India Program/Indic Languages/Summary of initial discussions - 2011

2011	Assamese	Bengali	Gujarati	Hindi	Kannada	Malayalam	Marathi	Nepali	Odia	Sanskrit	Tamil	Telugu	Summary of discussions

This is the mail that I sent to India list [1] and Indic language community mailing lists after I concluded my initial introductory discussions with different Indic wiki communities in November - December 2011.

Apologies beforehand for a rather long and winding mail - but there is so much that I want to say! I want to share how my thoughts are being crystallized. I want to try and cross-pollinate ideas from some Indic language communities across to all communities. I want to reach out and ask your views and suggestions. I want to understand how best we can help each community in a manner that is most appropriate to that community.

I have now completed sharing initial, introductory, exploratory discussions with a host of community members from across Indic language communities. I have shared these for 12 languages (Assamese [2], Hindi [3], Tamil [4], Telugu [5], Kannada [6], Nepali [7], Malayalam [8], Marathi [9], Odia [10], Sanskrit [11], Bengali [12], and Gujarati [13].) I haven't (yet) got any response from 7 other communities (Bhojpuri, Kashmiri, Punjabi, Urdu, Bhisnupriya Manipuri, Pali, and Sindhi).

> **Update**: One response from the Urdu Community can be found here [14]. Hindustanilanguage 06:15, 7 January 2012 (UTC).

At the very outset, I want to thank all of you who took time out and shared your experiences and thinking. It has been really useful and I hope you found it is as productive and constructive as I did. The purpose behind this exercise was to hear, learn, and understand the evolution of the various communities - and to therefore suggest ideas going forward. I urge everyone to go through all the other languages (even if they are not personally involved in those specific communities) because there are learnings for everyone from everywhere.

I have been reflecting on the various insights and inputs and ideas I have got from all these folks as well as subsequent discussions on mailing lists and talk pages. Here are my initial thoughts.

Community

It sounds like a self-evident and very basic thing but the single biggest priority for all communities (even for relatively bigger communities like Tamil and Malayalam) is community building. What has struck me from the various language communities is that everyone agrees that this is very much required but very few are aware of what needs to be done or how it needs to be done. I wanted to share some thoughts about this.

When I consider community building, I think of 5 broad aspects:

1. Editor retention
2. Attracting newbies
3. Community communication
4. Community collaboration
5. Community celebration

I would like to detail what I mean by each of these.

1. Editor retention: Like most language wiki communities we also have an editor retention issue in all Indic language communities. This is particularly an area of concern for us considering the fact that all our Indic language communities are really tiny and community buiding efforts in Indic wikis are very less. A dramatic case in point is Kannada where active editor numbers (that is, editors who do at least 5 edits a month) have declined from 25 members to just 9 members over the past 10 months. It is essential that all of us reflect on why this is happening and

what can be done to avoid it in future and to resurrect lapsed editors. Existing editors and old editors understand our projects and community and can play a huge role in community building and project quality improvement. Many times, they have become inactive because of changing personal priorities. However, sometimes, they leave because they are no longer excited by the projects. The lack of interest in a project or users not feeling proud about a project might be due to multiple reasons. Some of the reasons that old community members shared with me are poor quality of articles (driven by BOTs and Google translation project), dominance of wiki by one or two members, the huge amount of clean up and other administrative tasks required, and so on. We must reach out and welcome these editors back and we must encourage them to do what they love doing most - editing articles and making them regain their pride and ownership over their articles and projects. We must foster an environment that welcomes old editors back and gives them the space to follow their passions.

