Evaluating Digital Citizen Engagement

A PRACTICAL GUIDE

WORLD BANK GROUP
Governance

DIGITAL ENGAGEMENT
Evaluation Team | DEET

Evaluating Digital Citizen Engagement

A PRACTICAL GUIDE

February 2016

APTIVATE
www.aptivate.org

IDS Institute of Development Studies
www.ids.ac.uk

ica:uk
www.ica-uk.org.uk

© 2016 International Bank for Reconstruction and Development / The World Bank

1818 H Street NW,Washington, DC 20433
Telephone: 202-473-1000
Internet: www.worldbank.org

Some rights reserved
1 2 3 4 18 17 16 15

This work is a product of the staff of The World Bank with external contributions. The findings, interpretations, and conclusions expressed in this work do not necessarily reflect the views of The World Bank, its Board of Executive Directors, or the governments they represent. The World Bank does not guarantee the accuracy of the data included in this work. The boundaries, colors, denominations, and other information shown on any map in this work do not imply any judgment on the part of The World Bank concerning the legal status of any territory or the endorsement or acceptance of such boundaries.

Nothing herein shall constitute or be considered to be a limitation upon or waiver of the privileges and immunities of The World Bank, all of which are specifically reserved.

Rights and Permissions

This work is available under the Creative Commons Attribution 3.0 IGO license (CC BY 3.0 IGO) http://creativecommons.org/licenses/by/3.0/igo. Under the Creative Commons Attribution license, you are free to copy, distribute, transmit, and adapt this work, including for commercial purposes, under the following conditions:

Translations—If you create a translation of this work, please add the following disclaimer along with the attribution: This translation was not created by The World Bank and should not be considered an official World Bank translation. The World Bank shall not be liable for any content or error in this translation.

Adaptations—If you create an adaptation of this work, please add the following disclaimer along with the attribution: This is an adaptation of an original work by The World Bank. Views and opinions expressed in the adaptation are the sole responsibility of the author or authors of the adaptation and are not endorsed by The World Bank.

Third-party content—The World Bank does not necessarily own each component of the content contained within the work. The World Bank therefore does not warrant that the use of any third-party-owned individual component or part contained in the work will not infringe on the rights of those third parties. The risk of claims resulting from such infringement rests solely with you. If you wish to re-use a component of the work, it is your responsibility to determine whether permission is needed for that re-use and to obtain permission from the copyright owner. Examples of components can include, but are not limited to, tables, figures, or images.

All queries on rights and licenses should be addressed to the Publishing and Knowledge Division, The World Bank, 1818 H Street NW, Washington, DC 20433, USA; fax: 202-522-2625; e-mail: pubrights@worldbank.org.

Table of Contents

Foreword

Digital citizen engagement can be a compelling approach to increasing citizen participation and voice. However, it is not a silver bullet and does not work in all situations. The capacity to evaluate and generate knowledge is critical to the success of the field, though there are few benchmarks available to evaluate digital citizen engagement methods and the outcomes they produce.

As part of the ongoing effort to better understand digital citizen engagement, the Governance Global Practice commissioned this practical guide on evaluating digital citizen engagement. This guide is structured in a way that allows for cumulative learning and builds on pre-existing knowledge in three fields—technology, citizen engagement and evaluation. While this is not the first guide to evaluate projects at the intersection of technology and citizen engagement, it is structured in a broad manner that draws on international experts, practitioners and literature in these fields.

In recent years, the World Bank Group (WBG) has structured its knowledge and experience in these fields to focus on how and under what conditions these approaches improve development results. The Strategic Framework for Mainstreaming Citizen Engagement in WBG Operations adopted in 2014 offers a roadmap for the ways we as an institution mainstream and scale up citizen engagement across our operations to improve results. In a similar vein, the 2016 World Development Report (WDR) examines how digital technologies can be a force for development by generating economic growth, social and economic opportunity, and greater efficiency of public service delivery.

This guide serves as a useful starting point for those who seek to evaluate digital engagement efforts and contribute to cumulative learning.

Jeff Thindwa
Practice Manager, Open and Collaborative Governance

Acknowledgements

Evaluating Digital Citizen Engagement: A Practical Guide was commissioned and overseen by Tiago Peixoto, Marieta Fall and Fredrik Sjoberg – respectively task team leaders and lead researcher of the Digital Engagement Evaluation Team (DEET) – and funded by the Global Partnership Facility (GPF).

The guide was made possible through a partnership comprising Aptivate, the Institute of Development Studies, and ICA: UK. It was developed and written by Matt Haikin (lead author), Savita Bailur, Evangelia Berdou, Jonathan Dudding, Cláudia Abreu Lopes, and Martin Belcher. The authors are thankful for the valuable support received from Sara Gwynn, Claire Hughes, Nathaniel Whitestone, Patricia Haikin and Elly Boulboul.

Additional contributions and quotations were provided by Duncan Green (Oxfam), Lily Tsai (MIT Governance Lab), Rebecca Rumbul (mySociety), Peter John (University College London), Tim Unwin (University of London), Linda Raftree (Kurante), Renee Wittemyer (Intel), Adebusoye Anifalaje (BAO/HISP), Thomas Bjelkeman (Akvo), Carla Bonina, (ILDA), Robert Chambers (IDS), Levi Goertz (VOTO Mobile), Leslie Groves Williams and Eleanor Harrison (Global Giving UK), Eric Meerkamper (RIWI Corporation), Amy O'Donnell (Oxfam UK),Valerie Oliphant (SIMLab), Nimmi Rangaswamy (Indian Institute of Technology), Loren Triesman (Indigo Trust), and Leah Rosenzweig (MIT Governance Lab).

The team is thankful to those who accepted DEET's invite to offer input at different stages throughout the development of the guide: Jonathan Fox (American University), Raul Zambrano (UNDP), Simon Dalferth (Council of the European Union), Susan Stout (Georgetown University), Vonda Brown (Open Society Foundations) Alexander Furnas (Sunlight Foundation), Roberto Rocha Coelho Pires (Institute for Applied Economic Research), and Anthony Zacharzewski (The Democratic Society). We are also grateful to those who provided input as staff of the World Bank Group including Astrid Manroth, Bjorn-Soren Gigler, Elena Georgieva-Andonovska, Harika Masud, Helene Grandvoinnet, Mary McNeil, Michael Woolcock, Stephen Davenport and Utpal Misra, and in particular to DEET members Jonathan Mellon (Senior Data Scientist), Daniel Nogueira-Budny (researcher) Sanna Maria Ojanpera (researcher), and Kate Sunny Henvey (researcher), who also contributed with the final editing of this guide. The team is also thankful to Claudio Mendonca for the graphic design.

Thank you to the many leaders of our institution that have championed digital engagement and cumulative learning, notably Mario Marcel, former Senior Director, and Robert Hunja and Hassane Cisse, Directors, in the WBG's Governance Global Practice.

Acronyms and Abbreviations

CAQDA	Computer-assisted Qualitative Data Analysis
CDR	Call Data Record (from Mobile Network Operator)
CE	Citizen Engagement
CSO	Civil Society Organization
CSV	Comma Separated Value (file type for sharing data between systems)
DCE	Digital Citizen Engagement
GPS	Global Positioning System
GRM	Grievance Redress Mechanism
ICT	Information and Communication Technology
ICT4D	Information and Communication Technologies for Development
IVR	Interactive Voice Response
LFA	Logical Framework Approach
MNO	Mobile Network Operator
NGO	Non-governmental Organization
PB	Participatory Budgeting
PDO	Project Development Objective (within World Bank)
PII	Personally Identifiable Information
R	A programming language and software environment for statistical analysis
RCT	Randomized Control Trial
RDIT	Random Domain Intercept Technology
SAS	Statistical Analysis Software
SMS	Short Message Service (text messages)
SPSS	Statistical Package for the Social Sciences
STATA	STAtistics and dATA (analysis software)
TIC	Technologie de l'information et de la communication (ICT)
TTL	Task Team Leader (within World Bank)
USSD	Unstructured Supplementary Service Data (mobile messaging protocol)
WASH	WAter, Sanitation and Hygiene
WB	World Bank
WBG	World Bank Group
WBI	World Bank Institute

Tables

Executive Summary

With growing demand for transparency, accountability and citizen participation in policy making and service provision, engagement between citizens and their governments, as well as with donors and the private sector that deliver government services, is increasingly important. Increased use of technology brings both opportunities and challenges to citizen engagement processes, including opportunities for collecting, analyzing and evaluating data about these processes. This guide provides practical steps to assess the extent to which digital tools have contributed to citizen engagement and the help to understand the impact that the introduction of technology has had on the engagement processes.

With examples and lessons from case studies from Brazil, Uganda, Cameroon and Kenya, the guide provides practical tools and guidelines for use in evaluating the expanding field of digital citizen engagement (DCE). This guide can be used at many stages— to inform project design, as a tool for continual learning and improvement, and for undertaking mid-term or post-hoc evaluations. Written primarily for practitioners—including task team leaders (TTLs) at the World Bank Group (WBG), project or program delivery staff at Civil Society Organizations (CSOs), and internal or external evaluators or consultants throughout the project cycle—the guide is also a helpful resource for anyone seeking to better understand the role of digital technology in citizen engagement.

For the purposes of the guide, Digital Citizen Engagement (DCE) is defined as **the use of new media/digital information and communication technologies to create or enhance the communication channels that facilitate the interaction between citizens and governments or the private sector.** DCE may be used to improve decision making through more effective citizen participation, representation and expression, or to improve intermediate and final development outcomes (See Appendix B: Results Indicators for Citizen Engagement). In order to explore the full and nuanced range of perspectives involved, the guide presents five lenses— different perspectives through which DCE interventions might be viewed while undertaking an evaluation. **Table 6** (from **Section 3.3**) is included below, showing each lens and some key questions that they raise.

Question	Evaluation
OBJECTIVE	
What are the goals of the initiative, and how well is the project designed to achieve those goals?	Clarify the goals and planned changes, assessing the existence and appropriateness of those stated goals.

Question	Evaluation
CONTROL	
Which actors exert the most influence over the initiative's design and implementation, and what are the implications of this?	Explore the levels of influence on the engagement process, the dynamics of decision making, and levels of fairness and equitability among citizens.

Question	Evaluation
PARTICIPATION	
Which individuals participate in the initiative, and to what extent is their participation in line with their needs and expectations?	Examine who is included/excluded in the process, and how the program enables or discourages different opportunities for participation.

Question	Evaluation
TECHNOLOGY	
How appropriate was the choice of the technology, and how well was the technology implemented?	Take a practical look at the technology choices, technical delivery and management of the engagement process itself.

Question	Evaluation
EFFECTS	
What effects did the project have, and to what extent can this impact be attributed to technology?	Seek to understand the ultimate impact on citizens, government, collectives and service delivery/development outcomes.

Considering an evaluation through each lens can help to uncover important themes from the outset of the evaluation and used to directly inform the choice of evaluation questions and be used as guides throughout the evaluation. While recognizing that evaluation is necessarily an iterative process, the guide follows the stages of an evaluation lifecycle:

Scoping - This stage lays out the groundwork for the design and implementation of the evaluation by investing time and resources into understanding the project and its context, the operating environment and the recent developments and insights from the DCE evaluation field. This section is important for both commissioners and evaluators as it sets the parameters, develops focus and highlights opportunities for the evaluation itself and ensures that the evaluation process is suitably informed by and grounded in reality.

Designing - This stage builds on the information and knowledge gathered during the Scoping stage to begin the high-level and strategic design of the evaluation. This means agreeing the focus, goals and objectives, designing the evaluation questions, and deciding on an appropriate approach and method to achieve those goals in a way that is feasible and grounded in the reality of the project, the people involved and the wider environment. Detailed design decisions over subjects such as data collection are made in the subsequent Planning section.

Planning & Implementing - This section describes how the design process now moves to a more detailed level to decide what tools to use within the broad method for collecting new data, whether or not to use digital tools to collect new data, and how data collection can be implemented. Implementation of a DCE evaluation is broadly the same as for any evaluation so this is not covered in depth, but some specific tips are included that are of specific relevance to technology and citizen engagement.

Analyzing - This stage discusses how the DCE data can be analyzed and provides pointers for quantitative, qualitative and mixed methods of analysis. Challenges such as ensuring rigorous data and understanding bias are discussed, and suggestions offered as to how these can be addressed. It is also recognized that after initial analysis, there may be a need to iterate the process and re-visit the design or collect further data.

Sharing, Reflecting & Learning - This final section focuses on testing the findings, writing up the results and analysis of a DCE evaluation, considers methods of sharing findings (including discussing opening up evaluations and their data), and reflecting and learning on the lessons from evaluations.

Each section outlines the relevance and intent of the stage and the applicability of the five lenses and provides guidance and information on the key issues and provides an opportunity to self-assess progress before proceeding to the next section.

The guide offers two toolkits:

DCE evaluation bank: examples of primary (assessment/analysis) and supplementary (information gathering) evaluation questions—grouped by lens—and some 'satisfaction' questions. While these are not specifically evaluation questions, they may be useful in framing the evaluation questions for a particular context or as a starting point for conducting surveys or interviews with participants.**Using the lenses in scoping and design**: a set of considerations and questions that an evaluator might ask during the scoping and design stages, again grouped by lens.

Additional Appendices include further reading, global examples of DCE interventions, results indicators for citizen engagement and an overview of the Brazil, Uganda, Cameroon and Kenya case study data collection methods and costs.

The authors hope that readers of this Guide find it helpful in their work. It has been written as a guide to help practitioners identify questions to ask and methods to explore and is not intended to be followed rigidly or dogmatically. The quote from Duncan Green of Oxfam below captures its spirit and intent:

Guidance should open people's minds not close them down

People working in aid and development appreciate help – this is especially true for those seeking to evaluate relatively new areas of work such as Digital Citizen Engagement. Many of them don't take kindly to being told 'the world is complex, everything is context specific, so you're on your own'. The challenge is to design help and guidance so it can harness their initiative, imagination and appreciation of context.

There is a risk that advice and guidance on how to use technology for citizen engagement (and for many other areas) is turned into a checklist – a neat set of tick-boxes for what to do or not to do, closing down people's minds and restricting the options available to handle complex situations.

One way to avoid this is to limit guidance to "mind-openers" such as those included in this Guide – things like what kinds of questions to ask; sets of case studies that might spark new ideas, context-specific advice and reflections from mentors who have 'been there' themselves.

The aim should be to empower people to innovate, take risks, and learn from the inevitable failures.

Duncan Green
Strategic Adviser, Oxfam GB
oxfamblogs.org/fp2p

1. Introduction

With growing demand for transparency, accountability and citizen participation in policy making and service provision, engagement between citizens and government, as well as with donors and the private sector that deliver government services, is increasingly important.[1] Within this, the rapid proliferation of digital tools is opening up a new era of Digital Citizen Engagement (DCE). Initiatives such as online participatory budgeting, SMS voting and the use of handheld digital devices for beneficiary feedback are growing in use. Increased use of technology brings both opportunities and challenges to citizen engagement processes, including opportunities for collecting, analyzing and evaluating data about these processes.

This guide offers a means of assessing the extent to which digital tools have contributed to citizen engagement[2] and to help understand the impacts—positive or negative, intended or unintended—that the introduction of technology has had on the engagement processes. It addresses specific questions: Does adding digital technology to the citizen engagement process really provide quicker, cheaper, easier ways for citizens to engage with the state or other service providers? Can digital technologies lower interaction costs for governments and deliver improved, more targeted development outcomes? What risks come with this new technology—have certain citizens been excluded (intentionally or unintentionally) from the engagement process? Has the way in which people engage and communicate altered, for better or for worse? Has the technology affected the previously existing groups and institutions that were intermediating engagement processes before the technology was introduced?

The guide is designed to help people understand when the use of DCE is appropriate and under what circumstances, how to use it more effectively and what to expect from its use. It introduces the key issues relating to Digital Citizen Engagement and offers advice and guidance on how to evaluate it— including methods, indicators, challenges and course corrections that apply to the ***digital*** aspect of citizen engagement.

The guide complements the existing work on mainstreaming citizen engagement across the World Bank's operations (World Bank, 2014a) and seeks to add value to

1–Although the word 'citizen' is used throughout this guide, it is recognized that for some people (e.g. migrant populations or those living under certain non-democratic forms of governance), the beneficiaries, stakeholders, participants or end-users may technically not be citizens. Similarly the majority of this guide is applicable to any project seeking to improve communication and decision-making processes between groups of individuals and the institutions which affect them.

2–For the purposes of this guide, beneficiary feedback is viewed as a component of broader citizen engagement, with 'beneficiaries' being defined as *"a subset of citizens which are directly targeted by and expected to benefit from a development project"* (World Bank, 2014a, p4) and 'engagement' referring to consultation, collaboration and empowerment. In this way, all references to CE and DCE also refer to beneficiary feedback.

this work by focusing on those programs where technology plays a significant role in citizen engagement activity, and focusing on the evaluation of the effectiveness of both the citizen engagement activity overall, and the role of the technology within it. The value of methods such as Randomized Control Trials (RCTs), surveys and interviews are assessed, and the DCE-specific challenges for each of these are discussed, including appropriateness of the technology used, accessibility and quality of data and privacy.

As with citizen engagement, DCE needs to be looked at comprehensively in order for evaluators to fully understand the efficiency and the effectiveness of the program under review. To aid this, the guide introduces five lenses that provide a useful tool when scoping and designing evaluations by highlighting a range of perspectives relating to technology and citizen engagement, helping to ensure that important issues are not accidentally overlooked.

Additionally, recognizing there is much to be learned from previous experiences in CE and DCE, the guide incorporates a range of first-hand experiences. A variety of tips from a range of experts are provided in supplementary boxes throughout the text.

This guide was written in parallel with conducting a series of field evaluation of digital citizen engagement projects in Brazil, Uganda, Kenya and Cameroon. The learnings from these evaluations have been incorporated into this guide and reflect ***in situ*** on real-world programs.

While citizen engagement programs are not new phenomena, the major role that digital technology plays in these programs is relatively recent. DCE is a field that continues to develop rapidly as new technologies are introduced, and familiar technologies are used in new ways and as people continue to experiment with what is possible. This framework seeks not to contain such innovation and growth, but to bring support and rigour to it in the interests of learning better practice and, ultimately, more sustainable development outcomes.

1.1. Who will find the guide helpful?

This guide is written primarily for development professionals who already have some knowledge of the concepts of citizen engagement and evaluation and who are interested in understanding more about the contribution that a digital approach can bring to citizen engagement and how that contribution can best be evaluated.

More specifically, the guide is designed for World Bank Group (WBG) Task Team Leaders (TTLs) who may be evaluating programs or commissioning others to do so, and evaluators and consultants throughout the project cycle—from concept note to Proj-

ect Development Objectives (PDOs) and mid-term review to final evaluation. It is also intended to be a helpful resource for Civil Society Organizations (CSOs), researchers or anyone else seeking to understand the role of technology within citizen engagement.

Although the guide takes a step-by-step approach (see Section 1.2.), it allows readers to access those parts that are most relevant and useful to them, e.g.:

Practitioners and advisers/consultants working with technology in citizen engagement programs in the field and seeking to evaluate and improve their work, may find the practical guidance in **Section 4** to be the most useful when conducting or managing monitoring and evaluation activities.

- **Evaluators new to the fields of technology and/or digital citizen** engagement may find the overviews of DCE and its evaluation in **Sections 2 and 3** will be a useful learning resource, together with the DCE-specific suggestions for scoping and design in **Sections 4.1 and 4.2** (although more experienced evaluators may also find useful tips throughout Section 4).
- **Commissioners of evaluations** are likely to find the general introduction to DCE in **Sections 2 and 3** will be useful when framing the job requirements and the DCE-specific suggestions in **Sections 4.1 and 4.2** on scoping and design will be valuable when agreeing a statement of work with the evaluator, and ensuring the evaluation does not miss important aspects of DCE and that any Terms of Reference are based on a better understanding of what might be involved.

The challenge of evaluating Digital Citizen Engagement

A regular complaint among practitioners and academics alike is that we do not really know how effective technology interventions have actually been. All too often high quality monitoring and evaluation are ignored, underfunded, or left as an afterthought. Moreover, even when it takes place, the design of evaluation activities often means that they are more expressions of wishful thinking, rather than rigorous reviews of why different elements of a program might have been successful or not.

Three particular problems are pertinent with evaluating Digital Citizen Engagement: first, actually identifying the extent to which it is the technology, rather than anything else, that has had the impact; second, the use of generalised 'official' statistics, be they from governments or operators, which may not sufficiently differentiate between ownership of a device, and actual usage thereof; and third, getting the balance right between expected and unexpected outcomes. Digital engagement need not always be a positive outcome!

Professor Tim Unwin
Secretary General of the Commonwealth Telecommunications Organization, UNESCO Chair in ICT4D at Royal Holloway, University of London
www.unwin.wordpress.com

1.2. Structure of the guide

The guide offers:

Section 2 Digital Citizen Engagement: defines DCE and the benefits and challenges within it, particularly in areas with low technological access.

Section 3 Evaluating Digital Citizen Engagement: outlines existing evaluations of Digital Citizen Engagement and introduces the construct of five 'lenses' which can assist in focusing on the most important aspects of DCE when conducting an evaluation.

Section 4 A practical guide to evaluating DCE: how to conduct an evaluation of Digital Citizen Engagement, following standard stages of an evaluation lifecycle from scoping and high-level design, through detailed planning (and some tips on implementing the evaluation) and analyzing of the data, and advice on sharing, reflecting and learning from the evaluation findings.

Section 5 Evaluating the evaluation—reflecting on the process: provides some final thoughts and reflections on the evolution of Digital Citizen Engagement and what this means for a guide such as this along with a call for wider involvement in the continued evolution of this work.

Toolkit 1 DCE evaluation bank: examples of primary (assessment/analysis) and supplementary (information gathering) evaluation questions—grouped by lens—and some 'satisfaction' questions. Whilst these are not specifically evaluation questions, they may be useful in framing the evaluation questions for a particular context or as a starting point for conducting surveys or interviews with participants.

Toolkit 2 Using the lenses in scoping and design: a set of considerations and questions that an evaluator might ask during the scoping and design stages, again grouped by lens.

Appendices: including global examples of DCE interventions, results indicators for citizen engagement and an overview of the Brazil, Uganda, Cameroon and Kenya case study data collection methods and costs.

1.3. Navigating this guide

The guide uses a number of conventions to help the reader to navigate and to highlight the content for different purposes.

Detailed information collated into tables: to help the narrative flow, detailed explanations, examples and additional information have been collated into tables.

	Column	Column
Row	Information/Data	Information/Data
Row	Information/Data	Information/Data

Contributions from experts in the field: thoughts, advice or case-study reflections on topics related to the evaluation of the role of technology in citizen engagement. These are placed at the end of the section of the guide most relevant to their theme.

Introductions to evaluation lifecycle stages: each of the evaluation lifecycle stages in **Section 4** is introduced with a cover diagram showing the stage, followed by a box containing a brief explanation of the purpose of the section and what content it covers.

Reflections from the four related field evaluations: each stage in **Section 4** includes reflections of key lessons and learning from the four field evaluations that were undertaken as part of developing this guide.

Key lessons learned from the field evaluations

- Evaluation Lifecycle Stage -

Key lesson

Quotation from field evaluator

Further Reading: Curated lists of suggested online resources relevant to each section of the Guide

Section
Resource Name – *http://www.url.com*
Resource Name – *http://www.url.com*

Checklists for practical guide sections: each stage of the evaluation lifecycle in Section 4 ends with a brief checklist of the key points that an evaluator should have ad-

dressed (or at least considered) before moving on to the next stage of the evaluation.

Moving on from stage to stage?

- Summary of key point
- Summary of key point

1.4. Primary fieldwork – Brazil, Uganda, Kenya, Cameroon

The development of this guide included conducting four field evaluations of DCE projects in Brazil, Cameroon, Kenya and Uganda. These evaluations have informed the guide and lessons from each will be shared throughout the document. A brief summary of each is provided here.[1].

Brazil – This evaluation investigated the impact of online voting on the state-wide Participatory Budgeting process in Rio Grande do Sul, Brazil. It explored differences between online and offline voters: the extent to which online voting impacted on overall voter turnout; if online voting reached different sections of the population from those who traditionally engage in the face-to-face process; whether online voters interact with the participatory budgeting process in a different way from those engaged face-to-face.

Uganda – This evaluation looked at a long-standing and well-funded crowdsourcing platform *U-Report*. It considered the nature of the data collected through the crowdsourcing platform and highlighted the implications of this for contributors and policymakers. It had a particular focus on how representative U-Reporters are of the wider population of Ugandan youth and under what circumstances it is, or is not, appropriate to use U-Report as a tool for eliciting representative views.

Kenya – MajiVoice (Kenya) is a service that enables Kenyans to easily and conveniently give feedback to their water supply company through a mobile phone or the Internet and a back-end complaints handling system. This evaluation looked at the extent to which the digital engagement component of the feedback system was used and how such feedback contributed to people being able to influence the availability and quality of services provided to them. It also examined the impact such feedback mechanisms had within the service provider in terms of improved service delivery.

Cameroon – The Centre d'Appel Citoyen et TIC (TIC-GOUV) in Cameroon is a relatively new and local form of participatory budgeting that has been experimenting with the use of SMS technology to increase participation levels and, more recently, cast votes in budgeting decisions. This evaluation set out to investigate the extent to which short message service (SMS) managed to reach out to a diverse range of citizens and the impact of the use of SMS on the propensity of people to participate in the participatory budgeting process and the nature of their participation. Interestingly however, the program data did not make such an evaluation practical and instead the team proceeded to explore cost and data quality issues related to the program in order that it could improve its use of technology, its understanding of the role of data within its work and its evaluability in future years.

1–The full evaluation reports are being published separately.

TABLE 1. EVALUATION METHODS USED IN THE FIELD EVALUATIONS

Country	Evaluation methods
Brazil	Online web-form survey for online ballots **33,758 respondents** Face-to-face survey in physical polling stations **1,923 respondents** Interactive Voice Response randomized automated dialling telephone survey **2,173 respondents** Supplementary interviews with field staff, government officials and local academic experts
Cameroon	Systems data analysis, including SMS contacts database, meeting participants lists and call centre transaction logs A small selection of supplementary interviews with program staff, citizens and local officials
Kenya	Online surveys (conducted through intermediaries) with users (complainants) of MajiVoice **1,064 respondents** 106 paper based surveys (conducted through intermediaries) MajiVoice water company staff users MajiVoice system data (transaction logs) analysis
Uganda	SMS survey with U-Reporters **5,763 respondents** SMS U-Reporter poll **286,800 respondents** Household survey **1,188 respondents** RIWI (randomized internet survey software) **2,884 respondents** 20 interviews with U-Reporters and UNICEF staff

2 Digital Citizen Engagement

2. Digital Citizen Engagement

2.1. What is Digital Citizen Engagement?

WBG citizen engagement strategy defines citizens as the ultimate client of government and/or development institutions and private sector interventions in a country (World Bank, 2014a, p4) and citizen engagement as: the two-way interaction between citizens and governments or the private sector which give citizens a stake in decision-making with the objective to improve intermediate and final development outcomes (World Bank, 2014a, p5).

Building on this understanding of Citizen Engagement, this guide is built on a definition of DCE as **the use of new media/digital information and communication technologies to create or enhance the communication channels which facilitate the interaction between citizens and governments or the private sector**.

DCE is related to concepts such as 'civic technology' (Knight Foundation, 2014; Steinberg, 2014) and 'technologies for transparencies and accountability' (Kelbert et al., 2013; Leavy, 2014; McGee and Carlitz, 2013), but different in that it pays particular attention to the non-technical dimensions of participation and adopts a critical view of technology selection, implementation and delivery.

The empowering potential of Digital Citizen Engagement

While mobile and web-based technologies are not a panacea to all social problems, they enable information to be conveyed at a cost and scale never before possible. We believe that if people are able to access, share and create information, they are empowered to create positive change in their own lives and communities.

Citizens are now able to access information on their elected representatives, national and state budgets and the laws that govern them so that they can make more informed choices. They are also able to report challenges in service delivery using simple SMS-based solutions. Social media can also be used to amplify the voices of marginalised communities. When integrated into well devised programs, tailored to the local context, these initiatives can help enable citizens to hold government to account.

We are starting to see the real impact that technology can have in holding governments to account. As an example, when poor mining practices resulted in thousands of children dying due to lead poisoning in Bagega in Northern Nigeria, the Follow the Money team amplified their stories using social media and stimulated a campaign which resulted in the government releasing $5.3 million dollars for critical healthcare in the region.

This is just one of many examples of how citizens are utilising technology to amplify their voices and hold government to account.

Loren Treisman
Indigo Trust
www.indigotrust.org.uk

2.2. Typology for Digital Citizen Engagement

As a relatively new, expanding and increasingly diverse field, a sufficiently nuanced typology for DCE has yet to be fully developed, despite the existence of a multitude of different ways of categorising different dimensions of the DCE process and the agendas and priorities that these express.

Existing typologies can be helpful in providing insight into the main attributes of DCE. As illustrated in ***Table 2*** below, each typology emphasizes different aspects of DCE. Taken together these dimensions provide an overview of the wide spectrum of approaches and methodologies to DCE.

TABLE 2. TYPOLOGIES RELEVANT TO DIGITAL CITIZEN ENGAGEMENT.

Typology	Explanation	References
By IAP2* Spectrum of public participation	Categorizing DCE by the level of decision-making authority. For example, informing with static websites, webcasts etc.; consulting through interactive sites such as *www.fixmystreet.com*; involving with polling sites such as *www.wikiplanning.org*; collaborating with online 'town halls' and empowering through online participatory budgeting.	IAP2 (2007); Nabatchi (2012)
By stage in the policy cycle	The different uses/objectives of DCE at the policy cycle stages of political vision, policy formation, policy proposals, decision making and implementation.	Warburton et al. (2006)
By direction of engagement	This involves 'upwards' and 'downwards' accountability or flows of information. Upwards is where the state gains the ability to hear and observe its citizens (e.g., through participatory budgeting). Downward is when citizens can observe the conduct and policies of those in power (e.g, through freedom of information sites).	Davies and Fumega (2014)
By democratic model/function	Role of DCE in strengthening the public sphere to achieve collective purposes (e.g. in fixmystreet.com), for direct digital democracy (e.g., in mobile voting for participatory budgeting, online feedback systems), for social monitoring.	Fung et al. (2013)
By outcome	The extent to which there is more representative and powerful voice, more effective transparency and increasing accountability.	NDI (2014)

* International Association for Public Participation.

