INTERACTIVE VOICE RESPONSE SYSTEMS FOR LOW-LITERATE FARMERS

Eli Berk-Rauch
Benjamin Manning
Master in Public Policy Candidates
Expected Graduation: 05/27/2021

Precision Agriculture for Development

Client: Precision Agriculture for Development
Faculty Advisor: Dr. Rema Hanna
PAC Seminar Leader: Dr. Dara Cohen

Harvard Kennedy School of Government
April 6th, 2021
Submitted in partial fulfillment of the requirement for the degree of Master in Public Policy

Both authors contributed equally to both the research and writing for this report.

Executive Summary

Precision Agriculture for Development (PAD) works with government and non-governmental organizations to support the use of information communications technology (ICT) interventions for agricultural development. PAD has partnerships in nine countries providing support for the development, operation, and impact analysis of a variety of mobile phone intervention programs. PAD works with Ethiopia's Agricultural Transformation Agency (ATA) on the 8028 Farmers' Hotline, an automated interactive voice response (IVR) call system that allows farmers to access content on a variety of agricultural practices.

To support PAD, this report looks to improve the performance and usability of the 8028 system. We make design recommendations to make the 8028 system easier to navigate and improve performance without meaningfully changing user behavior. Within this work, we consider what specific design choices must be made to accommodate callers with less digital literacy and from disparate information cultures.

We find at the heart of the 8028 system a trade-off between optionality and simplicity. In order to make a staggering amount of information available, the 8028 design must be so complex that it may prevent any of the content from being accessible to many users. Using operations analysis, data analysis, elite interviews, and machine learning techniques, we provide two main recommendations to simplify the system. The first looks to restructure the call flow through which users navigate to reduce options and perform optimization based on past user performance, providing incremental gains in performance for all users. Our second recommendation uses a statistical model to predict potential weak users, attempting to focus performance gains where and with whom they would be most impactful.

Without sufficient access to outcome variables, we are unable to validate our findings beyond making clear and logically sound reasoning for why they will work. These techniques could be more formally tested through A/B Tests. We also provide a recommendation to engage users in the further design and optimization of these functions.

The 8028 system speaks to a greater need to consider the importance of simplicity in these intervention schemes. Mobile phone intervention schemes have exploded across the developing world, with ever greater functionality but also complexity, often designed with the administrator more in focus than the user. The recommendations provided throughout this paper look to refocus efforts on the user while providing a framework for improvement that could be replicated in any ICT program, regardless of the underlying technologies and delivery method.

Table Of Contents

Acknowledgements

A Thanks to PAD

We could not have completed this project without the guidance and initiative of **Ilana Kessler** and **Jonathan Faull**. They supervised our work, helped us navigate the entire process, and were always understanding of our occasionally sporadic communication, as HKS graduates who wrote their own PAEs.

We would also like like to thank **Habtamu Yesigat** and **Yabebal Ayalew** for their patience in working with us across continents and tireless explanations of the data.

A Thanks to the ATA

Although we did not explicitly work with anyone in the ATA on a consistent basis, we would like to thank them for providing the resources for us to complete this project. We are proud to be a part of their mission of improving agriculture in Ethiopia.

A Special Thanks to our HKS Professors

This report is truly a testament to our disparate and sometimes random set of skills, mostly acquired during our time at HKS. We thank and acknowledge the professors whose concepts formed our methodology.

We learned how to use descriptive analysis and machine learning techniques after taking API-202z: Empirical Methods II with Professor **Teddy Svoronos,** API-222: Machine Learning with **Dr. Soroush Saghafian,** and APC 209a/b: Topics in Data Science with Professor **Pavlos Protopapas.**

The idea of subspace manipulation without knowing the probabilities of actor decisions was inspired by Problem 2 of Problem Set 1 of API-302: Analytical Policy Frameworks with Professor **Richard Zeckhauser.**

In API-101z: Markets and Market Failure, Professor **Christopher Avery** taught us sit in a room and argue about a variety of economic and statistical applications. His course and office hours were essential to our development, and we appreciate his putting up with us every week.

Many aspects of the operations analysis were formulated using ideas like who/what and 4Ps from MLD 601: Operations Management with **Professor Mark Fagan.**

Professors **Rema Hanna** and **Dara Cohen** guided us immensely in our work and (hopefully) ensured we are not the group discussed at the beginning of next year to motivate their cohort (unless it's because we crushed it).

Introduction

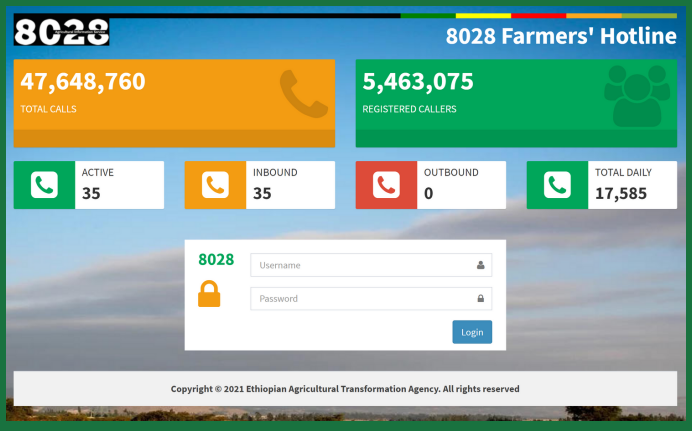
The 8028 Farmer Hotline (8028), operated by Ethiopia's Agricultural Transformation Agency (ATA), is an automated messaging service that provides content and advice on best farming practices in rural areas.

We attempt to improve the usability and accessibility of the 8028 system, with a focus on low-literate and high-frequency users. We form motivating evidence for our study through operations analysis, a literature review of the field and previous research on the 8028 system, and exploratory data work. Using a sample of 16445 calls from 1000 callers, we employ machine learning techniques to predict who will reach content and determine the most important variables in predicting user outcomes within the hotline.

This study is specifically interested in increasing the number of users who reach content, or pre-recorded pieces of farming advice, when calling into the 8028 system. In order to realize 8028's full value, farmers must be able to reliably, quickly, and accurately access information. These central goals are contrasted by the magnitude and variety of options provided through the system, which necessitate a complex interface requiring considerable user aptitude, especially for users with little digital experience.

To motivate our work, we distill the different demands on the 8028 system to identify key performance metrics as well as their implications on the broader performance of the system and methods of measurement. While these metrics did not represent quantitative measures, they provided a dimension to qualitatively consider potential interventions.

What is 8028?



Source: Ethiopian ATA 8028 Farmers Hotline Dashboard

The 8028 farmers' hotline began in 2014 and has received over 46 million calls from 5.6 million unique callers in five languages. The service functions primarily as a repository for a vast and diverse set of recorded advice, ranging from 20 seconds to two minutes. Farmers are able to receive advice on best practices in pre-planting, planting, crop protection, fertilizer use, and post-harvest /processing for a variety of crops.

System Overview

Aspects of 8028

Recruitment

Recruitment for the service is mainly through radio advertisement campaigns. The service's 5.6 Million unique users represent over 10% of all mobile phone subscriptions in Ethiopia (World Bank, 2020) and most likely a larger proportion of rural mobile phone users.

Content

Farmers hear pre-recorded advice on practices and processes for 20 different crops. Recordings range from 30 seconds to 2 minutes and include specific, specialized terminology. Content is customized to the user's language:

Sample Content

"Currently, two sources of nutrients are recommended: Urea (46% N) and NPS Ammonium Sulfo Phosphate (19% N + 38P2O5+ 7S). The recommendations for nutrient levels vary from place to place and additionally depend on the amount of rainfall. In areas with a long history of cultivation and rainfall above 800mm during the crop season, the optimum level of fertilizer application is most likely in the range of 200 to 300 kg/ha of Urea (92-138 kg of N/ha) and 100 kg/ha of NPS. However, if farmers find these levels to be high and difficult to afford, then 150 kg/ha of Urea and 100 kg/ha of NPS can be applied with the understanding that grain yield will be somewhat reduced but still attractive provided these inputs are applied efficiently. "

Management & Maintenance

The ATA directly oversees and operates all aspects of the 8028 program; managing calls through Ethiotelecom, the national telecom service. The service has been improved several times since its inception in 2014, including the reduction of first-time user registration requirements, adding caller IDs for better analysis, and rotating contentoptions seasonally. PAD serves as a consultant to the system.

Evaluation & Assessment

Improvements have been made based on analysis of the call logs, which represent a rich data source of user behavior within the system. The ATA runs a real-time dashboard that publicly reports current outbound and inbound calls, as well as the call service's legacy usage data. Several A/B tests, the process of testing two user interfacings on different segments of users simultaneously, are used to assess the impact of potential improvements.

The Problem

How Can We Increase User Performance ?

In order to realize 8028's full value, users must be able to reliably, quickly, and accurately access their desired content. User's only direct engagement with the 8028 system is when they call in for content (besides occasional external surveys), so the system must accommodate users of all skill levels, including low-literate individuals and users with limited digital experience, through its design or system functions.

These goals are contrasted by the magnitude and variety of options provided throughout the system, which necessitates a complex interface. In order to be accessible to its 5.6 million users, 8028 is further challenged to ensure its experience is equitable to all users.

This section of our study examines how the call flow process and system can be redesigned to improve user performance, based on the metrics defined below, through the intentional structuring of the 8028 system. We look to answer the question:

Are there system-wide changes that can broadly improve performance, or are there opportunities to provide more customized and supportive content?

How Do We Measure Success?

Table 1 8028 KPIs.

Key Performance Metrics	***Impact to User***	***Measurement***	***System Implications***
Completion	Does the user reach content?	Content reached per call/callers who reach content in their lifetimes	Reaching any content is likely to increase users' understanding of the system and likelihood to call again.
Accuracy	Does the user reach the content they want?	Are selection patterns rational?	Users can only apply advice heard if relevant and they understand its context.
Speed	How long did it take the user to reach content? A second piece?	Time to content and time on call.	The user investment, in terms of time on call, cannot be greater than the value of the experience.
Equity	Is the caller experience equitable?	Required actions needed are the same for all callers. Low barriers to entry and robust user support.	Can all farmers who want to use the service, use the service?
Education	Do farmers take knowledge away from their interaction?	Users listening to the right, or relevant, content. Users learn how to use the system better.	Users can only gain benefit from the service if they can still implement it upon ending the call.

Recommendations Summary

Restructure Call Flow

How

- Eleveate the crop menu to the top menu
- Eliminate HHI/Rainfed menu from call flow

Why

- Users seem to best understand the crop menu.
- Correctly answering the top menu puts users on a positive path dependency, ensuring they reach content that is relevant.

Methods

- Literature Review
- Data Analysis
- Elite Interviews
- Operational Analysis

Utilize Predictive Algorithm

How

- Predict weak users
- Understand differences between users

Why

- Weak users can be identified and diverted to supportive interventions or tutorials.
- Content predicted to be of interest to users can be elevated for easier access.

Methods

- Machine Learning
- Theoretical Applications
- Proof of Concept Testing

Project Scope

Returning Users

The scope of our quantitative analysis is limited to relatively high frequency and returning users. This is partially due to the limitations of our data; our sample was not random, but a sample of convenience consisting of strong users. The median user in our sample logged 7 calls in 2020 while only 24 calls were from first-time users. The average user in our sample made their first call into the system over 3 years ago. It is important to note, however, that our findings suggest even these "strong" users struggle to navigate the system. We believe that solutions that help experienced users avoid confusion should improve novice caller outcomes. Especially since our user-facing recommendations mostly involve simplification of the system.

All quantitative analysis was based on users' previous use of the system, and so by definition is only representative of those with experience in the system.

Table 2: Call Log Data Sample Key Caller Characteristics.

Sample - Sample of Callers and their Call Histories for the Year 2020		
1000 Unique Callers	16,445 Calls	42% of Callers Reach Content
97.6% Return Callers	Median User Made 7 Calls	Mean User Made First Call 3.25 Years Ago (from 10/2020)

Regardless, improvements to the experience of returning users represents the most impactful way to improve the system. 8028 has received over 46 million calls from 5.6 million callers (PAD, 2020), implying that the average user calls into the system 8.65 times. From a per-call perspective, this means that 88.4% of calls into the system are from returning users. In addition, for a user to get the full value of the 8028, they are likely to call frequently over the year, as the seasons compel them to work on different aspects of the agricultural process. Finally, users have been known to call into the system over and over again in a single day to either access different pieces of content or because they are having trouble navigating the call architecture*.

Table 3 The 8028 Farmers' Hotline Usage Overview.

Population - Cumulative System Usage from 2014 - 2021		
5,469,219 Unique Callers	47,720,554 Calls	37% of Callers Reach Content
88.4% Return Callers	Mean User made 8.65 Calls Median User unknown	Mean User First Call Unknown

*Information told to us by PAD Ethiopia team members

Literature Review-Highlights

A more extensive literature review is undertaken in the appendix, but for the sake of brevity, a couple of key points for our analysis are summarized below.

Availability vs. Accessibility

Throughout the scholarship on Information and Communications Technology (ICT) in the development context, there is a contrast between available and accessible services. Many authors, like Akers, Tadesee, Batchelor, and Scott, highlight that simply because mobile services are made available to remote communities, it does not mean that they have sufficient access to the services. Issues like awareness, value, the gender divide, the power dynamics of phone ownership, and capacity issues like literacy were repeatedly named as barriers to making use of services even if they were technically available to marginalized groups (Akers, 2016; Tadasee, 2105; Batchelor and Scott, 2014).

Information as a Resource

Regardless of their findings, scholars' characterization of ICT development work as solving a failure in the information market or economy was an important realization to our work. Authors like Bugger and Hellstrom use this to illustrate the incentives from the user perspective and the importance of content having true value (Bugger, 2011; Hellstrom, 2010).

The Immensity of the ICT Development Space

The size and diversity of ICT interventions and research for development are overwhelming. Interventions using ICT technology are being utilized for every purpose, using every delivery method, and for every different type of phone user imaginable (Nakasone, 2014). Its popularity in both public and private development speaks to its immense potential, but it also becomes apparent that these interventions are mostly targeting the same groups. In assessing 8028, it was helpful to contextualize the system as part of an ICT revolution that is very much happening to, rather than with, the most marginalized groups in Africa. Authors like Nakasone (2014 & 2016) and Akers (2016) note this.

Trust

Within the exponential growing sources of information for farmers, the idea of trust in the information provided was consistently identified as critical to achieving real impact. Users must trust the agricultural content they receive in order to implement said content because implementation has a direct effect on user's livelihoods.

User-Centered Design

A burgeoning field in the ICT world in general, and increasingly in its uses for development, is the need to focus more on user-centered designs, although considerable disagreement about its exact meaning pervades (Gulliksen, 2003).

User Participation and Co-design

There is a clear need for user participation in the design of systems (Lindblom, 2017). Other scholars and practitioners have begun to consider these key user-centered considerations in ways particularly relevant to the 8028 system. In building an IVR system similar to 8028 in nearby Tanzania, a user-centered approach was adopted to make key optimizations. Key insights like using audio recordings from people with local accents and the potential to use the popularity of the radio by matching information delivery style came from users, while researchers found they were able to form an effective relationship through the design process with farmers (Ortiz-Crespo, 2020). It should be noted, however, that after initial interest the system was not successful (Ortiz-Crespo, 2020). In the appendix, we present some opportunities to incorporate these ideas into the 8028 system in the future.

Oral Users

Authors like Sherwani and Ali highlight the importance of understanding users not only as literate or illiterate, but as coming from an oral or literary tradition. Those from oral traditions tend to have different ways of structuring and understanding information (Sherwani and Ali, 2009). To connect with these users, interfaces must be designed specifically for them. They highlight the need to display information with relevant examples, create narratives and redundancies, and avoid abstract categories to support oral-grounded interfaces (Sherwani and Ali, 2009). They also point to individuals from an oral tradition placing more weight on the source of information than those from literary societies(Sherwani and Ali, 2009).

