INCREASE VOICE INPUT IN CHAT GPT

INDIAN VOICE-TECH AND MOBILE USAGE

There is a rapid growth in voice technology and mobile usage due to the internet and smart phone penetration

MASSIVE GROWTH

Over 1 billion voice searches monthly worldwide.

8.4 billion+ voice assistants already in use.

By 2026, more than 50% of internet users expected to use voice assistants.

INDIAN SCENARIO

700+ million smartphone users and low-cost internet are driving voice usage.

65% of Indians use voice search, especially in Tier-2/3 cities with regional languages.

Helps overcome linguistic diversity and accessibility barriers.

INDUSTRIES

E-commerce: Flipkart (voice shopping).

Banking: HDFC (voice banking).

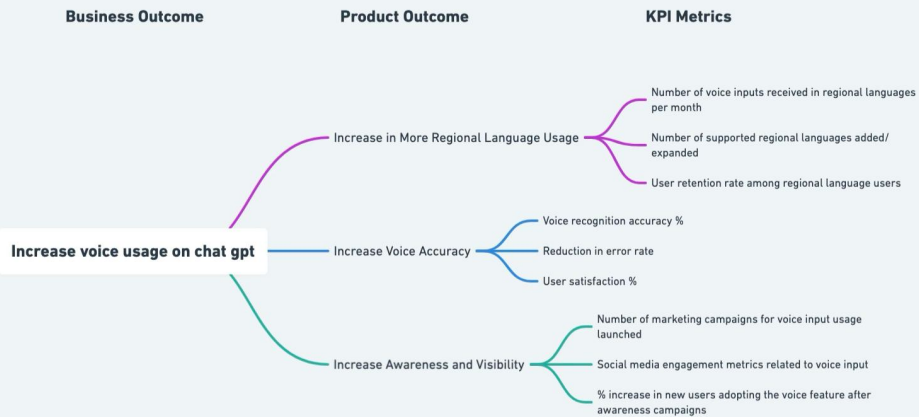Agri-tech: Ninjacart (farmer queries).

Google Assistant & JioSaavn adapting to regional languages and accents.

Source:- www.exchange4media.com

CHAT GPT UX AND OTHER VOICE APPS - COMPARISON

UX features	Chat gpt	Google assistance	Alexa
Integration	App only	Google search integration	Smart home device
Usage	Learning, brainstorming	Search, navigation	Home assistance, playing music
Language accent	multilingual/ struggles with strong accent	Strong regional language support	Device depended
Contect memory	Long term and dynamic	Short term	Short term
Advantage	Deep insight of the content	Hands free and local information	Eyes free and quick information

MAPPING PRODUCT OUTCOME TO BUSINESS OUTCOME

Business Outcome	Product Outcome	KPI Metrics
	Increase in More Regional Language Usage	Number of voice inputs received in regional languages per month
		Number of supported regional languages added/expanded
		User retention rate among regional language users
Increase voice usage on chat gpt	Increase Voice Accuracy	Voice recognition accuracy %
		Reduction in error rate
		User satisfaction %
	Increase Awareness and Visibility	Number of marketing campaigns for voice input usage launched
		Social media engagement metrics related to voice input
		% increase in new users adopting the voice feature after awareness campaigns

CHOSEN SEGMENT AND JUSTIFICATION

BUSINESS AND WORKING PROFESSIONALS - WHY?

1) This segment is more prone to using technological solutions like ChatGPT to improve productivity.

2) They are comfortable with technology, making them ideal early adopters and consistent users.

3) They already rely heavily on digital platforms and tools in their work environment.

4) This segment of users frequently use ChatGPT for tasks such as exploration, analysis, brainstorming, and even personal productivity.

INSIGHTS FROM USER RESEARCH

This segment users do not use voice input on chatgpt because

User PainPoint

1) 78 % of the users have encountered the problem of word mismatch when similar sound words are used while using voice input.

2) 64% of the users says ChatGPT AI engine capability efficiency in English is way too good than other regional languages.

3) 92% of users are not comfortable while using voice input feature in work place and other places like restaurant, park, metro, public transport due to privacy.

4) 68% of people feels Voice conversation mode in chat gpt response is too early before user completes the question.