Typology	Explanation	References
By theme / purpose of engagement	Categorizing DCE according to whether it falls under 'Open Government' (advancing government transparency, accessibility of government data and services, and civic involvement in democratic processes) and 'Community Action' (projects catalysing peer-to-peer information sharing, civic crowdfunding and collaboration to address civic issues).	Knight Foundation (2014)
By initiator	Type A: Principal-initiated and managed feedback system, reporting directly, e.g., to a Presidential Delivery Unit, the Prime Minister's Office, a Governor or Mayor. Type A feedback systems enable Principals to get real-time feedback on problem hot-spots, the effectiveness of their departments and to initiate appropriate remedial actions. Type B: Manager-initiated and controlled feedback systems, e.g., for a specific health service, school district, or a major infrastructure development project. Type C: User-initiated and owned feedback systems actively engage citizens in sharing responsibility for resolving the problems that affect them.	Galtung (2014)

Table 3 below attempts to group real-world examples of DCE using a simple one-dimensional categorisation derived from some of the above. It helps to show a diversity of tools and technologies are being used in development projects all over the world, to achieve a variety of different citizen engagement goals. Some of these tools, such as citizen score cards, have a long history in development. However the central question for DCE is the extent to which the new digital engagement tools are more effective than established non-tech approaches in enhancing interactions between citizens and governments to influence policies and improve delivery of services.

TABLE 3. EXAMPLES OF CITIZEN ENGAGEMENT AND DCE GLOBALLY[1].

'High-tech' | e.g. crowdsourcing, interactive mapping, web interfaces		
Social Monitoring		
Public service delivery (incl. beneficiary feedback)	**Corruption reporting**	**Citizen-driven election monitoring**
FixmyStreet (UK); DevTRacker (UK); SeeClickFix (USA); Open Data Kenya; CheckmySchool (Philippines); Map Kibera; g0v (Taiwan); IChangeMyCity; Map Tandale (Tanzania); Alaveteli; WhatDTheyKnow.	Ipaidabribe (India and Pakistan); Magyar Leaks (Hungary)	Ushahidi (Kenya); Mumbai Votes (India)
Direct democracy		
Interaction with political representatives	**Participatory budgeting**	
Mzalendo (Kenya); WritetoThem (UK); Mi Medellin (Colombia) Governo Escuta/Responde (Brazil)	d-Brain (South Korea)	
A more robust public sphere		
Consultation, discussion, deliberation		
NotemyVote (UK); Avaaz; g0v (Taiwan)	OpenTownHall (USA) NoteMyVote (UK) CodeForAmerica (USA)	

1–See Appendix B for links to more information on these projects.

'Medium-tech' | e.g. SMS or call center		
Social Monitoring		
Public service delivery (incl. beneficiary feedback)	**Corruption reporting**	**Citizen-driven election monitoring**
Maji Voice (Kenya) Jaankari (India) Hello Sarkar (Nepal) M4W (Uganda)	Transparency International (Zimbabwe)	Mobile Monitors (Nigeria)
Direct democracy		
Interaction with political representatives	**Participatory budgeting**	
Online voting (Brazil); SMS voting (Cameroon, DRC);	U-Report (Uganda, Nigeria, Zambia)	

'Low-tech' | e.g. community radio
Social Monitoring
Public service delivery (incl. beneficiary feedback)
CGNetSwara (India); TracFM (Uganda) Namma Dhwani (India) RadioActive (India)

2.3. Benefits and challenges of Digital Citizen Engagement

Key factors in both Citizen and DCE have been well documented elsewhere (e.g. Nabatchi, 2012). They include: consideration of who to engage with and how (participants and recruitment); how many to engage with (scale); why the engagement is taking place and what it is intended to achieve (purpose and goals); the nature of communication between all the parties concerned (mode of communication); which tools/processes to use (participation channels); the extent of the link to decision-makers (connection to policy process).

It has been argued that, if implemented thoughtfully and contextually, citizen engagement, including beneficiary feedback, can result in transparency, accountability and more targeted outcomes (Bertot et al., 2010; Nabatchi, 2012; Warburton et al., 2001; World Bank, 2014a), although assumptions that it will necessarily lead to sustainable improvements in people's lives have also been critiqued (Davenport, 2013; Guijt, 2014).

The volume, variety and velocity of data that is a feature of modern digital technologies is something never before experienced (UN Global Pulse, 2012). While handling and analysing this data poses significant challenges, if done successfully it holds great potential benefit for DCE to make communication for all parties faster, easier, cheaper, more diverse and at scale, cutting across time, space and hierarchy. Table 4 sets out some of these benefits and challenges.

TABLE 4. BENEFITS AND CHALLENGES OF DCE.

Benefits
• Expands and diversifies opportunities for citizens to engage with the state and other service providers in ways that bypass traditional intermediaries. • Reduces the costs of participation by tapping into existing technology, reducing the need to, e.g., travel, be present at fixed times, incur venue costs both for the citizens and the intervention (Bertot et al., 2010; Grossman et al., 2013; Wittemyer et al., 2014). • Quickly produces and disseminates accessible data by taking advantage of cloud-based technologies where results can be collated in a central database near instantaneously, and that same central database can be queried by managers, stakeholders and even end-users in real-time. • A 'glare effect' of using digital tools to draw attention to causes.

Challenges

- DCE inherits 'traditional' CE challenges, such as a lack of agreed methods due to diversity of projects, questions of who is a 'citizen' and what is meant by 'engagement' and the long-term results of evaluation vs the short-term timelines usually available (Nabatchi, 2012).
- The 'digital' aspect of citizen engagement being seen as a panacea, particularly in the case of applying DCE to contexts where Information & Communication Technology (ICT) and ICT knowledge and skills are not prevalent. Despite the spread of mobile phones DCE may still pose significant barriers to participation and, therefore, inclusiveness.
- Enforcing norms in digital spaces.
- Increased data availability does not guarantee good data quality for meaningful analysis.
- High volume of data requires specialist expertise to collect, handle and analyse effectively.
- High-quality information and analysis is not enough on its own—to influence decisions/policies it needs the right delivery channel, in the right place at the right time.
- Vast collections of citizen data pose increased risks for anonymity and confidentiality (even more so for cloud-based applications).

Every approach and technology comes with its own opportunities and pitfalls that need to be taken into account if the appropriate questions are to be formulated and methods selected. In some cases it is not the technology that is the primary barrier to engagement, but wider issues with the enabling environment–for example, a lack of trust in officials or the regime (McGee and Carlitz, 2013).

In summary, it is particularly important to question whether a *digital* approach to citizen engagement is a sensible, practical and fair option in a particular environment. There may be opportunities to combine digital with non-digital methods, or in some cases it may be more appropriate to use entirely non-digital methods of interaction – particularly in environments where technological infrastructure and ICT literacy are weak and where a digital approach to citizen engagement may exclude exactly the people it is intended to help.

Benefits: a DCE success story, SMARTerWASH Ghana

Digital Citizen Engagement is happening in a new initiative in Ghana to make rapid improvements to water and sanitation coverage in rural and small town areas.

Ghana's Community and Water Sanitation Agency (CWSA) is engaged in a program called SMARTerWASH, to enable a shift from counting facilities to monitoring the services actually provided. Services are measured against indicators for functionality, service level, service provider performance and service authority.

Local citizens, repair and maintenance businesses and local and regional government staff are linked through two partners, Akvo and SkyFox. Between June and December 2014, CWSA collected more than 24,000 data points, to assess and publish the state of local water facilities. This rapid surveying was possible by replacing paper-surveys with Akvo FLOW, an Android smart phone and web-based survey and monitoring tool. Such intense point monitoring was combined with improvements and incentives to the local and regional repair networks, with the goal of building a sustainable network of repair businesses, closely integrated by commercial partner SkyFox, making smart use of communication technologies (SMS, USSD and a call centre) to organize mobile payments, spare parts ordering, book mechanics and manage deliveries.

The results have triggered repairs and other remedial actions that benefit over 11 million water users, leading to a scaling up supported with additional funding of around $3.9 million from the Government of Ghana, the Netherlands Government, World Bank, Unicef and Conrad N. Hilton Foundation.

Thomas Bjelkeman
Co-founder and director, Akvo
www.akvo.org

Selected readings and online resources on Digital Citizen Engagement
Strategic Framework for Mainstreaming Citizen Engagement in World Bank Group Operations – https://consultations.worldbank.org/consultation/engaging-citizens-improved-results
Bang the Table – "engaging communities worldwide" http://bangthetable.com/
Citizen Participation and Technology – https://www.ndi.org/node/21589
Engagement DB – http://engagementdb.org/
Participation Compass – http://participationcompass.org
Participatory Budgeting: Core Principles and Key Impacts – http://www.publicdeliberation.net/cgi/viewcontent.cgi?article=1236&context=jpd
Participedia – http://www.participedia.net
Tech Change – http://techchange.org/
What Does the Civic Tech Landscape Look Like? – http://www.knightfoundation.org/features/civictech/
World Bank Social Accountability Sourcebook – http://www.worldbank.org/socialaccountability_sourcebook

Challenges : social complexities of evaluating Digital Citizen Engagement in Nigeria

The Governor of Jigawa state offered citizens and residents of the state his personal mobile number to report dysfunctional public services and other public needs by sending a text message to him directly. This was meant to redress the perception that the state is unresponsiveness to community needs. It was also intended to provide a sense that the citizens were valued and visible. Individuals applauded this effort:

"For six months we had no power supply. Now we have power supply as a result of many texts to the governor: government is now responding. The governor also acknowledged that the people have been very patient."

It was a different story with state officials who found it to be impractical, reactive and not always respecting constitutionally defined responsibilities across the three tiers of government. A state health official reported:

"The governor has given everyone–the public–his mobile number so that people can send him messages directly about public services in their communities. This is not effective because we are spread so thin, as our health care delivery remit is now effectively taking over what is Local Government function."

In countries where there is a palpable disconnect between those who govern and the governed, engagement and interaction with the state is conceived as a worthwhile goal to evaluate in its own right. However, this betrays an assumption that these interactions by themselves will transmogrify into sustainable institutional arrangements that create accountability structures for the delivery of essential public services.

The gulf between citizen and state provides a compelling narrative upon which a case is built for bridging the governance divide. Responsive governance, albeit virtual, through digitalising citizen engagement provides a platform for politicians to be seen as actively engaged in issues relating to the poor.

If we evaluate these interactions solely as outputs, we risk the danger of providing simplistic analyses of the social and institutional complexities underpinning these interactions. Evaluation of digital citizens' engagement within political society should therefore include robust analysis of the mutually reinforcing relationship with institutional arrangements and structures that undergird the process through which political actors are held accountable for their citizens' developmental needs.

Adebusoye Anifalaje
Director, BAO Systems and Consultant, Health Information System Program (HISP)
www.baosystems.com / www.hisp.org

3

Evaluating Digital Citizen Engagement

3. Evaluating Digital Citizen Engagement

3.1. An emerging practice

Improving transparency, accountability and effectiveness is a fundamental but complex development challenge for donors. Technology is constantly evolving, as is our understanding of the role of government and the nature of the government–citizen relationship. This means the field of DCE—at the juncture of these evolving disciplines—is also changing rapidly, illustrating that DCE does not operate in isolation, but is one of many options (both digital and non-digital) within a wider attempt to engage citizens more effectively. It also calls for new skills and capacities to be developed (e.g. a higher level of data literacy).

As the field develops, consideration of how best to evaluate it is also developing. Evaluations of DCE need to look at the intervention both in its own terms and within the broader field of citizen engagement. This includes considering comparisons against non-engagement or a non-digital alternative, the over-arching developments in technology and the application of technology to democracy and to aid and development.

Evaluating Digital Citizen Engagement at mySociety

mySociety is an international UK-based NGO with the mission to "invent and popularise digital tools that enable citizens to exert power over institutions and decision makers". These digital tools are open source and include: FixmyStreet, where citizens can report local problems to the relevant public authority, adopted in 13 countries; Alaveteli, where freedom of information requests can be made in over 20 countries, and Pombola, a parliamentary monitoring software enabling citizens to monitor the work of their elected representatives. These platforms are designed to promote government transparency at low cost through maintaining easy-to-use portals for citizen-institution communication, and providing an archive of requests and responses for any individual to search.

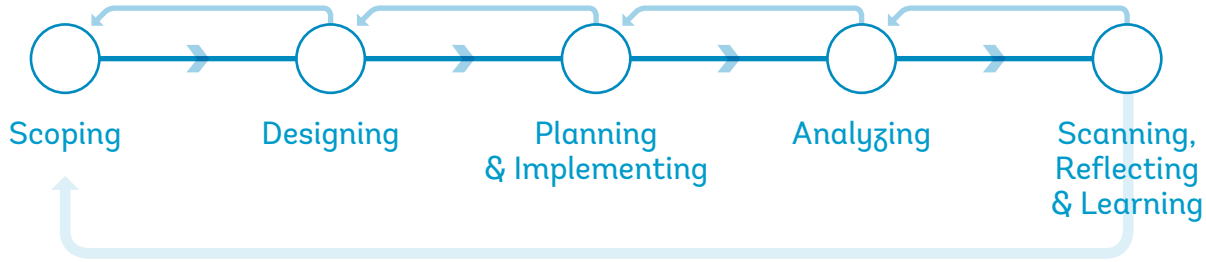
mySociety knows that its tools are being used by a large number of citizens within the UK and a growing number around the world. It is now turning its attention to how effective those tools are in having a meaningful impact upon citizens, decision-makers and institutions. Is requested information being provided in a timely and accessible format? Are local issue reports being fixed? Are institutions equally responsive to citizens? Are these tools genuinely making a difference? These are questions that mySociety is now asking in the hope of understanding where such tools can have meaningful impact upon citizens and institutions.

We are taking a mixed methods approach to analyzing its 'real world' impact. Several comparative quantitative activities are being conducted using surveys and online analytics tools to understand user demographics and attitudes. Randomized Control Trials are being run to understand how the appearance of the site and the information provided to users can influence their level of engagement. In-depth qualitative interviews are also being conducted with users, institutions and implementers to understand the motivations, frustrations and operations of those individuals and organizations using digital tools. These are some of the very first studies of this size in this area, and will provide a broad and rich understanding of the operation and impact of Digital Citizen Engagement.

Dr Rebecca Rumbul
Head of Research, mySociety
www.mysociety.org

Selected readings and online resources on Evaluating Digital Citizen Engagement
A Manager's Guide to Evaluating Citizen Participation – *http://www.businessofgovernment.org/report/manager%E2%80%99s-guide-evaluating-citizen-participation*
Designing Initiative Evaluation: a Systems-oriented Framework for Evaluating Social Change Efforts – *http://www.wkkf.org/resource-directory/resource/2008/04/designing-initiative-evaluation-a-systems-orientated-framework-for-evaluating-social-change-efforts*
Evaluating Citizen Engagement in Policy Making – *http://iog.ca/publications/evaluating-citizen-engagement-in-policy-making/*
Evaluating eParticipation Projects Practical Examples and Outline of an Evaluation Framework – *https://joinup.ec.europa.eu/sites/default/files/91/da/cc/ePractice%20Journal-Vol.7-March%202009.pdf*
Evaluating Participatory Projects – *http://www.participatorymuseum.org/chapter10/*
Evaluating Participatory, Deliberative and Co-Operative Ways of Working – *http://www.sharedpractice.org.uk/Downloads/Interact_Working_Paper.pdf*
Making A Difference: a Guide to Evaluating Public Participation in Central Government – *http://www.involve.org.uk/blog/2007/06/26/making-a-difference/*
Monitoring and Evaluating Advocacy (UNICEF) – *http://www.unicef.org/evaluation/files/Advocacy_Toolkit_Companion.pdf*
The True Costs of Public Participation—a Framework – *http://www.sharedpractice.org.uk/Downloads/TC_Framework.pdf*
Toward Metrics for Re(imagining) Governance: the Promise and Challenge of Evaluating Innovations in How We Govern – *http://thegovlab.org/wp-content/uploads/2013/04/GovLabMetrics.pdf*

3.2. The life cycle of a DCE evaluation



Throughout this guide the emphasis is on aspects of evaluation that are relevant to the *digital* aspect of citizen engagement, using the standard evaluation lifecycle of scoping, designing, planning and implementing, analyzing, and finally sharing, reflecting and learning on the evaluation process.[1]

This diagram, and indeed the guide, presents the stages of an evaluation in a linear fashion but, as the arrows show, the reality of designing and implementing an evaluation is less clearly delineated. Stages may be iterative and include any number of feedback loops (possibly involving small-scale pilots initially). The process often requires revisiting earlier stages as a development at a subsequent stage can have implications not only for the next stage, but also for previous ones. Different stages of evaluation can also happen in parallel and run concurrently, for example, design might still be going on while planning begins on some aspects. Some issues (for example involving beneficiaries throughout an evaluation, see box below) are cross-cutting and are relevant at every stage.

Another aspect to consider is the timing of the evaluation and the relationship between the evaluation process and the wider project. At one (preferred) end of the spectrum, the evaluation of the project is considered at the project Design stage and is integral to the implementation of the project (and is in the budget), at the other end of the spectrum evaluation is not thought about until after the project is over, with no resources set aside.

1–Of course not all evaluations or evaluation guides/frameworks use these exact names and stages, but the broad flow is consistent across a vast majority of evaluation work.

Feedback is about more than data collection: involving beneficiaries in design, validation and dissemination of learning and findings

Feedback for me is about a conversation. This is what distinguishes it from simple data collection. It is all too common for this conversation to only occur in the data collection phase. However, what about the other stages of the evaluation process? What about a conversation with beneficiaries during design or even before? What about a conversation to ensure that our provisional evaluation findings, including judgments, are on track? That 'they' haven't been subject to our world view to the point that we may have missed crucial cues or been unable to break through entrenched power relations? What about ensuring that findings, including lessons learned, are shared to ensure that beneficiary groups involved in a global program can learn from and adapt successful practice in other parts of the world? How about feedback at all these stages?

- *Feedback as part of evaluation design: e.g. sharing of/ consultation on/ participatory design of evaluation*
- *Feedback as part of data collection: could be extractive/ interactive/ participatory collection of information*
- *Feedback as part of joint validation and or/ analysis of information: could be extractive or participatory*
- *Feedback on end product/ response and/or follow up: could be simple dissemination or participatory engagement for future actions.*
- *This isn't just about manners, ethics, respect. It is also about ensuring we have robust evaluation findings.*

Dr Leslie Groves Williams
Expert, Participatory and Inclusive Approaches to Evaluation
www.beneficiaryfeedbackinevaluationandresearch.wordpress.com

Selected readings and online resources on effective evaluations

Better Evaluation – *http://betterevaluation.org*

Better Evaluation: Rainbow Framework – *http://betterevaluation.org/resources/download_the_Rainbow_Framework*

CIVICUS Toolkit for Monitoring and Evaluation – *http://civicus.org/index.php/en/resources/toolkits/228-monitoring-and-evaluation*

Conducting Quality Impact Evaluations Under Budget, Time and Data Constraints – *http://lnweb90.worldbank.org/oed/oeddoclib.nsf/DocUNIDViewForJavaSearch/757A5CC0BAE22558852571770059D89C/$file/conduct_qual_impact.pdf*

Framework for Assessment of ICT Pilot Projects – *http://www.infodev.org/articles/framework-assessment-ict-pilot-projects*

Impact Evaluation in Practice – *http://siteresources.worldbank.org/EXTHDOFFICE/Resources/5485726-1295455628620/Impact_Evaluation_in_Practice.pdf*

Monitoring and Evaluation: Some Tools, Methods and Approaches – *http://siteresources.worldbank.org/EXTEVACAPDEV/Resources/4585672-1251481378590/MandE_tools_methods_approaches.pdf*

UKES Guidelines for Good Practice in Evaluation – *http://www.evaluation.org.uk/assets/UKES%20Guidelines%20for%20Good%20Practice%20January%202013.pdf*

UNDP Handbook on Planning, Monitoring and Evaluation for Development Results – *http://web.undp.org/evaluation/handbook/documents/english/pme-handbook.pdf*

Why Evaluations Fail: the Importance of Good Monitoring – *http://enterprise-development.org/page/download?id=2484*

3.3. Five lenses for evaluating Digital Citizen Engagement

There is a range of interconnected issues and factors that need to be considered when evaluating DCE, such as the goals of the program, power dynamics and control, who is included or excluded, the planned and achieved impact and, of course, the choices and use of technology.

Learning from the field studies carried out in parallel with the development of this guide and the literature on CE, DCE and ICT for Development (ICT4D) has been used to develop five lenses. These lenses capture the different concepts and ideas involved in DCE (Section 3.1) and help in recognizing, organizing and evaluating them. Applying the term literally, the lenses provide five distinct yet overlapping ways of looking at the DCE, and we recommend that every evaluation considers every lens to begin with. This multifaceted view will help to ensure that important and nuanced issues relating to both technology and participation are not overlooked.

The lenses can help to focus on important themes for consideration in the early stages of the evaluation and directly inform the formulation of the evaluation questions. However, the lenses may well have differing levels of relevance and importance depending on the specific project, context and evaluation, and not every perspective, for example, may be fully represented in the evaluation questions.

Table 5 below sets out the five lenses, their perspectives and some key questions. They are described in more detail in **Section 3.4** and their application to evaluation is discussed in **Section 3.5**.

TABLE 5. FIVE LENSES FOR USE WHEN EVALUATING DCE

Question	Evaluation
OBJECTIVE	
What are the goals of the initiative, and how well is the project designed to achieve those goals?	Clarify the goals and planned changes, assessing the existence and appropriateness of those stated goals.
Question	**Evaluation**
CONTROL	
Which actors exert the most influence over the initiative's design and implementation, and what are the implications of this?	Explore the levels of influence on the engagement process, the dynamics of decision making, and levels of fairness and equitability among citizens.

Question	Evaluation
PARTICIPATION	
Which individuals participate in the initiative, and to what extent is their participation in line with their needs and expectations?	Examine who is included/excluded in the process, and how the program enables or discourages different opportunities for participation.

Question	Evaluation
TECHNOLOGY	
How appropriate was the choice of the technology, and how well was the technology implemented?	Take a practical look at the technology choices, technical delivery and management of the engagement process itself.

Question	Evaluation
EFFECTS	
What effects did the project have, and to what extent can this impact be attributed to technology?	Seek to understand the ultimate impact on citizens, government, collectives and service delivery/development outcomes.

3.4. The five lenses in more depth

This section provides an overview of the five lenses, defining them and outlining their importance in the evaluation of DCE. Although a thorough discussion of all the issues contained within each lens is outside the scope of this document, further reading on key topics relevant to each lens can be found at the end of each section. The application of the lenses is explored in more depth in **Section 3.5**.

3.4.1. Lens 1: Objective

Is the program objective reasonable and appropriate, and to what extent does the logic of the program lead to it?

This lens calls for the examination of both the objective and the design of the program. It is not sufficient to assess the objective in isolation– it is also important to consider whether the objective is sensible, reasonable and practical in the particular circumstances, and whether the theory or logic underpinning the program's design could reasonably be expected to meet this objective, assuming everything was delivered well. A program with a relevant objective built upon well thought-out logic creates the potential for success. The remaining lenses then consider whether this success was actually realized in practice.

Making this assessment requires an understanding of how the program was designed, who it seeks to benefit, and the thinking behind how its inputs and activities can be reasonably expected to deliver the intended results.

The objective is defined differently by different funders and organizations, but typically includes approaches such as a Logical Framework Approach (LFA or Logframe), Theory of Change or Theory of Action.

DCE goals can be multiple and impacts might be expected on varied dimensions/grounds, so an understanding of both the objective and the logic of the program is important in providing a reference point for the evaluation. Exploring a DCE program through this lens gives two crucial perspectives.

First, it provides an opportunity to explore and understand the project's goals (stated or otherwise) and the means and steps by which it hopes to meet those goals. Without this understanding, it is difficult to evaluate the activities themselves. In this way the lens can support the identification of researchable outcomes using the program's own terms.

Second, it offers an opportunity to reflect on these goals from an external perspective, to consider the program's terms—did the DCE program set out to do something realistic or was it unrealistic, overly ambitious or did it miss opportunities to do much more?

With both the depth and reach of technology developing rapidly, project managers are becoming increasingly literate in its uses and, in a prevailing culture of experimentation and lesson learning, DCE projects often develop and change both their approach and their goals. In other cases, citizens' needs only become fully apparent through the process of engagement, and so a level of responsiveness is needed.

From an evaluation perspective, being able to track these changes and understand the objective that is being sought and the logic that underpins them is important. In some cases, this logic is inherently complex and may benefit from reference to some theories/tools to help understand it[1]. From the project manager perspective, the refinement of the program's objective and logic based on insights from the evaluation, is a valuable gain.

Selected readings and online resources on Lens 1—Objective
Global Partnership Social Accountability—Results Framework – *https://www.thegpsa.org/sa/about/results*
How-to-guide: The logical framework approach – *http://www.bond.org.uk/resources.php/49/how-to-guide-the-logical-framework-approach*
Theory of change: The essentials – *http://www.bond.org.uk/theory-of-change-the-essentials*
W.K. Kellogg Foundation Logic Model Development Guide – *http://www.wkkf.org/resource-directory/resource/2006/02/wk-kellogg-foundation-logic-model-development-guide*

1–A simple example is Anthea Watson Strong's modification of Riker and Ordeshook's work on voting: *http://www.antheawatson-strong.com/writing/2014/6/8/the-three-levers-of-civic-engagement*

3.4.2. Lens 2: Control

Who controls and influences the digital engagement process?

This lens provides an opportunity to explore and understand who drives and controls the design and delivery of the DCE program, the choices of technologies used and the use of its outcomes. It is a chance to reflect on the degree to which citizens/beneficiaries are involved throughout the process, and whether this degree of participation seems appropriate in the context. It provides a reminder to consider and to examine the individual, group and state relationships within the project and between the project and its wider institutional and development context. It is also an opportunity to look at issues of transparency and accountability, and the extent to which decision-making is open and under public scrutiny and control.

This is an important perspective as these subtleties of citizen participation, citizen control and public scrutiny are often not core goals of a program, but in many cases are vital components to its success or failure, its actual impact and its sustainability.

Selected readings and online resources on Lens 2—Control
Mixed incentives: Adopting ICT innovations for transparency, accountability and anti-corruption – *http://www.u4.no/publications/mixed-incentives-adopting-ict-innovations-for-transparency-accountability-and-anti-corruption/*
Technology for Transparency – *http://globalvoicesonline.org/wp-content/uploads/2010/05/Technology_for_Transparency.pdf*

3.4.3. Lens 3: Participation

Who participates and how?

This lens brings into focus who participates, and how, when and why they participate. It involves understanding: who is included and excluded; people's desires and their ability to engage; incentives, expectations and barriers that may exist for different groups; the nature of their engagement. It is an opportunity to consider the wider dynamics of how representative those who engage are and whether this representation is suitable or problematic for the program goals.

This perspective is critical when evaluating a DCE program as claims are often made that a program is enhancing democracy, or representing one or more groups of citizens, or using the views of those engaged with to change policy or improve delivery, without those claims being substantiated. Knowing who is or is not involved and the dynamics of their engagement allows the validity of these claims to be discussed.

Selected readings and online resources on Lens 3—Participation
A Ladder of Citizen Participation (Arnstein, 1969) – *http://lithgow-schmidt.dk/sherry-arnstein/ladder-of-citizen-participation_en.pdf*
IAP2 Public Participation Spectrum – *http://www.iap2.org.au/resources/iap2s-public-participation-spectrum*
Integrating Information and Communication Technologies into Communication for Development Strategies to Support and Empower Marginalized Adolescent Girls – *http://www.unicef.org/cbsc/files/ICTPaper_Web.pdf*
Localizing Development: Does Participation Work? – *https://openknowledge.worldbank.org/handle/10986/11859*
Participation: The new tyranny and Participation: From tyranny to transformation – *http://www.zedbooks.co.uk/node/21816 and http://zedbooks.co.uk/node/21248*
The Participatory Museum – *http://www.participatorymuseum.org/read/*
Using Online Tools to Engage - and be Engaged by - the Public – *http://www.businessofgovernment.org/report/using-online-tools-engage-public*
Whose Reality Counts? Putting the First Last – *http://www.ntd.co.uk/idsbookshop/details.asp?id=355*
World Bank Participation Sourcebook (1996) – *http://documents.worldbank.org/curated/en/1996/02/696745/world-bank-participation-sourcebook*

3.4.4. Lens 4: Technology

How effective and appropriate is the choice and delivery of the technology?

This lens brings attention to the digital / technological aspects of a DCE program. Technology features in all the lenses as it is a core aspect of DCE, but this lens focuses specifically on the choice, effectiveness and implementation of the technology. Was it delivered well and appropriately? Was it a good choice? Was the choice made appropriate for the level of capacity amongst the participants? What was the impact of this choice? Could the same or better have been achieved using a different technology or non-digital approaches?

This perspective looks in more depth at technical and at wider delivery and management issues and considers whether the delivery process itself has impacted on the activities or results. Such a focus allows failures to be explored and learning and improvement to be identified. It also looks at the dimensions of technology, service development and delivery (such as data monitoring), that are essential whatever the goals of the program.

Selected readings and online resources on Lens 4—Technology
Barriers and Solutions in using M4D: Connecting Directly to Citizens for Scalable Impact – *http://www.votomobile.org/files/M4D_Barriers_Whitepaper.pdf*
Big data for Development - Challenges and Opportunities – *http://www.unglobalpulse.org/BigDataforDevWhitePaper*
Connect! - A practical guide to using ICTs in PLAN projects – *http://www.plan-academy.org/connect-a-practical-guide-to-using-icts-in-plan-projects/*
Global Mapping of Technology for Transparency and Accountability: New Technologies – *http://www.transparency-initiative.org/reports/global-mapping-of-technology-for-transparency-and-accountability*
Information Lives of the Poor - Fighting poverty with technology – *http://www.idrc.ca/EN/Resources/Publications/Pages/IDRCBookDetails.aspx?PublicationID=1275*
Insights into Participatory Video: a Handbook for the Field – *http://insightshare.org/resources/pv-handbook*
Insights into the Role of Technology in Addressing Development Challenges – *http://www.accenture.com/us-en/Pages/insight-role-technology-addressing-development-challenges.aspx*
Integrating Mobiles into Development Projects – *http://www.usaid.gov/documents/1861/integrating-mobiles-development-projects-handbook*
Making Mobile Feedback Programs Work – *https://openknowledge.worldbank.org/handle/10986/18712*
Mobile Data Solutions (online course) – *http://techchange.org/media/mobile-data-solutions/*
Mobile Technology Handbook – *http://pactworld.org/sites/default/files/Mobile%20Technology%20Handbook%202014.pdf*

3.4.5. Lens 5: Effects

What effects do citizens have on processes and outcomes?

The final lens is usually the most important as retaining an explicit focus on the actual impact of the engagement, and of the technology within it, is vital. Taking a step back from the details and the delivery, the DCE program can be looked at through the perspective of what it has actually achieved— what difference has been made in the lives of the citizens with whom it engaged, to the civil society groups or other collective groups involved, to government, those delivering public services and other decision makers. Of course, not all DCE programs set out to impact all these groups, but many will have an impact on every group anyway, so this lens encourages consideration of the program's ultimate impact in the wider world, above and beyond that of the initial program objective, theory and logic.

The lens also offers an opportunity to consider the impact of the Digital Citizen Engagement as a whole, and the impact the technological component specifically has had on citizen engagement.