Problems with Touch-tone

In testing a real-world application of IVR technology in rural Uganda, Lerer et. al were able to provide more detailed performance statistics from 150 teachers from rural areas. While generally finding a low success rate with the system, an issue of particular interest was that most users could not operate the touch-tone or touchtone-voice hybrid interface without prior training. They consistently received feedback that it felt like an obligation to press buttons and suggestions that a voice interface might be more effective (Lerer et. al, 2010)

Stakeholder Analysis

The Different Users of 8028

Our report will only focus on returning users, but Table 4 provides a brief stakeholder analysis performed on other specific sets of users, as well as some potential ideas for these groups that we did not have a chance to fully consider. While we did not look to specifically improve the experience of these groups, we did consider any harm or detriment our recommendations may have on them. Even though they are outside of our data sample, we are confident that our qualitative analysis is applicable to all groups and hopeful that our quantitative analysis is externally valid.

Table 4: Stakeholder Analysis.

8028 User	***Returning User***	***First Time Users***	***Community Users***	***Non-Users***
Composition	-The average user calls into the system 8.65 times -88.4% of calls into the system are from returning users.	- 2.4% of sample and 11.6% of population -May be inexperienced with interfaces.	-Some accounts are hypothesized to be community phones, where the same phone is used by a larger community.	-Farmers who might gain utility from the system aren't using it as they are unaware, unwilling, or unable (due to not having a phone).
Experience	-Must progress through a minimum of 4 menus before to reach the content. -While only reaching content on 41% of calls, the high number of calls per user indicates a real demand for 8028 service.	-Required to register a profile regarding farmer demographics and region. -Experience only customized by Language	-Experience may be similar to returning users, but some users may find difficulties. -Community phone usage presents a terrific opportunity to promote positive spillover effects within the community.	-ATA has run effective radio campaigns, implying a high awareness of the service. -Complexity of content and navigation is hypothesized to be difficult to describe in sufficient detail, so spillover effects to groups are limited.
Challenges	-If a user does not understand how to navigate the system, how can they learn?	-Balancing the need to record important information without alienating users.	-How to identify and quantify community phone usage w/o affecting the general user?	-How to reach groups indirectly who do not have phones?
Needs	- Correctly navigate to desired content or learn how to. -Reach content that's relevant to them	-Most users call 8028 multiple times (Min of 3 calls in sample); understanding the system is important.	-Community implies that many callers do not own their own phone and so may have less digital literacy.	-Opportunity
Possible Intervention	-Call Flow Restructure and Predictive Algorithm	-Elimination of registration process and intro of tutorial.	-Targeted analysis to better understand this group.	

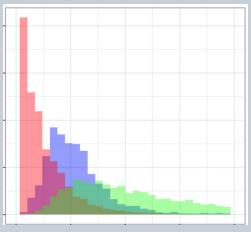
Methodology Summary

Our Approach

To analyze issues within the call flow and refine our recommendations, we examined patterns and descriptive statistics, conducted expert interviews with scholars in the field, performed a literature review, had frequent discussions with our client, and performed operational analysis. Given the global pandemic and the general difficulties of engagement with groups so remotely located, we did not have any direct contact users, limiting our ability to truly understand the impact on their lives. We recognize that our project would have been greatly enhanced by speaking with end-users. It was simply not possible. By mixing qualitative and quantitative analysis, we attempt to overcome these limitations by finding solutions that have a strong theoretical basis and robust quantitave support.

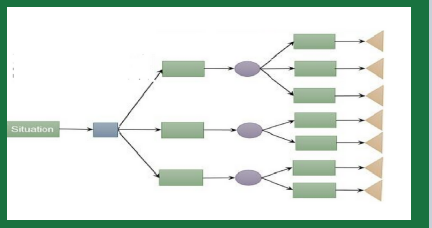
Elite Interviews

We spoke with two experts in the ICT for development field to better contextualize our findings



Data Analysis

Identified key selection and performance metrics from a sample call log of 16,445 Calls



Operations Analysis

We looked at the system in a new way; reimagining the call flow using operations analysis.

Elite Interviews

To better understand the previous work in support of rural communities, and with the 8028 system itself, we conducted a series of interviews with elite experts in the field. Based on our review of key scholar's work and discussions with our client, we reached out to ten scholars, receiving responses from Indrani Medhi Theis and Ofir Reich. Our interviews with both scholars provided us with considerable guidance for our own analysis. We present a summary of our discussions below, focused on a key insight from each.

Interviews were conducted over Zoom because of the coronavirus pandemic. Participants were informed of the goal of both the report and the interviews. They spent approximately 60 minutes each discussing both their published work as well as anecdotal insights that they had acquired over years as practitioners of tech interventions for development.

User Interface Best Practices

Our discussion with Ms. Theis focused on best practices in user interface design, particularly for low-literate users. The importance on **simplicity** and **fewer options** were key insights.

The Importance of the Crop Menu

Researchers and Mr. Reich discussed the seeming **importance of the crop menu** as the best understood menu.

Elite Interviews

User Interface Design

Low-literate users remain a key audience ICT for development scheme's and this remains true in Ethiopia, where literacy is around 51% (World Bank, 2020) . While literacy is particularly important for SMS and text-based applications, and so where most research is focused, there are still important lessons for design through a better understanding of the unique characteristics of low-literacy users that create barriers to navigating the system.

Indrani Medhi Thies, a researcher at Microsoft, has worked and published extensively on low-literate users' ability to navigate systems using text, graphical, and voice interfaces. While her work does not specifically focus on IVR systems, her focus on how the presentation of options affects user understanding was relevant for 8028's call flow. Key insights from our conversation with Ms. Theis are highlighted below.

Listing vs. Hierarchical Interfaces

- Participants more easily navigate listing UIs, that immediately provide final options in lists (where a user selects next if they don't find their option), than UIs that rely on hierarchical structures with nested options.

- While studying graphical interfaces, Ms. Thies found participants were significantly faster (25 vs. 65 seconds) and seemed to better understand system features, “On the phone list UI, we observed that even though the items were spread across 7 pages and not all of the items were visible at once, participants did not hesitate to move about quickly through the pages. Once they realized that the task item was not available on the first screen, unlike in the hierarchy UI.... they would promptly press the arrow (“forward” in this case) to move to the next page to find the task item” (Thies, 2018)

Fewer Options per Menu

- Ms. Thies anecdotally highlighted the memory strain that some IVR systems require. The more options per page, especially complex concepts, the greater burden on the user.

- This has important implications between the intention of the designer and the experience of the user. A system that seems simple from a design perspective because the information in condensed may be actually harder to navigate.

*All insights are from personal Interview on zoom with Ms. Thies on 1/22/2021

Elite Interviews

Importance of the Crop Menu

A team of top PAD scholars are in the process of publishing their own study on the 8028 hotline, which served as an important starting point for our research. They performed exploratory and motivated research before running several A/B tests. Outside of identifying some key characteristics of the 8028 user base, which have been referenced above, a key finding of their work was the demand for the service that existed, implying that most failures to reach content were due to the user having challenges with the system (Torsten, 2021).

In their efforts to simplify the system and make it more accessible, they attempted four experiments; removing registration for first time users, removing the option to save crops to receive push notifications, playing a help message if no option is selected and creating pauses between the language options. They found that removing the registration and adding pauses to be successful, with removing the crop profile to be moderately successful, in getting users to more content. Perhaps more interesting than their experimental findings, however, was that users seem to understand the crop menu better than others (Torsten, 2021). We discussed key findings with a co-author of the paper, Ofir Reich.

Selection Pattern Implications on Understanding

A prevailing finding throughout different analyses of the 8028 system was the over-selection of the first option on each page. While some over-selection of earlier options is to be expected, the distribution has been seen to be too consistent to ignore. We explore this issue in greater detail in our data analysis.

The **crop menu**, however, displays more typical selection patterns, indicated a **higher selection accuracy** for this page.

*All insights are from personal Interview on zoom with Mr. Reich on 2/4/2021

Operations Analysis

The 8028 Process

To better understand the 8028 organization, and how it affects performance, we tried to find unique frameworks to explain the system. We considered how the people and technologies behind the system affect the relationship users have with 8028 and the possibilities for system development. Finally, we looked at the 8028 call flow structure as a more general process akin to customers waiting on a series of lines or a doctor making a diagnosis based on a series of tests.

Call Flow: A Complex Web of Options

8028's basic call flow relies on a complex web of menus (see the call modules in the appendix). Users select their farming style (irrigation, rainfed, etc.), agricultural process of interest (pre-planting, planting, etc.), a planting process sub-menu, and finally crop type. After hearing any piece of content, users can replay the same piece of content, choosing a different crop, or return to the main menu*.

We can reimagine the call flow as a series of tests; in order for the user to get utility from the system, they must pass through each menu of options, which acts as a process node. After moving through this process, users receive a payout from the content they listen to, the utility of which is based on its relevance to their work. While a user will only get full value from the content they desired to hear, they may receive some utility from content that they can use in the future.

The current call flow works on a hierarchical structure, where a user's selection on each menu affects the subsequent menus and options seen. Nodes have different levels of complexity and requirements on the user; some require knowledge of technical terms (although likely more common to farmers) while others tax memory (with 10+ options to consider).

Menu Recordings & Navigation

8028's menus form the backbone of the call system, allowing for the diversity of content offered but also serving as the main barrier to reaching content for the user. Menus vary from five to twelve options, with some options that are not fully related sharing the same menu (i.e. rainfed farming techniques with livestock rearing and COVID-19 information), which may confuse a user listening in quick succession. These menus can also be seen in the appendix.

Navigating the system relies on a touch-tone interface, requiring users to listen to a recorded message and press the numerical button corresponding to their preferred choice.

*8028 often refers to irrigation as HouseHold Irrigation (HHI) and Rainfed as Rain. We will use them interchangeably.

Operations Analysis

Most experienced phone users will have experience with these systems, but it may be less apparent for those who have not used IVR systems previously (Lerer, 2010).

Navigation may be easier moving forward in the system than backward. There is not a consistent escape or help button for users to press if confused. On some pages, the * button returns the user to the main menu, while on other menus its brings the user to edit their registration profile. After listening to a piece of content, users are brought to a page that allows them to re-enter the call flow at any of the menus.

The Process of Impact/Application

The more general process of farmer understanding and application of content was also an area of key interest, although harder to fully envision given our lack of in-person engagement. Below we attempt to provide the key process questions related to user behavior after hearing content in the form of a Yes/No questionnaire. For those continuing to design the system, we believe that this could be a helpful tool to measure user acceptance. The more affirmative answers to these question, the more the system should be designed to maximize the diversity and breadth of content on the 8028. The more negative answers, the more ease of retrieval and simplicity should be prioritized.

Designer Questionnaire

Users Can: *Application*	**Users Can:** *Navigation*
Apply content after hearing it once? Effectively teach another farmer the content?	Navigate back to a help or top menu reliably? Navigate back to the same content reliably?
Listen to the content and then apply the content 1 hour later? 2 hours? A week? a month?	Navigate back to content by barging after retrieving it multiple times?
Describe the difference between different pieces of content in the same subspace?	Stay on a call without interruptions reliably for multiple minutes (signal issues)?
Explain why a piece of content is relevant to themselves but not a farmer of the same crop?	Call back to the system often at a later date (payment or technical issues)?

Top Menu Importance

An immediate insight that arose from our analysis was the importance of the top menu (the first menu in the call system). Since the system essentially flows one way (returning to higher levels is difficult), the top menu will define the entire user experience and subsequent options. The current top menu, which gives users the choice between HHI or rainfed farming styles, livestock, etc. has mutually exclusive subsequent choices, giving it even greater importance, as it is impossible to reach content relevant to the crop or planting process of interest (unless through complete luck) if the user fails to correctly navigate the top menu. The importance of the top menu has key implications for the optimal structure of the 8028 system.

Placing the Best Understood Menu First

If there is reason to believe that there are substantial differences in comprehension of disparate menus, it is important that the best understood is the first choice the user makes. Maximizing the accuracy of the top menu will ensure that the most users reach a second menu with their option of interest on it. The user's correct response to the top menu gives designers another key piece of data, allowing for better division of subspaces to maximize the amount of relevant information.

Matching Farmer's First-Order Associate for Calling

When a farmer calls into the service, what is their primary reason for calling? While they may be looking for help on harrowing Maize, would they describe themselves as primarily needing help with harrowing or Maize? Would they characterize their farming (I am a "blank" farmer) with the crop or the planting process? Matching selection order to association order, and particularly the top menu with the caller's first-order association, should improve both understanding of specific information and how information relates to other content. This might increase accuracy and relevant information reached by farmers.

Determining with certainty callers' first-order associations without direct engagement is simply not possible. Based on our conversations with administrators, and the patterns in the data, we hypothesize crops to be this top order association (biased by our own American perception of agriculture), but this should be robustness checked through future research. Two potential studies are presented in the appendix for those who, after the coronavirus pandemic, are better placed to directly engage with the user. We will also further support this hypothesis through our data analysis.

Reduction of Options Heard

In thinking about the call flow as a process, we identified that even without any intervention or change from users, simply reducing the number of options heard will improve accuracy and performance. If we think about each option as a test, with the user having to correctly answer select or not select, fewer questions will inherently lead to better results (since only wrong answers are penalized), not to mention quicker completion times. While this may seem apparent, there are actually several ways that the number of options heard per call can be reduced.

The HHI/Rainfed Options are not necessary

We identified that having users select irrigation or rainfed farming is not necessary to direct them to the appropriate content.

There is no content within the system that is only differentiated by these two options, so a user's choice on all other menus implies their selection of rainfed or irrigation. While including it in a separate menu makes sense from a design perspective, as it allows for a call structure which branches outward, it hampers the user experience.

Empty options

There are also opportunities to reduce the number of empty options, or options that a farmer must listen to where content is missing for their crop. An example of an empty option can be seen in figure 1. Removing empty options by having users select crop first is a clear benefit of this strategy. For example, a Maize farmer would only hear arrive at the top menu and only hear 3 options, all of which there was content about Maize.

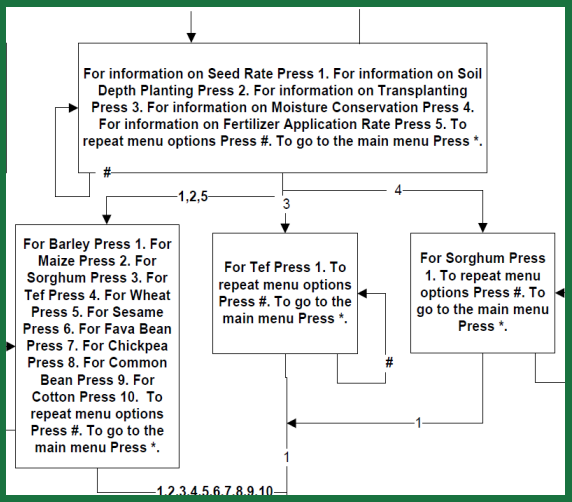


Figure 1. Empty Options

A part of the call flow within the rainfed menu. Because options 3 and 4 are only applicable to Tef and Sorghum farmers, respectively, these options are empty for other crop farmers. If these menus were reversed, the subspaces could by customized by crop so that all users would only hear options that are relevant.

Opportunities for Further Optimization

Saved Hot Keys

Allowing users to save certain content on their registration, which then can be found on a hot key on the top menu would drastically reduce the number of options heard by most returning users. High priority or seasonal content could be similarly elevated.

