Cláudia Brito

claudia.v.brito@inesctec.pt claudiavmbrito.github.io 0000-0003-4293-9887
ertclDsAAAAJ Rua da Universidade, Gualtar (4710-057), Braga, Portugal 253604477

Short Bio

I am an assistant researcher at INESC TEC and an invited assistant professor at the University of Minho. My current research rests on distributed machine learning, addressing several fronts. In detail, in the last few years, I have been delving into privacy-preserving distributed machine learning and federated learning, the application of distributed machine learning to healthcare challenges, and storage optimizations for machine learning and distributed machine learning. More recently, I have expanded my research lines towards energy-aware storage systems for deep learning workloads. I am starting to grow a publication record on these subjects, having already some relevant papers in reputable journals, conferences, and workshops (e.g., IEEE Access, ACM SAC, EPIA). I have been an active researcher on multiple international and national projects, such as Compete2020 BigHPC and AIDA, PT-UTAustin PAStor, PT2030 BCD.S+M, International consortiums (CENTRA) and European projects such as Green.Dat.AI. Currently, I am the co-PI of the PT2030 BringTrust project, with a total budget of 1 654 244€.

Experience

Aug 2024 - Curr **INESC TEC,** *Assistant Researcher*.

- Research on the intersection of AI and Systems: systems optimizations (storage and energy) for AI workloads, AI for optimizing and configuring systems, and distributed AI applied to the healthcare use-case. Leading the two latter topics under the distributed storage research group (`https://dsr-haslab.github.io/`).
- Actively involved in writing funding proposals and securing project grants from EU and national sources.
- Scientific peer-reviewed publications in high-impact journals and conferences (IEEE, ACM, Elsevier).
- Student supervision (MSc) and teaching activities.
- **Technologies | Skills:** HPC, Traditional DL and LLMs, Distributed and Federated Machine Learning, PyTorch, DeepSpeed, NVIDIA-smi, NVML, Intel RAPL.

Oct 2018 - 2024 **INESC TEC,** *Research Assistant*.

- Development of a privacy-preserving distributed machine learning system.
- Benchmark and evaluation of the privacy-preserving distributed ML system developed.
- **Technologies | Skills:** Apache Spark, Intel SGX, Cryptographic Primitives, Machine Learning.

Jul 2021 - Dec 2021 **Ubiwhere.** *Data Researcher* - Internship

- Development of a privacy-preserving federated system for urban mobility.
- Development of ML algorithms for urban transportation modalities.
- Benchmark of the privacy-preserving federated system for urban mobility developed.
- **Technologies | Skills:** Differential Privacy, Flower, Deep Learning.

Education

2018 – 2024 **Ph.D. in Informatics,** University of Minho.
Thesis title: *Towards a Privacy-Preserving Distributed Machine Learning Framework*
Final grade: Very Good.
Advisors: *João Tiago Paulo* and *Pedro Gabriel Ferreira*

2013 – 2018 **Integrated Masters in Biomedical Engineering,** University of Minho.
Thesis title: *Cloud-based analytics for monitoring and classification of arrhythmias*.
Final grade: 18/20.
Advisors: *António Luís Sousa*

Languages

English	Professional proficiency (C2)
Portuguese	Native proficiency
Spanish	Basic proficiency (B1)
French	Basic proficiency (B1)
Italian	Initial proficiency (A1)

F1 - Scientific and Technological Production

F1.1 - Research Lines

My research is currently divided into three main topics.

Adaptive and Intelligent Systems for AI. This topic addresses the design of adaptive infrastructures for AI workloads across heterogeneous environments, including HPC clusters, datacenters, and edge systems. The primary goal is to develop self-aware storage hierarchies, adaptive resource schedulers, and fault-tolerant mechanisms that dynamically adjust to workload characteristics. These solutions aim to improve energy efficiency, I/O performance, and resilience while ensuring scalability on modern and emerging architectures.

AI for Systems Optimization. This research line explores how machine learning can optimize systems themselves, not only in HPC contexts but also in large-scale distributed storage and compute platforms. Applications include energy-aware job scheduling, resource allocation, performance prediction, and proactive fault tolerance. The long-term objective is to enable self-optimizing computing infrastructures that can adapt autonomously to evolving workloads and operational constraints.

AI-Driven Security and Adaptive Protection in Systems. This line focuses on leveraging AI techniques to enhance security and resilience in complex computing infrastructures. It includes AI-based ransomware and anomaly detection, behavioral analysis of system components, and adaptive security strategies that dynamically reconfigure based on workload characteristics and the sensitivity of stored data. The goal is to build intelligent, self-defending systems that prevent and mitigate security threats while maintaining performance and availability.