The importance of this perspective is that, while it is not always possible reliably to evaluate impact, it is vital to consider both positive and negative impacts, intended or unintended, final or intermediate impacts a program may have had—and may continue to have. While it may be difficult to trace cause and effect directly from DCE activity to specific developmental outcomes, it is certainly possible to explore specific impacts (e.g. the extent to which the process changed participants' expectations of, and willingness to participate in, future DCE projects and other 'democratic outcomes'), intermediate results and indicators, and to speculate in an informed manner on the wider developmental impact of programs.

Selected readings and online resources on Lens 5 – Effects
Impact Case Studies from Middle Income and Developing Countries: New Technologies – *http://www.transparency-initiative.org/wp-content/uploads/2011/05/impact_case_studies_final1.pdf*
Measuring Impact On-The-Go – *www.theengineroom.org/wp-content/uploads/engnroom_monitoringguide_finalmay14.pdf*
So What Difference Does It Make? Mapping the Outcomes of Citizen Engagement – *http://www.ids.ac.uk/files/dmfile/Wp347.pdf*
World Bank Impact Evaluation Toolkit – *http://go.worldbank.org/IT69C50GLo*

3.5. The five lenses in practice

The lenses are a particularly useful device during the Scoping stage (**Section 4.1**) where the different perspectives may help guide who to talk to and what information to gather about the project itself and the wider environment in which it is being implemented. The data, knowledge and understanding developed through the Scoping can then inform the Design of the evaluation (**Section 4.2**) where the lenses are particularly valuable in helping decide on the evaluation questions, enabling the right balance to be found between a narrow scope and a broader evaluation, touching on wider-ranging issues. During the Planning, Implementation and Analysis stages of the evaluation (**Sections 4.3–4.4**), the lenses remain in the background, with the focus and guidance being provided by the chosen evaluation questions. The lenses come to the fore again in Sharing, Reflecting and Learning (**Section 4.5**) as a helpful way to summarize an evaluation and enable easier comparisons across different programs or across time.

Table 6 below sets out examples of the key areas that are covered by each lens, and how they can be applied at the Scoping and Design stages. Clearly there are areas of overlap between the lenses, reflecting the overlap in some of the issues within DCE. This overlap in the lenses acts a reminder of the key topics which run across the lenses.

The lenses can also be useful in exploring aspects of DCE that emerge from two lenses being looked at together. For example, by Control and Participation lenses together, it is possible to explore the two-way relationship between citizen and government and how people's decision on their level of engagement is influenced by their trust in the government and the process.

A more thorough set of considerations and questions that might be asked during the Scoping and Design stages can be found at the end of this document in **Toolkit 2**, grouped under each of the five lenses.

TABLE 6. EXAMPLES OF HOW TO USE THE LENSES DURING SCOPING AND DESIGN

Lens | Objective		
AREAS OF INTEREST	**TO EXPLORE AT THE SCOPING STAGE**	**TO CONSIDER DURING THE DESIGN STAGE**
The objective and goals of the engagement Specific goals related to the technology aspects of the program Sound objective and theory grounded in reality Different views of the program goals The counterfactual	Seeking to understand the explicit objective and underpinning assumptions of the program and the wider environment, including the planned impact	The gaps in our knowledge of the objective (is the objective explicit, or do we need to recreate one?), the arenas we want to explore further, and what counterfactual to use (if any).
Lens | Control		
AREAS OF INTEREST	**TO EXPLORE AT THE SCOPING STAGE**	**TO CONSIDER DURING THE DESIGN STAGE**
Mechanisms for citizen engagement are different depending on target group and at different stages of the program (or operation/intervention?) The extent to which the DCE program's processes are fair and equitable Factors that influence the ability of different stakeholders to influence the process How the sphere of influence of the program is being decided	Who is involved in decision-making at what stages; the mechanisms that exist to ensure fairness and equitability; the attention paid to historically marginalized groups; what evidence of stakeholders' influence already exists	How to evaluate the extent of different stakeholders' influence on the program and the implications of that
Lens | Participation		
AREAS OF INTEREST	**TO EXPLORE AT THE SCOPING STAGE**	**TO CONSIDER DURING THE DESIGN STAGE**
Recruitment and targeting Reasons for groups engaging or not How the program promotes (or prevents) inclusivity How the technology has changed the engagement dynamics The existence (or not) of times/spaces for meaningful engagement	The target audience(s), their characteristics, how they are reached; the opportunities provided by the program for them to participate and at what level	How to assess the effectiveness of the program in enabling participation, and whether the level of participation achieved met their objectives

Lens | Technology		
AREAS OF INTEREST	**TO EXPLORE AT THE SCOPING STAGE**	**TO CONSIDER DURING THE DESIGN STAGE**
Choice of technologies Data management and privacy Use of time and resources Overall program management and learning	The technology used and the reasons for its selection, the cost; How privacy issues are managed; How the overall program was managed	How effective the technology is and the quality of how it (and the overall project) is managed, cost-effectiveness in comparison to alternative approaches, quality of data safeguarding
Lens | Effects		
AREAS OF INTEREST	**TO EXPLORE AT THE SCOPING STAGE**	**TO CONSIDER DURING THE DESIGN STAGE**
Changes already noted in citizens / participants as a result of the program Changes brought about by the program in the way citizens organize collectively The impact on decision-makers Any tangible effects the program has made to the nature of life in its area To what extent any effects can be attributed specifically to the choice of technology used	The evidence (even anecdotal) that already exists of intended or unintended impacts; whether a 'control group' was identified or not; availability of baseline data; nature of the DCE project (e.g. designed as an RCT?)	How to establish whether the intended impact materialized, how to notice and assess unintended consequences, the cost of collecting data on the outcome of interest, the contribution of technology to the identified changes

Gender: A critical element of evaluating Digital Citizen Engagement

There is no doubt that it is critical to analyze gender participation in digital citizenship engagement projects with ICT tools. However, it is difficult to do that well. Evaluating and analyzing gender participation holds inherent challenges such as identifying issues of causality and isolating gender specific variables. Longitudinal quantitative research coupled with in depth qualitative methods can help to reveal some of the gender issues embedded in these projects in a meaningful way.

Renee Wittemyer
Director of Social Innovation, Intel Corporation
www.blogs.intel.com/csr/author/rkuriyan

4

A practical guide to evaluating DCE

Scoping 4.1

- Using the five lenses
- Finding scoping information
- Focus & Parameters

Designing

Planning & Implementing

Analyzing

Scanning, Reflecting & Learning

This stage lays out the groundwork for the design and implementation of the evaluation by investing time and resources into understanding the project and its context, the operating environment and the recent developments and insights from the DCE evaluation field. This section is important for both commissioners and evaluators as it sets the parameters, develops focus and highlights opportunities for the evaluation itself and ensures that the evaluation process is suitably informed by and grounded in reality.

4.1.1. Identifying scoping activity using the five lenses

As outlined in ***Section 3.5*** above, the five lenses are particularly useful at the early stages of an evaluation, to ensure important aspects of DCE are not being unintentionally ignored. The key issues to explore within each lens are repeated below and may be a helpful framing device when planning what information to explore during scoping:

TABLE 7. SCOPING USING THE FIVE LENSES

Lens	To explore at the Scoping Stage
Objective	Seeking to understand the explicit objective and underpinning assumptions of the program and the wider environment, including the planned impact
Control	Which actors are involved in decision-making at what stages; the mechanisms that exist to ensure fairness and equitability; the attention paid to historically marginalized groups; what evidence of stakeholders' influence already exists
Participation	The target audience(s), their characteristics, how they are reached; the opportunities provided by the program for them to participate and at what level
Technology	The technology used and the reasons for its selection, the cost; How privacy issues are managed; How the overall program was managed
Effects	The evidence (even anecdotal) that already exists of intended or unintended impacts; whether a 'control group' was identified or not; availability of baseline data; nature of the DCE project (e.g. designed as an RCT?)

4.1.2. Useful sources of scoping information

The suggestions above using the five lenses will guide the gathering of information about project goals, the wider environment, data that already exists or that may be needed, and the target group (see **Section 3.5** for an overview of using the lenses, and **Toolkit 2** for more detail). Potential sources of this information include:

Existing internal documents that define the project such as a Logframe or Theory of Change.

- Larger-scale **plans and statistics** in the public and private sectors.
- **Pilot field visits and/or stakeholder** interviews with, e.g., project commissioners, designers, managers, or participants.
- **Initial assessments** undertaken when framing the evaluation work (but attention should be paid to the potential drivers and influences of those who commissioned and conducted these)
- **System data** from DCE software tools and platforms.

- The **counterfactual**, such as a non-digital approach to engagement or no citizen engagement at all. The Scoping stage should reveal whether there is an obvious counterfactual to work with, whilst the Design stage can determine the extent of its use.
- Desk research should help to develop an understanding of the extent to which the goals and objectives are clear enough for an evaluation to be devised. In DCE programs, the goals are often not clearly articulated and sometimes they may not even be clearly understood by those involved, e.g., if the programs have evolved in response to wider policy, technological developments, or political or public demands. However, is important to have goals against which to evaluate the program so where there are no articulated goals or where goals have changed over time, some 'reverse engineering' may be required. In many cases, initial interviews with key stakeholders will help uncover the goals, even if they are not stated publicly, and in other cases goals may be inferred from historic information, wider political motivation, or broader activities into which the DCE work under evaluation falls.

In terms of **system data** one of the key, and potentially most valuable, assets of DCE is the data generated through the engagement process itself. The Scoping stage should explore what data already exists, its quality, and how technically and procedurally accessible it is. For example, privacy and security of data may need to be considered, including assessing the data against emerging best practices on responsible data (e.g. *responsibledata.io*) and identifying potential barriers to analyzing existing and collecting new data.

Macro- and micro-data about countries and populations, openly available on websites such as Gapminder, data banks or national bureaus of statistics, are also useful for understanding the context, meaning and limits of engagement. Some specific examples of data sources of particular relevance to DCE evaluations are provided in **Table 8**.

TABLE 8. POTENTIALLY AVAILABLE DATA SOURCES FOR DCE EVALUATIONS.

Information sought	Data sources	Examples
Citizen engagement metrics and content	System data generated by the DCE platform	SMS or email received/sent, complaints received, freedom of information requests made, categories of complains etc
Communications/ transactional data	Technical platform analytics	Mobile operator Call Detail Records, website use analytics, Google trends/ analytics
Population socio-demographics	Census Data	Population and housing census
Standardized surveys on socio-political attitudes and ICT use	Micro-data from national and cross-national surveys based on representative samples	AfroBarometer, Latinobarómetro, European Social Survey, World Values Survey, Global Corruption Barometer
Country data on economy, ICT penetration, governance and transparency	Macro aggregated data on social and economic statistics, perception indexes, ICT statistics	World Bank indicators, Human Development Index, Transparency International Corruption Index, Mo-Ibrahim Governance Index, Open Data Barometer, ITU statistics, Web Index
Online information on sentiments and opinions	News media and social media interactions	Twitter and Facebook trends, groups and reports, and other analysis of social media data using, e.g., sentiment analysis (Engine Room 2014, UN Global Pulse 2012)
Broader citizen reporting or crowd-sourced data	Information actively produced or submitted by citizens through mobile phone-based surveys, hotlines, user- generated maps, Web sites, etc.	Customer (beneficiary) complaint data submitted to water companies through the MajiVoice system in Kenya supplemented by Ushahidi data or Uganda U-Report SMS surveys supplemented by TracFM survey
Other project information	Previous evaluation surveys	Publicly available evaluation reports and supporting data, such as ***Participatory Budgeting in Cameroon: Booklovers, Mayors and Citizens*** (Paice 2014)

4.1.3. The focus and parameters of the evaluation

The information gathered and/or generated about the project and the external environment, deliberately kept broad at the Scoping stage by the use of the five lenses, now has to be set against the specific purpose, focus and goals of the evaluation.

Funders or commissioners of an evaluation may want to limit or specify areas of interest, and putting in place parameters is considered best practice for effective design or commission of evaluations (World Bank 2011, Better Evaluations 2014b).

These boundaries are often driven by a combination of clearly focused purpose and goals weighed against the related constraints of the time, budget/resources and data available:

Budget: financial and other resource constraints affect the number of interviews that can be conducted, the ability to combine quantitative and qualitative data collection and analysis, the size and professional experience of the research team, and the analysis that can be conducted.

Time: the start and duration of an evaluation will be affected by external time constraints—such as funder requirements, stakeholder availability and the length of time evaluators can spend in the field—as well as the nature of the project itself and the stage of the project lifecycle at which the evaluation needs to take place.

Data: when new surveys or data collection are conducted, data constraints can affect the ability to collect information from a suitable comparison group and obtain baseline information on the project population, or to collect sensitive information and interview difficult-to-reach groups (Bamberger 2006). In terms of human resources, DCE evaluation involves specific data analysis and technical skills that need to be addressed if the evaluator is not experienced in data or technical analysis.

However, the availability of system and transactional data ready for analysis and the ability to design and deliver cost-effective and rapid experimental models of engagement (and accompanying data ready for analysis) may militate against some of these constraints.

Given the range of issues that are relevant to DCE and its evaluation (see Sections 3.1 and 3.3), there is a danger that tight boundaries may mean important aspects or perspectives of the program being evaluated are missed. It is suggested that commissioners of evaluations consider each of the five lenses when framing requests for evaluation proposals, and that evaluators discuss the implications of the focus, limitations and lenses with the evaluation funder, who may agree that a broader focus is warranted.

4.1.4. Who evaluates?

In addition to the questions raised by Robert Chambers (see box below) the design of the evaluation needs to include consideration of what skills are needed, who is best placed to evaluate the program and the implications of such a choice. Whether the primary evaluator is internal or external, formally trained or not, when handling the large quantities of data that are typical in DCE programs, additional specialized skills are likely to be required – such as a data analysis, political science or statistical modelling for example.[1]

Evaluators and/or evaluation teams require objectivity, experience, digital and technical expertise, and an understanding of the organization or context in which the program operates. Local knowledge and context is also central to the ability to evaluate participatory value.

Depending on the evaluation approach the team might be comprised of local citizens and/or program, external specialist evaluators, or a combination of the two. It is important to consider the knowledge and experience (or lack of) each of these groups may bring, as well as the independence and objectivity or prior assumptions – this is as true for external evaluators as for internal.

Selected readings and online resources: scoping an evaluation
Beneficiary feedback in evaluation and research – *https://beneficiaryfeedbackinevaluationandresearch.wordpress.com/*
Decide who will conduct the evaluation – *http://betterevaluation.org/plan/manage_evaluation/who_conducts*
Frame the boundaries of an evaluation – *http://betterevaluation.org/plan/engage_frame*
Plan and Manage an Evaluation – *http://betterevaluation.org/start_here/plan_manage_evaluation*
Planning Evaluability Assessments – *https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/248656/wp40-planning-eval-assessments.pdf*
Writing terms of reference for an evaluation: a how-to guide – *http://siteresources.worldbank.org/EXTEVACAPDEV/Resources/ecd_writing_TORs.pdf*

1–For example Arthur Lupia's work with Climate Central and the Brennan Center *http://www.arthurlupia.com/*

Biases and distortions: critical who? and whose? questions when designing a DCE evaluation

As with all research, findings when evaluating Digital Citizen Engagement are affected by who exercises power and choice, who does not, who takes part and who does not, and potential biases and distortions at each stage of the process. Critical reflection requires asking who? and whose? questions such as:

- *Who decided what to try to find out about?*
- *Who determined the issues to be investigated and the questions to be asked, and why?*
- *Whose questions and issues were included, and whose left out?*
- *Who decided the medium or mechanism?*
- *Who had access to the medium or mechanism, and who did not?*
- *Who took part and who did not?*
- *(Gender, age, poverty, political, ethnic or religious group, class, caste, technical competences, other?)*
- *Why did non-participants not take part?*
- *(Lack of access, unaware, systematically excluded, unwilling, other reasons?)*
- *How were findings affected by who took part and who did not?*
- *What were the likely views of citizens who did not take part?*
- *What influenced or distorted the responses of those who did participate?*

(the interview or response situation—including who might be present when responding, distrust of how responses would be shared and used, fear of bad consequences from negative or critical responses, wanting to appear good, the presentation of the self, knowing the expected response, shortage of time and impatience, difficulty of using, cost of using, or misuse of the mechanism or medium, political, ethnic, religious or other group loyalty or other factors?)

- *Who had access to the data?*
- *Who owned the data?*
- *Who analyzed the data?*
- *Who shared the data and the analysis with whom?*
- *Who gained and who lost from the process?*
- *What steps were taken to reflect on and correct for possible or likely biases and distortions?*

Professor Robert Chambers
Institute of Development Studies
www.ids.ac.uk/person/robert-chambers

Key lessons learned from the field evaluations

- Scoping -

Focused scoping will result in a more targeted and easier to execute evaluation but may risk not looking at broader, equally important issues such as the impact of technology on participation and social inclusion

From Kenya: "The scope was nicely limited, looking at feedback from those complaining about the water service. That was an advantage because it helped keep things focused and well defined. It also helped that we were talking to someone who, whilst interested in wider issues, also had an initially limited scope for their internal evaluation requirements. They could also make decisions about actioning access, data sharing and data collection around the evaluation–working with someone like that good advice for anyone thinking about undertaking a DCE evaluation project. Basically, the focus was on "who is complaining and were the complaints being resolved?" It was very clear that it was not going to be about the wider social demographics and inclusion issues." – *Martin Belcher, Aptivate*

The importance of an iterative process – there may be valuable opportunities to reconsider the evaluation questions

From Brazil: "The original goals were very clear (asking whether people interact online in participatory budgeting and whether different people interact in different ways online than those who interact off-line). However, when we started the fieldwork and later the analysis, new, equally interesting questions arose–questions around opportunities for abuse and corruption, and data relating to the wider participatory budgeting not just the technological aspects. This gave us an opportunity to explore more of the five lenses and in more depth than we had originally hoped." – *Matt Haikin, Aptivate*

The importance of identifying well in advance potential barriers in data collection and analysis

From Uganda: "The study involved interviewing male and female U-Reporters across urban and rural environments. The lead researcher assumed that with UNICEF's help the process of identifying and recruiting interviewees would be straightforward. The original plan was to go through the database and directly contact interviewees asking them whether they wished to partake in the study. However, UNICEF's strict privacy policy, which was unclear at the time of preparation, prohibited the researchers to access to individual's phone numbers. To address this, UNICEF sent SMS messages to hundreds of U-Reporters asking them if they were interested in participating in the research and, if so, to text their phone number, so that they could be contacted directly. This additional step immensely complicated the interviewee recruitment process resulting in significant time delays" – *Evangelia Berdou, IDS*

Moving from Scoping to Designing?

Is there now sufficient information and understanding to design the evaluation?

Have the various sources of information been explored and utilised?

Considering each of the five lenses during scoping will have increased awareness of:

- The aspects of the wider environment: e.g. political and legal context, social/ cultural environment, information and communication technologies landscape;
- The nature of the DCE project: history and background; scale; whether it is stand alone, or a component part; what stage it is at; what its purpose is (i.e. what it is being evaluated against);
- The target group's significance and characteristics, the extent to which they've been involved in the project, their standing in relation to the project, dynamics and influences of the wider environment;
- An understanding of data availability, quality and accessibility (including from previous evaluations);
- Relevant resource constraints and time limitations.

Given the learning from the scoping exercise, can the DCE project or component can be meaningfully evaluated within the parameters and constraints identified.

- If yes, what are the key factors to keep in mind on the move into design?
- If not, is there further data that can be gathered in order to make the evaluation feasible? Are there other planned evaluations that could be linked with? Are there any narrower aspects of the DCE which could be usefully focused on, recognizing there may not be a complete picture, but may present valuable learning on one aspect?

Scoping

4.2

Designing

- Purpose & Goals
- Designing Evaluation Questions
- What Data
- What Methods

Planning & Implementing

Analyzing

Scanning, Reflecting & Learning

This stage builds on the information and knowledge gathered during the Scoping stage to begin the high-level and strategic design of the evaluation. This means agreeing the focus, goals and objectives, designing the evaluation questions, and deciding on an appropriate approach and method to achieve those goals in a way that is feasible and grounded in the reality of the project, the people involved and the wider environment. Detailed design decisions over subjects such as data collection are made in the subsequent Planning section.

4.2.1. Purpose and goals of the evaluation

It is important that the main purpose of the evaluation is decided and agreed on by key stakeholders as early as possible in the design phase. A key question is: what are the primary purposes and intended uses of the evaluation?

The purposes of an evaluation are important as they will inform, and be informed by, the evaluation timelines, resources, stakeholders involved and choice of evaluation options taken. They can and often do vary, but care should be taken not to define the purpose too vaguely, eg the evaluation will be used for 'learning' or 'accountability' or for examining 'value for money'. It is important to consider whether the primary purpose is related to using the *findings* of the evaluation or for using the *process* of the evaluation.

TABLE 9. PURPOSE OF EVALUATION: FINDINGS VS PROCESS

Using findings	Using process
• Contribute to broader evidence base • Inform decision making aimed at improvement (formative) • Inform decision making aimed at selection, continuation or termination (summative) • Lobby and advocate	• Build trust and legitimacy across stakeholders • Ensure accountability • Ensure diverse perspectives are included, especially those with little voice • (Better Evaluation, 2015)

Evaluations focused on learning need to identify who will be learning, about what and through what means? Will it be supporting ongoing learning for service delivery improvements or learning about 'what works', 'best practice' or 'what works for whom and in what circumstances' to inform future policy and investment?

It may be possible to address several purposes in a single evaluation design but often there needs to be a choice about where resources will be primarily focused. Reference to possible constraints on the evaluation is important to ensure *realistic* purpose and *achievable* goals. Agreeing the purpose of the evaluation is critical at this stage.

In addition to the information and understanding of the project developed, the Scoping stage should also have resulted in a clearer understanding of what the focus of the evaluation needs to be. Usually, as this understanding emerges, it will become clear that the evaluation will focus mostly on issues relating to a limited number of the lenses.

Returning to the five lenses is helpful when starting to design the evaluation and will help ensure important DCE considerations are borne in mind as the evaluation questions are decided upon. ***Table 10*** below repeats the key design considerations for each lens, outlined in ***Section 3.5*** above.

TABLE 10. DESIGN CONSIDERATIONS USING THE FIVE LENSES

Lens	To consider during the Design Stage
Objective	Determining and clearly outlining the objective including target audiences and outcomes.
Control	How to evaluate how citizen feedback directly or indirectly impacts outcomes
Participation	How to assess whether the level of participation met objectives, and what factors contributed to the level of participation their objectives
Technology	How effective the technology is and the quality of how it (and the overall project) is managed, cost-effectiveness in comparison to alternative approaches, quality of data safeguarding
Effects	How to establish whether the intended impact materialized, how to notice and assess unintended consequences, the cost of collecting data on the outcome of interest, the contribution of technology to the identified changes

Digital mapping throughout the evaluation cycle

Digital mapping can be used in different stages of the evaluation cycle. Perhaps the most interesting uses are when community groups are engaged from the very start and they themselves map their communities and get involved in measuring changes. Once a base map is established and agreed by the community it can serve as a basis for identifying issues, planning actions, requesting and allocating resources and doing before/after comparisons, including everything from changes in forest coverage, location of latrines and changes in the frequency of open defecation, incidence of violence, overall tracking of completion of community-led or government-funded projects.

As global positioning systems (GPS) become more common features of mid-range mobile phones it has become easier to include geo-location in simple mobile-based surveys and to then create geographic visualizations with the data. The POIMapper application, for example, has allowed Plan Kenya staff to collect geo-located data on community programs and upload it to a monitoring system complete with photos and survey data. Pact has used maps in its work with community forestry groups in Cambodia, where community sketches are transferred onto digital maps which community members use to patrol the forest and monitor forest use, watersheds, timber resources, boundaries, conditions and conflict areas.

SlumDwellers International supports local organizations to survey and map informal settlements, complete with profiles and boundaries. The information has helped local governments secure accurate digital maps of the settlements and has influenced plans to provide budget for upgrades. Having these maps can help communities to follow up and advocate for plans and promises to be made reality.

Mapping and GPS are not without concerns, however, including the potential for those who wish to remain below the radar to be suddenly put into the light and counted by the authorities or identified as a particular ethnic group, or for sensitive information to be linked to individual locations. Tracking violence, conflict or corruption can put citizens at risk if the potential for harm is not fully analyzed and mitigated. Furthermore, it is difficult to measure and track the direct relation of tools such as mapping and geo-located citizen reporting to actual outcomes. At best, contribution of mapping and GPS to overall efforts could be found, but attribution is difficult, meaning that the actual impact of GPS and mapping is difficult to isolate and prove.

Linda Raftree
Co-Founder, Kurante
www.lindaraftree.com

4.2.2. Designing evaluation questions

A key part of the Design stage is the formulation of the evaluation questions. Although certain lenses may become more or less important, it is recommended that every lens is considered, as all five lenses have a direct bearing on both the type and the breadth of the questions asked. The number of evaluation questions per lens may also be significant. A perspective that carries more importance will often require a wider range of questions or greater probing to uncover findings at a deeper level.

The evaluation questions can also carry different degrees of importance depending on the defined goal and objectives, the relative importance of the each lens, and the expectations of the audience and other stakeholders. This hierarchy also points towards the type of data needed and its relative importance, and may point to further work being required at the Scoping stage (illustrating the importance of the iterative approach in this regard).

Once agreed, the evaluation questions take over from the lenses as the principal source of focus and guidance for the following stages of the evaluation—the analysis of existing data and the collection of primary data. Given this, another important aspect to consider is the representation of perspectives, e.g. the perspective of citizens, civil society, government agents and funders, in the questions. When making an evaluation single or multiple perspectives about the objectives and impacts of DCE processes should be accounted for. Therefore, it is important to consider how the evaluation questions reflect (or require in their answers) different perspectives and, crucially, to be clear on which perspectives are being left out or not considered.

Table 11 shows examples of the sort of evaluation questions that can emerge from using the five lenses at the Design and Scoping stages. Whilst some of these are questions restricted to this stage, many of them are also carried over to form the basis of final research questions. Careful scoping and design can often form part of the actual evaluation work, so effort is seldom wasted at this stage.

TABLE 11. SAMPLE EVALUATION QUESTIONS ACCORDING TO THE 5 LENSES APPROACH.

Lens	Sample of typical evaluation questions
Objective	• What are the goals and objectives of the DCE? • Do the goals appear reasonable, practical, sensible? • Is there a clear objective in the project linking, for example, activities, objectives and goals? • What is the external reality of the program and how does this impact on the program?

Lens	Sample of typical evaluation questions
Control	What actors define/defined the project goals and objectives? To what degree are citizens, stakeholders, beneficiaries, end-users… engaged in the initial decisions, the design of the DCE process, the design of the technical platform, the delivery, the monitoring and evaluation..? Who participates at each stage of the DCE? Are there vigilance mechanisms in place and suitable levels of transparency to protect against corruption?
Participation	What characterizes the target audience in terms of availability, environmental/ societal influences, access to the engagement technology, desire to participate? Who is engaging? Are those who engage in DCE different from those who engage in CE? How are they engaging in DCE? Which interests and groups in society participants claim to represent?
Technology	How successful is the DCE? How is this measured? What are the weaknesses and fractures of the DCE process? What is the potential for abusing/manipulating/co-opting the DCE process? What can be improved upon? How does the program handle privacy issues resulting from citizen data being kept on a technical platform?
Effects	Has the DCE resulted in a change of government/citizen behaviour? Do the methods used have both internal and external validity? What indicators do we use to measure "Effects"?

These kinds of scoping and design research questions were used extensively during the field evaluations that inform this framework. Looking at the specific questions used in those studies can help clarify how and where this influence and framing against the lenses has been most useful. Specific examples from those four field evaluations are in **Table 12** below and a longer list of potential evaluation questions can be found in **Toolkit 1**.

Note that before moving on to consider what types of data are needed to answer the evaluation questions, it is worth taking a step back to ensure that the focus, goals and objectives of the evaluation are clear and appropriate; that the questions formulated can realistically be answered (given the findings that emerged during the Scoping stage and the focus and constraints of the evaluation) and that the questions will likely generate answers that will enable the evaluation goals and objectives to be achieved.

TABLE 12. CASE STUDY EVALUATION QUESTIONS (USING THE 5 LENSES APPROACH)

Evaluation questions			
BRAZIL	CAMEROON	UGANDA	KENYA
Lens | Objective			
Are the goals of the *Sistema* clear and appropriate?	How effective is the SMS platform in achieving the DCE objectives of the program?	What is the purpose of the U-Report?	Does the MajiVoice platform effectively handle customer complaints and feedback?
Lens | Control			
Which actors control the budget being allocated to the participatory process?	Which actors control the participatory budgeting process?	Which actors drivethe U-Report and during what stages?	Which actors control the platform use?
Lens | Participation			
Does online voting affect the level of turnout? Do online and offline voters have different demographics? Do online and offline voters engage with the participatory process in different ways?	Who does the SMS platform engage with? How do they compare to the general population, how do they compare with the people that engage in the budgeting process without engagement via SMS?	Who are the U-Reporters and how do they compare to the Ugandan population? If not representative, what could be done to obtain a more representative sample of the population?	To what extent have the different digital feedback mechanisms have been used, by whom and for what purpose?
Lens | Technology			
What opportunities for abuse exist in online / offline processes? What transparency and oversight do citizens have of the *Sistema* and of the implementation of the results?	How effective and efficient is SMS as an engagement tool compared to other media? What benefits/ drawbacks does it provide in its current form?	What are the limitations and opportunities for expression and representation supported by the platforms? How does the data collected through U-Report compare to those obtained by traditional means?	What is the impact of the digital feedback mechanisms on the propensity of people to provide feedback?

Evaluation questions			
BRAZIL	CAMEROON	UGANDA	KENYA
Lens | Effects			
Do online and offline voters vote differently and how does this affect government spending? What opportunities for individual citizenship and empowerment are available to online and offline voters?	To what extent does the SMS engagement strategy increase participation in the process? Does it change the nature of participation?	What types of change does U-report bring about? Does U-Report affect the decisions of government officials?	What is the effect of the digital feedback process on participants' attitudes, perceptions and performance (providers of feedback and receivers of feedback– organizational and individual level)?

4.2.3. What types of data are needed to answer the questions?

Once the evaluation questions have been agreed, the approach itself can be designed. This involves assessing what data is needed (remembering the importance of obtaining different perspectives), identifying and assessing whether any existing system data is sufficient, deciding whether new data must be collected, and deciding the broad approach to be taken in collecting the data. An evaluation matrix or simple decision tree can be extremely useful in this process.

Example: Creating an Evaluation Matrix

When you have identified options that might be suitable for answering key evaluation questions, create a matrix of the questions and potential options for data which may help to answer them. This matrix can help you check that the planned data collection will cover all the questions sufficiently, and enable you to better see if there is sufficient triangulation between different data sources (Better Evaluation 2014c).

