The challenges of implementing generic web and mobile apps for managing and monitoring community scorecards and social audits: lessons from Mozambique

Author - Mr. Robert Worthington
Organisation - Kwantu
Presenter - Mr. Robert Worthington
Email addreess – rob@kwantu.net
Phone number - 021 531 9120

Introduction

Social accountability approaches to effective public service delivery in Mozambique are very new features of the country's development landscape. However, they are beginning to be recognised by citizens and governments as a valid means for improving the efficient delivery of services and ensuring transparency in governance.

The first organized attempts to promote social accountability in Mozambique were focused on citizen engagement around the debt relief mechanism. This was led by a new coalition called *Grupo Moçambicano da Dívida* (Mozambique Debt Group) that was established in 1996. It aimed to promote citizen engagement to hold the Government to account on the debt relief process to the people of Mozambique.

The number and scale of social accountability projects have grown since then. Tools like community dialogues, social audits, citizen report cards and community scorecards are now used by a wide range of actors. These are now being applied across different sectors (health, agriculture and education for example), at different levels of government (from facility level to municipality, district and province to national level) and at larger scales (in over 100 facilities a year for some programmes).

Challenges to scaling social accountability projects

As the scale of these interventions grows, new challenges emerge. Implementing large numbers of social accountability activities is time-consuming and requires careful planning, coordination and management. These challenges are heightened as several organisations begin working in the same districts and municipalities, using the same social accountability tools. Without coordination they risk targeting the same service delivery outlets (schools, clinics, water outlets etc.) and potentially the same citizen groups. Scheduling and tracking the meetings and workshops required to facilitate citizen engagement is also more challenging as the scale grows.

Monitoring and evaluating these activities presents similar challenges. Citizen engagement activities typically generate large amounts of qualitative or loosely structured data. Collecting, aggregating and analyzing this data is very time-consuming. Without care and attention there is also a risk of significant data quality issues. For example, double counting the same activities, data transcription or aggregation errors and invalid responses to questions.

Donors also face challenges to justify the investment in social accountability. They require data to assess the effectiveness and efficiency of the interventions. This data should be comparable to enable donors to benchmark similar projects by time, place or implementer. This in turn would guide decision making to increase value for money.

Can standardized apps address these challenges?

Technology has helped private sector organisations to tackle similar challenges. There are many lessons to learn in particular from the franchising approach. The franchise owner commonly offers franchisees a set of standardized business processes that have been demonstrated to work. These are accompanied by technology to help manage and monitor activities required to implement these business processes. UPS shipping outlets in other shops are an example of franchises. A visit to several different UPS

shipping outlets (each of which are run by different people) will illustrate how the UPS standardised business process and technology ensures the delivery of consistent and high quality activities.

However, processes and activities in the social accountability sector are often complex and un-predictable. Different implementers have different ways of working and different information needs. We cannot assume that each implementer could use a generic app that models a standardized social accountability (business) process. The level of standardization imposed may be a disadvantage.

The traditional pattern has been for implementers to design systems (incorporating technology or otherwise) independently and in isolation. Data on indicators is shared to donors in the form of reports. However, more detailed data needed to address challenges of coordination (who is doing what, where and when?) and value for money (how effective and efficient are these activities?) is not recorded in a standardized way and hence is not easy to share and aggregate.

This paper explores the tensions between standardization and individual requirements. It seeks to test this assumption through research with three different implementers, each of which are using the same social accountability tool – community scorecards.

Research context

Kwantu partnered with two NGOs and one consortium implementing social accountability programmes in Mozambique to explore how ICT can help address these management and monitoring challenges. Each of these partners are using community scorecards (CSC) The table below provides an overview of each partner or programme and the type and scale of work that they implement.

