

# Virtual Assistant for Hiring Last-Mile Workforce

Piyush Makhija



[@pmakhija3](https://twitter.com/pmakhija3)



[/in/piyushmakhija/](https://www.linkedin.com/in/piyushmakhija/)



<https://piyushmakhija.com>

# Immediate Requirement

ET, Oct 31st

## Demand for delivery jobs much higher availability: Report

Festive boost to hiring

26

Swiggy, Zomato hike delivery boy salaries as  
competition grows

Wednesday, 07 November 2018 | Pioneer Avenues

# Demand for delivery jobs on rise: Report

Indeed in its report says there is a rise in hiring by e-commerce companies,  
particularly for temporary workers like delivery and warehousing personnel  
than openings: Online data

companies riding the wave in India

According to data from job site Indeed, jobseekers' interest in delivery jobs is twice as high as  
the number of available job postings for such roles.

ET, Oct 31st

## Delivery jobs demand twice number of festive season

E-tailing giants on hiring spree as festive season grips  
country

# Burning Need

- ❖ 500K delivery jobs in India \*
- ❖ 30% YoY increase \*
- ❖ Exploding on-demand & e-commerce industry
  - Heavy investment in scaling delivery & logistics operations
- ❖ Example
  - 2nd half, 2018: Dunzo increased its frontline workforce by 10x

# Last-Mile vs Standard Recruitment

	<i><b>Last-Mile</b></i>	<i><b>Standard</b></i>
Per Month Hire	500~10K	1~20
Attrition Rate	High (~75%)	Low (1~3%)
Hiring Process	Manual and Unorganized	Streamlined
Hiring Demand	Seasonal and Sporadic	Well planned and Long-Term

Jobs space is highly **fragmented**, **inefficient (manual)** and **hard to scale**, especially for **high-volume** last-mile delivery jobs

# Problem Statement

Build an *efficient & scalable* system for high volume recruitment of last-mile workforce

# Design Considerations

- ❖ Reduce manual effort by *automation*
- ❖ One stop shop
- ❖ Must work at *scale*
- ❖ Account for high *churn & attrition*
- ❖ There is *NO* “LinkedIn” for Blue-Collar Workers
- ❖ App?
  - Adoption challenge
  - Uninstalls



