Conversational interfaces and the long tail of languages in developing countries

The opportunities and challenges in the application of Natural Language Processing in low income countries

Adam Wills and Georgia Barrie of Farm.ink and Jake Kendall of the DFS Lab
Executive Summary
Executive Summary

In recent years, researchers and technology companies have made enormous advances in “Natural Language Processing,” the ability of computers to understand the human voice and language. Amazon’s Alexa, Apple’s Siri, and Google’s Home assistant have shown an incredible ability to understand people’s voice commands. They’ve also shown that major companies can create business models around selling voice-activated tools and services.

While Amazon and Apple were developing Alexa and Siri, people in the developing world have continued to connect to digital platforms in increasing numbers. Many of these people have made messaging platforms, including WhatsApp and WeChat, the primary methods for accessing digital communication.

This paper explores the ability to use Natural Language Processing to create conversational interfaces for digital products and services used by low-income people. Companies may be able to use artificial intelligence and natural language processing to improve agriculture, disaster response, health, and financial services for low-income people. Governments and NGOs may also be able to use these technologies to improve service design and delivery.

To achieve these goals, researchers and companies will need to overcome different barriers than the technological problems facing natural language processing in developed countries. These include:

• The fact that many low-income people use feature phones and low-end smartphones, rather than the high-end devices used in high-income countries.

• Commercial incentives driving NLP researchers and companies to focus on common languages like English, rather than the languages spoken by low-income people. This creates an increasing divide between the NLP products targeted at high- and low-income consumers.

• Informal language patterns used by low-income people, including code mixing, non-fluent English speakers, and non-standard spellings.

International donors and policymakers have a role to help address these barriers and unlock the power of NLP to improve the lives of low-income people. They can support the creation and sharing of datasets that can be used to improve NLP in languages more common among low-income people around the world. They can also engage private-sector actors to help direct more NLP research for low-income settings and support innovative experiments testing NLP in low income settings.
Opportunity:
“Alexa, how can I protect my child from malaria?”
“Siri, please send money home to my family in Ethiopia.”

In recent years, technology companies such as Amazon, Apple, and Google have released hugely popular products that use natural language processing and artificial intelligence. Amazon’s Alexa service, for example, can analyze voice commands with high degrees of accuracy.

In principle, anyone with a connected device should be able to interact with these systems. However, the reality is that, “Alexa/Siri/Cortana: tell me how to protect my children from malaria” works better, uttered in English, in Minnesota (where there is no malaria) than it does uttered in a language spoken where malaria actually exists. In fact, it appears that the gap between the aspirations for Artificial Intelligence in the English-speaking north and the low-resource language south is not only large but widening.

This paper explores why. It is based on desk research and discussions with experts in academia, business, and international development. We explore the state of “Natural Language Processing” (NLP) technology — the ability of computer software to interpret human language — across the developed and developing world and we analyse the reasons behind the emerging language divide. We also examine potential use cases for conversational technology in international development, and case studies of early successes. Finally, we outline what is holding back the application of NLP in international development, and how policy makers and international donors could reduce these barriers.
The conversational interface opportunity

The way people access the internet is changing fundamentally. While many people today rely on menus and buttons, voice interface may soon become a standard way for people to interact with mobile and web tools. The proliferation of connected devices and the availability of new AI technologies to create intuitive conversational experiences is what is driving this change in the developed world. These same technologies alongside the rapid growth of messaging apps and proliferation of phones — could create a similar opportunity in the developing world. Though there are a number of fundamental new challenges we document below that will have to be negotiated before this opportunity will be realized.
Conversational interfaces and the long tail of languages in developing countries

The conversational interface opportunity

“Pretty much everyone today who is building applications whether they be mobile apps or desktop apps, or websites will build bots as the new interface, where you have a human dialogue interface versus menus of the past”

