BIG DATA SOLUTIONS

Innovative approaches to overcoming agricultural challenges in developing nations by harnessing the power of analytics

June 2015

WORLD BANK GROUP
Leadership, Learning and Innovation

Booz | Allen | Hamilton

Contents

Introduction

The food system is fundamental for human life. It provides energy, nutrition, an income source for billions of people, and is the largest user of the world's natural resources. Yet about 800 million people still go to bed hungry every night, and many more suffer from the "hidden hunger" of malnutrition. Furthermore, assuming current trends continue, the world's growing population is expected to increase by another 2.4 billion people by 2050 , while the increasing variability in weather patterns negatively impacts growing seasons and the abilities of farmers to produce and scale their operations to accommodate that size of a population.

In response, improvements in agriculture operations have to be made at both the policy and individual farmer level in order to realize gains in efficiency and productivity. Further, these improvements need to be realized most in low and middle income developing countries, where there is currently a lack of traditional infrastructural (farming equipment, financing, distribution systems) support and an increased vulnerability to food security and climate change.

Feeding the growing population in the coming years will require ingenuity and innovation to produce more food, on less land, in more sustainable ways. The recent worldwide diffusion of new technologies, combined with Big Data and Analytics, is providing the opportunity for developing countries to leap frog some of the intermediate development phases providing farmer's in the developing world with greater access to timely, cost effective, and personally relevant information on best practices, markets, prices, inputs, weather and news of impending disaster.

In particular, the penetration of mobile phones, and the innovative applications of geospatial and sensing technologies are providing opportunities to use Big Data in support of agriculture initiatives, including the ones funded by the World Bank. The implementation of these technologies offers the potential to provide much needed information to develop a comprehensive and real-time global agricultural statistical database that bolster availability of tertiary services for farmers, identify and mitigate the spread of agriculture diseases, proactively respond to challenges from climate change, and deliver real-time information to farmers that can help them to optimize their operations.

This Solution Brief defines what Big Data is in the context of the developing world, presents a series of case studies on how Big Data has already been used to date, and identifies some lessons learned and potential opportunities for the use of Big Data in supporting the achievement of agricultural outcomes in the developing world.

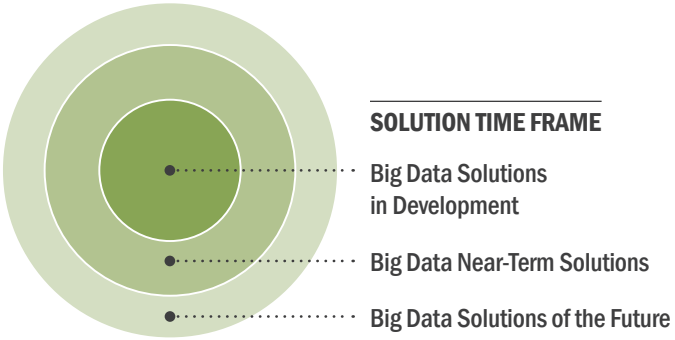
1. UN Reports: http://www.un.org/apps/news/story.asp?NewsID=45165#.VY0dxvlVgSV

What is Big Data In a Development Context? Smaller and Less Diverse, But Still Powerful.

Big Data is normally defined as high-volume, high-velocity and high-variety (the so-called "three Vs" — volume, velocity and variety) data sets that can be analyzed to identify and understand previously unknown patterns, trends and associations.

To date, most attention regarding Big Data has been focused on the developed world. However, the diffusion of low-cost data producing devices (e.g., cells phones, internet-enabled cameras), easy access to wide-area data collection systems (e.g., satellite systems, drone-based data collection), global access to very-large data systems, and the spread of data analytics expertise and low-cost data analytics packages means that Big Data analytics is a reality for developing countries. In fact, developing countries throughout the Americas, Africa and Asia are increasingly making use of Big Data to help analyze trends and guide policy.

It should be noted however, Big Data in the developing world is typically much smaller and less diverse than in the developed world. In the developing world, Big Data is generally in the range of terabytes rather than petabytes or exabytes, and normally comes from a single data source (e.g., cell phone data, satellite data) due to the lack of diverse data collection tools, standards and coordination and infrastructure as well as high costs associated with the setup of broad-based data collection systems. However, the breadth of information even with these smaller, less diverse data streams, when combined with advanced analytics, can reveal insights that were previously unknown. Furthermore, collected data from passive systems like cell phones or satellite imagery reduces the costs of data collection and analysis, and promotes an investigative approach to data in contrast to the static nature of merely creating descriptive reports.



STAKEHOLDER KEY

Farmers | Business/Agricultural Support Operations | Policy Makers

How is the Solutions Brief Organized?

This brief presents three major case studies that have been selected to reflect different levels of availability of Big Data solutions. These case studies (noted below) represent solutions that are either already in use in developing countries or will be ready for deployment in a development context in the near-medium term (within the next 3 years).

- **Using Big Data to Provide Dynamic, Targeted Advice to Rural Farmers.** *Examining the value of crowdsourcing, data integration, analysis and dissemination of information to help support individual farming yields and prioritize development and stakeholder operations. Famers contribute data on which descriptive and predictive analytics are performed, resulting in a dynamic feedback loop where the knowledge of one is merged with the knowledge of many and shared, providing valuable insights to the entire community.*

- **Using Big Data to Support Climate Smart Agricultural Management for Smallholder Farmers.** *Exploration of the use of big data to understand climatic and weather data sourcing and analysis along with an assessment of best practices for analyzing climate risks using crop-weather interaction models. Practical considerations such as the formulation of actionable smart climate management insights for farmers and dissemination of relevant information are also discussed.*

- **Using Big Data to Facilitate Financial Inclusion in Agriculture.** *Examining the value that big data brings to incentivize participation of private enterprise in the agricultural economy and share the burden with policy makers of providing financial access to farmers. The creditworthiness of a farmer is linked to crop yields and future prices in an extremely volatile commodities market. Data is collected, used to define, and measure individual, community and agricultural market risk to encourage participation from multiple market and development stakeholders.*

Complementing these case studies, this brief also contains a section on the future of Big Data in the context of agriculture – summarizing promising technologies, tools and techniques related to Big Data and associated analytics that could become available to the developing world within the next 5-10 years.