# Our Solution

**Automate job-seeker engagement on existing messaging apps**

Job seeker texts “hi” to  
***800-200-7000***



# Addressing Design Considerations


## ❖ Circumvent Adoption Challenge

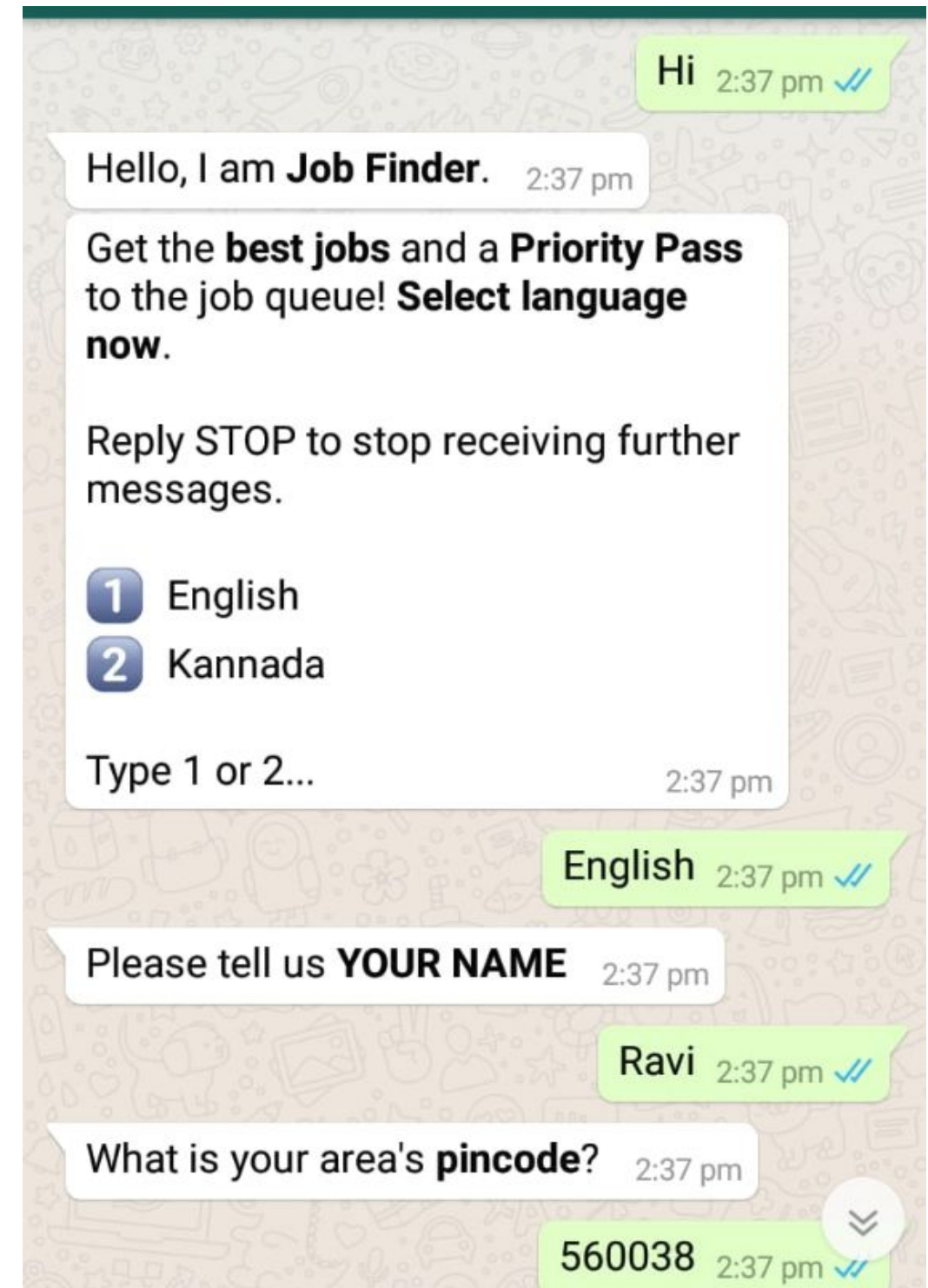
- India runs on whatsapp

## ❖ Automate

- Job pitch
- Gauge user's interest
- Basic screening
- Schedule walk-in

## ❖ Virtual assistant

- IVR 
- Voice based 
- Text based (Chatbot) 





# Addressing Design Considerations

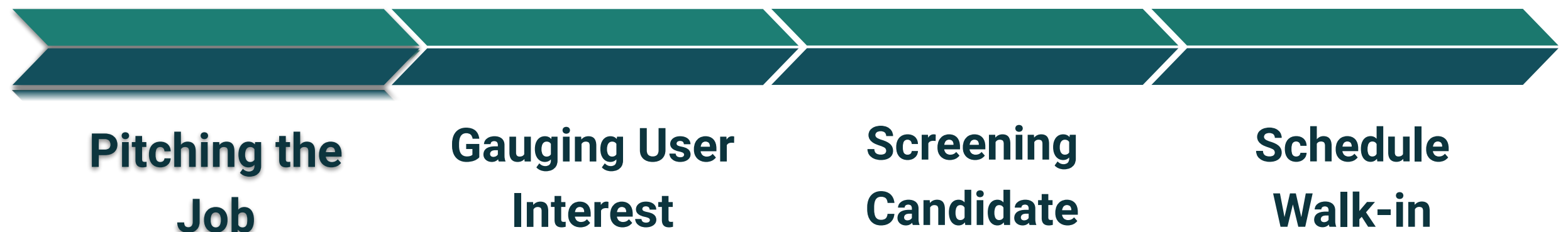
- ❖ Avoid hopping across different channels
  - Lead generation: Classifieds, Naukri, Quikr, etc
  - Processing: Excel Sheets
  - Engagement: Telecalling
  - Screening: On-site
- ❖ Scalable:
  - Engage 1000s of people simultaneously
- ❖ Counter churn
  - Increase top of the funnel→Adoption + Automation + Scalability
- ❖ Build rich user profiles
  - Laying the foundations of “LinkedIn” for Blue-Collar Workers