2. Attracting newbies: Attracting newbies is the only way our communities and projects can grow. I have to be honest and say that none of our language communities have achieved critical mass. According to me unless a project has 500 or more active editors, it can never be said to be in a state where organic growth is secured and momentum is ensured. Attracting newbies requires impactful outreach. By impactful, I mean outreach that is done frequently and to as a large a group of potential newbies as possible. However, it also means that we need to be much more systematic about how we do outreach. This covers everything from identification of the most appropriate target audience as well as doing outreach in a manner where we don't scare off newbies by information overload. We must make sure that our outreach sessions adequately convey the passion and love for our projects that we feel while working on them. Also, we need to critically look at how we reach out to attendees of outreach sessions (after the sessions) as well as other newcomers and see that we are providing an adequate helping hand to them. The Nepali community - though tiny - does very well in terms of posting personal talk messages to welcome new folks, having FAQs spaces and problem boxes, etc. - all with the objective of supporting newbies. All Indic languages are at a state where every single newbie should be identified and reached out to and given intensive help and warmly welcomed to the community. We must also look at both newbies to editing as well as existing **English Wikipedia editors** who have inclinations and abilities on Indic languages. Remember that many Indic editors initially started off in English Wikipedia and we must actively seek them out. I know some communities - like Marathi - who look for editors who have Marathi sounding names or edit Marathi/Maharashtra centric topics and quietly invite them to contribute to Marathi Wikipedia. Another aspect, and I am sure is this a bit of a controversial statement, but can we get few existing Indic editors to reduce their emphasis on editing and divert their time on outreach. (I know Tamil, Odia, and Malayalam communities are already doing this. But this needs to be replicated in other languages also). It is really tough and not everyone might have the interest to do outreach but the best outreach can be done by existing community members. However, as we know, volunteer time is limited. This is a challenge because what we love doing most is editing - but the reality is that the greatest need of the hour, and the area where we can contribute maximum, is attracting and training and supporting newbies. We should also look at digital outreach - by which I mean look at the existing internet activities in Indic languages (blog, facebook, google plus, and so on) and see if we can get newbies from there. For instance, many Indic languages have very active blogging. Can we reach out to bloggers and ask them to contribute to our projects, or at least evangelise about our projects and invite their readers to read Indic projects and contribute to them? Can we similarly look at social media like facebook and twitter to promote our Indic projects?

3. Community Communication: Community communication is an area which varies by community. There is a direct co-relation between the health and growth of the community and the inclusiveness, intensity, and warmth of the communication amongst that community. Community communication takes place on mailing lists, village pumps, meetups, and so on. With the exception of Malayalam and Bengali mailing lists, and to a lesser extent, Tamil, Odia, Mumbai, and Pune mailing lists, most others are virtually non-functional. Having said that, many village pumps are active across language communities. It really doesn't make a difference whether the communication is on mailing lists or village pumps. However, it is of paramount importance that it happens somewhere. Anywhere! To that extent, I encourage everyone to be more active wherever they are more comfortable

- but ideally in public spaces like village pumps or mailing lists. Reach out and ask for help or suggestions. Offer advice or inputs. Simply be friendly and accessible. Just talk! Community meetups are happening but not as frequent as one would like and with very limited attendance. Often, it is just 3 or 4 people who meet up everytime. Nothing wrong with that per se. Meetups are voluntary and the majority of wikipedians are happy to edit in the privacy of our homes and not meet up with others but even in this situation, we can and should be encouraging more people to attend meetups. People will attend meetups more regularly if they find them productive and inspiring. Too often, the feedback from community members has been that they don't find meetups useful or they find them dominated by 1 or 2 individuals. It is essential to have 1 or 2 individuals with the drive and hard work to organise meetups - but it is equally important that meetups are not centred exclusively around these 1-2 people but more about what the larger group want. How about meetups where all we do is spend an hour or two just editing a few articles? How about meetups where we plan a newbie outreach program involving everyone in the meetup? How about a meetup where that meetup is run by those folks who usually never speak up and that the entire meeting is devoted to what they are interested in? It is alarming when one looks at the situation in some Indic communities where there is virtually no communication at all amongst community members. It leads to a very cold and impersonal environment - which is not healthy to foster growth. Like plants and flowers, communities too need breeze and air and water and food and activity and earthworms and manure.

4. Community Collaboration: When I consider community collaboration, I think of 2 things. The first is ownership and the second is editing. On ownership, it is really critical that every one of us as individual community members believe and are made to believe that we own our projects. Every project is owned by all members of that community. Equally. We should all become more proactive in enforcing this ownership - whether it is in terms of coming up with initiatives or proactively participating in community discussions - whether it is about technical matters or content elements or community aspects. Every single individual counts and every single individual's voice must be encouraged. On editing, something that drives all of us is the thrill of collaborative editing. Wikipedians love it more than anything else to work together on an article and make dramatic improvements to it. Of course it happens even now, but this is something that we need to encourage much more and participate more actively in. This can be done in varied ways - but ideas like Collaboration of the Month or Editathons or whatever other idea should be organised. One can start with a handful of people working on a few articles - but one must try as hard as one can to make larger scale mini-events around this basic idea. It will help build personal relationships, project ownership, and drive community bonds.

