

# Towards multilingual and intercultural libraries: natural language processing and user experience design for native languages from Peru, state of the art

Paola Ascención Morales

Rendered drawing of Raimondi stele iconography  
“El Peru”, Antonio Raimondi, 1875  
Image: Wikipedia

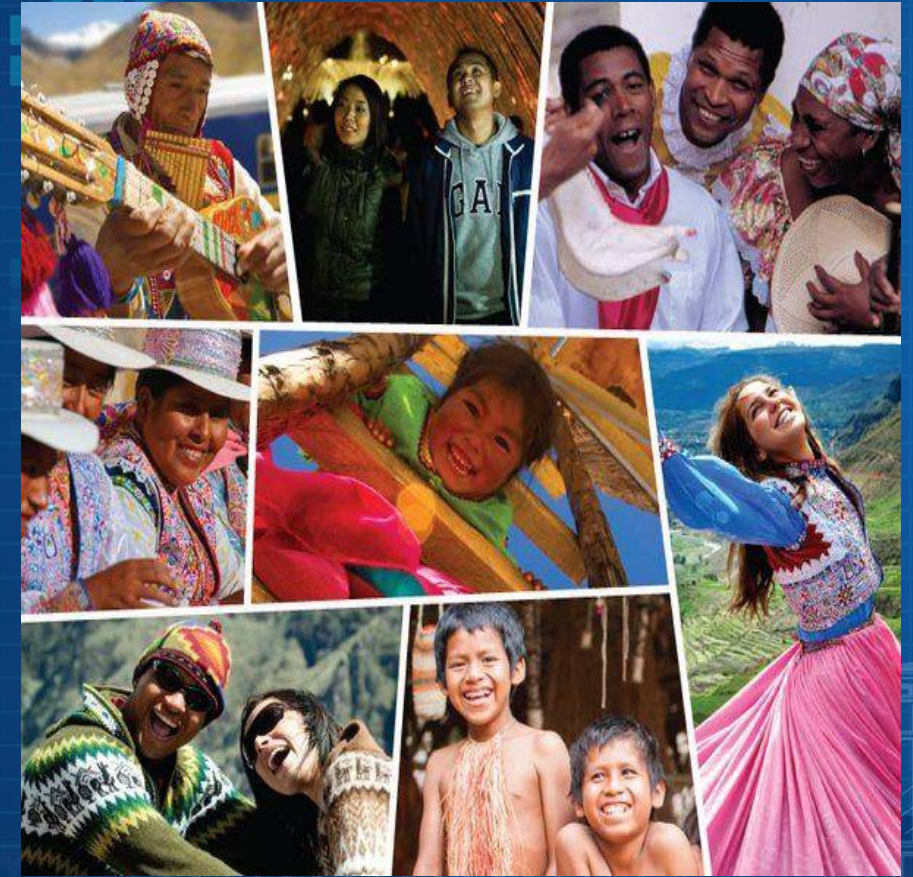


An Inca Quipu, Larco Museum, Lima. Credits image: Wikipedia

- # Quipu
- Record-keeping devices fashioned from knotted cords

# Contents

- 1) Institutional framework and public policies
- 2) Indigenous languages from Peru
- 3) Natural Language Processing
- 4) User experience design and interculturality
- 5) And... the libraries?
- 6) Future opportunities and challenges



## PERU

Multicultural and multilingual  
country

Total population:  
34.35 million  
(mid 2025 estimate)

Official languages:

Spanish 84%

Quechua 13%

Aymara 2%

...



# Institutional framework and public policies

Legal basis	Main public policies	Institutional framework
<p>Political Constitution of Peru</p> <p>Law No. 29735 (Language Law)</p> <p>Supreme Decree No. 012-2021-MC: Approves the updated National Policy to 2040, coordinating multi-sectoral actions.</p>	<p>National Policy on Native Languages, Oral Tradition, and Interculturality to 2040 (PNLOTI).</p> <p>Priority objectives:</p> <ul style="list-style-type: none"> <li>a) Multicultural relevance</li> <li>b) Reduction of discrimination</li> <li>c) Intergenerational transmission</li> <li>d) Oral and written proficiency</li> </ul>	<ul style="list-style-type: none"> <li>a) Ministry of Culture: Directorate of indigenous languages</li> <li>b) Interpretation and Translation Center (CIT)</li> <li>c) Ministry of Education (MINEDU): Intercultural Bilingual Education (IBE)</li> <li>d) Registration entities: the Presidency of the Council of Ministers (PCM) - National Registry of Native Languages.</li> </ul>

# Indigenous languages from Peru

- Languages that existed before the spread of Spanish and that are preserved and used in the national territory



# Indigenous Languages of Peru: A Comparative Overview

This infographic compares four official indigenous languages of Peru—**Aimara**, **Ashaninka**, **Awajún**, and **Quechua**—using data from the Peruvian Ministry of Culture and the 2017 National Census. It highlights the diversity in linguistic families, speaker populations, and official recognition.

