# Diego **Di Carlo**

Postdoctoral researcher
in audio processing and machine learning

*Kyoto, Japan*
*diego.dicarlo89@gmail.com*
*diegodicarlo.com*
*Chutlhu*

## Work experience

9/2022 – present
**Post-doc in Audio Signal Processing**, RIKEN AIP, Kyoto, Japan
Supervised by K. Yoshii.

7/2023 – 9/2023
**Visiting researcher**, ADASP groud (S2A, LTCI) at Télécom Paris, Paris, France
Mobility grant: KAKENHI project (JSPS).
collaboration with Mathieu Fontaine.

5/2021 – 7/2022
**Post-doc in Physics and Deep Learning**, Univ. of Rennes 2, UMR 6554, LETG, Rennes, France
Project: *CominLabs DynaLearn (2020–2024)*
Supervised by T. Corpetti and N. Courty.

12/2020 – 05/2021
**Career break**, due to COVID which slowed down administrative hiring procedures.

10/2017 – 12/2020
**Doctorare in Audio Signal Processing**, University of Rennes 1, Rennes, France
Scholarship grant: INRIA CORDI-S grant.

11/2019 – 1/2020
**Visiting PhD student**, Faculty of Engineering Bar-Ilan University, Tel Aviv, Israel
Mobility grant: Rennes Metropole.
Supervised by S. Gannot.

11/2016 – 2/2017
**Research internship**, *Multispeech team at Inria*, Nancy, France
Supervisor: A. Liutkus. – Erasmus Traineeship Exchange Program

2014 – 2016
**R&D external part-time consultant**, *Zamperla s.r.l*, Vicenza, Italy
`www.zamperla.com`.

## Education

10/2017 – 12/2020
**PhD in Audio Signal Processing**, University of Rennes 1, Rennes, France
Thesis title: *Echo-aware signal processing for audio scene analysis*
Supervised by A. Deleforge and N. Bertin.

2/2016 – 7/2016
**Master in Sound and Music Computing**, Aalborg Univ., Copenhagen, Denmark
Erasmus for Study Exchange Program.

10/2014 – 7/2017
**Master's degree in Computer Engineering**, University of Padova, Italy, grade 108/110
Thesis title: *Guassian Framework for Interference Reduction in Live Recordings.*
Supervised by A. Liutkus and N. Orio

10/2008 – 6/2014
**Bachelor's degree in Information Engineering**, *University of Padova*, grade 99/110
Thesis title: *Sequential Feature Selection: Algorithms And Applications for Audio Information Retrieval.*
Supervised by A. Rodá

## Teaching Experience

1/25
**Audio signal processing (1.5h)**, *M2 level*, Kyoto University (Eng. department), Japan
*Audio Signal Processing, Auditory Scene Analysis (Spatialization, Localization, Separation)*
1.5 hours lecture

10/2020 – 12/2020
**Module VAI "Vocal and Acoustic Interactions (10h)**, *M2 level*, Univ. Rennes 1, France
*Audio Signal Processing, Auditory Scene Analysis (Spatialization, Localization, Separation)*
6 hours lecture, 4 hours laboratory, and evaluation exam.

10/2019 – 11/2019
**Module VAI "Vocal and Acoustic Interactions (10h)**, *M2 Level*, Univ. Rennes 1, France
*Audio Signal Processing, Auditory Scene Analysis (Spatialization, Localization, Separation)*
4 hours lecture and 4 hours laboratory

The icon denotes geographical mobility (with respect to current or previous affiliation).

## Supervision activity

4/2023 – present
**Y. Fujita**, *Master M2 student*, Kyoto University, Kyoto, Japan.
Peer-reviewed contribution published at the IEEE APSIPA 2025 international conference.
Co-supervised by A. A. Nugraha, Y. Bando, M. Fontaine, K. Yoshii

5/2023 – 11/2023
**L. Kelley**, *Master M2 student*, Internship at RIKEN AIP, Kyoto, Japan.
Peer-reviewed contribution published at the INTERSPEECH 2025 international conference.
Co-supervised by A. A. Nugraha, Y. Bando, M. Fontaine, K. Yoshii

4/2023 – 4/2024
**Y. Sumura**, *Master M2 student*, Kyoto University, Kyoto, Japan.
Peer-reviewed contribution published at the IEEE IWAENC 2024 international conference.
Co-supervised by A. A. Nugraha, Y. Bando, M. Fontaine, K. Yoshii

3/2021 – 6/2021
**H. Zafar**, *Master M2 student*, University of South Brittany, Vannes, France.
Co-supervised by T. Corpetti.