Linked or Directly played Content

Another possibility is for some content to be played together in succession instead of presented as separate menus. Figure 2 shows a part of the call flow menu for rain-fed content which is already three user selections into the system. In this situation, it might be more user-friendly to eliminate the binary choice and simply play both pieces of content, so farmers spend more time listening to content instead of menus. Users can then have the option to have the content played in the opposite order for future calls. Tactics like this could be particularly effective in simplifying the call flow design should the crop menu be elevated to the top menu.

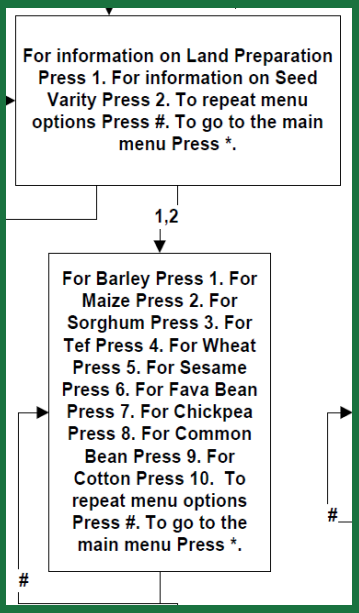


Figure 2. Linked Content

A piece of the call flow from the rainfed module that could be linked.

A Chance for Greater Assessment

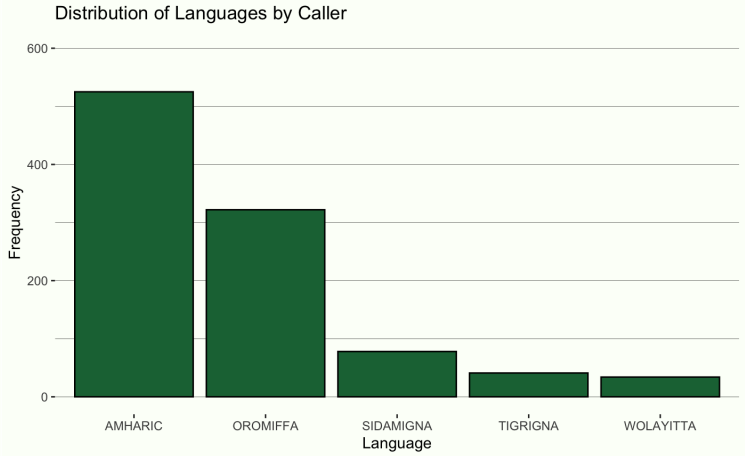
Another potential benefit of elevating the crop menus would be the ability to perform a more fundamental assessment of the different performance aspects of the 8028 system. Many of the key performance questions that have been discussed, like the pressing one phenomenon, are difficult to measure due to a lack of outcome variables and variation. This type of massive change within the system may allow for more intricate Regression Discontinuity/DID study methods.

Data Analysis

Further Data Demographics

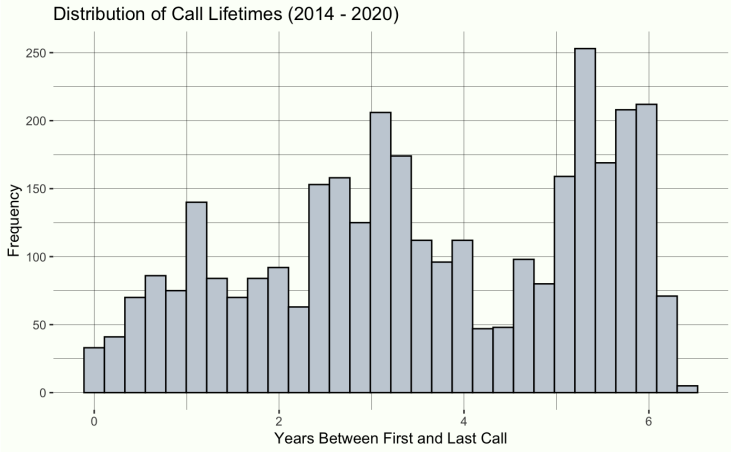
Beyond the 16,445 calls, 1000 individual callers, and 24 first-time callers, we have a relatively limited set of demographic information. The majority of users use only two languages, which are consistent across all calls for a given user. This is relatively consistent with the national spread of languages across the country (Oromiffa and Amharic accounting for more than 50% of the population)*. It is important to note that distributions of behavior (i.e., use of the 8028 system) were fairly consistent across all languages - showing a similar distribution to graphs presented below (Encyclopædia Britannica, 2020).

Figure 3.



The average user in our sample joined the system more than three years ago and many have used the system since its close to its inception. If these "experienced" users are having trouble navigating the system, it is likely that less experienced users outside our sample would have even more difficulty. Please note that all calls in our sample were from 2020, but the 8028 system repeatedly logs the time stamps for the user's first-ever call into the system. The only information we have prior to 2020 on callers is their first-call date and time.

Figure 4.

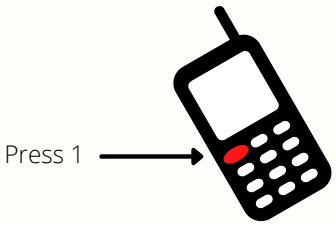


Our access to user information was limited, more of which could have improved our analysis and better targeted our work, but we believe that PAD could add this richness to our work, and incorporate more variables such as location and survey information into implementation.

*For those familiar with Ethiopian Languages, Oromiffa is synonymous with Oromo.

Data Analysis

Pressing One: A Phenomenon



Staff and researchers consistently noted that users seem to engage in a strange phenomenon of over-selecting the first option in a menu. We also identified this issue throughout our analysis and of the sample data. The architecture of the call system has changed along with the order of options that are listed since 2014. No matter what is listed first (and usually 2nd), this button almost always gets selected with the highest frequency. For example, we can see the distributions below in figures 5a, 5b, and 5b in the first two levels of menus a caller accesses within the call system. In each case, the option corresponding to the number 1 is pressed the most. This trend continues across dozens of other menus in the call system too, which we have chosen not to share for the sake of brevity.

Figure 5a.

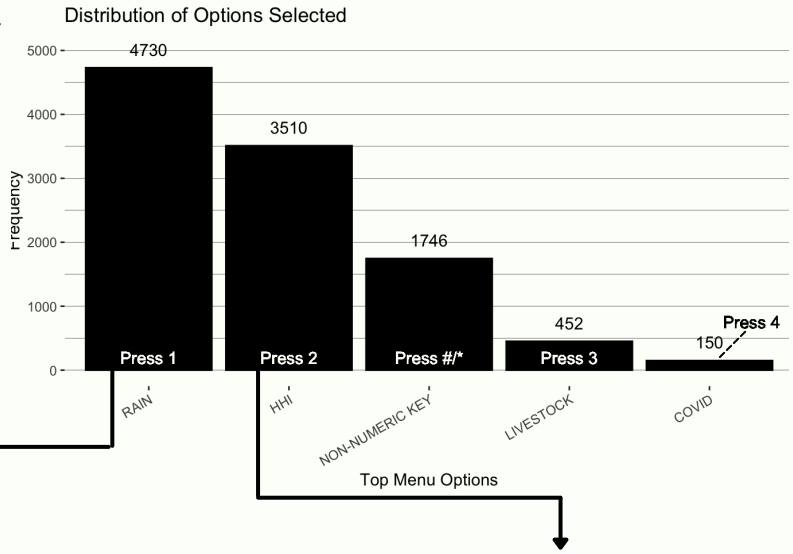


Figure 5b.

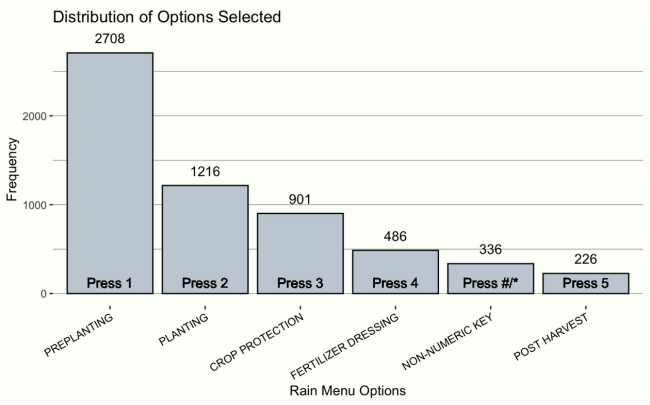
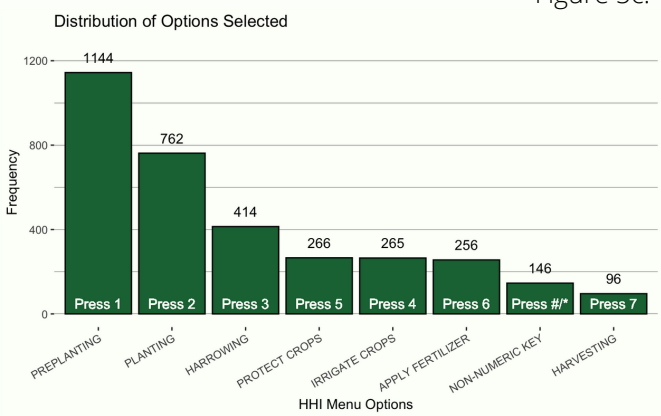


Figure 5c.



Data Analysis

We cannot definitively ground our belief, but we hypothesize users who did not know which button to press would default to the first one in their confusion or frustration, in consensus with the other staff we discussed this issue with. Terms such as “Fertilizer Dressing” and “Harrowing and Earthing” (example menu options) could be confusing - we have no idea what they mean. We do not know how these terms translate into the local languages nor do we know the agricultural vocabularies of all farmers. It is possible that these are common agricultural terms that all farmers are familiar with, but we have one major piece of evidence that suggests that this is not the case*.

The only menus in the entire system (from our sample) with a significant amount of access that did not follow the trend of the option corresponding to one being pressed the most were the crop menus. As can be seen in figures 6a and 6b, there does not appear to be any visible pattern corresponding to the option number in the selection choice.

Figure 6a.

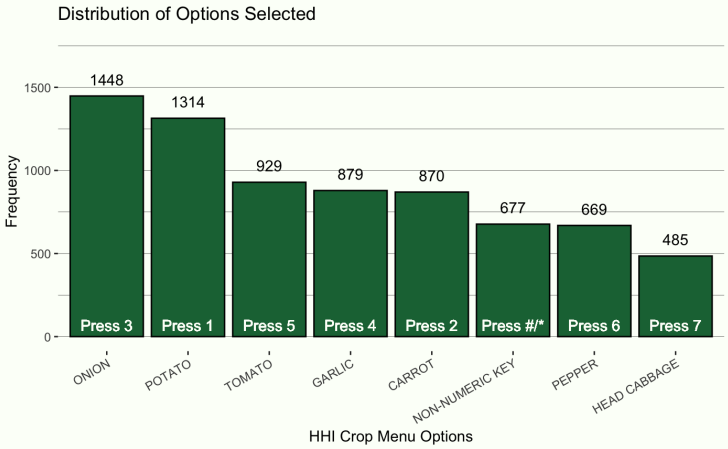
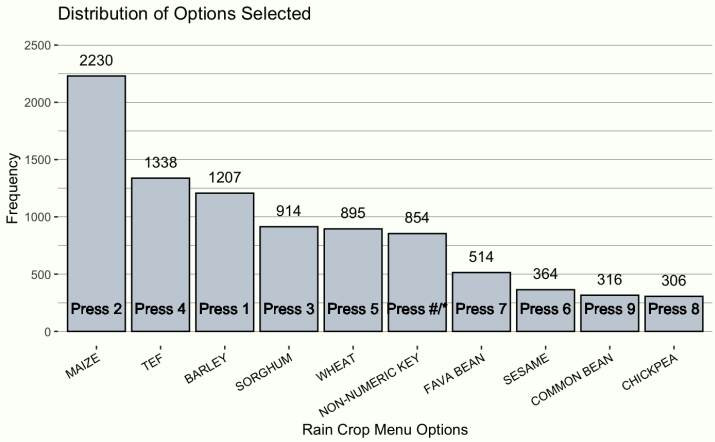


Figure 6b.



As previously mentioned, we think that users generally press one because they are confused. It would make sense that a farmer who was unfamiliar with an IVR system would more easily know what "Wheat" means as opposed to "Fertilizer Dressing". The call system is incredibly complex (see appendix AA) and it seems quite unlikely that these would be the only two menus to break the patterns for no reason. Although it could be due to statistical noise and we cannot prove the cause in the difference in the crop menu number selection compared to the other menus without a controlled experiment, this is an important factor in our assessment of how the call architecture *should* be designed.

*This is a perfect example of where being able to speak with real users would have been highly beneficial to our work. If we could determine what users consistently understand, we could more effectively make recommendations.

Key Recommendation

Elevate Crop Selection to the Top Menu

The key finding from our call-flow analysis is the need to elevate the crop menu to the top menu, integrating many of our smaller insights and best exploiting the patterns observed. Placing the crop menu first will allow users to first engage with the easiest to understand menu, building confidence in navigating the system from the start. It reduces the potential options in subsequent menus, mostly eliminating empty options. This would both reduce the numbers of options most users hear and increase the probability of reaching content of some relevance. Most importantly, when users choose their crop, they are implicitly choosing their agricultural method (irrigation vs. rainfed), eliminating the need for that menu and letting the system better guide the user.

Elevating the crop menu to the top menu will not only solve a lot of the issues identified in our analysis, but also create a call flow structure that is better suited to the incremental improvements covered. We hypothesize that having crop menus first matches the user's first association with the issue they are calling about, although this needs to be researched further.

Implementation

Implementing this change may require considerable effort to manage the organizational and user friction, but from a technical standpoint, should not be overly complex as the changes would be similar to those performed previously for A/B tests and experiments. Some potential barriers are detailed below, with strategies to overcome them. Ultimately, it requires buy-in from PAD, and the ATA, and then work by the developers who code the 8028 system.

Too many Crops

The largest barrier to implementation is the number of crops, which has the potential to make the first menu more complex. Instead of four options (with the addition of COVID-19 & Livestock Rearing), there are as many as 20 crops, meaning that the first menu could have as many as twenty options. There are two potential strategies to overcome this:

Elevate Crop Menu

One strategy would present the crop menu through a listing interface. Users could select up to 5 options or select "next" to hear the next five options (the system would move to the next page if no selection was chosen).

While this may require the users to listen to more options, our analysis suggests that such an interface may ultimately be easier to use for those struggling with the system (Thies, 2018). Additionally, crops can be prioritized (placed on an earlier menu) based on popularity, season, or user targeting, ensuring that most users find their crop of interest within these first five options.

Alternatively, users could be first prompted to select a crop supra-group (i.e. grain crops, root vegetables, etc.) before choosing their crop of interest. While this will not reduce the number of menus, it will create a more intuitive and closer connection between the top and sub-menu, perhaps leading to a higher accuracy rate. By carefully choosing groupings, with ample input from users, this process would also reduce the complexity of the system. A potential experiment involving users to help identify appropriate supra-groups is presented in the appendix.

User Change Acceptance

Another barrier may be user resistance to change, as many are at least somewhat familiar with the current structure. Given the seemingly low acceptance rate of the current structure among users, however, this may be a smaller effect. This could still be mitigated by making the change incrementally, potentially by allowing users to indicate interest in trying a new call flow structure.

Better Understand User Preferences

Developing a better understanding of user preferences is critical to the success of any intervention, but in particular to the creation of a successful top menu that includes crops (and livestock). Currently, the system is structured to maximize navigation for a user who would be interested in all the content, but this is clearly not realistic. A key question that we always asked, but could never determine, was what the actual percentage of the information available does each, or for that manner any user, really use. The narrower a farmer's interest, the greater the impact a restructuring would have on performance, especially if farmers tended to be interested in only a few crops. Qualitative interviewing would again help shed light on this, but the ATA could also build data by asking farmers to input the crops they grow occasionally (possibly with a lottery for mobile credit as a possible incentive).