## QUECHUA (Quechua)



**3.8 Million**  
Speakers

Most spoken indigenous language.

**Regional Distribution**  
Primarily Andean

**Vital Status Maintained**  
Officially categorized as "Vital"

### Linguistic Framework



**34 Characters in Alphabet**  
Largest alphabet among the four.

**205 Registered Quechua Interpreters**

\*Standardized Written Systems recognized since 1985.

## AIMARA (Aru)



**450,010**  
Speakers

Second most spoken, after Quechua.

**Regional Distribution**  
Primarily Andean

**Vital Status Maintained**  
Officially categorized as "Vital"

### Linguistic Framework



**32 Characters in Alphabet**  
Second largest alphabet among the four.

**29 Registered Aimara Interpreters**

\*Standardized Written Systems recognized since 2015.

## ASHANINKA (Arawak)



**73,567**  
Speakers

Represent Amazonian regions.

**Regional Distribution**  
Amazonian flora

**Vital Status Maintained**  
Officially categorized as "Vital"

### Linguistic Framework



**19 Characters in Alphabet**  
Smallest alphabet among the four.

**30 Registered Ashaninka Interpreters**

\*Standardized Written Systems recognized since 2015.

## AWAJÚN (Jíbaro)



**56,584**  
Speakers

Represent Amazonian regions.

**Regional Distribution**  
Amazonian flora

**Vital Status Maintained**  
Officially categorized as "Vital"

### Linguistic Framework



**21 Characters in Alphabet**  
Third largest alphabet among the four.

**24 Registered Awajún Interpreters**

\*Standardized Written Systems recognized since 2015.

\*Standardized Written Systems recognized since 1985 (Quechua) or 2015 (Aimara, Ashaninka, and Awajún).

NotebookLM

Infographic generated in NotebookLM from the data of the Database of Indigenous or Native Peoples BDPI, Ministry of Culture.

# Natural Language Processing

System or resource	Language	Tool	Reference
Spell checker (finite-state)	Quechua (Southern)	Spelling checker	Ríos, 2011.
Siminchik speak corpus	Quechua (Southern)	Speech corpus (transcribed)	Camacho, Zavallos, Cárdenas, 2019
ASR with HTK	Quechua (Southern)	Automatic voice recognition	Zevallos et al., 2019

# Natural Language Processing

System or resource	Language	Tool	Reference
QuBERT (model BERT and monolingual corpus )	Quechua (Southern)	Pre-trained language model corpus	Zevallos et al., 2022
Wav2vec2 XLS-R fine tuned ASR	Quechua (Southern)	Automatic voice recognition	Patrón Romero et al., 2024
Americas NLP shared task datasets	Quechua (Southern)	Resources and tasks of machine translation	Mager et al., 2021

# User experience design and interculturality

Beyond the linguistic process, user experience design in multicultural contexts becomes an exercise in technological sovereignty and cultural respect.

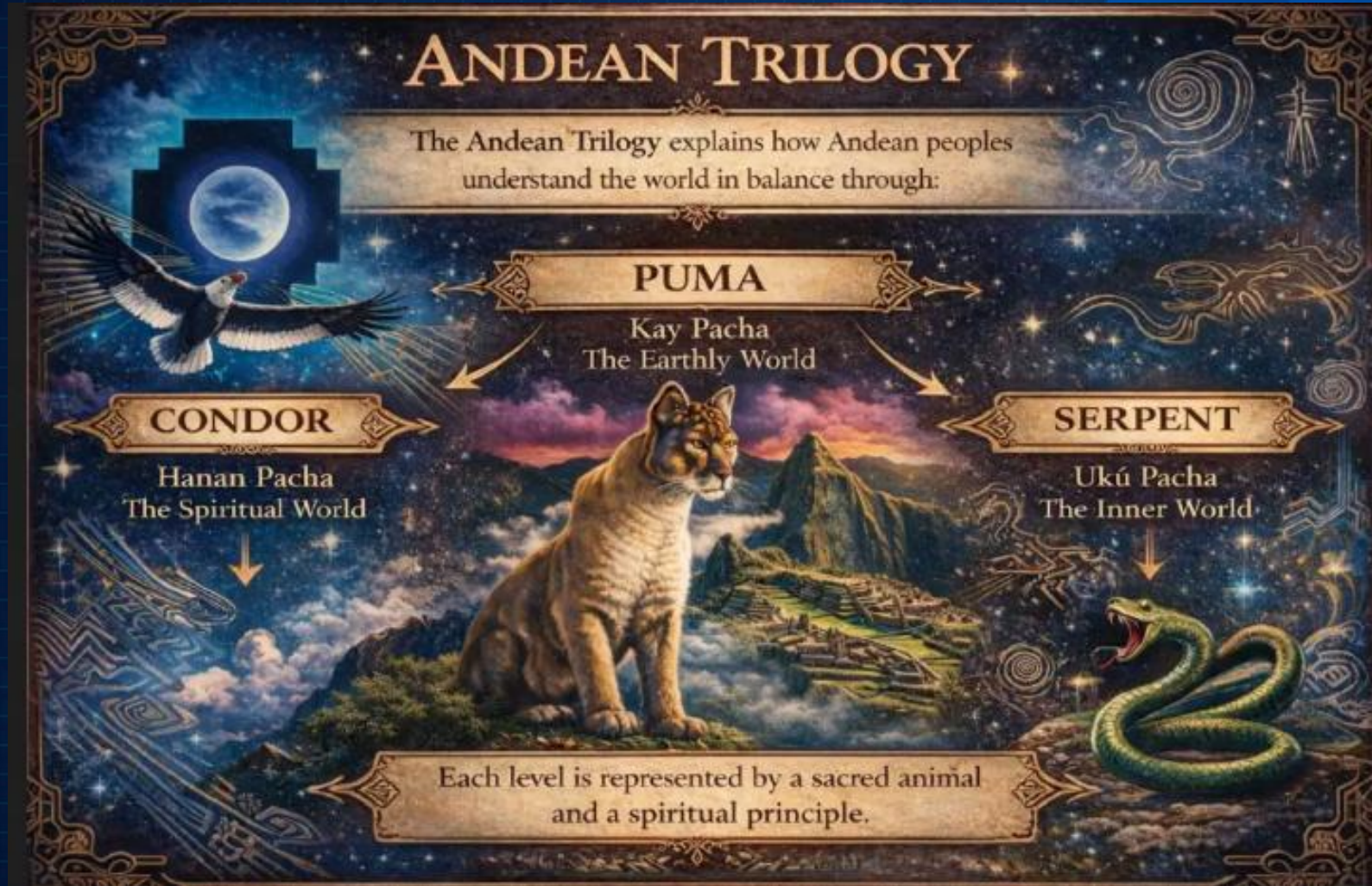
Digital inclusion—beyond technology and language—requires integrating the user's worldview to validate the digital space for speakers of indigenous languages.

