen relationships with stakeholders, or by sharing data with government partners to build trust and strengthen coordination.
- **Make sure that data review processes have enough time** to significantly influence how the next round of data collection is conducted, and that the people reviewing the data have a real influence on how project strategies are designed and adjusted.

In project management, the process of collecting, analysing and using information goes on informally all the time. Monitoring and evaluation also requires the "formal" and systematic collection and analysis of information. The information needs to be reliable and trustworthy. It is the evidence on which decisions are made about the work.

(Monitoring and Evaluation, Louisa Gosling)

27 Failing this means that monitoring processes might feel top-down and non-productive. This can be bad for data quality. You might even end up with critical data disappearing, only to find it used to wrap food outside the project office (that actually happened, see Goldstein (2012) - http://blogs.worldbank.org/impactevaluations/node/835).

Once you've mapped out a timeline for monitoring on-the-go, and aligned that with the project cycle, you'll need to make sure that every task has a specific person assigned to it. This has likely already happened when discussing the timing of data collection and review, but it's important to thresh out data collection points into specific tasks, and associate those with roles and responsibilities. Do log files need to be saved in a specific folder at a specific time? Do paper documents need to be scanned? Does travel need to be booked for field missions? Do licences need to be purchased for data streams? These are the types of small tasks that can be easily forgotten in the big picture, but on which a functioning framework will rely. Complete this exercise by mapping out roles and responsibilities for all people involved in the framework. What will they be doing to collect, analyze and review monitoring data?

When all steps are assigned to specific people, you will have a finished timeline for monitoring processes and allocation of tasks. How you package this depends entirely on how your individual project and organization works. Whether it is a 50 page document, a project on Basecamp, or a drawing on a whiteboard in the staff meeting room, this is the blueprint for a functioning monitoring framework.

Rolling it out

Who's involved: Project staff and consultants responsible for data collection and analysis, as well as a team responsible for reviewing monitoring data and making decisions in response to it, ideally composed of project management, organizational management and external counterparts.

What you need: A complete framework for monitoring on-the-go, or at least a blueprint for that framework that identifies specific timelines, roles and responsibilities, as described in step 14.

How to do it: Collect data and review data as planned. Adjust the framework based on challenges and insights. Rinse and repeat.

Steps:

15. Start an iterative process
16. Analysis
17. Communicating and learning

15. Start an iterative process

Based on the preceding design steps, you should have all the components you need to plan and implement data collection and review processes. Maintaining the flexibility necessary for iteration when first rolling out the framework can be challenging, however. There are bound to be bumps and hurdles in first efforts. Data might be missing or poorly formatted. The results might be counter-intuitive and throw the entire effort in question.

It is important that challenges such as these do not delay the monitoring timeline so much that the project is unable to respond strategically to

Target Audiences:

- A small group of people intimate with the project's work and goals
- Broad and indistinct local communities, with little technical literacy, but for whom the project is important
- Statisticians and researchers with an interest in documented results
- People familiar with the project or organization who want to see the program succeed and limit risk

Each of these groups represents a different kind of audience, for whom analysis should be tailored to the degree possible, independent of what kind of decisions will be made with that analysis. Each also implies different types of language and packaging, independent of the methods used.

monitoring results. Be sure to allocate enough time to data collection, data analysis and data review processes, that even with unforeseen delays in all three areas, the review team is able to make sense of the data and act accordingly.

This means that data review needs to be scheduled well before any major decision points in project activities (rolling out new project phases or initiating new partnerships). If the project's monitoring framework relies significantly on periodic data collection such as regular surveys or community consultations, it's important to have time enough after reviewing data from the first round of data collection, to adjust the frameworks, questions and indicators before the second round begins.

Data review needs to be scheduled well before any major decision points in project activities.

In early iterations of a new framework, this means that it might be wise to review results together with the design team that was responsible for developing the framework. If some of the indicators turn out to be inappropriate, or some of the methods impractical, this group will be best positioned to identify alternatives. This too will take time, however, and should be budgeted accordingly.

16. Analysis

The first step in making use of monitoring data is an analytical process. This is where the rubber first hits the road for turning measurement into learning, and it can take a variety of different forms. Methods for analysing monitoring data will vary as greatly as the data itself, and have hopefully been identified well before data was collected (see common approaches referenced in Step 10).

As with all aspects of measurement, the decisions about analysis- who conducts it, what timeframes, which methods - should be informed by how that analysis is to be used. This should flow directly from the work done in Step 9, but now you'll want to think a bit more about who analysis should target and what kind of analytical output will be most appropriate.