F1.2 - Publications

This section describes the complete list of publications in journals, conference proceedings, and workshop proceedings, all of which are peer-reviewed articles. Then, other publications are identified, such as theses, posters, communications, and articles in non-peer-reviewed archives.

Journal Articles

1 L. Brito, B. Cepa, **C. Brito**, Â. Leite, and M. G. Pereira, "Risk assessment profiles for caregiver burden in family caregivers of persons living with alzheimer's disease: An exploratory study with machine learning," *European Journal of Investigation in Health, Psychology and Education*, vol. 15, no. 3, 2025, ISSN: 2254-9625. DOI: 10.3390/ejihpe15030041.

2 **C. Brito**, P. Ferreira, and J. Paulo, "Exploiting trusted execution environments and distributed computation for genomic association tests," *IEEE Journal of Biomedical and Health Informatics*, pp. 1–8, 2025. DOI: 10.1109/JBHI.2025.3562364.

3 **C. Brito**, N. Pina, T. Esteves, R. Vitorino, I. Cunha, and J. Paulo, "Promoting sustainable and personalized travel behaviors while preserving data privacy," *Transportation Engineering*, vol. 19, p. 100 237, 2025. DOI: 10.1016/j.treng.2024.100237.

4 **C. Brito**, P. G. Ferreira, B. L. Portela, R. C. Oliveira, and J. T. Paulo, "Privacy-preserving machine learning on apache spark," *IEEE Access*, vol. 11, pp. 127 907–127 930, 2023. DOI: 10.1109/ACCESS.2023.3332222.

5 N. Pina, **C. Brito**, R. Vitorino, and I. Cunha, "Promoting sustainable and personalised travel behaviours while preserving data privacy," *Transportation Research Procedia*, vol. 72, pp. 2768–2775, 2023. DOI: 10.1016/j.trpro.2023.11.819.

6 **C. Brito**, M. Esteves, H. Peixoto, A. Abelha, and J. Machado, "A data mining approach to classify serum creatinine values in patients undergoing continuous ambulatory peritoneal dialysis," *Wireless Networks*, pp. 1–9, 2022. DOI: 10.1007/s11276-018-01905-4.

Conference Proceedings

1 I. Chrysakis, E. Agorogiannis, N. Tsampanaki, *et al.*, "Multi-partner project: Green. dat. ai: A data spaces architecture for enhancing green ai services," in *2025 Design, Automation & Test in Europe Conference (DATE)*, IEEE, 2025, pp. 1–7.

2 **C. Brito**, P. Ferreira, B. Portela, R. Oliveira, and J. Paulo, "Soteria: Preserving privacy in distributed machine learning," in *Proceedings of the 38th ACM/SIGAPP Symposium on Applied Computing*, 2023, pp. 135–142. DOI: 10.1145/3555776.3578591.

3 B. Cepa, **C. Brito**, and A. Sousa, "Generative adversarial networks in healthcare: A case study on mri image generation," in *2023 IEEE 7th Portuguese Meeting on Bioengineering (ENBENG)*, IEEE, 2023, pp. 48–51. DOI: 10.1109/ENBENG58165.2023.10175330.

4 J. Alves, B. Soares, **C. Brito**, and A. Sousa, "Cloud-based privacy-preserving medical imaging system using machine learning tools," in *EPIA Conference on Artificial Intelligence*, Springer, 2022, pp. 195–206. DOI: 10.1007/978-3-031-16474-3_17.

5 R. Macedo, C. Correia, M. Dantas, *et al.*, "The case for storage optimization decoupling in deep learning frameworks," in *REX-IO Workshop at 2021 IEEE International Conference on Cluster Computing (CLUSTER)*, IEEE, 2021, pp. 649–656. DOI: 10.1109/Cluster48925.2021.00096.

6 **C. Brito**, A. Machado, and A. L. Sousa, "Electrocardiogram beat-classification based on a resnet network," in *MedInfo*, 2019, pp. 55–59. DOI: 10.3233/SHTI190182.

7 J. Rei, **C. Brito**, and A. Sousa, "Assessment of an iot platform for data collection and analysis for medical sensors," in *2018 IEEE 4th International Conference on Collaboration and Internet Computing (CIC)*, IEEE, 2018, pp. 405–411. DOI: 10.1109/CIC.2018.00061.

8 C. Peixoto, **C. Brito**, M. Fontainhas, A. Abelha, H. Peixoto, and J. Machado, "Continuous ambulatory peritoneal dialysis: Business intelligence applied to patient monitoring: Capd study and statistics," in *2017 5th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW)*, IEEE, 2017, pp. 178–185. DOI: 10.1109/ficloudw.2017.91.

