2023-06713 - Fine-grained, multimodal speech anonymization

Contract type: Fixed-term contract
Level of qualifications required: PhD or equivalent
Function: Temporary Research Position
Level of experience: From 5 to 12 years

Context
This researcher position is part of the “Personal data protection” project of PEPR Cyrannetérit, which aims to advance privacy preservation technology for various application sectors. It will be co-supervised by Emmanuel Vincent and Marc Tommasi. The hired researcher will have the opportunity to spend time in both the Multispeech and Magnet teams, to collaborate with 9 other research teams in France and with the French data protection authority CNIL, and to contribute to the project’s overall goals including the organization of an anonymization challenge.

Assignment
Large-scale collection, storage, and processing of speech data poses severe privacy threats [1]. Indeed, speech encapsulates a wealth of personal data (e.g., age and gender, ethnic origin, personality traits, health and socioeconomic status, etc.) which can be linked to the speaker's identity via metadata or via automatic speaker recognition. Speech data may also be used for voice spoofing using voice cloning software. With firm backing by privacy legislations such as the European general data protection regulation (GDPR), several initiatives are emerging to develop and evaluate privacy preservation solutions for speech technology. These include voice anonymization methods [2] which aim to conceal the speaker's voice identity without degrading the utility for downstream tasks, and speaker re-identification attacks [3] which aim to assess the resulting privacy guarantees, e.g., in the scope of the VoicePrivacy challenge series [4].


Main activities
The first objective is to improve the privacy-utility tradeoff by better disentangling speaker identity from other attributes, and better decorrelating the underlying dimensions. Solutions may rely on suitable generative or self-supervised models [5, 6] or on adversarial learning [7]. The resulting privacy guarantees will be evaluated via stronger attackers, e.g., taking metadata into account.

The second objective is to extend the proposed audio-only approach to multimodal speech (audio, facial video, and gestures). Solutions will exploit existing facial anonymization technology [8]. A key difficulty will be to preserve the correlations between modalities, which are essential for training multimodal voice processing systems. Depending on the hired researcher's skills, additional directions may also be explored, e.g., evaluating the proposed anonymization solutions in the context of federated learning.


Skills
PhD and research experience in speech processing.
Strong programming skills in Python/Pytorch.
Prior experience in speech anonymization will be an asset.

Benefits package
- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking (after 6 months of employment) and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

Remuneration
From 3724€ gross/month according to experience