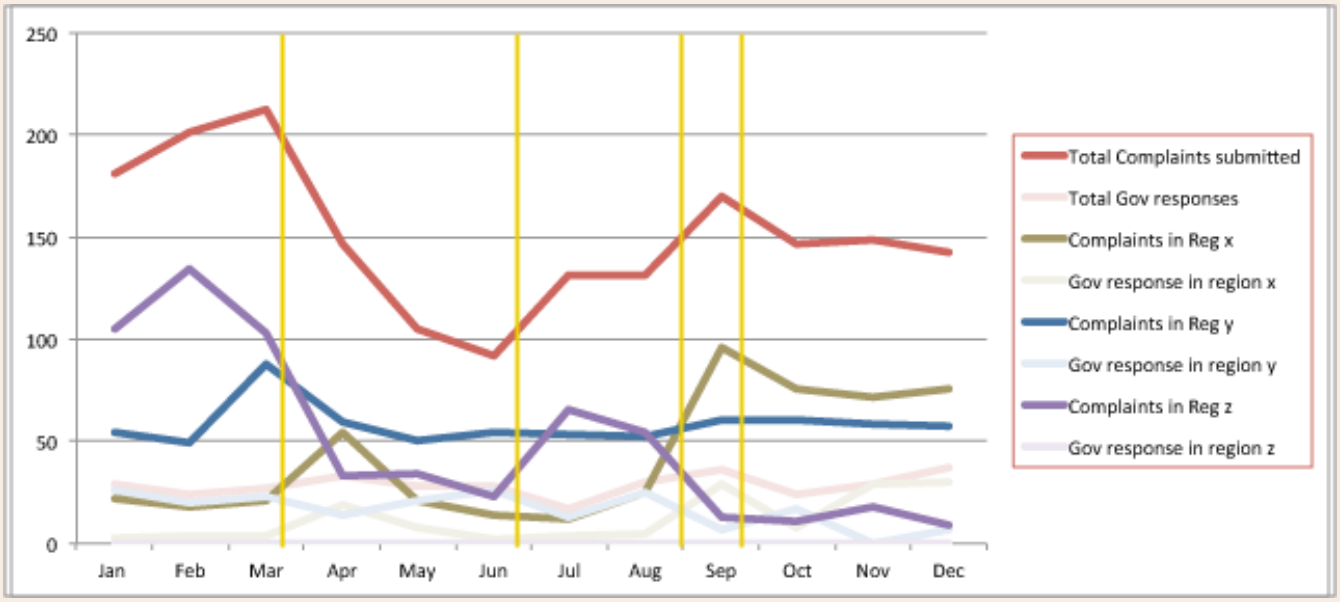
Don't be afraid to make the raw data available.

Sit down with the data in a dedicated workspace and begin applying the analytical methods selected in previous steps. As patterns emerge, consider sharing them informally with the project team, and listen carefully to reactions, objections and concerns that might point to additional analyses or hidden factors.

When the analytical team feels there are clear and defensible findings about trends and consequences, begin writing this up in formats that will gain traction with relevant audiences. This may involve fifty page reports, visualizations, statistical tables, or slide presentations.

Linking incremental changes and key events

It can be useful to plot out all data on timelines against contextual events. First map out monitoring data along an x axis that represents time, and a y axis that represents the data characteristics (% of qualitative or quantitative data matching a given value). This should produce a line graph, that will be more jagged the greater the frequency of data collection. Then place vertical lines across the x access to represent important contextual happenings (municipal elections, a groundbreaking prosecution is announced, a major media merger). See if any interesting patterns emerge.



Consider the following example, in which a hypothetical citizen complaint platform wishes to consider the impact of national elections on project performance. The graph lines represent the number of complaints and government responses. The vertical lines represent the start of electoral campaigns, two notable campaign promises relating to the complaint platform topic, and finally a key debate in which the project topic was addressed. This can provide insights on how national politics affect the willingness of citizens to complain (negatively, until a specific campaign promise was made), and the responses of government (little).

However you package analysis, don't be afraid to make the raw data available, even to those who might not obviously be able to interpret it. Don't be afraid to treat the analytical process as iterative, sharing analysis in a way that makes it clear how people can offer feedback. As feedback comes in, consider alternative findings and explanations. Consider alternative methods; prepare these considerations for sharing in conversations about learning.

Common Challenges

Regardless of how analysis is organized, there are a number of common challenges associated with monitoring for tech and accountability initiatives.

Attribution

Attribution refers to the problem of knowing how and if you can say that a specific output takes place "because of" a specific activity. This is especially challenging in advocacy contexts, and measurement experts tend to prefer measures of ***contribution*** over attribution. In tech for accountability projects as well, it is likely much easier to demonstrate (or argue) that project activities have contributed to accountability outcomes, than to claim such outcomes are directly attributable.[28]

Changes over time

Transparency and accountability work tends to progress at a slow place, and projects that can claim to have achieved their full outcomes are rare. In this context, it's important for monitoring data to be able to track incremental changes over time, however modest. To a large degree, whether or not this is possible will be determined by the data collected and the availability of baseline and contextual data, potentially collected by third parties. Documenting incremental change will generally be easier with more data and more regular data collection, which can help to demonstrate project impact if project objectives and outcomes are broad and difficult to document.

Aggregation and disaggregation

More demographic and contextual data will make it easier to disaggregate data and draw conclusions about specific phenomena, groups, regions or events. Consider the example of compliant platforms and electoral cycles, where the complaint data is disaggregated according to three regions that are politically, ethnically or socio-economically distinct. This may reveal that the regions receiving the most complaints also have the fewest government responses, or that political pressure causes an increase of response in some local governments, while causing a decrease in others. These types of

28 For further discussion of the methodological issues and challenges implied by attribution/causation, ILAC (2008), Mayne, J. (2011) - http://www.cgiar-ilac.org/files/publications/briefs/ILAC_Brief16_Contribution_Analysis.pdf. CDI (2013) - http://bigpushforward.net/wp-content/uploads/2011/01/CDIPracticePaper01finalR.pdf argues that moving from measures of attribution to contribution also implies "shift in emphasis from focusing on indicators, data collection tools and methods, to more critical thinking about choosing MEL strategies and designs that enable appropriate exploration of cause and effect relationships."

insights can be strategically important for monitoring frameworks (should the reporting platform include questions on political affiliation, should it request geodata for complaints, or do large numbers of complaints in a small region suggest over-reporting?). They can also be important for program strategies (do differences in local government responses indicate an institutional resistance within a specific party, and how does this affect advocacy strategies?).

Conflicting data

Monitoring data (especially qualitative data and narratives by stakeholders) may well provide conflicting accounts of project performance. At first look this may raise doubts about the quality of data, but it is important to explore underlying causes of conflicting data. Returning to the source of that data when reviewing monitoring outputs may reveal inappropriate indicators or data collection instruments, and suggest meaningful alternatives. This process may also reveal unanticipated social dynamics that the project wishes to address or account for.