# User experience design and interculturality

In participatory design, the user is not a passive recipient but a co-designer. This is indispensable in critical scenarios such as health and nutrition, where cultural relevance ensures that legal or medical concepts (informed consent) are truly understood.

To achieve socioeconomic impact, inclusive platforms demonstrate that when technology respects identity, it acts as a driver against poverty and digital isolation. In essence, the lack of cultural relevance is the primary barrier to the digital divide in Peru; without an interface that 'speaks' the user's culture, the tool is simply ignored

# Example: Andean cosmovision



Credits image: Inkayni Peru  
Tours

# Example: Andean cosmovision

Good bye, bye = Adiós

In **Quechua** there is no direct word for "good bye" because the Andean cosmovision (worldvision) focuses on reunion, to meet again

**Tupananchiskama** = Until life brings us together again

In Quechua, saying goodbye always implies the hope of a future meeting.

# Yupana    Yupay = Count

- Counting board used to perform arithmetic operations, dating back to the time of the Incas.

A quipucamayoc depicted in El primer nueva corónica y buen gobierno. A yupana, an Inca calculating device, is also visible in the lower left. Wikipedia.



# Ideas

- Technical ability vs. linguistic complexity
- Orality vs. digital textuality
- Ethics and data sovereignty

# In the library/ Bibliotecapi (Chanca Quechua)

Carnetniykita , apumurqankichu, Juana?  
Did you bring your library card, Juana?



Zariquiey & Córdova, 2008.



Ari, Ichaqa manam hurquymanchu ima  
qillqa maytutapas  
Yes, but I couldn't borrow any books.

Imasnasqa?  
Why?



Taniska kachkan. Paqarinraqmi  
kichakunqa  
t's suspended. It will be reactivated  
tomorrow.

# A library (**Quechua**) assistant test

- [Enlace](#)



# Future opportunities and challenges

- Language policy development: It is essential to establish public policy frameworks that facilitate the systematic creation of linguistic corpora and the implementation of voice technologies in public services.
- Promoting interdisciplinarity: strengthening collaboration among documentary linguists, NLP engineers, and the communities themselves is necessary to bridge existing disciplinary and cultural gaps.
- Focus on Amazonian languages: while Quechua has been the primary focus of research, there is a critical opportunity to expand studies toward less-documented Amazonian languages by employing knowledge transfer techniques.

# Future opportunities and challenges

- Language policy development: It is essential to establish public policy frameworks that facilitate the systematic creation of linguistic corpora and the implementation of voice technologies in public services.

# And...the libraries?

- Rimanakuy Comunnity Cultural Association. Quechua virtual library
- <https://comunidadrimanakuy.org/>
- Biblio Punku Aplicación. Biblioteca Nacional del Perú.
- <https://play.google.com/store/apps/details?id=pe.gob.bnp.bibliopunku&hl=es&pli=1>

BEYOND BOUNDARIES:  
New Modes of Access and Accessibility through Emerging Technologies Symposium

Thank you!

多謝你！

Gracias!

Riqsikuyki!

Paola Ascención-Morales  
paola.ascencion@unmsm.edu.pe

Organised by:

