Internet
to the Rescue of Unscripted
Languages

Preface

Internet to the Rescue of Unscripted Languages

Report Prepared by: Osama Manzar, Udita Chaturvedi & Abner Manzar

Design & Layout by: Satish Kumar

Read it online at www.defindia.org/publication-2

Contact:
Digital Empowerment Foundation
House No. 44, 2nd Floor
Kalu Sarai, (Near IIT Flyover)
New Delhi – 110016
Tel: 91-11-42233100
www.defindia.org

The Best Practice Forum (BPF) on Local Content at the United Nations Internet Governance Forum explored how the Internet can be used to preserve local language and cultural heritage, particularly in current contexts where cultural and linguistic diversity, artefacts and histories are at risk as a result of political and social shifts and upheaval.

The BPF had released a call for contributions to help gather examples and best practices of how digital technologies and the Internet can be used to promote, preserve and share local culture and content. It wanted to identify best practices of how to manage and promote the digitisation of existing analogue content (printed and electronic media, cinema, music, visual arts etc.) and services.

The Internet Governance Forum (IGF) serves to bring people together from various stakeholder groups as equals, in discussions on public policy issues relating to the Internet. While there is no negotiated outcome, the IGF informs and inspires those with policy-making power in both the public and private sectors. At their annual meeting delegates discuss, exchange information and share good practices with each other. The IGF facilitates a common understanding of how to maximize Internet opportunities and address risks and challenges that arise.

The fourteenth meeting of the United Nations Internet Governance Forum (IGF) was held in Berlin, Germany, from November 26-29, 2019. The fourteenth IGF was hosted by the Government of Germany under the overarching theme: "One World. One Net. One Vision". It aimed at bringing together more than 2,000 participants from different stakeholder groups and some of the most prominent individuals in the global Internet governance discussion.

New Delhi-based Digital Empowerment Foundation submitted a response to the BPF on the issues and challenges faced by India on the preservation of local language, history and/or cultural heritage.

Introduction

There are close to 7,100 living languages in the world but as many as 90% of them are spoken by less than 100,000 people in the world, according to Ethnologue, an annual publication on world languages. In India, where there are more than 780 languages, over 220 languages have died in the past 50 years and 197 others are categorized as endangered by UNESCO. Despite India's cultural, geographical and linguistic diversity, only 22 languages have the official status in India. What is even more disappointing is that India has lost nearly 250 languages in the last half century, and 196 more have been declared endangered by UNESCO. As many as 120 of these 196 languages are spoken in the North-Eastern parts of the country alone. With most of these languages spoken by tribes and lacking a script, it has been particularly difficult to preserve them.

These languages are at risk. More than 1,500 languages vanished from the 1971 Census (1961 census had listed 1,652 mother tongues and the next census listed only 108 mother tongues).

The Internet remains predominantly an English language platform with over 800 million users English language overs on the Internet and over half the world's websites in English. However, languages like Spanish, Arabic and, especially, Chinese are fast catching up. Chinese, particularly, grew by 1277.4% between 2000 and 2010.

Out of a roughly 7,100 languages in use today, the following languages (in that order) make up 82% of the total of the content on the Internet — English, Chinese, Spanish, Arabic, Portuguese, Japanese, Russian, German, French and Malaysian.

Even as India's Internet users and literacy rate are growing, about 30% Indians are officially illiterate (the unofficial figure of illiterates would be even higher) and millions of others can't read or write in English, which makes up for more than 55.7% of the content available online. As of 2019, India had about 500 million Internet users but only about 175 million of them used the Internet in English.

In a predominantly oral culture country like India, which is home to thousands of tribal communities and hundreds of languages, how do you preserve languages? How do you document languages that have only been sung in folk songs or narrated in folk stories without a script? How do you document the culture, traditions and history which are embedded in everyday life?

In such a scenario, it would be beneficial to address the issue of underrepresentation of lack of regional languages (which are often the foundation to represent a culture, tradition and geography) on the Internet, and empowering the marginalised communities to leverage digital tools and the Internet to document their own language, culture, tradition and folklore.

