Alice Plebe

September 14, 2026

E-mail: alice.plebe@gmail.com
Website: www.aliceplebe.com
Pronouns: She/Her
ORCID: 0000-0001-8567-0553

I am a researcher working at the intersection of artificial intelligence and engineering. In my work on autonomous driving, I developed deep neural networks inspired by the cognitive processes that enable humans to drive. More recently, I have investigated physics-encoded neural networks for modeling and control of mechatronic systems, using physical knowledge and mathematical models to structure their architectures. My research on large language models and vision-language models investigates human-like behaviors through behavioral experiments and interpretability analyses. In particular, I study how LLMs form trust toward other agents and how susceptibility to misinformation influences their reasoning and decisions.

I have contributed to collaborative projects funded through Horizon 2020, Horizon Europe, UKRI EPSRC, and the Italian Science Fund, involving industrial partners such as Centro Ricerche Fiat and the validation of AI methods on real vehicles. My teaching experience spans bachelor's, master's, and doctoral education, including specialist workshops for doctoral and postdoctoral researchers, and is complemented by formal training in student-centered and technology-enhanced teaching methods.

CURRENT POSITION

Junior Assistant Professor (RTD-A) *10/2025 – present*
Department of Industrial Engineering, University of Trento, Italy

- Investigating physics-encoded and model-structured neural networks for modeling and control of mechatronic systems, in which physical knowledge and mathematical models directly shape network architectures.
- Contributing to the FIS project "Neu4mes" (neu4mes.github.io) by investigating the conceptual and methodological relationships between model-structured neural networks and physics-encoded, physics-guided, and physics-informed approaches across robotics, industrial applications, materials engineering, chemistry, geophysics, and medical imaging.

PREVIOUS ACADEMIC POSITIONS

Senior Postdoctoral Researcher (Research Fellow) *05/2024 – 09/2025*
Department of Computer Science, University College London, United Kingdom

- Led a study showing that vision-language models exhibit susceptibility to image-driven misinformation comparable to that observed in humans, and analyzed how visual input and persona conditioning influence linguistic reasoning.
- Investigated linguistic trust and belief formation about others' trustworthiness in teams of LLM-based agents, including dispositional tendencies to trust; ongoing work uses mechanistic analysis to examine and interpret the associated internal representations.
- Contributed to the UKRI EPSRC project "Satisficing Trust in Human-Robot Teams," exploring trust-based mechanisms for collaboration applied to multi-agent teamwork in high-stakes scenarios.

Postdoctoral Researcher (Assegnista di Ricerca) *11/2020 – 04/2024*
Department of Industrial Engineering, University of Trento, Italy

- Led a systematic survey and critical classification of human-inspired approaches to autonomous driving, reviewing models grounded in cognitive science, neuroscience, and psychology, and organizing the field according to underlying cognitive principles.
- Investigated a framework for human–autonomous system collaboration based on a distributed cognition paradigm, implementing shared affordance-based representations to support coordinated decision-making between human drivers and autonomous agents.
- Contributed to the Horizon Europe project "Sunrise" (ccam-sunrise-project.eu), developing machine learning methods for Operational Design Domain analysis and scenario-based safety assessment in autonomous systems.

Visiting PhD Student *02/2020 – 06/2020*
Department of Cognitive Robotics, TU Delft, the Netherlands

- Developed a neuroscience–inspired occupancy grid mapping algorithm grounded in cortical magnification, incorporating non-uniform representational allocation across the visual field and improving obstacle prediction accuracy in safety-critical regions.

PhD Student *11/2017 – 10/2020*
Department of Information Engineering and Computer Science, University of Trento, Italy

- Conducted research on cognitively inspired neural architectures for visual mental imagery, supporting prediction and generation of novel visual scenarios in complex environments.
- Contributed to the EU Horizon 2020 project "Dreams4Cars" (`www.dreams4cars.eu`), exploring dream-like learning mechanisms in artificial agents, where simulated and imagined experiences were used to support learning in scenarios not yet encountered in the real world.

EDUCATION

PhD in Information and Communication Technology *20/04/2021*
Department of Information Engineering and Computer Science, University of Trento, Italy
Thesis: "Cognitively guided modeling of visual perception in intelligent vehicles".

Master's degree in Computer Science, 110/110 cum laude *29/11/2016*
Department of Mathematics and Computer Science, University of Catania, Italy

Bachelor's degree in Computer Science, 110/110 cum laude *25/07/2014*
Department of Mathematics and Computer Science, University of Catania, Italy

TEACHING

Lecturer, "Artificial Intelligence for Industrial Systems" (40 hours). *2025/26*
BSc in Industrial Engineering, University of Trento, Italy.

