Managing Data Quality with Negative Reporting

Kevin Zeigler
Special Projects Manager

ICT4D CONFERENCE
NAIROBI KENYA 2016
Today I’ll show you how we transformed our approach to data quality and keep our data systems clean.

1. Overview of TechnoServe and our organization.
2. Past data quality challenges;
3. Overview of our M&E database setup;
4. Using automated “negative” reporting for data cleaning;
5. Successes, challenges, and additional tips.
TechnoServe works with enterprising people in the developing world to build competitive farms, businesses, and industries.
Our Latin America region used a variety of M&E systems which were expensive and time consuming to create.
Our old data lifecycle required 11 steps to be completed before any data could be cleansed in the database.
Disparate M&E systems and lack of guidance created data quality problems that each team had to tackle on their own.

• Only skilled project/M&E managers and IT oriented resources could run queries to identify problematic data.
• Related records often with no counterpart (orphaned) and unusable in reporting.
• Participant tracking difficult or IDs mismatched.
• Training session dates often incorrect.
• No learning from mistakes as it was often difficult to tell when the mistake was made and by whom.
Our regional approach to M&E databases allows projects to have a robust system created for them in less than three weeks.
Our regional M&E database allows us to automate our reports and identify problematic data for the entire region at once.
Your reports can be customized to meet the needs of each project individually or managed at a regional level.

- Filters set in the report can help you to identify errors in fields such as dates, timestamps, Participant ID’s, and Household IDs.
- Always include the unique ID in your reports so that you can easily identify records when you correct them in the database.
Scheduling reports allows you to easily remember to clean your database on a periodic basis (*hint*: frequently!).

In Salesforce, reports can be sent to any user, creating a flexible data cleaning utility that you can customize to your needs. There are also apps in the AppExchange that allow you to send reports to non-Salesforce users.
Using templates (we use Google Sheets) can help you quickly fix issues when uploading corrections to your M&E database.
Keeping track of errors over time helps to facilitate field staff training and to identify common data entry issues.

- Each time an error is detected by our automated reports, we correct the error in Salesforce and log the error in our tracker sheet.
- Tracker sheet enables us to identify field staff with repeated data quality issues and design training interventions to help field staff better understand how to use the technology.
When confronting data quality issues, it’s usually always best practice to start at the data collection source.

1. Identify Data Quality Issue
2. Check for Solution at the Field Level
3. Check for Solution at the Database Level
4. Monitor Data Quality Post-Change
Since starting automated reporting and data cleaning on a regional level, we’ve seen a significant improvement in our overall data quality.

- Time spent correcting database errors is reduced.
- Overall accuracy of reports has increased.
- Can spot data issue trends much quicker across similar projects and teams in the region, allowing us to provide additional targeted training to field staff.
- Teams have more faith in data and are more willing to use reports generated by the system.
Remember to start at the source: survey design, strategy, and field staff training are just as important!

- Training field staff properly is an important part of mitigating data quality issues before they happen.

- Designing surveys with less room for error:
  - Case management / hierarchy drill down features;
  - Cascading selects;
  - Single and multi-select questions;
  - Barcodes and QR codes for IDs.

- User centric design for surveys – get everyone involved in helping you to create surveys!

- Use the right tools, mapping your surveys directly to your M&E system can have a huge effect on the quality of your data.
Data quality is everyone’s responsibility, get your teams involved early and you’ll reap the benefits!

• Spend time with training your project teams, field staff, and everyone involved in the data collection, analysis, and reporting.
• Don’t wait to start cleaning your data! Cleaning on a regular basis is much more manageable.
• Use the power of automated reporting to help catch data quality issues for you – it’s like auto-pilot for your reporting.
• A properly structured database and good survey design practices can help your teams focus on their data quality:
  – Case management / hierarchy drill downs;
  – Role hierarchies;
  – Permissions;
  – List views, dashboards, and reports.
• Data quality is everyone’s responsibility, from the field staff to the project manager.
Thank you for your time today! Please contact me with any questions and/or comments.

Kevin Zeigler
Special Projects Manager, TechnoServe
kzeigler@tns.org
kevin.zeigler [Skype]