	Participant questionnaire	Key informant interviews	Project records	Observation of program implementation
KEQ1 What is the quality of implementation	✓	✓	✓	✓
KEQ2 To what extent were the program objectives met?	✓	✓	✓	
KEQ3 What other impacts did the program have?	✓	✓		
KEQ4 How could the program be improved?	✓	✓		✓

As with all evaluations, the nature of the evaluation questions determines whether the data needed is quantitative and/or qualitative (and often it will be both). In DCE, quantitative data could include participant demographics, self-reported perceptions and number of exchanges (such as online corruption complaints or freedom of information requests). Qualitative data could be from, e.g., interviews to get the government or citizen perspective and perceptions of 'success'.

In addition to considering the broad nature of the data required, decisions around what data is needed will also be informed by decisions on the indicators to be used[1]. In doing so, it is also worth considering choosing universally applicable (standardized) indicators in order to gain greater consistency across DCE project evaluations. The Strategic Framework for Mainstreaming Citizen Engagement in World Bank Group Operations proposed standardized results indicators for Citizen Engagement at both intermediate and final outcome stage (World Bank, 2014b). Some examples include: percentage of beneficiaries that feel project investments are reflecting/reflected their needs; percentage of beneficiaries satisfied with specified dimensions, e.g. access, delivery time, responsiveness to needs, transaction costs, operational efficiency, bribery experience, staff attitude, quality of facilities; reports/allegations of abuse/waste/corruption; investigations resulting in remedial actions and/or sanctions. The full draft list can be found in **Appendix C**.

A digital element would measure both these indicators and the impact and effectiveness of the ICT medium, e.g. number of SMS sent/percentages responded to etc. Whichever indicators are selected, where possible, these should be tested at small-scale or with a sample group, and adjusted if necessary, before full deployment commences.

4.2.4. Is new data necessary?

Once it is clear what data is needed, the decision needs to be made as to whether the existing and available data identified in the Scoping stage is sufficient to answer the evaluation questions, or whether new data needs to be collected.

At this stage, an initial, superficial sweep of existing data will indicate how useful it will be. In reality, there is rarely perfect data—a key variable might be missing, it doesn't cover the full period in question, data quality is an issue, etc.—but the question should ask 'is the data sufficient for the purposes of the evaluation?' rather than 'is the data perfect?'.

Once data gaps have been identified, a decision can be made as to what (if any) new

1–This guide uses the definition of an indicator as the measurement of *"an objective to be met, a resource mobilized, an effect obtained, a gauge of quality or a context variable. An indicator produces quantifiable information with a view to helping actors concerned with public interventions to communicate, negotiate or make decisions"* (Tavistock Institute, 2003). For indicators to be most effective and informative then they benefit from being SMART (Specific, Measurable, Achievable, Realistic and Time Bound).

data will be needed and the time/budget available to collect this. It is also worth considering how any new data will be validated, stored, published and shared. An extended evaluation matrix or decision tree can be useful in mapping data needs, sources and collection requirements.

Thinking about the whole data lifecycle from the beginning will help ensure that the right type of data is being collected, potentially saving a lot of work later.

4.2.5. Deciding on a method for collecting new data

Choosing the right methodological approach for an evaluation is a critical decision and involves a high degree of technical knowledge. Although a detailed discussion of the different methods is beyond the scope of this guide, this brief introduction is provided to compare the most common and relevant methods to facilitate discussion between commissioners of evaluation and experienced evaluators. More advanced guides to research methods are included in the **Further Reading** at the end of this section. Practitioners with less experience in some methods may find the 'key factors' boxes at the end of each method more helpful than the broader discussions.

This section reviews eight common evaluation methods (adapted from Nabatchi, 2012, Shadish et al., 2002, and Shen et al., 2012). Each method has strengths and weaknesses, often depending on the nature of the data required to examine if and how the objective of the intervention connects with key outputs to outcomes. Not all of these methods are necessarily of particular relevance to DCE, but they are important for *all* evaluations, and all of them benefit from data collection via digital tools such as SMS, multimedia platforms and online questionnaires. The first four sit on a spectrum of how causality can be established they are[1] (true experiments; field experiments; ex-post-facto design; quasi-experimental: non-experimental), while the following four are types of non-experimental studies with a more qualitative nature (theory-based; participatory; ethnographic; case study). However, all of these methods can, in fact, include both quantitative and qualitative analysis.

For each method a brief description is provided, followed by suggestions of when it is more or less suitable for use, and key factors to consider when used for evaluating DCE programs.

4.2.5.1. RANDOMIZED CONTROL TRIALS

Randomized trials (or randomized control trials–when a control group is present) are often seen the most rigorous methodological approach to study impact. Ran-

1–A true experiment has two main characteristics: (1) randomly assigned groups–every participant has an equal chance of being in the experimental or control group(s); (2) manipulation of a variable where different groups are subject to different treatments or absence of treatment (control group). Natural experiments or quasi-natural experiments are studies where subjects were grouped by nature or by other factors outside the control of the investigators.

domized trials involve randomly assigning interventions (treatments) to subjects, for instance different persons get different outreach messages for a particular citizen engagement project. Randomized trials that take place in the real world as part of an ongoing intervention with real participants are called field experiments. However, randomized trials can also take place in more controlled or laboratorial settings, involving small-scale treatments with recruited participants.

Experiments can help us understand how effective an intervention really is by creating the conditions for comparing two or more scenarios (treatments or levels of independent variable). Duflo et al. (2012), for example, conducted an experiment whereby teachers in randomly selected schools were monitored daily using cameras. Attendance was rewarded with a salary increase. The researchers compared the attendance rate of this (treatment) group with teachers from schools that went about their business as usual (control group) and discovered that the attendance in the treatment group increased by 21%. The random selection of schools makes it possible to directly attribute this increase to the treatment being evaluated.

RCTs are regarded by many as the most rigorous appraisal tools to establish cause-effects relationships. This is based on the view that without random assignment of cases (e.g. individuals or schools) into different groups and direct manipulation of independent variable(s), one can never be certain that the observed effects are the result of the intervention and not pre-existing differences between the groups (classificatory factors) or situational or environmental variables (pseudofactors). Quasi-experiments share with true experiments the manipulation of independent variables, but fail at assigning participants at random to the treatment groups (they use natural occurring groups). For instance: without random assignment, how can one ensure that participants in a DCE initiative feel empowered because of the intervention or because the project had features that attracted well-educated citizens confident about their rights and voice? It is very difficult to answer this question confidently after the intervention has taken place.

Experimental designs can also be useful to test particular aspects of a DCE intervention. For example, an RCT study in Uganda (Grossman et al., 2014) varied the price of using an SMS system that allowed citizens to reach Members of Parliament to test if and how the cost affected usage.

KEY FACTORS WHEN CONSIDERING USING AN RCT

Despite the positive aspects of RCTs, there are important potential drawbacks. RCTs that test the effectiveness of the intervention in its entirety (rather than certain aspects of it or design features) can be expensive. However, the digitally mediated character of DCE interventions opens up new opportunities for testing and improving upon key design features quickly and cheaply and means *digital* random experiments can sometimes be a highly cost-effective option.

Furthermore, RCTs work better when: (Stern et al., 2012:38–39):

- There is only one primary cause and one primary effect. This might not be the case in more complex interventions;
- The control group is not 'contaminated' by the intervention (i.e. the 'non-treated' individuals need to be unaware of the treatment and with no contact with 'treated' individuals) so comparisons between treatment and control groups is valid;
- The focus lies on the success of a particular intervention. Generalisation to other individuals, times, contexts or interventions are not feasible (the problem of external validity);
- Understanding what works is more important than understanding why it works.
- The group sizes support statistical analysis (at least 30 cases per group)

Recruiting experiment partners for a Randomised Control Trial

To do an RCT, a partner usually needs to be recruited, often a small organization or voluntary group. Researchers rely on the organization agreeing to the experiment and understanding what is needed. It can be costly if the partner starts to doubt the experiment once it has started. The partner might also be wondering whether they are working with the right researchers. Partnerships can be like a dating game—early contacts are important in determining whether the partners like each other and where exit strategies are available if the liaison is not going to work.

How best to start? Do you go in at the top, say at the political level, and write a letter saying 'I would like to do an experiment, please help me'? Such a letter might work or might not be answered or be answered negatively. In other cases, it is better to approach personnel lower down the organization who deliver services. If they become enthused about the experiment they can seek higher-level authorization when the time is right. Informal contacts also play a role if the researchers already know the people involved by going to meetings and social gatherings, which is where ideas can be discussed and followed up.

The first planning meeting between researchers and partners is very important. To have got that far is often a good sign that internal conversations have taken place. But such meetings do not necessary go well as the organization can realize the costs of the intervention. But sometimes the worst meetings are where it goes too well: everyone is excited by the trial and there is a warm glow all round. The cost of the RCT only becomes apparent much later on. It is a good idea to balance out creating enthusiasm with conveying what is involved. It is important to go through each stage of the trial and work out what the researcher and the partner need to do.

Professor Peter John
Professor of Political Science and Public Policy, University College London
www.ucl.ac.uk/spp/people/peter-john

4.2.5.2. 'EX POST FACTO'

In ex post facto design, investigators begin their study after the intervention has taken place without manipulating independent variables or assignment participants at random (Silva Carlos, 2010). This type of design allows evaluators to weakly identify cause and effect relationships. In this type of strategy the researcher takes in the effect (dependent variable) and looks back in time to investigate possible causes, relationships and associations between natural occurring groups.

In this type of design the investigators compare groups who were exposed to a certain intervention (cause or independent variable) in terms of an outcome of interest. For example, comparing the voter turnout in provinces where people have access or have not access to online voting in participatory budget initiatives. It is also possible to use the reverse strategy to look back in time to investigate possible causes of the differences between two groups. For example, users and non-users of a complaint platform could be compared in terms of satisfaction with the service provided or other characteristic (e.g. level of education). Along these lines, subjects who differ on the dependent variable can be the starting point, and inspect difference on an independent variable (age, education, occupation). It may also be necessary to understand why an intervention affected the participants in a certain way (Cohen et al., 2011). Because the ex post facto design is aimed at inspecting dynamics that occurred in the past, the most common method for data gathering is structured questionnaires about past events, personal values or motivations or socio-demographical characteristics.

KEY FACTORS WHEN CONSIDERING EX POST FACTO EVALUATIONS

Ex post facto designs are useful when:

A more rigorous experimental approach is not possible for ethical or practical reasons

- Studying conditions
- Weak cause-and-effect relationships are being explored to be tested later through a RCT
- Studying the effectiveness of an intervention on naturally in existing groups

One of the main weaknesses of ex post facto designs is that they often cannot help establish the direction of causality (what is the cause and what is the effect) and rule out other explanations for the outcome that may have co-varied with the intervention, leading to potential confounding of causes (third variable problem). In addition, it is prone to self-selection bias as users and non-users may differ a priori on a number of characteristics.

4.2.5.3. NON-EXPERIMENTAL DESIGNS (CORRELATIONAL DESIGNS)

Unlike the designs presented thus far, non-experimental or correlational designs are not set up to examine cause and effect but only associations between events or characteristics of participants. These type of methods do not rely on manipulation of independent variables, but on measurement of potential independent and dependent variables to understand how they are associated. Surveys are the prime method for data collection in correlational designs to explore associations between different variables (e.g. age and frequency of participation). Their key strength lies in that they can be used to explore a broader range of hypotheses than experimental designs.

Surveys are commonly used to measure the characteristics, knowledge, attitudes opinions and practices of a given population. In a DCE context, a population can be defined broadly (to refer, for example, to all users of a platform) or narrowly (to include for instance only high-frequency users). A significant correlation between two variables might hint to a cause but is not in itself enough to establish a cause and effect relationship. There are also statistical strategies (e.g., partial correlation or multiple regression analysis) to control for the effects of other variables (measured and included in the analysis) when an association between two variables is found. Correlational designs can, however, be taken a step further to examine how different variables contribute to an outcome. For instance, what factors can better predict whether a participant will contribute more? Is it mostly gender, or is it an interaction of gender and education? However, it is important to think in advance about what kind of analysis might be required for a particular evaluation since this has a bearing on how variables are measured.

An important aspect of a correlational design is *sampling*. Sampling is the objective according to which the evaluator selects who will respond to our survey is selected to maximize the validity of the results. Although in the context of a DCE initiative, it might be possible to conduct a survey of all registered participants (called a census) this might want to be avoided. It will be too costly, time-consuming and could lead to bias, if for example, only high-frequency users responded. Sampling strategies can vary depending on the time and human resources and whether a lists for the population (sampling frame) or accurate population figures exist.

In the context of DCE survey designs, it should also take into account important contextual factors that might render some of the variables meaningless. For example, when conducting a study in an African country it is not advisable to use the employment categories that are used in surveys in western countries. This is because very few poor people are waged employees and make do by doing a little bit of everything-having a small stall at the market, working occasionally in construction, etc.

In the context of the DCE, where surveys can be conducted relatively cheaply over SMS the Internet and with automated voice systems, this temptation becomes even more appealing. Another good strategy to follow when designing a questionnaire for a DCE initiative is borrow questions and adapt from other sources in areas relevant for the evaluation (some examples are included in **Further Reading** at the end of the section).

KEY FACTORS WHEN CONSIDERING USING CORRELATIONAL DESIGNS

- Correlational designs are not suited to examine cause and effect relationships, but they are ideal for examining a broad range of hypotheses.
- The generalizability of findings will largely depend upon the nature of the sample. A rigorously constructed and sizeable sample is therefore an important aspect of this research design.
- In designing their questionnaires evaluators need also think about the types of analyses they would do as this affects how indicators are measured.
- Survey questions need to be informed by an understanding of the context of the evaluation.
- ICTs create new channels for delivering their questionnaires (SMS, face-to-face, telephone). However its option should be considered carefully as it will introduce its own biases in the data collection process.

4.2.5.4. THEORY-BASED DESIGNS

In this approach, the intervention is considered as a conjunction of 'causes' that follow a sequence like the pieces of a domino. The 'objective' lens is a variant of this research strategy (section 3.4.1). This approach develops an understanding not only whether an intervention works but also what are the conditions that make it work. Theory based methodologies are making a come-back in evaluative inquiry. This is because being able to say that x caused y often does not allow us to understand why some things work and do not work. Theory-based designs allow us to explain why and how exactly an intervention led to the desired change.

There are weak and strong versions of theory-based designs (Stern et al., 2012). Weak theory-based or program theories are usually no more than objective models that depict, usually diagrammatically, how an intervention is understood to contribute to intermediate outcomes and long-term impacts. Richer objective models might include the views of different stakeholders. Stronger theory-based designs go into more detail to identify not just the steps that lead from goals to outcomes and impact but the exact mechanisms that make things happen, that can be highly contextual. A program's theory usually combines a *theory of change*, an understanding of the process through change is seen to come about and a *theory of action*, the way in which an intervention is implemented to realise the theory of change.

Theory-based designs form an integral part of most evaluations because of their usefulness in enabling experienced evaluators to develop quickly some understanding of the main strengths and weaknesses of an initiative and to adapt the evaluation according. Stronger theoretical designs, especially when they are informed by relevant literature, can be helpful in generating hypotheses for the evaluation and helping to better understand cause and effect relationships.

KEY FACTORS WHEN CONSIDERING USING A THEORY-BASED DESIGN

Theory based designs are to a lesser or greater degree part of any evaluative strategy as they help evaluators appreciate how a program is understood to translate goals into intermediate outcomes and long-term impacts. Whereas RCTs help answer the question of whether an intervention worked or not, a theory-based design, through the development of nuanced hypotheses can help explain why it worked and why it didn't.

Theory based designs are particularly relevant for DCE where it is often assumed that the benefits of using digital technologies will flow automatically.

A Note on Ethnographic project evaluation

Ethnography is both scientific (to understand the world through the empirical senses) and expressive (to share evidence from this engagement through evocative writing and multi-media techniques techniques). The method contributes to evaluating research in two important ways: to understand grounded knowledge embedded in the praxis of social life; to provide a credible account of context and its relationship to lived experiences of individuals.

Ethnography can illumine the relationship between social context and digital content from the perspective of the digital citizen. An example of unraveling digital citizen engagement would be the following: Ethnography can contribute in the understanding of urban slum youth's engagement with social media, particularly Facebook. With extended field immersions and thick descriptions of the slum context and youth social media activity, insights on a number of social relationships can be established: how do slum youth make meaning out of Facebook and the ensuing digital affordability at once unique and challenging? How is Facebook configured by youth in the context of extreme lack of privacy in an urban slum context while allowing a certain control over ones representation and expression? How is Facebook a gateway to many unattainable expressions of one's personal and social compass? Is Facebook the first digital engagement with global experiences?

Ethnographic methods can help explore and answer deeper questions that more traditional methods offer little insight into. Results from such an ethnographic engagement can influence developmental engagements with non-elite sections of society, guide policy to better serve these groups.

Nimmi Rangaswamy
Xerox Research Centre India, Indian Institute of Technology Hyderabad
www.iith.ac.in

4.2.5.5. ETHNOGRAPHIC DESIGNS

Here the emphasis is placed on understanding the context in which the intervention takes place and the perceptions, views and ideas of stakeholders and the spaces that shape it. Ethnographic research prioritizes understanding the context of people's everyday lives, their shared social norms, priorities and assumptions.

Compared to experiments and survey-based designs which often adopt narrow definition of why, what and who, ethnographic data provides rich and holistic insights into people's views, habits and actions and the spaces that they inhabit. Ethnography relies more on individual and focus groups interviews, observations, and presupposes the immersion of the researcher into the setting where participants are located, often for long periods of time. Although this may be impractical, the spirit of ethnographic research can be adopted and adapted to the quick turnaround time of evaluations. Ethnography makes researchers an integral part of the study, inviting them to reflect on their own biases and assumptions and those made by the project. Key challenges of ethnography include that findings are context specific and may not be readily generalizable, and it can be time intensive both in terms of data collection and analysis and in many cases, especially when English is not spoken widely, requires the use of translators. In the context of DCE, ethnographic research can be invaluable in supplementing or informing the analysis of big datasets generated through the digital platforms by helping evaluators develop a sense of how the data is being generated and used, what practices and ideas drive participation and data use.

KEY FACTORS WHEN CONSIDERING USING AN ETHNOGRAPHIC DESIGN

- Ethnographic designs and methods are invaluable in developing a sense of the context of the evaluation and in informing the design of other methodologies of data collection and analysis.
- Ethnography need not be an all or nothing proposition. Used strategically, ethnographic methods such as observations / interviews can help refine the design of other evaluative tools and analyses.
- The selection of interviewees is important in ethnographic designs as responses, reactions, and the expressions of views can be influenced by power dynamics between interviewees and the interviewers.
- The use of experienced local researchers can be help to mitigate some of the costs in data collection and analysis and help overcome language problems.

4.2.5.6. PARTICIPATORY EVALUATION

Participatory evaluations involve the participation of project participants from the start to define the goals, agendas, measures against which a program will be evaluated. It often incorporates tools such as story-telling, scoring, social mapping, trend and change diagramming. The basis of causal inference here lies on the validation of a programs outcomes and impact by participants. By having program participants as partners in the research and privileging their ideas and opinions, it helps clarify the context and challenges that the initiative seeks to address, improving a DCE's relevance and increasing ownership. Participatory designs involve a panoply of methods that include the collaborative creation of matrices, maps, flowcharts and timelines, or questionnaires to review existing information, such as assessing program goals and outcomes, to plan and assign roles and responsibilities (Chambers, 2008). Digital storytelling can be an especially powerful tool for expression and learning.

Some aspects of participatory evaluation may be blended with other research strategies. Participants might be asked to define, for example, what success means in the context of the initiative and their definitions might inform a questionnaire design. However, it can also be costly, requiring a significant degree of commitment on the part of the participants.

KEY FACTORS WHEN CONSIDERING USING PARTICIPATORY EVALUATION

In addition to helping define locally relevant evaluation questions and indicators, participatory research can increase participants' sense of ownership of the project. However, it also requires a significant degree of commitment on the part of the participants and the evaluators in the process to ensure that the views and opinions raised in the process are taken into account.

The use of technology which involves the communication of highly complex terms and processes to participants could be a challenging but also potentially rewarding exercise. What do participants make of the new data flows? How do they address issues around anonymity?

Similarly to ethnographic designs, participatory designs are sensitive to power dynamics and elite capture in particular.

4.2.5.7. CASE STUDY DESIGN

The term 'case study' is often used as an umbrella term for a collection of research methods that are combined to examine an 'instance' of a project, group, organization, event. This 'instance' in research terminology is often referred to as a 'unit'. In the context of the DCE a 'unit' can be defined in a number of ways. One can chose to define an entire project as a unit (holistic design). Such a design would treat the project as a whole in terms of the five lenses. Alternatively the project can be treated as consisting of a number of different components, different organizations (including partner organizations) different participant groups, different processes. This embedded research design can serve to build a richer and fuller picture of an initiative or different initiatives by incorporating different perspectives and levels of analysis.

A case study research design is by far the more pluralistic of the designs presented thus far in terms of tools for data collection. Although it usually involves some type of ethnographic work (e.g. interviews and observations) it can also incorporate surveys. A key aspect of this type of design which distinguishes it from ethnography is the importance of theory development. This can be in the form a simple hypothesis at the beginning of the evaluation which is informed by relevant literature on what works and does not that can be tested further.

Although case studies are often considered as less scientifically credible due to their perceived limitations for drawing causal inferences and generalising findings, proponents of case studies have argued that, although the case study findings might not be statistically generalizable, they can be *analytically* generalizable " by having the theory developed for the study compared against the empirical findings" (Yin, 2003:32). New methods for the systematic causal analysis of cases such as Qualitative Comparative Analysis (Stern et al., 2012) are attracting more attention as a valid alternative to experimental research designs. Comparative case studies are an effective tool, for example, when one's own case is too small to engage in quantitative analysis, when other methods are unavailable or inappropriate, and/or when the variables are difficult to disentangle from each other.

In a DCE context, a well thought out case study research design combines the advantages of ethnographic work in terms of its ability to yield rich contextual data with some of the key elements of theory-based designs.

KEY FACTORS WHEN CONSIDERING USING A CASE STUDY DESIGN

A case study need not be a study of the intervention as a whole. This design can may be useful in identifying components of the program whose closer examination might bring generate important insights for the evaluators, especially when it is theoretically grounded. Combined with new methods for systematic causal analysis, theoretically sophisticated case study design can be a powerful tool for the evaluators.

4.2.5.8. CHOOSING AND COMBINING METHODS

The boxes at the end of each discussion above should help in identifying how suitable the different methods may be when designing a DCE evaluation. Although, traditionally, some of these methods have been considered as more scientifically rigorous than others, especially when it comes to the examination of cause and effect relationships, current debates suggest that if there is a 'gold standard' to be achieved, it is through methodological pluralism (mixed methods), blending elements of different research designs depending on what is appropriate for meeting the requirements of a particular evaluation and for increasing the evaluation's usefulness for funders, program designers and implementers and participants.

Here the order in which the methods are used (*sequencing*) is something that needs to be considered. "Mixed methods" is not only about using multiple methods to highlight different perspectives, it is about using the results from one method to inform designs of another. *Triangulation*, the process through which one method is used to check the results of the other can also be valuable in a mixed methods approach.

For DCE, such an innovative, mixed-methods approach is supported by the digital character of the intervention. The inherent characteristics of DCE also mean that collection of new data via both experimental and non-experimental methods is often a realistic option in terms of cost and timescale. The ability to reach participants through SMS, for example, may lower the costs for conducting representative surveys, at the expense of exclusion of low-income groups or those with no digital access. Similarly, digital tools allow for low-cost experiments that would have been impractical without large time and budgets using traditional methods.

Low-cost experimentation using social media

Can lower-cost ICT channels increase participation among those who already have economic and political capacity to participate? Over the past year MIT Governance Lab (MIT GOV/LAB–web.mit.edu/polisci/research/govlab.html), mySociety, and Mzalendo (info.mzalendo.com) have collaborated on a research program exploring ways to engage Kenyan citizens and to galvanize political action online. The researchers have developed an 'iterated experimentation' approach–relatively short and small-scale investigations using rigorous social science techniques that build on findings from previous rounds to test which operational and design choices are most effective in particular contexts.

From conversations with citizens and reading news articles, we believed that the information Kenyans receive about government is generally negative and threatening. We anticipated that Kenyans may experience 'threat fatigue' and instead of being motivated to act when there is a perceived threat, may be more likely to engage when there is a perceived opportunity or they receive information about the engagement of others.

We recruited Kenyan citizens interested in finding out about county government management of public funds using Facebook ads. We bought four different ads targeting men and women over and under 30 years old in order to explore how treatments affected subpopulations differently. Each person who clicked on the advertisement was randomly assigned to one of the treatment pages, which presented participants with an article about how county governments have misspent taxpayer dollars. Different treatment pages either framed the information in terms of the threat of lost resources, the opportunity of gaining resources, or neutrally. Additionally, we included a 'bandwagoning' interaction with each presentation that included information about how many other people had also taken action. At the end of the article, we provided four actions viewers could take: share the article on Facebook, share it on twitter, sign a petition, and send a comment to the Senate Majority leader. We compared the differences in number and types of actions taken by viewers who saw each treatment page.

Preliminary findings suggest that groups who have higher socioeconomic status–specifically older men – are more likely to take action via this online platform, and that people are more likely to take action when they receive information about how many others are participating. Findings from this iteration will inform the next round of experiments by testing hypotheses that expand on these results.

Lily L. Tsai and Leah Rosenzweig
MIT Governance Lab (MIT GOV/LAB)
www.web.mit.edu/polisci/research/govlab.html

Key lessons learned from the field evaluations

- Designing -

Try to design the evaluation at the same time as designing the engagement

From Cameroon: "There was great interest in the possibilities of a more detailed evaluation following earlier preliminary evaluation work by the World Bank, as there is a large body of citizen engagement data and information available going back several years. Unfortunately much of this data are not easily accessible. Other than the actual digital engagement platform system data (how many people have been sent SMS messages about the vote for example), it's all largely non-digital, on paper and generally not of great quality. So for example, data are being collected on paper when people are going to participatory budget meetings—in 2014, we were looking at 4 communes, 15 meetings per month each, 7 months, average of 30 participants per meeting = 12,600 participant detail records but all this is being collected on paper forms. If the program had the capacity to collect it in digital form and used digital tools to control the quality of the data collected from early on, then that would have made the analysis, quality-checking and evaluation much quicker and easier. There had been plans to do just this but due to resource constraints within the local program, these were not implemented. That was disappointing." (Martin Belcher, Aptivate)

Different approaches and tools can be used in the same evaluation and within different budgets to reach a range of audiences

From Brazil: "We decided early on we wanted to focus on three groups of people—those who voted online; those who voted using traditional voting booths and those who didn't vote at all. We used SurveyMonkey which popped up on the voting website, asking them if they could answer a few questions. For the face-to-face survey, we had around 50 enumerators around the various polling booths. For those who didn't vote, we did an automated random-digit-dial IVR survey which was an affordable compromise from a full door-to-door household survey—which we did want to do but would have been much more expensive." (Matt Haikin, Aptivate).

Local partners are an invaluable resource for testing and improving surveys

From Kenya: "There was a customer complainant survey—the local water company advised on the questions and gave some useful feedback, for example we wanted to ask about income and they advised that people would be sensitive around that and refuse to answer and so reduce completion rates significantly. So we decided to use a couple of proxy indicators that we could infer income from (rent levels and geographic zoning of complainants—which water company office the complaint would be directed to, regardless of how the complaint came in) to help with our understanding in this regard." (Martin Belcher, Aptivate)

Moving on from Designing to Planning and Implementation?

- Are the purpose and goals of the evaluation clearly defined and agreed?
- Have the evaluation questions been formulated which, when answered, will achieve the goals of the evaluation?
- Has the data been identified that will be needed to provide evidence for the answers to the evaluation questions?
- Is there clarity on the accessibility and quality of relevant existing data?
- Is there an agreed strategy for achieving the goals of the evaluation, based in the reality revealed by the Scoping stage?
- Is the chosen method/combination of methods for gathering the evidence needed to answer the evaluation questions placed appropriately on the spectrum between experimental and non-experimental, quantitative and quantitative?
- Is there an evaluation matrix to support planning?

Selected readings and online resources on designing an evaluation

Accenture Digital Engagement Pulse Survey (2012) – *http://www.accenture.com/us-en/Pages/insight-digital-government-digital-citizens-ready-willing-waiting.aspx*

Afrobarometer economic, social and political opinions/characteristics of Africans across several countries – *www.afrobarometer.org*

Civic engagement in the digital age – *http://www.pewinternet.org/2013/04/25/civic-engagement-in-the-digital-age*

Civic Plus Digital Citizen Engagement survey – *http://go.civicplus.com/l/9522/2012-10-26/8t542*

Conducting quality impact evaluations under budget, time and data constraints – *http://www.oecd.org/derec/worldbankgroup/37010607.pdf*

Mixed-Method Impact Evaluation for a Mobile Phone Application for Nutrition Service Delivery in Indonesia – *http://www.ids.ac.uk/publication/a-mixed-method-impact-evaluation-design-of-a-mobile-phone-application-for-nutrition-service-delivery-in-indonesia*

Decide which evaluation method to use – *http://betterevaluation.org/start_here/decide_which_method*

Fools' gold: the widely touted methodological "gold standard" is neither golden nor a standard – *http://betterevaluation.org/blog/fools_gold_widely_touted_methodological_gold_standard*

How useful are RCTs in evaluating transparency and accountability projects? – *http://www.makingallvoicescount.org/news/how-useful-are-rcts-in-evaluating-transparency-accountability-projects/*

Overview of current Advocacy Evaluation Practice – *http://www.innonet.org/client_docs/File/center_pubs/overview_current_eval_practice.pdf*

Participatory Approaches – *http://www.participatorymethods.org/sites/participatorymethods.org/files/Participatory_Approaches_ENG%20Irene%20Guijt.pdf*

Participatory Evaluation – *http://betterevaluation.org/plan/approach/participatory_evaluation*

Participatory Monitoring and Evaluation: Learning from Change (1998) – *http://www.ids.ac.uk/publication/participatory-monitoring-and-evaluation-learning-from-change*

Qualitative Evaluation Checklist – *http://www.wmich.edu/evalctr/archive_checklists/qec.pdf*

Randomized Evaluations: Methodology Overview and Resources – *http://www.povertyactionlab.org/methodology*

Scoping

Designing

Planning & Implementing

- Collecting new data
- To use Digital Tools?
- Choosing the right tools

Analyzing

Scanning, Reflecting & Learning

This section describes how the design process now moves to a more detailed level to decide what tools to use within the broad method for collecting new data, whether or not to use digital tools to collect new data, and how data collection can be implemented. Implementation of a DCE evaluation is broadly the same as for any evaluation so this is not covered in depth, but some specific tips are included that are of specific relevance to technology and citizen engagement.

4.3.1. Collecting new data

The focus and overall design of the evaluation are now clear. Existing data has been initially scanned for quality and usefulness, new data needs have been identified, and a method for collecting it has been selected. While the existing data which has been assessed being of potential use can move straight to Analysis (Section 4.4.1), more detailed planning is needed for collecting new data.