Challenges and Limitations

Failure to Reach Weakest Users

While restructuring the call flow looks to improve performance by incrementally improving the experience for all users, it does not successfully target weak users with greater support, which we believe is also necessary to truly improve system performance in terms of equity. It is possible that these changes will actually be harder for the weakest users. Given that the vast majority of our calls are from returning users who have used the system, and our access to only the performance of relatively strong users, reaching this group was always going to be a challenge. This recommendation would be best implemented as a complement to a specific intervention for the weakest users in the system.

Unclear True Impact

Our analysis can only comment on the amount of content heard, not whether the information will be ultimately better understood and implemented by the user. While we rely on the work of many other studies to assess this connection, it is ultimately unclear how changing the process of retrieving information will affect user's understanding and usage of the content. We present a possible experiment to test this in the appendix.

A Path to Potential Full Optimization

Pages 56 and 57 of the appendix contain the entire call flow for the rainfed and irrigation content module. Upon reaching the highest menu showed in each module, users would have already gone through one menu. As can be seen, most farmers will listen to at least four menus before reaching content, although some have three menus. By adopting the optimization strategies discussed in our analysis, the follow number of menus to content could be achieved for each option.

Table 5-Potential Gains for Implementing Call-Flow Restructure.

Content Module	*Process*	*Menus to Content*	*Average/Range of Options to Content*	*Extra Reduction Method*	*Pieces of Content Listened To*	*Other*
Rainfed	Pre-Planting	2 (All Crops)	9-13	LC	2	
	Planting	3 (All Crops)	11-16	EO	1	Content 3x Redux**
	Crop Protection	3 (All Crops)	12-16	LC	2	Content 3x Redux**
	Fertilizer Use	2 (All Crops)	12	-	1	Content 3x Redux**
	Post Harvest	3 (All Crops)	13-16	LC	1	
HHI	Pre-Planting	2 (All Crops)	9-13	LC	2	
	Planting + Transplanting	3 (All Crops)	11-16	-	1	Content 3x Redux**
	Harrowing and Earthing	2 (All Crops)	11	-	1	
	Irrigation Crops	Varies*	-	-	-	
	Protecting Crops	2 (All Crops)	13	LC	2	
	Applying Fertilizer	2 (All Crops)	14	-	1	
	Harvesting + Processing	3 (All Crops)	15	-	1	Content 3x Redux**

LC = Linked Content; **EO** = Empty Options
***Variant due to selection complications and not meaningfully affected by changes
****Content could be tripled up to reduce on menu

Implementing these strategies would reduce the number of menus users are forced to listen to by one for all farmers, by two for certain farmers, and even theoretically by three depending on a user's willingness to listen to three pieces of linked content. Inherently, reducing the number of decisions the user must make without increasing the complexity of those decisions should lead to better performance without any change in user competence. The effect on the users' time is harder to measure (based on the information researchers have available), but even if users' call time in this system increases (listening to two pieces of content instead of one), the time would be spent on content, not navigation.

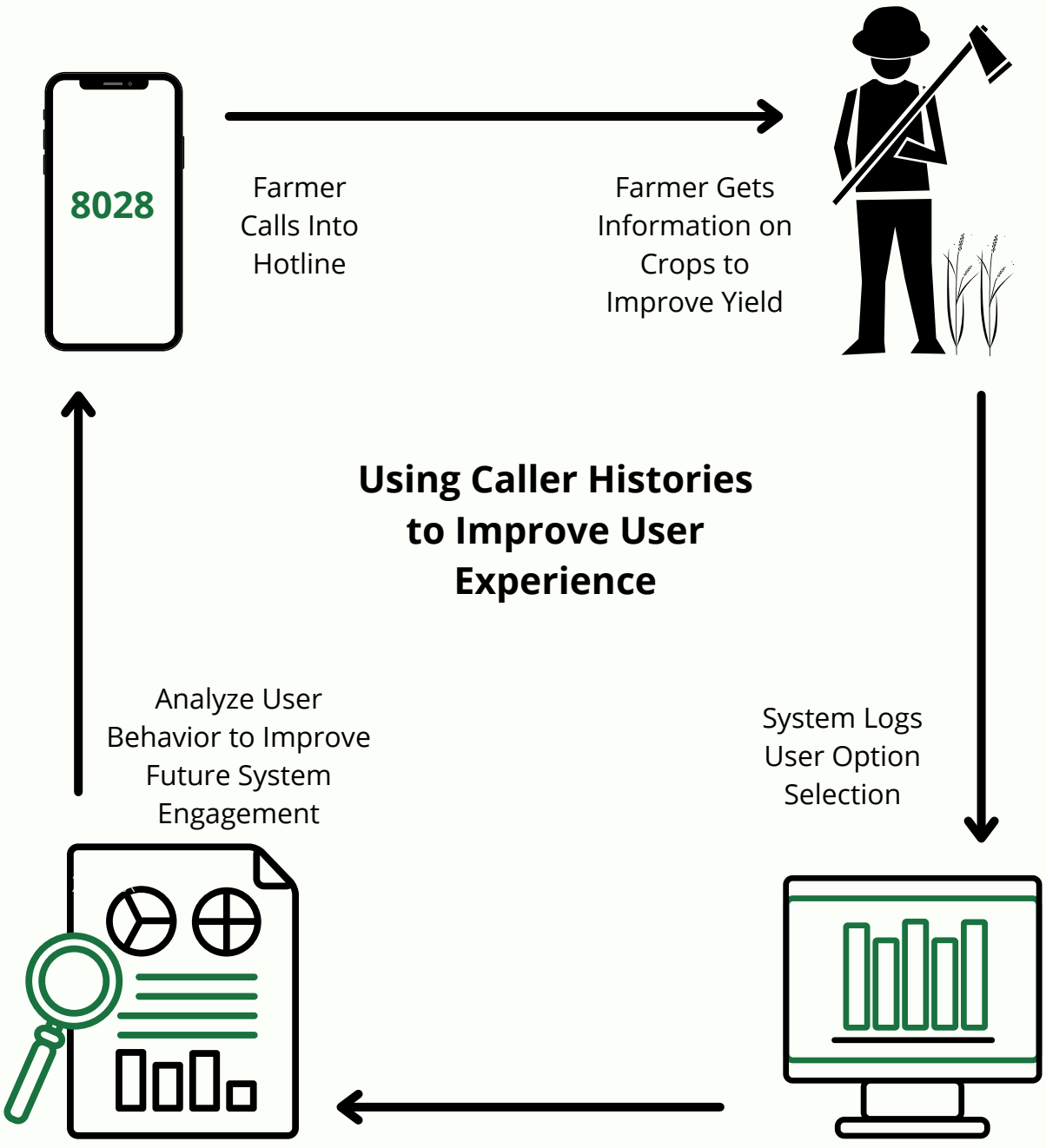
Putting It All Together

The effect on the number of options heard is also more complex to understand due to its dependence on crop of interest, and its importance more clear. For simplicity, we present an average range, assuming that the user listened to the average number of crop options (8) on the top menu. This is a flawed assumption for several reasons, but also highlights an issue already within the system, inequality for different crop farmers. Already, farmers whose crops are later in the options face this issue, but the restructuring now makes it an issue for all farmers.

Some strategies for reducing the number of crops that need to be listened to were presented above and we are confident that the system could be designed with customizable features so that the vast majority of callers hear their crop within the first five options.

Predicive Algorithm

Predictive Algorithm



Our Approach

Our second recommendation looks to better identify and assist weak users, providing a more targeted approach through the use of a predictive algorithm. Motivated by underlying differences in the behavior of users who do and do not reach content, we constructed a model trained on our data sample. There are still several technical and implementation challenges that need to be solved, so we present our model only as a proof of concept. While the model is not sufficiently realized to implement directly into the 8028 system at this moment, we think it is very possible PAD and the ATA could and should do so in the future.

Even with these limitations, our model predicted whether a call will reach content with 65% accuracy. We were forced to make some simplifying assumptions, but we believe that this performance is a strong indicator that there is value in this approach, especially with the addition of other explanatory parameters like demographic information from the caller registration profile. In particular, we believe that a model trained on a considerably larger sample of the caller population should perform better (i.e., the 5.6 million callers that have used this system as opposed to our sample of one thousand callers).

To build the prediction algorithm, we applied several machine learning techniques to the call log data. To prepare the data for training and testing, we cleaned it and created visuals to better understand its underlying characteristics. We considered traditional applications of machine learning and the specific context to come up with practical applications for the model. Finally, we developed the next steps for the project based on conversations with the PAD data science team.

The Task: What to Predict

The raw data, as can be seen in the appendix on page 52, had no clear outcome variable to utilize for classification, requiring considerable time and effort just to engineer usable features from the data. Another key challenge was the system's inconsistent method for logging actions, which made it difficult to determine which calls were successful. For example, both of the following observations in the data signify the caller reaching content:

- CONTENT PLAYED - PLOT PREPARATION - HEAD CABBAGE - OROMIFFA
- CONTENT PLAYED - LAND PREPARATION - BARLEY - SIDAMIGNA

In each of these cases, a different sub-menu, crop, and language were selected. To consider these observations together, we had to string match all instances of the term "CONTENT PLAYED" within a call. To filter for multiple pieces of content being listened to on the same call, we created a binary outcome variable. We defined a positive outcome as reaching content at least once within a given call and a negative outcome as never reaching content in a call.

Model Preparation

In order to prevent our model from "cheating" and looking into the future to predict call outcomes (from the perspective of the call in question), we used the outcome of the most recent call as the dependent variable, only using calls that predate it to train and make predictions.

A note on First-time Callers

The 24 First-time callers within our data sample were excluded from our model; as they had no previous behavior in the system, the model had no information with which to consider their performance (the first-time caller experience is also different because of profile registration). We believe, however, that by using descriptive statistics of first-time callers and their outcomes, PAD and the ATA could create a suggestive profile of first-time users from their average behavior. Given that only 37% of calls reach content in the 8028 system population, with an even lower percentage for weak callers (strong users bring this average up), it is reasonable to expect that the vast majority of first-time callers will fail to reach content (based on discussions with PAD employees) - information that could be used for a model if PAD pursues the idea further.

An Advanced Approach for PAD

After multiple meetings with PAD's data science team, we also have a proposition for a more comprehensive and challenging model that would be outside the scope of our work, but could improve the accuracy of predictions. We would slightly change the training of the model to use the outcome of every single call within the data set as our dependent variables (instead of just the last call), and then use the previous calls from that point backward as the independent variables. In other words, if an individual has made 6 calls into the system, our current model uses the outcome of the 6th call (content or no content) as the predicted variable, and the average information from the first five calls to train on. In this advanced approach. We would use each of the previous call outcomes independently as data points. For the 5th call outcome, we would use the first 4 calls as predictors. Then for the 4th call, we would use the first 3 calls as predictors and so on. This a much more time-consuming data engineering project, but we believe that with more time and effort, it would greatly improve model performance.

Model

Model Selection

The task of selecting a machine learning model is not simple and we ended up testing many before settling on a Random Forest Classifier. Our model ultimately predicts with 65% accuracy in our sample group of hold-out calls (a test set) will or will not reach any content when calling into the system

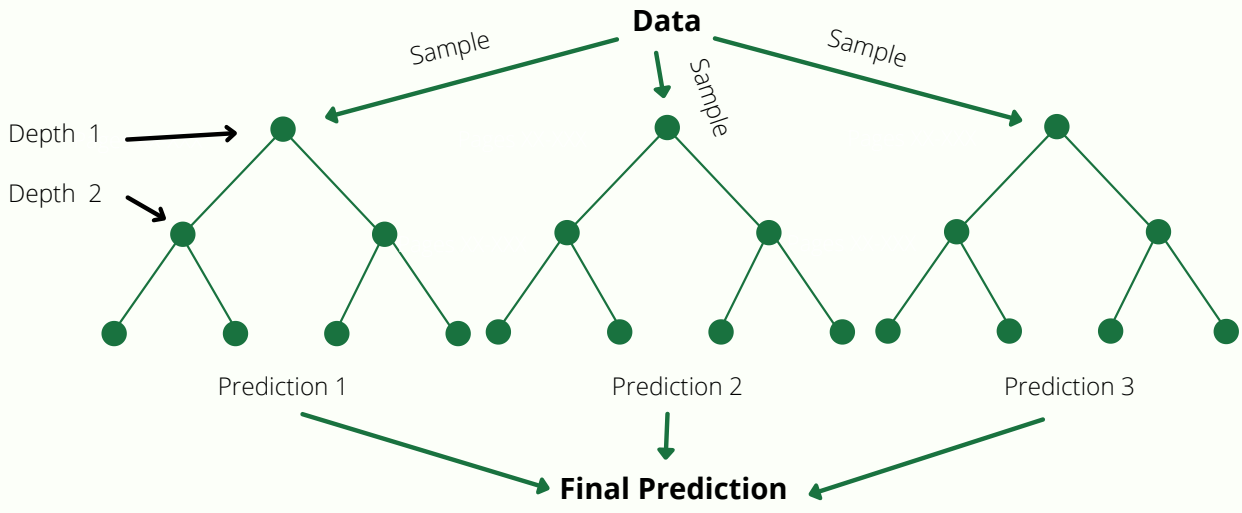
Table 6: Models Compared Throughout Our Analysis

Models Tested		
Logistic Regression	Lasso-like Logistic Regression	Ridge-like Logistic Regression
K-Nearest Neighbors Classifier	Random Forest Classifier	Decision Tree Ensemble Classifier
Feed-Forward Neural Network	Decision Tree Classifier	Logistic Generalized Additive Model

For a policymaker who is smart, but not well versed in machine learning, a random forest uses a decision tree structure to predict an outcome by splitting observations into bins with observations where the outcome is known (the training data). The algorithm will then predict the class that is the majority in the bin that observation was placed in. A unique feature of random forest is to restrict the importance of any given factor (so as to avoid overfitting). The process is repeated a predetermined number of times with random samples with replacement and variables from the data. The final prediction is the average across repetitions.

In our case, the tree depth was 7 and the number of repeats was 75. We used cross-validation to determine these as the optimal prediction hyperparameters. For reference, an example model with depth equal to 2 and 3 repeats can be seen in figure 7. Although this explanation is an oversimplification (there are other "attributes" to the model), it should be helpful for those with less modeling experience.

Figure 7: Example of A Simplified Random Forest.



Model & Features

Our model was fairly balanced in predicting the different outcomes. It predicts those who reach content in reality with 67% accuracy and those who do not with 64% accuracy. Hopefully, if extrapolated onto a larger and more comprehensive data set, this means that the model has a minimal bias in terms of prioritizing one group over the other. We could also tweak the model to better predict one group over the other if PAD was so inclined. This process would adjust the threshold probability at which the algorithm predicts users will reach content. We believe that it is more important to accurately predict those who won't reach content (minimize type 1 error), but we want to further discuss prioritizing certain callers with PAD and the ATA*.