17. Communicating and learning

No matter how elegant, efficient and robust a monitoring framework is, it will only help a project to the extent that it is communicated and learned from. Some general guidelines for communicating results were outlined in Step 14. This section will describe some of the key audiences for monitoring outputs, and potential strategies for engagement. The rules of thumb across all groups are to communicate openly and often, and to create spaces for meaningful review and feedback on monitoring outputs.

Talking to donors

Simply having a functional monitoring framework can be a huge advantage when dealing with donors. It demonstrates a commitment to learning and impact, which can strengthen credibility. Of course collecting data on what works and what does not may also highlight a lot of things that are not working. After all, a framework for monitoring on-the-go is primarily a tool for changing the things that don't work, and adding more of the things that do.

When monitoring data demonstrates project failures, it's important to internalize the narrative of learning both in project management and dialogue with donors.

When monitoring data demonstrates project failures, it's important to internalize the narrative of learning both in project management and dialogue with donors.[29] Internally, this means that there are mechanisms for meaningful iteration between data collection and review. In dealing with donors this implies emphasis on the financial cost and opportunity cost saved by identifying less impactful activities early, and pivoting out of them.

29 Rakesh Rajani of Twaweza offers guidance on building donor relationships on the basis of learning and evidnce, in this video, at https://www.youtube.com/watch?v=dSnQ6gFGyZc .

It is also worth recalling that most donors are eager to learn about what works and what does not work. In many situations this can be used a more explicit conversation with donors about project realities, and eventually bringing donors' measurement and reporting frameworks into line with "realities on-the-ground."[30]

Simultaneously, many donors are interested in evidence in a more abstract sense, and may welcome a discussion about shared monitoring frameworks and sharing indicators between grantees. If a project's peers area also supported by the same donors, this can be an entry point for strengthening data sources and pooling resources, and potentially receiving financial support to do so.[31]

Talking to staff

Hopefully the monitoring framework includes several opportunities for project staff and others in the organization to reflect on monitoring results. This can be accomplished through formal or informal mechanisms, and can be useful for surfacing additional insights on what the data means, as well as challenges for collecting data and opportunities for additional data or novel analysis. Most importantly, staff involved in the monitoring process should experience a sense of ownership with the process, in order to ensure sustainability and data quality. It is important that staff understand the way in which monitoring outputs influence strategic decision-making, and ultimately influence their own workloads and incentives to contribute to monitoring.

Talking to management

At the end of the day, the most powerful potential for monitoring on-the-go may well be to ground and inform broad overhauls and experiments in programming, to reinforce activities that contribute to objectives, and avoid those that don't. Doing this can be challenging institutionally and politically, and monitoring data may provide clear, evidence-based arguments for reallocating staff or resources. When this makes programmatic sense, such data can be useful for advocating to management. These arguments are nevertheless much more likely to be successful if there is management buy-in for the monitoring framework, and a conviction that the framework targets the right kinds of information. Ideally, management will have been involved from the beginning of the process, deciding to measure, mapping resources and developing the framework. At this point, it's important that management is kept up to date on monitoring outputs and lessons, and has an opportunity to contribute to the project's strategic response.

30 For a discussion on donor and grantee perspectives and collaboration around learning and results, see http://tech.transparency-initiative.org/at-tabridge-2012-donors-and-grantees-compare-challenges/ .

31 For more on shared measurement, see Sommerfeldt (2012) - http://cima.ned.org/sites/default/files/final_0.pdf produced by CIMA, and Kramer et al. (2011) - http://www.fsg.org/tabid/191/ArticleId/87/Default.aspx?srpush=true by FSG.

Talking to stakeholders/beneficiaries

Monitoring data can be a useful tool for engaging with project stakeholders during data collection. Reviewing monitoring outputs with stakeholder communities can be equally powerful, especially if their input leads directly to strategic changes in project implementation. Stakeholder feedback may well contradict monitoring outputs, offering alternative explanations or questioning assertions. This can be a unique opportunity to uncover previously hidden power relationships or contextual dynamics. It can also provide additional data that was not explictly collected, but which will strengthen analysis.

Often, open, structured mechanisms for this type of feedback will be helpful (consultations, town halls), facilitating engagement. This can also be a more effective way of publicly demonstrating the impact of stakeholder contributions to monitoring and learning processes, strengthening the relationships that underpin effective programming. Of course, this process can also provide useful information for adjusting and iterating the monitoring framework itself, identifying alternative metrics, indicators or data collection processes.

Resources

Toolkits & Measurement Frameworks (ready to use)

Framework for the assessment of ICT pilot projects Beyond Monitoring and Evaluation to Applied Research, InfoDev. Batchelor, S. & Norrish, P. (2005).
Retrieved from http://www.seachangecop.org/sites/default/files/documents/2005%2005%20%20Framework%20for%20assessment%20of%20ICT%20pilot%20projects.pdf.

IBP's Super Duper Impact Planning Guide, IBP. van Zyl, A. (2011).
Retrieved from http://internationalbudget.org/wp-content/uploads/Super-Duper-Impact-Planning-Guide.pdf.

Advocacy Evaluation Mini-Toolkit: Tips and Tools for Busy Organizations. LFA Group. (2013).
Retrieved from http://www.lfagroup.com/wp/wp-content/uploads/2013/04/Advocacy-Evaluation-Mini-Toolkit.pdf

Equal Access Participatory Monitoring and Evaluation Toolkit. Better Evaluation. (2014).
Retrieved from http://betterevaluation.org/toolkits/equal_access_participatory_monitoring.

Results Based Monitoring and Evaluation Toolkit, Local Livelihoods. Spreckley, F. (2009).
Retrieved from http://mymande.org/sites/default/files/Results%20Based%20Monitoring%20Evaluation%20Toolkit.pdf.