PROBLEM DEFINITION CANVAS

What is the real problem?
Business and working professionals face difficulty while using voice input due to words mismatch when similar sound words are used,privacy concern while using it in work & public environment and too short response time in voice conversation mode

.

AD Arthi Deepak

Who are the customers facing problem?
Business and working professionals are the users who mostly face these problems

AD Arthi Deepak

How do we know it is a real problem?
31 respondents User research and 5 interview shows following data
78% says-issue mismatch of similar sound words
92% feels it as a privacy concern to use it in their work place and public places.
68% says - too early response time in voice conversation mode

AD Arthi Deepak

What is the value generated by solving this problem for the business?
1) Above mentioned customer pain point will come down.
2) Advantage over rivals.
3) Increased customer retention
4) Increased brand image and trust.

AD Arthi Deepak

What is the value generated by solving this problem for the target customers?
1) Increased user experience.
2) Increased the users' productivity in their workplace.
3) Increased in satisfaction level of all the users.

AD Arthi Deepak

Why should we solve this problem now?
1) Pain points are mostly an improvement and extension of the existing product feature.
2) Competitors are improving voice features, so delay risks losing market share.
3) Early fix = higher adoption and customer trust

AD Arthi Deepak

Problem Definition

- What is the problem?
    - Words mismatch when similar sound words are used,
    - Privacy concern while using it in work & public environment

- Who is facing the problem?
    - Business and working professionals are the users who mostly face these problems

- What is the business value that will be unlocked by solving the problem?
    - Enhanced competitive advantage
    - Improved customer retention
    - Increased market share and revenue growth
    - Strengthened brand image and customer trust

- How will the target users benefit if the problem is solved?
    - Increased user experience.
    - Increased the users' productivity in their workplace.
    - Increased comfort of the users to use voice input in public environment
    - Increased satisfaction level of all the users.
    - Reduced words mismatch issues.

- Why is it urgent to solve this problem now?
    - The core pain point of the users can be solved easily by adding an extension to the already existing mic feature
    - Competitors are improving voice features, so delay risks losing market share.
    - Early fix of the problem will lead to higher adoption and customer trust.

Goals
- High level Goals
    - Increase the voice mode of the user to use voice input in work and public environment comfortably
- Functional Metrics
    - % of increase in voice input usage / adaption
    - % of increase in duration of voice input usage
    - % of increase in retention of users
    - % of user satisfaction level

- Non - Functional Metrics
    - Increase in analogy of accuracy of mismatch words
    - Increase in the bandwidth of voice input to optimize faster processing

Non-Goals

- **I**ncrease the focus to more regional languages.
- Marketing and visibility campaign.

Validation of the problem

- Insights from user research/data

  31 respondents User research and 5 interview shows following data

    - **78% say** -issue with **mismatch of similar sound words**
    - **92% feel** it as a **privacy concern** to use it **in their work place and public places**.

Competitive landscape

- Google being big player in the market, can easily be the early one to introduce this feature and capture more wide market, google voice assistance may overlap chat GPT in low voice input feature and attract more users

Understanding the target audience
- User segment - **Business and Working Professionals**
- size of the segment - 31 survey and 5 interview respondents

