Research Project Proposal

Digital Advice through AI in the Context of Smallholder Farming in the Global South

Clara Bersch[1], Charles Efferson[2], Nicole Harari[3], and Sonja Vogt[2]

1 Max Planck Institute, Berlin - 2 University of Lausanne - 3 WOCAT

15 April 2022

1. Background

The world's 500 million smallholders produce two-thirds of the world's food. However, they struggle with low productivity and income, vulnerability to climate change (CC), and poor access to information on practices that could help enhance their production and secure their livelihoods. Trends suggest that the world will need significantly more food while facing an increasingly hostile environment due to CC and diminishing resources. Moreover, reducing the yield gap of major food crops in small-scale agriculture will significantly contribute to food security. A major issue limiting agricultural progress in developing countries is a lack of access to agricultural knowledge (e.g., Oladele, 2011). However, accessing personalized solutions to help people with their farming practices has proven difficult, especially during the pandemic. Information is a crucial resource for socio-economic and sustainable development because it enables individuals to make educated decisions.

2. Proposed Project

In Work Package (WP) 1 of AgriPath, we aim to explore the most efficient ways to provide farmers in the Global South with digital agriculture solutions. The solutions will be most relevant to them based on their current farming situation, geolocation, and livelihood preferences. However, development research has shown that simply sharing information is not enough; one needs to engage people in order to trigger change. Therefore, we are using multiple versions of a chatbot to experimentally test the effects of digital solutions on productivity and well-being.

In order to contribute to this overall aim, we will **design an application-agnostic chatbot and experimentally manipulate its features.** This will allow us to identify if and when chat-based digital solutions support sustainable land management (SLM). We will specifically test whether chat-based interaction is a viable solution for people with low digital literacy levels and for marginalized groups like women and vulnerable youth. First, the chatbot will allow users to record the farming challenges they face in a simple and engaging way. In this way, the chatbot should lead users to feel comfortable and to feel as if they are being taken seriously. Second, after effectively eliciting information about the problem, the chatbot will

provide solutions, tailor-made for specific farmers, in an accessible fashion. Among other advantages, this design approach will allow us to test if the chatbot lowers entry barriers for women, who may have limited exposure to digital tools and not feel skilled enough to use such tools. We will use behavioral insights to develop features that specifically encourage women within agriculture to use digital tools.

- A chat-based interaction will allow us to collect data from the individual farmer and provide solutions without complicated navigation within an app. We are developing the **chatbot to be app-agnostic** so that it can be integrated into many different messenger services (e.g., WhatsApp, Telegram, WeChat), smartphone apps, and websites. **Our app does not build on the functionality of existing apps, and the code will be made fully open source so that other agriculture services can use and adapt it for their needs.** Especially for farmers in rural areas, the possibility to reach the chatbot via messenger is desirable. They do not have to download an additional app that consumes storage space. Also, they are often already familiar with messaging apps, which should make it easy to get started with our chatbot.
- The chatbot will allow farmers to decide on which kind of information related to SLM practices they want (e.g., costs, impacts on production), which will facilitate their ability to see if the SLM solutions provided suit their needs. Further, the app will be able to readily provide follow-up information, e.g., how fertilizers affect the farmer's soil.
- Because our chatbot will interact with farmers in natural languages, it can support interactions and provide information in ways generally associated with strictly human interactions, and in this way it can fulfill a social purpose. Based on the assumption that knowledge of sustainable agriculture practices and how to implement them is not the only factor driving adoption, **we aim to increase the feeling of agency among farmers by incorporating coping strategies based on growth mindset**. Growth mindset is the simple yet effective idea that succeeding at something is not a matter of innate talent, but rather of practice and effort. Interventions based on growth mindset have been widely tested and validated as a way to help people achieve their goals. Growth mindset interventions are specifically effective among the less privileged segment of the population as these people often lack a sense of agency.
- We will utilize Natural Language Processing (NLP) because it provides the ability to classify farmer input through machine learning and act accordingly. Although rule-based systems can also control chatbots, NLP offers the possibility to enable a broader range of interactions. In particular, it allows interactions that the programmer could not have anticipated. The free text input is an additional advantage, especially when a farmer's problems cannot be clearly outlined from the beginning. Data can be collected that can be used to improve understanding of individual challenges, and this can be leveraged later to improve the solutions provided. Apart from text and speech, farmers will have the option to answer with predefined buttons and we will be able to measure the added value of an NLP-based chat/voice-bot.

The planned chatbot experiment is divided into two phases: Phase 1 will focus on an online chatbot experiment with English-speaking farmers, while Phase 2 will focus on repeating the online chatbot experiment in local languages (Hindi for India and Luanda for Uganda). The two experiments are explained further in the following section.