Gender Evaluation Methodology (GEM) for Internet and ICTs. MyM&E.
Retrieved from http://mymande.org/node/1248

Call for Feedback: How-To Note on a Framework for Evaluating the Impact of ICT Programs, World Bank. Vasdev, S. (2012).
Retrieved from http://blogs.worldbank.org/ic4d/call-for-feedback-how-to-note-on-a-framework-for-evaluating-the-impact-of-ict-programs.

Guide to Measuring Information and Communication Technologies (ICT) in Education. UNESCO (2009).
Retrieved from http://www.uis.unesco.org/Library/Documents/ICT_Guide_EN_v19_reprintwc.pdf.

How To Measure Social Media: A Quick and Easy Guide to Choosing Metrics, The Measurement Standard. Anonymous. (2007).
Retrieved from http://kdpaine.blogs.com/themeasurementstandard/2007/06/how_to_measure_.html.

How-to Guides, Primers & Reference Documents

Monitoring and Evaluation, Bond. Gosling, L. (2013).
Retrieved from http://www.bond.org.uk/data/files/resources/50/Monitoring-and-evaluation-How-To-guide-June-2013.pdf

A Guide to Actionable Measurement. Bill & Melinda Gates Foundation. (2010).
Retrieved from https://docs.gatesfoundation.org/Documents/guide-to-actionable-measurement.pdf.

A Developmental Evaluation Primer. Gamble A.A., J. (2008).
Retrieved from http://tamarackcommunity.ca/downloads/vc/Developmental_Evaluation_Primer.pdf

Measuring Internet Activity: A (Selective) Review of Methods and Metrics. Faris, R. & Heacock, R. (2013).
Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2353457.

Mobile-based technology for monitoring and evaluation, CLEAR, et. al. Mitesh Thakkar, et. al. (2013).
Retrieved from http://www.seachangecop.org/node/1989#!.

Core Concepts in Developing Monitoring and Evaluation Frameworks. Markiewicz, A. (2013).
Retrieved from http://betterevaluation.org/sites/default/files/ME%20Framework%20Resource%20Guide%20Jan%202013.pdf.

The Community Builder's Approach to Theory of Change: A Practical Guide to Theory Development, The Aspen Institute. Anderson, Andrea A. (2013).
Retrieved from http://www.aspeninstitute.org/sites/default/files/content/docs/rcc/rcccommbuildersapproach.pdf.

#SMMStandards: A Cross-Industry Effort to Simplify and Unify the Measurement of Social Media, The Conclave. Marklein, T. & AMEC. (2013).
Retrieved from http://www.smmstandards.com/wp-content/uploads/2013/06/Complete-standards-document4.pdf

Social Media Monitoring Tools. Robert Wood Johnson Foundation. (2013).
Retrieved from http://www.rwjf.org/content/dam/files/rwjf-web-files/Resources/1/SMMonitoring_2013.pdf

Global eHealth – Measuring Outcomes: Why, What, and How, WHO. Scott, Richard E. & Saeed, A. (2008).
Retrieved from http://www.ehealth-connection.org/files/conf-materials/Global%20eHealth%20-%20Measuring%20Outcomes_0.pdf.

Good, But How Good? Monitoring and Evaluation of Media Assistance Projects, CIMA. Mosher, A. (2009).
Retrieved from http://cima.ned.org/sites/default/files/CIMA-Monitoring_and_Evaluation-Report.pdf.

Ubaldi, B. (2013). "Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives", OECD Working Papers on Public Governance, No. 22, OECD Publishing. doi: 10.1787/5k46bj4f03s7-en

Event Report: Next Generation Evaluation: Embracing Complexity, Connectivity, and Change. Includes recorded speeches at conference hosted by FSG and Stanford Social Innovation Review.
http://www.fsg.org/nextgenerationevaluation.aspx.

Communities & Online Resources

The Big Push Forward (http://bigpushforward.net/)
An effort by an informal international network of practitioners seeking discussion, debate and the exploration of appropriate approaches for assessing transformative development processes

Innovation Net (http://www.innonet.org/)
A nonprofit evaluation, research, and consulting firm that helps nonprofits and funders learn from their work to improve their results.

Better Evaluation (www.betterevaluation.org)
An international collaboration to improve evaluation practice and theory by sharing information about options (methods or tools) and approaches.

My M&E (www.mymande.org)
Interactive learning resource to share knowledge on country-led M&E systems worldwide – E Learning based.

Intelligent Measurement (http://intelligentmeasurement.net/)
Blog that focuses on evaluation and measurement in communications, training, management and other fields.

TheoryofChange.org (http://www.theoryofchange.org/)
Non-profit organization established to promote quality standards and best practice for the development and implementation of Theory of Change, with a particular focus on its use and application in the areas of international development, sustainability, education, human rights and social change.

INTRAC (http://www.intrac.org/)
An organization providing research, training, consultancies and programmes for Civil Society Organizations.

GSDRC Topic Guide on Measuring Results (http://www.gsdrc.org/go/topic-guides/measuring-results/applying-monitoring-and-evaluation-tools)
Global Social Development Resource Centre has produced a Topic Guide on Applying Monitoring and Evaluation Tools.

Online Tools and Tool Catalogues

Impact Story (http://impactstory.org/)
Finds where assorted work has been cited, viewed, downloaded, tweeted, calculates impact and shares the impacts of the articles, slides, datasets, and software etc.

Chart Beat (https://chartbeat.com/)
Real-time analytics tool that analyze how users interact on websites, focusing on unusual behaviors rather than simple click metrics standard with most web analytics options.

Tools and Resources for Assessing Social Impact (http://trasi.foundationcenter.org/)
Catalog of tools for assessing social impact created by Foundation Center.

