

TidyLang Challenge:

Speaker-Controlled Language Recognition

The TidyLang Challenge is a Speaker-Controlled and Zero-Shot Language Recognition challenge. The only permitted data from Mozilla Common Voice is the official Tidy-X training and validation partition; all other Common Voice data is strictly forbidden. The core task is spoken language recognition at the utterance level under controlled speaker-overlap conditions.

Primary evaluation conditions:

- Closed-set identification: Predict the correct language label among 40 languages seen during training (macro-averaged F1 / Accuracy).
- Open-set / zero-shot recognition: Handle additional languages not seen during training; evaluated with Equal Error Rate (EER) following standard protocols.

The challenge uses the Tidy-X dataset, a curated partition from Mozilla Common Voice featuring:

- Over 4,474 speakers across 40 languages
- Each speaker with utterances in between 2 and 10 languages
- Approximately 321,711 utterances totaling 457 hours
- Clearly defined training and validation splits
- Pseudonymized speaker identities for privacy

Evaluation set: Details about the evaluation data (including size, languages, and trial structure) are not disclosed before the evaluation phase to ensure a fair and unbiased benchmark. Information will be released when the evaluation phase opens.

More Information: <https://tidylang2026.github.io>

Timeline



Register here:

Collaborating Organizations:



Universität
Zürich UZH



OTTO VON GUERICKE
UNIVERSITÄT
MAGDEBURG



Australian
National
University

INDIANA UNIVERSITY

Mozilla
Data
Collective
Create. Curate. Control.