# Goal Oriented Dialog System

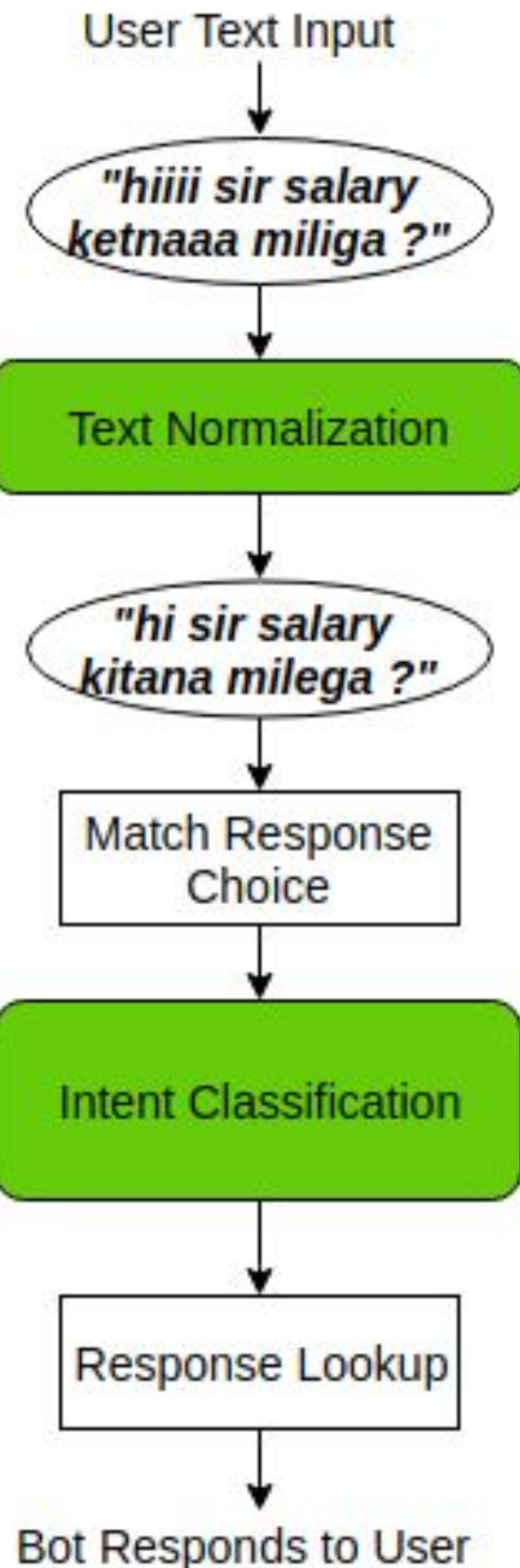
Carry out conversations to complete tasks to reach a goal.

❖ Goal: Qualify a user for a given job

❖ Tasks



# Our Solution: Bird's Eye View



**Method:** Neural Machine Translation  
**Performance:** >50% reduction in Word Error Rate (WER)

**Method:** Regex Match for Option Choices  
**Statistics:** Handles ~20% of the input user utterances  
**Current Case:** Input doesn't match any of the expected responses

**Method:** Rasa NLU Classifier  
**Performance:** Precision = 0.90, Recall = 0.82, F1 score = 0.85, Coverage = 0.91  
**Current Case:** "Salary Information" request Identified with 93.5% confidence

**Method:** Index Lookup  
**Current Case:** "Salary is upto Rs. 25,000 per month"

# "Your Model is Only as Good as Your Data"

## Comprehensive Dataset Issues

### Mixed Languages

"My Hinglish is different from your Hinglish"

Regional Influences

" ha "

### Poor Data Availability

New / Dynamic Phenomenon

Insubstantial Research

" howdu "

### High Variability

No Standardization

Cultural Influences

" s "