Each case focuses on how the information is relevant for a different set of target stakeholders. That is, we describe how information can be used at the national and regional levels (e.g., city, state, or district governments/policy makers), by individual users (e.g., smallholder and commercial farmers and their communities) and supporting operations for agriculture (e.g. development organizations and businesses).

In general, this Solutions Brief is focused less in the particular technology used to collect the data and more on how it has been used and value of the potential outputs. In particular, the case studies focus on challenges and issues that the World Bank can support, highlighting both what has already achieved along with some of the less well known stories of Big Data's use for development.

CASE STUDY

Dynamic, Targeted Advice to Rural Farmers

Agricultural practices, especially in the developing world, have traditionally relied on a model of conventional wisdom that is mostly static and often sub-optimal in the face of the dynamic conditions faced by most farmers today. Today there is immense pressure on the agricultural community to yield output commensurate with the growing population's need to be fed, while maintaining sustainability and ensuring stable livelihoods for farmers.

In this scenario, the use of Big Data and Analytics can be a key component in empowering farmers with a collective knowledge base that involves swift and prolific sharing of information and insights, filling in the gaps left by conventional wisdom.

BIG DATA SOLUTIONS

The Grameen Foundation's Community Knowledge Program has managed to achieve this objective in developing countries like Uganda and Colombia (among others) by leveraging technology and data analytics - disseminating and collecting information to and from those who have the least access to it.

The Community Knowledge Worker (CKW) program, started in Uganda in 2009, consists of a peer network of farmers in remote rural communities. These communities have had limited access to traditional infrastructure, and have a high cost for last-mile manual or infrastructural connectivity, making it difficult to set up the necessary data collection and information broadcast environment. The peer advisors, or "Community Knowledge Workers," are in many cases farmers themselves who are chosen by their peers and respected in their communities. To facilitate communication, the program provides the CKW's with Android phones running an open source, custom agricultural information app with an entire suite of data collection tools called TaroWorks (a product of the Grameen Foundation) behind it. Everything from market pricing, best practices, and disease information is included in the application, and cached for offline access when the phone is off the grid.

The first time CKWs meet a smallholder farmer,

Timeline: Ready to be deployed (Already in place in some developing countries)

Decision Makers: Farmers, Policy Makers, Development Organizations

Countries Involved: Uganda, Columbia, Ghana

Data type: Mobile phones, geolocation, crowdsourcing

KEY FINDINGS

- For agricultural development objectives to be met, a lack of access to infrastructure and technology in the developing world needs to be overcome with the dynamic collection, analysis and sharing of information for growth of collective community knowledge.
- Big Data and analytics, combined with accommodative technologies can empower policy makers and development stakeholders to collect relevant information quickly, understand ground realities and identify problems, prescribe appropriate measures, and broadcast relevant information.
- Data is a powerful tool for optimizing efforts from an operational and policy perspective and mature systems of data collection greatly facilitate the advancement of development objectives.