Features

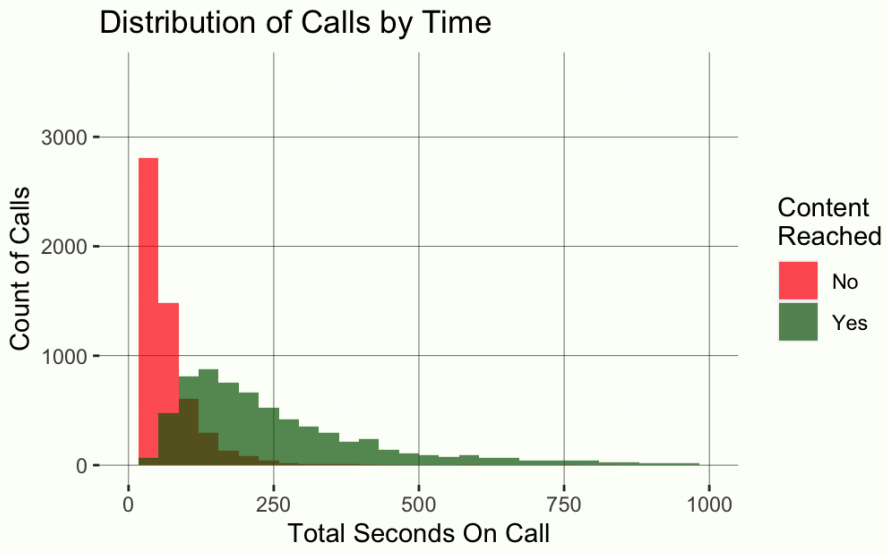
The model type was only part of the problem in creating good predictions. We also had to determine which independent variables would give us the highest accuracy. We tested all models, but ultimately ended up using the features listed in figure 6.

Table 7: Features used in the prediction model.

Predictors				
Average Length of Previous Calls	Average Number of Button Presses in Previous Calls	Percentage of Previous Calls reaching Content	Number of Previous Calls	Content Ever Reached in Previous Calls
Language	Year of First Call	Month of First Call	Year of Current Call	Month of Current Call

We found that the most impactful features (those that most affected model performance and accuracy) were the average length of previous calls and the percentage of previous calls that reach content (using feature importance). This is not surprising as there is a clear difference in the amount of time spent on calls between those that do and do not reach content. Additionally, it makes sense that individuals that have reached content in the past are more likely to reach it in the future (and vice versa).

Figure 8.



*This parameter adjustment comes as a direct tradeoff. The minimizing of type 1 error would be at the expense of type 2 error. Outside of overall accuracy, the optimal cutoff is a subjective decision.

Features

The data is time-stamped with written descriptions of each action in the call log architecture made during a call (as can be seen in the appendix on page 52). Along with engineering our outcome variable, we had to engineer all the other covariates besides language by pulling out information for the descriptions of user selections. This was one of the most time-consuming and intellectually challenging aspects of our analysis.

Further Feature Engineering Opportunities

Because engineering features was so time-consuming and difficult, we were limited in our capacity to take full advantage of all the information within the data. For example, we did not attempt to see if callers who tried to access certain crops or parts of the call structure, in general, were more or less likely to reach content. Other potential sources of data mining that could be used in the future with more time and resources to improve model accuracy include: call time of day, time since the previous call, number of replays selected in previous calls, and content amount accessed in the lifetime of calls. These factors all have the potential to be correlated with reaching content as they could be representative of behavior from a more experienced user or a user who simply has a better understanding of IVR.

Additional Data

Due to ethical concerns that we agree with, we did not have access to phone numbers nor profile registration data of callers into the 8028 system. However, this data does exist and includes both personal demographic information and location data; this additional information has the potential to yet further improve the accuracy of the models. The 8028 system records location data (from user input) at the equivalent of state, county, and district levels.

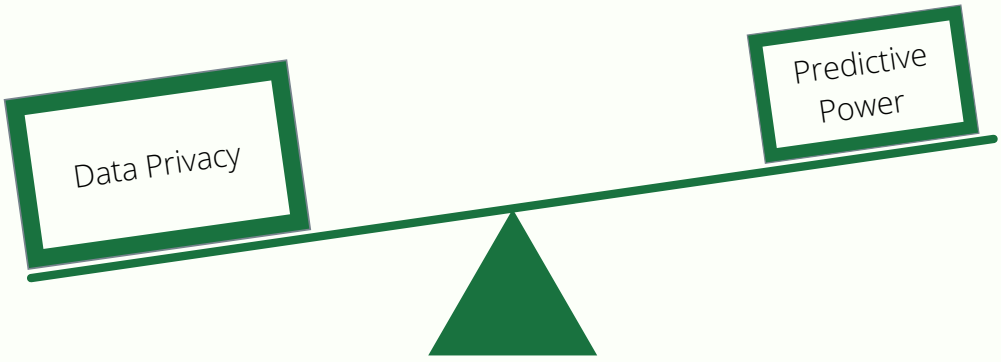
We suspect that location data in particular would be a powerful explanatory variable of the likelihood a caller reaches content. Different locations within any country have concentrations of higher education and wealth. Those people might be more familiar with the concept of an IVR system and technology in general, which would make the 8028 system easier to navigate. Additionally, the ATA has changed the profile registration system in the past and could do so again in the future. They have the ability to customize the data which they try to collect from callers. With some targeted thought and consideration, this data collection could be leveraged to explicitly gather data that could improve accuracy for a prediction model. For example, the registration process could ask callers if they have ever used a call system before - an experience that is likely valuable for successfully navigating the system in the future. Spillover effects, like how they were introduced to the system, could also provide powerful explanatory information.

Ethical Considerations

Imperfect Ethical Tradeoffs

In the previous sections, we described how we created our model, attempted to optimize the model, and laid out some ideas for future improvement for PAD and the ATA to consider. However, we did not discuss what *should* (in a meta sense) be done in terms of creating a model that uses caller data to predict caller behavior. If implemented, we would like the accuracy of our model to be as high as possible. We believe it can be a tool to help Ethiopian farmers gather important information to improve their crop yields, which could further decrease food insecurity and increase economic returns. However, using more personal data, such as phone numbers and location, comes at a cost to privacy.

The implications of using big data in any capacity on a personal level without clear and informed consent are frightening and must be fully considered before any implementation is taken. Some users of the 8028 may not even have the current capacity to provide informed consent around their data. Without a more specific effort to inform farmers, simply asking "can we use all your personal information in ways that you do not fully understand?" seems ethically untenable. Neither public nor private overreach of data privacy is acceptable in our opinions. Please note, this does not mean PAD or the ATA have infringed on anyone's privacy (they haven't). We simply feel obligated to discuss this tradeoff as we are the ones proposing the idea of a predictive model. In general, greater support for digital literacy and empowerment is needed to truly undertake this idea ethically.



We do not have the answer for what data *should* be used if the ATA implements a predictive model. However, given the aforementioned problems, we think that they should be cautious in using personal information that is not volunteered (but that can be identified by the phone number). We recommend that there be a ethical process or standard established to regulate what data would be acceptable for the 8028 to use. There is a balance to be struck between privacy and predictive power as the model does have the potential to help people with more information. With a slightly larger focus on privacy, the model could be implemented ethically and produce useful outcomes.

Prediction Summary

Ultimately, this predictive model is a proof of concept, which requires further time, effort, and a complete data set to fully realize. With the data available, we engineered both an outcome variable and covariates to predict whether users in the sample will reach content with reasonably high accuracy. We are confident that, using all of the data from the 5.6 million callers (as opposed to the sample of 1000 callers provided to us) and our methodology, the ATA could improve the model accuracy substantially. Training the model on an exponentially larger data set should alone create a much more powerful model.

PAD and the ATA can also engineer more features from the data and use additional data that was not accessible to us while balancing ethical responsibilities with accuracy. As the PAD data scientists are able to improve the model performance, the model can be applied in a variety of different ways to improve system performance, which leads us to the following question:

What can we do with the model to help farmers when using the 8028 call system?

1. IDENTIFY KEY DIFFERENCES BETWEEN GROUPS

2. CHECK ROBUSTNESS OF EXPERIMENTS

3. ONE-TIME INTERVENTION

4. REAL-TIME INTERVENTION

5. PREDICT USER PREFERENCES

Given the hurdles to actually using our model or creating a new model for this purpose, these applications should be considered an "extended implementation." They are two steps removed from the current state of affairs, but they should offer further reasons for PAD and the ATA to consider creating a usable prediction model. In relation to our key performance indicators, a working and implemented model could help users quickly get relevant information from the system with a focus on improving equity by targeting gains towards weaker users.

Potential Applications

These potential applications are not mutually exclusive. For example, both a one-time intervention and real-time intervention could be used together - improving both understanding of and efficiency in the system. Additionally, we believe that none of these interventions are invasive on farmer privacy (assuming ethical data collection) and would simply improve the user experience with the system.

1. Identify Key Differences Between Groups

Description:
Instead of trying to actively adjust the system for a current call in real-time, the algorithm can be used to understand the differences between users who will and will not reach content. We can introduce these users' demographic information from the registration profile or other sources to determine if there are groups that need to be targeted. Including demographic information would also allow us to better understand within-group variation in success.

Implementation:
The implementation would come with a successful model. Once the model is working, anyone could use statistical software to analyze the data and its predictions.

2. Robustness Check on Experiments

Description:
The prediction model could be used to provide a basic analysis of experiments, identifying if it increases callers' likelihood of reaching content or creating pools for experiments (i.e., experiments for predicted "strong" or experiments for predicted "weak" users)*. This might be helpful when A/B tests are not feasible or for exogenous effects to the system.

Implementation:
An analyst would have to select the proper outputs in order to create data groupings.

3. One-Time intervention

Description
The model could have the most impact as a way to better target interventions on weak users to improve performance. The model could be run asynchronously to develop a pool of weak users for a pre-planned intervention. This could vary from general IVR education to focus group creation to assist weak users.

*"Strong" being those whom the model predicts will reach content and "weak" being the users whom the model predicts will not reach content.

Potential Applications

Implementation
Implementation for this application is two-fold. The first stage simply involves setting the "cutoff" to determine groups of interest for an intervention. The second stage is more variable and complex. It could involve an extra text message or call to a user, or the logistical organization of a focus group.

4. Real-Time Intervention

Description
Real-time interventions could take two forms. First, if the model predicts that a user will have trouble reaching content, the system could provide an additional menu that guides the user through an extremely simple tutorial. Even if the model falsely identified a strong user as weak, they should have minimal difficulty navigating out of this additional short menu. We believe the benefits for a weak user (the vast majority) would greatly outweigh the detriment of a false negative (assuming "positive" is a strong user). Alternatively, the system could put weak users in a different architecture that was either far simpler than the main call path or that flashes seasonally relevant content to the farmer. Given that weak callers often receive no information when using the system unsuccessfully, giving them anything could be an improvement.

Implementation
In order to implement a real-time intervention, PAD and the ATA would need an adaptive system that could use the prediction model to "tag" a user into a group to see a specific architecture. Then, the ATA would have to design these alternate call-path architectures for the different groups. This would require significant software investment (both time and money), but could greatly improve the system for predicted "weak" users as callers would be appropriately funneled through menus that would maximize the probability that they get useful content.

5. Predict User Preferences

Description
If the model performed better in implementation, with more parameters and data, it could be also used to predict user preference like crop or agricultural process of interest, and elevate these options to the top menu. This could also be used as a way to target users for outreach programs or new content marketing.

Implementation
This application would require a new model that predicted the preferred type of content as opposed to whether a user reaches content. Similar to a real-time intervention, additional call-path architectures would be required to show predicted preferred content that could adapt based on the predicted preferences.

Methodology-Outside of Approach

A Sample of Convenience

The data provided from the call log was not random, but rather a sample of convenience drawn from a group of strong users who had used the system at least once in 2020. This may have affected our analysis in several ways, and certainly could have made our findings biased against groups of weaker users.

We believe that the effect of these biases may be limited in our analysis of high-frequency and return users, whom we believe are over-represented in this group. Given that these users are not random because of how recently they called into the system, they may also be more representative of the current 8028 hotline experience than a random sample taken from seven years of calls. These users also better represent the experience with the system as currently constructed, as more substantial changes to the system occurred in previous years (particularly in how the data is logged).

Areas of Bias

It is important to recognize that our analysis is susceptible to considerable personal bias. Both researchers are white men, who, due to the coronavirus pandemic and tight timelines for this project, did not have a chance to meaningfully engage with Ethiopian culture nor the people we hope to impact within this study. This drastically limited our ability to understand how local culture affects users' interaction with the system. To combat this, we tried to keep our analysis as culturally neutral as possible. How Ethiopian farmers associate these different agricultural concepts was one area that we felt particularly limited, and so we strongly recommended further engagement with the end-user before implementation of our recommendations.

Justification of Sources

To provide grounding for our paper, we attempted to use a diverse range of sources from all information available. The World Bank keeps an updated document on ICT for development which served as a source and its references an excellent starting place. We searched through the Harvard Library, The Web of Science, and Google Scholar, picking articles that were highly cited and, where possible, peer-reviewed; snowballing our sources based on their citations. How we used this process to find elite interviews was consistent with this method.

Conclusion

The 8028 hotline represents an important resource to complement traditional agricultural extension work in Ethiopia. The interest and demand in the program are exemplified in the willingness of farmers to repeatedly use the system despite a low rate of reaching content. While more work must be done to assess the impact of the advice provided, increasing the usability of the service will undoubtedly increase its impact, as measured by the number of users who reach relevant content.

This report presents two strategies for increasing the usability of the service, which could be used individually or as a connected approach. Given the current levels of user ability, we have recommended a new call flow structure to guide users to their desired content. We also present a predictive algorithm that could be used to identify and support weak users. We further detail possible next steps toward implementation for both recommendations, while detailing some incremental improvements that can be explored in the short run.

The success of programs like the 8028 hotline are critical for technology to be used as a tool to promote equality. It provides a resource to people in areas traditionally unreachable, and so under-supported, from support programs. By increasing its usability based on this report's recommendation, we believe 8028 can yet further realize its incredible impact on Ethiopian agricultural development by improving crop yields, lowering levels of food insecurity, and increasing economic productivity.

Bibliography

Aker, Jenny C, Ghosh, Ishita, & Burrell, Jenna. (2016). The promise (and pitfalls) of ICT for agriculture initiatives. Agricultural Economics, 47(S1), 35-48.

Jock R. Anderson, Gershon Feder, Chapter 44 Agricultural Extension, Editor(s): R. Evenson, P. Pingali, Handbook of Agricultural Economics, Elsevier, Volume 3, 2007, Pages 2343-2378, ISSN 1574-0072, ISBN 9780444518736, https://doi.org/10.1016/S1574-0072(06)03044-1. (http://www.sciencedirect.com/science/article/pii/S1574007206030441)

Barnett, Inka, et al. "Dial 'N' for Nutrition? A Landscape Analysis of What We Know about m-Nutrition, m-Agriculture and m-Development." HEART, Institute of Development Studies, 12 Dec. 2016, www.heart-resources.org/doc_lib/dial-n-nutrition-landscape-analysis-know-m-nutrition-m-agriculture-m-development/.

Batchelor, Simon, et al. "Is There a Role for Mobile to Support Sustainable Agriculture in Africa?" *2nd International Conference on ICT for Sustainability (ICT4S 2014)*, 2014, oxfamilibrary.openrepository.com/bitstream/handle/10546/325879/rr-icts-sustainable-agriculture-300714-en.pdf?sequence=1&isAllowed=y.

Berta Ortiz-Crespo, Jonathan Steinke, Carlos F. Quirós, Jeske van de Gevel, Happy Daudi, Majuto Gaspar Mgimiloko & Jacob van Etten (2020) User-centred design of a digital advisory service: enhancing public agricultural extension for sustainable intensification in Tanzania, International Journal of Agricultural Sustainability, DOI: 10.1080/14735903.2020.1720474

Boekestijn, Veerle, et al. The Role of Mobile Phone Services in Development. United Nations Policy Analysis Branch, sustainabledevelopment.un.org/content/documents/12862Policybrief_Mobile.pdf.