- Designed and delivered a course introducing the design and implementation of neural networks to students with an industrial engineering background.

Workshop Lead, "Introduction to LLM-based Agents" (4 hours). *06/2025*
Summer Institute in Computational Social Science,
UCL Computer Science and Imperial College Business School.

- Designed and delivered an interdisciplinary workshop for researchers in behavioral economics and the social sciences, combining an introduction to LLMs with practical methods for building LLM-based agents and applying them within participants' research domains.

Lecturer, "Vision-Language-Action models for Robotics and Autonomous Vehicles" (12 hours).
Doctoral School in Materials, Mechatronics and Systems Engineering, *2023/24*
University of Trento, Italy.

- Proposed and secured approval for a new doctoral module on vision-language-action models for robotics, which was formally incorporated into the doctoral school curriculum. Designed and delivered the course by combining theoretical foundations with hands-on Python programming for students from diverse academic backgrounds.

Guest Lecturer, "Sensors for Intelligent Vehicles and Autonomous Driving" (10 hours). *2022/23*
MSc in Mechatronics Engineering, University of Trento, Italy.

- Delivered a specialist module on computer vision and sensing technologies for autonomous vehicles. Adopted a flipped-classroom approach based on independent preparation and student-led seminar presentations.

Teaching Assistant, "C++ programming for Numerical Analysis" (40 hours). *2021/22 and 2022/23*
BSc in Industrial Engineering, University of Trento, Italy.

- Designed and delivered programming exercises translating numerical methods into practical C++ implementations, and provided individual support to reinforce students' conceptual understanding and programming skills.

CERTIFICATIONS AND TRAINING

Training in Higher Education Teaching Methodologies *06/2026*
FormID Teaching and Learning Center, University of Trento, Italy

- Completed a 20-hour program covering student-centered teaching approaches, including active learning, blended learning, challenge-based learning, and flipped classroom methods, as well as syllabus design, learning outcomes, constructive alignment, and assessment.

CapoCaccia Workshop for Neuromorphic Intelligence *05/2023*
Institute of Neuroinformatics, University of Zurich and ETH Zurich

- Completed a two-week workshop on neuromorphic engineering, covering biological foundations and hardware implementations.

Training on Deep Learning for Autonomous Vehicles – Perception *10/2018*
NVIDIA Deep Learning Institute, Munich, Germany

- Completed an intensive 8-hour certification course on perception for autonomous vehicles, leveraging NVIDIA architectures and platforms.

International Summer School on AI and Games *05/2018*
University of Crete, Chania, Greece

- Completed a 40-hour program on AI techniques for procedural content generation and player modeling in video games.

International Summer School on Deep Learning *07/2017*
University of Deusto, Bilbao, Spain

- Completed a 50-hour program covering fundamentals of deep learning and applications in computer vision, translation, and language processing.

PUBLIC ENGAGEMENT AND VOLUNTEER WORK

Public talk at *Pint of Science* Trento *05/2026*

- Delivered a public engagement talk on the emergence of mind-like characteristics in large language models and on computational approaches for studying reasoning, trust, and internal representations in AI systems.

Published for *The Conversation* *07/2024*

- Authored a public-facing article on the potential role of foundation models in supporting common-sense reasoning in autonomous vehicle technology [link].

Mentorship *2021 – 2024*
Voxel Community (`github.com/voxel-community`), *Trento, Italy*

- Organized mentoring and educational activities supporting individuals identifying as women in pursuing careers in technology, with a focus on inclusivity and wellbeing in academic and professional environments.

AWARDS

Best Student Paper Award *05/05/2019*
5th International Conference on Vehicle Technology and Intelligent Transport Systems.
For the paper "Mental Imagery for Intelligent Vehicles" (A. Plebe et al.).

PROFESSIONAL EXPERIENCE IN 3D VISUALIZATION

Forensic Reconstructions *2014 – 2021*

- Produced commissioned high-fidelity 3D reconstructions of crime scenes and incident dynamics, supporting analysis and communication in forensic and legal contexts.

Industry Visualization Projects *2015 – 2017*

- Developed commissioned 3D visualizations and prototypes for industrial and technology products, supporting design communication and technical showcasing.

TECHNICAL SKILLS

Programming languages and Development Environments

- Python (PyTorch, TensorFlow, HuggingFace Transformers, TRL, NumPy, SciPy, Pandas, Matplotlib), C, C++, C#, Wolfram Mathematica.
- Linux/Unix systems (Bash shell, Vim).