Business and Working Professionals - why?
- This segment is more prone to use technology like ChatGPT to improve productivity.
- They are comfortable with technology, making them ideal early adopters and consistent users.
- They already rely heavily on digital platforms and tools in their work environment.
- This segment of users frequently use ChatGPT for tasks such as exploration, analysis, brainstorming, and even personal productivity
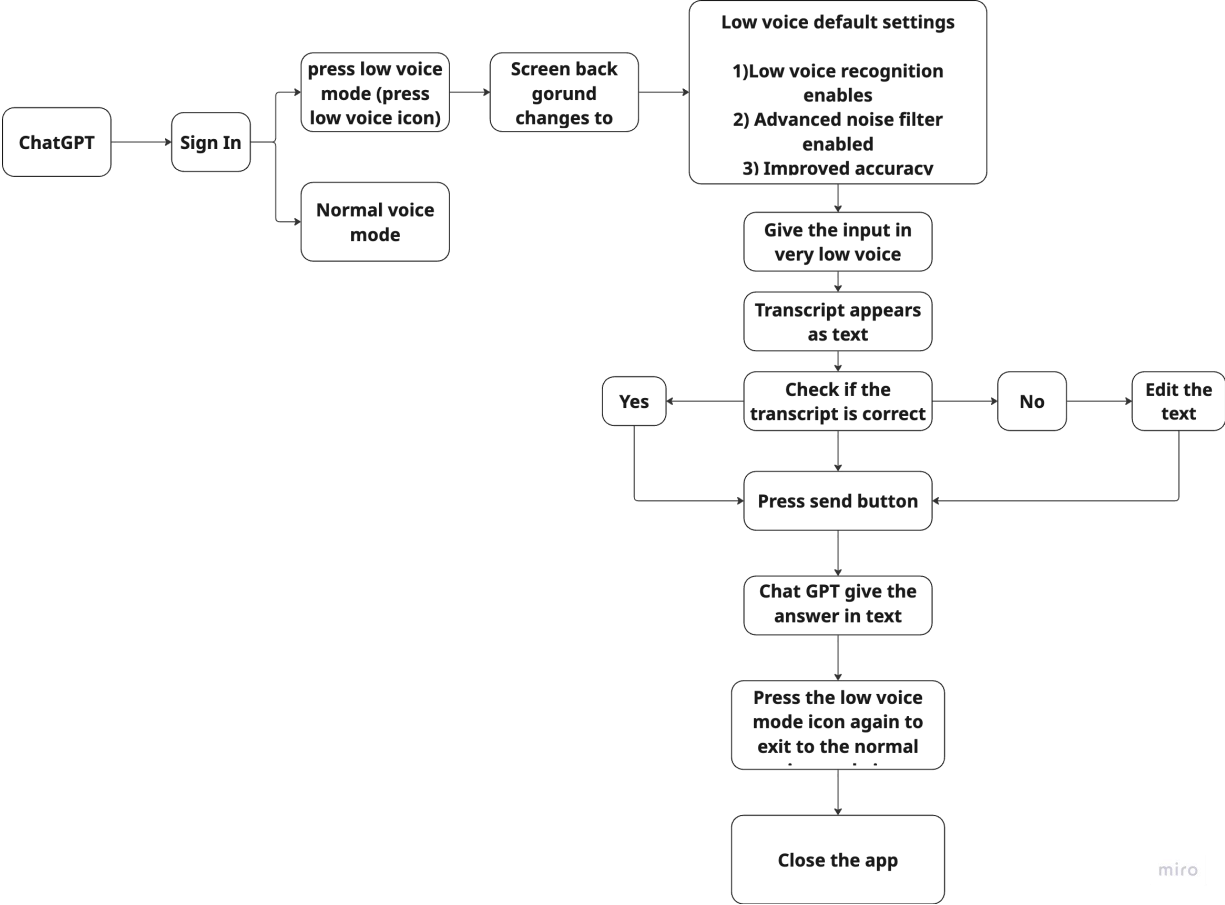
Key personas

- Maya is a Working Professional of age 28 years, she feels there is a lot of mismatch of similar sound words when using voice input while she was asking a query regarding their work related doubts. This she is been facing often and she feels very inconvenient to use voice input

- Suresh is a Working Professional of age 40 years feels embarrassed to use voice input on chat gpt during working hours in his work place in front of his colleagues. As his work place layout is very closely laid, he is unable to use voice input in Chat GPT while in office, hence he avoids it.

- Deepa who is Business person of age35 years says she feels inconvenient to use voice input mode on Chat GPT, while in a meeting or among other stakeholders, so she is not using voice input mode