Preprints

1 **C. Brito**, P. G. Ferreira, and J. Paulo, "A distributed computing solution for privacy-preserving genome-wide association studies," 2024. DOI: 10.1101/2024.01.15.575678.

2 B. Cepa, **C. Brito**, and A. Sousa, "To fid or not to fid: Applying gans for mri image generation in hpc," 2024. DOI: 10.1101/2024.09.27.615343.

3 A. Oliveira, B. Cepa, **C. Brito**, and A. Sousa, "Mastering artifact correction in neuroimaging analysis: A retrospective approach," 2024. DOI: 10.1101/2024.08.02.606374.

4 N. Pina, **C. Brito**, R. Vitorino, and I. Cunha, "Promoting sustainable and personalised travel behaviours while preserving data privacy," Available at SSRN: https://ssrn.com/abstract=4564774, 2022.

5 **C. Brito**, P. Ferreira, B. Portela, R. Oliveira, and J. Paulo, "Soteria: Privacy-preserving machine learning for apache spark," Available at IACR: https://eprint.iacr.org/2021/966, 2021. DOI: 10.1145/3555776.3578591.

Thesis

1 **C. Brito**, "Towards a privacy-preserving distributed machine learning framework," PhD Thesis, University of Minho, 2024.

2 **C. Brito**, "Cloud-based analytics for monitoring and classification of arrhythmias," MSc Thesis, University of Minho, 2018.

Posters

1 **C. Brito**, N. Pina, R. Vitorino, I. Cunha, and J. Paulo, *Emission-aware federated learning: A case study on transportation and carbon footprint*, EuroSys 2023, Poster Session, 2023.

2 **C. Brito**, *Let's go, private! towards a privacy-preserving and distributed machine learning system*, EuroSys 2022, Poster Session, 2022.

F1.3 - Prototypes

Open Science Prototypes. The following prototypes represent the foundation of my research work, developed within the scope of various projects and publications. These tools are openly available to the community, promoting transparency, reproducibility, and collaboration in scientific research. This goes in line with the principles of open science, which INESC TEC has been strongly advocating for.

- **EnerFrame.** A framework with an UI that allows the energy consumption of deep learning models to be estimated offering statistical information regarding the energy consumption of the models. Work developed by Adriano Novo for GreenData.AI while leveraging the energy profiler developed by José Fernandes, Sara Pereira, Mariana Amorim and Diana Rodrigues (MScs). Available at: https://enerframe-greendatai.haslab-dataspace.pt/. 2024.
- **PowMon.** An energy monitoring tool for understanding the energy patterns of concurrent applications. It enables fine-grained energy measurements, including thread-, process-, application-, and system-level, and correlates energy consumption patterns with application specific events (e.g., compactions, replication). Work originally developed by Diana Rodrigues, José Fernandes, Mariana Amorim, and Sara Pereira. Currently led by Alexandre Fernandes. Available at https://github.com/dsrhaslab/powmon. PowMon was recognized in the EU's Innovation Radar (https://innovation-radar.ec.europa.eu/innovation/59285), as part of the above-mentioned testing framework for AI energy efficiency (**EnerFrame**) developed by INESC TEC. 2024.
- **MOANA.** A framework developed to improve the accuracy and reliability of Deep Learning models by addressing the presence of artifacts in medical imaging data in a collaborative environment. This work, built by leveraging NVIDIA NVFlare, MOANA is designed to be deployed on High Performance Computing (HPC) systems. This work was developed by Alícia Oliveira (MSc). Available at: https://github.com/dmlrhaslab/MOANA. 2024.
- **Gyosa.** a privacy-preserving distributed GWAS framework that leverages Soteria to perform GWAS in a privacy-preserving and scalable way. Gyosa redefines the GWAS pipeline to be executed in a distributed fashion by resorting to Glow while ensuring that sensitive data is not leaked during the execution of the pipeline. Work developed in my PhD. Available at https://github.com/claudiavmbrito/gyosa. 2023.
- **MedGAN.** A Generative Adversarial Network (GAN) that generates synthetic medical images to be used in the training of deep learning models. MedGAN was developed by Beatriz Cepa (MSc) and is available at: https://github.com/dmlrhaslab/GAN-MRI-image-generation and https://github.com/dmlrhaslab/Distributed-GAN. 2023.
- **MedCloudCare (MCC)**. A cloud-based medical imaging system that leverages machine learning tools to provide a more efficient and accurate diagnosis. MCC allows researchers and medical staff to share and analyze medical images while ensuring the privacy of the patients' data by encrypting data at-rest and in-use as researchers may not have permission to access sensitive data in plaintext. This work was developed by João Alves and Beatriz Soares (MScs). Available at https://github.com/dmlrhaslab/MedCloudCare. 2023.
- **Soteria.** A privacy-preserving distributed machine learning framework that leverages Apache Spark as its underlying computing mechanism and MLlib API as the ML engine. Soteria introduces a hybrid scheme, combining computation done inside and outside TEE's enclaves. Work developed in my PhD. Available at https://github.com/claudiavmbrito/soteria. 2022.
- **PRISMA.** System storage that includes optimizations in data access for deep learning applications. The proposed optimizations allow to understand which data will be accessed and provide it more quickly to the applications, thus improving the training time for deep learning models. Work developed by Cláudia Correia (master's). Available at https://github.com/dsrhaslab/prisma. 2020.