## Languages

Fluent — Italian (native), English (TOEFL iBT: 82).

Beginner — French (conversational), Japanese.

## Grants and awards

6/2023 **IEEE ICASSP "Outstanding Reviewer Recognition"**
By the Organizing Committee of IEEE ICASSP 2023

3/2023 – 3/2026 **Grants-in-Aid for Scientific Research (KAKENHI) –Early-Career Scientists**
Grant No. 23K16912: *User-controllable and Physics-informed Neural Acoustics Fields for Multichannel Audio Rendering and Analysis in Mixed Reality Application*
from JSPS, the Japan Society for the Promotion of Science
Grant: 4.550.000 Yen gross (about 28.000 Euro)

10/2019 – 1/2020 **Rennes Metropoles: Mobilité internationale sortante**
from the Collége doctorale de Bretagne
Grant: 2.400 Euro

11/2018 **Main prize from Microsoft for the "best use of AI"**
Our project RAPPLE won the Abbey Road Red Hackathon
Award: 5 Xbox

10/2017 – 9/2020 **CORDI-S PhD scholarship Grant**
from INRIA, the French National Institute for Research in Computer Science and Automation
Grant: covering 3 years of salary in France

11/2016 – 2/2017 **Erasmus+ traineeships Grant**
from the University of Padova and European Commission
Grant: about 2000 Euro

1/2016 **Winner of the Oticon Audio Explorer 2016 edition**
from Oticon, Denmark.
Award: trip to New York for 3 days.

2/2015 – 7/2015 **Erasmus+ for Study Grant**
from University of Padova and European Commission
Grant: about 2000 Euro

## Comprehensive publication list

▷ Archives ouvertes and in preparation


[2025 ] D. **Di Carlo**, S. Koyama, A. A. Nugraha, M. Fontaine, Y. Bando, and K. Yoshii, "Novel view synthesis of steering vectors with physically consistent machine learning," work in preparation. Link to preview (do not distribute), 2025.

[2025 ] D. **Di Carlo**, A. A. Nugraha, K. Yoshii, and R. Scheibler, "Does it rake? analysis of data-driven beamformers in multipath conditions," work in preparation., 2025.

[2025 ] A. A. Nugraha, D. **Di Carlo**, Y. Bando, M. Fontaine, and K. Yoshii, "Sampling-rate-agnostic speech super-resolution based on gaussian process dynamical systems with deep kernel learning," work in preparation. Link to preview (do not distribute), 2025.

### ▷ International peer-reviewed journals

[2021 ] D. **Di Carlo**, P. Tandeitnik, C. Foy, N. Bertin, A. Deleforge, and S. Gannot, "Dechorate: A calibrated room impulse response dataset for echo-aware signal processing," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2021, pp. 1–15, 2021, Link to pdf ↗ SCIMAGO: Q2 - *Acoustics and Ultrasonics*. Reference JCR ↗ . DOI: 10.1186/s13636-021-00229-0.

[2021 ] C. Foy, A. Deleforge, and D. **Di Carlo**, "Mean absorption estimation from room impulse responses using virtually supervised learning," *The Journal of the Acoustical Society of America (JASA)*, vol. 150, no. 2, pp. 1286–1299, 2021, Link to pdf ↗ SCIMAGO: Q1 - *Acoustics and Ultrasonics*. Reference JCR ↗ . DOI: 10.1121/10.0005888.

[2019 ] A. Deleforge, D. **Di Carlo**, M. Strauss, R. Serizel, and L. Marcenaro, "Audio-based search and rescue with a drone: Highlights from the IEEE signal processing cup 2019 student competition," *IEEE Signal Processing Magazine*, vol. 36, no. 5, pp. 138–144, 2019, Link to pdf ↗ ; SCIMAGO: Q1 - *Applied Math*. Reference JCR ↗ . DOI: 10.1109/MSP.2019.2924687.

### ▷ International peer-review conference articles

[2024 ] Y. Fujita, A. A. Nugraha, D. **Di Carlo**, Y. Bando, M. Fontaine, and K. Yoshii, "Run-time adaptation of neural beamforming for robust speech dereverberation and denoising," in *2022 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, Link to pdf ↗ , IEEE, 2024.

[2024 ] Y. Sumura, A. A. Nugraha, D. **Di Carlo**, Y. Bando, M. Fontaine, and K. Yoshii, "Joint audio source localization and separation with distributed microphone arrays based on spatially-regularized multichannel NMF," in *International Workshop on Acoustic Signal Enhancement (IWAENC)*, IEEE, 2024.