Programme	Type of work	Scale of work
CEP – Citizen Engagement Programme Implemented by a consortium of 7 organisations (Save the children, CESC, N'weti, Kwantu, IDS, OPM and Cowi)	Improving the quality of health and educations services through increased citizen influence on policy making and service management Support citizens and communities to actively monitor service delivery and to engage with services providers to implement action plans that address the identified problems at the facility level. Scale up scale up issues that need responses at higher level through evidence based advocacy efforts.	4 provinces: Gaza (Chokwe and Bilene-Macia), Manica (NManica Sussundenga), Nampula (Monginqual, Liupo and Murrupula) and Zambezia (Mocuba and Lugela): covering more than 100 facilities each year
Empowering citizens in health rights and engaging education providers with service users Implemented by CESC	Informing and empowering citizens about their rights in the area of health and engage right-holders, in particular children, and members of school councils in collecting evidence that will contribute to the improvement of decentralized planning and quality education	4 provinces: Gaza, Cabo Delgado Zambézia and Tete.
Tua Cena, Tua vida, tua escolha Implemented by N'weti	Improve engagement between youth, civil society organisations and service providers to improve the quality of sexual and reproductive health, family planning and sexual transmitted infections including HIV/AIDS service delivery for adolescents and youth between 15 and 24 years of age	4 provinces: Gaza, Nampula and Maputo and Maputo City

The partnerships focused on developing web and mobile apps that assist with both the management and monitoring of social accountability interventions. Some partners have been using these apps for over a year now, others are just starting.

However, the process of analysing and documenting the requirements for each partner provided a unique opportunity: the chance to compare how different organisations implement community score cards. Through this experience the authors have collected data to try and answer the following research question:

Is it possible to implement one standardised App that could be easily adapted to meet each partner's specific requirements?

Alongside this central question, we also explored the following sub-questions:

- How consistent is the implementation of community scorecards across different partners?
- What are the similarities and differences between the data collection tools used by the different partners?
- Is it feasible to agree on a minimum core data set for each data collection tool?
- What additional data would be needed to facilitate joint activity planning, benchmarking and value for money comparisons across partners?

This paper presents findings related to these research questions.

Methods

Our research was in two parts. First, assisting our partners to analyse and document their requirements. This was necessary to implement systems to help manage and monitor their activities. It also generated the data needed for a comparative analysis. The second part of the our research method was a comparitive analysis of each partner's documentation to answer the research questions.

Analysing and documenting partner requirements

Since January 2014 Kwantu has assisted the partners above to implement systems to monitor social accountability activities. This technical assistance took place in four phases.

First, process mapping workshops to analyse and document the activities. Process mapping is an approach developed by Kwantu to help document programme activities in a standardised way. The approach is designed to identiy and address data quality issues. It looks at the holistic information needs of the programme, including both management and monitoring and evaluation requirements. Where possible data collection tools are designed to be completed as activities are carried out.

The process mapping workshop starts by breaking the programme activities down into a series of steps. These steps are linked to key points where management information is required or data quality reviews are needed. The following areas are agreed and documented for each step:

- Pre-conditions necessary before the step can begin
- The role player responsible for the step
- Which data collection forms needed to collect data for management or monitoring needs
- Tasks which must be completed during the step
- Documentation and guidance for field staff
- Which steps the process may move too once this step has been completed

The activity data collection forms linked to the process steps were then documented in detail. This included guidance and help text, field level validation rules and taxonomies to code data.

The output from this workshop was a detailed process manual covering the programme activities of each partner. This provided a means to compare the activities across partners and assess similarities and differences.

The second phase of the assistance focused on configuration of a monitoring and evaluation system for each partner. The requirements documented in the process manual were used to:

- **Configure a workflow engine** - This followed the steps defined in the process manual, specifying rules for which user role can take actions and what actions are required or optional during that step. This is designed to ensure that good practice related to data and implementation quality is followed.
- **Create web and mobile based forms** – This defines the data collection tools required. These tools are purposefully designed to collect data at the activity level to avoid any data quality issues associated with data aggregation.
- **Define taxonomies** – Taxonomies are closed lists of keywords that are used to provide closed responses to specific questions. They are also used to code the data, adding categories needed for analysis.
- **Configure reports** – Reports define a query to extract data from data collection forms. They also define which fields from the form are included and what labels are used. Taxonomy filters may also be applied to facilitate analysis. Once configured users can run reports to extract aggregate data in real-time for analysis.

