

Community Crowdsourcing Framework for Endangered Language Preservation

Executive Summary

This framework provides community groups with practical tools, technology recommendations, and step-by-step processes to launch crowdsourced endangered language preservation projects. It emphasizes community ownership, minimal technical barriers, and sustainable participation models.

1. Getting Started: The 30-Day Launch Framework

Week 1: Community Mobilization and Planning

Day 1-3: Stakeholder Identification

- **Elders and Fluent Speakers:** Identify 3-5 community elders willing to participate
- **Youth Champions:** Recruit 5-10 younger community members (ages 16-35) as project coordinators
- **Cultural Authorities:** Engage traditional leaders, cultural committees, or tribal councils
- **Tech-Savvy Volunteers:** Find 2-3 people comfortable with smartphones and basic digital tools

Day 4-7: Community Agreement

Create a simple one-page community charter covering:

- Who can participate in recording and validation
- What types of content are appropriate to collect (avoiding sacred/restricted materials)
- How the collected materials will be used and stored
- Basic consent protocols for recordings

Template Community Charter:

[Language Name] Preservation Project Charter

Purpose: To collect, preserve, and share our language for current and future generations.

Participants: Any community member may contribute recordings, stories, or translations.

Content Guidelines:

- ✓ Daily conversations, greetings, family terms
- ✓ Traditional stories approved by elders
- ✓ Names of plants, animals, places
- ✗ Sacred ceremonies or restricted knowledge

✗ Personal/private information

Data Use: All recordings remain community property. Materials may be used for:

- Educational apps and websites
- Community language classes
- Academic research with community approval

Contact: [Project Coordinator Name and Phone]

Signed: [Elder Representatives] [Date]

Week 2: Platform Setup and Tool Selection

Choose Your Primary Platform (Pick One):

Option A: Udio for Communities (Recommended for Beginners)

- Web-based, works on any device
- Simple record-and-upload interface
- Built-in community validation features
- Free for community use
- Setup time: 2 hours

Option B: Mozilla Common Voice

- Established platform with 100+ languages
- Requires community application but provides ongoing support
- Professional-grade data management
- Integration with larger preservation networks
- Setup time: 2-4 weeks

Option C: Custom Community Platform

- Use combination of Google Drive + Simple recording apps
- Full community control but requires more management
- Can migrate to professional platforms later
- Setup time: 1 week

Essential Hardware Checklist:

Minimum Setup (\$200-500):

- ☐ 2-3 smartphones with good microphones (iPhone 8+ or Android equivalent)

- External clip-on microphones for better quality (\$20-40 each)
- Portable phone stands (\$10-15 each)
- Power banks for extended recording sessions (\$25-35 each)

Upgraded Setup (\$800-1500):

- Digital audio recorder (Zoom H1n or similar, \$120-200)
- Tablet for larger screen validation work (\$200-400)
- Wireless microphone system for group recordings (\$150-300)
- Basic lighting for video recordings (\$50-100)

Week 3: Content Strategy and Collection Framework

Priority Content Categories (Start Here):

1. Essential Vocabulary (Target: 500 words in first month)

- Numbers 1-100
- Family relationships (mother, father, grandmother, etc.)
- Daily objects (food, clothing, tools)
- Colors, body parts, weather terms
- Common verbs (eat, sleep, walk, talk)

2. Conversational Phrases (Target: 100 phrases)

- Greetings and farewells
- "How are you?" and responses
- "What is this?" and answers
- Common questions and responses
- Polite expressions (please, thank you, excuse me)

3. Cultural Knowledge (Target: 20 recordings)

- Traditional stories (non-sacred)
- Song lyrics (if appropriate)
- Place names and their meanings
- Traditional practices and their terminology
- Seasonal activities and associated vocabulary

Recording Guidelines Template:

Recording Session Checklist:

Before Recording:

- ☐ Quiet space with minimal background noise
- ☐ Phone in airplane mode (to prevent interruptions)
- ☐ External microphone connected and tested
- ☐ Project app or recording software ready
- ☐ List of words/phrases to record prepared

During Recording:

- ☐ Speaker states their name and the word/phrase they'll say
- ☐ Pause 2 seconds, then say the target word/phrase clearly
- ☐ Pause 2 seconds after
- ☐ Record each item 3 times for quality options
- ☐ Take breaks every 20-30 recordings

After Recording:

- ☐ Listen to recordings for clarity
- ☐ Note any cultural context or usage notes
- ☐ Upload to project platform within 24 hours
- ☐ Log what was recorded in project spreadsheet

Week 4: Quality Control and Community Validation

Validation Workflow:

1. **Peer Review:** Every recording reviewed by 2 other community members
2. **Elder Approval:** Cultural content reviewed by designated elders
3. **Accuracy Check:** Transcriptions verified against audio
4. **Usage Notes:** Context and cultural information added

Community Validation App Interface Design:

Simple validation screen shows:

- Audio player with recorded word/phrase
- Phonetic transcription (if available)
- English translation

- Cultural context notes
 - Validation buttons:
 - ✓ Accurate pronunciation and meaning
 - ⚠ Needs minor correction (with comment box)
 - ✗ Incorrect (requires elder review)
 - ? Unsure (flag for elder review)
-

2. Technology Stack: Tools and Platforms

2.1 Recording and Collection Tools

Mobile Apps for Data Collection:

1. Open Source Recorder (Free)

- Basic audio recording with metadata
- Offline capability for remote areas
- Simple file naming conventions
- Export to multiple cloud platforms

2. Community Voice Collector (\$5-15/month)

- Custom forms for recording sessions
- Automatic audio quality checking
- Built-in community validation workflows
- Progress tracking and gamification

3. LinguaRecorder (Research Tool)

- Designed specifically for linguistic fieldwork
- Automatic transcription attempts
- Integration with academic databases
- Free for community use

2.2 Data Management and Storage

Storage Solutions:

Community Google Drive Setup (Free - 15GB, \$2-6/month for more)

Folder Structure:

📁 [Language Name] Preservation Project

📁 01_Raw_Recordings

 Vocabulary

 Phrases

 Stories

 Songs

 02_Validated_Content

 Approved_Audio

 Transcriptions

 Cultural_Notes

 03_Project_Management

 Participant_Consent_Forms

 Progress_Tracking

 Community_Agreements

 04_Educational_Materials

 Word_Lists

 Learning_Exercises

 Community_Resources

Backup Strategy:

- Primary: Google Drive or Dropbox
- Secondary: External hard drive kept by project coordinator
- Archive: Submit to ELAR or PARADISEC for permanent preservation

2.3 Validation and Quality Control Platforms

Community Validation Portal Features:

- Random assignment of recordings to validators
- Progress tracking for individual contributors
- Dispute resolution workflow for disagreements
- Cultural sensitivity flags and elder review queue
- Simple mobile-friendly interface

Validation Gamification Elements:

- Points for completed validations
- Badges for different types of contributions
- Community leaderboards (with privacy options)

- Monthly recognition for top contributors
-

3. Participant Recruitment and Engagement

3.1 Community Outreach Strategy

Launch Event Template:

Community Language Preservation Gathering

Agenda (2-3 hours):

1. Welcome and traditional opening (30 min)
2. Project presentation: Why preserve our language? (30 min)
3. Technology demonstration: How easy it is to participate (30 min)
4. Hands-on recording session (45 min)
5. Sign-up and next steps (15 min)
6. Community meal and informal discussion (30 min)

Materials needed:

- Demonstration devices (tablets/phones)
- Simple handouts in community language
- Sign-up sheets
- Contact information cards
- Traditional food/refreshments

3.2 Incentive and Recognition Programs

Participation Rewards:

Individual Recognition:

- Digital certificates for milestone contributions
- Featured contributor spotlights in community newsletter
- Priority access to finished educational materials
- Invitation to present at language conferences

Community Benefits:

- Progress toward collective goals (e.g., "500 words recorded")
- Development of custom learning apps for the community

- Integration with local schools and cultural programs
- Potential for small grants based on successful participation

Micro-Incentive Ideas:

- Coffee shop gift cards for recording 25 words
- Community t-shirts for validators who check 100 recordings
- Elder appreciation lunches sponsored by project
- Youth workshop opportunities for tech-savvy participants

3.3 Addressing Common Participation Barriers

Technology Anxiety:

- Pair tech-nervous participants with tech-comfortable youth
- Provide one-on-one training sessions
- Create laminated quick-reference cards
- Offer phone/in-person support

Time Constraints:

- Flexible participation: 5 minutes or 2 hours both valuable
- Mobile-friendly tools for recording during daily activities
- Group recording sessions that combine social time with data collection
- Integration with existing community gatherings

Cultural Concerns:

- Clear protocols about what content is appropriate
- Elder oversight and approval processes
- Option to mark recordings as "community only" vs "public"
- Regular community meetings to address concerns

4. Data Collection Workflows

4.1 Systematic Recording Process

Daily Vocabulary Collection (15-20 minutes/day):

Week 1 Focus: Family and Relationships

Monday: Immediate family terms (10 words)

Tuesday: Extended family terms (10 words)

Wednesday: Life stages and ages (10 words)

Thursday: Relationship descriptions (10 words)

Friday: Community roles and titles (10 words)

Weekend: Review and validation of week's recordings

Conversation Recording Sessions (1-2 hours/week):

- Set up natural conversation scenarios
- Record elder-youth exchanges
- Capture storytelling sessions
- Document traditional activity descriptions

4.2 Quality Assurance Protocols

Three-Tier Validation System:

Tier 1: Community Peer Review

- Any community member can validate recordings
- Focus on basic accuracy and clarity
- Simple thumbs up/down or 1-5 star rating

Tier 2: Language Expert Review

- Most fluent speakers in community
- Focus on pronunciation and cultural appropriateness
- Can suggest corrections or additional context

Tier 3: Elder Council Approval

- Final review for cultural sensitivity
- Approval for public sharing vs community-only use
- Resolution of disputes or concerns

Quality Metrics to Track:

- Average number of recordings per participant per week
- Percentage of recordings that pass each validation tier
- Time from recording to final approval
- Community satisfaction with process (monthly surveys)

5. Funding and Sustainability Models

5.1 Crowdfunding Campaign Framework

Campaign Title Examples:

- "Save Our Language: [Language Name] Digital Preservation"
- "Voices of Our Ancestors: Technology Meets Tradition"
- "[Community Name] Language Legacy Project"

Campaign Structure Template:

Goal 1: Basic Operations (\$5,000)

- Recording equipment for 5 community coordinators
- Platform setup and first year hosting
- Community appreciation events and incentives
- Basic app development for learning tools

Goal 2: Enhanced Features (\$15,000)

- Professional audio editing and cleanup
- Custom mobile app development
- Integration with educational institutions
- Elder documentation project (video interviews)

Goal 3: Long-term Sustainability (\$30,000)

- Two-year operating budget
- Professional development for community coordinators
- Archive submission and long-term preservation
- Replication toolkit for other communities

Campaign Content Elements:

- 2-3 minute video featuring elders and youth
- Story of why language preservation matters to community
- Demonstration of technology in action
- Clear breakdown of how funds will be used
- Regular updates showing progress and impact

5.2 Grant Application Support

Quick Grant Opportunities (\$500-5,000):

Community Foundation Grants:

- Many local foundations fund cultural preservation
- Application deadlines typically quarterly
- Emphasis on community benefit and engagement

- Simple application process (often 2-3 pages)

Corporate Diversity and Inclusion Funds:

- Technology companies often support digital inclusion
- Banks and utilities sometimes fund community development
- Local businesses may sponsor for community goodwill
- Focus on measurable community impact

Academic Institution Partnerships:

- University linguistics departments need fieldwork sites
- Graduate students may provide free technical assistance
- Potential for small research grants (\$1,000-10,000)
- Access to professional equipment and expertise

Template Grant Budget:

Language Preservation Project Budget - Year 1

Personnel (Community Coordinators): \$8,000

- 2 part-time coordinators at \$4,000 each

Equipment and Technology: \$3,000

- Recording equipment: \$1,500

- Platform setup and hosting: \$800

- Mobile devices and accessories: \$700

Community Engagement: \$2,000

- Launch event and workshops: \$800

- Participant incentives and recognition: \$600

- Elder appreciation activities: \$600

Administrative Costs: \$1,000

- Insurance and legal fees: \$400

- Communications and outreach: \$300

- Miscellaneous supplies: \$300

Total Project Budget: \$14,000

Community In-Kind Contribution: \$6,000

(Volunteer time, meeting spaces, meals)

Grant Request: \$8,000

5.3 Revenue Generation and Self-Sustainability

Potential Revenue Streams:

Educational Products:

- Custom language learning apps (\$2-5/download)
- Printed materials and children's books
- Online courses and workshops
- Cultural tourism components

Service Provision:

- Consulting for other communities starting similar projects
- Translation services for local government/healthcare
- Cultural education programs for schools
- Traditional knowledge documentation for external organizations

Intellectual Property Licensing:

- Carefully managed licensing of language content for media
- Educational content licensing to universities
- Technology platform licensing to other communities
- Trademark registration for language-based products

6. Technology Implementation Guide

6.1 Platform Setup Walkthroughs

Mozilla Common Voice Community Setup:

Step 1: Application Process

1. Visit commonvoice.mozilla.org/contribute
2. Click "Request new language"
3. Complete community information form
4. Provide proof of community support (signatures, letters)

5. Wait for approval (typically 2-4 weeks)

Step 2: Community Coordinator Training

1. Complete Mozilla's online coordinator course (4 hours)
2. Set up community validation workflows
3. Create recording guidelines in local language
4. Test platform with small group before full launch

Custom Platform Development Using Open Source Tools:

Technical Requirements:

- Web hosting account (\$5-15/month)
- Domain name (\$10-15/year)
- SSL certificate (free with Let's Encrypt)
- Basic HTML/CSS knowledge or developer volunteer

Recommended Stack:

- Frontend: Simple HTML5 with Web Audio API
- Backend: Node.js with Express framework
- Database: SQLite for small projects, PostgreSQL for larger ones
- Storage: AWS S3 or Google Cloud Storage for audio files

Development Timeline:

- Week 1: Basic recording interface
- Week 2: User registration and audio upload
- Week 3: Validation workflow and admin panel
- Week 4: Testing and community feedback integration

6.2 Mobile App Development

No-Code App Solutions:

Adalo (Recommended for Beginners)

- Drag-and-drop interface builder
- Built-in audio recording components
- User management and validation workflows
- \$50/month for published app

Glide Apps

- Google Sheets backend

- Simple setup and maintenance
- Good for basic recording and validation
- \$32/month for custom domain

Custom Development Options:

React Native Framework:

- Single codebase for iOS and Android
- Good community support
- Can integrate advanced AI features
- Requires experienced developer (\$5,000-15,000)

Progressive Web App (PWA):

- Works on all devices through web browser
- No app store approval needed
- Easy updates and maintenance
- Can work offline
- Development cost: \$2,000-8,000

6.3 Integration with Existing Platforms

Social Media Integration:

- Facebook groups for community coordination
- Instagram for showcasing progress and culture
- TikTok for engaging younger participants
- WhatsApp for daily communication and reminders

Educational Platform Integration:

- Google Classroom for structured learning modules
- Khan Academy integration for gamified progress
- Duolingo course creation (requires significant content)
- Local school system integration

7. Success Metrics and Evaluation

7.1 Quantitative Metrics

Data Collection Metrics:

- Total hours of audio recorded

- Number of unique words/phrases documented
- Number of active participants per month
- Percentage of recordings that pass validation
- Average time from recording to validation

Community Engagement Metrics:

- Number of community events held
- Participation rates across age groups
- Social media engagement and reach
- Feedback survey responses and satisfaction scores

Technical Performance Metrics:

- Platform uptime and reliability
- Mobile app download and usage statistics
- User retention rates
- Average session duration and frequency

7.2 Qualitative Impact Assessment

Community Health Indicators:

- Increased pride in cultural identity
- Improved intergenerational communication
- Greater participation in cultural activities
- Enhanced community cohesion and cooperation

Language Vitality Measures:

- Number of new speakers attempting to learn
- Use of language in new contexts (social media, texting)
- Integration of preserved content into community education
- Elder satisfaction with representation of their language

Cultural Preservation Outcomes:

- Documentation of previously unrecorded cultural knowledge
- Renewed interest in traditional practices
- Connection between language and cultural activities
- Transmission of cultural values through language

7.3 Long-term Sustainability Indicators

Financial Sustainability:

- Diversified funding portfolio
- Community capacity for ongoing fundraising
- Development of revenue-generating activities
- Reduced dependence on external grants

Technical Sustainability:

- Community members trained in platform maintenance
- Backup and archival systems functioning
- Integration with permanent preservation repositories
- Scalability for content growth

Social Sustainability:

- Leadership development within community
 - Succession planning for project coordination
 - Integration with existing community institutions
 - Adaptability to changing community needs
-

8. Risk Management and Troubleshooting**8.1 Common Challenges and Solutions**

Low Participation: *Problem:* Community members reluctant to participate *Solutions:*

- Start with influential community members as early adopters
- Simplify participation requirements
- Increase incentives and recognition
- Address specific concerns through one-on-one conversations
- Partner with existing popular community activities