The struggle to survive

India's diversity lies in its richness of languages. There is an Indian proverb that goes, "Every two miles the water changes in India, every four miles the speech".

Unfortunately, several of once thriving tribes in India have fewer than 1,000 members living today. Based on the yardstick set by our modern education system, these tribesmen and women are considered illiterate and backward. However, one has to visit their homes, interact with them and see their craft to truly understand the depth of their knowledge, art, traditions and sensibilities across wide sectors of medicine, architecture, agriculture, medicine, art and more.
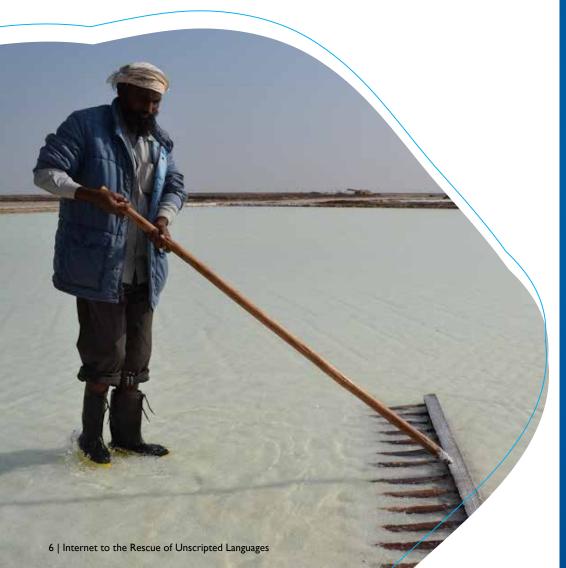
The Indian Toda tribe, for example, is largely dependent on buffalo herding and embroidery for its livelihood. Its members are incredibly skilled artisans known for the red-and-black embroidery on white fabrics that has even earned them a GI (geographical indication) tag. They live sustainable lives, in harmony with nature where all their resources are available. The Toda tribals have their own language, which does not have a script. Over the last century, their numbers have been dwindling. The sharp decline in their population is largely related to the decline in agriculture land, much of which has been lost of afforestation. With their dwindling numbers, their art, craft and traditions are facing a slow death. If not preserved, the day is not far when their unique embroidery, for instance, is lost forever.

The Todas are an extremely closed community, barely connected to the rest of the world and, thus, deprived of the opportunities connectivity offers. The Todas are not alone in this struggle. Several other tribes, which speak no other language than their mother tongue, and have either no script or aren't trained to understand their script, will never find way to the Internet, if service and content providers continue to focus on the English speaking populations of the world. Unfortunately, in India, even a large share of government websites is only available in English or Hindi. Only a handful are available in state languages. Other languages are completely missing, pushing the

marginalised communities further into marginalisation by disallowing them to access government records and information.

Around the world, the rural and villages' folk societies have depended on the oral traditions for centuries for information dissemination from generations to generations. Interestingly, more than 55 per cent of all Web content is in English even though only around 20 per cent of the world's population speaks English, and just five per cent of the world speaks English as their native language.

In such an English-dominating virtual world, where technology, too, is largely developed and designed by native English-speaking persons, the oral and illiterate communities struggle to survive.

Internet to the rescue

In the absence of digital tools and the Internet in the past, documentation through text was the only medium of preservation of history, culture, art and even languages. Not everyone was accomplished to write a book or influential enough to publish it. The Internet, too, for most part of its initial years represented only those who were on the top of the pyramid — both in terms of financial capability and English speaking skills. Then came the proliferation of digital tools, allowing history, culture, art and even (scripts of) languages to be photographed and preserved. We now have audio visual formats, which eliminate the need to have a script to document or preserve a language.

Linguists from National Geographic's Enduring Voices project have already produced eight talking dictionaries to document struggling languages. Besides containing 32,000 word entries in eight endangered languages, the dictionaries hold more than 24,000 audio recordings of native speakers—many of who are among the last fluent individuals in their native tongues—pronouncing words and sentences, and photographs of cultural objects. The first project under this initiative was to initiate the documentation of Koro, a Sino-Tibetan language spoken by less than a thousand people in Arunachal Pradesh, in 2010.