The workflow, forms, taxonomies and reports were configured as a Community Scorecard App in BetterData, Kwantu's configurable performance management platform. This enables it to be adopted and used by any organisation with permission to use it. A separate App was created for each partner to accommodate their specific requirements.

The third phase focused on extensive user testing and feedback to review and revise the Apps where needed. Since several partners had previously used notebooks and paper forms to collect data, this was a significant transition. In some cases several rounds of changes to data collection tools, as partners became more familiar with the benefits and limitations of electronic data collection forms. Once the Apps were finalised user training and administrator training was provided.

The fourth and final phase began once partners began using the Apps to actively collect data. Monthly management reports were generated to track each community scorecard against pre-agreed deadlines for completing specific implementation steps. This helped to establish where activities were on track and where they were delayed. It also highlighted where data expected to be entered for that step was missing. These tools were designed to help address the challenges of managing large scale community scorecard activities.

Comparative analysis of partner requirements

The requirements for each partner were documented as a process manual. This is a standardised way of documenting requirements for performance management app. It includes a detailed description of the following areas:

- **Process** – the process of implementing, reviewing and evaluating activities is documented as a series of steps
- **Data** – the data needed for planning, management and montioring and evaluation are documented as a series of data collection forms. These are linked to the steps in the process were they are collected
- **Taxonomies** – the vocabularies used to code the data for analysis are documented

We compared the process documentation to understand which steps each partner followed when they implement community scorecard activities. This helped us assess the consistency of implementation across different partners. A comparison table is included in the results section.

We compared the data collection forms documented for each partner to assess which and how many forms they each used. For similar forms we also compared the fields on each form to assess the similarity and differences. We use this analysis to construct a hybrid form that shows core fields used by all partners.

Finally, we compared taxonomies developed by each partner. This helped develop an understanding of the different criteria used to code data for analysis and commonalities among those criteria.

Methodological limitations

This research should not be considered representative of social accountability across Mozambique. It focuses on one tool (community scorecards) and three partners only. This is not a representative sample. The partners were also not selected randomly and have worked together before. This introduces the potential for selection bias. It provides data that can be used to make preliminary conclusions and included in future comparisons with other data.

Results

The comparative analysis covered implementaton steps, data collection forms and taxonomies. These results are presented below.

Comparison of implementation steps

The table below provides an overview of the implementation steps followed by each partner. In some cases sub-steps have been hidden to simplify the presentation of this data. For example, all partners include a sub-step during the monitoring step to track sub-steps related to monitoring cycles.

Step	CEP	N'weti	CESC
Planning	Yes	Yes	Yes
Preparation / community mobilisation	Yes	Yes	Yes
Review / sign-off	Yes	No	Yes
Scoring by community and providers	Yes	Yes	Yes
Review / sign-off	Yes	No	No
Generate a single list of issues	No	Yes	No
Legitimise issues with community	No	Yes	No
Dialogue with the facility	No	No	Yes
Review / sign-off	No	No	Yes
Preparation for interface meeting	No	Yes	No
Interface meeting	Yes	Yes	No
Review / sign-off	Yes	No	No
Institutionalisation	No	Yes	No
Monitoring the joint action plan	Yes	Yes	Yes
Review / sign-off	No	No	No
Completed	Yes	Yes	Yes

While the names may differ, the implementation steps followed by the partners are broadly the same. In some cases partners have choosen to include formal sign-off or review steps. At this point the app will re-assign the project to a different person to review data that has been collected so far. This relates to internal procedures and has no bearing on the comparability of the data collected.

We found that the overall activities of partners were broadly similar and comparable. In some cases their workflow includes additional steps. This relates to internal requirements for a greater focus on tracking specific parts of the process. It does not affect the comparability of the data.