Computer Graphics and Game Engines

- Blender: extensive professional experience and advanced training.
- Unity: working knowledge, gained through a bachelor-level module on game development.

Languages: Italian (native), English (near-native proficiency, C2).

SELECTED ORAL PRESENTATIONS

Invited talk: "Trust as a collaboration mechanism in teams of language-based agents", *26/03/2025*
Financial Computing and Analytics Group seminars, UCL Computer Science [video]

Conference presentation: "Autonomous vehicles inspired by human brain and cognition", *17/07/2024*
UCL NeuroAI Annual Conference [video]

Invited talk: "Challenges and opportunities in autonomous driving", *09/11/2022*
USERN (Universal Scientific Education and Research Network) Congress

Conference presentation: "A cognitively informed perception model for driving", *08/11/2019*
International Conference on Intelligent Robots and Systems (IROS), Cognitive Vehicles workshop

Conference presentation: "Mental imagery for intelligent vehicles", *09/06/2019*
IEEE Intelligent Vehicles Symposium (IV), Challenges for Autonomous Driving workshop

Conference presentation: "Lighting design by genetic evolution", *28/10/2017*
Blender Conference [video]

Conference presentation: "Simulating fire outbreaks in industrial plants", *30/10/2016*
Blender Conference [video]

Complete List of Publications

JOURNALS

1. A. Plebe, "Propensity to trust in large language models," *PLoS ONE*, vol. 21, no. 5, p. e0347328, 2026
2. X. Dong, M. Cappuccio, H. A. Jassmi, F. Alnajjar, E. Debie, M. Ghasrikhouzani, A. Lanteri, A. Luqman, T. McGregor, O. Molloy, A. Plebe, M. Regan, and D. Zhang, "Why autonomous vehicles are not ready yet: A multi-disciplinary review of problems, attempted solutions, and future directions," *Journal of Field Robotics*, 2025
3. T. Dorigo, G. D. Brown, C. Casonato, A. Cerdà, J. Ciarrochi, M. Da Lio, N. D'Souza, N. R. Gauger, S. C. Hayes, S. G. Hofmann, R. Johansson, M. Liwicki, F. Lotte, J. J. Nieto, G. Olivato, P. Parnes, G. Perry, A. Plebe, I. M. Rao, N. Rezaei, F. Sandin, A. Ustyuzhanin, G. Vallortigara, P. Vischia, and N. Yazdanpanah, "Artificial intelligence in science and society: the vision of USERN," *IEEE Access*, vol. 13, pp. 15993–16054, 2025
4. A. Cherubini, G. P. Rosati Papini, A. Plebe, M. Piazza, and M. Da Lio, "Bootstrapped neural models for predicting self-driving vehicle collisions with quantified confidence: Offline and online applications," *IEEE Transactions on Intelligent Vehicles*, vol. [early access], 2025
5. A. Plebe, H. Svensson, S. Mahmoud, and M. Da Lio, "Human-inspired autonomous driving: A survey," *Cognitive Systems Research*, vol. 83, p. 101169, 2024
6. A. Plebe and M. Da Lio, "Bio-inspired circular latent spaces to estimate objects' rotations," *Frontiers in Computational Neuroscience*, vol. 17, 2023
7. M. Da Lio, A. Cherubini, G. P. Rosati Papini, and A. Plebe, "Complex self-driving behaviors emerging from affordance competition in layered control architectures," *Cognitive Systems Research*, vol. 79, pp. 4–14, 2023
8. A. Plebe, G. P. Rosati Papini, A. Cherubini, and M. Da Lio, "Distributed cognition for collaboration between human drivers and self-driving cars," *Frontiers in Artificial Intelligence*, vol. 5:910801, 2022
9. M. Da Lio, R. Donà, G. P. Rosati Papini, and A. Plebe, "The biasing of action selection produces emergent human-robot interactions in autonomous driving," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 1254–1261, 2022
10. G. P. Rosati Papini, A. Plebe, M. Da Lio, and R. Donà, "A reinforcement learning approach for enacting cautious behaviours in autonomous driving system: Safe speed choice in the interaction with distracted pedestrians," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, p. 8805 – 8822, 2021
11. A. Plebe and M. Da Lio, "On the road with 16 neurons: Towards interpretable and manipulable latent representations for visual predictions in driving scenarios," *IEEE Access*, vol. 8, pp. 179716–179734, 2020
12. A. Plebe, M. Da Lio, and D. Bortoluzzi, "On reliable neural network sensorimotor control in autonomous vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, pp. 711–722, 2020
13. A. Plebe and G. Grasso, "Conceptual integrity without concepts," *International Journal of Software Engineering and Knowledge Engineering*, vol. 28, no. 7, pp. 955–981, 2018