F2 - Applied or practice-based research activities

Next, I highlight different indicators associated with the visibility and recognition of my scientific work. The following sections depict the international collaborations, awards and distinctions related directly with my research, the programme and artifact evaluation committees that I have been, or I am part of, the journal reviewer activities by invitation, the talks and webinars to showcase my work or my research group work, the volunteer activities on conference organization. In brief, this section highlights my research activities with practical impact and visibility, showcasing their contributions to real-world applications and the broader scientific community.

International Collaborations

Here, I highlight some of the external collaborations that I have been involved in, which have been developed in the context of the projects and publications in which I have been involved.

2025 – cur. **Yuta Seino, Nozaki Kazunori, Le Van An.** Federated Learning for Disaster Victim Identification. Collaboration within the CENTRA collaborative network, specifically in the CENTRA-ESDM project. The goal of this collaboration, for INESC TEC side, is to integrate privacy-preserving measures to a federated learning system that allows the training of machine learning models for disaster victim identification.
Le Van An, Truong Nguyen, Jason Haga. Optimizations under Federated Learning settings. Collaboration within the CENTRA collaborative network, specifically in the CENTRA-ESDM project. The goal of this collaboration, for INESC TEC side, is to develop optimizations for federated learning systems that allow the training of machine learning models.

2024 – cur. **Vera Rimmer and Lieven Desmet.** Ransomware Dataset Generator Pipeline for Machine Learning. Collaboration with the supervision team of MSc student Bruno Pereira (Tânia Esteves and João Paulo). INESC TEC and KU Leuven.
Rahma Nouaji. Optimization of Large Language Models. Collaboration with the supervision team, Ricardo Macedo (INESC TEC) and Oana Balmau (McGill University, Canada), under the Visiting Researcher Programme of INESC TEC, from May to July 2024.

2023 – cur. **Raju Rangaswami, Janki Bhimani, Yangzhao Wu.** Storage and Checkpointing Optimizations for Deep Learning Workloads. Collaboration alongside Ricardo Macedo and João Paulo, INESC TEC, and Florida Internation University (FIU), USA.

2018 – cur. **Jason Haga and Yusuke Tanimura.** Storage optimizations for AI workloads. Collaboration within the CENTRA collaborative network, specifically in the CENTRA-ESDM and CENTRA-PaSM projects between INESC TEC and AIST, Japan.

Awards and Distinctions

2024 **European Commission Recognition** – AI energy efficiency framework from GreenDat.AI listed in **EU Innovation Radar** for a breakthrough in AI sustainability.

2023 **Travel Grant, Eurosys'23.** Awarded by *Encontro Nacional de Sistemas Distribuídos (ENSD)*, Portugal.

2019 **Doctoral Scholarship (SFRH/BD/146528/2019).** Awarded by the Foundation for Science and Technology (FCT), Portugal

Programme and Artifact Evaluation Committee

2026 *The 32nd International Conference on High-Performance Computer Architecture.* HPCA'26 (Technical Program Committee).

2025 *The 31st Symposium on Operating Systems Principles.* SOSP 2025 (Artifact Evaluation Committee).
Workshop on Reliable Large-scale Data Management. W-RLDM, SRDS'25 (Technical Program Committee).
The 19th USENIX Symposium on Operating Systems Design and Implementation. OSDI'25 (Artifact Evaluation Committee).
The 5th International Conference on Electronic and Electrical Engineering and Intelligent Systems. ICE3IS'25 (Technical Program Committee).
The 20th ACM European Conference on Computer Systems. EuroSys'25 (Artifact Evaluation Committee and Shadow PC).
The 24th International Conference on Distributed Applications and Interoperable Systems. DAIS'25 (Artifact Evaluation Committee).

2024 *The 4th International Conference on Electronic and Electrical Engineering and Intelligent Systems.* ICE3IS'24 (Technical Program Committee).