- **Unmet needs (goals and pain points)**
  - Lack of accuracy and mistaken words
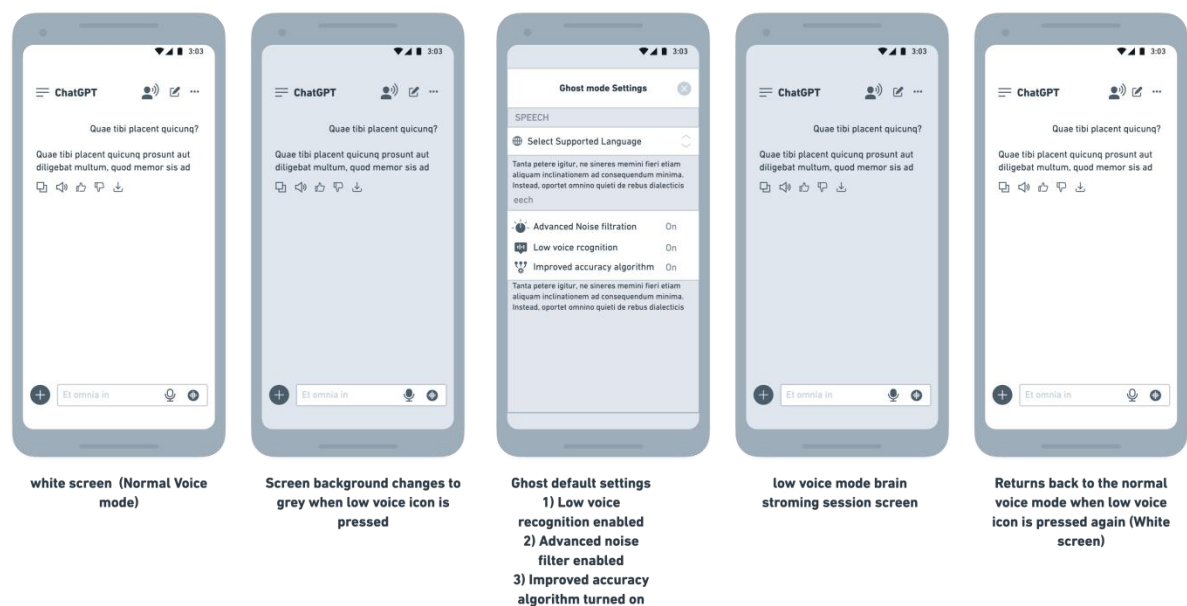 - Unable to use the voice input in work and public environment due to privacy and noisy environment

Solution to the defined problem:
- High Level Solution - Introduce super low voice mode - feature to enable the user to use voice input option in work and public environment comfortably, this feature will receive low voice as an input and enables advanced voice filtration settings with improved accuracy algorithm
- Introducing this feature will have a moderate reach to the target segments and it will create a strong impact with less effort and reduce the core pain point of the users.

User Flow Diagram (With new feature)

```
ChatGPT → Sign In → press low voice mode (press low voice icon) → Screen back gorund changes to → Low voice default settings

                                                                    1)Low voice recognition enables
                                                                    2) Advanced noise filter enabled
                                                                    3) Improved accuracy

          Sign In → Normal voice mode
```

Low voice default settings

1)Low voice recognition enables
2) Advanced noise filter enabled
3) Improved accuracy

↓

Give the input in very low voice

↓

Transcript appears as text

↓

Yes ← **Check if the transcript is correct** → No → **Edit the text**

↓

Press send button

↓

Chat GPT give the answer in text

↓

Press the low voice mode icon again to exit to the normal

↓

Close the app

miro

Wire frames

white screen (Normal Voice mode)

Screen background changes to grey when low voice icon is pressed

Ghost default settings
1) Low voice recognition enabled
2) Advanced noise filter enabled
3) Improved accuracy algorithm turned on

low voice mode brain stroming session screen

Returns back to the normal voice mode when low voice icon is pressed again (White screen)

Key Features

- Key Features: that will be developed for the users
    - super low voice mode - allows and receives voice input even in low volume while in work and public environment which will have default setting of low voice input adaption, advanced noise filtration and improved accuracy

- Benefits that will be developed for the users
    - ○ Better low volume voice input received
    - ○ Convenient for the users to use voice input even in work and public place privately
    - ○ Improved accuracy and reduction in noise
    - ○ Better user experience

- Key Logic: algorithm changes, schema changes, new data types etc.
    - Update voice recognition algorithm for low-volume accuracy.
    - Bandwidth optimization for faster processing