Satya Nadella, CEO, Microsoft

The opportunity for messaging

Many new opportunities for NLP will be tied to new messaging apps which are gaining significant traction with developing world consumers. Today, the top four social messaging apps (WhatsApp, WeChat, Messenger and Line) have over three billion active users.\(^1\) The growth of social messaging apps has outpaced the growth of social networking apps (such as Facebook and Twitter), making messaging the most prevalent user experience across the globe.\(^2\) This phenomenon is more pronounced in the developing world than it is in the developed world. In India the average time spent on messaging apps is over three quarters that of social networking apps while the reverse is true in the US.\(^3\)

When consumers interact with the internet primarily through messaging apps, their experience is fundamentally different from those who experience the internet through traditional desktop browsers and social media. One reason is that many messaging apps are mobile first, rather than desktop first. This creates a premium for language-based interfaces associated with chat, rather than more graphics rich-interfaces on browsers. Among low-income consumers, these pared-down interfaces are especially beneficial, as many people use low-end phones and may be severely limited in data usage.

Consequently, messaging apps are growing fastest in the Global South: India is WhatsApp’s biggest market, with over 200 million monthly active users.\(^4\) By the end of the year China’s WeChat could reach one billion active users.\(^5\) Many providers are taking advantage of the ubiquity of these apps to build new services and functionality over the top of traditional messaging.

China’s WeChat is leading the charge. The company added mobile payments and opened APIs to chatbot developers as early as 2013. WeChat has moved from a messaging app to a platform, providing everything from gaming to ecommerce. On an average day, WeChat makes up approximately 30% of all time spent on apps in China.\(^6\) China’s mobile payments market is 50 times that of the US, and WeChat now has a 40% market share.\(^7\)

---

3. Social Networking Apps or Chat Apps – What’s Your Pick?
5. “WeChat: Number of Users 2018.” Statista.
Other global messaging platforms are following in WeChat’s footsteps. Facebook opened a Messenger API, launched a payments integration, and is now testing payments for WhatsApp. Telegram opened a messaging API in 2015, and Line followed in 2016.

With messaging platforms already dominating smartphone use in many emerging markets, the addition of payments and chatbot APIs are likely to further this trend. One study found as many as two million people in India are already using WhatsApp and Facebook to run small-scale e-commerce businesses.

“Early on, we named Facebook, Twitter and similar sites ‘social media’ because ‘media’ was the only framework we had for mass communication... We may be headed down a similar path with messaging apps. We’ve named WhatsApp and the like ‘messaging apps’ because that’s our current context for them... But what if the larger story behind messaging apps is not that they remove barriers to sending messages but that they remove all barriers to digital interactions and transactions”

Meghan Keaney Anderson, VP Marketing, Hubspot
Conversational interfaces and the long tail of languages in developing countries

The conversational interface opportunity continued

What accounts for the dominance of messaging apps in the developing world?

This trend may be due, in part, to a lack of storage space on lower-cost smartphones. It may also be due to a lack of familiarity with the Google Play Store, which creates barriers to using apps that are not pre-installed on the handset. The cost of data is also a significant barrier to use among lower-income users, which may result in greater attention placed on data-light platforms such as WhatsApp. Mobile operator efforts to offer free data on services such as Facebook, Twitter and WhatsApp may also be skewing usage in favor of these apps.

The opportunity for voice

In May 2016, voice commands initiated an estimated 1 in 5 searches on mobile in the US, and this share continues to grow. By 2020 it is estimated that at least 50% of searches will be made either through images or speech.

The opportunity for voice as a computer interface may be even greater in the developing world than it is in the US. A survey in the US found that 24% of users of voice assistants said the primary reason for use was due to “difficulty typing on certain devices” and 12% said “to avoid confusing menus”. These percentages could be considerably higher among users of cheaper, less-user-friendly handsets and users with lower levels of tech literacy.

Voice is also a more accessible interface than written text for populations with lower levels of education or who are not fluent in one of the languages that dominate online text. Ease of use may be the reason why Interactive Voice Response (IVR) has been popular with developing-world users in recent years. Sophisticated, high-accuracy voice interfaces could help people with difficulty reading and typing access digital services.

The cost of data could be a barrier to the widespread use of voice-controlled interfaces, as a voice message is a larger data package than a text message. However, as we have seen with messaging mobile network pricing structures will likely affect behavior. Offering free data for voice-controlled interfaces or bundling of services could significantly reduce the cost barrier to use.