[2024 ] L. Kelley, D. **Di Carlo**, A. A. Nugraha, M. Fontaine, Y. Bando, and K. Yoshii, "RIR-in-a-Box: Estimating room acoustics from 3D mesh data through shoebox approximation," in *INTERSPEECH*, Link to pdf ↗ , 2024.

[2024 ] D. **Di Carlo**, A. A. Nugraha, M. Fontaine, Y. Bando, and K. Yoshii, "Neural steerer: Novel steering vector synthesis with a causal neural field over frequency and direction," in *2024 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW)*, Link to pdf ↗ , 2024, pp. 740–744.

[2024 ] P. Naylor, D. **Di Carlo**, A. Traviglia, M. Yamada, and M. Fiorucci, "Implicit neural representation for change detection," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, Link to pdf ↗ , 2024, pp. 935–945.

[2023 ] A. A. Nugraha, D. **Di Carlo**, Y. Bando, M. Fontaine, and K. Yoshii, "Time-domain audio source separation based on Gaussian processes with deep kernel learning," in *2023 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, Link to pdf ↗ , IEEE, 2023, pp. 1–5.

[2022 ] M. Fontaine, D. **Di Carlo**, K. Sekiguchi, A. A. Nugraha, Y. Bando, and K. Yoshii, "Elliptically contoured alpha-stable representation for MUSIC-based sound source localization," in *2022 30th European Signal Processing Conference (EUSIPCO)*, Link to pdf ↗ , IEEE, 2022, pp. 26–30.

[2020 ] D. **Di Carlo**, C. Elvira, A. Deleforge, N. Bertin, and R. Gribonval, "Blaster: An off-grid method for blind and regularized acoustic echoes retrieval," in *2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Link to pdf ↗ , IEEE, 2020, pp. 156–160.

[2019 ] D. **Di Carlo**, A. Deleforge, and N. Bertin, "Mirage: 2D source localization using microphone pair augmentation with echoes," in *2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Link to pdf ↗ , IEEE, 2019, pp. 775–779.

[2018 ] D. **Di Carlo**, A. Liutkus, and K. Déguemel, "Interference reduction on full-length live recordings," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2018, pp. 736–740.

[2018 ] R. Scheibler, D. **Di Carlo**, A. Deleforge, and I. Dokmanic, "Separake: Source separation with a little help from echoes," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Link to pdf ↗ , IEEE, 2018, pp. 6897–6901.

[2017 ] D. **Di Carlo**, K. Déguernel, and A. Liutkus, "Gaussian framework for interference reduction in live recordings," in *Journal of the audio engineering society: 2017 AES International Conference on Semantic Audio*, Link to pdf ↗ , Audio Engineering Society, 2017.

[2016 ] F. Grani *et al.*, "Gestural control of wave field synthesis," in *Proceedings of 13th Sound and Music Computing Conference, Hamburg*, Link to pdf ↗ , 2016.

▷ International non-peer-review conference articles:

[2022 ] D. **Di Carlo**, D. Heitz, and T. Corpetti, "Post processing sparse and instantaneous 2D velocity fields using physics-informed neural networks," in *20th International Symposium on Application of Laser and Imaging Techniques to Fluid Mechanics (LXLASER)"*, Link to pdf , 2022.

[2018 ] R. Lebarbenchon, E. Camberlein, D. **Di Carlo**, C. Gaultier, A. Deleforge, and N. Bertin, "Evaluation of an open-source implementation of the SRP-PHAT algorithm within the 2018 LOCATA challenge," *IEEE-AASP Challenge on Acoustic Source Localization and Tracking (LOCATA)*, 2018, Link to pdf .

## Collective tasks at research community level

### Tutorials

- (tentative) IEEE EUSIPCO 2025 tutorial – submitted, preview available here .
  *Physics-Informed Machine Learning For Audio Processing* (3 hours)
  Co-organized with M. Pizzoli (Politecnico di Milano, Italy) S. Koyama (NII, Japan).

### Reviewer

- Conferences IEEE ICASSP, IEEE WASPAA, IEEE EUSIPCO, INTERSPEECH, and ISMIR.
- Journals EURASIP Journal on Audio, Speech, and Music Processing (JASM) and IEEE Signal Processing Letters (SPL).

