A review of recent Voice Based Applications in developing countries

Marie-Lou David
Vrije Universiteit Amsterdam, The Netherlands
m.david2@student.vu.nl

ABSTRACT

Voice based applications have established themselves as hyper relevant in the ICT4D field in order to reach remote and largely illiterate communities. This paper reviews publications about voice based applications in developing countries and their observed impact on their target demographic. It was found that ongoing research was focused on improving different sectors of which agriculture, marketing and journalism are the most prominent. Moreover, the impact of voice based applications was focused towards economic, social and political improvement depending on the application type. Lastly this paper identified challenges to be aware of when designing such an application and encourages more discussion around the economic viability of a project and its long-term impact.

KEYWORDS

Systematic Literature Review, Voice based application, Interactive Voice Response (IVR), Interactive Voice Forum (IVF), Human Computer Interaction (HCI), Impact, Spoken Web

1 INTRODUCTION

With the growing digital divide between the global North and the global South, Information and Communication Technology for Development (ICT4D) is a popular field of research. It aims at bringing ICT to less developed countries in order to further some communities' development sustainably [13]. In the last 20 years, ICT4D has met a soaring success. Indeed, as technologies develop in the global north, there is a growing enthusiasm aimed at helping the global south with the same technologies. Many initiatives are taken to further development in less developed countries, and they more often than not include ICTs.

One particular challenge that remote populations in developing countries present is the low literacy rate, which prevents the use of written media for ICT as most of us use it. To circumvent that problem, researchers have focused on developing voice based applications, which rely on spoken commands, often over the phone, and takes as response either speech or number input through Dual Tone Multi Frequency (DTMF). Many initiatives have been taken to implement voice based applications in communities of developing countries in order to improve one aspect of their everyday lives. This is part of ICT4D projects.

This study aims at finding out the impact that such applications can have when implemented, what sector are they applied to and what are the challenges that developers faced. This is done through a systematic review of the existing literature. The papers are selected through a snowballing method, data is extracted and classified through standard classifications in order to best answer the research questions. The contributions of this paper include an overview of the different voice based applications deployed in developing countries, recommendations for evaluating such projects and factors of success.

2 RELATED WORK

The field of research around ICT4D projects is broad and can be scattered with the different initiatives taking place all over the world. Many literature reviews embarked on the process to assess the state of ICT4D research and to propose solutions to make it more cohesive. In this section, we present key papers to understand better the progress of this field, and of voice based interaction research.

Yim and Gomez in [35] underlined the foci, challenges, gaps, and limitations that the ICT4D actors encounter when evaluating the impact of their projects. They conducted a study over 222 reports and 24 interviews of relevant academics or professional. They found that ICT4D evaluation focused on continental or regional scope, method and research design, and notion of development. The authors identified several challenges and gaps in ICT4D evaluation, such as difficulties in assessing impact, lack of knowledge accumulation, project-oriented (short-term) evaluation practices, and a lack of domain-specific approach. Yim et al. evaluated the evaluation of the ICT4D field in general but in this paper, we will focus on the domain of voice based applications specifically and the result of their evaluation. Overall, they encourage the adaptation of the Capability Approach in order to standardize the evaluation of ICT4D projects.

In their paper [28], Ramadini et al. look at the state of holistic research in ICT4D. They note the divide in the focus of project development between technology adoption, economic impact, or social change. They highlight the need for a more holistic approach where researchers consider the complex interplay between technology, development context, and human factors. They encourage researchers to use multi level analysis to understand the impact of their projects, both on the macro and micro level. Our research is focused on examining the impact of voice based applications and it is evaluated and transmitted in the existing literature.

The existing reviews on Interactive Voice Response (IVR) applications are domain specific. In their study [15], Itorobong et al. compare the technology used to develop IVR applications. They briefly discuss their usefulness in fields such as healthcare, banking, telecommunication, education, and government. They found that a great majority of IVR interfaces are developed in VoiceXML, with the Voxeo speech engine and a MySQL database. Our study will focus on the effect of IVR applications in developing countries and how they can be an essential ICT4D tool.

3 STUDY DESIGN

3.1 Research Goal

Many voice based applications are developed for real word applications in developing countries. Those applications aim at improving an aspect of life for under served communities and people with low access to resources, especially ICT. In the last twenty years, projects have been implemented in order to help people in the Global South [35]. Voice based applications can facilitate access to ICT in regions of low literacy, and thus have been a focus for researchers trying to help in those areas. This study aims at analysing the impact of such applications. Indeed, due to limited hardware resources and low digital literacy, providing voice based services to remote communities in developing countries becomes a challenge and can often fail [5]. In this paper, we aim at examining key projects of the decade and analyse their areas of impact, how significant of a difference they can make and what challenges did they present. This will enable future researchers to learn from the mistakes of others and avoid pitfalls. Moreover, this will motivate the need for more voice based applications if they prove to be successful enough.

3.2 Research Questions

In order to conduct our analysis on current literature, we have the main research question: **RQ1: What are the key areas of impact of recent voice based applications in developing countries?**. This question will guide our research. To make it more precise, we identified three sub questions in order to fully answer the main question. They are as follows:

- **RQ1.1:** To what sectors are voice based applications the most applied?
- **RQ1.2:** What is the observed impact of voice based applications?
- **RQ1.3:** What are the main challenges faced when developing a voice based application?