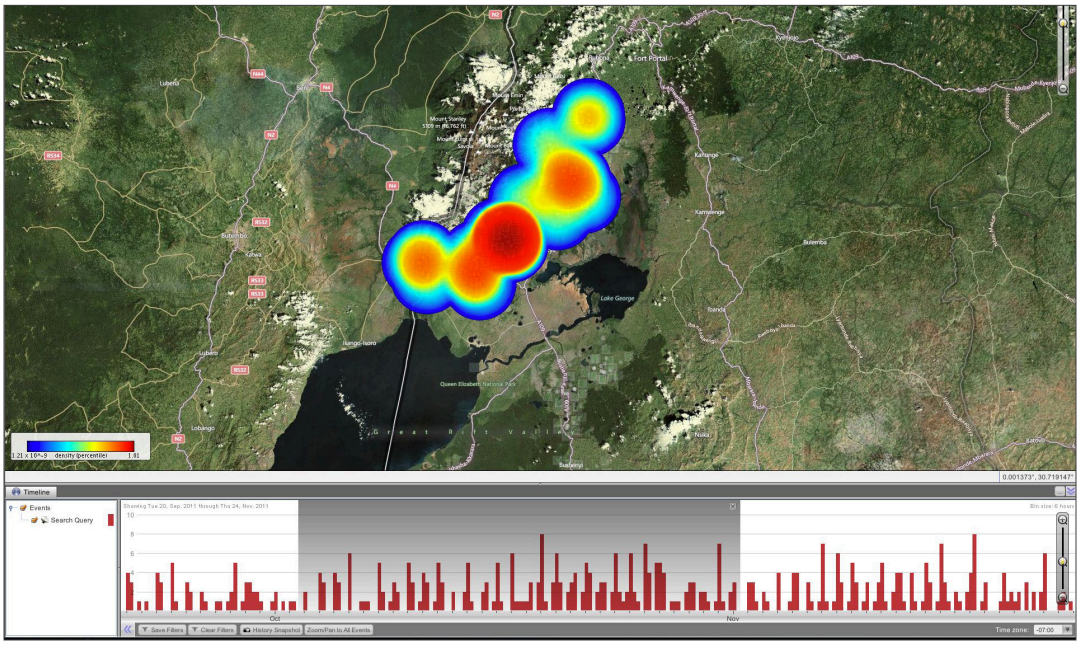


Fig. 1 Heatmap of Chicken Blight in Uganda

they register them using the application and conduct an in-depth baseline inquiry that contains questions about a broad variety of characteristics that include – socio-demographic information of farming household (Examples queries: "How many children under 11 do you have?", "Do your children have shoes? Clothes?", "What do you use for cooking fuel?"), farm information (Example queries: "How many of your livestock are currently ill?", "What pesticides are you using?", "How old are your tree crops?"), certification readiness indicators, good agricultural practices adoption metrics and details on access to finance/services . The collection and analysis of this information provides the Grameen Foundation the ability to examine the issues facing rural smallholder farming communities holistically. The insights that the data reveal are used by the Grameen Foundation to work with partners and facilitate delivery of services that are optimized and targeted to meet specific requirements of the farming communities. An example of this process in action is seen in the facilitation of 'mSourcing' – a process that helps farmer's associations source produce from farmers to sell to grocery chains, using the CKW data collection and information exchange mechanism to pay farmers quickly, and to plan production to meet seasonal demand and quality standards.

CURRENT APPLICATIONS

One of the most important things to note about the CKW program is that the communication and sharing of information is bi-directional. A rural farmer benefits from the information that is provided to him by the app on weather forecasts, market prices, and best practices. At the same

time, the information collected from him/her is a valuable crowd-sourced data set that is used to perform descriptive and predictive analytics, resulting in a dynamic feedback loop where the knowledge of one is merged with the knowledge of many to provide valuable insights to the entire community. The benefits of this bi-directional sharing of information are augmented by the CKW program's collaboration with the TECA (Technologies and Practices for Small Agricultural Producers) initiative of the United Nation's Food and Agriculture Organization. CKWs share information on agricultural practices from TECA's knowledge base with farmers and set up TECA exchange groups to review traditional farming practices that farmers have shared. After vetting, successful practices are added to the TECA content library for use by the greater community.

Another salient success of the CKW program is in its ability to optimally leverage technology even in the face of many challenges. Even though most of the farmers that this program supports live in communities that are outside of the coverage of Ugandan cell networks, the phones use GPS satellite signals to record the exact time and location of each query and data input. When the phones return to the grid, all of the data about the queries are uploaded to a central server. This results in the collection of a perfectly geo-coded dataset that maps out a large proportion of rural farming done in Uganda. The geo-coded and time stamped data is widely regarded as the best data on agriculture in Uganda where the lack of traditional infrastructure generally makes availability, completeness, reliability and relevance of agricultural data deficient.

The Grameen Foundation's CKW have compiled and answered over one million questions from over 200,000 thousand farmers in Uganda through the use of the mobile phone technology outlined above (as of 2014)[2]. This has already resulted in some very interesting and useful results. As an example, analysts at Palantir (partners of the CKW program) used the data collected by the program to evaluate the demographics and activities of the most vulnerable farmers (in the lowest 20th percentile of economic prosperity) and were able to drill down on questions relating to issues concerning farm animals asked between January and May 2012. This simple result led to an enhanced ability for different stakeholders (governments, non-profits, individual farmers) to detect early outbreaks of disease and parasites in chickens (concentrated in Northern Uganda). The coupling of demographic data with farm information revealed insights about the problems affecting most vulnerable farming communities – it

DATA COLLECTED BY CKW VOUNTEERS

- Social and Demographic – Characteristics of households, number of people reliant on farm for livelihood, economic prosperity etc.
- Current Farm Condition – Area under cultivation, productivity and yield, common pests and diseases, access to finance and services, etc.
- Crisis information – Adverse weather events, reports of degradation of soil/water, etc.
- Operational metrics – Number of people enrolled, time taken to complete survey, volunteer productivity.
- Best practices adoption and discovery

2. Palantir Blog Post: https://www.palantir.com/2012/10/grameen-foundation-palantir-partners-for-food-security/

was seen for example that during a period in 2012, farmers in southwestern Uganda were experiencing outbreaks of disease in cattle, and that these same farmers commonly did not own clothes and shoes (indicating poor economic conditions). Such data-driven insights were primarily made possible by the clean data collection practices of the CKW program. Traditional reporting and surveillance systems in the developing world typically cannot collect the data in the same manner as the CKW program reaches the remotest of farming communities due to absence of infrastructure. Further, the use of Big Data analytics on the collected data showed how economic conditions of rural farming communities can be combined with breakout information on plant and animal diseases, changes in weather patterns, soil fertility etc., empowering policy makers and other development stakeholders not only with early warning and detection signals of oncoming issues, but also with prioritization of development initiatives in real time – such that they maximize the effect of their efforts to solve problems for the people that require it the most.

As the CKW data collection and information sharing architecture has matured, the scope of the usage of analytics has also evolved – the substantial size of the collected data has allowed Palantir to develop multiple dashboards for the Grameen Foundation that cover operational statistics, economic viability assessments, discovery of optimum productivity practices, and good agricultural practices adoption rates/program effectiveness. This has transformed the way that foundation carries out its objectives in the agricultural domain to include:

- Operational statistics allow the foundation to monitor data collection/survey speeds and survey rates among other details so that they are always aware and can quickly respond to any challenges that their workers face.
- The economic viability assessment and impact tools allow the foundation to identify smallholder farmers along metrics of productivity, food security of household, ability to grow, etc. and use the insights from the analytics to accurately identify the needs and corresponding support that the farmers need (as an example, some farmers only face minor hurdles such as lack of access to good financing or outdated machinery, which can be acted upon by the foundation and its ground partners while other farmers face major challenges such as water contamination, high-dependence of household members, disease infestations, etc. that need to be escalated to government stakeholder level for resolution).
- The identification of good agricultural practices and tracking of adoption rates combines the knowledge of one with the knowledge of the community (program administrators

STAKEHOLDER INTEREST IN INSIGHTS FROM DATA

- Individual Farmers – Learns best practices and gets access to services that are at the right place at the right time
- Non-profits/non-government development partners – Look at effectiveness of programs, increase efficiency of delivery of services, develop effective communication and information sharing.
- Policy makers/Government Departments – Identification of targeted issues and food security risk mitigation, analysis and predictability of cost-benefit of development efforts.

use the analytics dashboards to identify the most effective solutions to common/uncommon problems faced by the farmers and can disseminate the information widely and effectively) while tracking the realities of how many farmers are adopting these practices to optimize their outreach.

FUTURE APPLICATIONS

The use of data has also allowed the Grameen Foundation to look inward and develop strategies to improve the CKW program. When the community volunteers share best practices and insights disseminated to them through cell phones, they also have the opportunity to collect data on the adoption of these practices at a later stage, giving program managers the ability to examine correlations between specific sets of volunteers and workers and high/low levels of best practices adoption. Insights from the successes and failures of their efforts allow the foundation to improve the program and fine-tune activities.

Originally initiated in Uganda, the CKW initiative has now been expanded to Colombia and Ghana. In Colombia, the Grameen Foundation is working with farmers and government stakeholders in Salgar, Antioquia, an area that has been hit especially hard during the country's internal conflicts. The government of the state of Antioquia in Colombia used the CKW model to deliver its backyard gardening and entrepreneurship program. Using the information provided, the government helped improve service delivery by

- using data to improve targeting: Identifying 25% of the families in the program who were relatively food secure and whose socio-economic conditions could be lifted by providing targeted commercial services and replacing them with families who are food insecure;
- changing the services package to maximize nutrition outcomes for the population;
- linking families and children who qualify to receive benefits but are not currently enrolled in school feeding and healthcare programs combining demographic data collected for different purposes across different government policies;

The government was able to show that families saved on average $300/year on food and that severe food insecurity was reduced by over 75%. This evidence helped make a broader case for the intervention and is being considered as a model for replication by other policy making organizations in the country.

The Community Knowledge Program is a people-involved, relatively low cost model that can transform the information-sharing paradigm in some of the most underprivileged environments in the world. This initiative can be set up and supported either by governments, aid organizations, or non-governmental organizations with the only significant requirement of effort being the recruitment of community participants and building of a peer network for farmers.

CASE STUDY

Climate Smart Agricultural Management for Small Holder Farmers

Weather is the single greatest factor influencing a smallholder farmer's productivity. Agricultural output and efficiency is extremely dependent on weather and climate to achieve food production at reasonable costs. Weather variability may directly affect production levels if farmers are not able to adapt their systems and practices in time. Around the world, precipitation patterns, maximum and minimum temperature cycles, etc. have changed differently in each region and the climate is becoming less and less predictable.

Traditionally, farmers have used conventional wisdom such as calendar references (when to sow, what to sow, when to harvest, etc.). However, due to increasing weather variability, this knowledge is constantly being challenged and farmers lack the information to make the right decisions in a rapidly changing environment.

It is important for agricultural development stakeholders/decision makers to extend weather smart agricultural management based on big data, especially to smallholder farmers in developing countries, such that optimal and stable agricultural yields that support food security are ensured along with the protection of livelihoods of some of the poorest sections of their population.

Timeline: 1-3 years

Decision Makers: Farmers

Countries Involved: Columbia, Ghana

Data type: Survey Collected, Mobile, Sensors

KEY FINDINGS

- Agricultural Output is highly reliant on the climate and hence, the livelihood of some of the poorest sections of the population are vulnerable without smart, dynamic management of weather effects on farming
- Combining data from multiple sources provides a longitudinal view of climate effects and produces insights that can be extremely beneficial to farmers
- Simply collecting the data and generating insights is not enough; the use of technology to make sure that the message reaches the right stakeholders at the right time is also important.

SOLUTIONS

aWhere has empowered smallholder farmers by providing real-time, localized weather and agronomic data covering all agricultural geographies. This global dataset provides accurate, consistent and current observed weather data, as well as forecasted data up to 8 days into the future, and historical weather data for up to 20 years in the past for some regions. The consistency and availability of this data, along with the ease of access that the platform and the application programming interface (API) provide, enables last-mile technology partners to pull together insights and convey them to farmers in a cost effective, yet meaningful way. Researchers and commercial practitioners are able to combine weather data with historical information on crop yields to generate field-specific agronomic models, as well as weather smart agriculture management recommendations. These models and recommendations generate new agricultural intelligence which enhance traditional agricultural practices and guide farmers on how to mitigate the risks of adverse weather events and climate variability.

CURRENT APPLICATIONS

One crucial component of agricultural intelligence is the ability to create localized recommendations or advisories that are matched to growth stage models specific to the region in which the farmer operates or even specific to the farmer's field. In this effort, it is important to be able to combine data from disparate sources (e.g. agricultural yield and weather changes) in order to predict future agricultural outputs.

An example of the research necessary to inform localized agricultural recommendations was conducted by the Center for Tropical Agriculture (CIAT). CIAT studied commercial harvest of rice in Colombia and its relationship with certain key climatic components from 2007-2013. For each climate sequence between sowing and harvest (around 4 months), weather variables were gathered from data provided by a climate research institute of Colombia and cross-referenced with harvest yield information, national rice surveys, and sowing experiment information provided by the Colombian Rice Growers Federation (FEDEARROZ), revealing underlying correlation patterns between climate factors and yield variability.

Fig. 2 Farmers in Malawi, using aWhere and Esoko's Agricultural Data SMS insights

This analysis identified the main growth and limiting climate factors for each region, and allowed researchers to identify why high or low yields were observed in a particular crop cycle. This insight, along with local seasonal forecasts, could allow farmers to infer possible future scenarios and determine which rice variety and which sowing date would result in optimal yield.