Social Media Monitoring Tools (http://www.rwjf.org/content/dam/files/rwjf-web-files/Resources/1/SMMonitoring_2013.pdf)
Sampling of tools available to produce social media metrics to be used as a method for assessing impact.

IRIS (http://iris.thegiin.org/)
Catalog of generally-accepted performance metrics that leading impact investors use to measure social, environmental, and financial success, evaluate deals, and grow the credibility of the impact investing industry.

SPLUNK (http://www.splunk.com/view/SP-CAAAG6E)
Online tool for collecting and analyzing social media and big data. Splunk offers Nonprofit Licensing Options.

TOCO Software (http://www.theoryofchange.org/toco-software/#5)
Theory of Change Online or "TOCO", is an accessible, easy-to-use learning tool for creating and implementing a Theory of Change. It provides users with a flexible drawing canvas for building, editing, and soliciting feedback.

Relevant Examples of Evaluations

Reimagining Governance in Practice: Benchmarking British Columbia's Citizen Engagement, GovLab. Young, Andrew. , Rogawski, C. , Rahman, S. & Verhulst, S. (2013).
Retrieved from https://s3.amazonaws.com/images.thegovlab.org/wordpress/wp-content/uploads/2013/10/Reimagining+Governance+in+Practice+%281%29.pdf

Global Focus: Haiti Creative Problem Solving in Complex Systems, OpenBox. Welaratna, D. & Choko, G. (2013).
Retrieved from http://opnbx.net/haiti/haiti_report.pdf.

List of ICT evaluations at the International Initiative for Impact Evaluation
http://www.3ieimpact.org/en/evidence/impact-evaluations/?sector=Information+and+Communications+Technology

Literature and Works Cited

Anonymous (2007). How To Measure Social Media: A Quick and Easy Guide to Choosing Metrics, The Measurement Standard.
Retrieved from http://kdpaine.blogs.com/themeasurementstandard/2007/06/how_to_measure_.html.

Batchelor, S. & Norrish, P. (2005). Framework for the assessment of ICT pilot projects Beyond Monitoring and Evaluation to Applied Research, InfoDev.
Retrieved from http://www.seachangecop.org/sites/default/files/documents/2005%2005%20%20Framework%20for%20assessment%20of%20ICT%20pilot%20projects.pdf.

Better Evaluation. (2014). Equal Access Participatory Monitoring and Evaluation Toolkit.
Retrieved from http://betterevaluation.org/toolkits/equal_access_participatory_monitoring.

Bill & Melinda Gates Foundation. (2010). A Guide to Actionable Measurement.
Retrieved from https://docs.gatesfoundation.org/Documents/guide-to-actionable-measurement.pdf.

Collier, D. (2010). Process Tracing: Introduction and Exercises.
Retrieved from http://dmeforpeace.omnidev3.com/sites/default/files/Collier_Process%20Tracing.pdf

Collier, D. (2011). Understanding Process Tracing.
Retrieved from http://www.ukcds.org.uk/sites/default/files/uploads/Understanding-Process-Tracing.pdf

Croll, A. & Yoskovitz, B. (2013). Lean Analytics: Use Data to Build a Better Startup Faster. O'Reilly Media, Inc.

Eyben, R. (2013). Uncovering the Politics of 'Evidence' and 'Results' A Framing Paper For Development Practictioners.
Retrieved from http://bigpushforward.net/wp-content/uploads/2011/01/Uncovering-the-Politics-of-Evidence-andResults-by-Rosalind-Eyben.pdf

Fine, Allison H. (2013). Evaluating the Impact of Information Technology.
Retrieved from http://www.innonet.org/client_docs/File/Eval_tech.pdf

Fodden, S. (2011). The Anatomy of a Tweet: Metadata on Twitter.
Retrieved from http://www.slaw.ca/2011/11/17/the-anatomy-of-a-tweet-metadata-on-twitter/

Gamble A.A., J. (2008). A Developmental Evaluation Primer.
Retrieved from http://tamarackcommunity.ca/downloads/vc/Developmental_Evaluation_Primer.pdf

Goldstein, M. (2012, July 3). Notes from the field: Making the most of monitoring data.
Retrieved from http://blogs.worldbank.org/impactevaluations/node/835

Gopalakrishnan, S. , Preskill, H. & Lu, SJ. Next Generation Evaluation: Embracing Complexity, Connectivity, and Change.
Retrieved from http://www.fsg.org/tabid/191/ArticleId/964/Default.aspx?srpush=true.

Gosling, L. (2013). Monitoring and Evaluation, Bond.
Retrieved from http://www.bond.org.uk/data/files/resources/50/Monitoring-and-evaluation-How-To-guide-June-2013.pdf

Granickas, K. (2013). Understanding the Impact of Releasing and Re-Using Open Government Data.
Retrieved from http://www.epsiplatform.eu/sites/default/files/2013-08-Open_Data_Impact.pdf.

Guijt, I. (2008). Critical Readings on Assessing and Learning for Social Change: A Review, IDS.
Retrieved from http://bigpushforward.net/wp-content/uploads/2011/05/guijt-assessing-social-change.pdf

Hearn, S. , Leborgne, E. & Brown, Valerie A. (2011). Monitoring and evaluating development as a knowledge industry: ideas in current practice, ODI.
Retrieved from http://wiki.ikmemergent.net/files/IKM_Working_Paper_No_12_Aug_2011.pdf.

Holland, J. , Thirkell, A. , Trepanier, E. & Earle, L. (2009). Measuring Change and Results in Voice and Accountability Work, DFID.
Retrieved from http://www.oecd.org/derec/unitedkingdom/44463612.pdf

Jones, H. (2011). Guide to monitoring and evaluating policy influence, ODI.
Retrieved from http://www.odi.org.uk/sites/odi.org.uk/files/odi-assets/publications-opinion-files/6453.pdf.