3.3 Initial search

The initial search for relevant articles to our research questions was conducted on Google Scholar because it is a neutral search engine, not linked to any specific publisher. We picked articles from a search with the keywords: *Voice based* and *developing countries*. These are very broad terms but are necessary for the scope of the search, which aims at examine the current state of voice based applications in developing countries and what their areas of impact are.

3.4 Application of selection criteria

Articles will be chosen according to the following inclusion and exclusion criteria in order to focus our analysis on relevant papers.

I1- Studies addressing the implementation of a voice based application. This inclusion criterion is utilized to select exclusively studies considering voice based application.
I2- Studies taking place in a developing country. This inclusion criterion is utilized to filter out studies which did not take place in our specific context.
I3- Studies discussing the impact of implementing a voice based application. With this inclusion criteria, we ensure that only papers which examine their impact on the community they focused on will be included.
E1- Studies in the form of editorials and tutorial, and poster, as they are deemed to not provide the required level of detail and information.
E2- Studies that have not been published in English language, as their analysis would result to be too time consuming.
E3- Studies that have not been peer reviewed, in order to ensure the high quality of the studies considered.
E4- Duplicate papers or extensions of already included papers, in order to avoid possible threats to conclusion validity.
E5- Papers that are not available, as we cannot inspect them.
E6- Papers that are published before 2007, as they are deemed not part of recent research anymore.

3.5 Snowballing

To ensure that no relevant articles are missing on our topic, we adopt a snowballing approach to select relevant articles. Indeed, our initial keywords are broad and does not encompass every paper that could be interesting for us. The approach is done following the guidelines in [33], which has proven to be reliable.

This method is an iterative process from a set of articles. The initial set of articles is selected from a Google Scholar search according to the criteria established above. For each paper of the selected set, we carefully examine the studies referenced and the studies that cite the considered paper. If any of those studies fit our criteria, they are included in the next set of papers. The process is repeated until no new papers is included in the next set.

This method ensures that the relevant literature is identified, it enables "famous" papers to be included because they will be referenced multiple times and "unknown" papers to be chosen because they most likely cite relevant literature already identified. This also prevents us from restricting our search to a specific application domain, area of impact or publisher.

3.6 Data Extraction

Once every relevant articles has been gathered for our review, the data will be manually extracted. The author will read through them and take notes in order to isolate answers to the research questions. The following information are extracted from the papers:

- Year of publication
- Country of action
- Sector of action
- Origin of funds
- Method of evaluation
- Main goal of application
- Evaluation results
- Challenges encountered

3.7 Data Synthesis

After reading through every articles, we will conduct a thematic analysis on our findings. This means that we will identify themes or patterns across the selected articles in order to see what the most recurring conclusions are when implementing voice based applications. Moreover, this will enable us to compare and contrast the results of different studies.

The year of publication and country of action are metadata which do not need to be classified. The sector of action of the application will be classified following the International Standard Industrial Classification of All Economic Activities (ISIC). The origin of funds will be annotated following the organizations providing funds for the project described in the paper. The method of evaluation is classified by qualitative and quantitative methods. Data regarding the main goal application goal, evaluation results and challenges will be annotated by the author for each article.

3.8 Study Replicability

In Table 1 are identified all selected articles, along with the iteration they were selected at. It is interesting to note that the papers P3 and P21 present multiple voice based applications. One application for each paper is also represented through its standalone paper (P8 and P9 respectively). It was decided to keep P3 and P21 and during data extraction, parts about the applications already described in P8 and P9 will be ignored.

Table 1: Selected articles per iteration following the snowballing method

ID	Iteration	Title	Publication Year	Citations	References	Selected Papers
P1	Start set	Developing voice-based information sharing services to bridge the information divide in marginalized communities: A study of farmers using IBM's spoken web in rural India [6]	2021	6	71	1
P2		An alternative information web for visually impaired users in developing countries [27]	2008	9	9	1
P3		Spoken web: using voice as an accessibility tool for disadvantaged people in developing regions [19]	2012	8	12	2
P4		Voice-Based Marketing for Agriculture in Northern Ghana [9]	2013	11	14	3
P5	Iteration 1	Power to the peers: authority of source effects for a voice-based agricultural information service in rural India [24]	2012	39	27	1
P6		VOISERV: Creation and Delivery of Converged Services through Voice for Emerging Economies [20]	2007	16	17	0
P7		Designing a voice-based employment exchange for rural India [32]	2012	21	21	3
P8		Avaaj Otalo: a field study of an interactive voice forum for small farmers in rural India [23]	2010	217	24	5
P9		RadioMarché: Distributed Voice- and Web-Interfaced Market Information Systems under Rural Conditions [7]	2012	9	23	0
P10		Tibaηsim: Information Access for Low-Resource Environments [8]	2021	0	21	0
P11		E-Service Innovation in Rural Africa Through Value Co-Creation [4]	2020	4	27	1
P12	Iteration 2	Supporting treatment of people living with HIV / AIDS in resource limited settings with IVRs [17]	2014	32	37	0
P13		Personalized weather information for low-literate farmers using multimodal dialog systems [26]	2021	3	50	1
P14		Learnings from deploying a voice-based social platform for people with disability [10]	2019	7	27	0
P15		A Real-Time IVR Platform for Community Radio [18]	2016	38	35	0
P16		User-centred design of a digital advisory service: enhancing public agricultural extension for sustainable intensification in Tanzania [3]	2020	32	56	0
P17		A Hybrid Multi-Modal System for Conducting Virtual Workshops Using Interactive Voice Response and the WhatsApp Business API [25]	2021	1	31	0
P18		Collective Support and Independent Learning with a Voice-Based Literacy Technology in Rural Communities [21]	2020	14	88	1
P19		Revisiting CGNet Swara and its impact in rural India [22]	2015	33	34	1
P20		Ila Dhageyso: an interactive voice forum to foster transparent governance in Somaliland [11]	2013	14	12	0
P21		Voice-based Web access in rural Africa [12]	2013	15	20	0
P22	Iteration 3	Baang: A Viral Speech-based Social Platform for Under-Connected Populations [30]	2018	23	33	0
P23		Use of interactive voice response for professional development in Kenya [16]	2017	2	17	0
P24		Karamad: A Voice-based Crowdsourcing Platform for Underserved Populations [29]	2021	8	67	0

