

Despite the rapid progress in Natural Language Processing (NLP) research – propelled by the advent of neural networks and the subsequent pre-training wave – there are concerns on how **inclusive** or **impactful** this research is to the masses. As an **AI Center Fellow** at Microsoft Research India (MSRI), my primary research focus over the past two years has revolved around the notion of “Language Technologies for All”, in particular, on understanding the disparity which exists between research & its deployment, and on developing technologies to mitigate this gap using collaborative interfaces between humans and Artificial Intelligence (AI). I intend to continue working in this area of **Human AI Interaction**, through **Human AI Collaboration** and **Human-Centered AI** by leveraging recent advances in NLP and Human Computer Interaction (HCI).

### The Inclusion Problem in NLP

An important and ongoing problem in the field of NLP is how its major breakthroughs tend to disproportionately favor resource-rich languages like English. The resource-poor languages, although spoken by millions around the world, tend to lag behind significantly.

I came to realize the gravity of this problem when I got an opportunity to interact with the Gond community, first-hand, as part of an MSRI research visit to Raipur, a city in India. Gondi, the language spoken by this 3 million-strong community, is severely under-resourced. Even basic digital technologies like mobile keyboards, or translation systems to facilitate information exchange are absent, which is now forcing the youth of this community to shift from their mother tongue to more popular and “useful” languages. Naturally, this has sparked a fear about whether their language might wither, or even disappear entirely. This experience got me thinking about the myriad challenges facing similar communities that have been marginalized by the lack of adequate language resources. I led a survey along with my colleagues at MSRI on understanding the many challenges faced while building and deploying language technologies for such resource-poor communities. We characterized the basic necessities that need to be fulfilled to ensure that such languages prevail and continue to thrive. This study was accepted at **ICON 2019** [1]. Further, I was inspired by the efforts of MSRI in addressing some of these enlisted challenges through innovative technological interventions spanning several years, and through active engagement with the Gond community. We documented these efforts in our **LREC 2020** paper [2], in the interest of sharing some insights from our successes and failures with the broader research community. Through this experience, I learnt that the introduction of a new technology should always focus on the interests of its end-users, and thus, be ‘human-centered’ in its very essence.

Deployment of language technologies to marginalized communities is certainly left wanting, but is there adequate research being conducted or models being developed for resource-poor languages? To dig deeper, we characterized the disparity of language inclusion in NLP research at top-tier conferences, which are important drivers of research progress. To quantify how certain languages are under- or over-represented, we conducted a comprehensive analysis of how skewed the resource distribution was for different languages, and how the different NLP conferences scored in terms of language inclusion. In this work, which was accepted at **ACL 2020** [3], I was specifically involved in creating a data model for the conferences, and devising statistical experiments and metrics to highlight the extent of this problem. We found that while certain moderately resourced languages are reaping the benefits of massive multilingual models, a lot of languages are still left behind, and this requires renewed attention and focus. Working on this problem helped me gain valuable insights into performing large-scale contrastive analyses often involving complex data wrangling.

### Impact through Human AI Interaction

The disparities between high and low resourced languages are significant, and would require tremendous research and engineering efforts to mitigate. The performances of several NLP tools and models wear off rapidly after the top few languages. In the absence of sophisticated language resources, re-introducing humans into the process might prove to be the key to bridging this gap. For instance, interactive systems that enable humans to collaborate with subpar AI, can automate certain tasks and also allow for feedback systems to enhance the AI.

My very first work at MSRI was in this direction in collaboration with Pratham Books, a non-profit organization that publishes storybooks for children in around 280 languages. The task at hand was to look at ways to improve productivity of amateur community translators by providing assistance. Off-the-shelf machine translation (MT) were not likely to work in this context as (i) the quality quickly deteriorates beyond the top-n languages and (ii) translation of stories require a more creative and contextual translation, as opposed to a literal translation. To address these challenges, we came up with a human-AI styled solution that uses MT to aid translators with cues in the form of gists and dropdowns, enabling them to select an option if they liked, thus helping with their productivity as well as with quality control. While such studies have been attempted before, they have seldom been explored in resource-poor contexts where MT systems struggle to generate fluent or coherent translations, and where minimal human inputs can prove to be game-changing. I developed a constrained decoding mechanism on top of OpenNMT in order to take partial inputs from users and provide suggestions accordingly.

This work was accepted as a demo paper at **EMNLP 2019** [4] and was eventually open-sourced. Through this experience, I also learnt how to work with a heavily optimized code in sync with a complex and interactive front-end interface that I had developed. Currently, we are exploring the integration of this system with Bing Translator and Microsoft Office 365.