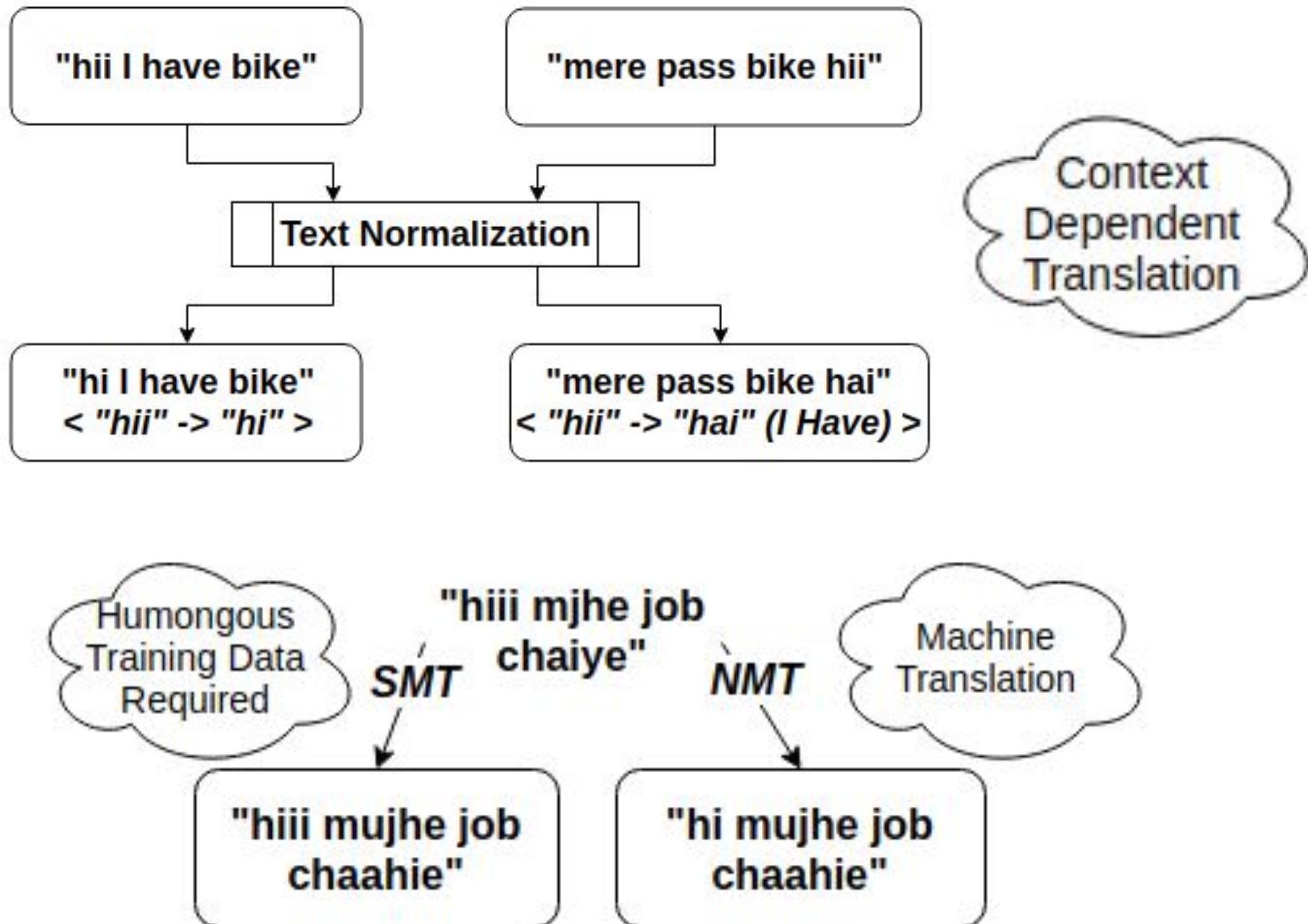
# Data Collection Exercise

- ❖ Jokes Bot ✗
- ❖ Daily Quotes Bot ✗
  - Too Much Spam
- ❖ Translation Bot ✗
  - Sporadic Usage
- ❖ English Learning Bot ✗
  - Narrow Content
- ❖ FriendFinder Bot ✓
  - Inspiration: Chatroulette
  - 2M messages per month