Beyond performing analyses that have the potential to provide weather smart agricultural management to the smallholder farmer, it is necessary to ensure that these insights are well received and implemented. The barriers of access to information technology infrastructure and poor rates of literacy in developing environments are important considerations. Further, the current paradigm of data sharing for weather smart agriculture is largely unidirectional – with the data providers either supplying the data or derivative insights to their consumers. This model works well in the developed world due to a greater sophistication of infrastructure and availability of tools for automated collection of feedback, but fails to be valuable from both a development and a business perspective in the developing countries.

The aWhere platform was developed to overcome the hurdles of providing analytics to smallholders. Its platform consists of open API driven systems that have the ability to track information automatically at a low cost, creating a data-driven feedback loop, which can be used to iteratively improve the sharing of insights and information while maintaining the appropriate balance of abstraction in the context of poor technological access and farmer literacy in the developing world.

A key example of this system's use is aWhere's partnership with Esoko in Ghana. Esoko has developed a robust communications solution on top of aWhere's weather and agronomic data platform that incorporates mobile messaging, voicemail, and a farmer call center. aWhere's API metrics for its Esoko customers provide daily baselines on actual usage of weather information services by the individual farmers broken down by location and timing. As of 2015, aWhere is providing data to and collecting usage metrics from over 200,000 fields per day[3].

The primary advantage offered by aWhere as compared to the traditional "ground" meteorological data collection systems common in developing countries lies in the use of Big Data analytics by aWhere to combine automatically collected API metrics with satellite data and remote weather sensing information, creating a dynamic, accurate cheap, customizable and scalable system that can support the weather management needs of smallholder farmers in the most efficient way possible.

FUTURE APPLICATIONS

Combined with contextual information about individual farmers, metrics collected by aWhere on the usage of its API can be analyzed to provide insights into the smallholder's adoption of weather smart practices from a development policy perspective. Gathering statistics on the number of farmers subscribing to weather forecasts (broken down by gender, location, growing choices), number of farmers subscribing to agronomic advice, number of farmers participating in sustainable agricultural practices, and the gap between optimal potential yields vs. realized yields can be used to track vulnerabilities in the climate management strategies of agricultural communities with implications for protection of livelihood, poverty risk reduction, food security and crisis intervention practices.

Ongoing research shows that farmers using services from Esoko have seen significant improvements in yield compared to control plots. However, the current process for disseminating information to rural farmers may hinder its deployment and democratization in the near term. Although some of the more advanced smallholder farmers are able to benefit from the automatic extracts through API's and delivered through mobile technologies, the majority may still require more traditional dissemination methods like manuals, reports, or slick sheets delivered in person during the introductory stages. Moving forward, aWhere and other providers are not only considering more ways to extract value from their data for weather-smart agriculture to smallholders in developing countries but also working on ways and means to ensure that useful and sufficiently abstracted information is broadcasted in a cost-effective yet prolific manner.

Promoting weather smart agricultural management has the potential to compensate for the rapidly declining reliability of traditional knowledge in the face of increasing climate and situational variability that many smallholder farmers face today by guiding efficient decision making on what, when, and where to sow, and how to manage farms. The impact of better-managed crops can include increased farm profit and productivity, higher overall regional food security, and general social stabilization in developing countries.

3. aWhere Inc. Monthly API metrics report (July 2015)

CASE STUDY

Facilitating Financial Inclusion in Agriculture

In order to feed the growing population of the world in the near future, some estimates reveal that overall food production will need to grow by over 70% between now and 2050[4]. The availability of land for farming is already a strained resource, making any increase in food output increasingly reliant on optimizing farm practices and maximizing the value that can be extracted from land. In order to achieve this, both small and commercial farming operations all around the world need access to capital, seeds and fertilizers, crop insurance, storage, and distribution.

Agriculture in the developing world today is obstructed by a lack of reliable infrastructure. As an example, the worldwide average number of tractors per 100 sq.km of arable land in 2012 stood at 200, whereas in Asia it is closer to 129 and in Africa, it is only 13[5]. Further, despite having close to one-fifth of the world's (currently utilized) arable land, the agricultural sector in Africa only accounts for around 3% of the world's fertilizer consumption[6]. Central in this deficit of infrastructure is the unavailability of financing for farming in the developing world. On average, agriculture accounts for 35% of the gross domestic product and employs close to 65% of the labor force in African countries, yet less than 1% of commercial lending in Africa goes to the agricultural sector[7]. Farmers in the developing world today do not have the financial means, such as crop insurance, to ensure their livelihood and are often forced to arrange borrowing at high interest rates because of the lack of fair financing options, severely exacerbating the strenuous grip of poverty over a large section of the population.

From a policy perspective, moving infrastructure towards supporting the needs of the farmers is an especially daunting task that involves large-scale investment in machinery, including the expenditure of considerable time and effort. Further, micro-financing and social cooperative efforts work well but are often resource strapped, limiting the number of people they can serve. Big Data can play a vital role in easing the process and making financial inclusion in agriculture a possibility.

Timeline: 1-3 years

Decision Makers: Farmers, Businesses, Policy Makers

Countries Involved: Ethiopia, Multiple East-African Countries

Data type: Government Authorities, Development Organizations, Crowdsourcing

KEY FINDINGS

- Agriculture is perhaps one of the most underserved economic sectors. Lack of financial and business services prevent farmers from growing at the rates necessary to feed an increasing and undernourished world population.
- Big Data from multiple sources and stakeholders can help provide the essential backbone for the advent of these support services in the absence of traditional infrastructure, providing use to farmers, businesses and even policy makers.

