DECOLONIALIZING AI:
NATURAL LANGUAGE PROCESSING FOR UNDER-RESOURCED LANGUAGES

The Case of Pashto

Lecture by Jandad Jahani

What is the lecture about?

D. AI While global AI development continues to privilege so-called "high-resource" languages, many languages spoken across Asia and Africa remain digitally underrepresented. This imbalance reinforces linguistic inequality and restricts participation in emerging digital infrastructures. The talk will question the role of "under-resourced" languages in the field of Artificial Intelligence, particularly within Natural Language Processing (NLP).

Using the case of Pashto, the lecture will showcase how community-driven digitalization efforts can challenge these hierarchies. Drawing on a recent women-led speech data initiative, the project demonstrates that under-resourced languages are not inherently technologically limited.
The lecture will also highlight the broader socio-political implications of such initiatives, including questions of linguistic sovereignty, access, and the ethical dimensions of digital labor.

Who is the lecturer Jandad Jahani?

D. AI **About the lecturer:**
Jandad Jahani leads and coordinates the data collection for Pashto on the Mozilla Common Voice platform.

When is the lecture?

D. AI **Date**:
Thursday, February 5, 2026 at 4:15 pm
Place:
Central Asian Seminar, room no. 507

Is participation via Zoom possible?

D. AI Yes, it is. Here is a QR code.

Typ a message ...

+ ⚙ Tools

Transregional Central Asian Studies