Samaahar

Knowledge Spoken in Hinglish

AI-Powered Podcast Generator

Production Ready Application

Problem Statement

Why we built Samaahar

- Complex information suffers from **low accessibility** for the masses.
- High-quality resources are predominantly **English-heavy**.
- **Hindi-first users** are significantly underserved online.
- Rising demand for **audio-first learning** experiences.

TARGETING 4 KEY SEGMENTS:

Kids **Teenagers** **Adults** **Elderly**

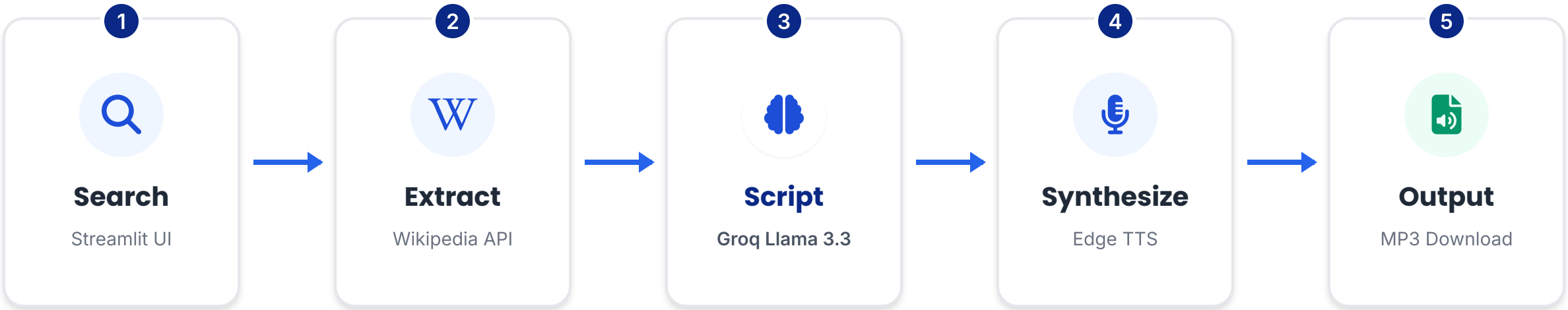
Solution Overview

Bridging the gap with AI

- Transforms **Wikipedia articles** into engaging **Hinglish podcasts**.
- Dynamic two-host format featuring **Rajesh & Priya**.
- **Audience-adaptive** content: Adjusts language, tone, and pacing.
- Fully **automated pipeline** from search to synthesis.
- Instant output: **Downloadable MP3** ready for sharing.

System Architecture

End-to-End Content Generation Pipeline



Resilience Layer

Robust Error Handling & Recovery

Auto-Retries

Rate Limit Backoff

Graceful Fallbacks

Why Groq Llama 3.3 70B?

Model Selection Rationale

Our Choice: Groq

Llama 3.3 70B Versatile

Ultra-Fast Inference

~130 tokens/sec enables near-instant scripts

Dialogue Quality

Excellent conversational flow & natural tone

Multilingual

Strong English + Hindi mixing capabilities

Generous Free Tier

Ideal for hackathons & scaling

14.4k TPM

Alternatives

Why we skipped them

OpenAI GPT-4

Too expensive for scale; latency issues

Google Gemini

Slower output generation; stricter filters

Verdict

Groq offers the best speed/cost ratio for real-time apps.

Why Microsoft Edge TTS?

Text-to-Speech Selection Rationale

Our Choice: Edge TTS

Microsoft Azure Cognitive Services

High-Quality Indian Voices

Natural sounding voices like Prabhat & Neerja

Natural Prosody Control

Fine-tune pitch, rate, and pauses via SSML

Generous Free Access

No strict API limits for development use

Unlimited Synthesis

Perfect for generating long-form podcasts

FREE

Alternatives

Why we skipped them

Google Cloud TTS

High cost per character; complex billing

ElevenLabs

Very restrictive free tier; low character limits

Verdict

Edge TTS provides the best quality/cost balance for Indian voices.