SOLUTIONS

The case of GroVentures (originally started in Ethiopia and currently operating as GroIntelligence in Kenya), showcases how private enterprise can utilize Big Data to participate in the agricultural economy and share the burden with policy makers of providing financial access for farmers.

Commercial lending is missing from agriculture because of the difficulty in quantifying the risk of variability in production. Traditionally, banks quantify the risk of lending by using multivariate mathematical credit models which predict the ability of a borrower to repay loans (and determine the appropriate interest rate). The creditworthiness of a farmer is therefore always linked to crop yields and future prices in an extremely volatile commodities market. Data can be provide crucial understanding of risk factors involved and encourage participation from multiple market and development stakeholders. At the same time, a lack of data creates an environment in which providing support services in agriculture becomes difficult.

CURRENT APPLICATIONS

When GroVentures started operations in 2014, they found that data on historical yields, regional practices, and general sector related statistics was missing. As an example, the United Nations Food and Agricultural database (perhaps one of the only long-standing sources of data) only contained static, abstracted, high-level information relevant for policy and even today is missing information from many East African countries since 2011. GroVentures overcame this problem by tracking down and aggregating data from a large range of nonprofits, social enterprises, and developmental agencies working in the agricultural space, augmenting this with previously unavailable harvest data from authorities in the African Union. This resulted in a consolidated, granular and ubiquitous information set for all things pertaining to agriculture – from soil quality, to historical yield, to correlations between yield patterns and elemental changes. In 35 of 54 countries, this meant that, for the first time, stakeholders had information on the risk profile of each 100 square kilometers of arable land. The advent of data and analytics to support agriculture had a positive impact on businesses, individual farmers, and policy makers. Businesses are now able to effectively allocate capital and loans to the agricultural sector while using a data-driven approach to fairly evaluate risk and offer affordable crop-insurance, alleviating the financial strain that farmers and farm laborers currently experience

Did you know?

Agriculture accounts for 32% of the GDP and employs 65% of the labor force in Africa, yet **less than 1% of commercial lending goes to agriculture**

Africa would need 3.5 million more tractors to be on par with other areas of the world, since it only has **13 tractors per 100km^2 of arable land**

Africa lags behind most of the of the world in fertilizer consumption, accounting for only **2.9% of overall fertilizer consumption in 2011**

in the absence of the data. GroVentures hopes to facilitate loans of $25 million to approximately 73,000 farmers in cooperatives and informal farmer organizations via pooled lending by 2015. This will bring financial inclusion not only to the farm owners, but also to the workers and families attached to the agricultural businesses, affecting directly or indirectly approximately 290,000. Apart from the positive economic impact on a large underserved portion of their population, up-to-date information on market activity and production estimates from the data platforms would also equip Policy Makers to identify vital transportation and infrastructure needs. Furthermore, providing this up-to-date information will help policy makers fulfill immediate objectives while bolstering their ability to better manage grain reserves, and ensure that production shortfalls and market disruptions do not hinder broader population food security.

FUTURE APPLICATIONS

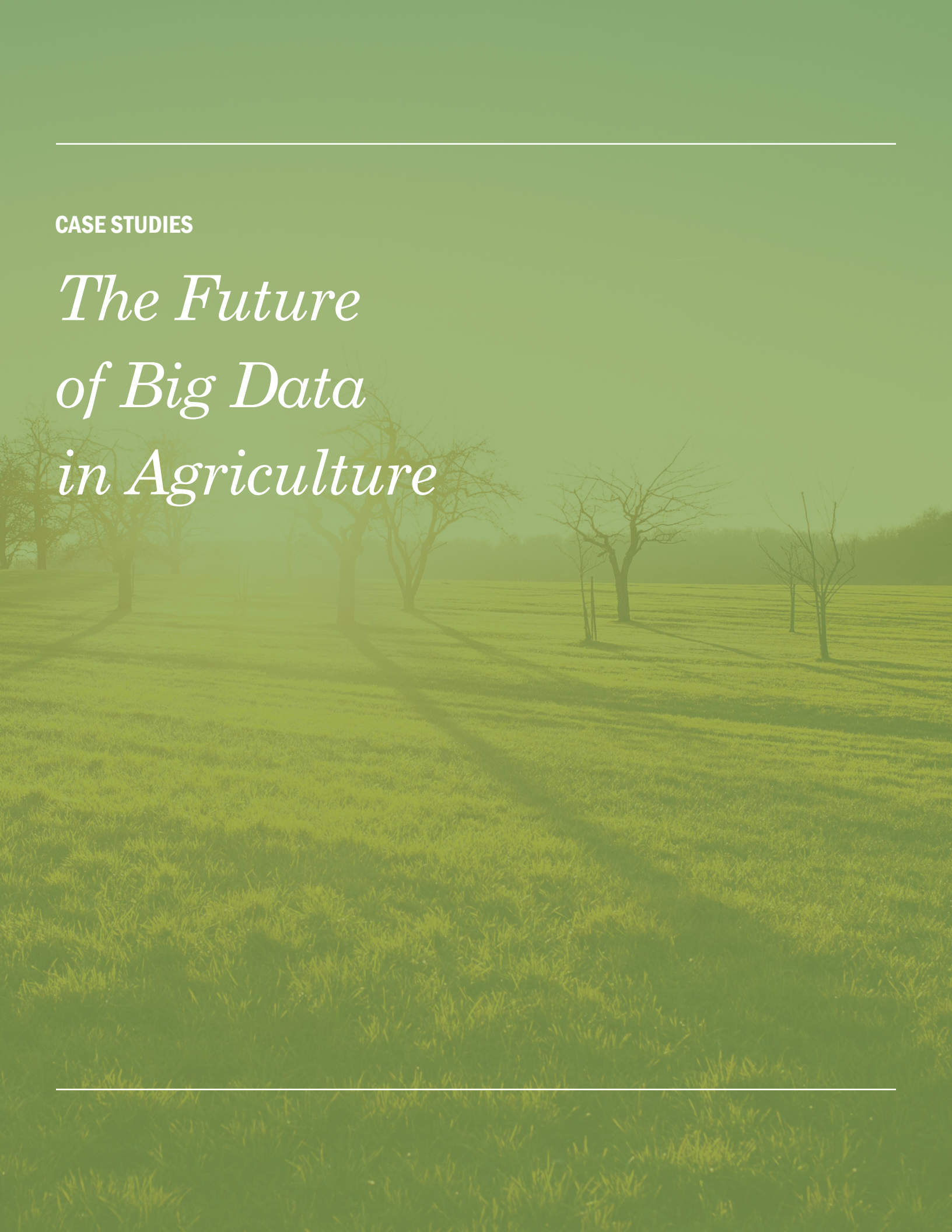
GroVentures is now embarking on crowdsourcing additional data directly from farmers in order to fill in the gaps that their analytics platform requires. Farmers are beginning to manually self-report information in exchange for mobile airtime credit or access to weather and meteorological services for free. This bi-directional exchange of data and information has an immense potential for optimizing collection practices and monitoring crucial components of a lending profile, while pushing out best practices and ensuring their implementation for a more prosperous outcome for all parties involved.