2023 *The Sixth Annual Conference on Machine Learning and Systems.* MLSys'23 (Artifact Evaluation Committee).
The 23rd International Conference on Distributed Applications and Interoperable Systems. DAIS'23 (Artifact Evaluation Committee).

Reviewer

The following list does not include the articles reviewed as a conference program committee member, which have been presented previously.

2025 *Progress in artificial intelligence.* Springer Nature.

2024 *Transactions on Dependable and Secure Computing.* IEEE
Computers in biology and medicine. Elsevier.
NeuroImage. Elsevier.
Science of computer programming. Elsevier.
Sustainable Computing: Informatics and Systems. Elsevier.

2021 *Progress in artificial intelligence.* Springer Nature.

Talks and Webinars

2025 **Efficient and Secure Data Management for HPC and Cloud Computing.** *In CENTRA 8 event.* Hsinchu, Taiwan.

2024 **Distributed Systems Research.** *In Cloud Computing Services and Applications, MEI MSc Course in Informatics Engineering, University of Minho.* Braga, Portugal.
Distributed Storage Research. *Talk given during a faculty research visit at Florida International University (FIU).* Miami, USA.
Preservando a Privacidade da Computação e Outras Coisas. *In Women in Engineering Celebration, Agrupamento de Escolas de Ponte da Barca.* Ponte da Barca, Portugal.

2023 **Privacy-Preserving and Distributed Machine Learning.** *In Women in Engineering Celebration, Agrupamento de Escolas de Ponte da Barca.* Ponte da Barca, Portugal.

2022 **Privacy-Preserving Machine Learning for Apache Spark.** *In Encontro Nacional de Sistemas Distribuídos.* Évora, Portugal.
Protecting the Security of the AIDA Platform and the Privacy of its Data. *In a* Webinar *of AIDA Project.* Online.
Let's Go, Private! Towards a Privacy-Preserving and Distributed Machine Learning System. *In EuroSys Doctoral Workshop.* Rennes, France.
Privacy-Preserving and Distributed Machine Learning. *In Cloud Computing Services and Applications, MEI MSc Course in Informatics Engineering, University of Minho.* Braga, Portugal.

2021 **Introduction to Private and Secure Machine Learning**. *In an internal seminar at Ubiwhere*. Online.

2020 **SecureMLlib: Privacy-Preserving Distributed Machine Learning.** *In Ciclo de Eventos – Dinâmicas para a Inovação, Associação Nacional de Inovação (ANI).* Aveiro, Portugal.

Jury of Academic Evaluation

2024 **Jury of a master's thesis from the Master in Data Science.** *Active Inference against Federated Learning: Attacks and Solutions.* Ana Catarina Gomes, Faculty of Sciences of the University of Porto. Final grade: 20/20

Volunteer

2020 *International Conference on Learning Representations 2020 (ICLR'20).* Remote.

F3 - Teaching and Supervision Activities

In this section, I present my teaching and supervision activities. I have been teaching at the University of Minho since 2022, where I have been responsible for the development of class materials and exam and practical assignment evaluations in the area of Computer Science. I have also supervised several MSc theses and undergraduate research projects.

Teaching Activities

Since 2022, I have been an Invited Assistant Professor (25%) at University of Minho.

2024 – cur. **Big Data Analysis.** *Since 2024/2025.*

- Co-responsible with the class coordinator for developing theoretical and practical class contents, including using Jupyter notebooks, Apache Spark, and PyTorch for the practical classes.
- Lecturer of the practical classes to MEI and MIEI courses (4th year, 2nd semester).

Software Platforms. *Since 2024/2025.*

- Co-responsible, with the class coordinator, for restructuring the practical classes and their contents to allow for a better understanding of the course contents lectured in the theoretical classes (namely the basis of operating systems) and for developing exams and practical assignment evaluations.
- Lecturer of the practical classes to MIBiom, MBiom - MI course (4th year, 1st semester).

Object Oriented Programming. *Since 2023/2024.*

- Co-responsible, alongside the remaining class professors, for the practical assignment and exam evaluations. Responsible for maintaining the GitHub organization for the class and the student groups within the organization.
- Lecturer of the practical classes to LEI course (2nd year, 2nd semester).

2022 – cur. **Cloud Computing Application and Services** *Since 2022/2023.*

- Co-responsible, alongside the remaining class professors, for the practical assignment and exam evaluations and the continuous improvement and validation of practical class contents.
- Lecturer of the practical classes to MEI, MIEI, and MMC courses (4th year, 1st semester).

Supervision of MSc Theses

Supervision of MSc thesis from the 2nd year of MEI, the 5th year of MIEI and the 2nd year of MIEBIOM.