Key Features: Audience Profiles

Dynamic Content Adaptation & Voice Mapping

Kids

Energetic tone, simple vocabulary, high pitch

Rate: +10% **Voice: Neerja**

Teenagers

Trendy slang, faster pace, casual style

Rate: +5% **Voice: Aarav**

Adults

Balanced, professional, informative tone

Rate: Normal **Voice: Prabhat**

Elderly

Slower pace, clear diction, respectful

Rate: -10% **Voice: Prabhat**

Key Features: Workflow Efficiency

Streamlined Experience & Automation

3-Click Process

Reduced from 6+ clicks to just 3 steps

Auto-Generation

Instant script creation on selection

Instant Audio

One-tap synthesis after approval

Mobile-First UI

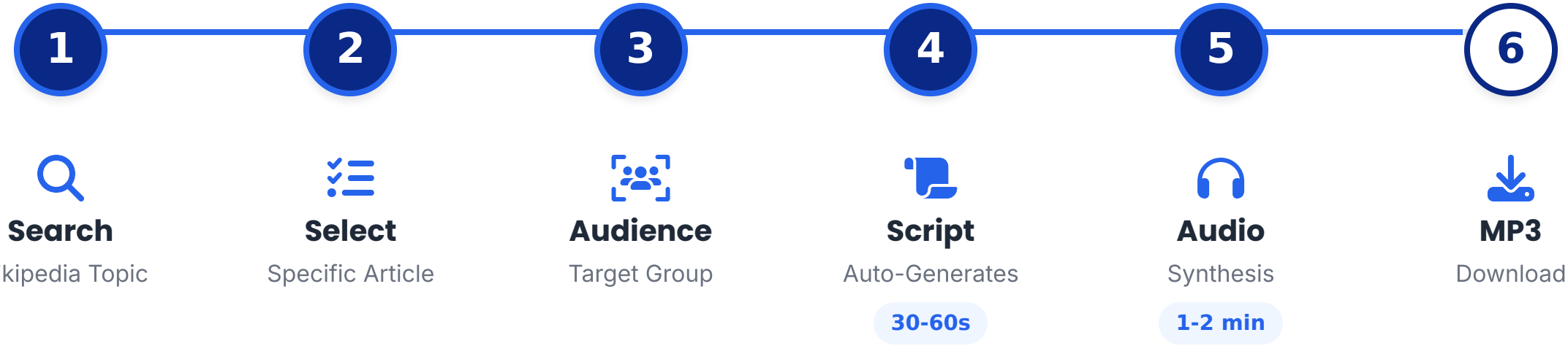
No sidebar, vertical touch layout

Jio-Inspired

Modern gradients & brand colors

User Journey

From Search to MP3 in 3 Clicks



Total Time: ~2.5 Minutes

Compared to 30+ minutes for manual podcast creation. Zero technical skills required.

What Worked Well

Part 1: Technical Wins

Rapid Script Generation

Full podcast scripts generated in just 30–60 seconds

avg 45s

Consistent Hinglish Tone

Seamless blending of Hindi vocabulary within English syntax

Distinct Audience Voices

Clear acoustic differentiation between Kids, Teens, and Adults

Rapid Iteration Speed

What Worked Well

Part 2: Optimization Wins

Token Usage Optimized

Limited input to 1500 chars to maximize processing speed

~3000 tokens

Resilient Rate Limits

Exponential backoff logic handles 429 API errors gracefully

Auto-Retry

Mobile-First UI Design

Removed sidebar and optimized layout for touch devices

Responsive

High Pipeline Reliability

Consistent end-to-end execution without crashes

95%+ Success

What Didn't Work

Part 1: Major Challenges

Python 3.13 Incompatibility

Streamlit default runtime broke 'pydub' audio library

Fixed: Force Python 3.11

Groq API Rate Limits

Frequent '429' errors during heavy testing bursts

Fixed: Auto-Retry Logic

Invalid Voice Names

Initial config used non-existent Azure voice IDs

Fixed: Verified IDs Only

What Didn't Work

Part 2: UX & Navigation Issues

TTS Pitch Parameter Errors

Edge TTS rejected pitch adjustments for certain Indian voices

Fixed: Use Rate Only

Complex User Flow

Initially required 6+ clicks to generate a single podcast

Fixed: Streamlined to 3 Clicks

Navigation Loops

App state reset unexpectedly during audience selection

Fixed: Step-Based State

Technical Challenges

Problem Solving & Resilience Strategy

Challenge	Implemented Solution
pydub runtime error	✓ Downgraded to Python 3.11
Rate Limits (429)	✓ Exponential Backoff & Retry Logic
Invalid Voice Names	✓ Verified Indian Voice List
TTS Pitch Errors	✓ Removed pitch, used Rate only
Navigation Loops	✓ Step-based State Tracking

Performance Metrics

Real-world Application Benchmarks

45s

AVG SCRIPT TIME

Groq Llama 3.3 generation

90s

AVG AUDIO TIME

Edge TTS synthesis

~2.5m

TOTAL TURNAROUND

Concept to MP3 file

~3000

TOKENS USED

Per 2-minute episode

~2 MB

FILE SIZE

Optimized MP3 output

95%+

SUCCESS RATE

After retry logic implemented

Use Cases

Real-World Application Scenarios

Education

Simplify complex topics into bite-sized audio lessons for students.

News & Media

Generate daily Hinglish briefings for local community engagement.

Corporate Training

Employee onboarding and safety guidelines in familiar dialects.

Accessibility

Instant information access for visually impaired or illiterate users.

Local Outreach

Public announcements and community alerts in spoken Hinglish.

Future Enhancements

Product Roadmap & Scale Strategy

01

More Languages

Expand beyond Hinglish to Tamil, Bengali, and Telugu support.

02

Voice Cloning

Allow users to clone their own voice for personalized hosting.

03

Long-Form Shows

Scale context window to support 10–15 minute deep-dive episodes.

04

Audio Production

Auto-mix background music beds and sound effects (SFX).

05

Series Mode

Generate multi-episode series from a single broad topic.

06

Enterprise Scale

Paid tier with higher quotas and API access for developers.

Conclusion

Key Takeaways & Project Status

End-to-End Pipeline Proven

Fully automated conversion from Wikipedia topic to MP3 podcast is functional.

Hinglish Audio Viable

Successfully achieved context-aware language blending for Indian audiences.

Resilient Architecture

Overcame rate limits and API instability with robust retry logic and error handling.

Production-Ready UX

Polished, mobile-first interface optimized for speed and ease of use.

Scalable Foundation

Architecture ready for multi-language expansion and longer content formats.

Thank You!

Samaahar

Knowledge Spoken in Hinglish