We conclude that the processes followed by each partner are comparable. However, to accommodate individual requirements a generic app would need to offer the option to modify the core steps followed, re-naming them, adding additional steps and defining review steps as needed to match internal procedures.

Comparison of data collection forms

The following table provides an overview of the types of data collection forms used by each partner.

Form	Explanation	CEP	N'weti	CESC
Project registration form	Registers a planned scorecard project. Includes data on it's location and implementation schedule	Yes	Yes	Yes
Facility survey form	Survey form to collect basic infrastructure data on the facility where the scorecard is being implemented	Yes	Yes	Yes
Group registration form	Registers each group (community or provider) that is taking part in the scorecard for a facility	Yes	Yes	Yes
Health scorecard form	Registers the scorecard for each group for a specific facility. This version includes standardised satisfaction questions related to the health sector	Yes	Yes	Yes
Education scorecard form	Registers the scorecard for each group for a specific facility. This version includes standardised satisfaction questions related to the education sector	Yes	No	Yes
Municipality scorecard from	Registers the scorecard for each group for a specific facility. This version includes standardised satisfaction questions related to a municipality	No	No	Yes
Action plan form	Registers a joint action plan for a specific facility	Yes	Yes	Yes
Meeting form	Records a meeting that took place in relation to a specific scorecard project	Yes	Yes	Yes
Contact form	Records a key contact related to a specific facility	Yes	Yes	Yes
Outcome form	Records an un-expected outcome related to a specific scorecard project using the outcome mapping tool	Yes	Yes	Yes
Action plan monitoring form	Records progress made towards resolving actions on the joint action plan	No	No	Yes

The types of data collection forms used by partners are very similar. However, some key individual requirements were identified. A generic app must be able to specify a list of standardised questions that the partner needs to use to track satisfaction. These will vary from sector to sector. Each partner may also use different questions for the same sector. One partner also choose to include an additional form to provide a stronger way of monitoring progress against actions on the joint action plan. Other partners monitor this by editing the action plan form. A generic app would need to mandate the use of core forms (needed to collect and share data for planning, coordination and value for money). It should also provide a list of optional forms that enable them to collect more data should they wish to.

Comparison of taxonomies

The following table provides an overview of the taxonomies used by each partner to code data.

Taxonomy	CEP	N'weti	CESC

Province	Yes	Yes	Yes
District	Yes	Yes	Yes
Facilities	Yes	Yes	Yes
Group type	Yes	Yes	Yes
Priority	Yes	Yes	Yes
Level	Yes	Yes	Yes
Common issues	Yes	Yes	Yes
Scorecard type	Yes	Yes	Yes
Donor	No	Yes	Yes
Year	No	Yes	Yes
Programme / Project	No	Yes	Yes
Partners	No	No	Yes

The types of taxonomies used by each partner are comparable. Some partners have added additional taxonomies to link scorecard projects to specific donors, partners and projects or programmes. These are internal criteria that do not affect comparibility. However, they provide data that the partners need. A taxonomy of years was added by two partners to track which year each scorecard project is being implemented in. This is an important enhancement that CEP will also include when they update their taxonomies next.

Most of the common taxonomies used are easy to standardise. Provinces, districts and facility names are all managed by the Mozambique government. The taxonomies used by the partners in this study already follow these standards making comparability simple.

Group type and common issues have been developed individually and would require a facilitated process to standardise. Group type refers to the types of groups that are recruited for each community scorecard project. These typically vary by sector and include around five or six different types of groups (for example, pregnant mothers, teachers, young men, young women). Standardising this list of group types would enable comparative analysis across the data of all partners.

The common issues taxonomy is more complex. This tracks a list of common issues that have been raised during scoring and in action plans. It is used to code the issues raised by groups (where they relate to a common issue) to enable frequency analysis. Developing this taxonomy takes time, as it requires an analysis of existing issues raised by groups. We do not have sufficient data at this time to assess the feasibility of developing a shared taxonomy of common issues.

Standard data needed to coordinate joint planning

Coordinating planning of community scorecard projects across different implementers would help increase efficient use of resources. It would also facilitate linkages between different projects. To facilitate this a generic app would need to collect the following data for each community scorecard project:

- Name of facility being monited
- GPS coordinates for facility (from which we can infer the administrative post, district and province)
- Planned start date and end date
- Implementer name

The project registration form used by each partner already collects these details (with the exception of the GPS coordinates, which are linked to the taxonomy of facilities). All that is required is a mechanism to share and aggregate this data and permission from each partner.

Standard data needed to compare value for money

Value for money is a complex concept. There are several frameworks and approaches to measuring value for money. The authors are not aware of any that have been applied to community scorecards. To assess the data needed to measure value for money in this context, we use the framework proposed by the UK National Audit Office. This has three criteria:

- **Economy**: minimising the cost of resources used or required (inputs) – **spending less**;
- **Efficiency**: the relationship between the output from goods or services and the resources to produce them – **spending well**; and
- **Effectiveness**: the relationship between the intended and actual results of public spending (outcomes) – **spending wisely**.

Using this framework to compare value for money will require data on:

- Inputs (ie the project budget)
- Outputs (ie number and type of different activities produced by the project)
- Outcomes (ie the intended and un-intended results of the project)

One partner (N'weti) already uses the project registration form to record the budget for each community scorecard. This is currently used for internal planning processes. Adding this as a standard field to this form would cover this requirement in a basic way. More sophisticated analysis would be possible if the budget were broken down in more detail.

All partners already utilise the same type of forms to track outputs. This includes data on the output of activities, such as number of meetings held, number of groups registered, number of scorecards produced, number of action plans produced. While the content of these forms are largely similar, there are a number of individual variations to include data required by each partner. This would not affect a comparison of the number of each output per project.

All partners collect outcome data in similar ways. They ask satisfaction questions as part of the scoring process and collect information on un-expected outcomes as they arise. However, the type of satisfaction questions are not always comparable between partners. Nor are these questions applied again at the end of the project to assess a change in satisfaction. Due to the longer time frames needed to assess the outcomes of social accountability, these changes are instead tracked using external evaluations. While a generic app could collect information on outcomes in a standardised way, external evaluations would still be needed to measure effectiveness.

Conclusions

We conclude that it is technically feasible to develop a generic app that different partners could use to manage and monitor the implementation of community scorecard projects. The technical changes required to the existing apps appear to be minimal.

Following is a discussion of the additional work needed to develop a generic app and the implications that this would have for implementers and donors using this app.

Flexibility to adapt aspects of the app to fit requirements of implementers

While key aspects of the app could be standardized, the implementers assessed in this study would need the flexibility to adapt the following specific areas to meet their individual requirements:

- Adjusting the workflow to modify the core steps followed, re-naming them, adding additional steps and defining review steps as needed to match internal procedures
- Selecting from a library of optional forms as to which they would like to use
- Adding additional questions to the optional or mandatory forms

- Creating and modifying taxonomies used to manage partner specific coding criteria, like donors, projects and partners
- Managing a list of standard satisfaction questions used by that partner

Most of these options are already configurable. However, they would ideally be presented in a set-up wizard that prompts each partner as to which areas they may wish to adjust to fit their own context. This would enable each partner to quickly configure a generic app to meet their own requirements.

It's likely that implementers outside of this sample would need to adapt the same areas. However, we cannot extrapolate from these results as to what other flexibility would be needed for other implementers.

Leadership to facilitate development of common information standards

Tackling the challenges of more coordinated planning and better data on value for money is depedendent on the ability to share and aggregate data across multiple implementers. While a standardised app will help to achieve this, equally important are common information standards. In the analysis of the results, we identify the following areas where national level information standards are required:

Agreeing on standardised taxonomies

The following taxonomies should ideally be developed as a national standard:

- Types of groups and definitions for these types (while these are sector specific, there may be overlaps between the sectors – for example, young men and women)
- Common issues raised by groups (developing at least general categories of issues would allow implementers to code the issues raised against these. This would enable an interesting analysis of the types of issues raised)
- Satisfaction questions (this is an important area to standardise as it provides one way of measuring the impact of community scorecard projects on satisfaction)

While not essential, this would facilitate joint analysis of data across different implementers.

Agreeing on standardised data to collect

Depending on the type of value for money analysis required, agreement may be needed on the budget categories for community scorecard projects. These could be very simple categories like staff time, overheads and expenses.

Govenrment, donors and multi-lateral agencies (such as the United Nations) are best placed to lead the development of these types of information standards.

Mechanism to facilitate data exchange

A final pre-condition for addressing these challenges is the ability to share and aggregate the data collected by implementers. The existing app has reporting tools to export the data into XLS or other widely used file formats. However, aggregating data from multiple implementers would be time-consuming and prone to human error.

This would be better tackled by a mechanism that can automatically share and aggregate data. Ideally such a mechanism should include:

- Governance controls that enable implementers to determine which data is shared and with whome
- Automated sharing and aggregation of data as and when an Internet connection is available
- Role based user accounts that check governance settings to see which data a user may access

- Analysis and reporting tools that enable users to query and analyse the data

Kwantu have begun documenting requirements for this kind of mechanism and plan to develop a prototype in 2016.

Discussion of broader implications

This research demonstrates that – in the narrow context of three partners implementing community scorecard projects in Mozambique – it is feasible to develop a standardised web and mobile app to help manage and monitor this activity. With additional work, it is feasible to establish the information standards and mechanism to share and aggregate this data. This has several broader implications.

Implementers

Implementing a standardised app to manage and monitor community scorecard projects offers the potential for economies of scale. Instead of each implementer creating bespoke systems, they would have the option to use one shared system. This would significantly lower the set-up and running costs.

A shared solution also provides the option to develop and incorporate guidance on good practice. This could take the form of a shared knowledge base of resources (for example documents, videos and frequently asked questions). Implementing partners would benefit from sharing knowledge around a specific activity.

Implementers would benefit from the ability to see which activities are planned, where and when. This enables better planning for new activities and better coordination of existing ones. Finally, implementers would benefit from the ability to automate reporting to donors using the data sharing mechanism.

Donors and Government

Donors and government would benefit in two key ways. First, the ability to share data easily across multiple implementers bring significant efficiencies to the accountability process, as manually prepared reports on activities would not be needed. Grantee reporting could instead focus on lessons learned, knowing that data is shared automatically with selected stakeholders. This would in turn give donors and governments access to larger and more detailed datasets that cover aggregated activity level data from multiple implementers.

As explored in the results section, the shared dataset would facilitate a range of comparative analysis. This includes more detailed value for money analysis, such as benchmarking inputs and outputs across projects and over time. This data would help guide policy decisions on how to best allocate funding for social accountability work.

Finally, shared planning and project data would facilitate joint approaches to impact evaluation. Knowing which projects are active where and when makes it easier to coordinate joint impact evaluation studies. If several donors are funding the same work, they have the option to share evaluation costs, knowing that the underlying approach used by the implementers is comparable.

Implications for community scorecard projects in other countries

We cannot draw strong conclusions for the feasibility to replicate this approach to other countries. Having carried out similar analysis with other partners in other countries, we are confident that the requirements could be clearly identified and documented. However, we cannot extrapolate from this data to determine the suitability of this app for projects in other countries. Other types of flexibility not covered in this paper may be required.

Implications for other types of projects and activities:

While this research focused on a specific type of activity, it could have implications for other activities, tools and approaches. We have carried out similar analysis for other types of projects and activities. This includes social audits, social protection, agricultural extension visits and community security activities. We lack sufficient data to determine the level of standardization feasible across these activities. Based on this analysis we assume it is likely that there are a range of development activities where there is sufficient standardization to implement a generic app for that activity. Further research is needed to establish which types of activities are suitable for this approach.