Kanter, B. (2013). What is your organization's "One Metric That Matters"?.
Retrieved from http://www.bethkanter.org/performance-management/

Karr, D. (2013). 13 Tools for Hashtag Research and Management.
Retrieved from http://www.marketingtechblog.com/hashtag-research-tools/

Kirkland, K. (2013). How to Measure a Successful Tech for Good Venture, Nominet Trust.
Retrieved from http://www.nominettrust.org.uk/knowledge-centre/blogs/how-to-measure-successful-tech-good-venture.

Kramer, M. , Parkhurst, M. & Vaidyanathan, L. (2011). Breakthroughs in Shared Measurement and Social Impact, FSG.
Retrieved from http://www.fsg.org/tabid/191/ArticleId/87/Default.aspx?srpush=true

Leborgne, E. , Brown, Valerie A. & Hearn, S. (2011). Monitoring and evaluating development as a knowledge ecology: ideas for new collective practices.
Retrieved from http://wiki.ikmemergent.net/files/IKM_Working_Paper_No_13_Aug_2011.pdf.

Lennie, J. (2005). Building community capacities in evaluating rural IT projects: Success strategies from the LEARNERS Project.
Retrieved from http://academic.research.microsoft.com/Publication/6835119/building-community-capacities-in-evaluating-rural-it-projects-success-strategies-from-the-learners

LFA Group. (2013). Advocacy Evaluation Mini-Toolkit: Tips and Tools for Busy Organizations.
Retrieved from http://www.lfagroup.com/wp/wp-content/uploads/2013/04/Advocacy-Evaluation-Mini-Toolkit.pdf

Markiewicz, A. (2013). Core Concepts in Developing Monitoring and Evaluation Frameworks.
Retrieved from http://betterevaluation.org/sites/default/files/ME%20Framework%20Resource%20Guide%20Jan%202013.pdf.

Marklein, T. & AMEC. (2013). #SMMStandards: A Cross-Industry Effort to Simplify and Unify the Measurement of Social Media, The Conclave.
Retrieved from http://www.smmstandards.com/wp-content/uploads/2013/06/Complete-standards-document4.pdf

Mayne, J. (2008). Contribution Analysis: An approach to exploring cause and effect, ILAC.
Retrieved from http://www.cgiar-ilac.org/files/publications/briefs/ILAC_Brief16_Contribution_Analysis.pdf.

Mayne, J. (2011). Addressing Cause and Effect in Simple and Complex Settings through Contribution Analysis. In Evaluating the Complex, R. Schwartz, K. Forss, and M. Marra (Eds.), Transaction Publishers.

McGee, R. and Gaventa, J. (2013). The Impact and Effectiveness of Transparency and Accountability Initiatives, Overseas Development Institute.
Retrieved from http://www.ids.ac.uk/publication/the-impact-and-effectiveness-of-transparency-and-accountability-initiatives

Mitesh Thakkar, et. al. (2013). Mobile-based technology for monitoring and evaluation, CLEAR, et. al.
Retrieved from http://www.seachangecop.org/node/1989#!.

Moehler, Devra C. (2013). Democracy, Governance and Randomised Media Assistance, Annenberg School for Communication.
Retrieved from https://downloads.bbc.co.uk/rmhttp/mediaaction/pdf/Democracy_gov_working_paper_4.pdf.

Morariu, J. , Athanasiades, K. & Emery, Ann. (2012). State of Evaluation 2012: Evaluation Practice and Capacity in the Nonprofit Sector.
Retrieved from http://www.innonet.org/client_docs/innonet-state-of-evaluation-2012.pdf.

Mosher, A. (2009). Good, But How Good? Monitoring and Evaluation of Media Assistance Projects, CIMA.
Retrieved from http://cima.ned.org/sites/default/files/CIMA-Monitoring_and_Evaluation-Report.pdf.

O'Neil, G. (2012, January 6). A pragmatic guide to monitoring and evaluating research communications using digital tools [Msg 1522].
Retrieved from http://onthinktanks.org/2012/01/06/monitoring-evaluating-research-communications-digital-tools

Pade-Khene, C., & Sewry, D. (2011, April). Towards a Comprehensive Evaluation Framework for ICT for Development Evaluation–An Analysis of Evaluation Frameworks. In 2nd International Conference on Information Management and Evaluation. Toronto, Canada.

Papendieck, A. (2011, April 12). Final Report: Independent Evaluation of the Ushahidi Haiti Project.
Retrieved from https://sites.google.com/site/haitiushahidieval/news/finalreportindependentevaluationoftheushahidihaitiproject

Pratt, B. (2012, 25 April). Obsessive Compulsive Measurement Disorder.
Retrieved from http://www.intrac.org/blog.php/22/obsessive-compulsive-measurement-disorder

Roche, C. and Kelly, L. (2012) "The Evaluation of Politics and the Politics of Evaluation" DLP Background Paper 11.
Retrieved from http://www.dlprog.org/news-events/advocating-a-mixed-methods-approach-to-the-evaluation-of-politics.php.

Robert Wood Johnson Foundation. (2013). Social Media Monitoring Tools.
Retrieved from http://www.rwjf.org/content/dam/files/rwjf-web-files/Resources/1/SMMonitoring_2013.pdf

Rogers, C. (2010, November 22). USAID: Obsessive measurement disorder?.
Retrieved from http://devpolicy.org/usaid-obsessive-measurement-disorder20101122/

Russell, J. & Greenhalgh, T. (2011). Why Do Evaluations of eHealth Programs Fail? An Alternative Set of Guiding Principles, PLoS Medicine.
Retrieved from http://www.comminit.com/global/content/why-do-evaluations-ehealth-programs-fail-alternative-set-guiding-principles.