Policy Makers can support the financial inclusion initiatives by providing access to open data and facilitating commercial participants like GroVentures, such that the pitfalls of a slow, linearized, undeveloped infrastructure are overcome by entities that can combine information from multiple sources and provide crucial support for services that the agricultural industry requires.

4. United Nations Food and Agricultural Organization Report – Global agriculture towards 2050 (http://www.fao.org/fileadmin/templates/wsfs/docs/Issues_papers/HLEF2050_Global_Agriculture.pdf)
5. FAO Statistical Yearbook 2012: Africa Food and Agriculture (Chart 20) http://www.fao.org/docrep/018/i3137e/i3137e.pdf
6. United Nations Food and Agricultural Organization Report - Current world fertilizer trends and outlook to 2016 (ftp://ftp.fao.org/ag/agp/docs/cwfto16.pdf)
7. World Bank Group International Finance Corporation (IFC) Report - IFC and Agri-Finance: Creating Opportunity Where It's Needed Most (http://www.ifc.org/wps/wcm/connect/44c74a0049585fb1a082b519583b6d16/IFC%2B%2Band%2BAgriFinance%2B-%2BGeneric%2BPresentation%2Bv-4.pdf?MOD=AJPERES&CACHEID=44c74a0049585fb1a082b519583b6d16)

CASE STUDIES

The Future of Big Data in Agriculture



Data Driven, Cost-Effective Climate Insurance

Data Source: Sensors, Cameras

Stakeholders: Farmers, Businesses

In developed economies all around the world, commercial farmers are turning to the use of a combination of sensors, cameras and probes to generate data and analytics for optimized risk management and cost effective selection of crop insurance. As an example, the Climate Corporation, recently acquired by the Monsanto Company as part of their Big Data and Analytics offerings suite, utilizes field sensors to collect crop yield, soil quality, and weather pattern data, along with farmer reported inputs, to help commercial operations make the best decisions with regards to choosing crop insurance. The technology and data analytics suite quantifies risk factors and combines historical data with predictive intelligence in order to suggest best practices that can mitigate specific risk elements. The platform is currently geared to benefit large-scale operations and requires significant investment in setting up the data collection infrastructure. However, as developing countries leapfrog established technology in favor of cost-effective alternatives (e.g. replacing costly sensor equipment with low cost cameras and software) more and more farmers will be able to incorporate smarter techniques for managing agricultural risks and insuring risks to their livelihood.

Mapping with Drones to Feed the World

Data Source: Drones, Cameras

Stakeholders: Farmers, Policy Makers

Remote sensing and monitoring via the use of drones is increasingly becoming a reality for developing countries as the underlying technology becomes cheaper. To take advantage of this increased accessibility that drones will have in the future, Raptor Maps a startup based in the United States, has created a novel platform that uses remote sensing (mainly drones, among other sensors) and associated data analytics to monitor crop and irrigation health, with the objective of managing environmental impact of farming while increasing crop yields. Through a web-interface, the company can provide insights and key statistics directly to farmers and commercial growers. Expanded, this same technology can be repurposed for the remote data collection for policy makers and governments with the associated analytics, resulting in early detection and monitoring of crop and animal disease outbreaks, irrigation and water level control, management of food security, and production deficits, among many others.

Using Data and Analytics to Practice 'Precision Agriculture'

Data Source: Sensors, Drones, Cameras

Stakeholders: Farmers

'Precision Agriculture' is the collection of real-time data on weather, soil and air quality, crop maturity, and even costs and availability of equipment and labor, followed by predictive analytics that guide smarter dynamic decision making. Traditionally, precision agriculture has been limited only to developed countries and the most large-scale operations due to significant requirements on technology and capital with control centers collecting and processing data in real time and providing direct action items for immediate implementation. This results in farmers making the best decisions with regard to planting, fertilizing, and harvesting crops at any given time. Expensive sensors that need to be placed throughout the fields to measure data points such as temperature and humidity of the soil and surrounding air, yield tracking, and microclimate responses are fast being replaced by insights derived from drones and satellite imagery which are low cost alternatives. Further, in developing countries, the availability and relative cheapness of human capital makes it possible for large-scale manual collection of data – providing low tech alternatives that still achieve the same level of results for precision agriculture. Data-driven precision agriculture - which is already increasing yields by close to 20% and cutting costs by close to 50% in many operations - will be a crucial component that will help the sector sustainably keep up with demand while advancing prosperity for all.