Technical Difficulties: *Problem:* Participants struggle with technology *Solutions:*

- Provide extensive one-on-one training
- Create simple visual instruction guides
- Establish peer support networks
- Offer multiple ways to participate (phone, in-person, etc.)
- Maintain dedicated technical support contact

Cultural Conflicts: *Problem:* Disagreements about appropriate content or methods *Solutions:*

- Establish clear cultural protocols from the beginning
- Create elder advisory committee with decision-making authority
- Regular community meetings to address concerns
- Option to participate in restricted/private modes
- Respect for traditional knowledge transmission methods

8.2 Data Security and Privacy Protection

Data Protection Measures:

- Encrypted storage for all audio files
- Secure user authentication systems
- Regular backup procedures
- Clear data retention and deletion policies
- GDPR compliance for international participants

Community Privacy Controls:

- Granular privacy settings for each recording
- Option to limit access to community members only
- Ability to withdraw consent and remove content
- Protection of children's voices with special protocols
- Anonymous participation options where culturally appropriate

8.3 Legal and Ethical Safeguards

Intellectual Property Protection:

- Clear community ownership agreements
- Creative Commons licensing with community control
- Trademark protection for language name and content
- Copyright registration for original educational materials
- Legal review of platform terms of service

Ethical Research Standards:

- Informed consent procedures in appropriate languages
- Community review board for research proposals
- Fair compensation for community expertise and time
- Transparency about all project activities and partnerships
- Regular ethical review and adjustment processes

9. Scaling and Replication Framework

9.1 Documentation for Other Communities

Replication Toolkit Components:

- Step-by-step setup guides
- Template documents and agreements
- Technology requirement specifications
- Budget templates and funding source lists
- Lesson learned reports and best practices

Community Mentorship Program:

- Experienced communities mentor new projects
- Regular video calls and knowledge sharing
- Shared resource libraries and troubleshooting guides
- Annual conference for community coordinators
- Online forum for ongoing support

9.2 Network Building and Collaboration

Inter-Community Cooperation:

- Shared technology platforms and cost savings
- Joint grant applications and funding proposals
- Exchange programs between communities
- Collaborative research and documentation projects
- Political advocacy for language rights and funding

Academic and Institutional Partnerships:

- University research collaborations
- Student internship and volunteer programs
- Professional development opportunities
- Access to advanced technology and expertise
- Publication and conference presentation opportunities

10. Quick Start Checklist

Week 1: Foundation

□ Identify 3-5 elders willing to participate □ Recruit 5-10 youth coordinators □ Hold initial community meeting □ Create simple community agreement □ Choose primary technology platform □ Set up basic equipment

Week 2: Technology Setup

□ Create accounts on chosen platform □ Test recording equipment □ Set up cloud storage system □ Create community social media presence □ Design simple recording guidelines □ Train initial coordinators

Week 3: Content Planning

□ Identify priority vocabulary (500 words) □ Plan first recording sessions □ Create validation workflow □ Set up participant recognition system □ Design community outreach materials □ Schedule regular recording sessions

Week 4: Launch and Refine

□ Hold community launch event □ Begin systematic recording □ Start validation process □ Gather feedback and adjust □ Plan next month's activities □ Apply for initial funding

Month 2: Expansion

□ Increase participant recruitment □ Add more content categories □ Implement quality improvements □ Begin educational material development □ Seek partnership opportunities □ Document lessons learned

Month 3: Sustainability Planning

□ Assess progress against goals □ Plan long-term funding strategy □ Develop community leadership □ Create replication materials □ Submit to preservation archives □ Celebrate achievements and plan next phase

Resources and Support

Essential Contacts

- Mozilla Common Voice Support: commonvoice@mozilla.com
- Endangered Languages Project: support@endangeredlanguages.com
- Living Tongues Institute: info@livingtongues.org
- Te Hiku Media (for community AI guidance): contact@tehiku.nz

Free Online Training

- Linguistic Society of America Field Methods Course
- Mozilla Community Portal Training Modules
- Google AI for Everyone Courses
- UNESCO Intangible Cultural Heritage Toolkit

Funding Databases

- Foundation Directory Online
- Grants.gov (US Federal Opportunities)
- UNESCO Culture Grants Database
- Corporate Social Responsibility Program Directories

This framework provides communities with everything needed to launch a successful crowdsourced language preservation project within 30 days, with pathways for long-term sustainability and impact.