Scott, Richard E. & Saeed, A. (2008). Global eHealth – Measuring Outcomes: Why, What, and How, WHO.
Retrieved from http://www.ehealth-connection.org/files/conf-materials/Global%20eHealth%20-%20Measuring%20Outcomes_0.pdf.

Shutt, C. (2011). Monitoring, Evaluating & Assessing the Impact of Governance Programmes: Summary of Literature and Practice Review.
Retrieved from http://governance.care2share.wikispaces.net/file/view/Governance+M%26E+Literature+Review+Final+6.10.11.pdf

Shutt, C. & McGee, R. (2013). Improving the Evaluability of INGO Empowerment and Accountability Programmes.
Retrieved from http://bigpushforward.net/wp-content/uploads/2011/01/CDIPracticePaper01finalR.pdf.

Sommerfeldt, E. (2012). Can Shared Indicators Improve Monitoring and Evaluation?, CIMA.
Retrieved from http://cima.ned.org/sites/default/files/final_0.pdf.

Spreckley, F. (2009). Results Based Monitoring and Evaluation Toolkit, Local Livelihoods.
Retrieved from http://mymande.org/sites/default/files/Results%20Based%20Monitoring%20Evaluation%20Toolkit.pdf.

Starling, S. (2010). Monitoring and evaluating advocacy: lessons from Oxfam GB's Climate Change campaign. Development in Practice, 20(2), 277-286.
http://www.tandfonline.com/doi/pdf/10.1080/09614520903564215

Stern, E. , Stame, N. , Mayne, J. , Forss, K. , Davies, R. & Befani, B. (2012). Broadening the Range of Designs and Methods for Impact Evaluations, DFID.
Retrieved from http://r4d.dfid.gov.uk/pdf/outputs/misc_infocomm/DFIDWorkingPaper38.pdf.

The World Bank. Beneficiary Assessment.
Retrieved from http://go.worldbank.org/3AIUQJ5WP0

The World Bank. (2004). Monitoring & Evaluation: Some Tools, Methods and Approaches.
Retrieved from http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2006/02/15/000012009_20060215093620/Rendered/PDF/246140UPDATED01s1methods1approaches.pdf

Twersky, F. & Lindblom, K. (2012). Evaluation Principles and Practices.
Retrieved from http://www.hewlett.org/uploads/documents/EvaluationPrinciples-FINAL.pdf

Twersky, F. & Nelson, J. & Ratcliffe, A. (2010). A Guide to Actionable Measurement.
Retrieved from https://docs.gatesfoundation.org/Documents/guide-to-actionable-measurement.pdf

Ubaldi, B. (2013). "Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives", OECD Working Papers on Public Governance, No. 22, OECD Publishing. doi: 10.1787/5k46bj4f03s7-en

UNESCO (2009). Guide to Measuring Information and Communication

Technologies (ICT) in Education.
Retrieved from http://www.uis.unesco.org/Library/Documents/ICT_Guide_EN_v19_reprintwc.pdf.

USAID Center for Development Information and Evaluation. (1996). Performance Monitoring and Evaluation TIPS Conducting a Participatory Evaluation.
Retrieved from http://preval.org/documentos/00487.pdf

Vasdev, S. (2012). Call for Feedback: How-To Note on a Framework for Evaluating the Impact of ICT Programs, World Bank.
Retrieved from http://blogs.worldbank.org/ic4d/call-for-feedback-how-to-note-on-a-framework-for-evaluating-the-impact-of-ict-programs.

van Zyl, A. (2011). IBP's Super Duper Impact Planning Guide, IBP.
Retrieved from http://internationalbudget.org/wp-content/uploads/Super-Duper-Impact-Planning-Guide.pdf.

Welaratna, D. & Choko, G. (2013). Global Focus: Haiti Creative Problem Solving in Complex Systems, OpenBox.
Retrieved from http://opnbx.net/haiti/haiti_report.pdf.

Whitty, B. (2013). Experiences of the Results Agenda Draft Findings for Discussion from the Crowd-Sourcing Survey.
Retrieved from http://bigpushforward.net/wp-content/uploads/2011/01/Experiences-of-the-Results-Agenda-by-Brendan-Whitty.pdf.

Young, A. , Rogawski, C. , Rahman, S. & Verhulst, S. (2013). Reimagining Governance in Practice: Benchmarking British Columbia's Citizen Engagement, GovLab.
Retrieved from https://s3.amazonaws.com/images.thegovlab.org/wordpress/wp-content/uploads/2013/10/Reimagining+Governance+in+Practice+%281%29.pdf.

Acknowledgements

This guide was originally conceived as a report to document best practices and useful frameworks for M&E of tech and accountability initiatives. It quickly became clear, that there was very little good practice out there, and almost no well-documented case studies.[32] This is at least in part to the fact that tech for accountability initiatives tend to operate with very limited resources and monitoring and evaluation does not always get prioritized. It also became very quickly clear, however, that tech for accountability initiatives are not significantly different than other types of projects when it comes to monitoring and evaluation. In fact, they tend to face some of the most challenges aspects of traditional monitoring (such as attribution and advocacy impact), and are further burdened with unrealistic expectations about "all the hard data" technology produces.

In the course of grappling with these challenges, faced with a lack of established practice and frameworks, and listening to what experts and project workers thought would be most useful, this report morphed slowly into a Guide intended to help initiatives develop their own good practice.

It has benefited from the comments and guidance of many along the way.

Thanks to:

Adam Groves, Bond UK
Anne Garbutt, Intrac
Brendan Halloran, TABridge
Heather Leson, Open Knowledge Foundation
Irene Guijt, The Big Push Forward
Isabel Vogt, independent consultant
Jed Miller, independent consultant
Jennifer Chan, Harvard University
June Wang, Hewlett Foundation
Kasia Kedzia, Social Impact
Kenny Meesters, Propus
Laura Walker Hudson, Social Impact Labs
Linda Raftree, Independent consultant
Mary Joyce, University of Washington
Rebecca Chiao, HarassMap
Sean McDonald, Frontline SMS
Theo Keane, Nesta
Varja Lipovsek, Twaweza

32 A notable exceptions to this is Oxfam GB's advocacy, see Starling, S. (2010).

Glossary

Adaptive management: Refers to an iterative process of decision-making with the aim to lessen uncertainty through system monitoring.

Baseline: This refers to a measure of a project's context against which to assess new improvements. Baselines are used as a logical basis for comparison.

Citizen Reports: Often juxtaposed with formal media outlets, these reports represent accounts from public citizens "on the scene" and are often employed to draw attention to political, social or environmental crises, or comment upon local problems, such as transportation. There are countless potential uses for citizen reporting, but often they function as a public watchdog on government behavior, advocating accountability and transparency.

Developmental evaluation: Refers to a method that combines the adaptive and iterative aspects of measurement on-the-go, with the assessment dynamics of an external evaluation. It is designed to track and support program innovation in situations of high complexity.

Dual-use Data: Dual-use data refers to any data which can satisfy more than one goal at any given time. In the case of this guide, it is most often data that functions as tool for monitoring and evaluating the project as well as a data source for backing up campaigning and lobbying efforts. A rigorous framework for data collection and analysis can enable powerful dual use data.

Ex-Ante: Formative Evaluations (measurement on-the-go), is typically applied to pilot projects, especially where there is a significant degree of unknown factors or causal relationships, to provide an opportunity to reconsider and adjust project strategies during implementation (adaptive management).

Ex-Post: Also known as Summative Evaluation, this term refers to an evaluation that takes place after projects are completed, and may be conducted by project teams, but are often conducted by external consultants. Summative evaluations generally aim for a holistic understanding what context and process factors contributed to project successes or failures. They often target donors and researchers.

Client Relationship Management platform (CRM): A database and interface system for managing information about customers. In advocacy, such platforms are often used to maintain information about colleagues, policy makers, recipients of mailing lists or other target groups. CiviCRM (**https://civicrm.org/**) is an open source CRM intended for use by non-profits.

Interactive Voice Response: A technology that allows users to interact with surveys or other question and answer formats on their phones, either

by speaking or using touch tones. Kubutana's Freedom Fone (**http://www.freedomfone.org/**) is one of the more notable examples of this technology being used for advocacy.

Formative Monitoring: Monitoring that informs decision-making and adaptive management on an on-going basis. In the world of measurement, this is distinct from "summative" monitoring and evaluation systems, which aim to summarize what h as worked and what has not, and which are primarily tools for accountability and review after a project or project phase is complete.

Harmonization: The process of comparing two or more data component definitions and identifying commonalities among them that warrant their being combined, or harmonized, into a single data component. This often occurs when different segments of the same project produce similar but incompatible indicators. Harmonizing these indicators can allow for more powerful diagnostics across the organization, and make data management a lot easier.

Iterative Process: An Iterative process in a project context may refer to the technique of developing and delivering incremental components of project, integrating new information and technique as the project develops.

Mapping data: Mapping refers to the action of collating all information into a usable format. More specifically, mapping data can mean gathering, assigning and examining values of various data in order to make better decisions about monitoring frameworks and existing resources.

Metadata: Defined as the data providing information about one or more aspects of the data, such as: the means of creation of the data, the purpose of the data, time and date of creation, creator or author of the data, location on a computer network where the data were created, and/or standards used. Metadata provides a tremendous amount of information about the people on social networks that are following or engaging in campaigns and can be useful for monitoring projects that already utilize social media data. In the context of this guide, metadata is the most useful for mapping automatically generated data (See Step 5).

"pressure to measure": The expression refers to the pressure to document project results often due to international austerity and the increased pressure on bilateral and multilateral donors for international development aid. This "pressure to measure" is experienced by many small projects and recipients of grants.

Purposive/purposive sample: respondents to a survey are not randomly selected, but rather selected purposively, because they fit a certain criteria. For example, interviewing stakeholders at locations such as a clinic for a health care survey.

Representativity/representative sample: A loose statistical concept indicating that data points in a sample have the same characteristics as a larger population. In other words, if a survey of Guatemalan ice cream preferences shows that 75% of respondents prefer chocolate, it is representative of Guatemalans only if 75% of all Guatemalans prefer chocolate. There are several elaborate methodologies for pursuing representivity, which generally involve randomly selecting survey participants and observing stringent selection rules. Obviously, the survey example above would not be representative if it was sent to people who often buy chocolate ice cream. For similar reasons, crowdsourced data, or citizen reports, will often not be representative.

Social Network Analysis: An analysis that views social relationships in terms of network theory, consisting of nodes (representing individual actors within the network) and ties (which represent relationships between the individuals, such as friendship, kinship, organizations, sexual relationships, etc.). Network analysis can also provide useful information about the relationships between people following or participating in a online campaign.

Tech and Accountability Projects: Any project, large or small, that is using technology in some meaningful way to work towards improved governance, transparency or accountability.

Theory of Change: A set of connected building blocks that provides a graphic representation of the change process that, when taken together, are required to bring about a given long-term goal.

Use Case: Borrowed from software development, this term refers to a specific description of how a tool or piece of information will be used, including details about the person using it, their motivations for doing so, and what they require to do so effectively.

This guide presents a series of steps that tech and accountability initiatives can use to develop monitoring frameworks. It will work as a comprehensive roadmap or as an *a-la-carte* menu. It is not comprehensive, but provides an introduction to a number of relevant methods, tools and strategies, with links and recommendations for further exploration.