---

12 Mozilla Digital Skills Observatory
13 "2016 Internet Trends Report." Kleiner Perkins
14 “Intelligent Voice Assistants Research Report - Q1 2016” MindMeld
Natural language processing (NLP) technology is necessary to deliver conversational services. In recent years, researchers and technology companies have vastly improved the ability for computer software to interpret human language. Some machines are now as good as humans in areas such as translation and speech recognition. In principle, anyone with a smartphone will be able to interact with these systems. However, there is a large and widening gap between the capabilities of NLP technology for English and European languages versus the long tail of low-resource languages.
The role of Natural Language Processing

Defining NLP and AI

Artificial Intelligence (AI) is defined as computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages. In this sense, Natural Language Processing is essentially a subset of Artificial Intelligence.

It is important to note that NLP techniques range considerably in their sophistication and purpose. The simpler end of the spectrum includes tools to identify specific, relevant terms in a body of text, or to label text extracts by category. Many of these techniques have existed for some time and are commonly used in services such as search engines.

The most sophisticated end of the NLP spectrum includes recent advances such as machine translation and natural language generation and understanding. These tools often require computer programs to understand both the meaning of each individual word and the meaning of an entire statement.

An ecosystem of major players

Natural language processing has traditionally been the domain of academics. Major advances have been made by teams such as the Stanford Natural Language Processing Group, who developed the Stanford Core NLP technology tools. However, while academia plays a key role in the advancement of NLP techniques, major tech players have begun to use their vast troves of computing power and data to advance this field.

The state of NLP technology

One of the most important advances in the increasing accuracy of natural language processing has been the advent of Deep Learning. Deep Learning aims to emulate the functioning of the brain through the creation of artificial neural networks.

Google has begun to use Deep Learning to improve the accuracy of its machine translation. The original Google Translate used Phrase-Based Machine Translation (PBMT), a classical NLP technique. Google asked people to evaluate translation results on a scale from 0 to 6 (where 0 means nonsense and 6 means perfect translation)
The role of Natural Language Processing continued

and found that English-to-Spanish translation was rated with an average score of 4.885. When Google Translate switched to using Deep Learning techniques, this average increased to 5.428. Language translated by humans was given an average of 5.504, meaning that Google’s translation engine is now nearly as accurate as humans for English-to-Spanish translation.\textsuperscript{16}

![Translation quality chart](image)

**Comparing translation models**

*Source: Google*

Technology companies have also begun using Deep Learning for speech recognition, which dramatically increased accuracy. In October 2016, researchers announced that human parity had been reached for conversational speech. The study compared Microsoft’s NIST 2000 automated system against professional transcriptionists and found for the first time a higher error rate among humans than computers.\textsuperscript{17}

Similarly, in May 2017 Google announced that the company’s speech recognition technology had achieved a 4.9% word error rate, a significant improvement from the 23% error rate reported in 2013.\textsuperscript{18}

![Word error rate chart](image)

**“Voice UI follows a U-shaped curve: one command is great and ten is probably OK, but 50 or 100 is terrible, because you still can’t ask anything but can’t remember what you can ask. The other end of the curve comes as you get closer and closer to a system that really can answer anything, but, again, that would be ‘general AI’”**

*Benedict Evans, Andreessen Horowitz*

While researchers and technology companies have achieved major advances in fields such as translation and speech recognition, a realistic human-style conversation is still outside of our reach. Natural Language Understanding, the ability for computer software to understand the meaning of a statement, is highly complex. Most NLP tools are designed so developers must train the technology to understand the different types of user questions or statements. In some cases, this means programming in thousands of different ‘intents’ and responses.