System diagram
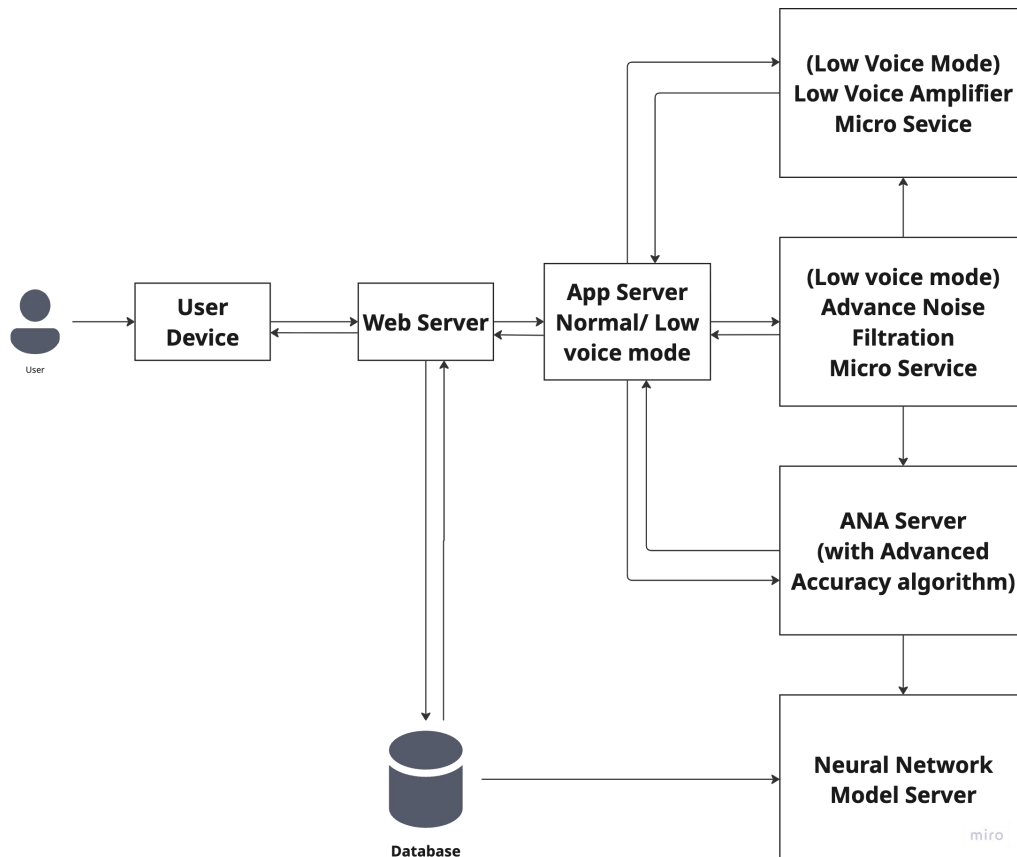
Data Instrumentation

- Data instrumentation for tracking key metrics
  - % of daily active users toggle to low voice mode
  - % users frequently using the new feature (repeat users)
  - % duration of users engaged with the low voice mode
  - % of reduction is errors in receiving low voice input

- Edge error case handling
  - Unable to recognize
  - No network connection
  - try again
- Product marketing
  - Display banner
  - Awareness campaign
  - Slow blinking of super low voice icon

Launch Readiness

- Key Milestones
  - Wireframes & PRD
  - Design complete
  - Discussion and Testing
  - Development
  - Internal testing done
  - Beta version
  - Complete version

- Launch Checklist
  - Feature ready to use.
  - Increased band width for faster voice input result.
  - All feature working flow.
  - Analytical department readiness to track increase in usage

- Experimentation plan
  - A/B testing if required

- Future Iteration
  - Deducts low voice and automatically switches to low voice mode(low voice mode)
  - More regional languages can be added.

- Risk and Mitigations
  - response / reply may get delayed
  - Latency issue
  - Lack of awareness of this new low mode option available

Open Questions & Decisions Taken

- Open questions
    - Should the increase in regional language support be considered
    - Should we conduct an awareness campaign?
- Outline what has been descoped
    - Regional language support
    - Awareness campaign

- List out the trade-offs made
    - Exclude regional language support as there are too many languages and accent and the cost involved is higher instead , pre-selecting one language from top 5 global trained language for voice input earlier will lead to better user experience increasing the use of voice input even in work and public places by the user

Appendix
For the curious reader

- https://www.surveymonkey.com/results/SM-6UcuEmVtnabCpEX2WYoVXw_3D_3D/