3.9 Threats to Validity

This section presents the different threads to validity that our study encounters and their mitigation strategy if one has been taken.

3.9.1 Internal validity. The main internal threat presented by the snowballing method is of incomplete coverage. Indeed, it might prevent us from finding a relevant article because it was not cited or referenced in other pre-selected articles. This threat can be mitigated by having a starting set of articles which covers different areas and enables the search to go in different directions.

Moreover, there is a subjectivity threat posed by the selection of articles in itself. If the author selecting the articles is subjective in their choice, then the result of the study would be skewed. This is alleviated by the presence of well-defined inclusion and exclusion criteria that are standard in this type of literature review.

3.9.2 External validity. By conducting a snowballing approach, we rely on existing published works and their interconnectedness. This might introduce a publication bias, by potentially overlooking unpublished or less-cited works. This external threat is accepted as we want to analyse peer-reviewed studies as they provide a more reliable discussion than unpublished papers.

3.9.3 Construct validity. The principal construct validity of this paper stems from the manual data extraction and synthesis performed by the author. A bias is introduced because the author will read every article and extract data themselves, meaning that it could lead to essential information being missed. This threat is mitigated by the author taking consistent notes on each paper that relate to the research questions. This will enable us to have a clear overview of the corpus and to discuss the different findings in a relevant manner.

3.9.4 Conclusion validity. Because the data extraction will be done manually by the author, there is a threat of inadequate synthesis. This threat is mitigated by reading the articles multiple times when looking for a specific information in order to ensure that the information is correctly found and processed. Claims of the result section will be backed up with the list of specific articles supporting that claim to ensure replicability and coherence.

Moreover, this study is carried out by one person, single-authored snowballing studies might lack diverse perspectives and interpretations, limiting the depth and richness of the analysis. It is mitigated through the supervision of experts in the ICT4D field who will bring nuance to the discussion.

4 RESULTS

After careful examination of all articles compiled, the results are presented in this section. The year of publication ranges from 2007 to 2021. 15 of the articles discuss an application developed either in India or Pakistan while the remaining 9 papers address implementations in the African continent. Every project had a similar target demographic composed of people with low literacy and digital skills, living in remote communities and without access to the latest technologies. Projects before 2020 focused on services using phone lines because it provided the most coverage, was cheaper to use and could be accessed by non smart phones [9], [19]. However, a trend has been observed that more people in remote communities of developing countries have a cheaper access to internet than to phone lines and that they have more and more smart phones [25].

Section - Division	Corresponding articles
Agriculture	P1, P3, P4, P5, P8, P13, P16, P17
Professional activities - advertising	P4, P6, P9, P11
Information and communication	P2, P19, P20, P21
Arts, entertainment and recreation	P10, P14, P22
Administrative and support service activities - employment	P7, P23, P24
Human health	P12, P15
Education	P18

Table 2: Papers categorized by sector of activity

Let us answer the sub research questions with the knowledge acquired from the articles.

4.1 To what sectors are voice based applications the most applied?

Implementing a successful voice based application requires knowledge of the needs of the community that it will help. Therefore, we categorized the papers per sector of action according to the ISIC. The results are reported in Table 2. A third of the papers touch on helping farmers to improve their agricultural techniques. The papers classified in the advertising division are also targeted towards the marketing of agricultural goods. Thus a majority of the voice based applications are aimed at helping and improving a country's primary economy [34]. The other sectors that we isolated are less targeted towards economic aid but more towards social well-being. This is seen through the different reporting tools which help citizen act on their rights (section Information and Communication), the employment tools which aim at alleviating unemployment, the social media applications (section Arts, Entertainment and Recreation) that aim at creating a community and linking people or with healthcare and education applications which try to raise living conditions for its users.

Overall, voice based applications can be used to improve economic performance through helping farmers with better yields and sales. It can also be used as a way to create social link, improve access to citizens' rights, education or healthcare.