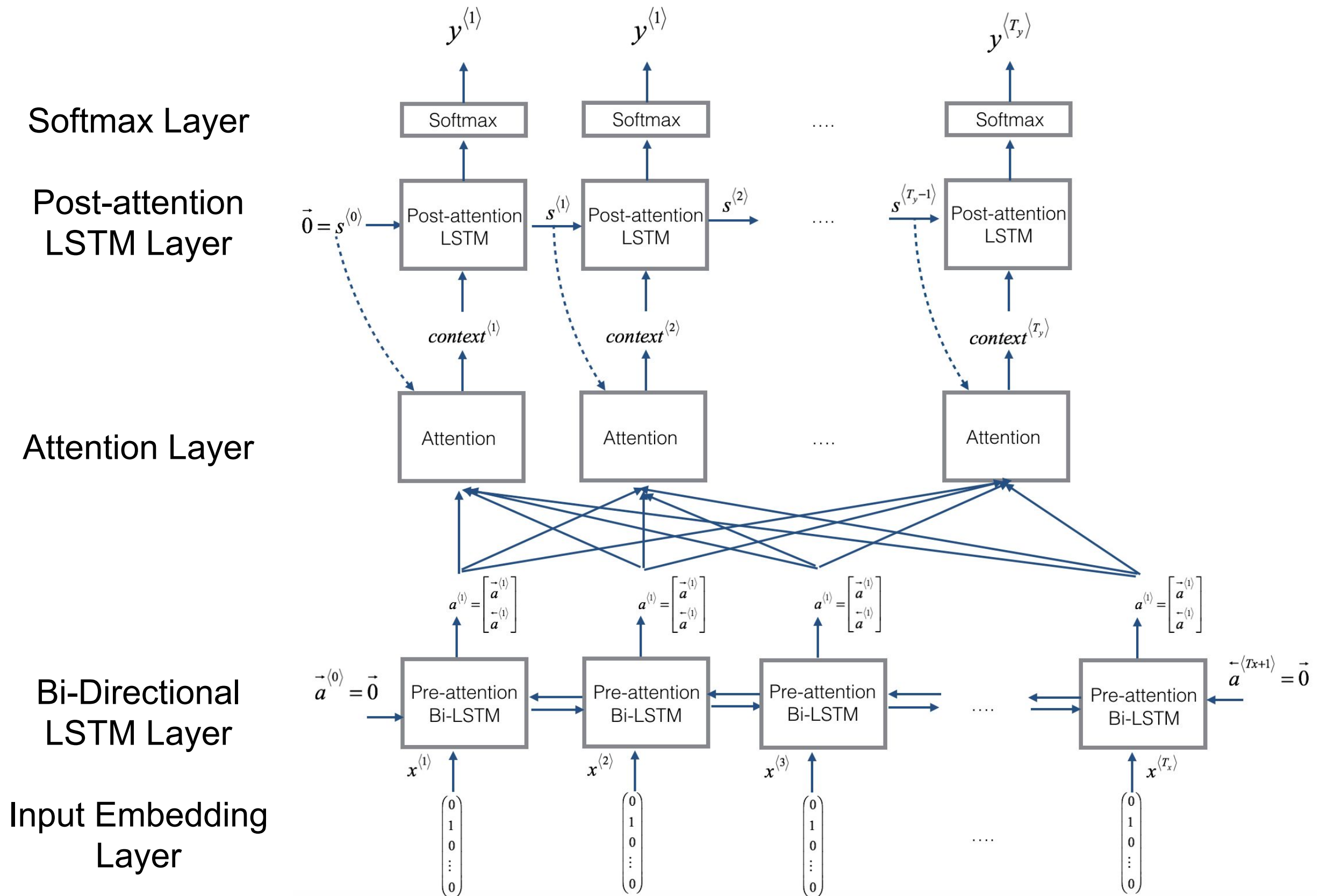




# Handling Nuances of Code-Mixed Data