At this stage there are two key decisions to be made. Firstly, which methods are going to be used to collect the data—these can be either non-digital or digital. The focus in this guide (sections 4.3.3 and 4.3.4) is on exploring the strengths and weaknesses of *digital* tools.

Secondly, the nature and quality of the data to be collected needs to be decided. This can both inform and be influenced by the choice of methods. Some of the tips and considerations below (Section 4.3.2) need to be considered at the planning stage, but some depend on practice during the data collection.

4.3.2. Tips to consider when planning and implementing DCE data collection

Some of these tips apply across the evaluation spectrum as the implementation and delivery of a DCE evaluation is not fundamentally different from the implementation of any other evaluation. However, there are some aspects relevant to the use of technology and/or the participatory nature of DCE evaluations that are particularly worthy of highlighting. It is useful to bear in mind:

- **Involvement of beneficiaries and stakeholders in the planning** as a way of building trust and demonstrating openness and transparency.
- **Responsible data collection** (see box below) and, within that, the question of data protection is an important principle. There is need to use data in a way that respect participants' trust and to be clear that data collected should only be used for the purposes for which it was collected.
- **Sampling techniques** where participants are selected through a random procedure (probabilistic technique) imply representative samples, although a high level of non-response (refusals to participate) may threaten the representativeness. Techniques that do not involve random selection of participants (non-probabilistic techniques) are less demanding but produce samples that are not necessarily representative of the population, prohibiting direct generalization of results (Trochim, 2006). Within non-probabilistic techniques, quota sampling that fixes the percentage of certain groups (e.g., gender and age groups) in the sample to be the same as in the population is less problematic, allowing some degree of extrapolation of results to particular popu-

lation groups. When dealing with non-probabilistic sampling techniques, it is important to bear in mind that digital tools for data collection, e.g., mobile phones or internet, are more prone to produce biased samples as their uptake varies across gender, age or education levels (as demonstrated in the Uganda U-Report's evaluation).

- **Match variables** if secondary data is used to complement primary research, for example, age or income bands in a survey should match with the bands used in census data.
- **Re-use best-practice survey questions** that have been designed and tested in other studies with similar populations or found in searchable in questions banks, e.g. Survey Questions Bank by UK Data Service (http://discover.ukdata-service.ac.uk/variables).
- **Pilot the method,** e.g. survey or interview guide, with the exact audience in which it will be undertaken in order to detect problems in questions or answering options, spontaneous reactions of respondents, refusal rates and timing response rates/drop-off points.
- **Make sure all the data needed is being collected** including any data related to the control group. It is easier to get this right first time than to go back and collect more data later.
- **Use of intermediaries** when conducting the DCE evaluation means they may bring in their own agenda/prejudices and become more than just a channel, introducing their own biases and dynamics with the respondents.
- **Unintended bias in interviews** as interviewees may frame answers in a particular way for a variety of reasons (hence the need for triangulation). There may be different effects in responses between one-to-one interviews and group scenarios—such as focus groups—where respondents may be influenced by others in the group, either openly or tacitly. Other influences can be whether the interviews are recorded or not, the language they are conducted in and, if an interpreter is involved, the biases they bring in.
- **Negative impact on participants** as those invited to participate in DCE are often time poor—especially if they are women—and may have historical reasons to expect little responsiveness from their governments. It is important to bear in mind that these factors may influence their response.
- **Consider the effect of using technology** and how it may affect the response, for example, when a recording device is switched off, an interviewee tends to

speak more comfortably; if a tablet is used to collect responses in the field it could be that the interest and questions from respondents are about the technology itself; the respondent may even take the artefact and start interacting with it. In some cases, people may not feel comfortable enough to admit being unable to use or understand the device.

- **New tools and technologies** appear constantly, and some are valuable and some are less so. It is important not to be fooled by the marketing of the technology vendors or by whether the technology is open source or proprietary. The tools need to be researched properly, the people using them need to be spoken to and there needs to be an understanding of similar tools and the benefits and potential pitfalls they might bring before deciding on their use.

Responsible data collection – from extraction to empowerment

Responsible data can be defined as the "duty to ensure people's rights to consent, privacy security and ownership around the information processes of collection analysis storage presentation and reuse of data while respecting the values of transparency and openness." (Responsible Data Forum: working definition, September 2014.) While data has invaluable potential to make organizations more needs driven and responsive, there are also huge risks to communities if related processes are not responsibly designed or managed. Factors to consider are:

Collection and use of data: *data should be collected in a culturally and contextually appropriate manner. Data collection should not put an excessive burden on participants. It is vital to maintain accurate and relevant data representative of populations and ensure to appropriately collect, analyze, utilize, and disseminate information.*

Consent: *it is important to gain informed, voluntary consent before obtaining information. Informed consent is a process for getting permission to collect data of any kind based upon a clear appreciation and understanding of the facts, implications, and consequences of engagement. Participants must be free to choose to consent or not, without inducement, and free to withdraw their involvement at any stage without negative implications, including to their participation in the activity. Consent should be based on transparent and accurate assessment of use of data and context and if the use or context changes, re-consent may be needed. Special considerations must be taken when working with children.*

Privacy: *the process of data collection must be conducted in an environment where privacy of the individual is maintained. We must anonymise data as early on the data collection process as possible and limit the collection of personally identifiable information (PII).*

Risk mitigation: *do not collect unnecessary identifying information that could put participants at risk without viable justification (religion, ethnicity, victims of assault etc.) and ensure that the effect of actions have no negative physical, psychological or political consequences on the participants.*

Oxfam and others are increasingly recognizing the responsibility to represent contributors of data and involve them in the process of how data are used by adopting less extractive and more empowering methods. Digitalisation of processes can ensure more transparency in the collection of reliable and accurate data which can be made accessible in near real time so that once duly anonymised, data can be presented back to those who contributed to it. It also assists with transparent reporting processes to stakeholders to ensure maintenance of high standards. Digitalisation of processes presents new risks, such as vulnerabilities in cloud storage, but also new opportunities such as ability to encrypt or build in good behaviours by prompting or reminding users about best practice, like regular password changes.

Amy O'Donnell
ICT in Programmes Lead, Oxfam UK
www.oxfam.org.uk

Digital innovation supporting cross-country citizen surveys

RIWI's Random Domain Intercept Technology (RDIT) offers an innovative way of surveying global citizens by randomly intercepting web users in every country and territory in the world (and on all web-enabled devices, from smartphones to desktops), it is enabling evaluators to capture the attitudes of citizens in hard-to-reach regions of the world.

This offers the potential to provide a new voice for global citizens that can otherwise often be left out of important discussions and decision making.

RDIT is particularly useful for evaluations requiring large-scale cross-country surveys—previous examples include innovative global pandemic work, in conjunction with Massey College at the University of Toronto, a 450,000 person Global Corruption Index with the International Association of Prosecutors; democratic engagement in Indonesia with the International Foundation for Electoral Systems; real-time tracking of citizen attitudes in West Africa towards Ebola with BioDiaspora; and the World Bank's Open Government Project in 63 countries (www.openinggovernment.com).

Eric Meerkamper
President, The RIWI Corporation

4.3.3. Weighing the use of digital tools in the evaluation of DCE

It is important to consider carefully the pros and cons of digital vs non-digital methods, given the inherent trade-offs associated with deciding between competing approaches. There should be no assumption that one is better than the other. **Table 13** below elaborates on some of the benefits and challenges related to using digital approaches.

TABLE 13. CONSIDERATIONS ON USING DIGITAL DATA COLLECTION.

Benefits	Challenges
• Lowers cost considerably, e.g. using different wordings for target groups and different user interfaces (Raftree and Bamberger, 2014) • Collects real-time data • Triangulation through date stamps/IP address/GPS tracking/location through mobile number if applicable • Greatly facilitates data collection (Raftree and Bamberger, 2014) • Can be conducted remotely • Video/audio recordings, for understanding body language including participatory video (Lunch and Lunch, 2006) • Accurate transcriptions of interviews; in-depth qualitative approaches such as narrative/linguistic analysis can be conducted (i.e. to understand what the speaker puts emphasis on)	• Cost-reduction may not be the main determinant of the evaluation, e.g. in qualitative research, better rapport may be built in asking questions face-to-face through a local evaluator. • Infrastructure challenges, e.g. power cuts (Farmer and Boots 2014) • Selectivity bias, excluding those who do not have access to the technology (Raftree, 2013) • While there is nothing inherently wrong with technology transfer per se, challenges can arise when insufficient attention is paid to consequences of transferring technology into a new context, or where the technology itself rather than the underlying need becomes the main driver • Technology may be viewed suspiciously and as a barrier to rapport building • Loss of privacy and increased levels of risk, especially once data enters the public domain • While technology is often seen to lower barriers of access, it can similarly lend itself to higher 'drop-off points' • Total costs of ownership to introduce the technology to the target group in the long term

A key argument in favour of collecting evaluation data digitally is cost saving, for example, "*one program in Zimbabwe recorded a $10,000 saving by switching to tablets to survey a sample of 5000 people, as compared with using a 25 page paper questionnaire*" (Raftree and Bamberger 2014, p23). The experience from the four evaluations that accompanied the development of this guide leads to a similar conclusion, demonstrating that using digital means can reach more people more quickly at less cost than using non-digital techniques (recognizing that bias may come in other ways, and that cost is often not the only consideration).

At the same time, there is a clear preference for non-digital methods (interviews, household surveys, etc.) where more detailed, nuanced information is required. More detail on the specific data collection methods and costs for these field evaluations can be found in **Appendix D**.

4.3.4. Choosing appropriate digital tools

It is tempting, and all too common, to group all digital tools as one and judge them accordingly. **Table 14** below shows some of the digital tools available and how their uses clearly vary depending on the context.

TABLE 14. DIGITAL TOOLS COMPARED.

Advantages	Disadvantages	Examples
SMS /USSD communications platforms		
• Continuous engagement (e.g., follow-up questions) • Two way-communication (open channel for feedback) • Low cost, easy to implement • Immediate access to data • SMS widespread in low-resource areas • Reach groups of participants through bulk messaging • Real-time data across different platforms (computer, smartphones) • Cloud-based storage and analysis • Flexible for small and large-scale communications • Automatically creates databases with telephone numbers of participants (e.g., for follow-up or linking to other data) • Useful for determining who may be excluded from your project (e.g., if the majority of the messages coming from urban area identified by phone codes)	• Only short surveys (limited to a 160 characters so questions need to be specific and easily answered, e.g. 'did the local government provide clean water in your area?') • SMS is not suitable for low-literacy populations • Participants need to have network signal and electricity to charge phones • Privacy issues with mobile phone numbers • Sample biased (limited to mobile phone and SMS users) • High non-response (particularly on non-personalized polls) • Usually requires shared shortcodes or virtual numbers (a dedicated number is needed) • Available only in some countries • Data management requires technical expertise (e.g. large text datasets) • Time and energy consuming to keep track of income messages and answer messages (even in small scale projects)	CommConnect EchoMobile Elva Platform FrontlineSMS/Cloud Groundsource RapidSMS Telerivet Textlt Text to Change VOTO Mobile Vumi (USSD)
Interactive voice response (IVR) platforms		
• Suitable for low-literacy populations • Easy to implement • Personalized messages • May require transcriptions as speech recognition is limited to certain languages	• High non-response (some people do not feel comfortable talking to a machine) leading to sampling bias • Privacy issues with mobile phone numbers and/or voice identification • Sample biased (limited to mobile phone users)	Asterisk CommConnect Ctalk FreedomFone Voicent IVR Studio Twilio VOTO Mobile

Advantages	Disadvantages	Examples	
Digital data collection apps (smartphones, tablets, netbooks)			
• Multiple data entry (e.g., text, multiple choice, multimedia) • Automatic geospatial and time data • Easy data collection in smartphones • Real-time data collection • Automatic visualization of charts and maps • Generate reports automatically (e.g., pdf) • Allows downloads of aggregate data • Text-to-speech functionalities available • Run in several operating systems • Some allow integration with SMS platforms	• Upfront equipment costs (good investment in the long term for organizations that regularly use data) • Training time (most are user-friendly) • Sample biased (limited to smartphone users)	Commcare EpiCollect Fieldata iFormBuilder Kobo Toolbox Magpi Nokia Data Gathering	Open Data Kit OpenXData Pendragon Forms PoiMapper RIWI TaroWorkz ViewWorld
Online surveys			
• Relatively cheap • Easy to design • Easy to implement at short notice • Fast data collection • Immediate access to data • Allows standardized and free text answers • Extremely useful when non-representative surveys are required (e.g., to gather information on participants, check stakeholder perceptions of program goals)	• High level of non-response • Easy for respondents to drop out • Non-coverage of groups with no internet access • Sample obviously biased (limited to literate, internet users)	FluidSurveys FormSite Google Forms KeySurvey LimeSurvey PollDaddy QuestionPro Qualtrics	SnapSurveys StatPac SurveyGizmo SurveyGold SurveyPro Survey Monkey Zoomerang

Advantages	Disadvantages	Examples
Social Media		
• Useful to share and gather feedback from a large audience • Leverage social networks with trust ties • Provides real time analysis of qualitative data (analytical dashboards available for users) • Easy to track topics (based on popularity and hashtags) • Useful to analyze feelings and thoughts of the target population • Provides a variety of information about the users (location, user history, gender, age, IP address)	• The platforms are public but with proprietary restrictions • Pose ethical issues (e.g., informed consent of target groups) • Users' opinions affected by self-presentation bias	Blogs Facebook Mxit (Nigeria and South Africa) Twitter YouTube WeAgent (Russia) WeChat (China)
Geospatial data tools		
• Useful for tracking information, analyzing data and presenting updates in real-time • Inexpensive • User-friendly interfaces • Allows exploration through web in static or interactive format • Allow to identify areas underserved by the program and uneven distribution of resources • Allows to link survey and qualitative data with geographical information • Shareable form of information	• Biased access (information submitted through smartphones or tablets) • Participants' location is identifiable making them vulnerable (e.g., to government, organizations, other individuals) • Map data may not be cross-comparable across different tools (e.g., misspellings of city names)	Crowdmap First Mile Geo Open Street Map Poimapper Quantum GIS Resource Map Ushahidi

As can be seen from this table, bias is an important issue to consider when using digital tools. If samples are employed in a quantitative approach, probabilistic techniques that rely on random selection are more rigorous (Groves, 2009). They produce representative samples whose results can be generalized directly to the original population from which the sample was drawn, e.g., identifying all users from the system data (sampling frame), and randomly selecting a group of these to be contacted. As such, random sampling requires more resources to invite selected participants and technical expertise (e.g. multi-stage sampling). It also runs the risk, however, of introducing bias through high non-response rates, unless greater effort is made to reach non-respondents, but this can be resource-intensive and beyond many budgets.

Getting better results from digital evaluation and engagement tools

SMS-based surveys should be designed carefully to ensure they inspire the respondent's interest and keeps them engaged through topics they care about. It is difficult to logistically make SMS free for users so at times airtime can be given to compensate for the cost of participation, or be sure to use toll free numbers. There are also concerns about data reduction to simple yes/no answers losing nuance, so they can be supplemented with other methods. If possible put the most important questions first to deal with drop-off rates. Where possible run tests before hand to evaluate best wording of the content and other factors like time of day.

IVR (Interactive Voice Response) is usually free to engage and provide answers through button presses to give their opinion. Using voice enables interaction with those speaking minority languages or lacking literacy skills. To maximize participation it is important to use local languages, make interactions short (e.g., 5 minutes), keep questions clear and simple and ensure the topic is of interest to the user.

Participants can be motivated when they get feedback through the same IVR channel where they gave input. After engaging with a group a follow up message can be sent to summarize the findings or share the impact of their input. This is especially important for regular IVR interactions such as quarterly citizen priority surveys.

It is important to test IVR surveys to ensure they work. Factors to test include gender of the voice, time of day of message, use of incentives, wording and order of questions and phrasing of introduction.

Using hybrid tools: Radio has massive reach but is often one way with no opportunity for citizen input. Through beep to vote, SMS and IVR surveys radio can be used to encourage participation. Farm Radio International often uses mobile technologies to run surveys during radio shows, sometimes getting thousands of calls ins during a single show. This creates interactive radio content and allows for the conversation to evolve based on listener input. It also provides a great channel for promoting citizen engagement and can be targeted at specific groups by show (e.g., on a women's health program ask about women's priorities for local government investment).

Running tests before launch: It can be hard to know what approach for soliciting input will work best. When you're unsure try an A/B test where you randomly split a piece of your target audience in half and use a different approach for each. For example offer a financial incentive to participate in giving input to one group and no incentive to the other. After running your test with a small sample, choose the better method to continue with. Advanced use of this will successively iterate on a number of dimensions of your engagement (e.g., incentive, tone, length, structure) to truly optimize.

Create a forum for on-demand citizen input: It is very simple to set up and promote a hotline where citizens can call free of charge and provide input on key policy or programs questions. This structure puts the power in the hands of the citizen for when and how they provide input. The hotline can be open ended where citizens provide broad input or it can be focused with specific questions to be answered that will inform tangible decisions. The phone-based input can be instantly shared online so there is transparency on what kind of feedback is coming in. An example could be a vote on priorities for infrastructure investment where citizens call a hotline to express their opinions and a website tallies and displays the votes in real time. For more tips, see VOTO Mobile's blog http://www.votomobile.org/blog

Levi Goertz
Chief Operating Officer, VOTO Mobile
www.votomobile.org/blog

Key lessons learned from the field evaluations

- Planning and Implementing -

Don't under-estimate the importance of interviewing and testing early findings with a wide range of stakeholders

From Brazil: "I would suggest embedding your evaluation with those on the ground conducting the Digital Citizen Engagement as early as possible (if possible build in time for a pilot study, especially if you're doing it at a large-scale). For us, the qualitative interviewes were not a core data source, but proved to be an invaluable source of contextual information which allowed a better understanding of the quantitative data we colelcted, and also helped develop useful relationships we could pursue when questions about unexpected outliers arose – we saw a huge variation in the percentage of citizens voting online in different municipalities in the state, which totally confounded us, until one of the interviewees explained that certain regions have been actively pushing their citizens online and phasing out the offline vote, while other regions are not. Without this insight we would have risked mis-interpreting the results.." (Matt Haikin, Aptivate)

Be prepared to think and act fast when things don't go according to plan

From Cameroon: "There was a lot of data entry needed - data are being collected on paper when people are going to participatory budget meetings – in 2014, we were looking at 4 communes, 15 meetings per month each, 7 months, average of 30 participants per meeting = 12,600 participant detail records - all being collected on paper. We hadn't known the data was all paper-based, so we had to very quickly identify additional local resource to do some rapid data-entry and quality checking before we could even begin to analyze anything." (Martin Belcher, Aptivate)

Moving on from Planning and Implementation to Analysis?

- Have decisions been made on how any primary data will be collected, from whom, using what approaches, and what digital and/or non-digital methods (including considering different perspectives/triangulation of data sources)?
- Do plans reflect the reality of the resources (time, budget, human resources) available for the evaluation?
- Have plans been checked with Section 4.3.2, with other evaluations or case studies in order to learn from their insights and avoid making similar mistakes?

Selected readings and online resources on planning and implementating an evaluation

12 Tips for using ICTs in Monitoring and Evaluation – *http://lindaraftree.com/2012/08/09/tips-on-using-icts-for-social-monitoring-and-accountability/*

Affordable, simple tools to collect data, communicate with clients and measure impact – *http://impacttrackertech.kopernik.info/*

A pragmatic guide to monitoring and evaluating research communications using digital tools – *http://onthinktanks.org/2012/01/06/monitoring-evaluating-research-communications-digital-tools/*

Handbook for participatory action research, planning and evaluation – *http://www.participatoryactionresearch.net/sites/default/files/sites/all/files/manager/Toolkit_En_March7_2013-S.pdf*

ICT for Data Collection and Monitoring and Evaluation – *http://documents.worldbank.org/curated/en/2013/12/18658539/ict-data-collection-monitoring-evaluation-opportunities-guidance-mobile-applications-forest-agricultural-sectors*

ICTs for Monitoring and Evaluation of Peacebuilding Programmes – *https://www.sfcg.org/wp-content/uploads/2014/05/CCVRI-SSP-_ICT-and-ME-_Final.pdf*

Mobile Data Collection in Africa – *http://webfoundation.org/projects/research-mobile-data-collection-opportunities-in-sub-saharan-africa/*

Mobile-based Technology for Monitoring and Evaluation – *http://www.theclearinitiative.org/mobile-basedtechnology.html*

Monitoring and Evaluation in a Tech-Enabled World – *http://www.rockefellerfoundation.org/blog/emerging-opportunities-monitoring*

Participatory Video for Monitoring and Evaluation – *http://www.insightshare.org/sites/insightshare.org/files/file/Video%20Girls%20For%20Change%20-%20Final%20Project%20Report.pdf*

Research Methods Knowledge Base: Sampling – *http://www.socialresearchmethods.net/kb/sampling.php*

Responsible Data Forum: Resources – *https://responsibledata.io/category/resources*

Responsible Development Data: Practitioner's Guide – *https://github.com/tanialee15/Responsible-Development-Data*

Sample size calculator for surveys – *http://www.surveysystem.com/sscalc.htm*

Scoping

Designing

Planning & Implementing

Analyzing

- Analyzing existing data
- Analyzing new data
- In-depth analysis

Scanning, Reflecting & Learning

This stage discusses how the DCE data can be analyzed and provides pointers for quantitative, qualitative and mixed methods of analysis. Challenges such as ensuring rigorous data and understanding bias are discussed, and suggestions offered as to how these can be addressed. It is also recognized that after initial analysis, there may be a need to iterate the process and re-visit the design or collect further data.

4.4.1. Working with data analysts

By the time the evaluation reaches this stage, the data needed to answer the evaluation questions should have either been identified from amongst existing data, or will have been collected using one or more of the methods described in section 4.2.5.

More than any other aspects of the evaluation process, analysis of this data requires specific expertise and will often be carried out by specialists and the inclusion of a data analyst/data scientist as part of the DCE evaluation team is recommended.

This stage of the guide is necessarily more technical than the previous stages, bus is not intended as a comprehensive introduction to data analysis (see ***Further Reading*** at the end for useful links on analysing data). It is primarily to serve as guidance to evaluators and commissioners around what is involved, to help identify the right questions to ask to experts and other collaborators, understand the pros and cons of different options and, importantly, to ensure that the factors that contribute to effective analysis of data on DCE are adequately considered and incorporated during the Design, Planning and Implementation stages.

4.4.2. Analysing existing DCE system data

In many cases DCE data may already be available and may be considered as providing suitable evidence for addressing the evaluation questions, at least after an initial scan. However, a more detailed assessment of the data is needed in order to accurately assess its value and relevance, weighted by cost considerations. For example, if the existing data needs cleaning or if there is a high cost attached to accessing it, does using existing data still compare favourably with collecting new data? Trustworthiness of the data is also important, considering how the data was collected and for what purpose, and whether respondents' privacy has been respected and permission/consent obtained from the respondents for their data to be utilised. When assessing existing DCE data, the following issues may also need consideration:

'Good enough' versus 'top quality' as the capacity required from people and systems to produce scientifically credible data may not be present in every context, there is a judgment to be made about whether to use what already exists or collect new data in the hope of improving quality. However, the challenge can be to make sure that the data really is 'good enough'.

- **Specialist tools and skills** may be needed to investigate the large data sets that DCE can result in. Simple things like call logs can be many thousands of records and office-based software may not be appropriate. Even simple tasks like cleaning data, querying, importing/exporting in appropriate formats can prove troublesome. This means that a data analyst may need to be factored into costs.

It is also important to consider the stage at which the analyst is brought in—it is better to involve them in data and measurement related issues as early as possible as missing out an important covariate at the Design stage, for example, may well prevent quality analysis later on?

- **Who the data was collected from and whether more people need to be targeted** is important to understand during the assessment of the secondary data. To understand the effects that a particular DCE project has had, data may need to be collected from the target group and from those outside the target group (a control group).
- **How quickly the data may age**. Consider this perspective on real-time data.

Collecting and working with real-time data

When working with real-time data, the key considerations are primarily human related, as the reason for collecting real-time data is often to inform quick decision making. For example, if you are creating an early-warning system, it will be critical that there are systems in place for verifying data and reports in a timely manner, as well as linking the data to a quick and appropriate response mechanism. This is less about analyzing a large data set and more about responding to each individual data point.

One way to verify data is to use a bounded-crowd method, where you only accept SMS reports from key trusted individuals. Community members will contact these individuals, who will then vet the information and can submit clear and concise data on their behalf. Over time, one could analyze all of the data points for trends. For instance, if this was an early warning system for conflict, you may want to see if there are certain triggers for violence (religious holidays, food shortages, etc.). To do this, you may need to supplement the SMS data with more in-depth research (i.e.- there was a spike in reports of violence in May, what else was going on in the community at this time?). This can then inform future programming activities and policy decisions.

Real-time data use and efficacy depends on the capacity of institutions and decision-makers to make real-time decisions with it, and to respond in turn. Without the tools, money, or systems to respond, there is a significant risk of undermining trust and contributing to 'development-fatigue' rather than enabling more dynamic and sustained processes.

Finally, the ethics and risks of using SMS, and any tool, are important to consider. SMS is not a secure medium and is often paired with mapping software. Institutions should think through the implications of collecting sensitive or geo-located information that may put individuals or communities at risk.

Valerie Oliphant
Projects Manager, Social Impact Lab (makers of FrontlineSMS)
www.simlab.org

4.4.3. Initial analysis of new data

The first step of the data analysis consists of inspecting the dataset in order to check if the data have been collected accurately. The initial analysis is necessarily linked to the evaluation questions (and, in turn, to the five lenses).

If conducting **quantitative** analysis, the initial checking of data consists of visualizing the data using pivot tables that display counts, percentages, totals and means for all variables. The variables of interest can also be analyzed using cross-tabulations, for example, by gender and age groups.

This procedure can provide an initial check for data validity, and it is particularly useful for detecting values outside the usual range and biases in the sample. It is important, therefore, to ascertain whether the unexpected findings or values are valid or whether they reflect issues related to the sampling procedure or data collection.

If conducting **qualitative** analysis, a first step is to listen/read through transcripts of interviews or watch any videos collected. Common themes can then be plotted. If a group of stakeholders has been interviewed (e.g., service providers, users, non-users) it may be interesting, at this stage, to compare interviewees perspectives. For example, what different interviewees focus on or omit. The decision on whether or not to use qualitative software for deeper analysis can also be made at this stage (***Tables 17-19*** contain more information on digital tools to help with data analysis).

If the initial analysis reveals that additional data is necessary, either due to remaining gaps or to poor quality data, there may be a need to return to the Design stage and collect more data by, e.g., correcting the sample or to double checking answers with the respondents. While time and budget restrictions mean that only a few evaluations have this level of flexibility, it is important to build this possibility into the design from the beginning.

4.4.4. In-depth analysis of quantitative data

The in-depth analysis of quantitative data involves several steps including assessment of data quality, data cleaning, and statistical analysis and interpretation of results. This process is not linear as the data visualization or statistical analysis may reveal that additional coding or cleaning is needed.

Before starting the statistical analysis, the dataset should pass through a careful **data quality assessment**. **Table 15** below provides a list of possible errors/bias that may be detected in the data, alongside with procedures for quality assessment and possible solutions to deal with these errors. It is worth noting that no data is free of error, but a good design and implementation minimize their occurrence.

Well-informed decisions about the definition of the population, the sampling technique, training of enumerators (if any), the choice of technology for gathering data, and ways to increase non-response will produce high-quality data. There should be a good trade-off between scientific rigor and cost considerations, often incompatible in field evaluations. More sophisticated analyses cannot compensate for poor-quality data, although in some cases it is possible, for example, to correct for sampling bias by applying weights to the data so the sample would be closer to the population in terms of key socio-demographical characteristics.

The assessment of data quality is necessarily followed by the stage of **data cleaning** in order to deal with repeated cases, missing values, outliers, data outside the range or inconsistent values. Digital methods for data collection are more prone to certain types of error such as sampling error and nonsensical data and may require extra resources for data cleaning and weighting, compared with non-digital methods. Again **Table 15** shows in details these procedures.

TABLE 15. STRATEGIES TO EVALUATE DATA QUALITY.

Description	Detection/Assessment	Solutions
Type of error/bias | Sampling bias		
Difference between the target population and the sample. Mobile phone and internet surveys are more prone to sample bias than face to face and random digital dialling.	Compare sample statistics with official population statistics by using appropriate tests to detect statistically significant differences between the sample and the population (e.g. z-test).	Weighting data based on official statistics or collecting additional data for underrepresented groups. Indicating margins of error (e.g. ±3%) to generalize to population when presenting results.
Type of error/bias | Non-response bias		
Percentage of people who were invited but refused to participate in the evaluation. Digital vs non-digital methods are more prone to non-response bias (it's more difficult to refuse in a face-to-face situation)	It is problematic if distributed unevenly across socio-demographical groups (e.g., higher non-response among women) or if non-response presents a particular pattern (e.g., higher non-response participants who are more satisfied with a service). Compare respondents with non-respondents in socio-demographics and other variables of interest. If there is evidence that they are similar, no further action is required.	Using incentives to motivate people to participate in the evaluation. Consider using multiple methods for data collection to reach non-respondents (e.g., face-to-face and mobile phone). Report the differences between respondents and non-respondents as an indication of the quality of the recruitment of participants. Weighting data to adjust for non-response bias.

Description	Detection/Assessment	Solutions
Type of error/bias | Missing values		
Answers not provided from respondents who agree to participate in the evaluation. Digital vs non-digital methods are more affected by missing values as the participants can easily drop out.	It is a problem when missing values occur only for particular questions, signalling poor question design (sensitive/ offensive, difficult to understand or to answer), or when missing values are associated with certain demographic groups. Detected through simple frequency tables per question and averages across all questions and across individuals	In some cases (e.g., if the statistical technique does not accept missing data), cases with missing values need to be excluded from the analysis; alternatively, missing values need to be replaced by predicted values using statistical techniques.
Type of error/bias | Outliers, data outside range		
Atypical values that fall beyond the distribution of other values. May be an 'interesting case' or may be a mistake when collecting or entering data. Mobile phone surveys are more prone to outliers and out of range values as uses free-text.	Outliers and atypical values can be detected using simple frequency tables or boxplots.	Ideally, double check with the respondent. If it is not possible for ethical or practical reasons, set outliers to missing values and run the analysis with and without outliers to find if the findings are robust to the inclusion of outliers
Type of error/bias | Nonsensical data		
Inconsistencies that suggest issues in the interpretation of the question, data collection or data entering.	These values are not outliers, as they fall within the distribution of other values, but are unfeasible when compared to other answers.	Set the inconsistent data to missing if it there is no doubt that the answer is wrong, or create rules to replace inconsistent data.

The **statistical analysis** starts with data exploration through frequencies, tables, charts and graphs in the first stage. The second stage focus on descriptive statistics and the third stage on statistical hypothesis testing if a sample is employed (not needed when analyzing the population). **Table 16** below shows the types of analytical tools that can be employed for each stage and some examples from Uganda and Brazil's analysis.