4.2 What is the observed impact of voice based applications?

After identifying the sectors of action of voice based application, it is important to evaluate their impact on the communities they are supposed to help. In the papers selected, the impact on people was assessed through qualitative interviews with a sample population of the users, quantitative data from a form or quantitative usage data of the application. In the quantitative form category was included the papers which conducted a System Usability Scale analysis through

Method	Corresponding articles
Quantitative usage data	P1, P2, P3, P5, P7, P8, P11, P12, P13, P14, P15, P16, P17, P18, P19, P20, P22, P23, P24
Qualitative interviews	P1, P2, P5, P8, P9, P11, P12, P13, P14, P15, P16, P17, P18, P19, P20, P21, P22, P24
Quantitative form	P4, P10, P15, P16, P21, P22, P23, P24

Table 3: Methods used for impact assessment of a voice based application

a questionnaire. Table 3 presents the methods chosen by each paper. Almost every paper used a mix of interviews and usage data from their system to evaluate their relevance to the context.

There is a clear distinction in the impact of the application depending on its knowledge sharing model. The first type of application crowd sources information and aims at freely sharing the information inside the same community. The information comes from the users and is shared to other users, and in some cases it can be brought for higher authorities to deal with but it is not the main focus of the project. The second type of application is more of a top down approach, where an authority broadcasts information to a community, feedback from the users can be taken into account but it is not the main feature of the application.

Those different types of voice based applications are important to understand the different impact they have on their target demographic. Let us examine such observed impacts according to the previously defined categories.

4.2.1 Crowd sourced knowledge sharing. This type of application includes the publications P1, P3, P4, P6, P8, P9, P10, P11, P14, P19, P22.

Because the main goal is these applications is to facilitate knowledge sharing inside of a community, its main benefit for people is the community building aspect. When using forums in the shape of voice based applications, users reported being pleased to link with their peers even when the purpose of the application was not explicitly to make social connections. This is observed directly in P1, P8, P10, P14, P19 and P22. Users of those applications showed enthusiasm at the idea to be able to interact with other people and in some cases, the use of the application increased their social standing in their village (P1 and P8). For people living in remote areas, it was a valuable feature to be able to share their experience and knowledge with their peers. It enabled them to extend their social circles, which is a beneficial social impact from interactive voice forums.

Concerning applications which implement a social media targeted towards people with low literate skills or for visually impaired people, the studies P7, P14 and P22 found that the users, in addition to socializing about diverse subjects, could advance their professional network through the media. Indeed, many people reported that they effectively used the media to find work, advertise their service or share opportunities with others, especially in the blind community. Therefore, users showed an ability and a tendency to diversify the use of a voice based application and make it relevant to their needs, if the system allows it.

Moreover, applications from the advertising sector have for goal to improve sales and they proved to be successful in that endeavour, especially P11 which highlights the economic viability of their model through an e^3 value analysis. This methodology is especially interesting because it enables to perform an in-depth assessment of the business model of an application. Unfortunately, it has not been applied to any other case study present in this selection. It is interesting to note that of every papers, only P2, P4, P8, P9, P11, P16 and P24, discuss with varying degree of preciseness the economic feasibility of their project and how to make it accessible to as many people as possible by providing value to every stakeholder involved.

4.2.2 Top down knowledge sharing. This type of application includes the publications P2, P5, P7, P12, P13, P15, P16, P17, P18, P19, P20, P21, P24.

Sharing knowledge in a top down manner implies that people have to trust the information given to them for the application to be successful. Indeed, the information given in those applications come from an authority figure, either an expert from a local NGO, a government official or a professional. Across studies, it was found that this approach had varying degrees of success depending on the sector of application. Indeed, in an agricultural context, P5 compares advice given by a peer, meaning another farmer, or by an expert from a more scientific background in India. The study found out that users of the system used it more if the advice was given by a peer. This is confirmed by P16 which took place in Tanzania and during feedback, users shared doubts about the reliability of the recommendations given to them by scientists. Publication P17 shares also similar findings, where farmers did not follow the formation given to them all the way because they wanted to receive education from people with first hand experience of farming.

However, in a sector like healthcare, people responded very positively to having access to advice from a professional healthcare physician (P12, P15). It was found to improve their access to health-related information, their trust in the information given and their overall health. This was especially true when the voice used for the voice based application corresponded to the voice of their own doctor or nurse, which they already trust. In publication P15, it was found that people with a health condition would rather have access to a doctor's input than to their peer's experience.

Three applications focused on bridging the gap between citizens and their government, they are described in P2, P19 and P20. Those applications all had a relatively high usage rate and in their interviews, users made good use of the access to government officials either for grievances (P2 and P19) or for transparent information on government proceedings (P20). The application of the paper P19, CGNet Swara has impressive results over 4 years of deployment, where more than 6 900 posts were published and the system helped at least 280 problems to be solved, including by going through government officials if necessary. These applications show that voice based applications can be a solution to bring citizens closer to their government and help them navigate their problems. It is essential

for a state to be reachable by its people to ensure prosperity of its population.

Another recurring theme during interviews for top down knowledge sharing applications is that the users felt an improved sense of autonomy (P2, P13, P18, P19, P22, P24). Indeed the voice based web provides an inclusive access to information and services that people could not access through the regular Web. People, especially blind users, reported that they were happy to gain autonomy on their phone and not having to rely on their family or friends for some tasks. Thus voice based applications provide a unique way for users to improve their lives and overall well-being.