### Challenges

- Co-organizer of the IEEE SPCup 2019 Challenge on *Search & Rescue with Drone-Embedded Sound Source Localization*, IEEE Signal Processing Magazine 36 (5), 2019.
  Co-organized with A. Deleforge, M. Strauss, R. Serizel, L. Marcenaro
- Participation at the IEEE AASP Challenge on *Acoustic Source Localization and Tracking (LOCATA)*, in proc. of the LOCATA Challenge Workshop - a satellite event of IWAENC 2018

### Scientific dissemination

- *From Neural Fields to PINNs and beyond* .
  at PRISM CNRS, (Paris, France, 2024); S2A TelecomParis (Paris, France, 2024); National Institute of Informatics (Tokyo, Japan, 2024).
- *Physics-informed Neural Network* .
  At CERAMA, (Strasbourg, France, 2022); CAp&RFIAP 2022 (Vannes, France, 2022).
- *How to kickstart your (Python) project*
  At LETG UMR 6554 (Rennes, France, 2021); Seminaire Au Vert – team building (Rennes, France, 2020); Bar-Ilan University (Tel-Aviv, Israel, 2019).
- *Echo-aware Signal Processing for Audio Scene Analysis* .
  CentraleSupelec (Rennes, France, 2020); Bar-Ilan University (Tel-Aviv, Israel, 2019), workshop "Parcimonie et acquisition compressée"a at (Roskoff, France, 2019).

### Vulgarization

- Sound designer and screenwriter of the video *"The Science Behind Music Technology" (2021)*

- *Sound to Data – aller et retour* (IGDR, Rennes, France, 2021).
  Seminar about music signal processing for audience in medicine.

### Softwares and datasets

- **code** `AV-SUARA`: General-purpose automated captioning system for the Hololens 2 in ROS.
  Code is private due to licensing, but it was presented at INTERSPEECH 2023 tutorial
- **code** `Pyroomacoustics`: actively contributing to an open-source coding library for audio signal processing.

- **code** `SPCup2019` : Baseline and evaluation code for the IEEE SPCup Challenge (2019)
- **code** `RAPPLE` : AI-powered Rap Battle Opponent, which listens to your freestyle on a beat and responds to you with a punch line to the same line.
  Winner of the Microsoft Main prize for "best use of AI" at Abbey Read Red Hackathon (2018).
- **code** spAAce : graphical application to control real-time movements of sound sources by drawing trajectories.
- **dataset** `dEchorate` : a calibrated dataset for echo-aware audio signal processing.
- **awesome-list** `awesome-sound-field` : a public curated list of papers and resources for Sound Field Reconstruction and HRTF upsampling

## Administrative and collectives task at team level

2024 **Kyoto University Gasshuku (coding camp)**

At Kyoto University, Kyoto, Japan.

- Context — Organization of "coding camp" for Master students of the Speech and Audio Processing Laboratory group of Kyoto University. The project consisted of implementing the complete pipeline of a real-time spoken dialog system.
- Involvement — Co-organization the event and assisting in the design of the system pipeline.

2022 - now **AV-SUARA Project (team project)**

at RIKEN AIP, Kyoto, Japan.

- Context — AV-SUARA is a general-purpose automated captioning system designed for the HoloLens 2, aimed at improving accessibility and human-computer interaction through audio and visual cues. This project is associated with the KAKENHI project of my colleague A. A. Nugraha (KAKENHI project 23K16912 ).
- Involvement — Improvement of the system's front-end acoustic processing by continuously integrating my research achievement.

2019 **Haru Robot Project with Honda Research Institute (socio-economic transfer)**

at Panama team, INRIA Rennes, Rennes, France.

- Context — The Haru robot, developed by Honda Research Institute, is designed to improve human well-being by using social and interactive features powered by advanced audio-visual technologies.
- Involvement — Collaboration with A. Deleforge in studying the feasibility of a novel sound source localization algorithm tailored microphone array installed on the robot.

2019 **"Remi and the Panama Papers" (coding challenge)**

at Panama team, INRIA Rennes, Rennes, France.

- Context — This coding challenge was created as part of the team-building event "Seminar Au Vert" of the PANAMA team (INRIA Rennes).
- Involvement — Co-organization of the challenge, and definition of the coding objectives.

2017 – 2018 **Journée Science et Musique event (dissemination event)**

at Panama team, INRIA Rennes, Rennes, France.

- Context — The Journée Science et Musique (Rennes, France) event was an annual event organized by the PANAMA team (INRIA Rennes) for scientists to engage with the public through audio- and music-related demonstrations.
- Involvement — Co-organization of the edition 2018 of the event, participation as a volunteer in the 2017 , supporting the technical and logistical aspects of the event.