2024 – cur. **Ana Sousa.** *Classifying Alzheimer's Disease through Federated Learning and 3D Brain MRIs.* Co-advised with António Sousa, Beatriz Cepa and Alícia Oliveira. University of Minho.
Ricardo Araújo. *Neuroimaging analysis pipeline leveraging open-source frameworks in distributed environments.* Co-advised with António Sousa and Beatriz Cepa. University of Minho.
Alexandre Fernandes. *Energy Monitoring System for Large-Scale Infrastructures.* Co-advised with Ricardo Macedo, University of Minho.
Carlos Machado. *An NVMe-oriented I/O scheduler for reducing energy consumption.* Co-advised Ricardo Macedo, University of Minho.
Miguel Braga. *GPU Energy Control for LLMs.* Co-advised with Ricardo Macedo, University of Minho.
Gonçalo Sousa. *Fault-Tolerant storage solutions for AI training in HPC.* Co-advised with João Paulo, University of Minho.
André Lucena. *Storage Optimizations for large-scale AI training in HPC.* Co-advised with João Paulo, University of Minho.

2023 – cur. **Francisco Neves.** *Heterogeneous Storage Solution for Deep Learning Frameworks.* Co-advised with João Paulo and Ricardo Macedo, University of Minho.

2023 – 2024 **Diana Rodrigues.** *GPU energy control for Deep Learning systems.* Co-advised with Ricardo Macedo and António Sousa. Final grade: 19/20. University of Minho.
Mariana Amorim. *Energy control system for disaggregated storage resources.* Co-advised with João Paulo and Ricardo Macedo. Final grade: 18/20. University of Minho.
José Fernandes. *Comprehensive study of the energy impact of key-value stores.* Co-advised with Ricardo Macedo and António Sousa. Final grade: 19/20. University of Minho.
Alícia Oliveira. *Correction of Motion Artifacts using Deep Learning in High-Performance Computing.* Co-advised with António Sousa and Beatriz Cepa. Final grade: 20/20. University of Minho.
Sara Pereira. *Energy Control System for Large-Scale Infrastructures.* Co-advised with Ricardo Macedo and Carlos Baquero. Final grade: 18/20. University of Porto.

2022 – cur. **Luís Branco.** *Integration of Distributed Deep Learning in MCC application.* Co-advised with António Sousa and Beatriz Cepa, University of Minho.

2022 – 2023 **Maria Beatriz Moreira.** *Distributed I/O Optimizations for Deep Learning Training.* Co-advised with João Paulo and Ricardo Macedo. Final grade: 18/20. University of Minho.

2021-2022 **Beatriz Cepa.** *Deep Learning for Image Generation Using HPC.* Co-advised with António Sousa. Final grade: 18/20. University of Minho.
Beatriz Soares. *A Web and Machine Learning-based DICOM Image Tool.* Co-advised with António Sousa. Final grade: 20/20. University of Minho.
João Alves. *Cloud-based Healthcare Application with Privacy-preserving Analytics.* Co-advised with António Sousa. Final grade: 19/20. University of Minho.

2020 – 2022 **Ana Cláudia Abreu.** *Distributed Deep Learning Approach for Epilepsy Detection on EEG signals.* Co-advised with António Sousa. Final grade: 17/20. University of Minho.

2019 – 2020 **Margarida Machado.** *Distributed Deep Learning For Sleep Apnea Detection on ECG signals.* Co-advised with António Sousa. Final grade: 18/20. University of Minho.

Undergraduate Research Projects

Supervision of projects from the curricular unit Project in Computer Engineering (15 ECTS, 2nd year of MEI and 5th year of MIEI).

2024 – 2025 **André Ferreira, Carlos Machado, and Gonçalo Sousa.** Characterization of DL Models and LLMs with eBPF. Co-advised with Tânia Esteves, University of Minho.

2023 – 2024 **Diana Rodrigues, José Fernandes and Mariana Amorim.** Monitoring Energy Consumption with Fine-Granularity in Data-centric Applications. Co-advised with António Sousa and Ricardo Macedo, University of Minho

2019 – 2020 **Daniel Fernandes, Mariana Miranda and Helena Poleri**. Adaptative Storage System with Reinforcement Learning. Co-advised with João Paulo, University of Minho
João Barreira and Rafael Costa. Study and Evaluation of Secure Deep Learning Frameworks. Co-advised with João Paulo, University of Minho

Supervision of Research Projects

Supervision of students in the scope of research projects and grants.

2025 **Nuno Matos and Manuel Carvalho.** Development of an end-to-end pipeline resorting to Apache Spark and PyTorch for the analysis of large-scale data. Co-supervised with João Paulo, University of Minho.