Moreover, in a few papers (P13, P22 and P24), authors were surprised by the demographic their app was able to reach, which was not the initial target demographic. Indeed, in P13, when surveying the average age of the users, they found out that older people 50 or more years old made regular use of the app and could navigate it properly. Similarly, in P22 and P24, the majority of users were visually impaired, which was not their intended audience but proved to improve accessibility to information and opportunity.

4.3 What are the main challenges faced when developing a voice based application?

Developing a voice based application presents a unique set of challenges, especially in a developing country where there is a limited access to hardware resources such as phones or servers. Over the years, the situation has changed in regards to phone line coverage and smart phone possession. Nowadays, most people have access to a smart phone and to internet but it was not the case 5 years ago. This affected greatly the development of voice based application. One challenge that still remains is to really understand the target demographic of users. One has to thoroughly understand what access to phones, internet or mobile coverage they have and more importantly how much they know how to use their device. Most people targeted for voice based application have low literacy skills, but one also has to take into account their low digital skills. There are instances of people not knowing how to unlock their phone (P12). It is paramount to take those parameters into consideration when developing an application so that the user can actually use them.

In terms of hardware limitation, a challenge that developers face is that the system can also be down for periods of time due to hardware or software malfunctions. In limited resources environment, things can easily go wrong because of unreliable material or internet connection for example. This is something that multiple papers mentioned and the system can stay down for up to multiple months, which hurts the app in the end (P13, P16, P20, P22).

A recurring problem when interviewing people is that they did not know how to fully use the system or might not be aware of every functionality that is offered. A big challenge is therefore to train users on the system so that they can become proficient in using it. This takes time and effort because it can be a repetitive task as some people need a lot of time to understand how to work certain systems. However, if people are actively aware of how to use the system, they will use it more and it will provide them with added value.

Another challenge that voice based applications face is a low use rate. It can be increased with better training as mentioned above or through a better user experience. Indeed, the system needs to be tailored to its users so that they feel comfortable navigating it. The developers have to be very culturally aware of the population that they intend to serve. It is often important for users to hear native speakers of their language for the system to be adopted (P4, P5, P9, P10, P13, P20, P21, P22) and for the whole system to be available in the local language as most people will not be proficient in English or the official language of the country. This is a big challenge for voice based application developed who wish to use voice input and Automatic Speech Recognition (ASR). Those technologies require a lot of fine tuning and are not usually available in small localized language and dialects without a lot of recordings publicly available. Providing reliable ASR for such languages is an ongoing field of research. It limits possibilities of development as opposed to mainstream languages like English for example.

In relation to this, some papers have raised the issue that they needed multiple rounds of development to make their application culturally aware enough to increase their use rate (P9, P10, P13, P16, P20, P22). For some users it is important to be greeted a certain way, as is the habit in their community. P13 saw a considerable increase in usage after adding proper greetings. Also every user need to be properly guided through the specific input format that it expected when using ASR and if an error happens, the system needs to correctly inform the user on what is the problem and how it can be solved.

Lastly, even with a well designed voice based application, one has to consider how will the user be aware that it exists and to start using it. One challenge related to a low use or retention rate is that it is difficult to acquire new users. Indeed marketing is complex in communities where ICT has not penetrated the lives of everyone living in it. Most publications relied on key actors to spread the word in their communities and get people around them to use the application in a word of mouth strategy. One application (P24) reused a discontinued phone number which used to connect to voice based application liked by blind users. This strategy also proved successful as they gained a significant amount of users in a short time just through people trying to reach the former application associated to that phone line.

5 DISCUSSION

It is important to keep in mind that the projects described in the papers selected have all been implemented for varying amounts of time before being evaluated. This affects the results of the assessment because it will be more difficult to evaluate the impact of an application that has only been implemented for less than 6 months. We included papers presenting applications ranging from proof of concept to an implementation of multiple years because it is interesting to compare the early stages challenges of designing a relevant application for its community with the later challenges that come which are more related to retaining the user and making sure that new users are acquired. We can classify our corpus depending on their runtime into three categories:

- **Proof of concept**: The application has been designed and validated by forms to potential users. This is essentially done

to highlight the need for a voice based solution in a community and give a prototype of how it can be achieved by taking into consideration the limited resources. It includes the papers P4 and P6.

- **Trial runtime of less than 6 months**: Many research papers implementing voice based applications do so for a fixed period of time, often with a total runtime of less than 6 months. This enables the researchers to design a relevant voice based application and to validate it for a short time on the field. Thus they can ensure that their solution is helpful, at least short term and they can draw conclusions on the short term impact of their work. It includes the papers P2, P3, P5, P7, P9, P10, P11, P12, P15, P16, P17, P18, P20 and P23.
- **Applications with a long runtime of 6 months or more**: Those applications have been proven to be useful and helpful to the community therefore they have a longer runtime in the study. Some of them are still running as of now. This includes the applications from papers P1, P8, P13, P14, P19, P21 and P22.

The majority of papers describe applications that have been implemented for less than 6 months. This shows that there is a demand that is well identified for voice based applications. However, it is complex to keep those applications running outside of an academic scope. We also have to account for the fact that research on voice based applications and ICT4D in general is decentralized, not everything is published and peer reviewed. Some applications in the corpus might still running but they just might not have any publications in their name to advertise their impact years after. This highlights the need for more consistent communication around voice based applications and their impact on communities they are part of. The added traction would benefit them by attracting more talented people to work on improving voice based applications.