The choice of analytical technique depends on evaluation questions, the types and quality of data, the expertise of the data analyst, and the timeline and resources of the evaluation. The choice of the analytical technique does not depend on the meth-

od, as the same techniques can be used to analyze data from a RCT or a survey, for example. The confidence in the results, the degree they can be generalized to other individuals or contexts, or whether an association implies causation depend on the methodological decisions rather than the analysis itself.

As a dynamic rather than an extractive process, analyzing quantitative data implies not only technical skills but also theoretical and contextual knowledge in order to detect inconsistencies, explore patterns, identify spurious correlations and follow clues. Along these lines, analyzing data is not an individual exercise but a team endeavour through which findings are discussed.

TABLE 16. TOOLS FOR QUANTITATIVE DATA ANALYSIS AND EXAMPLES FROM FIELD STUDIES

Description	Analytical tools	Examples *(U-Report, participatory budgeting Brazil)*
Objective of the analysis | Exploration		
Visualizing the data through percentages, tables, charts, and graphs	Frequency tables, cross-tabulations, histograms, boxplots, bar or line graphs, heat maps	Graph plotting in two separate lines the evolution of online and offline voting from 2005 to 2014 in Rio Grande do Sul, Brazil
Objective of the analysis | Description		
Summarizing variables using descriptive statistics and relations between variables using correlations/associations	Percentages, means, standard deviations, correlations, effect sizes	66.3% of U-Reporters are very occasional contributors, answering up to 40% of questions; 53.4% of U-Reporters surveyed think that U-Report has led to some or many changes in their districts (53.4%)
Objective of the analysis | Statistical hypothesis testing		
Generalizing results from the sample to the population using confidence intervals; Testing the significance of differences or associations based on p-values (smaller p values are better as they mean that the results are less likely to be encountered by chance); when testing multiples hypotheses with the same dataset it is important to use a correction for false discoveries such as Bonferroni method	Confidence intervals, chi-square, t-test, ANOVA, regressions, structural equation modelling	Between 11.4 and 17.2% of Ugandan citizens have heard of U-Report (for 95% confidence level) U-Reporters who work in government or civil society are more likely to raise a complaint to a local council leader in the last 12 months ($X2 = 16.8476$, $df = 4$, $p<.01$)

When analyzing data and interpreting results, some common errors can compromise the validity of the evaluation. The following golden rules and associated common errors in DCE evaluations can be used as a guide to assess the quality of the analysis:

- **Use unique entries**. Error: considering all the entries in system data and not unique IDs (e.g., mobile phones' numbers) to determine the characteristics of users of a digital platform
- **Ascertain sample bias**. Error: failing to compare the characteristics of the users on key variables with official population figures (e.g., household census or system data) whenever available (needed to ascertain degree of bias)
- **Use confidence intervals for statistics**. Error: omitting confidence intervals or margins of error of a statistic when presenting results from a sample (these are needed to generalize results to a population)
- **Correlation is not causation**. Error: interpreting correlation as causation when using non-experimental methods (that do not control for other possible causes or establish the direction or causality)
- **Include confounders in the analysis**. Error: ignoring important confounders and interpreting spurious associations as real ones, e.g., when a third variable *Z* is causing the association between *X* and *Y*. The statistical analysis should be controlling for those (e.g., multiple regression)
- **Look at patterns for separate groups**. Error: not considering whether the association between two variables *X* and *Y* holds for different groups (men/women, more active/less active contributors) and rely on associations derived from pooled samples (mixing different groups) that result from mere artefacts (e.g., Simpson Paradox)
- **Report effect sizes**. Error: Over-reliance on inferential tests and p-values without taking effect sizes into account (e.g., when comparing users and non-users groups to draw conclusions about impact).

Table 17 presents a selection of the most popular data analysis software available in the market, along with their strengths and weakness. The choice of particular software is related to the objectives of the analysis (e.g., exploration) and the expertise of the data analyst. Some tools are satisfactory for initial analysis (e.g., Google Charts, SurveyMonkey) but may not be adequate for more in-depth analysis or customized visualizations.

TABLE 17. SOFTWARE/PLATFORMS FOR QUANTITATIVE DATA ANALYSIS.

Strengths	Weaknesses	Examples
Survey platforms(exploration)		
Easy to use; immediate visualizations and pivot charts for initial inspection of data	Do not allow data manipulation; do not perform statistical analysis: Limited to web and mobile phone surveys	Survey Monkey (web); Qualtrics (web); Google forms (web); Echo mobile (SMS)
Data visualization platforms/software (exploration)		
Easy to use; Open Access; high-quality graphics; interactive visualizations; customized visualizations that can be exported as images or embedded on webpages	Data transformations are not allowed; do not perform statistical analysis; data need to be uploaded in their servers and may become available to the public	Tableau, IBM Many eyes, Google Charts, Vizualize Free, RapidMiner ('big data')
Data processing software (exploration, description)		
Easy manipulation of data; performs basic analysis and visualizations; some are open access or free	Not suitable for large datasets (>10,000 rows); do not process free text SMS data; some analysis require using/writing macros; data importation limited to only a few formats (e.g., csv)	Excel, Google Drive sheets, Calc (Libre office), Gnumeric, i-Work/Numbers, EpiData
Statistical computing software (exploration, description and hypothesis testing)		
Wide range of statistical procedures and graphs/ charts; reads multiple data formats; menu and command interfaces available to suit the level of the user; Some allow processing of free text SMS data (e.g., R, SPSS)	Steep learning curve; requires sound statistical knowledge use it efficiently;sSome require expensive licenses (e.g., SPSS, SAS)	SPSS, STATA, SAS, Minitab, R ,MPlus, Python StatsModels and Pandas
Advanced statistical analyzes		
Appropriate for particular techniques; full customization of the analysis; technique-specific visualizations	Requires in-depth statistical knowledge; steep learning curve; expensive licenses	NetMiner and UNICET (social network analysis), Amos and EQS (structural equation modelling), HLM and MlwiN (multilevel modelling, ArcGIS and QGIS (spatial analysis)

4.4.5. In-depth analysis of DCE qualitative data

One of the often-cited challenges of qualitative data is that it is considered as non-generalizable (as compared to quantitative data) because it can be very context specific (see also the description of different qualitative methods in section 4.2.5.). However, this characteristic is also part of qualitative data's strength.

It is important to distinguish here between structured, semi-structured and unstructured questions, providing varying parameters of how the interviewee is limited in expressing their perspective. In the first two, there is a framework of questions while in the latter the emphasis is on storytelling that allows for narrative analysis.

Narrative analysis refrains from treating interviews as data, and instead sees the storytelling itself as revealing in how interviewees see themselves and others—focusing on, what is left out, emphasis and tone. In development, storytelling is increasingly seen as a type of participatory evaluation and so would be a relevant method for analyzing DCE. The ways in which digital tools provide powerful opportunities for self-expression are the subject of recent pilot studies (Reitmaier et al., 2011; Frohlich et al. 2009). Extending qualitative analysis further, another valuable approach is ethnography that, as we have seen, involves observing individuals in their real-world context.

In all qualitative (and for that matter, quantitative) analysis it is also important to recognize personal and cultural bias. All these attributes of qualitative analysis underscore the importance of why, wherever possible, mixed methods and triangulation from diverse data sources are advisable (see for example Alvesson and Sköldberg, 2009; Raftree, 2014). The analysis of qualitative data involves:

- **Developing a coding scheme for organizing the data** on the basis of the key themes. For example, if exploring interviewee's motivations annotate all mentions of motivations with a 'motivation' tag
- **Coding data and refining your coding scheme** as during the process of coding new themes and subthemes may emerge. One might discover than when people talk about their motivation they may talk in fact about different things (altruistic, non-altruistic). Sometimes it is very useful to double check the validity of the coding scheme by asking another person to code the data
- **Exploring links between different codes** and forming hypotheses of how different codes/themes relate to each other: is it possible that certain types of motivations are strongly associated with certain types of expectations?
- **Interpreting associations and checking representativeness** by counting instances of particular themes, comparing them and associating them to the pro-

files of interviewees (do women talk about motivations differently than men? Is the age of interviewees a factor?)

- **Exploring alternative explanations and looking for negative cases**.

The approach described here is more suited to thematic and content analysis so it might need to be adjusted for other types of analysis (e.g. narrative/discourse analysis). The trustworthiness of qualitative data depends on the availability of quality data, (i.e., rich data, appropriate and diversified) and on the rigorousness of the data gathering and analysis and reporting. The criteria for trustworthiness of qualitative analysis (Lincoln and Duba, 1995) are:

- **Credibility**: does the analysis provide a true picture of the phenomenon being studied? Are the participants accurately clearly identified and described?
- **Dependability**: do the findings hold at different times and under different conditions?
- **Conformability**: what is the level of congruence of two or more independent researchers on the data's accuracy, relevance and meaning?
- **Transferability**: can the research be transferred to other settings or groups?
- **Authenticity**: to what extent does analysis fairly and faithfully show a range of realities?

Table 18 presents software for qualitative for computer-assisted qualitative data analysis (CAQDA). Most of these packages are very flexible and can be used across a range of qualitative methods and analytical strategies. Most of the software listed present functions for data and document management, developing code hierarchies, annotation of text with codes, writing memos, exploring data and text retrieval and visual representations of data codes and annotations.

TABLE 18. SOFTWARE FOR COMPUTER ASSISTED QUALITATIVE ANALYSIS

Analysis	Description	Types of data	Software
Content and thematic analysis	Categorization of textual data for classification, summarization and tabulation	Structured and unstructured text, multimedia data	HyperResearch, NVivo, Atlas, Alceste, QDA Miner, QSR N6, Max QDA, Dedoose (multimedia), Inforapid (html and rtf), Studiocode (video), Transana (video and audio), Overview, Weft QDA, Voyant Tools, NodeXL/Excel add-in, Ethnograph
Discourse analysis	Based on how language is used in spoken interaction and written texts	Unstructured text, audio data	HyperResearch, Tams Analyzer,
Conversation analysis	Analysis the sequential organization and details of conversation	Unstructured text, audio data	HyperResearch, NVivo, Atlas, Qualrus
Narrative analysis	Focused on how respondents impose order on the flow of experience in their lives and make sense of events and actions in which they have participated	Unstructured text, audio data	HyperResearch, NVivo, Atlas, Alceste, QDA miner, Max QDA, Qualrus
Qualitative comparative analysis	Based on several cases, identify a combination of conditions that need to be present to produce a certain outcome (Ragin, 1987)	Classification of cases	Fs/QCA, Tosmana, Kirq

Although aided by the software, it is a human interpreter who makes sense of data through manual coding and retrieving. Alternatively, analysis can be done simply in Microsoft Word or Excel, or open source alternatives.

A simple approach to data analysis

At Social Impact Lab, we focus on using inclusive technologies, which we define as accessible, easy-to-use, interoperable, and sustainable, in order to reach last-mile communities. This is why one of our flagship projects, now their own company, FrontlineSMS, was built to rely on SMS, mobile phones, and a modem- tools that many organizations and people already owned or had access to, and had the technical knowledge to use.

In keeping with a focus on using inclusive and low-end tech, we feel it is important to use simple data analysis and tools that people already have and know how to use. In reality, people primarily use FrontlineSMS because it's free, so it's fairly unlikely they have the resources to hire someone else to analyze their data for them. They also use FrontlineSMS because it is simple- it requires very little training and they don't have to spend a lot of time learning how to operate a new system or software.

This is why Microsoft Excel works well as a database for analyzing data from FrontlineSMS or FrontlineCloud, and why you can easily export your data as a CSV file:

- *Just about everyone has a basic command of Excel.*
- *Excel is available on most computers.*
- *Excel creates graphs and tables, and easily compiles data.*
- *Advanced functions like PivotTables allow you to dive deeper into your analysis, make nice visualizations, and are relatively easy to learn/self-teach.*
- *Excel files are easy to save and share with others.*
- *There are a lot of free resources for Excel, including how-to videos, written guides, and even free online courses.*

Valerie Oliphant
Projects Manager, Social Impact Lab

4.4.6. Large-scale text mining

It is important to distinguish between 'designed' data collected with a pre-specified purpose and with potential users in mind (Groves, 2011) and organic data, that is generated automatically through users' interactions with digital platforms, for example in social media.

Insights from organic data are often quality metrics that encompass opinions, sentiments, satisfaction ratings, conversations, number of shares, comments, re-tweets, replies, ratings, as well as the quality of engagement over time (Paine, 2011).

More innovative analytical strategies, including those that use organic data, do not fit easily into the distinction between qualitative and quantitative data. Organic data, usually larger than designed data is streaming data (rapid, continuous feeding), and it can be simultaneously quantitative (e.g., number of complaints) or qualitative (e.g., content of the complaints). Since organic data can be used for answering questions that were not considered when the platform was designed, advanced tools for data digging to find out relevant information (if any) are needed. This tools that extract information through sophisticated algorithms belong to a category of techniques called statistical or textual data mining.

Textual data mining techniques are used to extract information from large textual organic datasets in an automatic or semi-automatic way, i.e., with limited human input. For example, sentiment analysis automatically codes words or expressions of a large corpus of data based on their emotional value (positive, neutral or negative) derived from sentiment dictionaries. More advanced techniques (e.g., deep sentiment analysis) consider the adequacy to the context where the word/expression appears).

Topic discovery is another technique used for data mining that consists of classifying automatically documents or parts of documents based on words or expressions that appear on it. For example U-Report in Uganda used topic modelling to classify the unsolicited messages (around 10,000 texts a week) into 10 topics (education, emergency, employment, energy, family, health, orphan, social, violence, water) with an index of relevance of each message for each topic, so they could be routed to the responsible agency (Melville et al., 2013). The big limitation of these analyses is that they require language resources (dictionaries, lexica) that are available only for a limited number of languages.

The most suitable analytical strategy will be influenced by factors such as the availability and character of organic data (e.g., language), as well as by more traditional factors such as time and budget constraints to acquire software or to hire a data analyst. ***Table 19*** shows the two most common strategies for data mining along with examples of software.

TABLE 19. SOFTWARE FOR DATA MINING/TEXTUAL ANALYTICS

Analysis	Description	Types of data	Software
Sentiment analysis	Classification of the polarity of the attitudes, emotions and opinions in documents (positive, negative, neutral) using natural language processing	Unstructured text (big data)	NLP Stanford demos, Gate, ForSight platform (social media), Exact Target (social media) Mozdeh (Twitter)
Topic discovery	Machine learning techniques to discover latent meaning in documents	Unstructured text (big data)	DiscoverText, SAS Text Miner, Indico, RDataMiner

4.4.7. Handling and interpreting the results

There are some key factors that are crucial to bear in mind when handling the kind of sensitive data that typically emerges from DCE evaluations and these should be discussed with the people doing the data analysis at the very start of their work to ensure no potential problems as things develop, some of these are highlighted below:

Data handling and privacy: It is important during data analysis to bear in mind how data is handled: how is user's anonymity / privacy protected (by whom, and from whom)? How is confidentiality handled? Is usage data accessible? Who has access to project reports? All these issues are particularly important when there is currently such a great focus on open data–in the emphasis on open data it is important that the focus is not lost on confidentiality and privacy.

Critical analysis of results and trends: It is always advisable to question any trends or results critically. It may be necessary to bring in external subject-matter experts or local stakeholders who understand the reality on the ground to help make sense of what the data appears to be telling the analysts

Watch for misleading results: For example "false negatives" are common in SMS reporting where, for example, when crowdsourced conflict data is being collected, there is an absence in a particular area, simply because of lack of SMS use/mobile phone ownership (Fruchterman, 2011), and not because there is a lack of conflict. Similarly, with sentiment analysis or opinion mining, there are cautions about approaching this uncritically (UN Global Pulse, 2012). An expert data analyst who lacks detailed sector expertise could easily miss these kinds of misleading results, it is vital that the evaluator or another stakeholder with this kind of knowledge is involved in discussions to avoid this happening.

Key lessons learned from the field evaluations

- Analysis -

Even though many digital tools provide superficial analysis automatically and instantly, don't under-estimate the importance of using an experienced data analyst

From Brazil: "It was easy to do a first superficial analysis–it took about two hours to collect the top-level results from all the three surveys in Survey Monkey into different spreadsheets and even from that you could immediately get some level of analysis… I would recommend working with a data analyst if possible, which we did. For example, you may make naive assumptions, and then they might point out that you can't always trust the data at face value (Matt Haikin, Aptivate)

Mixed methods are likely to give a much better insight into data than simply one set of methods

From Uganda: "We used extensive mixed methods, and the qualitative approach of interviewing the 17 U-Reporters threw up some really interesting insights. The survey results may assume one response per person, but actually in interviews we came up with diverse findings–for example, individuals consulting colleagues before they responded, to seek what validation of their response (so is this an individual opinion, or peer pressure/groupthink?) Or one woman who said she had three phones and gave different responses on each phone. Although these are unique instances, they should be kept in mind so we constructively critique the quantitative analysis. In the trend for 'big data' in DCE, we should not forget the importance of people's stories–the 'small data' " (Evangelia Berdou, IDS)

Start analyzing as early as you can – it might point you towards changes you need to make to your data collection plans

From Kenya: "From initial data analysis, it became clear that there was a real preference to engage via non-digital channels (over the counter and telephone complaint submissions). This meant the number of people we could survey about their use of digital tools would be smaller unless we adjusted the sample strategy to reflect this. When surveying people we also found that more people were willing to complete the survey interviews (from all channel user categories). This meant that we could complete more surveys within the allocated time and resources. So we adjusted things as we went along to take account of rapid data analysis and digital tools monitoring." (Martin Belcher, Aptivate)

Moving on from Analysis to Sharing, Reflecting and Learning?

- Have all of the evaluation questions been answered?
- Has the data has been collected and analyzed in a sufficiently robust way to ensure its credibility (including some sort of triangulation)?
- Is there a need to collect more data or carry out further analysis using different methods?
- Is all the supporting information collated in order to give people confidence in the stated results?

Selected readings and online resources on Analysis
Research Methods Knowledge Base: Analysis – *http://www.socialresearchmethods.net/kb/analysis.php*
The "Real Book" for story evaluation methods – *https://chewychunks.files.wordpress.com/2012/05/storytelling-realbook-may-23-2012.pdf*
Who counts? The quiet revolution of participation and numbers – *http://www.ids.ac.uk/files/Wp296.pdf*
Quantitative and qualitative methods in impact evaluation and measuring results – *http://www.gsdrc.org/docs/open/EIRS4.pdf*
Evaluation Toolkit Data Analysis – *http://toolkit.pellinstitute.org/evaluation-guide/analyze/enter-organize-clean-data/*
Cookbook for R – *http://www.cookbook-r.com*
Online SPSS and STATA tutorials – *http://www.lse.ac.uk/methodology/tutorials/introduction.aspx*
StatSoft Electronic Statistic Textbook – *http://www.statsoft.com/Textbook*
Real Statistics Using Excel – *http://www.real-statistics.com*
Programme Development and Evaluation: Analyzing quantitative data – *http://learningstore.uwex.edu/assets/pdfs/G3658-6.pdf*
Participatory Research Methods: A methodological approach in motion – *http://www.qualitative-research.net/index.php/fqs/article/view/1801/3334*
Resources to help you learn and use SAS, UCLA – *http://www.ats.ucla.edu/stat/sas/*
Resources to help you learn and use SPSS, Institute for Digital Research and Education (UCLA) – *http://www.ats.ucla.edu/stat/spss/*
Resources to help you learn and use STATA Institute for Digital Research and Education (UCLA) – *http://www.ats.ucla.edu/stat/stata/*
Resources to help you learn and use R, Institute for Digital Research and Education (UCLA) – *http://www.ats.ucla.edu/stat/r/*
Subjectivity lexicon, University of Pittsburgh – *http://mpqa.cs.pitt.edu*

Scoping

Designing

Planning & Implementing

Analyzing

Scanning, Reflecting & Learning

Testing findings & Reporting
Influencing
Disseminating and opening up results

4.5

This final section focuses on testing the findings, writing up the results and analysis of a DCE evaluation, considers methods of sharing findings (including discussing opening up evaluations and their data), and reflecting and learning on the lessons from evaluations.

4.5.1. Testing findings

A common penultimate evaluation stage is the sharing of the report internally or with a smaller group before the results are made public. This could involve setting up an advisory group, a workshop with participants, and digital consultation on draft reports. It is important to include those who were initially participants on the research, to whatever extent it is possible, including making as much of an effort to translate into relevant languages (and budgeting for this).

Doing this provides a critical opportunity for respondents or participants to clarify their response or correct the analysis where they feel it was not an accurate representation (of course, this will not be possible with large-scale surveys, but addenda can be provided at the end of a report, also illustrating the dialogue involved in presenting the research).

4.5.2. The five lenses as an aid to reporting on an evaluation

The first step in writing up a report is to decide what results the evaluators wish to focus on and communicate. Here guidance comes from the purpose of the evaluation itself, the Terms of Reference from the commissioner, and any framework, such as a Theory of Change, describing the purpose of the evaluation and pointing towards the required dissemination strategy.

Whilst the detail of writing reports will not be covered here, one addition this guide suggests is the use of the five lenses as part of the reporting format. It is not proposed that the whole report should be structured this way, but rather that there could be a short section accompanying the Executive Summary where the five lenses are used to categorize and describe findings. Using a framework of the lenses and associated areas of interest would **contribute to the sharing of evaluation findings on DCE** and make them more comparable across projects, and provide the possibility of lessons and experiences being shared and applied across the DCE.

4.5.3. Influencing decision-makers

Some larger organizations have dedicated teams to assist with constructing dissemination plans and executing these. For independent evaluators too this activity should be factored into budgets.

While 'free' social media and networking sites can be optimally used for disseminating, it is also useful to consider setting extra budget aside for writing blogs and guest blogs on other websites, and also to disseminate into relevant languages and formats, e.g., to share with research respondents. Another factor to bear in mind is that responding to feedback should be a dialogue: e.g., if comments boxes are provided at the end of blogs, then resources need to be set aside for responding to these comments.

One of the challenges of DCE is that it can cross many disciplines and have many different audiences. Lessons and findings from DCE evaluations may be of interest to technology experts, social scientists, program managers, donors, other evaluators and, of course, citizens. As Caroline Heider (2014) of the Independent Evaluation Group at the World Bank points out, it is important to identify the diversity of audience which may read a DCE evaluation: from internal to external, and from senior levels (for policy buy-in) to operational levels (essential to make an actual difference on the ground). This not only means potentially presenting the data and findings in different ways, but also tailoring the findings when making presentations, etc. It is also always important at this level not to forget to include the beneficiaries or the 'citizens' in the Digital Citizen Engagement, and the respondents when sharing the evaluation results.

4.5.4. Incorporating digital technologies for creative dissemination

Digital technology can be used to disseminate evaluation findings. Examples include data visualizations, mapping, photo stories, and videos (Beardon, 2013). The use of participatory video (where people from the target community are trained and supported to use a video camera, allowing them to film what is important for them) as a tool for data collection and analysis provides a rich source of relatively unfiltered reporting and can be used in dissemination, bringing respondents voices directly to the audience, and potentially playing a role in empowering respondents in the process (InsightShare, 2015; Lemaire and Muñiz, 2011; Milne et al., 2012).

As well as being used on its own, digital technology can creatively complement and enhance more established, non-digital means of dissemination. Digitally based products from the evaluation (videos, etc.) can, for example, be screened and a face-to-face discussion facilitated with the selected audience in order to reflect on the findings, to learn from them and consider how they might best be acted upon.

4.5.5. 'Open evaluations'

There has been much emphasis recently on 'open data' and 'open aid', as well as admitting to failure to achieve this. These principles on openness can be extended to ensuring 'open' evaluations, i.e. those which follow the definition where "*open means anyone can freely access, use, modify, and share for any purpose (subject, at most, to requirements that preserve provenance and openness)*." (http://opendefinition.org/).

Opening up evaluations can increase the quality of the evaluations and provide new insights: by making data collected open for re-use, others can analyze the same findings from a different perspective, cross-check for potential effects or even suggest new indicators. Evaluation data in open format could also be 'mashed-up' with

other available datasets by collaborators, external evaluators or interested parties, making the results more robust.

Opening up data may also help to reduce future evaluation costs: while many programs are context and content specific, general background data could be shared among evaluators and thus reducing duplicating efforts. And finally, paradoxically, many evaluations of citizen engagement programs can be closed and distant from citizens. We often get to know results of evaluation projects in pdf reports that can be quite disengaging. Making results and collection mechanisms open is participatory in itself.

4.5.6. Reflecting on the process

As a closing activity to the evaluation, an important final step is to reflect on the overall evaluation process to identify what worked, what didn't work, what has been learnt, how cost-effective it was, and arguably most importantly, what should be done differently (or at least kept in mind) in the next evaluation. It is through this reflection and learning that practice and outcomes will improve.

Sharing, Reflecting and Learning: is the evaluation completed?
Have the findings been sufficiently tested with key stakeholders to give greater credibility to the work?
Have the five lenses been used to demonstrate the breadth of the evaluation and for it to provide comparable findings with other projects?
Has there been maximum transparency of the data and design/experience to enable others to benefit?
Has a suitable strategy been designed and implemented for sharing the final report, ensuring it contributes to the overall purpose of the evaluation and conveying the findings to the right people in an appropriate way?
Have the lessons learnt from the experience of evaluating DCE been captured, and are they accessible for future evaluations?

Peer-to-peer results sharing through real-time dashboards

GlobalGiving have recently introduced real-time effectiveness dashboards for the benefit of their non-profit partner organizations. It is a sophisticated tool that takes full advantage of all that digital technology offers in terms of statistical analysis, breakdown and benchmarking. It provides their partners with a detailed picture of all their online activity - fundraising, communications, finance, learning and volunteering—that is automatically maintained, in a presentation that is intuitive and requires no time at all to grasp and understand. It is hoped that by offering their partners this insight into the effectiveness of their activities, the dashboard will actively engage them in trying to improve their efforts. More than that, it is hoped it will encourage them to take a more dynamic interest in the efficiencies of all their activities – online and on the ground.

It is a variation on the Benjamin Franklin's axiom, 'If you want something done, ask a busy person'. In this case GlobalGiving's thesis is: 'Get someone keenly engaged in one activity and they will become enthusiastically engaged in all'.

Eleanor Harrison
CEO, Global Giving UK
www.globalgiving.co.uk

Opening up your evaluation data

To make your evaluation data truly open, make it accessible in a standard and structured format so it can be easily processed; make it reliable and guarantee that the data can be accessed consistently over time; make it linked to other data and also traceable so others can check the context of the data and where it originates. In more practical terms, look at the criteria and examples to make your open data five star: http://5stardata.info

- *Think about opening up the data before you start collecting the data. There are useful open source tools available that may help doing so, for example, check out the open data kit https://opendatakit.org*
- *Make your open data ethical: you should be careful not to harm the privacy of individuals.*
- *Ensure that time, human and financial resources are taken into account from the very beginning of the project as it may take longer than anticipated particularly if there is no previous expertise in opening up evaluations.*
- *Seek help from organizations familiar with both evaluations and open data. For example, in Mexico, a platform called "Datamx" (http://datamx.io) facilitates civil society organizations to open up their data–either by helping cleaning data, providing coding and insights about the potential of sharing records.*

Dr Carla Bonina

Research Associate, Latin American Open Data Initiative (Iniciativa Latinoamericana por los Datos Abiertos – ILDA idatosabiertos.org) and Assistant Professor in Entrepreneurship and Innovation, University of Surrey

Key lessons learned from the field evaluations

- Sharing, Reflecting and Learning -

Leave sufficient time and budget for dissemination

From Brazil: "I would say don't forget to leave both time and money in your budget for testing, sharing dissemination. We had hoped to take the results of the evaluation back to the stakeholders in Brazil to explore it with them, but the data collection took longer and cost more than initially hoped so we've had to achieve this by email and without a Portuguese translated version of the results – its not ideal. " (Matt Haikin, Aptivate)

Make every effort to make your report accessible to all

From Cameroon: "Those involved in participatory budgeting itself should be able to access and understand the evaluation reports–this is potentially an issue—we were hoping to produce comprehensive reports for the key stakeholders but also find resources to produce simpler, shorter versions for general public consumption. (Martin Belcher, Aptivate)

Be aware of publishing private sector data in DCE evaluations

From Kenya: "On publishing the findings, one interesting point is that everyone involved is in principle open to publishing the report publicly, but there is also caution as it does contain potentially commercially sensitive information, since we're talking about working with a private sector company. So a broader DCE issue may be that we have to make the distinction between public interest information and commercially sensitive information." (Martin Belcher, Aptivate)

Selected readings and online resources on Sharing, Learning and Reflecting

Embracing Evaluative Thinking for Better Outcomes: Four NGO Case Studies
http://www.theclearinitiative.org/EvaluativeThinkingReport_FINAL_online.pdf

Here is the evaluation report… so now what do we do?
http://capacity4dev.ec.europa.eu/article/here-evaluation-report-so-now-what-do-we-do

Improving the use of monitoring and evaluation processes and findings
http://www.managingforimpact.org/resource/cdi-conference-report-2014-improving-use-monitoring-evaluation-processes-and-findings

Influential evaluations: Evaluations that improved performance and impacts of development programs
http://lnweb90.worldbank.org/oed/oeddoclib.nsf/DocUNIDViewForJavaSearch/67433EC6C181C22385256E7F0073BA1C/$file/influential_evaluations_ecd.pdf

Use of impact evaluation results
http://www.interaction.org/sites/default/files/Use%20of%20Impact%20Evaluation%20Results%20-%20ENGLISH.pdf

What makes evaluations influential?
http://gendereval.ning.com/profiles/blogs/what-makes-evaluations-influential?xg_source=msg_mes_network

5

Looking forward: developing better practice

5. Looking forward: developing better practice

This guide has been produced in a spirit of openness and reflection. The field is evolving, understanding of how to evaluate it is evolving, and the advice and guidance contained in this document should be evolving alongside it.

While the framework has been helped by first-hand experience and draws on a wide range of knowledge and experience in the various fields that inform DCE, it will only remain useful and relevant if the different experiences of using it can be shared and learnt from, so that future iterations of the framework remain relevant and informed by reality. Being honest about what works and does not work is both essential and extremely valuable.

Such reflections are equally valuable in individual evaluations of DCE. It is not only about closing a feedback loop between citizens and governments, but also about feeding forward and closing the loop between the end of one evaluation and the beginning of the next. By demonstrating the value of reflection and learning, by being open with and pooling our knowledge and our experiences, by being more rigorous in our evaluations, we gain ideas for future research and improved practice and we can influence and improve organizational and program structures, systems and culture.

"Evaluation is an activity. Evaluative thinking is a way of doing business. This distinction is critical. It derives from studies of evaluation use. Evaluation is more useful—and actually used—when the program and organizational culture manifests evaluative thinking." (Interaction 2013)

In this context it is intended that this guide should evolve into an open and online resource that is owned by those practicing in the field and that is being continually developed in conjunction with the wider DCE sector and those evaluating it.