Bloomberg, J. (2017, December 27). Reinventing Interactive Voice Response For The Digital Era. Retrieved January 29, 2021, https://www.forbes.com/sites/jasonbloomberg/2017/12/27/reinventing-interactive-voice-response-for-the-digital-era/?sh=4244557e722e

Brugger, 2011 Brugger, F., 2011. Mobile Applications in Agriculture. Syngenta Foundation, Basel, Switzerland <http://www.syngentafoundation.org/__temp/Report_on_mAgriculture_abridged_web_version.pdf>.

Cieslik, Katarzyna, Cecchi, Francesco, Assefa Damtew, Elias, Tafesse, Shiferaw, Struik, Paul C, Lemaga, Berga, & Leeuwis, Cees. (2021). The role of ICT in collective management of public bads: The case of potato late blight in Ethiopia. World Development, 140, 105366.

Bibliography

Cole, Shawn Allen and Fernando, A., 'Mobile'izing Agricultural Advice: Technology Adoption, Diffusion, and Sustainability (June 3, 2020). Harvard Business School Finance Working Paper No. 13-047, Available at SSRN: https://ssrn.com/abstract=2179008 or http://dx.doi.org/10.2139/ssrn.2179008

Davison, R., Vogel, D., Harris, R., & Jones, N. (2017, December 05). Technology leapfrogging in developing countries – an inevitable luxury? Retrieved February 18, 2021, from https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.1681-4835.2000.tb00005.x

Fabregas, Raissa, et al. "Realizing the Potential of Digital Development: The Case of Agricultural Advice." Science, American Association for the Advancement of Science, 13 Dec. 2019, science.sciencemag.org/content/366/6471/eaay3038.

Getaw Tadesse, Godfrey Bahiigwa, Mobile Phones and Farmers' Marketing Decisions in Ethiopia, World Development, Volume 68, 2015, Pages 296-307, ISSN 0305-750X, https://doi.org/10.1016/j.worlddev.2014.12.010. (https://www.sciencedirect.com/science/article/pii/S0305750X14004082

Hellstrom, J. 2010. The Innovative Use of Mobile Applications in East Africa. Stockholm: Swedish International Development Cooperation Agency.

Jan Gulliksen, Bengt Göransson, Inger Boivie, Stefan Blomkvist, Jenny Persson & Åsa Cajander (2003) Key principles for user-centred systems design, Behaviour & Information Technology, 22:6, 397-409, DOI: 10.1080/01449290310001624329

Julius J. Okello , Oliver K. Kirui & Zachary M. Gitonga (2020) Participation in ICT-based market information projects, smallholder farmers' commercialisation, and agricultural income effects: findings from Kenya, Development in Practice, 30:8, 1043-1057, DOI: 10.1080/09614524.2020.1754340

Kraft, M. R., & Androwich, I. (2012). Interactive voice response technology: a tool for improving healthcare. NI 2012 : 11th International Congress on Nursing Informatics, June 23-27, 2012, Montreal, Canada. International Congress in Nursing Informatics (11th : 2012 : Montreal, Quebec), 2012, 224.

Kiiza, Barnabas, & Pederson, Glenn. (2012). ICT-based market information and adoption of agricultural seed technologies: Insights from Uganda. Telecommunications Policy, 36(4), 253-259.

Bibliography

Kruetz , Christian. “The Many Potential Channels for Mobile Services.” Blog, 2009, www.crisscrossed.net/2009/11/01/the-many-potential-channels-for-mobile-services/.

Lerer, Adam, Molly Ward, and Saman Amarasinghe. 2010. Evaluation of IVR data collection UIs for untrained rural users. In Proceedings of the First ACM Symposium on Computing for Development (ACM DEV '10). Association for Computing Machinery, New York, NY, USA, Article 2, 1–8. DOI:https://doi-org.ezp-prod1.hul.harvard.edu/10.1145/1926180.1926183

Lindblom, Jessica, Lundström, Christina, Ljung, Magnus, & Jonsson, Anders. (2017). Promoting sustainable intensification in precision agriculture: Review of decision support systems development and strategies. Precision Agriculture, 18(3), 309-331.

Miller, C.; Saroja, V.N. and Linder, C. (2013) ICT Uses for Inclusive Agricultural Value Chains, www.fao.org/docrep/017/aq078e/aq078e.pdf (accessed 8 November 2016)

Nakasone, Eduardo, & Torero, Maximo. (2016). A text message away: ICTs as a tool to improve food security. Agricultural Economics, 47(S1), 49-59.

Nakasone, Eduardo, Torero, Maximo, & Minten, Bart. (2014). The Power of Information: The ICT Revolution in Agricultural Development. Annual Review of Resource Economics, 6(1), 533-550.

Sherwani, J. & Ali, Nosheen & Rosé, C.P. & Rosenfeld, R.. (2009). Orality-grounded HCID: Understanding the oral user. Information Technologies & International Development. 5. 37-49.

Walter, Torsten F., et al. Using Data for Development Evidence from a Phone System for Agricultural Advice. Precision Agriculture for Development, 12 Oct. 2020, www.povertyactionlab.org/sites/default/files/research-paper/working-paper_9244_Data-for-Development-Phone-System-for-Ag-Advice_Ethiopia_Oct2020.pdf.

“Ethiopia: Ethnic Groups and Languages.” Encyclopædia Britannica, Encyclopædia Britannica, Inc., 2020, www.britannica.com/place/Ethiopia/Ethnic-groups-and-languages.

“World Bank. 2011. ICT in Agriculture : Connecting Smallholders to Knowledge, Networks, and Institutions. Washington, DC. © World Bank. https://openknowledge.worldbank.org/handle/10986/12613 License: CC BY 3.0 IGO.”

Appendix

Literature Review Continued

Context

Previous research into the specific process involved in the 8028 system is limited outside of the work PAD has done. Most scholarly work focuses more broadly on the adoption of mobile phone technologies for development, assessing both interventions that require smartphones or mobile phones. Christian Kreutz identifies these key categories for technologies used and how they affect availability/impact, which were then cited by the World Bank in their review of ICT technologies (Kreutz, 2009). In their 428 page report, the bank highlights the importance of these interventions to be a complement, not a substitute, for investment in traditional areas of development such as road, traditional education, financial services, etc. The report also demonstrates the importance of ensuring that mobile phone interventions do not further social inequality, as the owners of mobile phones tend to be richer males (World Bank, 2011).

Table A1: The ICT market by Technology.

Types and Availability of Mobile Technologies

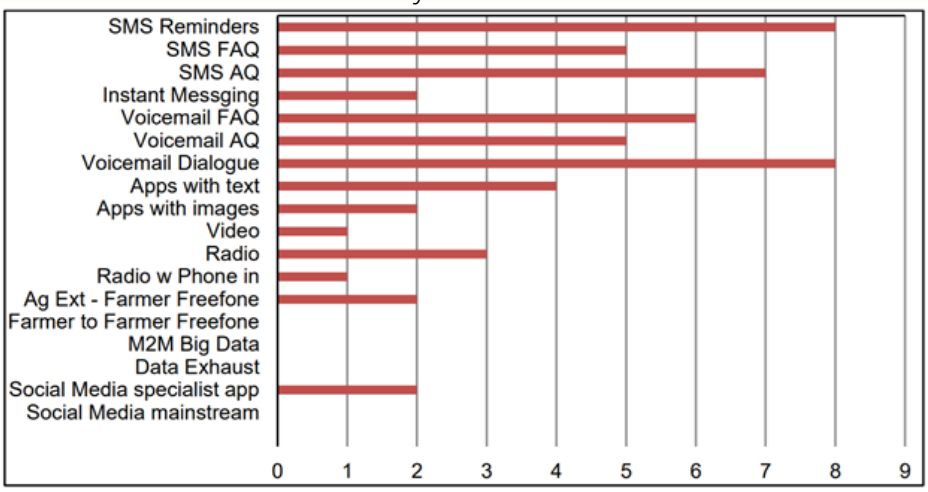
TECHNOLOGY	DESCRIPTION	AVAILABILITY
Voice	The most basic channel; avoids most literacy or linguistic barriers	Basic phones
Short Message Service (SMS)	Ubiquitous text-based messaging limited to 160 characters	Basic phones
Unstructured Supplementary Service Data (USSD)	A protocol used by Global Service for Mobile Communications (GSM) phones to communicate with the mobile network	Basic phones
Interactive Voice Response (IVR)	Computer programs that respond to the voice input of callers	Basic phones
General Packet Radio Service (GPRS)	Low bandwidth data service	Midrange phones
Software App (e.g., Java or iOS)	Preinstalled or downloaded software of varied sophistication	Midrange, but increased sophistication with smartphones
Mobile Wireless Application Protocol (WAP)	A limited manner of browsing the Internet	Midrange phones
Multimedia Messaging Service (MMS)	SMS-based technology to transmit multimedia (including images and video)	Midrange phones
Camera	For capturing still or moving images	Midrange phones
Bluetooth	Protocol for transmitting data over short distances	Midrange phones
Mobile Web	Full-fledged web access	Smart phones
Global Positioning System (GPS)	Technology allowing for location-based information	Smart phones

Source: Christian Kruetz,2009

In many ways, the dynamics underpinning 8028 are slightly more unique than these interventions in that it has some of the lowest user requirements; many interventions that do not require smartphones use message campaigns requiring literacy. In their 2016 IDS working paper, Inka Barnett, Nigel Scott et. all illustrate the current agricultural development market as a function of their information channels (Barnett, 2016).

Appendix

Figure A1. -Mobile Phone Intervention frequency by delivery method



Source: Inka Barnett, Nigel Scott 2016

The Appeal of Mobile Phones for Development

Simon Batchelor, Nigel Scott et. al point to developing countries' inability to meet an effective rate of one extension officer per 400 farmers, the focus of agriculture as a pro-poor approach from international development groups, and the knowledge-intensive nature of modern agriculture techniques as key reasons for the specific focus on exploiting the growing Telecom market for agricultural development (Batchelor, 2014). The focus of most of these efforts looking forward, especially those of the international community, is to make smartphones, broadband, or similar technologies available at a cheaper rate and to those living ever more remote (Batchelor, 2014).

In their working paper for the FAO (Food and Agricultural Organization for the United Nations), Miller et. al discuss the growth of mobile phone interventions since the technology was identified to have leapfrogging potential for development in the 1990s (Miller, 2013). They highlight the evolution from live call centers to an automated system to network connections schemes, which simply look to connect remote villages to each other to promote farmer-to-farmer education (Miller, 2013). In their report, the UN's sustainability report best identifies the appeal of mobile phone technology in development; as a tool to bring down costs and extend reach when bringing support to the most remote areas (Boekestijn, Not dated).

Design

Within the delivery of these programs, Miller identifies the direction of information flow (one way or two), the provision of tailored services, voice vs. graphics/text, and live vs. asynchronous information as key choices made. In designing these services, they recommend an approach with user input and which prioritizes simplicity, but at the same time looks to the future with a realistic understanding of how this market, and by extension, the accessibility of these technologies, will evolve in the next five to fifteen years (Miller, 2013).

Appendix

In his work on mobile applications in agriculture, Fritz Bugger makes the distinction between mLearning, where general knowledge about farmers is pushed out normally through SMS, and mFarming, where farmers receive information on their specific farming situation to influence their immediate decision-making based on their location and unique characteristics(Bugger, 2011). Bugger maps extensively almost all mAgricultural services in operation (which made an effective starting point for researchers to understand the industry). He highlights the importance of availability & accessibility, value/richness of content, and costs related to content generation and technical maintenance as they key factors in creating a sustainable and viable service (Bugger, 2011).

In his paper of the use of mobile phone applications, J. Hellstrom identifies these key applications with the agricultural space. Given the inherent complexity within the agricultural information space (Hellstrom, 2010), he highlights the importance in truly understanding how a mobile intervention will positively affect information flow, suggesting these questions as guiding principles.

Table A2-Various Roles for Mobile phones in Agriculture

GOAL	METHOD
Education and awareness	Information provided via mobile phones to farmers and extension agents about good practices, improved crop varieties, and pest or disease management.
Commodity prices and market information	Prices in regional markets to inform decision making throughout the entire agricultural process.
Data collection	Applications that collect data from large geographic regions.
Pest and disease outbreak warning and tracking	Send and receive data on outbreaks.

Source: Hellstrom 2010.

"How is information, broadly defined, produced, valued, exchanged, and consumed in a given context? What information is most valuable, to whom, under what conditions? How can mobiles and other ICTs promote innovation and efficiency in this information economy?" (Hellstrom, 2010)

Impact

Understanding the impact of mobile phone intervention is relatively complex, as often the very nature of the problem being solved, inaccessibility to the user, make testing outcomes challenging. In their review of ICT's effect on agriculture, Eduardo Nakasone highlights the demonstrable promise of ICT technologies, and their evidence of Macro impact, but highlights the micro or individual impact to be more mixed (Nakasone, 2014). Katarzyna Cieslik et. all studied the impact of ICT in our country of interest and with our population of interest, Ethiopian farmers, in preventing potato blight. Their findings "provide evidence that farmers use ICT to: (1) facilitate complex coordination, (2) establish collective norms, (3) detect and pressure 'free riders', and (4) manage reputation to increase trust." (Cieslik, 2021)

Appendix

In Uganda, Kiiza and Pederson found a positive impact for farmers who had access to ICT-based market information (Kirza, 2015). Okello et. al similarly find that ICT participation increases farmers' participation in markets and household income. (Okello, 2020)

Conversely, Getaw Tadesse and Godfrey Bahiigwa found limited impact for ICT technologies in affecting Ethiopian farmers market decisions, as they found few with these services and those who do are not successful in securing lower input prices (Tadesse, 2017). Nakasone and Torero clearly identify the promises of ICT interventions, but also highlight that they cannot be realized without solving the three barriers of valuable content, consistent connectivity, and users with sufficient capacity (Nakasone, 2016). They also compile this extremely helpful table of the different studies that they could find, illustrating both the number of different findings among the academic community.

In the Promise (and Pitfalls) of ICT for agriculture initiatives, Akers, Ghosh, and Burrell also identify the promise of ICT to provide information more timely and affordably, but point to mixed results in ICTs actual impact on welfare (Aker, 2016). They highlight the need to better understand the user and computer interaction, the goals of this report, as a key area that need to be better understood, if ICTs full potential is to be realized in development (Aker, 2016).

Table A3-Review of Studies of the Impact of ICT on Farmer's Income

Technology	Location/product	Effect (and outcome)	Study
Latin America			
Public pay phones	Peru/various crops	+16% on prices	Beuermann (2011)
Public phones	Peru/various enterprises	+13% on farm income	Chong et al. (2005)
Cell phones	Peru/various crops	+11% household consumption	Beuermann et al. (2012)
Cell phones	Peru/various crops	+11–14% on average prices	Nakasone (2016)
SMS	Colombia/various crops	No significant effect	Camacho and Conover (2011)
Africa			
Radio	Uganda/maize	+15% on prices	Svensson and Yanagizawa (2009)
Mobile phone coverage	Uganda/banana and maize	Somewhat positive relationship, but depends on distance to district center No effect for maize	Muto and Yamano (2009)
Grameen/MTN village phones	Rwanda/various products	No significant effect	Futch and McIntosh (2009)
Cell phones	Niger/cowpeas	No significant effect	Aker and Fafchamps (2010)
SMS	Ghana/maize and groundnuts	Price increases for maize (12.7%) and groundnuts (9.7%)	Courtois and Subervie (2015)
SMS	Ghana/various crops	7% price increase for yams. No effect for maize, cassava, and gari	Nyarko et al. (2013)
Asia			
Cell phones	Philippines/various crops	+11–17% on the growth rate of per capita consumption	Labonne and Chase (2009)
Cell phones	Kerala, India/fisheries	+8% in fishers' profits	Jensen (2010)
eChoupal	Madhya Pradesh, India/soybeans	+1–3% (average: 1.6%) on prices	Goyal (2010)
SMS	West Bengal, India/potatoes	No significant effect	Mitra et al. (2011)
SMS	Maharashtra, India/various products	No significant effect	Fafchamps and Minten (2012)

Source: (Nakasone, 2016)

PAD Specific Research

In its role, PAD has performed extensive research on its programs beyond their work on 8028. Pad has found their programs to be effect input choices through Random Controlled Trials in in East Africa (Fabregas, 2019) and India (Cole, 2020). Fabregas was able to use meta-analysis to find that their behavior did lead to yield increases of 4% across seven countries in Africa and India(Fabregas, 2019) .

