Food and Agriculture Organization of the United Nations

e-Agriculture Promising Practice

Open Data Kit: a new field data collection tool for bean breeders and researchers

Improving the common bean breeding process by providing accurate field trial data

Key facts

- **Location:** Uganda
- **ICT used:** Open Data Kit
- **Area of work:** bean value chain
- **Target group**: bean breeders, field technicians and researchers
- **Stakeholders:** bean value chain actors, bean farmers, researchers, CIAT, Uganda National Bean Program
- **Timeframe:** 2016-2017

The International Center for Tropical Agriculture (CIAT), through its Pan Africa Bean Research Alliance (PABRA) program, works in collaboration with country national programs in over 29 countries across the African continent to improve the varieties of beans available to the farmers through research. This promising practice factsheet focuses on the introduction of the Open Data Kit (ODK) in Uganda.

ODK is one of the ICT platforms that have been taken up by the PABRA program. It is believed that ODK can revolutionize the way data is managed by ensuring the seamless collection, storage and analysis of data. With five key themes, (breeding, markets, nutrition, seed systems and capacity building) the operationalization of the ODK platform in Uganda in the last two years focused on working with the National Bean Program in using ODK to manage trial data aimed at releasing climate smart bean varieties that are multiple stress resistant and drought tolerant.

Open Data Kit (ODK)

Open Data Kit is a free and open source set of tools, which help organizations author, field and manage mobile data collection solutions. ODK provides an out-of-the-box solution for users to build a data collection form or survey, to collect data on a mobile device and send it to a server, and to aggregate the collected data on a server and extract it in useful formats.

Context and problem addressed

The humble common bean has great potential to improve the health, food security and incomes of the most marginalized communities in Africa, not to mention making a significant contribution to national economies. As a result, beans are becoming an ever more important crop. The common bean is rapidly evolving from subsistence to a market-oriented cash crop. It is the most important, widely grown and consumed grain legume in Eastern, Central and Southern Africa, where about 6.3 million hectares of land is used to grow beans every year.

The International Center for Tropical Agriculture (CIAT), through its Pan Africa Bean Research Alliance (PABRA) program works in collaboration with country national programs in over 29 countries across the African continent to improve the variety of beans available to the farmers through research.

ODK is one of the ICT solutions that have been adopted by PABRA which is believed has the potential to revolutionize the way data is managed by ensuring the seamless collection, storage and analysis of data. With five key themes, (breeding, markets, nutrition, seed systems and capacity building) the operationalization of the ODK in Uganda focused on working with the National Bean Program in using ODK to manage trial data aimed at releasing climate smart bean varieties that are stress resistant and drought tolerant.

Poor data quality can have substantial scientific, social and economic impacts on an organization. Although the National Bean Program in Uganda is improving data quality with practical approaches and tools, there is still a need to channel the efforts to focus narrowly on accuracy so as to form a strong foundation for information in decision-making. It is a common practice for bean researchers and field technicians working with the National Program to use paper based tools for data collection and monitoring. This often results in several problems such as delays in curating data, wrong data entries, delays in data validation and problems with data integrity.

The use of a paper based field approach for monitoring breeding trials is not only cumbersome, but also increases the time needed to transfer the data to computer software (Excel, SPSS etc.) and the chances of error are quite high, especially when capturing data because validation at data capture is reliant on the vigilance of the data entrant.

Bean field experiments are important in bean breeding. At advanced breeding stages, field experiments are carried out in multiple locations and correspond to multiple environments. The purpose is to establish the performance of the bean type under investigation when grown in different environments. A bean type growing in most or all environments in a country is said to have superior qualities and is a candidate for consideration by the national release selection using committee and for national release and promotion to farmers. The importance of field

experiments cannot be underestimated. Monitoring these distant field sites using paper tools makes the collection and consolidation of data not only cumbersome but also time consuming. The data centralization architecture of the ODK platform allows the aggregation of distributed data sources and is time saving.

ODK provides an opportunity to improve and hasten the research process at trial data collection by ensuring the quality of data is good, stored in a centralized location and is validated for accuracy. ODK is an open-source, modular toolkit that enables organizations to build application-specific data/information services for use in resource-constrained environments, it further has the advantage of being able to work with low end android phones/tablets, easily customized with limited need of programming experience, ability to store in a cloud platform that is relatively secure and easily accessible by those that have access details, and has the ability to capture diversity of data types among other advantages.