More literature on voice based applications and their long term impact is very much needed in order to truly evaluate the lasting impact that ICT4D can have on the communities it aims at helping. It is also necessary to reflect on projects and their factors of success or failure. Indeed, many ICT4D implementations have failed in the past and revisiting them can avoid future failures [5], this includes voice based applications. New technologies have an immense impact when implemented, and it is the duty of their developers to ensure that the process ensures the prosperity of its community.

The need for impact assessment raises the question of methodology for impact assessment. In every paper studied, impact was evaluated through interviews and quantitative data from the applications. They did not follow a set methodology that could be used to objectively compare each project and conclude on their overall usefulness. The introduces a subjectivity to the evaluation which makes it more difficult to argue for more international aid to developing countries in the form of voice based applications.

There exists multiple ways to assess impact in the ICT4D context, of which the capability approach [31] is the most used and adapted to specific context [2] [14]. Some voice based applications are only part of an isolated academic project (e.g. P7 or P20), and therefore their impact is minimal because they are not made to last, only to provide an academic discussion on the need and potential effect of such application. Thus, they have no lasting impact on the community they originally aimed at helping with the developed product.

This raises the question of funding projects. Kick-starting ICT4D projects, especially voice based applications, requires good technical knowledge, some hardware equipment and a lot of background research. This often cannot be offered for free by the people carrying out ICT4D projects. The selected papers describe projects funded by different actors described below:

- **International government**: aid coming from other governments if frequent, especially from the USA, UK and the EU. The projects P15, P16, P17, P18 and P21 benefitted from international aid.
- **International NGO**: most projects are funded through non-government international aid. This can be done by either private foundations, corporations or public institutions giving grants. This is the case for P4, P5, P8, P9, P10, P11, P12, P18, P20, P21, P22, P23 and P24.
- **Local government**: projects can be funded by the the state in which they take place, however it is still rare for voice based applications as only two papers are partially funded by locally governments (P13 and P15).
- **Local NGO**: they are usually big stakeholders in the development of ICT4D projects but they are rarely the ones bringing funds. In our selected papers, the NGO granting resources was mostly IBM Research India which spearheaded many voice based applications with their VoiceSite Technologies (P1, P2, P3, P6, P7, P8) and one was funded by Microsoft Research India (P14). Even through those organizations are locally implemented, they remain international corporations from the United States of America and operate on a global scale.

Funds come from various sources, but mainly from international aid. Even though ICT4D projects are executed in close collaboration with relevant stakeholders to the application such as local NGOs or communities, the first steps of the projects are dependent on outside resources. This adds constraints for developers because the grant might only cover a set period of time and when the funds run out, the project cannot go on unless funds are found again. The issue of economic sustainability is therefore essential to consider when deploying voice based applications for the projects to keep having impact even when money runs out. However, this subject is touched on by less of a third of the articles. This means that few articles studied the durability of their project and they stay reliant on outside funds, often international help.

This paper encourages voice based application developers to design their project in such a way that it creates value for every stakeholder involved once international aid is subtracted from the equation. This will help developing countries gain more independence and thus enter a virtuous cycle of economic and social growth according to the Sustainable Development Goals [1].

When discussing economic viability of their applications, researchers all concluded that the cost incurred by users should be very small to none. Users come from low socio-economic background, thus they are unwilling or incapable of spending money to access a new service, especially if they are unsure of its benefits at

first. If the service provided is not free, it should bring increased value in the users' life.

Another important factor of success is the system's usability. Voice based applications are targeted towards people with low digital skills, therefore the designed system should be kept as simple as possible. Usability is an essential non-functional requirement for any voice based applications. User centric design is highly recommended with multiple iterations of feedback to facilitate the development. This will ensure that the application has a successful launch for its users.

6 CONCLUSION

This study considered recent research on voice based applications in the context of developing countries. They are becoming a popular tool for ICT4D projects and they have been implemented in various settings. During the last twenty years, an important topic in ICT4D has been to evaluate the impact that the projects can have. Through a systematic literature review, we aimed at finding out what kind of impact and challenges can those applications present for their developers and researchers. To select relevant papers, we followed a snowballing approach. This enabled us to identified the most prominent research in the field and to analyse its conclusions.

We found that voice based applications are applied mostly to the agricultural sector in order to improve farming practice or boost sales. In addition to the economic value that they can bring, voice based applications also can also improve social cohesion, health and strengthen the link between remote citizens and the state. However, the evaluation of voice based applications' impact does not follow a set standard and often fails to consider the economic sustainability of the project.

Most of the projects examined in this study were very young when written about. A big avenue for future research would be revisit those projects if they are still in use and to assess their long-term impact and factors of success or failures. Moreover, research into adapting popular approaches (e.g. Capability Approach) for the evaluation of impact of voice based applications could benefit the field.

REFERENCES

[1] [n.d.]. https://sdgs.un.org/goals

[2] Julian Bass, Brian Nicholson, and Eswaran Subrahmanian. 2013. A Framework Using Institutional Analysis and the Capability Approach in ICT4D. *Information Technologies & International Development* 9 (01 2013), 19–35.