References

CASE STUDY: Providing Dynamic, Targeted Advice to Rural Farmers

1. Gantt, W. (2014). Grameen Foundation and Learning with Community Knowledge Workers. The Grameen Foundation;
2. Palantir and Grameen Foundation; Kari Knox. (2014, November 25). FOOD FOR THOUGHT: IMPROVING FARMER LIVELIHOODS ON 3 CONTINENTS. Retrieved from The Palantir Blog: https://www.palantir.com/2014/11/food-for-thought-improving-the-lives-of-farmers-on-3-continents/
3. Palantir and Grameen Foundation; Ari Gesher. (2012, October 12). GRAMEEN FOUNDATION & PALANTIR: PARTNERS FOR FOOD SECURITY. Retrieved from The Palantir Blog: https://www.palantir.com/2012/10/grameen-foundation-palantir-partners-for-food-security/
4. The Grameen Foundation. (2010, March). COMMUNITY KNOWLEDGE WORKER PILOT REPORT. Retrieved from The Grameen Foundation Publications: http://www.grameenfoundation.org/resource/community-knowledge-worker-pilot-report

CASE STUDY: Climate Smart Agricultural Management for Small Holder Farmers

1. Accenture Institute for High Performance. (2012). http://www.accenture.com/Microsites/emerging-markets/Documents/pdf/Accenture-Esoko-Case-Study-Final.pdf. Retrieved from http://www.accenture.com/Microsites/emerging-markets/Documents/pdf/Accenture-Esoko-Case-Study-Final.pdf
2. Cariboni, D. (2014, September 30). Colombia rice growers saved from ruin after being told not to plant their crop. Retrieved from The Guardian: http://www.theguardian.com/global-development/2014/sep/30/colombia-rice-growers-climate-change
3. Esoko. (2014, August 28). Will it rain today? Weather Straight to your inbox. Retrieved from Esoko Blogs: https://esoko.com/will-it-rain-today-the-weather-straight-to-your-inbox/
4. International Center for Tropical Agriculture (CIAT): http://ciat.cgiar.org/. (2014). Proof of Concept: Big Data for Climate Smart Agriculture - Enhancing & Sustaining Rice Systems for Latin America and the World. World Bank Innovation Challenge Submission Documents.

CASE STUDY: Ensuring Growth And Prosperity by Facilitating Financial Inclusion in Agriculture

1. Byrne, C. (2013, September 14). Data-Driven Lending Could Help African Farmers Feed The World. Retrieved from Fast Co.Labs: http://www.fastcolabs.com/3019953/data-driven-lending-could-help-african-farmers-feed-the-world
2. GroIntelligence. (2015, May 29). African Fields to Whole Foods: Potential of Organic Trade in Africa. Retrieved from GroIntelligence Weekly analysis: https://www.gro-intelligence.com/weekly-analysis/Gro-Organic_Potential_May_29_2015.pdf
3. GroIntelligence. (n.d.). Original Analysis by Gro Intelligence. Retrieved from GroIntelligence Website: https://www.gro-intelligence.com/#newsletter
4. GroVentures. (2013, April 12). Big Data as a Tool for Financing African Farmers: Data-Driven Lending. Retrieved from SlideShare Net: http://www.slideshare.net/pafoafri/gro-ventures-cta-presentation-sara-maker

References

The Future of Big Data in Agriculture

1. Sherrick, B. J., Barry, P. J., Ellinger, P. N., & Schnitkey, G. D. (2004). Factors influencing farmers' crop insurance decisions. American Journal of Agricultural Economics, 86(1), 103-114.
2. Sharma, R. (2013). Growing the Uses of Drones in Agriculture. Forbes. Available online at:[http://www. forbes. com/sites/rakeshsharma/2013/11/26/growing-the-use-of-drones-in-agriculture/].
3. Saari, H., Pellikka, I., Pesonen, L., Tuominen, S., Heikkilä, J., Holmlund, C., ... & Antila, T. (2011, October). Unmanned Aerial Vehicle (UAV) operated spectral camera system for forest and agriculture applications. In SPIE Remote Sensing (pp. 81740H-81740H). International Society for Optics and Photonics.
4. McBratney, A., Whelan, B., Ancev, T., & Bouma, J. (2005). Future directions of precision agriculture. *Precision Agriculture*, 6(1), 7-23.
5. IBM Research. (n.d.). Precision agriculture: Using predictive weather analytics to feed future generations. Retrieved June 26, 2015, from IBM Research - Articles: http://www.research.ibm.com/articles/precision_agriculture.shtml
6. The Climate Corporation. (n.d.). *Climate Insurance*. Retrieved June 26, 2015, from The Climate Corportation Website: https://www.climate.com/mp-federal-crop-insurance/
7. Subbaraman, N. (2015, May 14). Raptor Maps' drone technology snags MIT's $100,000 prize. Retrieved June 30, 2015, from http://www.betaboston.com/news/2015/05/14/raptor-maps-drone-technology-snags-mits-100k-prize/
8. Raptor Maps. (n.d.). *Technology*. Retrieved June 26, 2015, from The Raptor Maps Website: http://raptormaps.com/technology.html

Contact

This solutions brief is one of several knowledge products delivered through the Innovation Labs and their program in Big Data Analytics. The solution briefs are produced in close cooperation with Global Practices across the World Bank Group.

The Innovations Lab sits in the Leadership, Learning, and Innovation vice presidency. Its big data program includes a core program team from the Development Economics Group, the Transport and ICT Global Practice, and Information and Technology Solutions. The purpose of the program is to accelerate the effective use of big data analytics across the organization, and to position the World Bank as a leader in the big data for development community.

For additional information about this solutions brief or to find out more about the program, please visit http://bigdata (WBG intranet) or contact Adarsh Desai (adesai@worldbank.org) or Trevor Monroe (tmonroe@worldbank.org).