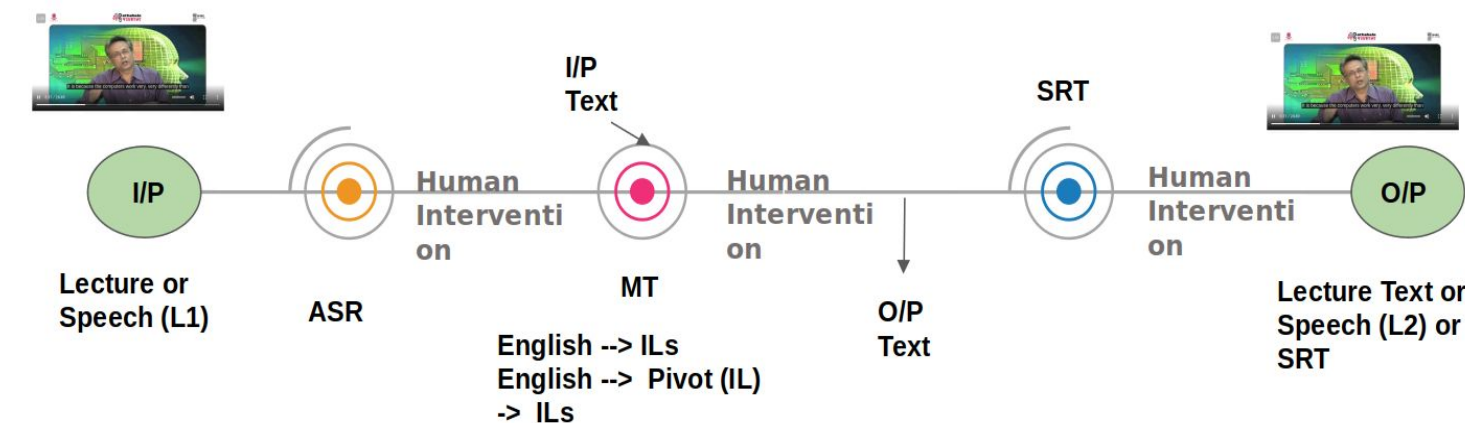# Speech to Speech Machine Translation (SSMT)

ABSTRACT

The Larger goal is to have capability which can transcribe, translate, subtitle and can convert translated subtitled text into audio for a given Audio/ Video.

Approach

I/P Text

SRT

I/P — Human Intervention — Human Intervention — Human Intervention — O/P

Lecture or Speech (L1)   ASR   MT   O/P Text   Lecture Text or Speech (L2) or SRT

English --> ILs
English --> Pivot (IL) -> ILs

Components

- **Automatic Speech Recognition (ASR)**
- **Pre-Processing/Editing on ASR**
- **Machine Translation**
  - **English to ILs**
  - **ILs to ILs**
- **Subtitling**
- **Text to Speech**

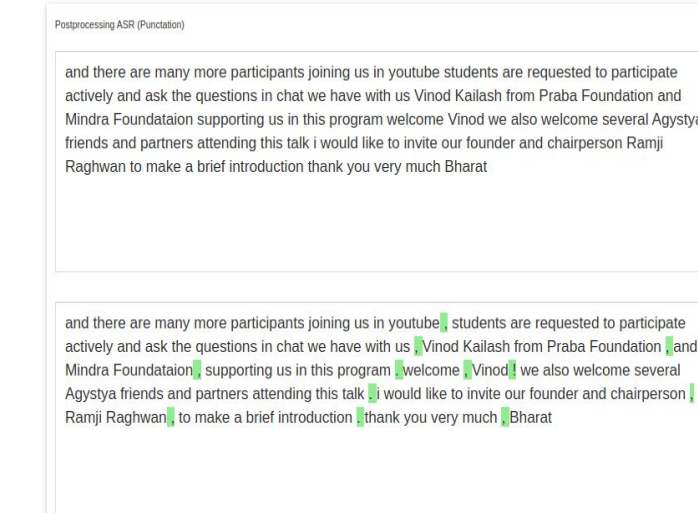## Challenges : Gaps (ASR to MT)