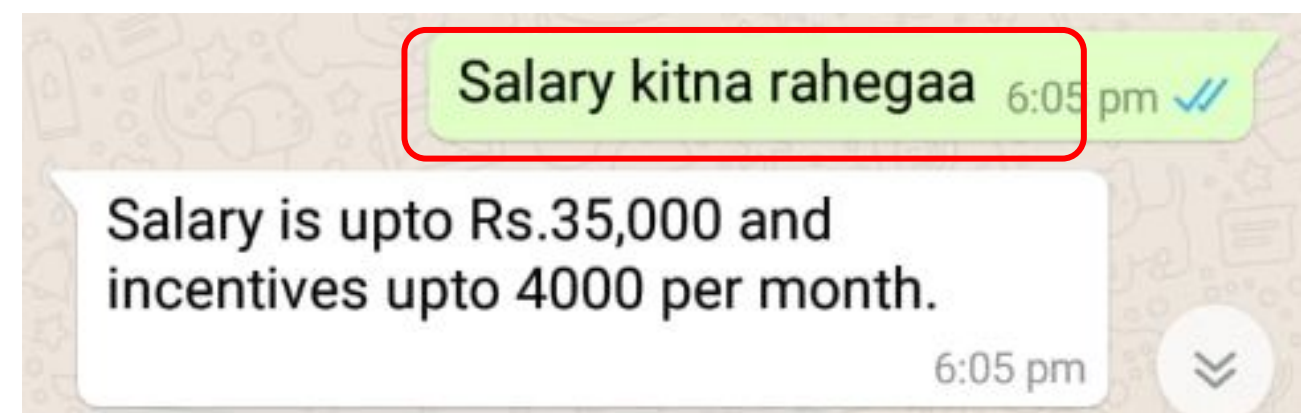
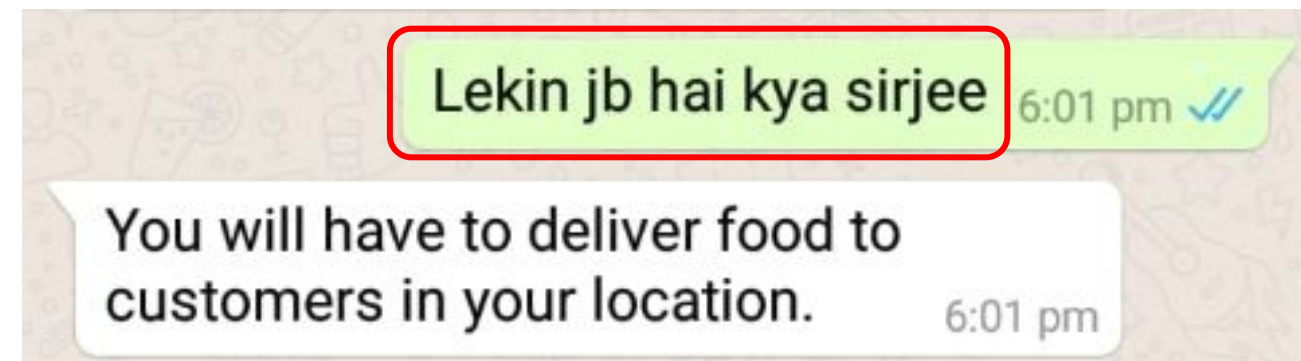
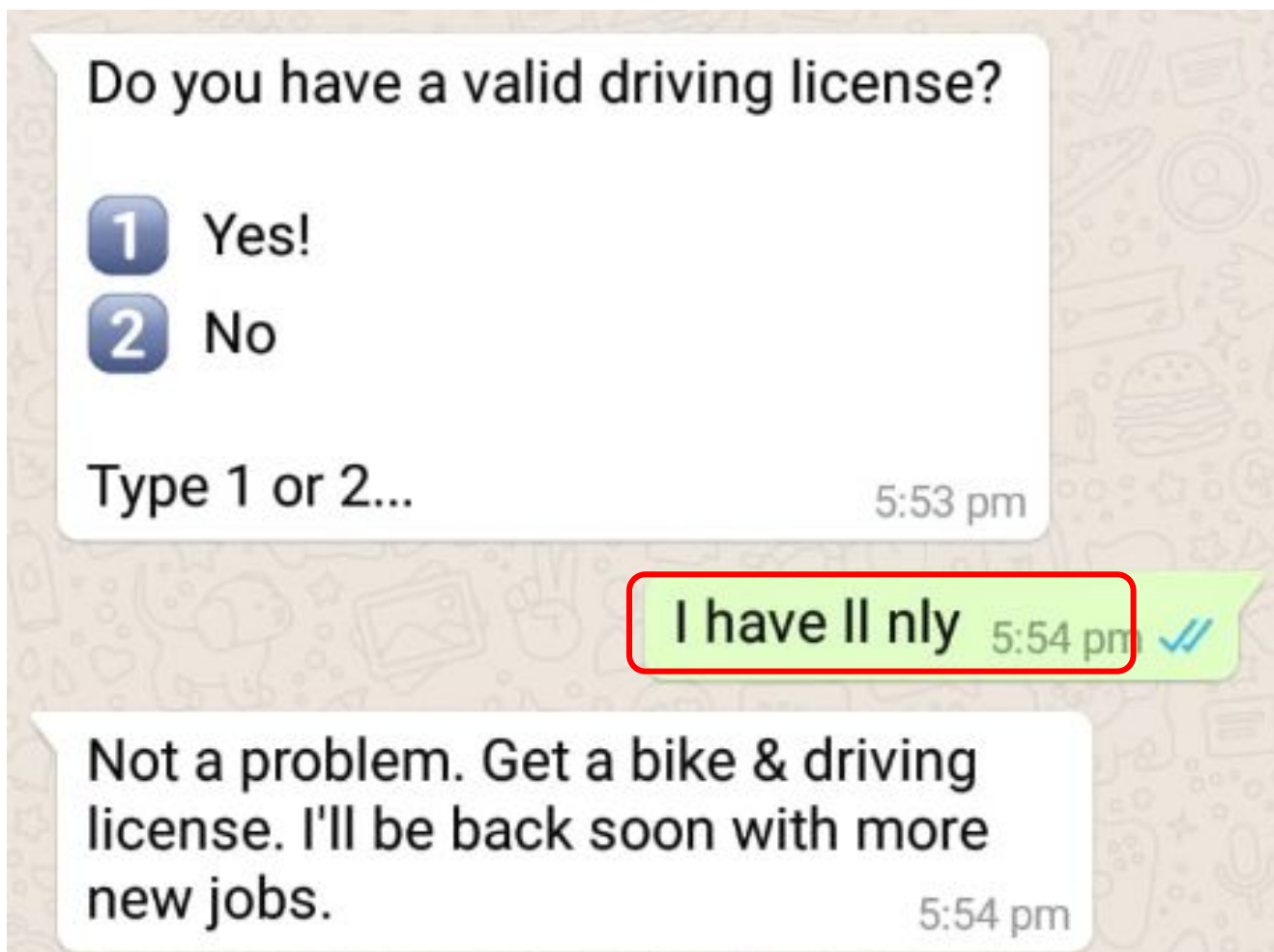