\textsuperscript{16} Google’s Neural Machine Translation System: Bridging the Gap between Human and Machine Translation
\textsuperscript{18} Pottalinsky, Emil. “Google’s Speech Recognition Technology Now Has a 4.9% Word Error Rate.” VentureBeat, VentureBeat, 01 May 2017
The role of Natural Language Processing continued

"We talked about the digital divide. Now we’re seeing a language divide... and that language divide is just getting bigger and bigger"

Kalika Bali, Microsoft Research India

The language divide

NLP researchers have focused much of their effort on the English language. Therefore, capability of NLP tools to interpret written and spoken English is fast outpacing the same capabilities in other languages. This research gap creates a large and widening language divide.

We can get a sense of the size of the language divide by looking at the language coverage of some of the major NLP platforms. Analysis shows there is a clear income gap in access. Google Translate is available in the languages spoken by just 54% of those living on less than $1.90 per day. The picture is even more stark for Natural Language Understanding frameworks such as Dialogflow, which supports languages spoken by only 3% of those living on less than $1.90 per day.

When you factor in the quality of language coverage, this picture will be even more weighted towards languages spoken by the higher income West. While it is difficult to find hard data comparing the quality of language provision, the quality of English-to-Spanish translation, for example, will almost certainly be higher than it is for Igbo-to-English translation. Similarly, the quality of support for English in NLP middleware platforms such as Wit.ai is likely much better than, say, support for Swahili.

Percentage of those living on less than $1.50/day whose language is covered by the tool

Source: World Bank, Ethnologue and software provider websites

When an NLP program, like a chatbot, understands some commands and not others, users may get confused or frustrated. User testing of chatbots built with Facebook’s DeepText early in 2017 found that only 30% of requests could be fulfilled without human agents.

For this reason, most successful chatbots have a very specific niche (e.g. weather forecasting) and explain to the user which commands can be understood and which cannot (e.g. some ask, “Why don’t you ask me what the weather will be like tomorrow?”). Facebook have stated that they plan to “narrow down the set of cases so users aren’t disappointed by the limitations of automation.”

FastText (Facebook)

Google Translate

Wit.ai (Facebook)

Microsoft Bot Framework

Dialogflow (Google)

169 languages 64%

97 languages 54%

49 languages 30%

47 languages 24%

12 languages 3%
There is a fundamental tension between the commercial and social value of languages. There is unclear return on investment in creating tools for languages spoken by smaller, lower-income populations. This dynamic is compounded by the fact that these languages suffer from a lack of data available to train machine learning systems.
The barriers to NLP for low-resource languages

The incentive problem

Discussions with platform providers highlight the fact that many struggle to see the return on investment in developing services for low-income populations or less-widely-spoken languages. While more common languages, such as Hindi and Swahili, are likely to be in the pipeline for many major NLP providers, business cases for less-common languages may never emerge.

We can calculate an approximate commercial value of a language by multiplying the number of speakers by the national GDP per capita of those speakers. Ordering languages by commercial value finds that the top 100 languages cover approximately 96% of global GDP. However, these 100 languages cover less than 60% of the total population living on less than $1.90 per day as shown in Figure 1. This highlights a fundamental tension between the commercial and social value of languages.

Currently large NLP platforms such as Google Translate often rely on volunteers to generate datasets in lower-resource languages. For example, Google saw there was demand from people in Kazakhstan to have Kazakh represented on Google Translate, but did not have enough labeled data available. They crowdsourced translations from English speakers in Kazakhstan through the “Translate Community” to produce two million validated English to Kazakh translations. According to Barak Turovsky, product lead at Google Translate, at least 14 other languages have now been added in this way.

“We were skeptical. We created a micro task—sent it to 300–400 people asking them to produce two million validated translations. To our surprise, we started getting 200,000 contributions a day. We wondered how. Apparently, there was a press conference in the office of the President who called everyone who knew English in Kazakhstan, and cajoled them to contribute. That’s how we got Kazakh.”

Barak Turovsky, Product Lead, Google Translate

Another issue is the lack of academic researchers working on NLP for lower-resource languages. Currently there appears to be a greater incentive for academics to improve the NLP technology that exists for English than there is to improve the resources available for lower-resource languages.

“Many academics just want to prove that they can get 2% more accurate than the current baseline. There’s little incentive to start from scratch with a whole new language.”