Appendix

More Background on Agricultural Extension and IVR

The 8028 system relies exclusively on Interactive Voice Response (IVR), which allows users to navigate through a hierarchical menu structure through a touch-tone interface. IVR is based on a tone-dial methodology developed by Bell Labs and unveiled at the 1962 world's fair (Kraft 2012). It first came into considerable use in the private industry during the 1970s and 1980s, primarily as a customer service tool for businesses to reduce the need for live communication between customers and employees. Its use has expanded dramatically as technological advancements have made it cheaper to operate and more useful for customers, the key advancement being computer integration (Bloomberg, 2012).

IVR is one of many mobile-phone-based interventions to help support agricultural development as governments focus on optimizing their agricultural growth to match the needs and pressures of industrialization. Agricultural extension has a variety of structures; including "training and visit" extension, decentralized systems, "fee-for-service" and privatized extension, and farmer-field-schools; but typically focused on in-person instruction to improve technology transfer and farm management (Anderson, 2007).

IVR cannot replace traditional forms of agricultural extension or in-person support from agronomists, but can act as an effective complement for reinforcing or expanding on new techniques. IVR has the advantage of minimizing staff needs and engaging with farmers in areas previously unreachable as mobile phone penetration has leapfrogged traditional development of, electricity, and roads (Davidson, 2017). It is also more accessible to low-literate users than SMS or written mediums (including online sources), although the use of graphical interfaces has become more prevalent as smartphone penetration increases.

Effective IVR management for development must manage the complexity and customization with accessibility and ease of use. IVR's history as a product of consumerism implies that it is not as familiar to this new target audience. Additionally, IVR has been traditionally structured with the primary goal to reduce expenses, not to make information more accessible. For this reason, we may need to consider call structures considerably different from the traditional customer service model to improve its performance within a development context.

Appendix

Potential Applications of User-Centered Design

In thinking about how we might have engaged end-users of 8028 had we the opportunity, we adapted the following potential games/ideas. To allow western readers to quickly grasp the concept, we use western examples, but acknowledge the importance of changing the examples to the local context.

The Supra-Grouping Game

To the right is the food pyramid in its most current form (we believe). While we think of agriculture groups as quite static, they are often actually fluid and dependent on the perspective of the user. In this pyramid, items are roughly split into 7 categories (ignoring the extras like exercise and alcohol). Can you come up with a way to group them into six categories so that any person would know where to find each? Five categories? A similar idea could be used with farmers to see how many categories they would place the different crops into and what would they name them. This could be achieved through graphic playing cards with the crops on them, with farmers telling the group names to facilitators. Given the different thought processes of oral and literary traditions, farmers may come up with very different groups. For the record, our five categories are at the bottom of the last page.

Touch-Tone Band

During our conversation with Ms. Thies, she discussed a few ICT projects effectively using social functions to increase user engagement and comfort, like creating a function that allows users to distort and record their voice. In order to increase user comfort with the touch-tone interface, 8028 could introduce "Touch-Tone Band", a function that plays different strings of a guitar by pressing different buttons. Such a function would be fun for users while helping reinforce how the different buttons will create contrasting reactions from the phone.

Appendix

Memory Games

To better understand the role of memory in how 8028 content is applied, researchers can engage farmers in different types of memory exercises in ways that are both fun and instructive. Researchers can ask farmers to do different actions, for instance, rearrange pieces on a chessboard in a certain order, through different mediums and methods of communication. Researchers could also vary time and space, asking the farmers to wait ten minutes before performing the actions, to see how that affects implementations or whether the farmer was looking at the pieces when being given instruction. All of these could help elucidate the barriers a farmer might face from the time that he hears content from 8028 to the time he actually applies it.

High-Frequency User Conference Calls

To improve the user participation in the ongoing design of the 8028 system, we propose that 8028 staff invite their highest frequency users to regular conference calls to hear feedback and take suggestions. This would provide key feedback and give users more agency in the system while also helping researchers answer burning questions about user actions within the system.

Appendix

Ethics and Transparency Statement

This study observed all of the ethical considerations prescribed byHKS policy, our faculty advisor, and IGA Leader. While we studied a vulnerable population, we did not have any contact with end-users, nor did we at any point receive information that could identify any users in the system. All phone numbers were removed and replaced with generic caller IDS before being provided to us.

All parties with whom researchers discussed this project were informed of the purpose of the study and the outcomes of our conversations. We only discussed specific details of the service with employees of the client (PAD) and did not have any direct interaction with the ATA.

IGA Applications/Areas for Further Development

We believe that our experience with 8028 has implications across PAD projects in multiple countries. The prediction model is a technique that could be used in many different contexts, and in particular might be more effective in PAD's projects in India, where more demographic information is collected. The introduction of user-centered, and hopefully user co-designed, programs was also a consistent theme explored in this report, which seemed relevant in similar programs throughout the developing world.

Similar mobile phone interventions exist across Africa, with mHealth perhaps being a prime example in neighboring Ghana. These programs all rely on the potential for mobile phones, as a leapfrogging technology, to provide services that have been traditionally done in-person. It is perhaps too early to understand the implications of these strategies on long-term growth, but our experience has made clear that the user experience is critical in getting users to make such a large technological jump in adopting these services.

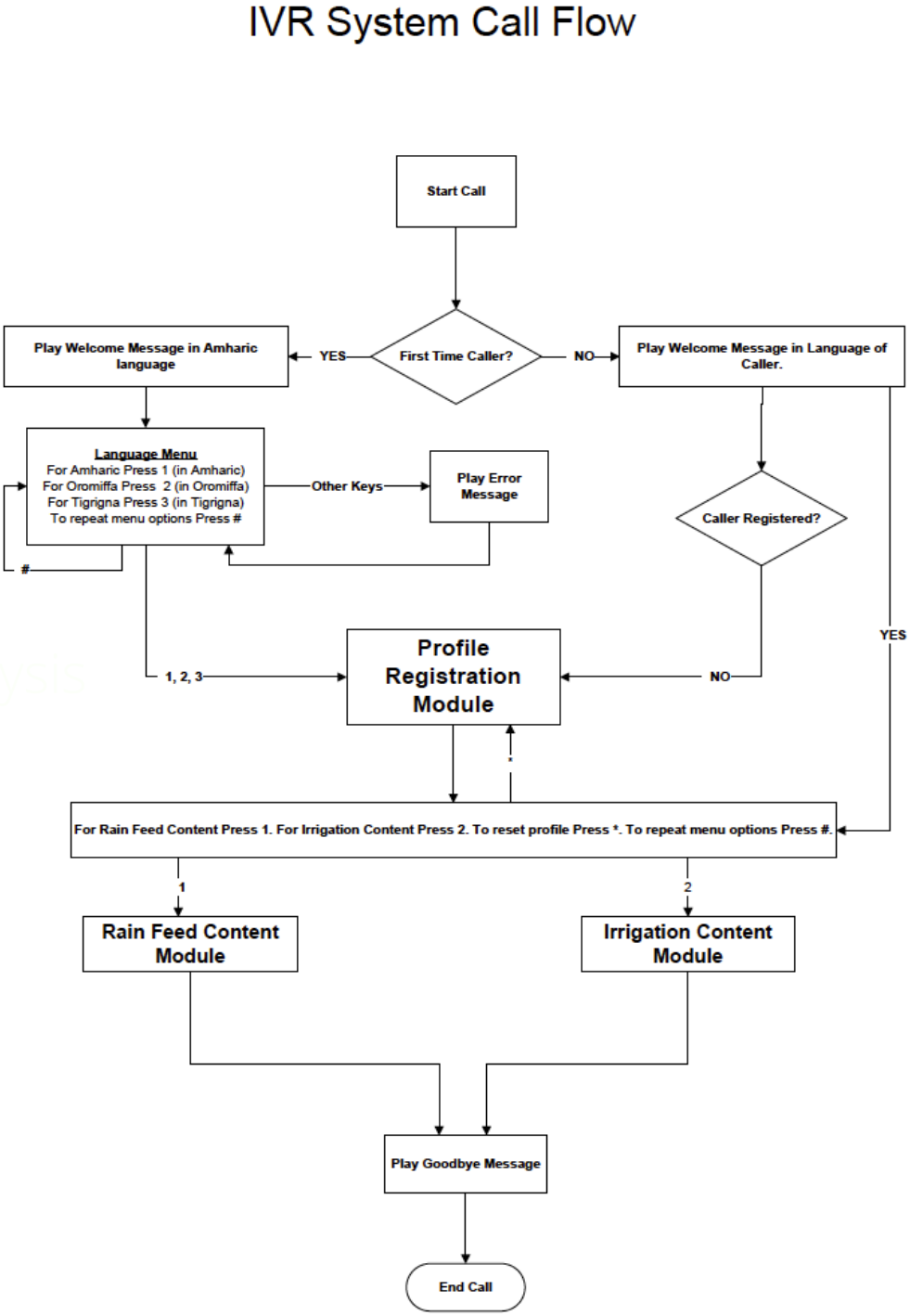
We also believe that the trade-off between system sophistication and approachability is one experienced throughout ICT for development schemes across the developing world. The level of complexity of the hotline, similar to many of its counterparts, and the number of functions is nominally impressive, but makes both management and use harder. Getting this trade-off right is specific to the program and the services it provides, but our experience highlighted the importance of customization. Complex systems that are unable to provide highly customized services to their users may create a gap between the demands and value of their service.

callerId	langId	callTime	lastCallTime	noCallsMade	noContentListened	callId	eventTime	logInfo	logInfoId
250878	2	2014-09-03 16:...	2020-02-25 20:34...	96	128	36ce39bd-94b8...	2020-01-02 08:...	CONTENT MENU – CONTENT MENU REPL...	53
250878	2	2014-09-03 16:...	2020-02-25 20:34...	96	128	36ce39bd-94b8...	2020-01-02 08:...	CONTENT MENU – GO TO MAIN MENU OP...	54
250878	2	2014-09-03 16:...	2020-02-25 20:34...	96	128	36ce39bd-94b8...	2020-01-02 08:...	TOP MENU – RAIN OPTION SELECTED	301
250878	2	2014-09-03 16:...	2020-02-25 20:34...	96	128	36ce39bd-94b8...	2020-01-02 08:...	MAIN MENU – PLANTING OPTION SELECTED	19
250878	2	2014-09-03 16:...	2020-02-25 20:34...	96	128	36ce39bd-94b8...	2020-01-02 08:...	MENU 2 – SEED RATE OPTION SELECTED	285
250878	2	2014-09-03 16:...	2020-02-25 20:34...	96	128	36ce39bd-94b8...	2020-01-02 08:...	CROP MENU – MAIZE OPTION SELECTED	35
250878	2	2014-09-03 16:...	2020-02-25 20:34...	96	128	36ce39bd-94b8...	2020-01-02 08:...	CONTENT PLAYED – SEED RATE – MAIZE –...	1000584
250878	2	2014-09-03 16:...	2020-02-25 20:34...	96	128	36ce39bd-94b8...	2020-01-02 08:...	CONTENT REPLAY REQUESTED	49
250878	2	2014-09-03 16:...	2020-02-25 20:34...	96	128	36ce39bd-94b8...	2020-01-02 08:...	CONTENT PLAYED – SEED RATE – MAIZE –...	1000584
250878	2	2014-09-03 16:...	2020-02-25 20:34...	96	128	36ce39bd-94b8...	2020-01-02 08:...	THE CALLER HUNG UP THE PHONE	32
4373509	2	2019-06-26 13:...	2020-02-13 14:28...	32	103	08ad1d6b-7a81...	2020-01-02 08:...	INCOMING CALL STARTED	162
4373509	2	2019-06-26 13:...	2020-02-13 14:28...	32	103	08ad1d6b-7a81...	2020-01-02 08:...	WELCOME MESSAGE PLAYED	1
4373509	2	2019-06-26 13:...	2020-02-13 14:28...	32	103	08ad1d6b-7a81...	2020-01-02 08:...	WOREDA MENU – INVALID KEY PRESSED	9999
4373509	2	2019-06-26 13:...	2020-02-13 14:28...	32	103	08ad1d6b-7a81...	2020-01-02 08:...	THE CALLER HUNG UP THE PHONE	32
4373509	2	2019-06-26 13:...	2020-02-13 14:28...	32	103	1e94c50a-0dc4-...	2020-01-02 08:...	INCOMING CALL STARTED	162
4373509	2	2019-06-26 13:...	2020-02-13 14:28...	32	103	1e94c50a-0dc4-...	2020-01-02 08:...	WELCOME MESSAGE PLAYED	1
4373509	2	2019-06-26 13:...	2020-02-13 14:28...	32	103	1e94c50a-0dc4-...	2020-01-02 08:...	THE CALLER HUNG UP THE PHONE	32
4784091	5	2019-12-25 12:...	2020-03-20 10:03...	199	84	e34495d1-8449...	2020-01-02 09:...	INCOMING CALL STARTED	162
4784091	5	2019-12-25 12:...	2020-03-20 10:03...	199	84	e34495d1-8449...	2020-01-02 09:...	WELCOME MESSAGE PLAYED	1
4784091	5	2019-12-25 12:...	2020-03-20 10:03...	199	84	e34495d1-8449...	2020-01-02 09:...	WOREDA MENU – INVALID KEY PRESSED	9999
4784091	5	2019-12-25 12:...	2020-03-20 10:03...	199	84	e34495d1-8449...	2020-01-02 09:...	TOP MENU – RAIN OPTION SELECTED	301
4784091	5	2019-12-25 12:...	2020-03-20 10:03...	199	84	e34495d1-8449...	2020-01-02 09:...	MAIN MENU – PRE PLANTING OPTION SEL...	15
4784091	5	2019-12-25 12:...	2020-03-20 10:03...	199	84	e34495d1-8449...	2020-01-02 09:...	MENU 1 – LAND PREPARATION OPTION SE...	280
4784091	5	2019-12-25 12:...	2020-03-20 10:03...	199	84	e34495d1-8449...	2020-01-02 09:...	CROP MENU – BARLEY OPTION SELECTED	34
4784091	5	2019-12-25 12:...	2020-03-20 10:03...	199	84	e34495d1-8449...	2020-01-02 09:...	CONTENT PLAYED – LAND PREPARATION ...	1001815
4784091	5	2019-12-25 12:...	2020-03-20 10:03...	199	84	e34495d1-8449...	2020-01-02 09:...	CONTENT MENU – GO TO MAIN MENU OP...	54
4784091	5	2019-12-25 12:...	2020-03-20 10:03...	199	84	e34495d1-8449...	2020-01-02 09:...	TOP MENU – RESET PROFILE OPTION SELE...	303
4784091	5	2019-12-25 12:...	2020-03-20 10:03...	199	84	e34495d1-8449...	2020-01-02 09:...	RESET MENU – INVALID KEY PRESSED	353
4784091	5	2019-12-25 12:...	2020-03-20 10:03...	199	84	e34495d1-8449...	2020-01-02 09:...	RESET MENU – GO TO MAIN MENU OPTIO...	352