Phase 1: Online chatbot experiment in English language

In the first experiment, which will be conducted by the end of 2022 in India and Uganda, we will test different text-based chatbot versions in English against each other. One version of the chatbot will only provide SLM solutions to problems described by the farmer using the app. This version of the app will not include messaging designed to address the psychological concerns of the farmer. The second version, in contrast, will address such concerns by augmenting the solutions present in the first version with additional information about growth mindset and coping strategies. For this initial experiment, it will be necessary to work with farmers who speak English and can thus interact with the chatbot. We expect these to be relatively well-situated farmers who farm to provide for their families but are also high in terms of market integration. These are farmers for whom it would often be beneficial in the long-run to adapt their farming practices to render them more sustainable. Further, such changes could lead to sizable impacts on the environment precisely because these farms will tend to be relatively large. Finally, relatively well-situated farmers could serve as role models for other farmers and therefore act as agents of change. Therefore, this first experiment will provide a first test of the chatbot among a relatively well-off set of farmers, who are probably in the best position to initiate new farming practices in a local population. This study is already funded by Agripath.

Phase 2: Online voice-bot experiment in local language

Training language models is a challenge for languages that are widely spoken, but for which there are few written texts that can be used as training samples. In order to train the chatbot to understand these languages, training data must be collected and labeled. In addition, the language models must be adapted to the syntactic structure of the specific languages. We are very excited and thankful that FairForward will help us train the chatbot for Uganda and India. Our overall objective is to test experimentally the different chatbot versions in all three languages (English, Hindi, Luanda) to provide empirical evidence on what works and what does not work with different types of farmers (age, size of farm, gender).

In the first experiment, explained above, we plan to work with relatively affluent farmers who speak English and probably run relatively large farms. However, many smallholders do not speak English, but together they account for a large proportion of the crop production. Therefore, in Phase 2 we aim to test the chatbots in local languages and dialects in order to test their effectiveness with smallholders who do not speak English and are perhaps less affluent than the farmers in Phase 1. Phase 2 will allow us to study specifically how DAS can

empower smallholders who belong to marginalized groups like women and youth. Since we aim to make the bots voice-based, as opposed to text-based, they will also be accessible to farmers with low literacy, which we expect to be the case significantly more often for female farmers. Therefore, phase 2 will further generate essential insights into how to provide digital solutions to smallholder farmers in ways that support the uptake of sustainable practices.

Because training the local language models will take several months, we plan to conduct the Phase 2 experiments in February/March 2023 (the respective months are based on the agricultural cycle of Uganda and India).

Analogous to Phase 1, in Phase 2 we will compare two versions of the local language apps. One version will provide SLM solutions only. The other will add growth mindset and coping strategies to empower farmers and build up psychological resilience. The most marginalized segments of a population often lack a feeling of agency. Even though they have information they could use to improve their situation, they may simply not believe they use this information precisely because they exist on the margins of society. This creates a kind of harmful feedback, but dozens of experimental studies have shown that marginalized decision makers can benefit from growth mindset messaging, which can lead them to actively pursue better solutions to the problems they face in spite of setbacks.

Our experimental studies will take several weeks and involve thousands of farmers. We will measure engagement, well-being, diffusion (forwarding solutions via text messaging to friends, family, and other farmers), as well as attitudinal and behavioral change regarding SLM.

Results and learnings

The data collected from both experiments will allow us to gain a comprehensive understanding of how technologies need to be designed to help male and female farmers of different age and social groups and to induce the adoption of SLM practices. We will be able to measure the added value of free text and speech input possible through NLP. In addition, we will be able to examine the impact of growth mindset messages in order to strengthen a feeling of agency. Further, we can evaluate what differences exist across groups (better-situated farmers vs. smallholder farmers, men versus women).

In WP4 of AgriPath, led by Grameen US, we will produce a toolkit for DAS providers that includes the main findings and results from the AgriPath research. The toolkit will target DAS providers to support them in successfully targeting farmers, especially women, and inducing behavior change via digital solutions. The **causal** insights from these chatbot/voice-bot experiments will form an essential part of the toolkit findings as the bots are fully portable. Results will be fully generalizable, and we aim to publish them in a general science journals to facilitate impact by reaching a wide audience. The study will specifically

provide insights into which types of farmers (size of the farm, gender, age, income, challenges) benefit most from specific types of bots (button-based, NLP based, with or without growth mindset messages). These kinds of insights are crucial because people everywhere in the world are easily saturated with information. The task is not simply to provide information, but to reach a specific farmer with the right information (including SLM practices) provided in an engaging and respectful way. Farmers must want to use the app because it provides benefits and actually helps them improve their well-being. As we can only reach a certain number of farmers on an individual basis, we need to make their experience with DAS valuable and enjoyable. Only then will farmers encourage others to use DAS, which is essential for scaling up SLM.