Punctuations , Disfluencies , Pet Phrases , Domain terms , Repair , Repeat, etc

okay , so , that is about working with incomplete information . the this thing , the ai techniques also are capable of uh working in a fault tolerant way . what do i mean by fault tolerant way ? suppose if i i am providing information which is less than uh sorry , i provide information which is partially incorrect . for example , i i have a friend who told me that he lives in a a society called ratlam society . i i just heard that thing over phone . i was looking for that uh um uh society . i i asked panwala , where this ratlam society is ? the panwala said there is no ratlam society nearby but there is a satnam society nearby . how could that panwala find out the real answer ? i have not provided him correct input but the pa... panwala could get that . how ? because his mind was capable to process the incorrect information and get the correct answer . are we able to do that ? are our programs... all of us are able to do that in fact , but can our program do it ? can my database , suppose if i have a database . i have indexed my database on customer name but if i provide an incorrect customer name , can my database return with right customer uh data ? think of it . if you can do that . if the program can do that . it is it is exhibiting ai technique .
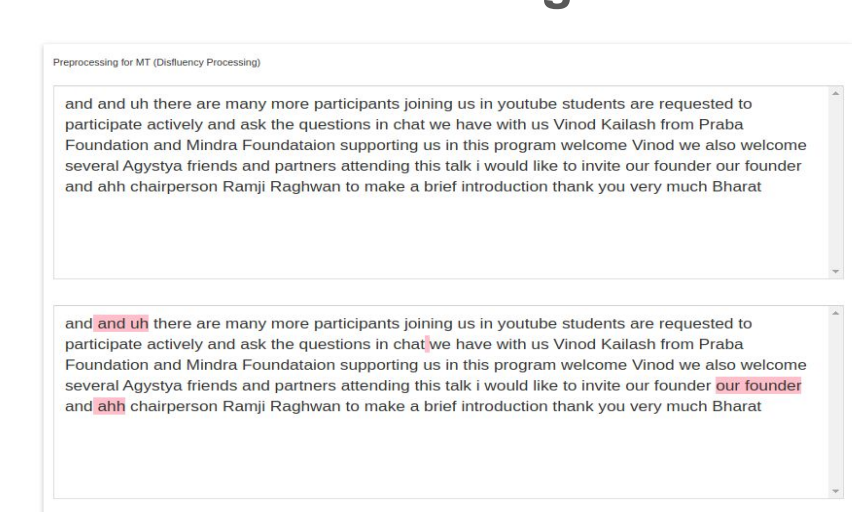
- **Disfluencies removal**
- **Punctuation marking**
- **Sentence/utterance segmentation**
- **Repair identification and removal**
- **Special notation identification and marking**
- **Domain Terms Identification**

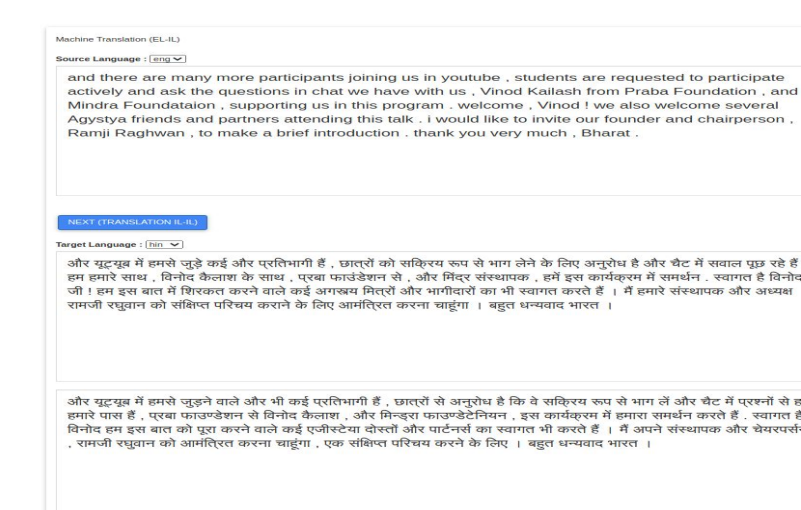## Process: Tools + Human In the Loop (http://ssmt.iiit.ac.in/ssmt_text)

Disfluencies removal

Punctuation Marking

EL to IL MT

IL to IL MT

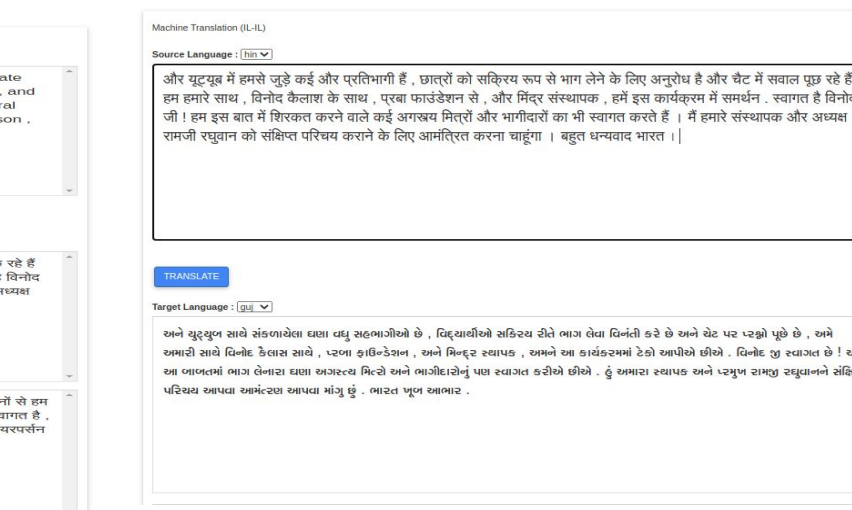## Future Work :

- **Capability for MT to handle noisy ASR output**
- **Capability for MT to Incorporate Spoken Prosody**