Folk Lore Hub - Project Report

# 1. Team Information

- Team Name: BharatiyaVaani  
- Project Created By: Dheeraj  
- Team Members:  
 1. Suraj – Full Stack Developer  
 2. Rahul – UI/UX Designer  
 3. Neha – AI Integration Specialist  
 4. Abhita – Community Outreach & Documentation

# 2. Application Overview

Project Title: Folk Lore Hub  
  
Problem Statement:  
Traditional Indian stories, proverbs, and local wisdom are at risk of being lost as oral transmission declines.   
There is no easy, tech-driven solution for people in villages and diverse regions to share and preserve these treasures,   
especially where internet connectivity is limited.  
  
MVP Scope (1 Week):  
- Simple, intuitive Streamlit interface (mobile-first)  
- "Submit a Story/Proverb" (text or voice input, converted to text)  
- Language selection (at least 6 Indian regional languages + English/Hindi)  
- Optionally upload a related local image (optional for corpus enrichment)  
- Instant, offline caching – works even if internet is briefly down  
- Public view of submitted stories with upvote/sharing features

# 3. AI Integration Details

- Speech-to-text (optional, client-side via open-source Whisper model for voice)  
- Automatic transliteration to Devanagari and Latin scripts  
- Language detection and tagging using Hugging Face models  
- Text normalization (remove duplicates, basic spelling corrections using OSM+AI)  
  
All AI components run as open-source and support offline-first execution on affordable edge devices, with fallback to serverless Hugging Face Spaces for heavy operations.

# 4. Technical Architecture & Development

- Frontend: Streamlit, optimized for mobile  
- Backend: Python, leveraging Hugging Face open-source AI models  
- Offline-First: IndexedDB/local browser storage for stories and images, with auto-sync when online  
- Deployment: Hugging Face Spaces (open access)  
- Version Control: GitLab

# 5. User Testing & Feedback (Week 2)

- 12 beta testers from rural communities in Telangana & Odisha  
- Tasks: Submit a folk tale or proverb in their language and vote on others' submissions  
- Feedback: Surveys, WhatsApp voice notes; focus on app clarity, offline mode, transliteration accuracy  
  
Iterations:  
- Improved font rendering for Telugu/Odia  
- Added storytelling audio snippet support  
- Increased offline cache size  
- Added optional anonymous contribution

# 6. Project Lifecycle & Roadmap

Week 1: Development  
- Build UI, integrate at least 2 local language models for input & transliteration  
- Set up offline caching (test in airplane mode)  
- Deploy MVP on Hugging Face Spaces  
  
Week 2: User Testing & Iteration  
- Distribute via regional WhatsApp groups, schools  
- Collect feedback; patch font/UX bugs, refine AI detection  
  
Weeks 3-4: User Acquisition & Corpus Growth  
- Target: School teachers, college clubs, regional WhatsApp and Telegram groups, NGOs  
- Messaging: “Preserve your grandmother’s stories for the world”  
- Methods: Digital posters in local languages, demo videos, partnerships with local radio/teachers  
- Results: [Enter target numbers] unique users, [X] stories/proverbs submitted (track language/city), feedback loop

# 7. Post-Internship Vision & Sustainability

- Add more languages, community leaderboard, recognition for top contributors  
- Scale to include short audio-visual tales and folk songs  
- Partner with universities, SHGs (Self Help Groups), and cultural bodies for corpus validation  
- Open research access for linguists and AI developers  
- Apply for cultural preservation grants

# 8. Repository Structure and Additional Files

- README.md: Clear usage, setup, deployment, contribution, contact info.  
- CONTRIBUTING.md: Guidelines for new languages, scripts, AI model enhancements.  
- CHANGELOG.md: Weekly updates and bugfixes log.  
- requirements.txt: Streamlit, Hugging Face Transformers, openai-whisper, Pillow, langdetect, etc.  
- LICENSE: MIT or Apache 2.0.  
- /app/: Core Streamlit code.  
- /test/: Unit/integration tests for AI modules & storage.

# 9. Example README.md

Folk Lore Hub  
  
An open-source Streamlit app to collect and preserve Indian folk tales and proverbs in multiple languages (offline-ready!).  
  
Features:  
- Submit stories/proverbs in 6+ languages  
- Speech-to-text (optional)  
- Works even with spotty internet  
- Open-source AI for transliteration & language tagging  
- Community upvotes, sharing, and feedback  
  
Install (local dev):  
pip install -r requirements.txt  
streamlit run app/main.py  
  
License: MIT  
Contact: folkteam@example.com