Robert Munro, VP of Machine Learning, Crowdflower
The barriers to NLP for low-resource languages continued

“Technology wise there isn’t really any bottleneck. We know how to build these systems, we know how to train these systems...but for many languages the data is just not there”

Kalika Bali, Microsoft Research India

The data problem

The amount of data required to train sophisticated NLP tools is vast. In the case of supervised Deep Learning, the neural network requires ‘labeled training data.’ For example, a speech recognition neural network needs to be fed labeled, recorded speech, which it analyzes and learns. The more data it is fed, the more accurately it can recognize words as they are spoken.

According to interviews with researchers working in this space, a human-parity speech recognition tool requires approximately 100 thousand hours of recorded speech. Building accurate machine translation for a language requires millions of validated translations, and as NLP tools increase in sophistication, the amount of data required increases, too.

A long legacy of internet usage in the West has generated large datasets available to train Deep Learning systems. Billions of searches and social media posts, for example, have enabled tech platforms such as Google and Facebook to build highly sophisticated systems. These datasets are much smaller for languages such as Swahili and Hindi and relatively non-existent for the long tail of languages spoken in much of Africa and Asia.

The rapid spread of messaging apps such as WhatsApp and social networking platforms in the developing world is growing the amount of data that exists in lower-resource languages. However, there are two key problems: Firstly, the data generated is typically locked inside the messaging app or social networking app in which it is generated, and app providers have, to date, been reticent to share this data. Secondly, the data that is generated through these apps is typically of low quality, meaning that it is not well labeled. High-quality, labeled data is required for training advanced neural networks.

English is a bad place to start

“A lot of people launch in English, because they want to get one language working and then expand from there. Then they realize that because of the way they’ve coded the system they’ve painted themselves into a corner”

Robert Munro, VP Of Machine Learning, Crowdflower

Another factor holding back the development of non-English NLP tools is that the particular quirks of the English language make it a very bad place to start when creating a multilingual service. A number of factors mark English as an outlier, including the fact that English has the largest vocabulary of any language. Verbs are largely uninflected and there is a complicated article system. Unlike in a highly inflected language, such as German, meaning in English is conveyed largely by word order, which requires a very different approach to NLP.25

One of the other key differences between English and other languages is that English spellings are far more standardized. While there are some differences, such as British vs. Americanized spellings, these differences

24 Google Translate
The barriers to NLP for low-resource languages continued

are minor compared to the variation found across other languages. For example, a medical service delivered in Chichewa language found users of the service used over 30 different terms and spellings for the word patient.26

Further challenges in the developing world

There are several other phenomena which make NLP provision in developing markets more difficult than in the West.

Code-mixing — the use of more than one language in a single conversation or utterance — exists in all multilingual societies. In a Facebook dataset analyzed by Bali et al. (2014), almost all sufficiently long conversation threads originating from India were found to be multilingual. As much as 17% of the comments contained code-mixing.

While code-mixing is commonplace in much of the developing world, it is much rarer in the West. That means that relatively little research is being done in this area. There are examples of efforts in this space, such as Microsoft’s Project Mélange research group based in India. However, these efforts are small compared with the investment in English and European languages.

Another common phenomenon is the use of a second or third language online. Many people whose native language is less widely spoken may prefer to use a more common language online such as English. For example, three quarters of Facebook users in Algeria frequently post in French,27 even though Algerian Arabic and Berber are the native languages of over 99% of Algerians.28 This can create challenges for NLP, as users may be less fluent in their ‘online language.’ Non-native users may use non-traditional spellings and sentence structures.

The use of non-traditional spellings becomes an even greater challenge when languages don’t use the same lettering system or script. Many users have resorted to “romanization,” the conversion of writing from a different writing system to the Roman (Latin) script. Analysis by Bali et al. (2014), found that Hindi is seldom written in the Devanagari script in social media posts. The language is far more likely to be loosely translated into Roman script, because it is easier to type on a standard phone. One study in India reported that approximately 50% of users surveyed found typing in English ‘very efficient’ or ‘extremely efficient,’ meanwhile only 20% of users said the same for typing in an Indic language.29
The use of NLP tools in development is at an early stage, but important use cases are starting to emerge. While it is difficult to find many large-scale implementations of NLP, there are organizations currently experimenting with this technology.
### NLP for development

#### Agriculture

The agricultural space in lower-income markets has seen huge growth in user generated activity online. Farmers from Kenya to India have self-organized in large, topic-focused groups, notably on Facebook and WhatsApp. For example, there are currently more than 19 million Facebook users with an interest in agriculture in lower, middle-income markets and below.

These online communities create new opportunities to deliver tailored information and smarter agronomic services to farmers. They also allow companies and researchers to collect and analyze data to identify market trends previously impossible to track.

“When we analyzed the most effective digital services for farmers in Kenya it was WhatsApp and Facebook. We realized the opportunity was to augment these experiences, not redesign them”

Georgia Barrie, Co-Founder, Farm.ink and co-author of this article

---

### Farm.ink

**Chatbots for small-scale farmers**

Across East Africa, farmers have begun to create large communities online. Some have formed “mega-groups” of 100,000s of active members on Facebook, others have joined smaller groups across services like WhatsApp. These groups are typically forums for farmers to ask and answer questions about farming practices. These groups also often serve as marketplaces for the buying and selling produce.

While these forums provide a tangible service to users, much of the value of the data is lost as content is not indexed or searchable.

Farm.ink developers use a range of NLP tools — including sentiment analysis, text-search engines, and text-based classifiers — to label social content and make it discoverable. Using a chatbot, farmers can ask a question about farming and receive the closest matches to their question from the tens of thousands of user generated posts.

A recent survey conducted by Farm.ink found that over 90% of active users reported that they had received information through the chatbot that had changed the way that they farm.

In 2018 Farm.ink partnered with the International Livestock Research Institute (ILR) with the aim of improving the ability to identify and track livestock disease in Kenya. Farm.ink is working with ILRI livestock experts to create a labeled data set of social media livestock disease reports. This data will be used to train a NLP algorithm to automatically identify new livestock disease reports on social forums.
Disaster response

NLP can be a powerful tool for disaster response efforts, including the creation of systems to interpret and respond to populations in crisis. Notably, in the 2010 earthquake in Haiti, while local emergency services were inoperable, 70–80% of cell-towers were quickly restored. People were texting, calling, and interacting with social media, primarily in the Haitian Kreyol language. The tech companies Crowdflower and Ushahidi built a platform for translating and mapping these reports. The idea was to bridge the gap between the Kreyol-speaking population and agencies tasked with delivering aid.

Health

Health practitioners already employ NLP tools in developed markets, from automating responses to patient enquiries to depression screening using social-media data. In low-income markets, where health information is lower quality and costs more to access, practitioners may be able to use some of the same approaches. NLP may be able to improve on digital rule-based health information services in these environments. To achieve this, however, the NLP tools will need to ensure quality control of content and accommodate users’ first language.

“When an issue is more serious the woman tends to switch to Swahili because she feels more comfortable explaining it in her first language”

Sathy Rajasekharen, Chief Innovation Officer, Jacaranda Health

CASE STUDY

Jacaranda Health
Kenya-based social enterprise providing maternity clinics

Jacaranda Health have set up WhatsApp groups for each of their maternity groups with a qualified doctor playing a moderating role. The groups have proven to be highly successful in improving the women’s access to maternal health information and in providing a way for professional moderators to identify potentially critical health problems outside of the clinics.

However, moderating the groups can be resource intensive, and there is a risk that important issues are missed. Jacaranda Health saw an opportunity to use NLP to identify when anyone in the group cites a potentially dangerous symptom and to flag this to a health professional.

Jacaranda Health is still experimenting with NLP tools. So far, they have used a WhatsApp tool to extract data from group conversations and have manually flagged messages in the conversations that relate to critical issues. The team have inputted text from these flagged messages into an online platform called Dialogflow, which is designed to use NLP to automatically identify similar messages.

Jacaranda Health believes NLP could help. However, there are a number of complicating factors: Firstly, WhatsApp does not yet have an API, which means that it is not possible to connect an NLP tool directly to the WhatsApp conversation. One solution is to move the groups onto a platform with an API, such as Facebook Messenger or Telegram, but such a move risks disrupting the group.