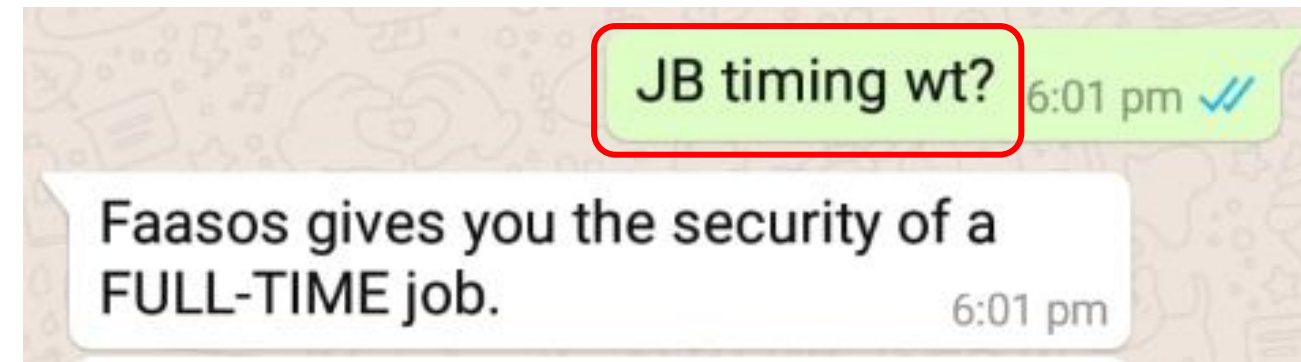
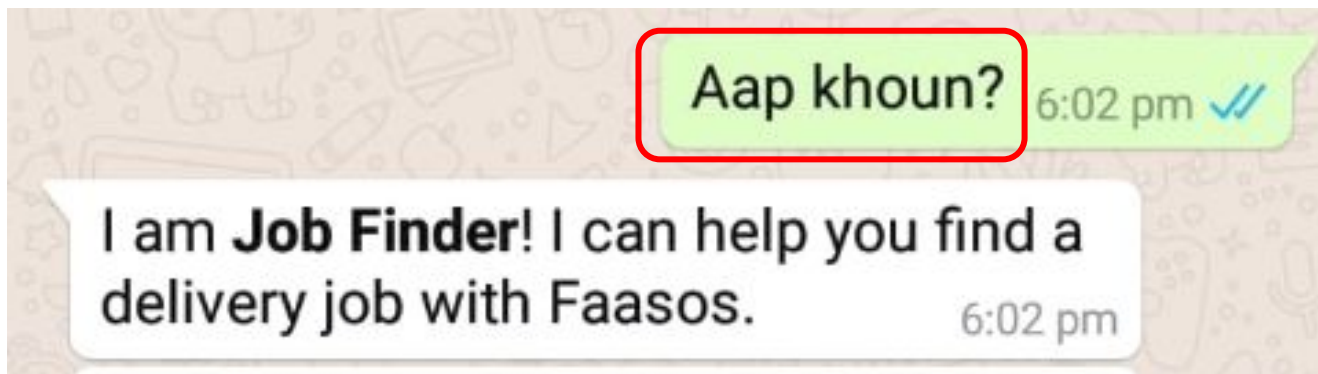
# NMT Architecture



