VAANI AI - PRD

🎤 VAANI AI - Product Requirements Document

Migrated from personal space

Executive Overview

VAANI is a voice AI feature for Vyapar App that enables Indian MSMEs to manage their business operations through natural voice commands in their preferred language. It's designed as a mobile-first solution targeting the next 500 million Indian entrepreneurs who prefer speaking over typing.

🎯 Strategic Vision & Market Opportunity

The Core Insight

- Not about 10x improvement in any single workflow
- About 50 small improvements that are each 2x better
- Real value: Capturing data that currently goes unrecorded, not just saving time

Market Reality

- 72% of Indian smartphone users actively use voice assistants
- 500M+ Indians prefer voice input over typing
- 40% of rural users rely on voice due to literacy constraints
- 63.3M MSMEs in India with only 12% digital maturity

Business Impact

- Transaction Capture: From 70% to 95%
- Expense Tracking: From 40% to 80% completeness
- User Retention: 51% improvement in D90 retention
- Premium Conversion: 2.4x higher for voice users

📋 Selected Use Cases (MVP)

Primary Use Cases

1. Quick Transaction Logging

- **Frequency:** 50+ times/day
- **Time saved:** 45 seconds → 15 seconds per transaction
- **Example:** "Ramesh ne 5 kg chawal 1250 rupees cash diya"

2. Intelligent Expense Tracking

- **Frequency:** 20+ times/day
- **Impact:** Captures ₹15,000-20,000 monthly leakage
- **Example:** "Delivery ke liye 200 rupees diye"

3. Real-time Inventory Alerts

- **Frequency:** 30+ times/day
- **Impact:** 40% stockout reduction
- **Example:** "Rice sirf 10 bag bache hain"

4. Instant Party Creation

- **Frequency:** 5+ times/day
- **Time saved:** 90 seconds → 20 seconds
- **Example:** "Add customer Sharma Electronics 9876543210"

5. Quick Catalog Building

- **Frequency:** 3+ times/day
- **Impact:** 3x catalog growth rate
- **Example:** "Add Tata Salt 1kg MRP 28 selling 26"

MVP Scope Limitation

- ✅ Create operations only
- ❌ No edit/delete via voice in MVP
- Focus on accuracy and trust building first

🏗️ Technical Architecture

Modular Design

Each use case has three modules:

1. Data Module

- Bucket A: Necessary fields (required)
- Bucket B: Additional fields (optional with smart defaults)

2. Prompt Module

- Initial extraction prompts
- Missing field prompts (contextual)
- Confirmation prompts
- Error handling messages

3. Execution Module

- Pre-execution validation
- Step-by-step execution
- Post-execution actions

Key Technical Features

- **Hybrid Processing:** On-device for common commands, cloud for complex
- **Offline Support:** Basic operations work without internet
- **Language Support:** Hindi-English code-mixing from day one
- **Accuracy Requirements:** 95% for amounts, 85-90% for general fields

---

🎨 User Experience Design

Entry Points

1. Floating Voice Button - Persistent on all screens
2. Home Screen Widget - Direct access from phone home
3. Contextual Icons - Within input fields
4. Empty State Prompts - Encouraging voice use

Core Interface States

1. **Listening** - Animated waves, real-time transcription
2. **Processing** - Understanding indication, showing interpretation
3. **Confirmation** - Visual summary before committing
4. **Success** - Quick actions for next steps

Design Philosophy

- Voice as natural extension, not separate feature

- Visual feedback for trust building
- Always provide manual fallback
- Progressive disclosure of complexity

📊 Success Metrics

Phase 1 (3 months)

- 10% of DAU try voice at least once
- 5% become weekly active voice users
- 25% increase in data capture for voice users
- NPS of 45+ from voice cohort

Long-term (12 months)

- Voice becomes primary input for 20% of users
- Voice transactions = 20% of platform total
- Category recognition as voice-enabled platform
- Foundation for regional language expansion

Go/No-Go Criteria

- 80% accuracy in real environments
- 70% task completion without training
- 30% time reduction vs manual
- 60% D30 retention

🛡️ Risk Mitigation

Technical Risks

- **Accuracy in noisy shops:** Start with number-heavy commands
- **Language mixing:** Specialized training on Hinglish patterns
- **Infrastructure costs:** Hybrid architecture, progressive rollout

User Adoption Risks

- **Trust in AI for money:** Visual confirmation for all financial data
- **Behavior change:** Position as addition, not replacement
- **Privacy concerns:** Clear indicators when voice is active

---

🚀 Implementation Approach

Development Priorities

1. Core voice recognition integration
2. Transaction logging (highest frequency use case)
3. Visual confirmation system
4. Offline capability
5. Performance optimization for ₹5,000-15,000 phones

Pilot Strategy

- Begin with 100-user pilot in Delhi NCR
- Single use case perfection before expansion
- Mobile-first development
- Continuous learning and adaptation

💡 The Vision

By 2027, a vegetable vendor in Varanasi runs her entire business by voice:

"Vyapar, aaj subah Sharma ji ne 2 kilo aloo liye the, Mishra aunty ne 500 ka udhar chukaya, aur mujhe kal ke liye 20 kilo pyaaz order karna hai"

And Vyapar handles everything - invoice, payment, inventory, purchase order.

This isn't just a feature. It's financial inclusion at scale.

Last Updated: August 4, 2025