A second complication that people often use languages that are hard for traditional NLP tools to interpret. While most of the conversation happens in English, users frequently switch to Swahili. This switch can even happen part way through a message. The conversation also often makes use of highly localized terminology, which can present a challenge for off-the-shelf NLP tools.
Financial services

Mobile messaging platforms could also integrate NLP into payment tools, possibly improving the reach and effectiveness of financial services. Chatbots may be well-equipped to guide users through transaction funnels or educational programs that could improve customer engagement for various financial institutions and products. The fintech company Teller have designed a chatbot to do precisely this, currently focused on users in West Africa and Madagascar. Bots could change the economics of service delivery for products like micro-insurance, loans, or customer support queries. Just as micro-health insurance services are now delivered through mobile telecom providers using airtime, there is likely a set of innovative use cases for delivering financial services at greater scale using bots.

Other use cases

There is a wide range of further NLP use cases that can be explored in other development contexts, too. These use cases include education legal support, and new types of impact measurement. Many of these use cases are nascent but show potential in lower-income markets in the next few years. Unfortunately, the lack of support for different languages will fundamentally limit service delivery in many markets. And while individual use cases will serve to raise awareness of the impact of NLP tools to lower income audiences, they are unlikely to move the needle when it comes to support of low-resource languages on an individual basis.

Teller
Chatbot for financial institutions

Teller began life focused on the American market with the aim of making money management easier and more understandable. Working with financial institutions Teller designed chatbots to enable users to type questions about finance e.g. “What is an overdraft?” and receive explanations in simple, human language. While Teller saw some success in the US, they quickly realised that the opportunity to deliver such a service was far greater in the developing world. Unlike in the US, many people in the developing world do not have easy access to customer service call centres or relevant websites. Teller has recently pivoted to focus on West Africa as a priority region.

Teller built a chatbot that uses Google’s natural language platform Dialogflow to extract entities and intents from user inputs. The NLP technology can then match these entity and intent values against a predefined list, created by Teller, to select the best response to the question. While Dialogflow has proven to be a useful tool, this requires a substantial amount of work. Teller created over 150 intents, relating to a wide spectrum of possible questions about finance. Due to the broad nature of human conversation there are still many questions that cannot be answered by Teller’s bot. Therefore, Teller is working on techniques to guide users through a more structured conversation.

One of the biggest challenges Teller faced in using the platform is the limited number of supported languages. While European languages, such as French, can be parsed by Dialogflow, no African languages are yet supported. Teller is currently working with a financial institution in Madagascar and has designed the chatbot to work in French due to the lack of support for Malagasy.

Another challenge has been the difficulty navigating a nascent and disparate NLP space. Information about the features and capabilities of different platforms has been difficult to find. So far, Teller has spent a fair amount of time testing different services. Finding and consolidating other relevant information — such as the features of SMS APIs, penetration of messaging platforms, and local financial terms across African markets — has also required a significant amount of time and effort.
The skew toward higher-income Western markets in the provision of NLP tools will slow its application in the development space. This skew is apparent not only in language coverage, but also in the design of many NLP platforms and tools. However, a more challenging issue may be the fact that NLP remains a nascent space and its successful application often requires a deep level of expertise.
What’s holding back NLP for development

Language provision

The lack of language support for many languages outside of the developed world is already hindering NLP adoption in the development space. Chatbot provider Teller built its chatbot in English and French, even though many of their users natively speak other languages.36 The tech startup Farm.ink has, to date, focused on the Kenyan market due to the widespread use of English.37

In the example of Crowdflower delivering a disaster response monitoring service in Haiti, humans were employed to translate messages received in real time from Haitian Kreyol to English.38 Better language provision could radically improve this situation.