5. Community Celebration: Lastly on the community aspect, let us bring some magic back to the community. Let us start celebrating successes - no matter how small. Let us start taking goals - no matter how seemingly unambitious. Let us spread cheer all around when we meet these objectives. Let us start publicly celebrating over the profiles of new or active editors (Tamil wiki community is already doing this)- whether because they are 12 years old or 80 years old or whether their article counts are 100 or 10,000! Let us celebrate when our wiki cross a major milestones, Let us celebrate when one our community member does some marvellous things for wiki. Let us celebrate when community able to engage in a relationship with state government... There are many reasons to celebrate. Let us celebrate all those and build the sense of pride about their projects among our community members. The most powerful fuel in our engines is passion - and we need to get more of it in our veins.

Projects

There is a constant debate of what should come first - article count or article quality. I don't think there is an answer to this that is equally applicable across all projects and communities. I had strong convictions on this based on my past experience with Malayalam wiki projects - which have been reinforced after my initial discussions with Indic Wikimedians from across the country. In this regard, I wish to share a provocative statement about bots. Bots can and should be used to do repetitive tasks (like adding categories) because that reduces wasting volunteer time - which is limited and precious. However, the use of bots for article creation is something that I would strongly

discourage. The current state of Newari wikipedia (which has nearly 70,000 articles but zero active editors) reinforces my argument.

The argument for using bots for article creation is that it provides placeholders for editors to start working on these articles. While there is some merit in this argument, the problem is that this kind of artificial intervention means that the volume of work required to improve quality far outpaces the community strength. It is like a sportsman using steroids. It is not natural or healthy. It results in large numbers of very poor quality articles - which are of such a basic nature that it might be better not to have them in the project. (For example, if the only information about a town is that "Abc town is in Abc district which is in Abc state and the population is 12345 according to the 2001 census", this article is so weak that it cannot honestly be said to exist.) If a project has thousands of these kind of articles, the whole project will be regarded as being of poor quality and will put off readers.) More fatally, if a project has thousands of such bot entries, it doesn't inspire editors to contribute - but instead makes them disillusioned because they feel that so many articles of such bad quality that they just give up on where to start! There are many who feel that, for example, Hindi wikipedia has been adversely impacted by the overusage of bots.

Another very important aspect I want to address is the kind of policies we adopt for Indic projects. Too often, tiny projects and communities are adopting too many of the policies of English Wikipedia. The policies of English wikipedia have evolved over years as English Wikipedia grew in community and article size. These policies are suitable for English Wikipedia given the size and breadth of its community. My view is that many of these are not appropriate for the current state of most Indic projects and communities given that the community sizes are 60,000 for English and ~25 for the average Indic community. If English Wikipedia policies are indiscriminately adopted, results in the feedback that I am seeing from many Indic editors that they are spending too much time doing "administrative" tasks like categorisation and not getting enough time for basic core editing. Let me elaborate. Something like NPOV is central to our overall philosophy. This cannot and must not be diluted. However, even if I take the larger Indic Wikipedias, it really is not such a major issue if the categorisation is currently weak. The focus has to be to build articles quality and content, and not necessarily having all the content neatly slotted into categories. Of course, something like categorisation is good, but not at the cost of article quality. I want to make an even more provocative suggestion. Verifiability is really really really important to all our projects. However, if one looks at how English Wikipedia evolved in the early days, it started with editors just adding content. Over a period of time, other editors came in and added and improved citations. Even today, as a recent Signpost article mentioned, there are 2.5 lakh articles in English Wikipedia [15] that don't have references. We should encourage editors to write, write and write! References will follow. Let us not chase away editors because we want every article to be perfect in a 20,000 article project. Of course we want quality but let us take it in stages - and let us prioritise what is most important to begin with. I think many editors would find it incredibly satisfying and inspiring and motivating to start and edit new articles, and they might get it 80% right. This will attract a much bigger community within which there will emerge a new generation of editors who love to add detail and citations.