CONFERENCES AND WORKSHOPS

1. A. Plebe, T. Douglas, D. Riazi, and R. del Rio-Chanona, "Images amplify misinformation sharing in Vision-Language Models," in *Proceedings of the 20th International AAAI Conference on Web and Social Media (ICWSM)*, 2026

2. A. Cherubini, G. P. Rosati Papini, A. Plebe, A. Giugliano, M. Muro, and M. Da Lio, "A subsumption scheme for emergent collaboration of self-driving vehicles in intersections," in *Proceedings of the 17th IFAC Symposium on Control of Transportation Systems (CTS)*, vol. 58, pp. 43–47, Elsevier, 2024

3. A. Cherubini, G. P. Rosati Papini, A. Plebe, and M. Da Lio, "Energy costs of safe speed policies in a pedestrian-crossing scenario," in *Proceedings of the 35th IEEE Intelligent Vehicles Symposium (IV)*, pp. 1–6, IEEE, 2023

4. E. Pagot, M. Piccinini, A. Plebe, E. Bertolazzi, and F. Biral, "Real-time autonomous parking in unstructured scenarios with an indirect optimal control approach," in *Workshop at the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2022

5. S. Mahmoud and A. Plebe, "A critical look into cognitively-inspired artificial intelligence," in *8th International Workshop on Artificial Intelligence and Cognition (AIC)*, 2022

6. A. Plebe, J. F. Kooij, G. P. Rosati Papini, and M. Da Lio, "Occupancy grid mapping with cognitive plausibility for autonomous driving applications," in *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops (ICCVW))*, pp. 2934–2941, 2021

7. A. Plebe and M. Da Lio, "Neurocognitive-inspired approach for visual perception in autonomous driving," in *Smart Cities, Green Technologies and Intelligent Transport Systems*, pp. 113–134, Cham: Springer International Publishing, 2021

8. A. Plebe and M. Da Lio, "Visual perception for autonomous driving inspired by convergence–divergence zones," in *Proceedings of the 11th International Symposium on Image and Signal Processing and Analysis (ISPA)*, pp. 204–208, IEEE, 2019

9. A. Plebe and M. Da Lio, "Variational autoencoder inspired by brain's convergence-divergence zones for autonomous driving application," in *Proceedings of the 20th International Conference on Image Analysis and Processing (ICIAP)*, vol. 11751 of *Lecture Notes in Computer Science*, pp. 367–377, Springer, Cham, 2019

10. A. Plebe, R. Donà, G. P. Rosati Papini, and M. Da Lio, "Mental imagery for intelligent vehicles," in *Proceedings of the 5th International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS)*, pp. 43–51, Science and Technology Publications, 2019

11. A. Plebe, G. P. Rosati Papini, R. Donà, and M. Da Lio, "Dreaming mechanism for training bio-inspired driving agents," in *Proceedings of the 2nd International Conference on Intelligent Human Systems Integration (IHSI)*, pp. 429–434, Springer, Cham, 2019

12. A. Plebe, V. Cutello, and M. Pavone, "Optimizing costs and quality of interior lighting by genetic algorithm," in *Computational Intelligence: 9th International Joint Conference, IJCCI 2017 Funchal-Madeira, Portugal, November 1-3, 2017 Revised Selected Papers*, pp. 19–39, Cham: Springer International Publishing, 2019

13. M. Da Lio, A. Plebe, D. Bortoluzzi, G. P. Rosati Papini, and R. Donà, "Autonomous vehicle architecture inspired by the neurocognition of human driving," in *Proceedings of the 4th International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS)*, pp. 507–513, Science and Technology Publications, 2018

14. A. Plebe, V. Cutello, and M. Pavone, "Evolving illumination design following genetic strategies," in *Proceedings of the 9th International Joint Conference on Computational Intelligence (IJCCI)*, pp. 289–296, Science and Technology Publications, 2017

15. A. Plebe and M. Pavone, "Multi-objective genetic algorithm for interior lighting design," in *Proceedings of the 3rd International Workshop on Machine learning, Optimization, and Big Data (MOD)*, vol. 10710 of *Lecture Notes in Computer Science*, pp. 222–233, Springer, Cham, 2017

16. A. Plebe and G. Grasso, "Particle physics and polyedra proximity calculation for hazard simulations in large-scale industrial plants," in *Proceedings of the 12th International Conference of Computational Methods in Sciences and Engineering (ICCMSE)*, pp. 090003–1–090003–4, American Institute of Physics Publishing, 2016