



Offre n°2024-07586

## PhD Position F/M Towards a comprehensive speech anonymization framework

Type de contrat : Fixed-term contract

Niveau de diplôme exigé : Graduate degree or equivalent

Fonction : PhD Position

### Contexte et atouts du poste

This position is part of the ANR SpeechPrivacy project, which aims to advance speech anonymization technology. The PhD student will be co-supervised by [Emmanuel Vincent](#) and [Mickaël Rouvier](#). He/she will have the opportunity to spend time in both [Multispeech](#) and [LIA](#), and to contribute to the project's overall goals including the organization of an international evaluation challenge [1].

[1] <https://www.voiceprivacychallenge.org/>

### Mission confiée

Large-scale collection, storage, and processing of speech data pose severe privacy threats [2]. Indeed, speech encapsulates a wealth of personal information in the speaker's traits and the verbal and non-verbal content (e.g., speaker's identity, age and gender, ethnic origin, personality traits, health and socio-economic status, etc.). Compliance with privacy laws such as the European general data protection regulation (GDPR) calls for privacy preservation solutions tailored to speech technology. This has led to the development of voice anonymization methods [3] which aim to conceal the speaker's voice in verbal speech, and evaluation methods [1] which quantify the resulting speaker re-identification risk against linkability and singling-out attacks [4] under simplifying assumptions. This PhD aims to extend these methods into a comprehensive speech anonymization framework, which quantifies the real-world risk and accounts for non-verbal speech and other speaker traits.

[2] A. Nautsch, A. Jimenez, A. Treiber, J. Kolberg, C. Jasserand, E. Kindt, H. Delgado, M. Todisco, M. A. Hmani, M. A. Mtibaa, A. Abdelraheem, A. Abad, F. Teixeira, M. Gomez-Barrero, D. Petrovska, N. Chollet, G. Evans, T. Schneider, J.-F. Bonastre, B. Raj, I. Trancoso, and C. Busch, "Preserving privacy in speaker and speech characterisation," *Computer Speech and Language*, vol. 58, pp. 441–480, 2019.

[3] B. M. L. Srivastava, M. Maouche, M. Sahidullah, E. Vincent, A. Bellet, M. Tommasi, N. Tomashenko, X. Wang, and J. Yamagishi, "Privacy and utility of x-vector based speaker anonymization," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 30, pp. 2383–2395, 2022.

[4] [https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216\\_en.pdf](https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf)

### Principales activités

Existing average-case evaluation metrics overestimate the re-identification risk due to the unrealistic assumptions that the attacker is able to set the optimal decision threshold and knows when re-identification has succeeded. At the same time, they underestimate the risk for certain speakers or utterances which are more easily re-identifiable [5, p.105]. Therefore, a first goal is to develop reliability measures [6] enabling the attacker to decide whether re-identification has succeeded, and to study what makes certain speakers or utterances more easily re-identifiable so as to adapt the anonymization parameters accordingly without running costly speaker verification experiments.

Existing anonymization and evaluation methods also do not account for non-verbal speech (laughter, breathing, cough, scream, crying, etc.) and speaker traits (age, gender, phonation, articulation, prosody, accent, etc.). While classifiers have been developed for several of these traits [7,8], their anonymization and their use for inference attacks [4] remain to be studied. Therefore, a second goal will be to study the effectiveness of simple speech transformations [9] or specific synthesis techniques [10] for non-verbal speech anonymization, and to quantify the re-identification risk resulting from the combination of several speaker traits.

[5] B. M. L. Srivastava, "Speaker anonymization : representation, evaluation and formal guarantees," PhD thesis, University of Lille, 2021.

[6] P.-M. Bousquet, M. Rouvier, and J.-F. Bonastre, "Reliability criterion based on learning-phase entropy for speaker recognition with neural network," in *Interspeech*, pp. 281–285, 2022.

[7] Y. Lin, X. Qin, N. Jiang, G. Zhao, and M. Li, "Haha-Pod: An attempt for laughter-based non-verbal speaker verification," in 2023 IEEE Automatic Speech Recognition and Understanding Workshop, pp. 1-7, 2023.

[8] N. Obin and A. Roebel, "Similarity search of acted voices for automatic voice casting," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 24, no.9, pp. 1642-1651, 2016.

[9] J. Patino, N. Tomashenko, M. Todisco, A. Nautsch, and N. Evans, "Speaker anonymisation using the McAdams coefficient," in Interspeech, pp. 1099-1103, 2021.

[10] D. Xin, S. Takamichi, A. Morimatsu, and H. Saruwatari, "Laughter synthesis using pseudo phonetic tokens with a large-scale in-the-wild laughter corpus," in Interspeech, pp. 17-21, 2023.

## Compétences

MSc degree in speech processing, machine learning, or a related field.

Strong programming skills in Python/Pytorch.

Prior experience with speech processing is an asset.

## Avantages

- 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

## Rémunération

2100€ gross/month (the 1st year)

## Informations générales

- **Thème/Domaine** : Language, Speech and Audio
- **Ville** : Villers lès Nancy
- **Centre Inria** : [Centre Inria de l'Université de Lorraine](#)
- **Date de prise de fonction souhaitée** : 2024-10-01
- **Durée de contrat** : 3 years
- **Date limite pour postuler** : 2024-05-26

## Contacts

- **Équipe Inria** : [MULTISPEECH](#)
- **Directeur de thèse** :  
Vincent Emmanuel / [emmanuel.vincent@inria.fr](mailto:emmanuel.vincent@inria.fr)

## A propos d'Inria

Inria est l'institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2600 personnes. Ses 215 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3900 scientifiques pour relever les défis du numérique, souvent à l'interface d'autres disciplines. L'institut fait appel à de nombreux talents dans plus d'une quarantaine de métiers différents. 900 personnels d'appui à la recherche et à l'innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 200 start-up. L'institut s'efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

**Attention:** Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.

## Consignes pour postuler

### Sécurité défense :

Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

**Politique de recrutement :**

Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.