Readership

One of my big discoveries I had was to see the total size of readership. I have often contemplated the Catch 22 situation of Indic language Wikimedians - where there is no awareness of the projects so there is no readership and even where there is readership, readers are not satisfied because of a low number articles or poor quality of articles. Conversely, editors don't find adequate motivation and satisfaction because they believe there are too few readers for their contributions. I often wondered how we would approach this problem - and which we should address first. I used to think that we should first focus on community building and article quality - and that readers will automatically follow. To that extent, I used to think that we shouldn't worry about readers because they will inevitably follow content. The fact that last month, we had more than 4 crore readers for our Indic language wikipedias means that the dilemma of what we need to do is no longer valid. We have readers. Lakhs and lakhs and lakhs of them for each Indic language wiki! We now need to focus singlemindedly on community building and

project quality. As internet penetration and mobile data access increase, we will get even more Indic readers. We don't need to do anything to attract them in the near future. However, we need to do *everything* to keep them coming back by increasing article count while religiously maintaining and increasing article quality and size of community.

Moving forward

I would love to hear your thoughts and views on the above suggestions.

The next stage of my work is going to be to speak directly with various communities in village pumps itself. I will try and make these as relevant and specific to individual communities - and also to share some ideas which have relevance across similar communities. For instance, some ideas will be similar to all communities with less than 25 active editors. I also want to try and identify potential areas of support that India Programs could work closely with communities on. The idea is to support community across languages. We would like to identify a very limited (1 or 2) pilots of a very controlled nature (in terms of scale) that we would like to collaboratively design with respective communities. Given the efforts that will be required in any pilot (even if it is of a relatively small scale), we believe that there needs to be a certain basic level of community size and collaboration to be able handle such pilots.

I will be sharing this mail on the various local language / local town mailing lists as well as the respective language village pumps. I look forward to hearing your views.

Shiju Alex

India Programs Team

References

[1] http://lists.wikimedia.org/pipermail/wikimediaindia-l/2011-December/005813.html

[2] http://meta.wikimedia.org/wiki/Wikimedia_Foundation_-_India_Programs/Indic_Languages/Assamese/Discussions/2011

[3] http://meta.wikimedia.org/wiki/Wikimedia_Foundation_-_India_Programs/Indic_Languages/Hindi/Discussions/2011

[4] http://meta.wikimedia.org/wiki/Wikimedia_Foundation_-_India_Programs/Indic_Languages/Tamil/Discussions/2011

[5] http://meta.wikimedia.org/wiki/Wikimedia_Foundation_-_India_Programs/Indic_Languages/Telugu/Discussions/2011

[6] http://meta.wikimedia.org/wiki/Wikimedia_Foundation_-_India_Programs/Indic_Languages/Kannada/Discussions/2011

[7] http://meta.wikimedia.org/wiki/Wikimedia_Foundation_-_India_Programs/Indic_Languages/Nepali/Discussions/2011

[8] http://meta.wikimedia.org/wiki/Wikimedia_Foundation_-_India_Programs/Indic_Languages/Malayalam/Discussions/2011

[9] http://meta.wikimedia.org/wiki/Wikimedia_Foundation_-_India_Programs/Indic_Languages/Marathi/Discussions/2011

[10] http://meta.wikimedia.org/wiki/Wikimedia_Foundation_-_India_Programs/Indic_Languages/Odia/Discussions/2011

[11] http://meta.wikimedia.org/wiki/Wikimedia_Foundation_-_India_Programs/Indic_Languages/Sanskrit/Discussions/2011

[12] http://meta.wikimedia.org/wiki/Wikimedia_Foundation_-_India_Programs/Indic_Languages/Bengali/Discussions/2011

[13] http://meta.wikimedia.org/wiki/Wikimedia_Foundation_-_India_Programs/Indic_Languages/Gujarati/Discussions/2011

[14] http://meta.wikimedia.org/wiki/Wikimedia_Foundation_-_India_Programs/Indic_Languages/urdu/Discussions/2011

[15] http://en.wikipedia.org/wiki/Wikipedia:Wikipedia_Signpost/2011-10-31/Opinion_essay

Article Sources and Contributors

India Program/Indic Languages/Summary of initial discussions - 2011 *Source*: http://meta.wikimedia.org/w/index.php?oldid=3242256 *Contributors*: Hindustanilanguage, Naveenpf, Psubhashish, Shiju, Shijualex

License