

Speech Resources in the Tamasheq Language

Marcely Zanon Boito¹, Fethi Bougares², Florentin Barbier³, Souhir Gahbiche³,
Loïc Barrault², Mickael Rouvier¹, Yannick Estève¹

¹LIA - Avignon University, France ²LIUM - Le Mans University, France ³Airbus - France

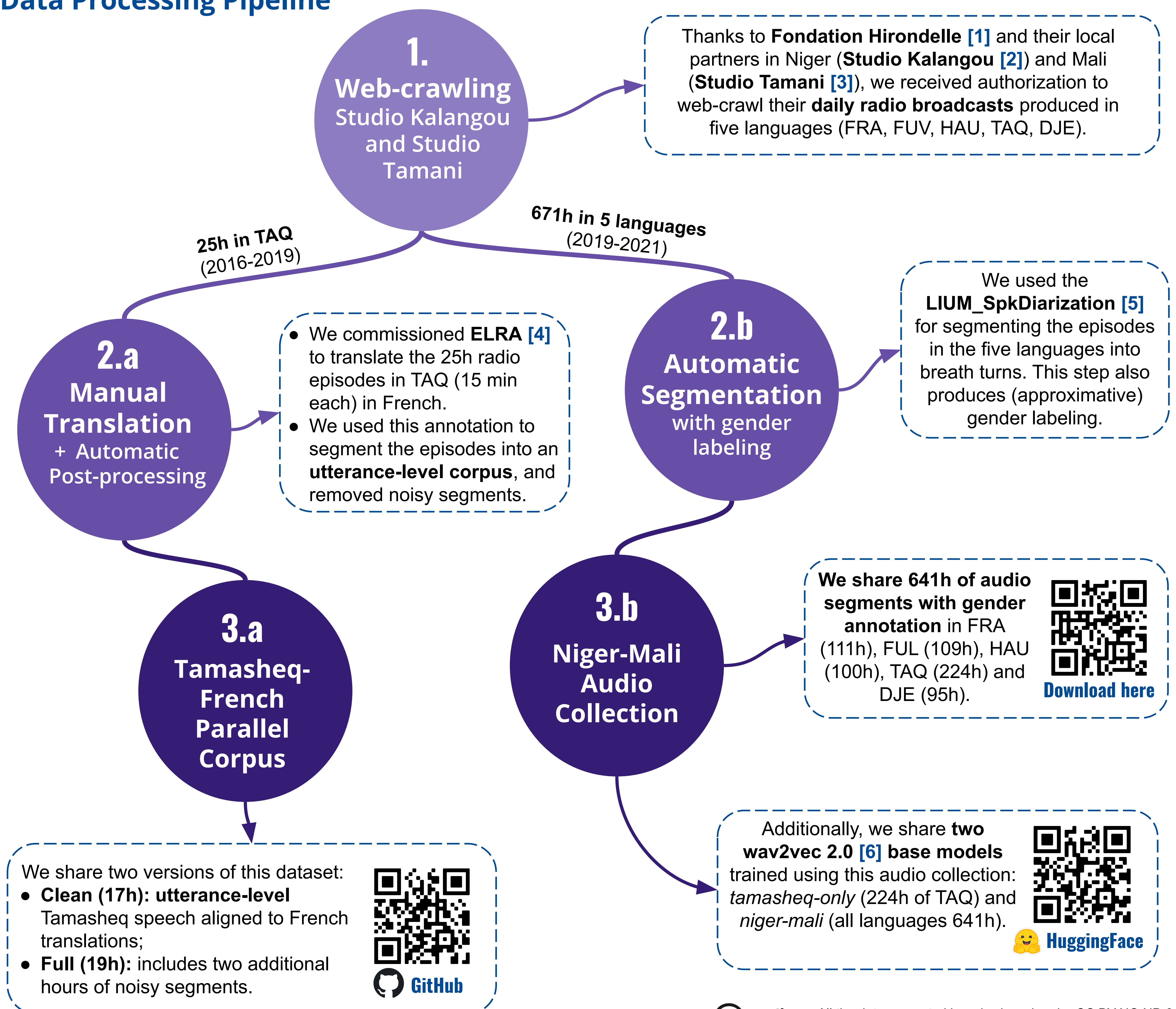
Contact: {marcely.zanon-boito, yannick.esteve} @univ-avignon.fr

In this paper we...

1. Produce a Tamasheq to French **speech translation corpus** (17h).
→ **The Tamasheq language:** ~500 thousand speakers, Tuareg variety, Berber macro-language spoken by nomadic tribes across North Africa in Algeria, Mali, Niger and Burkina Faso.
2. Distribute a 641h **audio collection** in five languages from Niger:
→ French (FRA), Fulfulde (FUV), Hausa (HAU), Tamasheq (TAQ) and Zarma (DJE).



Data Processing Pipeline



All the data presented here is shared under CC BY-NC-ND 3.0: Attribution-NonCommercial-NoDerivs 3.0 Unported