A US-first design

Another factor affecting the application of NLP in development is the fact that many platforms are designed with a US-first focus. For example, Facebook Messenger opened its API in April 2016 and launched payments for the US market a few months later. At the end of 2017 most payments features for Messenger are only available for US-based chatbots.39

Similarly, WhatsApp is the most widely used messaging app in much of the developing world, but Facebook has yet to open an API for WhatsApp chatbots. This means that bot developers in Africa must design for Facebook Messenger or Telegram, despite the fact that they have lower penetration and usage rates.

A nascent space

A fundamental challenge to those employing NLP tools for development is the fact that this is a highly fragmented space. A plethora of tools are available, each with its own features and constraints. Prospective users of these tools often have to spend considerable time researching what is available and testing potential applications. The fact that NLP is nascent means that the availability and capability of tools is also changing rapidly.

Over time, it is likely that we will see consolidation of tools and the emergence of dominant platforms. The purchase of Wit.ai by Facebook in 201540 and Api.ai (now Dialogflow) by Google in 201642 may be early signs of this trend.

Sources:
36 Teller
37 Farm.ink
38 Crowdflower
39 “Payments (Beta) – Messenger Platform – Documentation,” Facebook for Developers

Countries in green are those where WhatsApp is the most popular messaging platform

Source: Android App Data
What’s holding back NLP for development continued

Complexity of AI technology

The biggest challenge currently faced by NLP product developers is the technological complexity. Understanding the right NLP techniques to use in the right contexts requires not just a programming background, but a fairly deep knowledge of the NLP space.

It is no coincidence that most people implementing NLP for development have either a background in AI or have partnered with others who do. Tech startup Farm.ink, for example, has recruited a team of volunteer NLP experts based in the UK to advise on the application of NLP tools.43

Over the next few years, we will likely see more off-the-shelf tools reduce the complexity of implementing the technology. For now, however, expertise is still needed.
Donors and policy makers can support the development and use of NLP tools for lower-resource languages.
Suggestions for donors

Increase awareness of the social impact of languages

While accurate global data on the size and income levels of language populations does not exist, we can roughly approximate the social impact of including different languages in NLP platforms as the number of low income speakers of that language. Combining country-level poverty data with national language splits, we can estimate a language’s coverage of the world’s poor. This analysis is captured in Figure 1 above and shows that 30% of the world’s poor speak just eight languages, and 50% of the world’s poor speak just 21 languages.44

In the future, market forces will likely lead to better coverage of these popular languages. However, progress is slow. Increased awareness of the social impact of NLP tools for these languages could spur greater commercial and academic interest.

Encourage data sharing for low-resource languages

Much of the textual data that is created in the long tail of languages is locked inside social apps such as Facebook and WhatsApp. Currently there is very limited external sharing of this data.

There is an opportunity to encourage the large platform players (e.g. Google, FB/WhatsApp) to share data for low-resource languages, particularly with the academic and non-profit sector, especially as there is limited commercial value in building NLP tools for these populations.

“If there’s one thing you could just throw money at, it’s creating good datasets in low-resource languages”

Robert Munro, VP of Machine Learning, Crowdflower

Small amounts of funding can trigger innovative new ideas and approaches

The client demand for conversational interfaces to improve social safety nets and G2P programs, as well as things like credit reference bureaus and banks, is strong and solves a real need. Small amounts of money can trigger innovative, new ideas and approaches. However, more is needed to scale and trigger wide adoption.

Providing small amounts of seed funding to people building conversational technology for social good can be a low-risk and cost effective way to catalyze innovative new solutions. With funding from the Gates Foundation, the DFS Lab recently ran a small grants program supporting exploration of innovative use cases for conversational interfaces and found significant progress could be made with relatively limited funding due to the proliferation of off the shelf NPL tools.45

Opportunity to better share knowledge across providers

There is no existing resource center or community of practice focused on applying conversational interfaces in DFS or development more broadly. Conversations with social enterprises and startups have highlighted that currently little information is shared across these entities.

There is an opportunity to create a community of practice and resource center to help these initiatives learn from each other and propagate helpful tools.

Suggestions for donors
continued

DFS Lab is taking one step in this direction by creating a resource center for use of NLP in financial inclusion and associated Facebook community group. Visit here to learn more.
Suggestions for donors
continued