This wonderful initiative is not the only one; there are cross-language open source tools to orally document pronunciations. Several linguists around the world are working on other similar efforts. However, these efforts by linguists and researchers are not enough. There are far too many unwritten languages and far too few efforts to document all of them.

There is no denying that with the help of basic digital tools, people can be empowered to tell their own stories, beating long-set information exchange criterion of being able to read and write one or more script. By using the medium of spoken words and audio-visual story-telling, masses are better placed in the current information economy.

With an increasingly high penetration of mobile phones, we at Digital Empowerment Foundation have witnessed people learning to use digital tools to tell their stories. There are "illiterate" people who are using smartphones to become community journalists; who are sharing their folk music on YouTube; who are using WhatsApp to find buyers for their handloom and handicraft; who are using emojis and audio notes to communicate with their loved ones. It is interesting to see how people in rural, remote and tribal locations with no formal education and lack of knowledge of a script are engaging in fluent conversations in real time through audio notes, video calls and a bucket full of emojis to express an entire thought without any letter at all. And so, the written medium of communication is no more the parameter to define literacy.

Knowledge is for all. And that is what makes it imperative to include the excluded — those who are yet to get online are largely those who form the oral and illiterate society. In order to bring them into the digital world to both produce and consume content on the Internet in their regional languages—thus improving the scope and volume of traditions, cultures, art and language that populate the present day Internet—it is important to not just train them in digital literacy but in media and information literacy (MIL). MIL has the power to ensure that the new users of the Internet do not just become consumers of information; but can consciously and judiciously access, organise, analyse, evaluate and produce information.

It is interesting to note here that while none of the Indian languages make it to the top 10 languages on the Internet as of now, Indian language user base grew at a compound annual growth rate of 41% between 2011 and 2016 to reach 234 million users at the end of 2016, surpassing the English users in India, according to a recently released Google-KPMG report. With mobile penetration fast increasing in the country, Indian language Internet users are expected to account for nearly 75% of India's Internet user base by 2021. Further, the high penetration of mobile phones has given them the opportunity to connect with their loved ones because speaking over the phone does not require one to know a script unlike writing letters.

Digital Empowerment Foundation was started with the simple idea to take the Internet to the people in rural India and help them access the world of information and opportunities. And we've seen enough examples of the democratisation of information—and subsequently even power in many cases—around the world due to the proliferation of the Internet, the Web and even social media. Through our work, too, we've seen overwhelming instances of people accessing the Internet, exploring the Web and transforming their lives. I remember this young girl who had come to one of our digital literacy centres in Uttar Pradesh for her first day of class. She was scared to touch the laptop even on the insistence of the trainer. "What if I get an electric shock?!" she had said. When I visited the village again a few years later and met her, she was all smiles and confidence. She told me that she never even got a chance to complete her high school but she could access the world of information on Google.
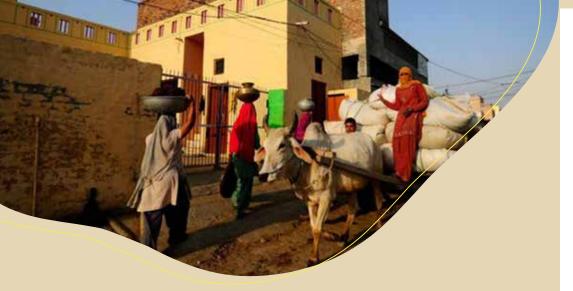
In Bihar, about 400 women camped at one of our centres from sunrise to sunset for three straight days with their passbooks in hand because a man at the centre had decided to help them match their MGNREGA attendance with the daily wage received in their bank accounts. Any discrepancy that was noticed, was filed as a complaint and sent to the block office.

A 62-year-old widow in Jharkhand did not exist according to the government records. A digitally literate woman, equipped with a tablet phone, in her village helped her arrange necessary documents, file an application and submit the papers. For the last two years, she's been receiving her old age pension as well as widow pension.

In Madhya Pradesh, weavers of beautiful silk sarees—that were sold for over Rs. 6000 in the market—barely even got Rs. 300 for putting in days of labour to make the product. It was until they were able to access the Web and see the market prices that they could demand the right wages from their master weavers.