The challenges data that bean researchers face, especially at collection stage emanating from data management practices, are slowly being addressed by exploring various ICT options available. The ODK tool currently provides a better way for the National Bean Program in Uganda to improve and streamline field data collection processes, both on farmers' fields and on experimentation plots located on bean research stations.

Role of the different stakeholders

The target groups are the bean breeders and field technicians involved in the breeding process. The use of ODK has been narrowed to the trial fields in Nakasongola, Arua, Lira, Sorori and Jinja districts, where the National Bean Program carries out drought field experiments.

Bean farmers who own land used for off station trials are the secondary stakeholders who benefit from the use of the ODK. Though they may not directly interact with the ODK tool, they play a critical role during the evaluation of the trials by providing the data needed to be fed into ODK.

The International Center for Tropical Agriculture (CIAT), through its Pan Africa Bean Research Alliance (PABRA) program works in collaboration with the Ugandan National Program. CIAT specifically provides the technical support needed to develop the ODK tool.

Similar to any new invention, the rate at which it is appreciated and adopted is dictated by various parameters, such as availability of sufficient smart phones and tablets to use and awareness of the new technology. During 2016-2017, the National Bean Program in Uganda has used ODK regularly. With every use, attempts are made to improve the ODK system by addressing the major hurdles that hampered it's use, such as the issue of light interference during data collection.

Jane Mukabalanga is a field assistant from the National Crops Resources Research Institute (NaCRRI). She is using a tablet with the Open Data Kit to monitor trials in a field from a farmer in the Arua district.

The development of the ICT used: Open Data Kit

The use of the ODK by the National Bean Program is based on a participatory consultative approach involving the bean researchers, field technicians, and the CIAT-PABRA database officer. To understand how the ODK was operationalized, CIAT categorized the development stage into two distinct categories: the definition of variables for monitoring trials and the ODK platform.

Definition of variables for monitoring trials

The steps in defining these variables were as follows:

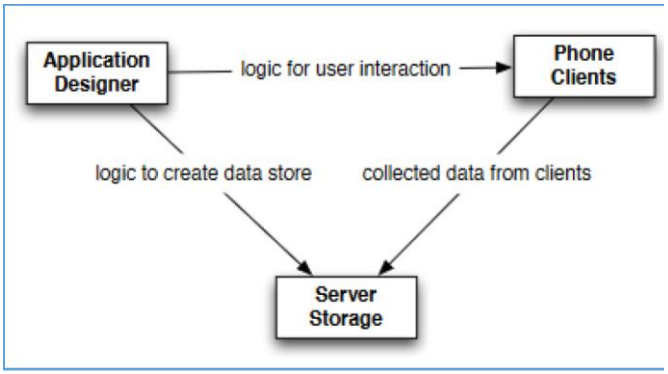
- Assessing the existing data collection infrastructure, data flows, interests and demands;
- Making sense of prior experience of researchers and technicians in digitized data collection;
- Identifying variables of interest for tracking change in farmers owned bean trials;
- Joint review of variables by a broader group of technicians and scientists to arrive at a set of variables for tracking change in farmers owned bean trials;
- Structuring variables in an Excel sheet and populating with dummy data. The dummy data helped to test the Excel sheet for completeness and accuracy;
- Developing a final Excel template for monitoring the variables and circulating to target users.

The ODK platform

The technical support to develop the ODK platform was provided by the CIAT-PABRA database officer as the ODK technical support officer. This was a participatory process involving the researchers whose role was to guide the ODK technical officer, especially in creating a common understanding of the variables and set the validation rules for the type of data to be collected.

- Joint review of the final variables involving the researchers and the ODK technical officer, to establish a common understanding on variables;
- Configuring of the template into the ODK structure by the ODK officer;
- Configuring the smart phones and tablets with ODK;
- Hands on training workshop for researchers and technicians;
- Pre-testing of the ODK solution on site at the bean research station in Namulonge to give the researchers a real field feel;
- Improvements were made to the pre-test version and a second test on site at the same station was conducted;
- Advancing the ODK to farmers' bean fields in trial sites at multiple locations in the districts;
- Transmitting data collected using the ODK tool to the cloud server at the end of every day's field activity where it was accessed by the authorized researchers and further analyzed;
- Making the backup of the collected data available using an export function to PDF, Excel or CSV.