Figure A2: Example View of Data from Call Log Files

Appendix

Figure A3: Entry point of 8028 Call Flow



Appendix

Figure A4: Profile registration for first time users

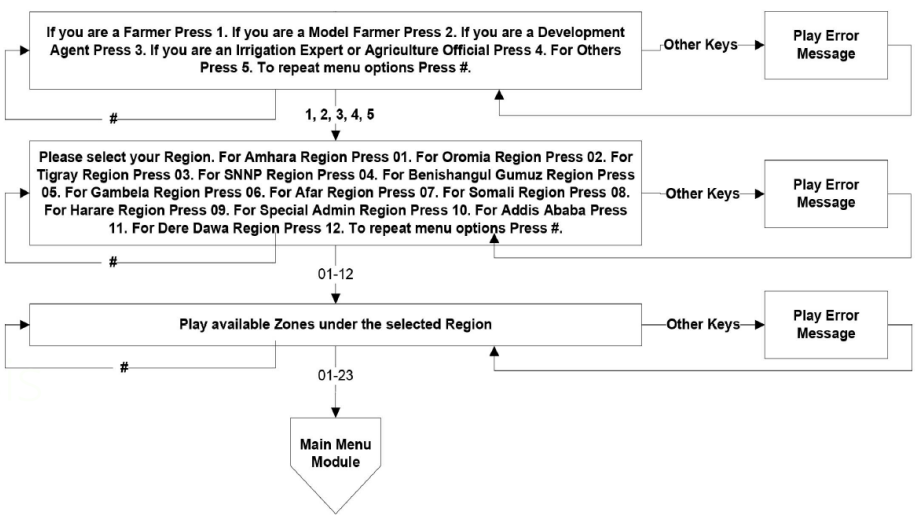


Figure A.2: Flow chart of profile registration module - part 1

Appendix

Figure A5: Profile registration for first time users

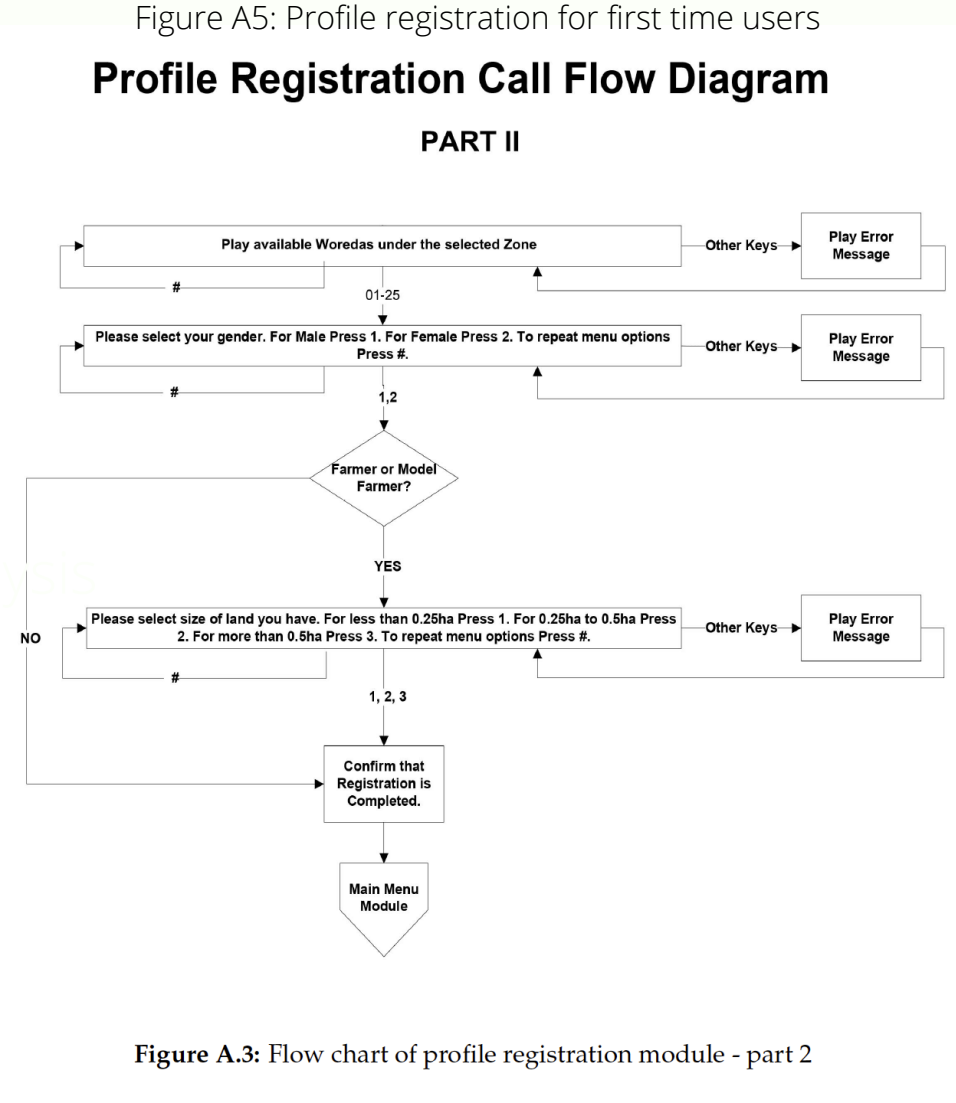


Figure A.3: Flow chart of profile registration module - part 2

Appendix

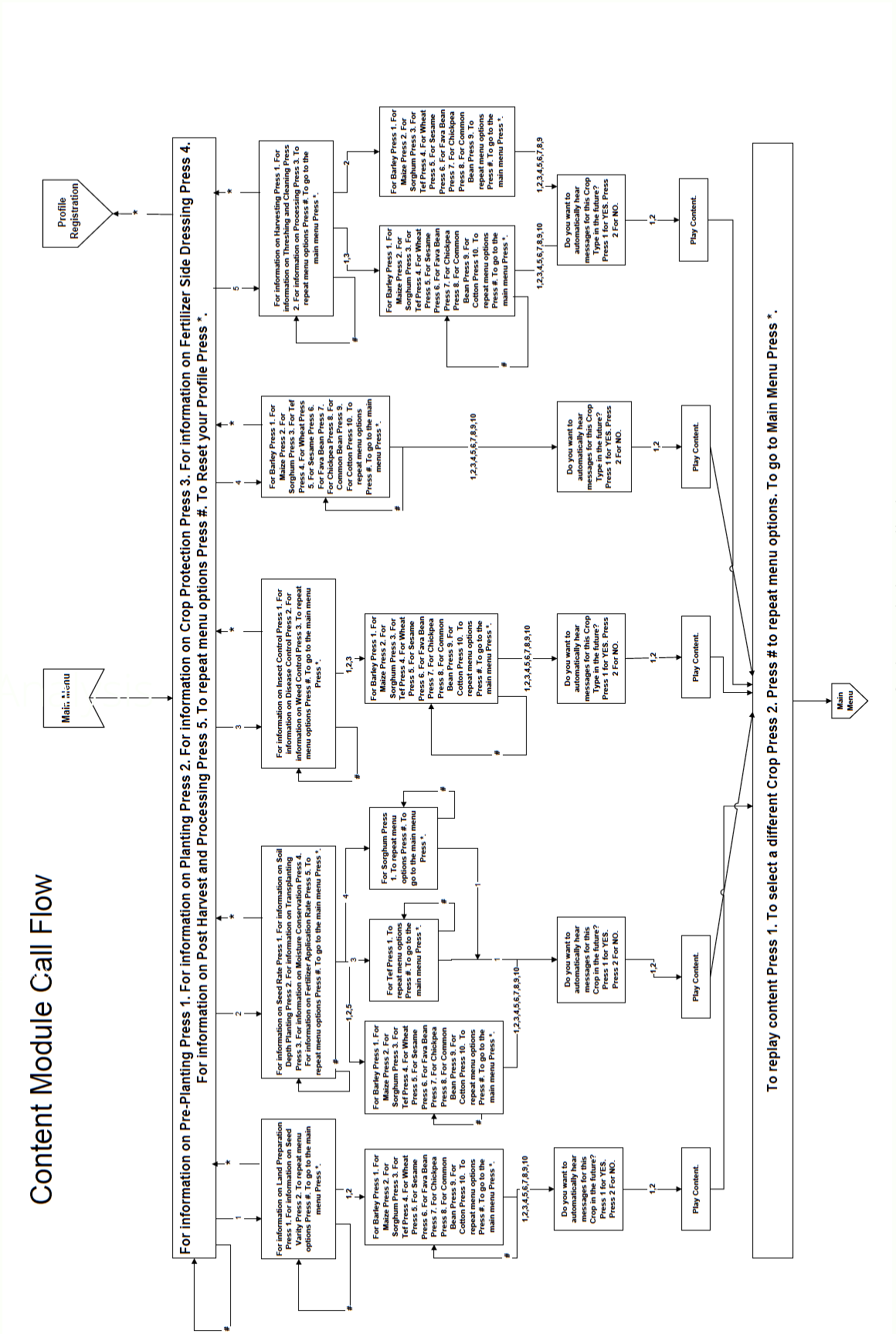


Figure A6: Rainfed Module Call Flow

Appendix

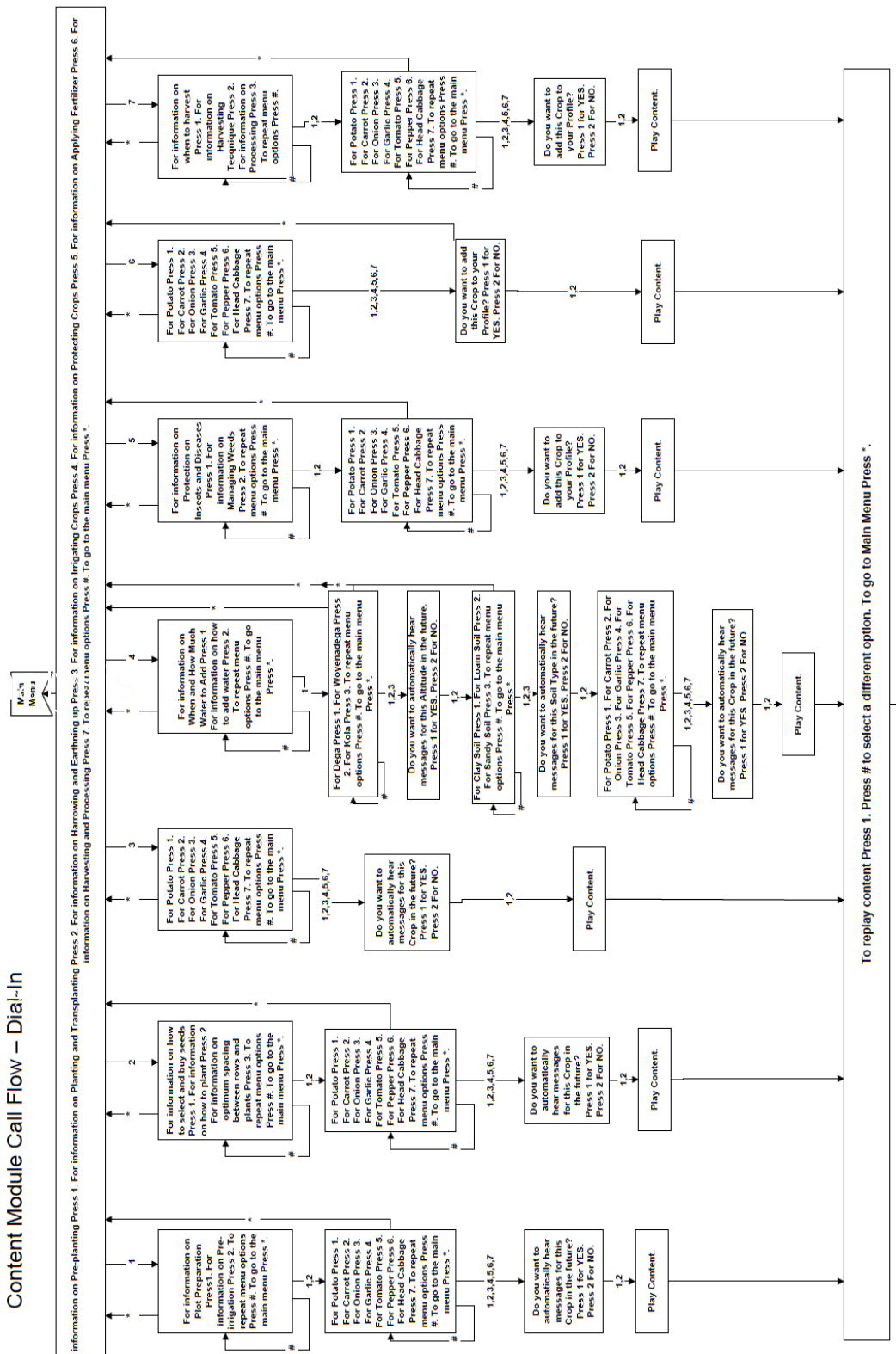


Figure A7: Rainfed Module Call Flow

Appendix

Interviewee's Information

Table A4. Interwee Information Overview

Interviewee	Organization	Field
Indrani Medhi Thies	Microsoft Research	Academic Expert
Ofir Reich	Center for Effective Global Action (CEGA), Google, PAD	Academic Expert, Data Scientist

Interview Questions

Interviews focused specifically on projects participants had worked on and mostly on points within their work. For instance, the majority of our conversation with Ms. Thies revolved around her finding on listing interfaces, while the conversation with Mr. Reich focused on the pressing one phenomenon. A sample interview plan for Ms. Thies below, but please note that some questions were not covered.

UI and Low-Literacy (30 Minutes)

- Even within a non-text IVR system, what special considerations should designers have when working with low-literate populations?
  - How do you think literacy affects the confidence of use of these systems, even when they do not include text?
- I really appreciated your comparison of lists to hierarchical UIs on mobile phones. To what extent do you feel that your findings carry over to IVR systems?
  - Could you discuss the issue that some participants had understanding nesting options? Did you feel that the content had an effect i.e. (if the next menu options were clearly related to one's previous selection)? Going back?
- In your work comparing voice and text UI, I was really taken with this graph and the plateau of subject completing with voice UI. While your paper discusses some reasons for greater attrition among voice feedback users, I am wondering why you believe there is such a drastic plateau at such a point? How would you characterize the external validity of this finding?
- I was also struck by your finding about the different experience between men and women using the voice response system, what do you think attributes to the difference in completion? Have you found other pervasive differences in the uses of the system by demographic?
- First touch-What might explain this phenomenon?

Working in the space/supply side (10 Minutes)

- How have you handled data collection when working with local practitioners? What tactics have you found most effective in getting what you need?
- To what extent do you believe that the composition of callers (in terms of their frequency calling the service) should affect design and call flow?
- How do you consider the trade-off between customization and quick access to information?
- From your experience, who were the strongest influencers with the communities to adopt, persist with, or reject ICT interventions or technology? Who would someone go to if they struggled with the system?

Working with the end user (10 minutes)

- A really unfortunate aspect of research at this time is the inability to interact with end users. Given your experience, I saw in a few of your papers you performed qualitative interviews, were there any surprising aspects of your interactions with end users.
- From an anecdotal perspective, how do end users respond to these new technologies? Does their familiarity, or lack thereof, affect their confidence in the content? Were there any responses to the experiment which stuck out as specifically surprising?

Appendix

Answer one	Answer two
Seedless Produce and Products	Green Foods
Seeded Produce and Products	Red Foods
Animal from the Land	Blue Foods (seafood)
Animals from the Sea	White Food
Non-Perishables	Other Colored Food