FONDAZIONE BRUNO KESSLER
MAGAZINE

The first family of open science models for speech recognition and speech translation

June 30, 2025

FBK builds the first large-scale open-science voice system for Italian and English

A speech recognition and translation system developed entirely from scratch—without relying on pre-trained models from major tech companies and built exclusively using open-source data and tools. This is the achievement of SpeechTek and **Machine Translation**, two research units at Fondazione Bruno Kessler, developed through the project **"FAMA: The First Large-Scale Open-Science Speech Foundation Model for English and Italian." The project stands out for** its innovative approach, clear vision, and significant impact, and is part of the broader efforts of the FAIR Foundation – Future Artificial Intelligence Research.

The true innovation lies not only in the model's performance but in its **complete transparency**: it was trained on over **150,000 hours of freely available audio data**, all distributed under permissive licenses. In addition to this open data, the team created a large volume of "synthetic data"—automatically generated transcriptions and translations in English and Italian—specifically for the project, and released through the MOSEL dataset.

"All the code, data, and procedures are fully public and well-documented, enabling anyone to replicate or build upon the system. *The expertise developed through this collaborative effort, along with the model's potential applications and ongoing evolution, makes FAMA a valuable asset for FBK,"*explain **Alessio Brutti**, head of the SpeechTek unit, and **Luisa Bentivogli**, head of the Machine Translation unit.

"We've demonstrated that Italy also has the capability to develop large-scale AI models," add project coordinators Sara Papi and Marco Gaido, "models that can compete globally while remaining fully compliant with the latest European regulatory frameworks."

Currently supporting Italian and English, the model lays the groundwork for a multilingual, fully open-source voice platform. The experience gained in managing data, computational loads, and

system resources provides a strong foundation for future expansion.

The training process was powered **by CINECA, which supplied the computing** infrastructure and technical support necessary to manage the large-scale operations involved.

Beyond the technical milestones, this project represents a significant step toward more **open, independent, and reproducible** artificial intelligence—paving the way for a fairer, more inclusive digital ecosystem.

Both small and medium versions of the model are now publicly available on HuggingFace.

FAMA: The First Large-Scale Open-Science Speech Foundation Model for English and Italian

In the picture:

Back from left: Mauro Cettolo | Matteo Negri | Alessio Brutti | Mohamed Nabih
Front from left: Sara Papi | Marco Gaido | Marco Matassoni | Luisa Bentivogli

Finanziato dall'Unione europea NextGenerationEU | Ministero dell'Università e della Ricerca | Italiadom PIANO NAZIONALE DI RIPRESA E RESILIENZA

PERMALINK

https://magazine.fbk.eu/en/news/the-first-family-of-open-science-models-for-speech-recognition-and-speech-translation/

TAGS

- #artificialintelligence
- #augmentedintelligence
- #dataset
- #digitalindustry
- #fair
- #fama
- #modello
- #mosel
- #open science
- #open source
- #speech recognition
- #speechtranslation
- #translation

AUTHORS

- Michela Antino