The figure below summarizes the architecture of the ODK solution:



The application designer is the environment where the variables of interest for tracking change in bean field trials are defined into a phone readable format. The stage requires technical understanding of the platform syntax.

The phone client is the ODK version that runs on the smart phone or tablet. This is the ODK section where the bean technicians interact with the collected bean trial data.

The server storage is a cloud platform where all the collected bean field data is saved. Access to this platform requires a username and a password in order to view the collected data.

Impact

The use of ODK for monitoring farmers owned bean field trials is in its second year of implementation. Five researchers have used it. Thirty farmers in six districts have had their bean trial farms monitored using the ODK. The accurate impact of ODK may not be easily ascertained but the accurate, clean and timely data collection process is evident, and is characterized by reduced time on data cleaning and analysis.

The ability of ODK to capture varying data types such as GPS and images provides options for researchers and an opportunity to widen the scope of experimentation field data available. This will improve the quality of research outputs in the long run.

Globally, the PRABA program had a direct impact on the bean value chain by introducing 550 new bean varieties in the member countries with CIAT assistance, of which 20 in Uganda. In the context of climate change, scientists have identified beans that can beat the heat and perform well under at least 3 degrees Celsius higher average temperatures. Climbing beans, yielding three times more that the familiar bush type, provide an especially eco-efficient solution for densely populated land-scarce places and specially bred, high-iron beans have reduced iron deficiency and anemia in young women.

Innovation and success factors

ODK is one of the leading data collection solutions available at this moment and has been deployed by a wide variety of organizations in dozens of countries around the world. Its introduction to the Ugandan National Bean Breeding program has improved data collection in farmer owned bean trial farms. It is a fundamental shift in the way field based trial data is managed when fully embraced.

The successful implementation of ODK for bean trial data collection in Uganda can be attributed to a number of factors, including:

- Availability of institutional phones and tablets for use by field researchers and technicians;
- Successful application of ODK in a related unit of the bean program for socio-economic surveys prior to the introduction in bean trials;
- The willingness by most field technicians to test and use new field data collection methods and scale to different locations;
- Technical support and facilitation by research and development partners.

Challenges and constraints

The implementation of the program and ODK also faced several challenges.

- Inadequate phones and tablets, lead to the fact that in some instances where researcher's personal phones needed to be used. This is quite risky for securing data.
- ODK is attempting to replace the long-standing paper based practice for field data collection. The adoption levels can be improved by carrying out several cycles of introduction in the same location.
- Monitoring trials is an outdoor activity. Using the tablets under direct sunlight introduces a glare effect, which affects the eyes of the user and can slow down the field activity. Anti-glare shields needed to be attached to the tablets.
- ODK relies on the Internet infrastructure to send data to the cloud. High cost and low penetration of Internet access in remote areas of Uganda affects the real time aggregation of data. The temporary solution is to use the offline ODK function with submission upon return of an Internet connection.

Lessons learned

- Awareness and training is very important when introducing the ODK tool. The technology attempts to replace the industry's practice of paper based data collection, and therefore extensive education and training of users was required to ease the transition process.
- Flexibility is required for ODK to be relevant to the ever-changing data needs of users. There is a need to build institutional capacity to be able to develop and manage their own ODK platforms, which includes being able to author forms.

- There is no one-size fit all approach to introducing ODK as a tool for monitoring. Depending on the audience, one should always adjust the approach accordingly.
- Understanding user ICT skills capacities is a stage that should not be ignored. Carrying a basic ICT skills assessment help one understand the special needs that a user group requires.
- Bean scientist and researchers are privy to their field trial data. A discussion on how the specific data will be used and the applicable data access and ownership rights is essential.

Sustainability

The advent of ICT's provides opportunities for stakeholders to better define, generate, aggregate and share research data in their respective domains.