# Text Normalization: Performance

Normalization Metrics	Baseline	Statistical MT	Neural MT
Word Error Rate (WER)	52.23 %	7.93 %	2.72 %
Average Sentence Bleu	0.62	0.93	0.97

# Intent Classifier





# Intent Classifier - Edge Cases

Meri bus book kra do 8:35 pm ✓✓

Please type 1, 2 or 3 to choose your answer to the question 8:36 pm

I don't know where to find a pikachu 8:42 pm ✓✓

You can work in your own location. Your interview will be at the closest Faasos Office. 8:42 pm

Do you own a bike and have a valid driving license?

- 1 Yes!
- 2 No

Type 1 or 2...

8:39 pm

Yep 8:39 pm ✓✓

I don't understand. 😊 8:40 pm



# Intent Classification: Performance

Classification Metrics	Rasa NLU	FastText
F1-Score	0.85	0.94
Precision	0.9	0.95
Recall	0.82	0.92
Coverage	0.91	0.92

# Ongoing Efforts

- ❖ Text Normalization
  - > 95% WER Reduction on Domain Data
  - Exploring BERT to improve performance
- ❖ Intent Classification
  - > 95% coverage, > 97% f1-score
  - Hierarchical Classification model
- ❖ Dialog System
  - Context-based humanly responses
  - Text Generation Models

# Key Takeaways

- ❖ Urgent need for a high volume recruitment solution
- ❖ Urgent need for a “LinkedIn” for Blue Collar Workers
- ❖ Ease of Adoption & Daily Usability
- ❖ Need for R&D on Code-Mixed Datasets
- ❖ Strategies to collect large amount of Code-Mixed Data
- ❖ Personalization for regional Persona

# Key References

- ❖ Philipp Koehn et. al., 2007, Moses: Open Source Toolkit for Statistical Machine Translation
- ❖ Minh-Thang Luong et. al. , 2015, Effective Approaches to Attention-based Neural Machine Translation
- ❖ Rasa NLU for Intent classification
- ❖ FastText library for text representation and classification
- ❖ Google's Rules for Machine Learning

# Vision - “LinkedIn” for blue-collar workers

Capturing profile data on blue-collar workers (80% of global workforce) and using it to unlock more LTV





# We are Hiring !!

- NLP Engineer
- Full-Stack / Backend Engineer
- Lead Engineer
- QA Engineer
- Mid-level Sales & Marketing
- Sourcing Operations Manager

GET IN TOUCH

- [piyush@vahan.ai](mailto:piyush@vahan.ai)