[3] Jonathan Steinke Berta Ortiz-Crespo, Jeske van de Gevel Carlos F. Quirós, Majuto Gaspar Mgimiloko Happy Daudi, and Jacob van Etten. 2021. User-centred design of a digital advisory service: enhancing public agricultural extension for sustainable intensification in Tanzania. *International Journal of Agricultural Sustainability* 19, 5-6 (2021), 566–582. https://doi.org/10.1080/14735903.2020.1720474

[4] Anna Bon, Jaap Gordijn, and Hans Akkermans. 2017. *e-Service Innovation in Rural Africa Through Value Co-Creation*. 359–378. https://doi.org/10.4018/978-1-5225-2084-9.ch018

[5] Suzana Brown and Alan Rolf Mickelson. 2019. Why Some Well-Planned and Community-Based ICTD Interventions Fail. *Information Technologies and International Development* 15 (2019), 13. https://api.semanticscholar.org/CorpusID:199762204

[6] InduShobha Chengalur-Smith, Devendra Potnis, and Gaurav Mishra. 2021. Developing voice-based information sharing services to bridge the information divide in marginalized communities: A study of farmers using IBM's spoken web in rural India. *International Journal of Information Management* 57 (2021), 102283. https://doi.org/10.1016/j.ijinfomgt.2020.102283

[7] Victor de Boer, Pieter De Leenheer, Anna Bon, Nana Baah Gyan, Chris van Aart, Christophe Guéret, Wendelien Tuyp, Stephane Boyera, Mary Allen, and Hans Akkermans. 2012. RadioMarché: Distributed Voice- and Web-Interfaced Market Information Systems under Rural Conditions. In *Advanced Information Systems Engineering*. Springer Berlin Heidelberg, 518–532.

[8] Francis Dittoh, Hans Akkermans, Victor de Boer, Anna Bon, Wendelien Tuyp, and André Baart. 2022. Tiba$$\eta a$$sim : $InformationAccessforLow-ResourceEnvironments$.*InProceedings of Sixth International Congress on Information and Comm* −683.

[9] Francis Dittoh, Chris van Aart, and Victor de Boer. 2013. Voice-based marketing for agricultural products: a case study in rural Northern Ghana. In *Proceedings of the Sixth International Conference on Information and Communications Technologies and Development: Notes - Volume 2*. Association for Computing Machinery, New York, NY, USA, 21–24. https://doi.org/10.1145/2517899.2517924

[10] Karn Dubey, Palash Gupta, Rachna Shriwas, Gayatri Gulvady, and Amit Sharma. 2019. Learnings from deploying a voice-based social platform for people with disability. In *Proceedings of the 2nd ACM SIGCAS Conference on Computing and Sustainable Societies*. Association for Computing Machinery, New York, NY, USA, 111–121. https://doi.org/10.1145/3314344.3332503

[11] Mohamed Gulaid and Aditya Vashistha. 2013. Ila Dhageyso: an interactive voice forum to foster transparent governance in Somaliland. In *Proceedings of the Sixth International Conference on Information and Communications Technologies and Development: Notes - Volume 2*. Association for Computing Machinery, 41–44. https://doi.org/10.1145/2517899.2517947

[12] Nana Baah Gyan, Victor Boer, Anna Bon, Chris van Aart, Hans Akkermans, Stephane Boyera, Max Froumentin, Aman Grewal, and Mary Allen. 2013. Voice-based Web access in rural Africa. 122–131. https://doi.org/10.1145/2464464.2464496

[13] Richard Heeks. 2017. *Information and Communication Technology for Development (ICT4D)* (1st ed.). Routledge. https://doi.org/10.4324/9781315652603

[14] Md. Rakibul Hoque. 2020. The impact of the ICT4D project on sustainable rural development using a capability approach: Evidence from Bangladesh. *Technology in Society* 61 (2020), 101254. https://doi.org/10.1016/j.techsoc.2020.101254

[15] Itorobong A. Inam, Ambrose A. Azeta, and Olawande Daramola. 2017. Comparative analysis and review of interactive voice response systems. In *2017 Conference on Information Communication Technology and Society (ICTAS)*. 1–6. https://doi.org/10.1109/ICTAS.2017.7920660

[16] Nishanth Jayarajan, Amanda Puckett BenDor, Norbert Boruett, L Duncan, M Kinyua, Lisa Mwaikambo, and A Lee. 2017. Use of interactive voice response for professional development in Kenya. https://api.semanticscholar.org/CorpusID:169099836

[17] Anirudha Joshi, Mandar Rane, Debjani Roy, Nagraj Emmadi, Padma Srinivasan, N. Kumarasamy, Sanjay Pujari, Davidson Solomon, Rashmi Rodrigues, D.G. Saple, Kamalika Sen, Els Veldeman, and Romain Rutten. 2014. Supporting treatment of people living with HIV / AIDS in resource limited settings with IVRs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, 1595–1604. https://doi.org/10.1145/2556288.2557236

[18] Konstantinos Kazakos, Siddhartha Asthana, Madeline Balaam, Mona Duggal, Amey Holden, Limalemla Jamir, Nanda Kishore Kannuri, Saurabh Kumar, Amarendar Reddy Manindla, Subhashini Arcot Manikam, GVS Murthy, Papreen Nahar, Peter Phillimore, Shreyaswi Sathyanath, Pushpendra Singh, Meenu Singh, Pete Wright, Deepika Yadav, and Patrick Olivier. 2016. A Real-Time IVR Platform for Community Radio. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, 343–354. https://doi.org/10.1145/2858036.2858585