The ODK platform is open source or "free", making it a relatively cheap piece of technology to acquire. Furthermore the ODK tool does not fully rely on having a full time active Internet connection allowing to be used in the field where Internet accessibility may be limited. However, the cost of acquiring a tablet/phone that is appropriate for use in a real field setup may be quit high (approximately 200 USD/piece) especially if there is need to have farmers use them as well.

There is need for continuous in-house mentoring and re-tooling of users in order to ensure a sustained capacity with institutions.

Licensing costs of the ODK platform need to be considered, so as to assure users on data safety in the cloud. There is a cost implication ensuring the platform is running on a licensed platform.

Technical contact Uganda :

Dr. Stanley Nkalubo

Head of Bean Research Program,

National Crops Resources Research Institute

P.O Box 7084, Kampala, Uganda

Email: *tamusange@gmail.com*

There is no known environmental effect of the ODK, however the rationale for introducing this platform is to reduce on the use of paper tools for data collection and reporting. The effects of replacing the paper tools with electronic data management go a long way in saving the environment.

The popularity of the ODK is slowly gaining ground in the Uganda National Bean Breeding Program. There are opportunities to use the same ODK for data collection by other crop breeding programs.

Replicability and upscaling

The efforts and methodology for entrenching the ODK as a field data collection tool for routine monitoring differs from using the system for a one off effort, such as surveys. The National Bean Breeding program in Uganda has used this platform for the last 2 years for bean field trial data collection. The scale up efforts will seek to promote ODK for harvest data collection.

Because most bean breeding efforts are similar for sub-Sahara Africa, the successful use of ODK for bean field trial data collection in Uganda means this tool and methodology can be replicated in other bean breeding programs throughout Africa with some minor adjustments to fit specific countries.

Testimonies

"The ODK platform, despite being relatively slow, compared to the manual field book we use at data collection has eliminated the time taken to capture this data from the field book to the computer. This is what makes this innovation interesting to work with. I am sure as we get used to it, our efficiency will also improve" - Jane M. Technician NaCRRI

"This field has been completely destroyed. If you gave us that tablet I see you using for collecting data, I could help you when you're very far, so that you don't miss any information."- Farmer in Arua *District.*

"Thank you team. All the data collected in the field has been well received and is ready for use, please come again during harvest so we use the tool again"- Eunice Kensiime - Researcher NaCRRI

Innocent Obilil
CIAT
i.obilil@cgiar.org

Rachel Muthoni-Andriatsitohaina
CIAT
r.muthoni@cgiar.org

About the authors:

Innocent Obilil is a trained ICT practitioner with a passion for ICT's as an enabler in business functions. He is the database manager for PABRA. His contribution is in ICT based monitoring and evaluation. He has been instrumental in digitizing systems for data capture in PABRA, where the introduction of ODK was key. He also leads the automation of responses to interests and demands for routine performance information. He holds a master's degree in Information Management, a PGD in project management and various professional certifications in M&E.

Rachel Muthoni-Andriatsitohaina leads the data and information group at the PABRA program. She oversees the implementation of the knowledge management strategy of the organization. She is a monitoring, learning and evaluation leader. She is involved in evaluating institutional capacity strengthening

Resources

- PABRA Database: http://database.pabra-africa.org
- PABRA 20 YEARS (magazine) https://cgspace.cgiar.org/bitstream/handle/10568/81196/WEB_PABRA20_AR16_17_SPREADS.pdf?sequence=2&isAllowed=y (Page 74)
- Open Data Kit: https://opendatakit.org

E-AGRICULTURE CALL FOR GOOD AND PROMISING PRACTICES

This document was developed in the framework of the 2017 e-Agriculture Call for Good and Promising Practices on the use of ICTs for Agriculture and Rural Development in collaboration with the Food and Agriculture Organization of the United Nations (FAO) and the Technical Centre for Agricultural and Rural Cooperation (CTA).

e-Agriculture is always happy to review your good or promising practices! You can submit a proposal, following the sections in this document to e-agriculture@fao.org
Would you like to stay informed? Register to become a member of the e-Agriculture Community of Practice: www.e-agriculture.org/user/register

Good and Promising Practices on the use ICT for agriculture in collaboration with