In addition to its intended use, this work sparked interest in the resource-poor language communities as a crowd-sourced data collection tool, especially in an active learning approach. For instance, Translators without Borders have started using this ecosystem for obtaining parallel sentences in English and Tigrinya. We are also carrying out similar efforts to obtain Hindi-Gondi sentences in an offline setup to counter limited internet connectivity. In this space of crowdsourcing interfaces, I had also briefly worked on developing an interface for collecting transcripts for code-mixed speech, which was a challenging task in cases with script differences. This interface, COSSAT [5], allowed the crowd workers to be able to select the suggestions guided by a suboptimal Automatic Speech Recognition system, thereby reducing the error rates in data collection significantly.

### Research Interests and Graduate Studies

Drawing on my research background, I am particularly interested in taking a human-centered approach towards how AI can be made **inclusive** and **impactful** to the broader community. During 1960s, in the race to build the best supercomputer, it was the focus on human computer interaction through building personal computers and hand-held devices that brought about a revolution in making computing pervasive. As we now find ourselves living through a rapid development of AI, there is a need for novel interactive paradigms and architectures that enable for synergy between humans and AI, beyond focusing on building increasingly sophisticated models. I am keen on exploring this through the lens of intelligent interfaces and human-in-the-loop systems, especially putting focus on aspects of building trust between human and AI.

At present, natural language ‘understanding’ in NLP models is highly contested[6]. We are increasingly aware that these models are often adept at solving tasks that deal with pattern recognition, and/or require substantial amounts of memory, or computing resources. However, they have not yet reached the levels of cognitive reasoning or pragmatic understanding to match that of humans. This is where both human and machine intelligence can complement each other, for the greater good (**Human-AI Collaboration**). NLP has an added dimension of varied resource-availability for different languages where interactive systems can be immensely helpful, by allowing for humans to collaborate with a subpar AI. Further, it also provides for a testing ground to understand optimal human AI collaboration strategies at different levels of resources, as well as in different socio-cultural contexts. While deploying NLP models in such settings, a common concern is that they are often done so without a thorough examination of its ethical implications. For example: Machine Translation can be used in fairly benign contexts like ordering food in a foreign country, as well as in high stake scenarios like conversing with local law enforcement. Measuring MT systems with just one metric (BLEU score) can lead to unintended consequences. Can these translations be explained? Or, can we understand how confident the MT is? (**transparency**). Who shares responsibility when the translations lead to adverse outcomes? (**accountability**). How do we guard against prejudicial and discriminatory translations? (**ethics**). For a start, human-centered metrics such as these can be made inherent to NLP leaderboards, which otherwise record just single-score model accuracies, and do not account for the criticality or stakes of the situation where it can be potentially deployed (**Human-centered AI**).

Through my research experience over the past few years, I am convinced that I want to pursue research as a Ph.D. student, and University of Washington (UW) offers me with a unique environment to do so. UW with its high-quality research faculty and productive research groups working heavily in NLP, AI and HCI, will provide me the platform to realize my interests in the continuum of these areas. I am particularly interested in working with **Prof. [REDACTED]** in his research efforts on combining human machine intelligence from different perspectives in sync with what I intend to pursue. I am also interested in working with **Prof. [REDACTED]** on intelligent interfaces from the socio-cultural aspects which aligns well with my work at MSRI on interactive interfaces. I am also excited about **Prof. [REDACTED]**'s works as her interests revolve around NLP for social good, resource-poor languages and ethical considerations in NLP, which has been a unifying thread for my work over the past two years. I am interested in working with **Prof. [REDACTED]** whose work on predictive translation memory laid foundations to my MSRI work on INMT. I am particularly inspired by the work of PhD students at UW working in similar areas, namely **[REDACTED]**. Overall, I find UW to be a great match for my interests with several esteemed faculty and bright students working in these areas. Pursuing a Ph.D. with them, will be a stepping stone towards my intent to pursue academic research as a career.

- [1] P. Joshi, C. Barnes, S. **Santy**, *et al.*, “Unsung Challenges of Building and Deploying Language Technologies for Low Resource Language Communities,” ICON, 2019.
- [2] S. **Santy\***, D. Mehta\*, R. Mothilal, *et al.*, “Learnings from Technological Interventions in a Low Resource Language: A Case-Study on Gondi,” LREC, 2020.
- [3] S. **Santy\***, P. Joshi\*, A. Budhiraja\*, K. Bali, and M. Choudhury, “The State and Fate of Linguistic Diversity and Inclusion in the NLP world,” ACL, 2020.
- [4] S. **Santy**, S. Dandapat, M. Choudhury, and K. Bali, “INMT: Interactive Neural Machine Translation Prediction,” EMNLP, 2019.
- [5] S. Shah, P. Joshi, S. **Santy**, and S. Sitaram, “CoSSAT: Code-Switched Speech Annotation Tool,” AnnoNLP, EMNLP, 2019.
- [6] E. M. Bender and A. Koller, “Climbing towards NLU: On meaning, form, and understanding in the age of data,” ACL, 2020.