If you would like to be involved – [contact or website URL here]

Toolkits & Appendices

Toolkit 1: Examples of evaluation questions relevant to DCE

Section 4.2.2 there are some ideas of types of evaluation question that might be particularly relevant to evaluating digital citizen engagement activity. These highlight how different questions help explore different aspects of the work, through focusing on the five lenses of **Objective**, **Control**, **Participation**, **Technology** and **Effect**.

The tables in the toolkit below give a more thorough comprehensive set of typical questions, examples of real questions, or ideas of helpful exploratory questions, related to each of these lenses. These questions may overlap and are intended as indicative only: it is not intended that they should be used verbatim, but they are intended to help think through the most appropriate questions and areas of explorations for a particular piece of evaluation work. They also indicate the overlap between lenses (e.g. there are several the supplementary questions under Effects that also relate to Objective)

These are grouped by Primary questions (questions which are core to the understanding of the success of the work under evaluation) and Supplementary questions (additional questions which might bring valuable information or data that can then be helpful in answering the primary questions or developing a more nuanced understanding of the work). The two categories should not be seen as watertight, and may well vary according to the scope of the evaluation.

Lens | Objective	
Example Evaluation Questions	
PRIMARY (ASSESSMENT/ANALYSIS)	**SUPPLEMENTARY (INFORMATION GATHERING)**
Was the digital engagement work appropriate for the specific/local social and political context? Were local technical, social and other pre-conditions fully understood and incorporated into the design of the program and the choice of technologies? To what extent does the actual impact meet the needs identified in the external environment? How appropriate was the monitoring in terms of scale, depth and program understanding? What were the goals of the digital engagement process? Were they achieved? What was the planned digital engagement activity? What happened? What were the gaps between the plan and the reality?	**External Environment for DCE** What does the information landscape for targeted groups look like? How do people normally stay informed and communicate? What does people's access to and usage of the internet look like? Are there significant differences along the rural-urban divide, between men and women across these dimensions? How widespread and reliable is access to electricity? What are the targeted citizen literacy levels and other basic capabilities required for participation (such as texting)? How widespread is mobile ownership and what form(s) does it take? How good is mobile coverage? How cheap is it? Are there specific socio-cultural norms regulating the use or sharing of ICTs and especially mobile phones (for example for women, the youth? **Program Purpose** What are the goals and objectives of the DCE (project and evaluation)? Is the core purpose to inform, mobilize, involve or empower? Does the project aim to increase citizens' engagement? If so, how? Does the project build capacity of citizens to engage? If so, how? Does it aim to increase citizen influence/control of the process? How? Does the project build capacity of citizens to influence/control engagement process? If so, how? **Program Objective** To what extent do the program activities fit within the wider context? To what extent do the activities flow from the inputs? To what extent do the program outputs fit with the wider context? To what extent do the outputs flow from the activities? To what extent does the actual impact relate to the planned impact? How are citizens/citizenship portrayed in the logic of intervention? What monitoring processes were implemented to guide implementation? What evidence exists of monitoring data informing or changing implementation plans?

Lens | Control	
Example Evaluation Questions	
PRIMARY (ASSESSMENT/ANALYSIS)	**SUPPLEMENTARY (INFORMATION GATHERING)**
• What are citizens actually deciding and/or influencing and how much of this is institutionalized vs. by goodwill (i.e. consultation masquerading as more)? • How much of the program do citizens have the capacity to influence or control, and to what extent does the program seek to build up this capacity? • How is the DCE process perceived by "decision makers" of DCE and how open are they to being influenced/ informed by the results • Did people feel as though they could be heard? Were they satisfied with the nature of their opportunities to participate? • Did the digital engagement process meet the individual's expectations? • Was there appropriate information provided? (Was it timely, technically and linguistically accessible and easy to understand?) • How satisfied were participants about the issues addressed in the process?	Who are the actors in the DCE process? To what extent were citizens involved in the early decisions and design of the digital engagement process? Who defines/defined the goals and objectives? Who participates at what stage of the DCE? What organizations are involved? How much were 'end users' involved in designing and building the digital engagement platforms themselves? Have the designers of the initiative identified and developed relations with the relevant policy and civil society stakeholders? Do participants have control over the data that is being produced, including vetoing the collection or publication of data that might put at risk themselves or their community? Do participants have access to the products of the projects, tools and know-how necessary for analysis? What opportunities exist for co-option of the process, abuse of the system or manipulation of the results? What transparency and vigilance mechanisms are in place to ensure the decisions are followed through? How does collective bargaining take place? Who takes action on the part of participants? Was the digital engagement process fair? How satisfied were participants and other stakeholders with their level of input and influence over the process and the outcomes?

Lens | Participation	
Example Evaluation Questions	
PRIMARY (ASSESSMENT/ANALYSIS)	**SUPPLEMENTARY (INFORMATION GATHERING)**
• Who are the non-participants/non-users, and why have they not participated? • Are those who engage in DCE different from those who engage in CE? • Are the participants genuinely representative of the groups they represent, in the context of the goals of this engagement activity? • How meaningful is the citizen participation? • To what degree is there informed deliberation? • Has the approach to inclusion been informed by an understanding of disparities in the local information landscape and in people's abilities to engage effectively? • Does the support offered by the DCE initiative take into account citizen's prior mobilization, confidence and political know-how? • What further, if necessary, provisions are being made so that reprimands against citizens are prevented or minimized? • To what extent has democratic consultation been institutionalized? • What impact does the extent to which democratic consultation has been institutionalized have on DCE? • What impact do aspects of the physical environment have on participation? • Were participants satisfied with their ability to contribute to discussions? • Were participants satisfied with the way they were treated? • Were discussions open, honest and understandable? Were differing viewpoints encouraged and respected? • Have the different dimensions of inclusiveness been considered and prioritized appropriately? • Have relevant power relations and social dynamics been taken into account when designing the engagement process? • How is the notion of transparency expressed in the initiative's logic of intervention? What form does it take and how is it justified? How are the data generated or published by the initiative presented and shared?	Who is invited to participate? What criteria were used? What is their ability to participate in terms of: - literacy/IT literacy/information literacy? - understanding of issues? - communication skills? - time? - power, young/old, male/female, urban/rural? Who participates/who doesn't? Why do they participate? What is their desire to participate? How do they participate? What is the awareness of the effect of their participation? Do citizens have a clear understanding of the policy processes that the initiative is meant to influence? What social groups do the people who are engaged claim to represent? How are privacy concerns addressed? What influence do power dynamics between and within groups have? Were all potential participants provided sufficient information to help them understand the engagement process and participate in it effectively? Were a diverse range of people, or diverse views and opinions involved in the process? Were the people who were targeted the people who were eventually engaged with? Were relevant actors missing from the digital engagement? What was the demographic characteristic of participants? How many participated, how many chose not to and why? What was participants' previous experience in digital and/or other forms of citizen engagement? What was the satisfaction of participants (before and after the digital engagement)? What were the participant's motivations for engaging digitally?

Lens | Technology	
Example Evaluation Questions	
PRIMARY (ASSESSMENT/ANALYSIS)	**SUPPLEMENTARY (INFORMATION GATHERING)**
• To what extent does the choice of technology reflect the capacity of the implementers? • How widespread is the understanding of the use of technology? • How well understood and used is the data internally generated? • Who was engaged (compared to who the intended targets were) and who was missed? • How are citizens' views of the process reflected in participation rates, intensity of debates, follow-up actions over time? • Has the method has been rigorous enough to ensure that the relationship between the independent and dependent variables are clear enough to make a causal inference? (internal validity) • Could the same methods applied to other DCE initiatives? i.e. are the methods generalizable/ replicable, and if not, why not? (external validity) • How satisfied were participants with the digital engagement technology? • How appropriate was the level of privacy and security of the technology? • Are citizens expectations about what can be achieved through the platform realistic? • How satisfied were participants with the way they interacted and engaged with the technology? Was it fair across different groups?	Was the process properly run from a technical perspective (did the technology function as planned)? Why was digital technology used at all? What was the process behind the Digital Citizen Engagement intervention design and management? What were the intended results of this process? What were the inputs that fed into the Digital Citizen Engagement? What Digital Citizen Engagement activities happened; when, how, what technology, where, with who? What methods of delivery were used? What were the Digital Citizen Engagement outputs of the program? What system level data exists from the technology used? How many "episodes" of engagement where there? How do citizens themselves feel about their participation in the DCE initiative? What do citizens say about their involvement, especially in relation to the broader political context? How are the outputs monitored? How is the output monitoring data used in program management and learning?

Lens | Effects	
Example Evaluation Questions	
PRIMARY (ASSESSMENT/ANALYSIS)	**SUPPLEMENTARY (INFORMATION GATHERING)**
• What effects did the program have in relation to planned impact? What unintended outcomes were there? • How does DCE support individuals to become more effective citizens so that they can realize their rights? • Is participation in DCE congruent with citizens' experiences of participation in other development initiatives? If yes, how? If not, how? • What collective dimensions of participation, if any does the DCE initiative support? • How has the initiative curtailed wasteful public expenditure and contributed to improved services and citizen satisfaction? • Has the initiative supported/challenged public officials and service providers? • Did participants change their attitudes to digital engagement or the wider engagement process? • Did participants understand the goals of the engagement? • How much did the participant's attitudes shift before and after the engagement? • Did participants believe the engagement was worthwhile? • How do these results contribute to improving the underlying driver of the engagement? • How was the information generated by the digital engagement used by policy makers? • Did participants end up engaging in other ways or disengaging in other ways? • How satisfied were participants and other stakeholders with the outcomes? • How satisfied were involved or affected communities with the degree to which the outcomes represent their interests? • What are the recommendations for planning future digital engagement? • Were there any unforeseen consequences that were a direct result of the digital engagement (i.e. increased engagement in other mediums, increase or loss of trust, changing perceptions, etc.)? • Did the digital engagement affect a policy decision and how? • What was the impact of the engagement in terms of people, processes or services? • What were the outcomes or results from the digital engagement activity or program?	What effects did the program plan to have? How does collective action, support for civil society feature in an initiative's objectives? Is there a clearly articulated view of why such an aspect is important/ unimportant/not relevant for the DCE initiative? How is government/service providers' action conceptualized? What does it mean? What would it look like according to the theory and logic of the program? What assumptions are being made in the objective about how the initiative would promote positive change through government action? Does the initiative have a clear understanding of the policy processes its designers try to influence, including the pressure points, barriers for supporting accountability, especially in contexts where corruption is widespread? What worked well? What could be improved? What was learned? Did participants learn anything?

Toolkit 2 : Using the lenses in Scoping and Design stages

Section 3.5 there is an introduction to how the lenses can serve as an aid to early scoping of an evaluation, and to designing the evaluation and its goals and questions, and ***Table 6*** highlights key areas of interest for each lens that an evaluator may wish to consider and explore during these stages. The tables below take these areas of interest for each of the lenses (**Objective**, **Control**, **Participation**, **Technology** and **Effects**) and goes into more depth about the typical issues and concerns that might be relevant to these areas of interest, and suggests areas that it may be useful to look into during both the Scoping stage and when Designing the evaluation and its questions.

LENS 1: Objective	
(Is the program objective reasonable and appropriate, and to what extent does the logic of the program lead to it?)	
Areas of Interest	
What to explore at the Scoping stage (seeking to understand the explicit objective of the program and the wider environment, including the planned impact)	**What to consider at the Design stage** (the gaps in knowledge of the objective (is the objective explicit, or does it need to be inferred?), the arenas to explore further, and what counterfactual to use (if any))
The objective and goals of the engagement	
Are there specific goals available that explain the project? Is there a clear objective in the project linking (e.g.) activities, objectives, goals? Have the goals and the activities changed/ evolved over time?	Do the goals appear reasonable, practical, sensible? Do the goals and related documents/ discussions demonstrate that those in charge of the program appear to understand the nuances required to set effective goals for DCE (e.g. consider questions in lenses 2-4 in Toolkit 1)?
Specific goals related to the technology aspects of the program (N.B. note link/overlap with Lens 4)	
Is there clear rationale for both why technology was chosen as an engagement tool, and for the specific choice of technology platform? Is it clear what the goals are for the technology itself and do these goals align with the wider engagement goals of the program?	The gaps/questions here are likely to tie to questions related to Lens 4 in Toolkit 1
Sound logical theory grounded in reality	
What is the external reality of the program? How does the program relate to/tie in with other initiatives? With technological developments? Is there evidence of lessons from other DCE projects being incorporated?	Are the links reasonable and well thought out between activities and expected results in the (stated or implied) objective of the program? Do these links and their underlying assumptions appear to be grounded in reality or based on evidence and research? Are there falsifiable hypotheses which can be tested?
Different views of the program goals (N.B. Link/overlap with Lens 3)	
How do the different stakeholders view the program goals?	What are the implications of this for the program (see also questions related to Lens 3 in Toolkit 1)?
What is the counter-factual?	
Is there a valid 'control' group or data which the program can be evaluated against?	What might have happened if no (D)CE engagement took place at all?

LENS 2: Control	
(Who controls and influences the digital engagement process?)	
Areas of Interest	
What to explore at the Scoping stage (who is involved in decision-making at what stages; the mechanisms that exist to ensure fairness and equitability; what evidence of stakeholders' influence already exists)	**What to consider at the Design stage** (how to evaluate the extent of different stakeholders' influence on the program and the implications of that)
Types of involvement of different groups at different stages of the program	
To what degree are citizens, stakeholders, beneficiaries, end-users... engaged in the initial decisions, the design of the DCE process, the design of the technical platform, the delivery, the monitoring and evaluation..?	To what extent do people feel involved or in control? How do expectations of the level of participation measure against reality? How does the level of control change over the lifetime of the project?
What is in place to ensure the DCE program's processes are fair and equitable?	
Are there vigilance mechanisms in place and suitable levels of transparency to protect against corruption? What mechanisms are in place to ensure that decisions are implemented and decision-makers are held to account?	How effective are these mechanisms? Is there evidence of abuse, corruption or cheating the system by participants, program staff or decision makers?
Factors that influence the ability of different stakeholders to influence the process	
Who selected the technology? Are intermediaries (e.g. civil society groups, technology intermediaries) involved? What data-points are used, influencing what is collected, reported and given importance?	How familiar are people with the platform? How does it affect their participation? How does the mediation affect the participation? How much control is actually wielded by distant funders, by private sector technology partners, by industry experts etc.? How were the data points defined? Who made those decisions?
How is the sphere of influence of the program being decided?	
What ambition does the program have to influence others and to what extent is this reflected in program design? (may tie in with Lens 1- objective)	To what degree are the decision-makers genuinely open to being influenced by the results? Which aspects of this process been institutionalized or enshrined in law? How much influence do citizen participants genuinely have and is this capacity being built if it needs to be?

LENS 3: Participation	
(Who participates and how?)	
Areas of Interest	
What to explore at the Scoping stage (the target audience(s), their characteristics, how they are reached; the opportunities provided by the program for them to participate and at what level)	**What to consider at the Design stage** (how to assess the effectiveness of the program in enabling participation, and whether the level of participation achieved met the expectations of participants and stakeholders)
Recruitment and targeting	
Does the program target the entire population or specific sub-groups? Are participants self-selecting or being recruited via intermediaries such as civil society groups?	Is the opportunity to engage promoted equally to all potential participants, or is it reaching more of certain groups? What are the implications of the chosen recruitment method on people's engagement?
Why are some groups engaging or not?	
What characterizes the target audience in terms of (e.g.) availability, environmental/ societal influences, access to the engagement technology, desire to participate? What incentives has the program used to encourage engagement?	What is affecting (positively or negatively) people's capacity to engage? What importance is the target audience attaching to the engagement (e.g. when compared to livelihood activities)? How effective are any incentive schemes utilized?
How has technology changed the engagement dynamics?	
What technology has been used for engagement by the program, and why (note overlap with Lenses 1 and 2)? What engagement strategies has the program developed?	What effect has the choice of technology had on representation of (e.g.) hard-to-reach groups? How does technology affect group dynamics and the power of collective voices, how do issues of low bandwidth, poor mobile coverage, and power outages disproportionately affect poorer communities? How has the program responded?
What times/spaces for meaningful engagement exist?	
What opportunities (spaces and times) does the program provide for people to become informed about relevant issues and deliberate them?	To what extent are people engaging meaningfully, with an understanding of their role and the issues they are engaging with? How suitable is the level of deliberation for the goal of the program, and is it realistically available equally to all participants?

LENS 4: Technology	
(How effective and appropriate is the choice and delivery of the technology?)	
Areas of Interest	
What to explore at the Scoping stage (the technology used and the reasons for its selection, the cost; How privacy issues are managed; How the overall program was managed)	**What to consider at the Design stage** (how effective the technology is and the quality of how it (and the overall project) is managed, cost-effectiveness in comparison to alternative approaches, quality of data safeguarding)
Choice of technologies	
Which specific technologies were chosen and why, were other possibilities explored, was a non-technical option assessed?	Do the technologies chosen seem suited to the goals of the project, the technical activities expected of it? What additional value has the choice of technology brought to the program? What other impact (positive or negative) has the technology had on the engagement process?
Data management and privacy	
How does the program handle privacy issues resulting from citizen data being kept on a technical platform?	How effective are these safeguards? Does the program understand the potential for abuse of the system or the data and have processes in place to mitigate against this?
Use of time and resources	
What is the total cost of engagement of the program?	How is the program cost different than it might've been using different platforms or no technologies? What trade-offs have been made between quality and scale due to the technology? What would the costs have been in the counter-factual and how many citizens might have been engaged with?
Overall program management and learning	
What capacity does the program have for managing the technology? When problems occurred (technical or otherwise), how were they handled? Are systems in place to extract meaningful data from the systems and use this to monitor and seek to improve activity? What quality and accountability mechanisms are in place?	How suitable is the technology, how well is it managed, how good is its user experience, how accessible is it? How well is the wider program managed and delivered? How well have the relevant institutions and individuals learned from their experience? How suitable are the systems and how well are they managed?

LENS 5: Effects	
What effects do citizens have on processes and outcomes?	
Areas of Interest	
What to explore at the Scoping stage (the evidence (even anecdotal) that already exists of intended or unintended impacts; whether a 'control group' was identified or not; availability of baseline data; nature of the DCE project (e.g. designed as an RCT?)	**What to consider at the Design stage** (how to establish whether the intended impact materialized, how to notice and assess unintended consequences, the cost of collecting data on the outcome of interest, the contribution of technology to the identified changes)
How have citizens / participants changed as a result of the program?	
What expectations does the program have for individual change (relates to Lens 1)? (e.g. DCE can make people more aware of their rights, give people a sense of possibilities and an opportunity to build their confidence, to challenge injustices) – have any of these improvements been observed?	To what extent have these expectations been met? Have people been given more than just information, are they (e.g.) able to build their own skills and resources to improve their lives, are they able to better navigate bureaucracies, negotiate tensions, build alliances?
Has the program changed the way citizens organize collectively?	
Did the program seek to build the capacity of civil society or advocacy groups? If so, how?	Has capacity for advocacy and/or collective social action been increased or diminished? Have new spaces been created (online or offline) for public dialogue, and informed exchange of ideas, or have such spaces declined? Has the power of amplifying collective voices been realized and if so, has this resulted in positive results (e.g. increased collective bargaining power) or negative ones (e.g. domination of the engagement process by more vocal groups representing elite and middle classes)? Were any changes found intended or unintended?
How have decision-makers been impacted?	
What evidence already exists of changes to decision-makers/decision-making processes?	Is the engagement process being incorporated into decision making processes – i.e. are citizens actually deciding anything? Has the process opened up a more transparent way of making decisions so public scrutiny can take place? If so, is this transparency actually translating into improved accountability, are decisions being changed, challenged? Is quality and responsiveness of governance changing – is the feedback loop closed so they are they willing to listen and act and share information about these actions?

LENS 5: Effects	
What effects do citizens have on processes and outcomes?	
Areas of Interest	
What to explore at the Scoping stage (the evidence (even anecdotal) that already exists of intended or unintended impacts; whether a 'control group' was identified or not; availability of baseline data; nature of the DCE project (e.g. designed as an RCT?)	**What to consider at the Design stage** (how to establish whether the intended impact materialized, how to notice and assess unintended consequences, the cost of collecting data on the outcome of interest, the contribution of technology to the identified changes)
What tangible effects has the program had on the nature of life in its area?	
What impact (even anecdotal) has been noted on bigger issues such as inequality and poverty reduction, impact on specific indicators in the relevant sector such as educational attainment or maternal health outcomes? (these impacts are potentially hard to measure, and extremely hard to attribute direct causal effects to). What direct impact has been noted on service delivery by the program?	If no impact has been recorded (or there are gaps in the information - either qualitatively or quantitatively) - is there evidence of such an impact that could be gathered by the evaluation? By following the feedback loop can we identify exactly if and how the engagement activities resulted in changes – whether the changes are the inclusion of a new clause in a legal policy, improved quality of delivery of healthcare, or result in specific tangible actions such as fixing a faulty water pump.
What effects can be attributed specifically to the technology?	
Have any effects specifically due to the technology been noted or suggested so far?	What evidence is there that any impact can be specifically attributed to the choice of technology? (This could be (e.g.) to amplify or dull the wider CE impact, or could be tangential, such as improving ICT literacy skills. Were there instances when technology prevented change? (e.g. DCE can seek to build a direct relationship between the state and every citizen, which can undermine potential collective dynamics)

Appendix A: Global examples of Digital Citizen Engagement

In ***Section 2.2*** (***Table 2***), a range of DCE initiatives, tools and websites are shown. Links to further information on these and other examples of DCE projects around the world are below. Evolving and maintained lists of cases can also be found at *www.participedia.net* and *www.participationcompass.org*.

Project	Country/ies	Link to further information
Agora em Rede Canoas	Brazil	*agoraemrede.com.br*
Alaveteli	Global	*alaveteli.org*
Avaaz	Global	*avaaz.org*
Barrio Digital	Bolivia	*barriodigital.cl*
CGNetSwara	India	*cgnetswara.org*
Change.org	Global	*change.org*
Check My School	Philippines	*checkmyschool.org*
Code For America	USA	*codeforamerica.org*
d-Brain	South Korea	*digitalbrain.go.kr*
Daraja Maji Matone	Tanzania	*blog.daraja.org/2012/02/so-what-have-we-learnt-summarising.html*
DevTracker	Global	*devtracker.dfid.gov.uk*
Fix My Street	Global	*fixmystreet.org/sites/*
g0v	Taiwan	*g0v.tw*
Hello Sarkar	Nepal	*doinepal.gov.np/home/feedback*
I Change My City	India	*ichangemycity.com*
I Paid A Bribe	India, Pakistan	*ipaidabribe.com*
Jaankari	India	*biharonline.gov.in/RTI*
M4W	Uganda	*m4water.org*
Magyar Leaks	Hungary	*atlatszo.hu/magyarleaks*
Maji Voice	Kenya	*majivoice.com*
Map Kibera	Kenya	*mapkibera.org*

Project	Country/ies	Link to further information
Mapa de Transparencia	Brazil	*mapa.rs.gov.br*
Map Tandale	Tanzania	*explore.ramanitanzania.org*
Mejora Tu Escuela	Mexico	*mejoratuescuela.org*
Mexico Como Vamos	Mexico	*mexicocomovamos.mx*
Mi Medellin	Colombia	*mimedellin.org*
Mobile Monitors	Nigeria	*mobilemonitors.org*
Mumbai Votes	India	*mumbaivotes.com*
Mzalendo	Kenya	*info.mzalendo.com*
Namma Dhwani	India	*communityvoices.in/directory/community-media-profile/1327*
Note My Vote	UK	*notemyvote.co.uk*
Observatorio Anti-corrupcion	Chile	*observatorioanticorrupcion.cl*
Open Data Kenya	Kenya	*opendata.go.ke*
On Our Radar Kenya	Kenya	*onourradar.org/kenya*
Open Town Hall	USA	*opentownhall.com*
Por Mi Barrio	Uruguay	*pormibarrio.uy*
RadioActive	India	*jgi.ac.in/radioactive/Aboutus.htm*
Redencao Park	Brazil	*redencao.cc*
Rio 1746	Brazil	*1746.rio.gov.br*
Sauti Za Wananchi	Tanzania	*twaweza.org/go/sauti-za-wananchi-english/#_*
SeeClickFix	USA	*seeclickfix.com*
They Work For You	UK	*theyworkforyou.com*
TracFM	Uganda	*tracfm.org*
Transparency International Election Monitoring	Zimbabwe	*tizim.org*
U-Report	Uganda	*ureport.ug/*
Ushahidi	Global	*ushahidi.com*
What Do They Know	UK	*whatdotheyknow.com*
Write To Them	UK	*writetothem.com/*

Appendix B: Results Indicators for Citizen Engagement

The table below builds on the approach to evaluation proposed in the Strategic Framework for Mainstreaming Citizen Engagement in World Bank Group Operations. When evaluating Digital Citizen Engagement a range of indicators can be used, however the exact measurements and sources may be different depending on the readily available sources of data.

Intermediate Outcome Indicators	Final Outcome Indicators
Citizen Engagement Activity | Consultation (one time or institutionalized as part of a process, e.g. budget formulation, legislation process, resettlement etc.) **Mechanisms:** focus groups, consultations, surveys etc.	
Citizens consulted reporting satisfaction with key aspects of consultation process (information available, facilities, facilitation, location, language etc.) (% rating 4 or above on a rating scale 1-7 with 7 being most satisfied) Share of participants in consultations that are from vulnerable and marginalized citizens (e.g. women, slum dwellers, youth, disabled indigenous people) (%) A qualitative assessment undertaken of how consultation feedback has impacted the matter subjected to consultation based on analysis of materials used to close the feedback loop (Yes/No)	Beneficiaries that feel project investments reflected their needs (percentage) Citizens who consider that design and/or implementation of policy or program or project subjected to consultation] is responsive to their views (% rating 4 or above on a rating scale 1-7 with 7 being most satisfied) Changes to policies subjected to consultation as a result of consultations (yes/no) Changes to program/project activities subjected to consultation as a result of consultations (yes/no)

Intermediate Outcome Indicators	Final Outcome Indicators
Citizen Engagement Activity | Citizen/community collaboration (on planning and/or execution of a policy, program or project) Mechanisms: User groups, Participatory budgeting, Participatory audit Etc	
Project-supported organizations implementing participatory methods (number) Government /service provider puts in place channels for collaboration that can be used on a sustainable basis (yes/no) Vulnerable and marginalized beneficiary population who participate in non-project consultations and decision making forums (percentage) Representatives in community-based decision making and management structures that are from the vulnerable or marginalized beneficiary population (percentage) A qualitative assessment undertaken of how collaboration process was implemented and citizen feedback incorporated in programs and projects based on analysis of materials used to close the feedback loop (Yes/No)	Citizens who consider that [design and/or implementation of policy or program or project subjected to collaboration] is responsive to their views (% rating 4 or above on a rating scale 1-7 with 7 being most satisfied) Change in time it takes to deliver a service (Hours/ Days) (%)
Citizen Engagement Activity | Beneficiary feedback Mechanisms: Satisfaction surveys, Citizen/community report cards, Hotlines, SMS/ online feedback, etc.	
Actions taken by program/project managers based on user/beneficiary feedback (yes/no) Feedback collecting programs/institutions publish reports on feedback received and how this feedback has been used (Yes/No) Proportion of total beneficiaries who provided feedback (%) Share of feedback providers that are from vulnerable and marginalized or other target groups (e.g. geographic group) (%) Project supported organizations that use feedback provided by independent monitoring (Number) Reports/allegations of abuse/waste/corruption received, vetted and referred to appropriate authorities (%) Waste/abuse/corruption allegations investigated (%)	Beneficiaries satisfied with specified dimensions [e.g. access, delivery time, responsiveness to needs, transaction costs, operational efficiency, bribery experience, staff attitude, quality of facilities] (%) Reports/allegations of abuse/waste/corruption investigations resulting in remedial actions and/or sanctions (%)

Intermediate Outcome Indicators	Final Outcome Indicators
Citizen Engagement Activity | Grievance redress mechanisms	
Affected households that filed grievances/complaints (%) Periodic reports on GRM published (Yes/No) Average time required to resolve complaints (days)	Grievances registered related to delivery of project benefits that are actually addressed (percentage) Grievances responded and/or resolved within the stipulated service standards for response times (%) Project management improvement actions triggered by grievances/complaints (yes/no) Complainants satisfaction with response and grievance redress process (% rating 4 on rating scale 1-7 with 7 being most satisfied)
Citizen Engagement Activity | Citizen monitoring, evaluation and oversight Mechanisms: Social audits, Public expenditure tracking surveys, Citizen juries, Third Party Monitoring (e.g. in procurement, service delivery), etc.	
Report on findings of citizen led monitoring published (Yes/No) Dialogue events organized to discuss monitoring/oversight findings with relevant authorities (number)	Beneficiaries satisfied with specified dimensions monitored [e.g. access, d livery time, responsiveness to needs, transaction costs, operational efficiency, bribery experience, staff attitude, quality of facilities] (%) Violations and issues identified in the monitoring/ oversight activities on which authorities act (%) Changes to policies as a result of citizen monitoring (yes/no) Changes to program/project activities as a result of citizen monitoring (yes/no)

Intermediate Outcome Indicators	Final Outcome Indicators
Citizen Engagement Activity | Empowering citizens/communities with resources and decision-making powers Mechanisms: Community-driven development projects, Community management of natural resources, (Un)conditional cash transfer programs, etc.	
Representatives in community-based decision making and management structures that are from the vulnerable or marginalized beneficiary population (percentage) [Core Indicator] Vulnerable and marginalized beneficiary population participating in non-project consultations and decision-making forums (percentage) Citizen groups that are complying with conditions for the delegated funds according to independent third party verification (% or number) Amount of delegated funds and/or resources being managed by citizen groups or households (amount in currency and/or physical units e.g. acres of pasture land) Community contribution to total project cost (%)	Beneficiaries that feel project investments reflected their needs (percentage) [Core Indicator] Sub-projects or investments for which arrangements for community engagement in post-project sustainability and/or operations and maintenance are established (percentage) [Core Indicator] Improved quality of service as measured by [specific project measure] (%) Improved community access to targeted services as measured by [specified project measure] (%)
Citizen Engagement Activity | Capacity building for engagement (both of public officials and citizens/CSOs)	
Increased citizen engagement knowledge and skills as a result of training (pass rate on quick test administered/share of training participants rating skills acquired as useful) CE initiatives launched/completed applying the skills learned (number) Citizen programs with trained persons involved in consultation and GRM activities (%) Monitoring committees [e.g. school management committees] trained in participatory monitoring (%)	Citizen engagement activities where trained persons are involved (number)

Appendix C: Field evaluation data collection methods and costs

The four field evaluations that were undertaken as part of the development of this Guide, each utilized different methods and data collection tools. These are summarized below along with indicators for cost and reach. These figures are indicative and should not necessarily be assumed to hold true in different regions, but may be a helpful guide when considering the data collection needs for an evaluation.