In addition, GIZ/FairForward will be able to utilize the learnings from the experiment for DAS and the use of NLP-powered voice-bots:

NLP generalizable insights:

- What features of a chat or voice bot specifically encourage farmers to engage with the respective bot?
- Can free-text/speech possibilities enabled through NLP help gather more accurate information, which in turn is useful for improving the services?
- Do certain groups (e.g., women) particularly benefit from this interface?
- How do growth mindset messages affect the sense of self-efficacy and the perception of control? Does growth mindset messaging bring particular benefits for women, who often have less access to technology and agricultural resources?
- Do farmers prefer to interact with buttons, or do they make use of the possibility to use free text and speech?

DAS generalizable insights:

- What information is particularly relevant for farmers to make a decision (costs, impact, labor) on whether to implement an SLM practice?
- Does the prioritization of information differ between smallholder farmers and farmers with higher market integration, between men and women, and between younger and older?
- Many SLM practices pay off in the long run with no immediate outcomes. In addition to the actual implementation, what specific information supports the transition to SLM practices?

3. Components and Outputs

We propose the following components, outputs, and related activities.

Component 1: Programming of voice-bot in Luganda and Hindi

Outputs:

1. Relevant SLM practices translated in Luganda and Hindi
2. A refined pool of possible intents[1] identified based on the experience of Phase 1
3. Voice-bot that can handle intents of farmers correctly developed

From the development of the English-based chatbot, intents will already be available. Based on the experiences gained, these will be adjusted to refine the functionality. The training samples required for intent recognition will be identified in Luganda and Hindi in the already existing pool of training samples. Missing samples need to be identified and created. Likewise, there will be an improvement of the entity types that can be extracted and an adaptation of the conversational flow. We intend to pilot the first version to test whether the voice-bot recognizes the intents correctly. After that, we will again add new intents and refine the functionalities.

Component 2: Research on agency, locus of control, and adoption of SLM practices with smallholder farmers who use the voice-bot for a period of four weeks

Outputs:

1. Data showing effects of two different types of voice-bot (Growth-Mindset vs. neutral) on the feeling of agency, locus of control, and SLM adoption from studies in Uganda and India

Farmers will take part in the experiment over four weeks in which they can interact with the voice-bot daily via one of the widely used messengers. Before the study starts, we will educate farmers about the study and obtain their informed consent. Farmers are asked to complete a 15-minute questionnaire at the recruitment stage. We will elicit their most significant challenges concerning farming and where they usually obtain information about farming practices. We will also ask them about their well-being, measure locus of control, as well as risk and time-discounting preferences. Farmers will then be randomly assigned to one of the two different conditions (neutral chatbot vs. growth-mindset chatbot). When the experiment ends, we will again collect data for the dependent variables, e.g. locus of control, agency, and well-being. In addition, farmers will be able to indicate if they have implemented any of the SLM practices or plan to implement practices in the future. We will also ask about their experience with the voice-bot, whether it was worthwhile, what troubled them, and whether they trust the technology.

Component 3: Analysis of results and publication of findings

Outputs:

[1] Intent: Intention, or purpose of the user in the conversational flow.

1. Presentation of results at leading scientific conferences related to agriculture, development economics, and development studies
2. Presentation of results specifically to GIZ, SDC, World Bank (behavioral science unit), DfiD, and the Max Planck Institute in Berlin
3. Aiming for a high-profile publication in a prestigious peer-reviewed journal such as *Nature Sustainability*. The advantage is that these journals invest a lot of resources in the communication and dissemination of the published articles and thus have an enormous reach within the scientific community and among practitioners
4. The results will be compiled in the WP4 Toolkit with practical guidelines for DAS

Budget

Description	Quantity	Unit	Eligible up to in Euro	Total in Euro
Component 1: Voice-Bot Programming				
Progammer chatbot[2]	*TBD*	months	*TBD*	*TBD*
Component 2: Experimental Testing				
Salary: Research assistants participant recruitment / 10 days	6	per assistant	25/per day	1500
Salary: Research assistants onboarding and follow up calls / 10 days	6	per assistant	25/per day	1500
Internet Bundles	2	per country	500	1000
Incentives and rewards for participation	2	per country	25000	50000
Experimental costs per country : 27000				

[2] The programming costs depend on which training samples are already available. FairForward already collects data. The training of the NLP models has to be estimated as well. Depending on the piloting, additional costs may be required.