Awards and distinctions of students under my supervision

Although the merit of the awards described below is of the students and the work developed by them, these are also a recognition of the impact of my guidance activities.

2024 **Diana Rodrigues** Master's award from the Ordem dos Engenheiros Região Norte.

F4 - Scientific and Technological Projects

This section describes the scientific and technological projects in which I have been involved, including the current and past projects, as well as the projects under review. I have been involved in more than 8 projects with national and international partners, with different topologies and funding sources and with different roles. I am actively involved in the scientific writing of the proposals (with currently 6 proposals under review) the project's management, and the research activities. I am currently co-leading (*i.e.,* sharing the Principal Investigator role on the INESC TEC side) the BringTrust project, a PT2030 project with a total budget of 1 654 244€.

These projects topics can be divided into three main topics: storage optimizations for AI, energy-aware systems and applied research in AI. The projects are funded by national and international funding sources, such as PT2030, FCT, Compete2020, and European projects.

Next, I present the list of the current and past projects, as well as the projects under review, in which I have been involved as a PI/Co-PI or Researcher (this last point includes the scientific writing of the proposal).

Name	Status	Start	Total	INESC TEC (€)	Role
BringTrust	Accepted	2025	1 654 244€	438 768€	INESC TEC Co-PI
DisaggregatedHPC	Accepted	2025	58 274.8€	29 899.59€	Researcher
BCD.S+M	Ongoing	2024	1 045 509€	463 144€	Researcher
GreenDat.AI	Ongoing	2023	6 056 207€	545 000€	Researcher
ATE	Ongoing	2023	274 195 106.32€	11 984 698.38€	Researcher
AIDA	Ended	2020	1 183 005€	414 109€	Researcher
BigHPC	Ended	2020	1 840 876€	415 000€	Researcher
PAStor	Ended	2020	90 136€	49 235€	Researcher
ESDM - Centra	Ongoing	2018	NF	NF	Researcher
Rescueware	In Review	–	900 227,84€	403 843.84€	Researcher
INOCULUM	In Review	–	174 096.00 €	174 096.00 €	Researcher

NF: No Funding; PC: Principal Investigator.

Current and Past Projects

2025 – cur. **BringTrust: Strengthening CI/CD Pipeline Cybersecurity and Safeguarding the Intellectual Property** *INESC TEC's Project Co-Coordinator*.

- The project aims to provide a secure and verifiable environment for the full CI/CD pipeline.
- I am the co-leading this project from the INESC TEC side, being co-responsible for the project's management and the research activities.
- My research tasks on the project also relate to research on privacy-preserving techniques for cybersecurity and trusted execution environments for CI/CD pipelines.
- I was actively involved on the scientific writing and submission of the proposal.
- *Partners*: INESC TEC, UAlg, SCALABIT.
- *Topology*: PT2030
- *Total Budget*: 1 654 244€
- *INESC TEC Budget*: 438 768€

DisaggregatedHPC: Towards energy-efficient, software-managed resource disaggregation in HPC infrastructures *Researcher*.

- This project aims to reconstruct the HPC infrastructure paradigms by disaggregating the resources and managing them in a software-managed way.
- Research on energy-efficient resource management for HPC infrastructures and on resource disaggregation for HPC infrastructures. Dissemination activities and writing of deliverables will also be part of my tasks.
- I was actively involved on the scientific writing of the proposal.
- *Partners*: INESC TEC and NarLabs.
- *Topology*: INESC TEC's Internal Calls
- *Total Budget*: 58 274.8€
- *INESC TEC Budget*: 29 899.59€

2024 – cur. **BCD.S+M: Modular Blockchain Data Storage and Management System with AI** *Researcher*.

- The project aims to develop a modular blockchain data storage and management system with an auto-configurable module based on AI.
- My tasks involve research on AI for systems optimization by automatic configuration, dissemination and writing of deliverables.
- I was involved on the scientific writing of the proposal regarding the AI aspects.
- *Partners*: INESC TEC and Invisible Lab.
- *Topology*: PT2030
- *Total Budget*: 1 045 509€
- *INESC TEC Budget*: 463 144€

PASM: Performance and Sustainability Management in Modern Data Centers *Researcher*

- Research of new storage, operating, and distributed systems building blocks for attending the performance and sustainability requirements of modern, large-scale infrastructures (e.g., cloud date centers, HPC supercomputers).
- *Topology*: Global CENTRA.
- *Partners*: INESC TEC, AIST (Japan), FIU (USA)
- *Website*: `https://www.globalcentra.org/projects/#pasm`.