[19] Arun Kumar and Sheetal K. Agarwal. 2012. Spoken web: using voice as an accessibility tool for disadvantaged people in developing regions. *SIGACCESS Access. Comput.* 104 (sep 2012), 3–11. https://doi.org/10.1145/2388818.2388819

[20] Arun Kumar, Nitendra Rajput, Dipanjan Chakraborty, Sheetal K Agarwal, and Amit Anil Nanavati. 2007. VOISERV: Creation and Delivery of Converged Services through Voice for Emerging Economies. In *2007 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks*. 1–8. https://doi.org/10.1109/WOWMOM.2007.4351733

[21] Michael A. Madaio, Evelyn Yarzebinski, Vikram Kamath, Benjamin D. Zinszer, Joelle Hannon-Cropp, Fabrice Tanoh, Yapo Hermann Akpe, Axel Blahoua Seri, Kaja K. Jasińska, and Amy Ogan. 2020. Collective Support and Independent Learning with a Voice-Based Literacy Technology in Rural Communities. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, 1–14. https://doi.org/10.1145/3313831.3376276

[22] Megh Marathe, Jacki O'Neill, Paromita Pain, and William Thies. 2015. Revisiting CGNet Swara and its impact in rural India. In *Proceedings of the Seventh International Conference on Information and Communication Technologies and Development*. Association for Computing Machinery. https://doi.org/10.1145/2737856.2738026

[23] Neil Patel, Deepti Chittamuru, Anupam Jain, Paresh Dave, and Tapan S. Parikh. 2010. Avaaj Otalo: a field study of an interactive voice forum for small farmers in rural India. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, 733–742. https://doi.org/10.1145/1753326.1753434

[24] Neil Patel, Kapil Shah, Krishna Savani, Scott R. Klemmer, Paresh Dave, and Tapan S. Parikh. 2012. Power to the peers: authority of source effects for a voice-based agricultural information service in rural India. In *Proceedings of the Fifth International Conference on Information and Communication Technologies and Development*. Association for Computing Machinery, New York, NY, USA, 169–178. https://doi.org/10.1145/2160673.2160696

[25] Vishnu Prasad, Richard Shallam, Alok Sharma, Delvin Varghese, and Devansh Mehta. 2021. A Hybrid Multi-Modal System for Conducting Virtual Workshops Using Interactive Voice Response and the WhatsApp Business API. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Article 408. https://doi.org/10.1145/3411763.3451820

[26] Muhammad Qasim, Haris Bin Zia, Awais Athar, Tania Habib, and Agha Ali Raza. 2021. Personalized weather information for low-literate farmers using multimodal dialog systems. *International Journal of Speech Technology* 24, 2 (June 2021), 455–471. https://doi.org/10.1007/s10772-021-09806-2

[27] Nitendra Rajput, Sheetal Agarwal, Arun Kumar, and Amit Anil Nanavati. 2008. An alternative information web for visually impaired users in developing countries. In *Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility*. Association for Computing Machinery, New York, NY, USA, 289–290. https://doi.org/10.1145/1414471.1414542

[28] Luthfi Ramadani and Christoph F. Breidbach Sherah Kurnia. 2018. In Search for Holistic ICT4D Research: A Systematic Literature Review. In *Proceedings of the 51st Hawaii International Conference on System Sciences*. IEEE. https://doi.org/10.24251/HICSS.2018.304

[29] Shan M Randhawa, Tallal Ahmad, Jay Chen, and Agha Ali Raza. 2021. Karamad: A Voice-based Crowdsourcing Platform for Underserved Populations. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery. https://doi.org/10.1145/3411764.3445417

[30] Agha Ali Raza, Bilal Saleem, Shan Randhawa, Zain Tariq, Awais Athar, Umar Saif, and Roni Rosenfeld. 2018. Baang: A Viral Speech-based Social Platform for Under-Connected Populations. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, 1–12. https://doi.org/10.1145/3173574.3174217

[31] Amartya Sen. 1985. *Commodities and Capabilities*. North-Holland, Amsterdam. New Delhi: Oxford University Press, 1987; Italian translation: Giuffre Editore, 1988; Japanese translation: Iwanami, 1988.

[32] Jerome White, Mayuri Duggirala, Krishna Kummamuru, and Saurabh Srivastava. 2012. Designing a voice-based employment exchange for rural India. In *Proceedings of the Fifth International Conference on Information and Communication Technologies and Development*. Association for Computing Machinery, New York, NY, USA, 367–373. https://doi.org/10.1145/2160673.2160717

[33] Claes Wohlin. 2014. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*. Association for Computing Machinery, New York, NY, USA, Article 38, 10 pages. https://doi.org/10.1145/2601248.2601268

[34] Martin Wolfe. 1955. The concept of economic sectors. *The Quarterly Journal of Economics* 69, 3 (1955), 402. https://doi.org/10.2307/1885848

[35] Moonjung Yim and Ricardo Gomez. 2022. ICT4D evaluation: its foci, challenges, gaps, limitations, and possible approaches for improvement. *Information Technology for Development* 28, 2 (2022), 251–274. https://doi.org/10.1080/02681102.2021.1951151