In Tamil Nadu, a young orphan girl wants to be an IPS officer like Kiran Bedi. She's read up so much about her online and almost memorised her biography. She knows general knowledge is important for a role like this, and so she spends 20 minutes every day taking daily news quizzes.

Over the years, millions of women have turned from homemakers to entrepreneurs. They skills and art of embroidery, pickle making, cake baking, handicraft and so much more are no longer confined to the appreciation and use of their family members,

but are available online for people around the world to buy. Thousands of women in urban and rural India have opened up their homes to app-based homestay services, introducing both national and international tourists to their lifestyle, their culture, their traditions and their languages.

Stories like these are many, and stories like these also showcase the link between technology and women empowerment. While some women are overcoming social barriers and fighting patriarchy with the help of the Web, others are finding economic opportunities and a platform for their views. And I truly believe that it is the information-packed mobile phone that will be the end of patriarchy one day.

Digital Empowerment Foundation visited Baran, a tribal locality in the Indian state of Rajasthan in 2007. Here lived the Sahariya Tribe with no telecom or Internet connectivity. Disconnected from the rest of the world but not in discontent with their self-sustainable lifestyle.
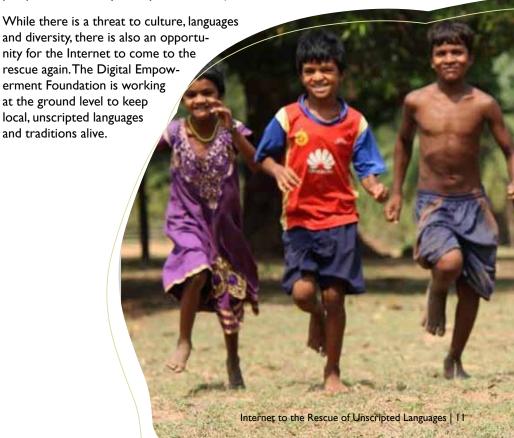
The Sahariya Tribe was financially not "rich" but they were extremely "rich" in terms of culture, traditions, art and local language. The tribe is also spread across in different clusters across three Indian states, but barely connected to each other.

Here, the Digital Empowerment Foundation established a 200 km community network and connected it to a local server. This way, even if the Internet is down, the community can share content and access content through the local server, thus creating an intranet or community network. This has also encouraged the community over the years to create a localized database and archive its oral and traditional knowledge, art as well as culture. This connectivity did not just connect the Sahariyas to each other across a 200-km geographic span, but also connected them to the rest of the world to both consume contextually relevant information and to populate the Internet with traditional knowledge and culture, too, in audio visual formats.

Connectivity can ensure better access to government schemes, entitlements and rights; digital market linkages can enable tribal communities to exhibit their craft and agricultural produce to the world for an improved livelihood; access to the Internet can keep them updated on government notifications and other relevant information; digital documentation can preserve and showcase their richness for posterity.

A low literacy rate, especially in rural India, has for long kept people devoid of any information. But with the advent of the internet, there is a possibility for the gap to be bridged.

Gondi, a South-Central Dravidian language in India, is spoken by close to three million people of Gond ethnicity. Although it is the language of Gonds, only a one-fifth can speak the language, putting on the verge of extinction the rich folk literature that survived through marriage songs and narrations till now. There is also certain unwillingness on the part of the younger generation in learning the native language as they mostly migrate to cities looking for better opportunities leaving the age-old livelihood practices. However, pages on social media sites like "Humans of Gondwana" (Facebook) and "Gondwana Events" (YouTube) are trying to give a new lease of life to the culture. Run by locals, these pages are telling the world stories from a communities' perspective, their day-to-day lives and major Gondi events.

While there is a threat to culture, languages and diversity, there is also an opportunity for the Internet to come to the rescue again. The Digital Empowerment Foundation is working at the ground level to keep local, unscripted languages and traditions alive.

Digital Empowerment Foundation
House No. 44, 2nd Floor
Kalu Sarai, (Near IIT Flyover), New Delhi – 110016
Tel: 91-11-42233100 | www.defindia.org