2023 – cur. **GreenDat.AI: Energy-efficient AI-ready Data Spaces.** *Researcher*. ID: 101070416

- The project aims to develop a framework for energy-efficient AI-ready data spaces.
- My tasks are directly related to WP2 where a novel framework for benchmarking the energy consumption of AI models was developed.
- I became responsible for the management of the framework's development and I am currently responsible for the dissemination activities and writing of deliverables.
- Distinction of the Testing framework for AI energy efficiency by its listing in EU's Innovation Radar.
- *Partners*: International consortium with academic and industry partners.
- *Website*: `https://greendat.ai`.
- *Topology*: Horizon Europe
- *Total Budget*: 6 056 207€
- *INESC TEC Budget*: 545 000€

ATE: Energy Transition Alliance. *Researcher*.

- This project (*i.e.,* in WP12) aims to develop an AI-driven and Data Analysis Pipelines for DevOps and energy transition.
- My tasks involve the architectural design of such pipelines and definition of possible use-cases for real-world deployment.
- *Partners*: Multidisciplinary Portuguese Consortium.
- *Topology*: National Funds - PRR
- *Total Budget*: 274 195 106.32€
- *INESC TEC Budget*: 11 984 698.38€

2020 – 2023 **AIDA: Adaptive, Intelligent and Distributed Assurance Platform.** *Researcher*. ID: POCI-01-0247-FEDER-045907

- My tasks involved research on secure and privacy-preserving machine learning techniques and also the implementation of cryptographic and secure computation techniques for federated and distributed machine learning.
- I was actively involved on the scientific writing of the proposal regarding the AI and privacy-preserving aspects.
- *Partners*: Mobileum, INESC TEC, U. Coimbra, CMU.
- *Website*: `https://aida.inesctec.pt`.
- *Topology*: Compete2020
- *Total Budget*: 1 183 005€
- *INESC TEC Budget*: 414 109€

BigHPC: A Management Framework for Consolidated Big Data and HPC. *Researcher*. ID: POCI-01-0247-FEDER-045924

- This project aimed to develop a management framework with high-performance and high QoS for Big Data and HPC.
- My task involved research on storage optimization for large-scale AI and HPC applications.
- I was actively involved on the scientific writing of the proposal regarding the AI and storage optimizations for AI workloads.
- *Partners*: Wavecom, INESC TEC, LIP, MACC, UTAustin, TACC.
- *Website*: `https://bighpc.inesctec.pt`.
- *Topology*: Compete2020
- *Total Budget*: 1 840 876€
- *INESC TEC Budget*: 415 000€

2020 – 2021 **PAStor: Programmable and Adaptable Storage for AI-oriented HPC Ecosystems.** *Researcher*. ID: UTA-EXPL/CA/0075/2019

- This project aimed to develop a storage system for AI-oriented HPC environments.
- My task involved research on storage optimization for large-scale AI and HPC applications.
- I was actively involved on the scientific writing of the proposal regarding the AI and storage optimizations for AI workloads.
- *Partners*: INESC TEC, UT Austin, U. Minho.
- *Topology*: PT-UT Austin
- *Total Budget*: 90 136€
- *INESC TEC Budget*: 49 235€

2018 – cur. **ESDM: Efficient and Secure Data Management for HPC and Cloud Computing.** *Researcher*. Global CENTRA Project

- Research on privacy-preserving distributed machine learning for large-scale cloud environments.
- *Topology*: Global CENTRA.
- *Partners*: INESC TEC and AIST.
- *Website*: `https://www.globalcentra.org/projects/#prv`.

Submitted (under review)

Rescueware: Intelligent and Self-configurable Cybersecurity and Data Recovery for Ransomware Resilience *Researcher*.

- This project aims to develop an AI-driven system for cybersecurity and data recovery for ransomware resilience.
- My tasks will involve the research on AI for cybersecurity configuration of storage systems and on AI for efficient data recovery. Dissemination activities and writing of deliverables will also be part of my tasks.
- I was actively involved on the scientific writing of the proposal.
- *Partners*: INESC TEC, Invisible Lab and ULSAM.
- *Topology*: PT2030
- *Total Budget*: 900 227,84€
- *INESC TEC Budget*: 403 843.84€

INOCULUM: Boosting the Immunity of Storage Systems Against Ransomware *Researcher*.

- This project aims to deliver a novel storage system that is highly resilient to ransomware attacks and guarantees the integrity and availability of the data.
- My tasks will involve the research on AI for systems optimization by automatic configuration and on AI for cybersecurity configuration of storage systems. Alongside this, it is also expected that dissemination activities and writing of deliverables will be part of my tasks.
- I was actively involved on the scientific writing of the proposal.
- *Partners*: INESC TEC.
- *Topology*: FCT - IC&DT
- *Total Budget*: 174 096.00 €