Evaluation methods	Cost	Reach
Brazil		
Online web-form survey for digital voters (using Survey Monkey)	$2-300 (Survey Monkey annual subscription)	33,758 surveys completed over 3 days
Face-to-face survey in physical polling stations	~$20,000 (enumerator training, logistics of being at physical locations of polling stations etc.)	1,923 respondents (50 enumerators required)
Interactive Voice Response randomized automated dialling telephone survey	~$5,000	2,173 responded (including 1,373 non-voters) over 3 days of calls
Supplementary interviews with field staff, government officials, local academic experts	Minimal telephone interview time	Ad-hoc basis to answer specific follow-up questions
Systems data analysis; SMS contacts database, meeting participant lists, call centre transaction logs.	~$1,000 on data entry and digitisation of selected paper-based data plus considerable staff time on system data quality assessment and analysis	NA
A small selection of supplementary interviews with program staff, citizens and local officials	Local research staff fees and limited telephone interview time	Ad-hoc basis to explore scoping and specific follow-up questions

Evaluation methods	Cost	Reach
Kenya		
Online surveys (conducted through enumerators using (using Survey Monkey) with complainants to Maji Voice	~$4,745	1,064 surveys by 4 enumerators, completed over 8 days
Uganda		
SMS Survey with U-Reporters	~$5,775	5,763 respondents
SMS U-Reporter poll	~$4,588	286,800 respondents
Household survey	~$18,932	1,188 respondents
RIWI RDIT	~$8,500	2,884 respondents
Face-to-face interviews with U-Reporters and government officials	~$4690	20 respondents interviewed by one interviewer over 24 days (including preparation on site etc.)

Appendix D: Links to useful evaluation software tools

The list below contains links to examples of digital tools which may be useful for evaluations. This list is not exhaustive and should not be treated as a recommendation for any particular tool, but a starting point for understanding what is available. Note that only some of the tools below are open source and/or use open standards – some are proprietary solutions or may limit the ease with which their data can be transferred to other systems.

Survey, data collection and visualization/mapping tools	
Asterisk	www.asterisk.org
Commcare	www.commcarehq.org
CommConnect	www.dimagi.com/commconnect
Crowdmap	crowdmap.com
Ctalk	www.ctalk.co.uk
EchoMobile	www.echomobile.org
Elva	www.elva.org
EpiCollect	www.epicollect.net
Fieldata	www.fieldata.org
First Mile Geo	www.firstmilegeo.com
FluidSurveys	fluidsurveys.com
FormSite	www.formsite.com

Survey, data collection and visualization/mapping tools	
FreedomFone	www.freedomfone.org
FrontlineSMS/Cloud	www.frontlinesms.com
Google Forms	www.google.co.uk/forms/about
Groundsource	www.groundsource.co
iFormBuilder	www.iformbuilder.com
KeySurvey	www.keysurvey.co.uk
Kobo Toolbox	www.kobotoolbox.org
LimeSurvey	www.limesurvey.org
Magpi	home.magpi.com
Mxit	www.mxit.im
Nokia Data Gathering	nokiadatagathering.net
Open Data Kit	opendatakit.org
Open Street Map	www.openstreetmap.org
OpenXData	www.openxdata.org
Pendragon Forms	www.pendragonsoftware.com
PoiMapper	www.poimapper.com
PollDaddy	polldaddy.com
Qualtrics	www.qualtrics.com
Quantum GIS	www.qgis.org
QuestionPro	www.questionpro.com
RapidSMS	www.rapidsms.org

Survey, data collection and visualization/mapping tools	
Resource Map	instedd.org/technologies/resource-map
SnapSurveys	www.snapsurveys.com
StatPac	www.statpac.com
Survey Monkey	www.surveymonkey.com
SurveyGizmo	www.surveygizmo.com
SurveyGold	www.surveygoldsolutions.com
SurveyPro	www.apian.com
TaroWorkz	taroworks.org
Telerivet	telerivet.com
Text to Change	ttcmobile.com
Textlt	textit.in
Twilio	www.twilio.com
Ushahidi	www.ushahidi.com
ViewWorld	www.viewworld.net
Voicent IVR Studio	www.voicent.com/ivr.php
VOTO Mobile	www.votomobile.org
Vumi	vumi.org
Zoomerang	www.zoomerang.com

Tools for Quantitative Analysis	
AMOS	www-03.ibm.com/software/products/en/spss-amos
ArcGIS	www.esri.com/software/arcgis
Calc (Libre office)	www.libreoffice.org/discover/calc
Echo mobile	www.echomobile.org/public/main
EpiData	www.epidata.dk
EQS	www.mvsoft.com/eqs60.htm
Gnumeric	www.gnumeric.org
Google Charts	developers.google.com/chart
Google Drive sheets	www.google.co.uk/sheets/about
Google forms	www.google.co.uk/forms/about
HLM	www.ssicentral.com/hlm/
IBM - Many eyes	www-969.ibm.com/software/analytics/manyeyes
i-Work/Numbers	www.apple.com/uk/mac/numbers
Minitab	www.minitab.com/en-us
MLwiN	www.bristol.ac.uk/cmm/software/mlwin/
MPLUS	www.statmodel.com
NetMiner	www.netminer.com
Pandas (Python)	www.pandas.pydata.org
QGIS	www.qgis.org
Qualtrics	www.qualtrics.com
R	www.r-project.org

Tools for Quantitative Analysis	
RapidMiner	rapidminer.com
SAS	www.sas.com/en_us/software/sas9.html
SPSS	www-01.ibm.com/software/uk/analytics/spss
STATA	www.stata.com
Statsmodels (Python)	statsmodels.sourceforge.net
Survey Monkey	www.surveymonkey.com
Tableau	www.tableau.com
UNICET	sites.google.com/site/ucinetsoftware/home
Vizualize Free	visualizefree.com

Tools for qualitative analysis and textual data mining	
Alceste	www.image-zafar.com/en/alceste-software
Atlas.ti	atlasti.com
DataWatch	www.datawatch.com
Dedoose	www.dedoose.com
Discover Text	discovertext.com
Ethnogragh	www.qualisresearch.com
ExactTarget	www.exacttarget.com/uk
ForSight platform	www.crimsonhexagon.com/social-media-intelligence/forsight-platform
Fs/QCA	www.u.arizona.edu/~cragin/fsQCA/software.shtml
GATE	gate.ac.uk/sentiment
HyperResearch	www.researchware.com/products/hyperresearch.html

Tools for qualitative analysis and textual data mining	
InfoRapid	www.inforapid.de/html/searchreplace.htm
LingPipe	alias-i.com/lingpipe/demos/tutorial/sentiment/read-me.html
MAXQDA	www.maxqda.com
Mozdeh	mozdeh.wlv.ac.uk
NodeXL	nodexl.codeplex.com
NLP Stanford demos	nlp.stanford.edu:8080/sentiment/rntnDemo.html
NVivo	www.qsrinternational.com/products_nvivo.aspx
Overview	www.overviewproject.org
QDA Miner	provalisresearch.com/products/qualitative-data-analysis-software
QSR N6	www.qsrinternational.com/products_previous-products_n6.aspx
Qualrus	www.ideaworks.com/qualrus
R DataMiner	www.rdatamining.com
Rapid Miner	rapidminer.com
SAS Text Miner	www.sas.com/en_us/software/analytics/text-miner.html
Studiocode	www.studiocodegroup.com
Tams Analyzer	tamsys.sourceforge.net
TranSana	www.transana.org
Tosmana	www.compasss.org/software.htm#tosmana
Voyant Tools	docs.voyant-tools.org
Weft QDA	www.pressure.to/qda

Bibliography

Accenture (2012). Build it and they will come? The Accenture Digital Citizen Pulse Survey and the Future of Government Operations. Available at: *http://www.accenture.com/SiteCollectionDocuments/PDF/Accenture-Digital-Citizen-FullSurvey.pdf*

Alvesson, M. and Sköldberg, K. (2009). ***Reflexive methodology: New vistas for qualitative research***. Sage publications ltd. [online]. Available from: *http://books.google.co.uk/books?hl=en&lr=&id=32G4M7-20xgC&oi=fnd&pg=PP1&dq=alvesson+skoldberg&ots=G1Nlr9Ys-RA&sig=8RbkumaRoCp5s6tZzBNstQNx4s8* [Accessed July 30, 2012].

Ashford, M (2012). Early Findings From the CivicPlus Digital Citizen Engagement Survey. CivicPlus blog. Available at: *http://blog.civicplus.com/blog/bid/203653/Early-Findings-From-the-CivicPlus-Digital-Citizen-Engagement-Survey*

Avila, R., Feigenblatt, H., Heacock R., Heller, N. (2010). ***Global Mapping of Technology for Transparency and Accountability***. London, UK: Transparency and Accountability Initiative.

Bailur, S. and Longley, T. (2014). ***Assessing the Impact of Alaveteli***. London: mySociety. [online]. Available from: *https://www.mysociety.org/2014/11/19/research-assessing-the-impact-of-alaveteli*

Bamberger, M. (2006). ***Conducting quality impact evaluations under budget, time and data constraints***. World Bank, Independent Evaluation Group/Poverty Analysis, Monitoring and Impact Evaluation Thematic Group.

BangTheTable. (2012). Managing consultation fatigue. [online]. Available from: *http://bangthetable.com/freebies/tips-sheets/managing-consultation-fatigue*

Bay, K. (2011). ***Does Participatory Governance Deliver? Citizen Participation and Social Service Delivery in Nicaragua***. Brown University.

Baykurt, B. (2012). Redefining Citizenship and Civic Engagement: political values embodied in FixMyStreet.com. ***Selected Papers of Internet Research***, 0(12.0). [online]. Available from: *http://spir.aoir.org/index.php/spir/article/view/16* [Accessed September 26, 2013].

Beardon, H. (2013). ***CONNECT! A practical guide to using ICTs in Plan projects***. Plan International.

Bell, M. and Flock, E. (2011). 'A Gay Girl in Damascus' comes clean. ***The Washington Post***. [online]. Available from: *http://www.washingtonpost.com/lifestyle/style/a-gay-girl-in-damascus-comes-clean/2011/06/12/AGkyHoRH_story.html* [Accessed January 10, 2015].

Benest, G. (2014). Solitary Squares to Collaborative Circles. [online]. Available from: *http://insightshare.tumblr.com/post/104665468060/solitary-squares-to-collaborative-circles* [Accessed January 8, 2015].

Bertot, J. C., Jaeger, P. T., Munson, S., and Glaisyer, T. (2010). Engaging the public in open government: The policy and government application of social media technology for government transparency. IEEE Computer, 43(11), 53-59

Bertot, J.C., Jaeger, P.T. &Grimes, J.M. (2010b). Using ICTs to create a culture of transparency: E-government and social media as openness and anti-corruption tools for societies. *Government Information Quarterly*, 27(3), pp.264–271.

Bertot, J.C., Jaeger, P.T. and Grimes, J.M. (2010a). Crowd-sourcing transparency: ICTs, social media, and government transparency initiatives. In pp. 51–58.

Better Evaluation (2014a), Plan and Manage an Evaluation. Accessed 20/01/2015. Available from: *http://betterevaluation.org/start_here/plan_manage_evaluation*

Better Evaluation (2014b), Decide Purpose. Accessed 20/01/2015. Available from: *http://betterevaluation.org/plan/frame/decide_purpose*

Better Evaluation (2014c), Decide which evaluation method to use. Accessed 20/01/2015. Available from: *http://betterevaluation.org/start_here/decide_which_method*

Better Evaluation (2014d), Accessed 7/2/2015. Fools' gold: the widely touted methodological "gold standard" is neither golden nor a standard. Available from: *http://betterevaluation.org/blog/fools_gold_widely_touted_methodological_gold_standard*

Blumestock, J.E. and Eagle, N. (2012). Divided We Call: Disparities in Access and Use of Mobile Phones in Rwanda. *Information Technologies and International Development*, 8(2), pp.1–16.

Carr, E.R. (2010). The place of stories in development: creating spaces for participation through narrative analysis. *Development in Practice*, 20(2), pp.219–226.

Chambers, R. (1997). *Whose Reality Counts? Putting the First Last*. London, UK: Practical Action Publishing.

Chambers, R. (2008). Revolutions in Development Enquiry. London: Earthscan.

CivicPlus (2012), Digital Disconnect: The gaps between Government-to-Citizen online interactions. *http://oregoncapitalprsa.org/wp-content/uploads/2013/06/101812_DigitalDisconnect2.pdf*

Cohen, L., L. Manion, and K. Morrison (2011). 'Ex Post Facto Research.' In *Research Methods in Education*, 303-11. London ; New York: Routledge.

Cook, T.D., Shadish, W.R. and Wong, V.C. (2008). Three conditions under which experiments and observational studies produce comparable causal estimates: New findings from within-study comparisons. *Journal of policy analysis and management*, 27(4), pp.724–750.

Corlazzoli, V. (2014), ICTs for Monitoring and Evaluation of Peacebuilding Programmes, Department for International Development. *https://www.sfcg.org/wp-content/uploads/2014/05/CCVRI-SSP-_ICT-and-ME-_Final.pdf*

Davenport, S. (2013). Prioritise citizen feedback to improve aid effectiveness. ***the Guardian***. [online]. Available from: *http://www.theguardian.com/global-development-professionals-network/2013/jul/22/feedback-loops-citizen-development* [Accessed December 16, 2014].

Davies, T., and Fumega, S. (2014). Mixed incentives: Adopting ICT innovations for transparency, accountability, and anti-corruption. ***U4 Issue***, ***2014***(4).

Duflo, E., Hanna, R., and Rya, S. P. (2012). 'Incentives Work: Getting Teachers to Come to School.' ***The American Economic Review*** 102(4): 1241–78.

Farmer, L and Boots, M. (2014). Barriers and Solutions in using M4D: Connecting Directly to Citizens for Scalable Impact. In Proceedings of 4th International Conference on M4D Mobile Communication for Development, 2014, Dakar, Senegal, April 7-9 2014

Fox, J. (2007). The Uncertain Relationship Between Transparency and Accountabilty. ***Development in Practice***, 17(4), pp.663–71.

Frohlich, D. , Rachovides, D., Riga, K., Bhat, R.,Frank M., Edirisinghe, E.,Wickramanayaka , D., Jones, M., Harwood, W. (2009). StoryBank: mobile digital storytelling in a development context. In ***Proceedings of the SIGCHI Conference on Human Factors in Computing Systems***. ACM, pp. 1761–1770. [online]. Available from: *http://dl.acm.org/citation.cfm?id=1518972* [Accessed January 12, 2015].

Fruchterman, J. (2011). Beneblog: Technology Meets Society: Issues with Crowdsourced Data Part 2. [online]. Available from: *http://benetech.blogspot.co.uk/2011/03/issues-with-crowd-sourced-data-part-2.html* [Accessed January 10, 2015].

Fung, A., Gilman, H. R., and Shkabatur, J. (2013). 'Six Models for the Internet+politics.' International Studies Review 15(1): 30–47.

Fung, A., Gilman, H.R. and Shkabatur, J. (2010). ***Impact Case Studies from Middle Income and Developing Countries***. London, UK: Transparency and Accountability Initiative. [online]. Available from: *http://www.transparency-initiative.org/wp-content/uploads/2011/05/impact_case_studies_final1.pdf*

Funnell, S. C., and P. J. Rogers (2011). ***Purposeful Program Theory: Effective Use of Theories of Change and Logic Models***. 1st ed. ed. San Francisco, CA: Jossey-Bass.

Galtung, F. (2014). The World Bank's 100% citizen feedback agenda: a daunting challenge and an amazing opportunity. ***Integrity Action***. [online]. Available from: *http://www.integrityaction.org/anonymous/world-bank%E2%80%99s-100-citizen-feedback-agenda-daunting-challenge-and-amazing-opportunity* [Accessed January 9, 2015].

Gaventa, J. and Barrett, G. (2010). So What Difference Does it Make? Mapping the Outcomes of Citizen Engagement.

Green, D.P. and Gerber, A.S. (2008). ***Get out the vote: How to increase voter turnout***. Brookings Institution Press. [online]. Available from: *https://books.google.co.uk/books?hl=en&lr=&id=LK-GaYyZqZbEC&oi=fnd&pg=PR7&dq=gerber+green+get+out&ots=_5wgOMjCFn&sig=YJ_uPr93Vghal-8M44W9y-4w3Mbc* [Accessed January 12, 2015].

Grossman, G., Humphreys, M., and Sacramone-Lutz, G. (2014). '"I Wld Like U Wmp to Extend Electricity 2 Our Village": On Information Technology and Interest Articulation.' ***American Political Science Review*** 108(03): 688-705.

Groves, L. (2014). 4 Steps to Beneficiary Feedback in Evaluation. ***Beneficiary feedback in evaluation and research***. [online]. Available from: *https://beneficiaryfeedbackinevaluationandresearch.wordpress.com/2014/12/08/4-steps-to-beneficiary-feedback-in-evaluation* [Accessed January 10, 2015].

Groves, R., Fowler, F., Couper, M., Lepkowski, J., Singer, E., Tourangeau, R. (2009). Survey Methodology. New Jersey: John Wiley and Sons.

Groves, Robert M. (2011). Three eras of survey research. Public Opinion Quarterly, 75(5), 861-871.

Guijt, I. (2014). Week 50: Feedback loops – new buzzword, old practice? | Better Evaluation. [online]. Available from: *http://betterevaluation.org/blog/feedback_loops_new_buzzword_old_practice* [Accessed January 12, 2015].

Haikin, M. and Duncombe, R. (2013). A Framework to Assess Participation and Empowerment Impacts of ICT4D Projects. Development Informatics Working Paper 55. Available from: *http://www.seed.manchester.ac.uk/subjects/idpm/research/publications/wp/di/di-wp55*

Heider, C. (2014). What makes evaluations influential? [online]. Available from: *http://gendereval.ning.com/profiles/blogs/what-makes-evaluations-influential?xg_source=msg_mes_network* [Accessed January 8, 2015].

Hisroshi, K, Tetsua, N., Hideo, W. (2004). Deeper sentiment analysis using machine translation technology. COLING '04 Proceedings of the 20th international conference on Computational Linguistics.

IAP2 (2007), IAP2 Spectrum of Public Participation. International Association for Public Participation. Available from: *http://c.ymcdn.com/sites/www.iap2.org/resource/resmgr/imported/IAP2%20Spectrum_vertical.pdf*

IBRD. (1996). ***The World Bank Participation Sourcebook***. The International Bank for Reconstruction and Development / The World Bank.

ICTworks. (2014). Fail Festival 2014: Failure Happens and Flowers Grow From It | ICT Works. [online]. Available from: *http://www.ictworks.org/2014/12/15/fail-festival-2014-failure-happens-and-flowers-grow-from-it* [Accessed January 8, 2015].

InfoDev (2015). ***Framework For The Assessment OF ICT Pilot Projects***. [online]. Available from: *http://www.infodev.org/articles/framework-assessment-ict-pilot-projects* [Accessed January 8, 2015].

InsightShare (2015). Capacity Building in Participatory Video – Myanmar. Insight Share. Available from *http://www.insightshare.org/watch/video/capacity-building-participatory-video-myanmar* [Accessed 7/2/2015]

InterAction (2013). Embracing Evaluative Thinking for Better Outcomes: Four NGO Case Studies. InterAction and CLEAR-AA. Available from: *http://www.interaction.org/sites/default/files/EvaluativeThinkingReport_FINAL_online.pdf*

John, P. (2011). ***Taking Political Engagement Online: An Experimental Analysis of Asynchronous Discussion Forums***. Rochester, NY: Social Science Research Network. [online]. Available from: *http://papers.ssrn.com/abstract=1762895* [Accessed January 12, 2015].

Johnson, P. and Robinson, P. (2014). Civic Hackathons: Innovation, Procurement, or Civic Engagement? ***Review of Policy Research***, 31(4), pp.349–357.

Kelbert, A., McGee, R. and Carlitz, R. (2013). Understanding 'the users' in Technology for Transparency and Accountability Initiatives. [online]. Available from: *http://www.ids.ac.uk/publication/understanding-the-users-in-technology-for-transparency-and-accountability-initiatives* [Accessed January 12, 2015].

Kessler, A. and Tanburn, J. (2014). Why evaluations fail. Donor Committee for Enterprise Development.

King, S.F. and Brown, P. (2007). Fix my street or else: using the internet to voice local public service concerns. In ICEGOV '07. New York, NY, USA: ACM, pp. 72–80. [online]. Available from: *http://doi.acm.org/10.1145/1328057.1328076* [Accessed September 30, 2013].

Knight Foundation. (2014). Trends in Civic Tech. [online]. Available from: *http://www.knightfoundation.org/features/civictech* [Accessed January 12, 2015].

Kube, M., Hilgers, D., Fueller, J. and Koch, G. (2013). Re-Thinking Citizen Participation in the Public Sector: Adopting Open Innovation for Collaborative Idea Generation. Available at SSRN: *http://ssrn.com/abstract=2336673*

LabsCivico. (2013). Laboratorio de Innovación en Transparencia del Consejo para la Transparencia. [online]. Available from: *http://www.consejotransparencia.cl/labcivico* [Accessed January 12, 2015].

Leavy, J. (2014). ***How useful are RCTs in measuring transparency and accountability projects***. University of East Anglia, Norwich: Making all Voices Count.

Lemaire, I. and Muñiz, S. (2011). ***Participatory Video for Monitoring and Evaluation of Community Based Adaptation to Climate Change***. Centre for Development Informatics (CDI), University of Manchester, UK With the support of the International Development Research Centre (IDRC).

Lincoln, S. Y., and Guba, E. G. (1985). Naturalistic inquiry. Thousand Oaks, CA: Sage

Long, J., Woodman, L. and Brindley, W. (2014). ***Technology in Development: Engaging with people in emerging markets***. Accenture Development Partnerships/NetHope. [online]. Available from: *http://www.accenture.com/us-en/Pages/insight-role-technology-addressing-development-challenges.aspx*

Lopes, C. and Srinivasan, S. (2014). Africa's Voices: Using mobile phones and radio to foster mediated public discussion and gather public opinions in Africa. Centre of Governance and Human Rights, working paper 9. University of Cambridge.

Lunch, N. and Lunch, C. (2006), Insights into Participatory Video: A Handbook for the Field. InsightShare. *http://insightshare.org/sites/insightshare.org/files/file/Insights%20into%20Participatory%20Video%20-%20A%20Handbook%20for%20the%20Field%20%28English%29%281%29.pdf*

McGee, R. and Carlitz, R. (2013). ***Learning Study on 'The Users' in Technology for Transparency and Accountability Initiatives: Assumptions and Realities***. Brighton, UK: The Institute of Development Studies, University of Sussex.

Melville, P., Chenthamarakshan V., Lawrence, R., Powell, J., Mugisha, M. (2013). Amplifying the Voice of Youth in Africa via Text Analytics. Proceedings of KDD Conference, August 11-14, Chicago.

Mitchell, C. and de Lange, N. (2012). ***Handbook of participatory video***. AltaMira Press. [online]. Available from: *https://books.google.co.uk/books?hl=en&lr=&id=-gI4Osn-bsYC&oi=fnd&pg=PR7&dq=participatory+video+milne&ots=R_XwIpyYdm&sig=_zzOGwmwtrMMO88CSnWINvjnRww* [Accessed January 12, 2015].

Nabatchi, T. (2012). ***A manager's guide to evaluating citizen participation***. Washington D.C.: IBM Center for the Business of Government.

National Democratic Insitute. (2014). ***Citizen Participation and Technology: An NDI Study***. Washington DC: National Democratic Institute. [online]. Available from: *https://www.ndi.org/node/21589* [Accessed January 12, 2015].

O'Donnell, A. (2014). Developing an Organizational Policy for Responsible Data, Oxfam. *https://responsibledata.io/developing-an-organizational-policy-for-responsible-data/*

Paice, E. (2014), THE BOOKLOVERS, THE MAYORS AND THE CITIZENS: PARTICIPATORY BUDGETING IN YAOUNDÉ, CAMEROON, Africa Research Institute.

Paine, K.D. (2011). Measure What Matters: Online Tools For Understanding Customers, Social Media, Engagement, and Key Relationships. Hoboken, NJ: Wiley.

Raftree, L. (2013). Benefits, barriers and tips for ICT-enabled M&E. ***Wait... What?***. [online]. Available from: *http://lindaraftree.com/2013/04/17/benefits-barriers-and-tips-for-ict-enabled-me* [Accessed January 12, 2015].

Raftree, L. (2014). Devsplaining. ***Wait... What?***. [online]. Available from: *http://lindaraftree.com/2014/07/22/devsplaining* [Accessed January 12, 2015].

Raftree, L. and Bachan, K. (2013). ***Integrating Information and Communication Technologies into Communication for Development Strategies to Support and Empower Marginalized Adolescent Girls***. UNICEF.

Raftree, L. and Bamberger, M. (2014). ***Emerging Opportunities: Monitoring and Evaluation in a Tech- Enabled World***. The Rockerfeller Foundation.

Reitmaier, T., Bidwell, N.J. and Marsden, G. (2011). Situating digital storytelling within African communities. ***International Journal of Human-Computer Studies***, 69(10), pp.658–668.

Sasaki, D. (2010). ***Technology for Transparency: The role of technology and citizen media in promoting transparency, accountability and civic participation***. Amsterdam: Technology for Transpareny Network. [online]. Available from: *http://ifap-is-observatory.ittk.hu/node/498*

Scriven, M. Coryn, C., Evergreen S. (2009). Contemporary thinking about causation in evaluation: A dialogue with Tom Cook and Michael Scriven. ***American Journal of Evaluation***. [online]. Available from: *http://aje.sagepub.com/content/early/2009/12/04/1098214009354918.short* [Accessed January 12, 2015].

Shadish, Cook and Campbell (2002), Experimental and Quasi-experimental Designs for Generalized Causal Inference. Houghton Mifflin. *http://impact.cgiar.org/pdf/147.pdf*

Silva Carlos, N. (2010). 'Ex Post Facto Study.' In ***Encyclopedia of Research Design-Volume 1***, edited by N. Salkind, J.: 465-66. Thousands Oaks, CA: Sage Publications.

Silverman, D. (2013). Doing qualitative research: A practical handbook. London, SAGE Publications Limited.

Smith, A. (2013). Civic engagement in the digital age. Washington: Pew Research Centre Report. Available at: *http://www.pewinternet.org/files/old-media//Files/Reports/2013/PIP_CivicEngagementintheDigitalAge.pdf*

Steinberg, T. (2014). 'Civic Tech' has won the name-game. But what does it mean? ***mySociety***. [online]. Available from: *https://www.mysociety.org/2014/09/08/civic-tech-has-won-the-name-game-but-what-does-it-mean* [Accessed January 12, 2015].

Stern, E., Mayne, J., Stame, N., Forss, K., and Davies, R. (2012). 'Developing a Broader Range of Rigorous Designs and Methods for Impact Evaluations, Final Report, London: Department for International Development (Dfid).'

Subjectivity lexicon, University of Pittsburgh: *http://mpqa.cs.pitt.edu*

Tavistock Institute (2003), Evaluation/Impact Assessment Frameworks for EU policies on Asylum and Human Rights

The Engine Room. (2014). ***Measuring Impact on-the-go: A users' guide for monitoring tech and accountability programming***. [online]. Available from: *https://www.theengineroom.org/wp-content/uploads/engnroom_monitoringguide_finalmay14.pdf*

Toyama, K. (2015). ***Geek Heresy: Rescuing Social Change from the Cult of Technology***. PublicAffairs.

Trochim, W. M. (2006). The Research Methods Knowledge Base (2nd Edition). Available at: *www.socialresearchmethods.net/kb*

Tufekci, Z. and Wilson, C. (2012), Social Media and the Decision to Participate in Political Protest: Observations From Tahrir Square. Journal of Communication, 62: 363–379

UN Global Pulse. (2012). ***Big Data for Development: Challenges and Opportunities***. New York: United Nations. [online]. Available from: *http://unglobalpulse.org/sites/default/files/BigDataforDevelopment-UNGlobalPulseJune2012.pdf*

UNESCO. (2014). The first Civic Hackathon in Chile awards three projects. [online]. Available from: *http://www.unesco.org/new/en/santiago/press-room/single-new/news/primer_hackaton_civico_de_chile_premia_tres_proyectos/#.VLP1n3umlN-* [Accessed January 12, 2015].

Wagner D., Day B., James T., Kozma R., Miller J., Unwin T. (2005). ***Monitoring and evaluation of ICT in education initiatives: A handbook for developing countries***. M. Trucano, ed. InfoDev / World Bank.

Walker Hudson, L. (2013). Mobile phones in M&E: towards sustainable and democratic practice. In OnTrac #55, September 2013. INTRAC.

Wampler, B. (2012). Participatory Budgeting: Core principles and Key Impacts. ***Journal of Public Deliberation***, 8(2).

Warburton D., Dudding J., Sommer F., Walker P. (2001). Evaluating participatory, deliberative and co-operative ways of working. InterAct. *http://www.sharedpractice.org.uk/Downloads/Interact_Working_Paper.pdf*

Warburton, D., Wilson R., and Rainbow, E. (2006) Making a Difference: A guide to Evaluating public Participation in central government, INVOLVE. Available from: *http://www.involve.org.uk/wp-content/uploads/2011/03/Making-a-Difference-.pdf*

Wiedemann, Gregor (2013). Opening up to Big Data: Computer-Assisted Analysis of Textual Data in Social Sciences. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 14(2), Art. 13. http://nbn-resolving.de/urn:nbn:de:0114-fqs1302231.

World Bank (2007). Social Accountability Sourcebook. World Bank. *http://www.worldbank.org/socialaccountability_sourcebook/*

World Bank (2011), Writing terms of reference for an evaluation: a how-to guide. Independent Evaluation Group

World Bank (2013b). Open Data Essentials. [online]. Available from: *http://data.worldbank.org/about/open-government-data-toolkit/knowledge-repository*

World Bank (2013c). World Bank Group President Jim Yong Kim's Speech at Annual Meetings Plenary: One Group, Two Goals: Our Future Path. [online]. Available from: *http://www.worldbank.org/en/news/speech/2013/10/11/world-bank-group-president-jim-yong-kim-speech-annual-meetings-plenary* [Accessed January 9, 2015].

World Bank (2014a). Strategic Framework for Mainstreaming Citizen Engagement in World Bank Group Operations. Available from: https://consultations.worldbank.org/Data/hub/files/consultation-template/engaging-citizens-improved-resultsopenconsultationtemplate/materials/finalstrategicframeworkforce.pdf

World Bank (2014b). Piloting Citizen Engagement in Projects - A Guidance Note for World Bank Staff Working in the Middle East and North Africa Region. World Bank.

Yin, R. K. (2003). ***Applications of Case Study Research Design***. Thousand Oaks, CA: Sage Publications.

Yu, P, de Courten, M., Pan, E., Galea, G., Pryor, J. (2009). The development and evaluation of a PDA-based method for public health surveillance data collection in developing countries. Int J Med Inform 78(8):532-42

WORLD BANK GROUP
Governance
DIGITAL ENGAGEMENT
Evaluation Team | DEET