

# Roberta Rocca

Skt Pauls Gade 41, 4th, 8000 Aarhus C, Denmark

☎ +45 93 92 03 66 | ✉ [rbrcc@gmail.com](mailto:rbrcc@gmail.com) | 🌐 <https://rbroc.github.io> | 📄 [rbroc](#)

## WORK EXPERIENCE

<b>Tenure-Track Assistant Professor</b>	12/2022 – Present
<i>Department of Culture, Cognition, and Computation, Aarhus University</i>	<i>Aarhus, Denmark</i>
<b>Postdoctoral Researcher</b>	03/2022 – 12/2022
<i>Department of Culture, Cognition, and Computation, Aarhus University</i>	<i>Aarhus, Denmark</i>
<b>Predictive Analytics Data Fellow</b>	06/2021 – 08/2021
<i>Centre for Humanitarian Data, United Nations OCHA</i>	<i>The Hague, The Netherlands</i>
<b>Postdoctoral Researcher</b>	11/2019 – 03/2022
<i>Psychoinformatics Lab, University of Texas at Austin</i>	<i>Austin, TX</i>

## ADDITIONAL RESEARCH EXPERIENCE

<b>Visiting Researcher</b>	06/2024 – 07/2024
<i>MilaNLP group @ Department of Computing Sciences, Bocconi University</i>	<i>Milan, Italy</i>
<b>Visiting Researcher</b>	08/2018 – 03/2019
<i>Department of Applied Mathematics and Computer Science, Technical University of Denmark</i>	<i>Copenhagen, Denmark</i>
<b>Visiting Researcher</b>	06/2019 – 07/2019
<i>Department of Psychology, University of East Anglia</i>	<i>Norwich, United Kingdom</i>
<b>Visiting Researcher</b>	01/2018 and 10/2018
<i>Institute of Cognitive Sciences and Technologies, National Research Council</i>	<i>Rome, Italy</i>
<b>Research Assistant</b>	03/2016 – 07/2016
<i>Department of Psychology, University College London</i>	<i>London, United Kingdom</i>
<b>Research Assistant</b>	08/2015 – 03/2016
<i>Department of Language and Cognition, Max Planck Institute for Psycholinguistics</i>	<i>Nijmegen, The Netherlands</i>

## EDUCATION

<b>PhD in Cognitive Science</b>	09/2016 – 10/2019
<i>Department of Linguistics, Cognitive Science, and Semiotics, Aarhus University</i>	<i>Aarhus, Denmark</i>
<b>MSc in Cognitive Science</b>	09/2014 – 07/2016
<i>Centre for Mind/Brain Sciences, University of Trento</i>	<i>Trento, Italy</i>
<b>BA in Philosophy</b>	10/2010 – 03/2014
<i>School of Humanities, University of Palermo</i>	<i>Palermo, Italy</i>

## TEACHING

<b>Data Science, Prediction, and Forecasting</b>   <i>MSc Cognitive Science, Aarhus University</i>	Spring 2024
<b>Natural Language Processing</b>   <i>MSc Cognitive Science, Aarhus University</i>	Fall 2023
<b>Applied Cognitive Science</b>   <i>BSc Cognitive Science, Aarhus University</i>	Spring 2023, Spring 2024
<b>Cognitive Neuroscience [Teaching Assistant]</b>   <i>BSc Cognitive Science, Aarhus University</i>	Spring 2018
<b>Cognitive Science</b>   <i>MA Cognitive Semiotics, Aarhus University</i>	Fall 2017

## SKILLS

**Methods:** Machine Learning; Data Science; Statistics; Experimental Methods; Agent-based Simulations  
**Programming:** Python, R, SQL, Matlab, bash, HTML/CSS, PERL, Git/GitHub, CI/CD, Docker, VSCode  
**Main Libraries:** PyTorch, TensorFlow/Keras, Transformers, Scikit-Learn, NumPy, SciPy, StatsModels, Pandas  
**Languages:** Italian (native), English (fluent), Danish (fluent), French (proficient)

## ONGOING PROJECTS (SELECTED)

---

### **Contrastive Learning for User-Aware Language Modeling** | *with Dirk Hovy*

- Applying contrastive learning protocols to tune LLMs to learn user-aware embeddings
- Creating datasets for training and downstream evaluation from pre-existing data sources (e.g., arXiv, YELP)
- Benchmarking results models on sentence-level and user-level NLP tasks

### **Evaluating LLM Safety in Mental Health Requests** | *with Paul Röttger*

- Assessing prevalence of mental health requests in real LLM user prompts (e.g., WildChat)
- Developing data-driven taxonomy of mental health requests involving safety risks
- Developing a safety evaluation dataset for mental health and used it to evaluate closed- and open-source LLMs

### **Biases in LLM-based Assessment of Mental Health** | *with Flor Plaza-del-Arco*

- Benchmarking multiple open-source models on diagnostics and severity assessment for mental health
- Creating novel datasets to evaluate demographic and intersectional biases in LLMs x Mental Health
- Evaluating intersectional biases in closed- and open-source LLMs across tasks and prompting strategies

## PAST PROJECTS (SELECTED)

---

### **Neuroscout: a platform for reproducible and scalable analysis of fMRI data** | *Psychoinformatics Lab*

- Developed and maintained the feature extraction library *pliers*
- Performed scientific validation of feature extraction and analysis workflows
- Contributed features and documentation, and provided input on API development
- Performed QA across multiple stages of development
- Authored Neuroscout's release paper and its open-source code

### **Benchmarking NLP methods on Differential Diagnosis for Mental Health** | *Aarhus University*

- Developed datasets and pipelines for benchmarking of NLP methods on text-based differential diagnosis
- Evaluated over 100 models on text- and speech-based differential diagnosis of psychiatric disorders
- Authored scientific paper reporting research results and related open-source code

### **Dynamical Modeling for Humanitarian Needs Forecasting** | *Centre for Humanitarian Data*

- Identified methods, data requirements, and partnerships for research project on cholera response
- Provided strategic recommendations on causal and dynamical modeling for humanitarian needs assessment
- Communicated findings to non-technical audience through reports, blog posts, talks, and webinars

### **Computational Modeling of European Identity from Social Media and Survey Data** | *Aarhus University*

- Collected social media datasets to track topics and sentiment towards European institutions
- Developed workflows for semantic and stylistic analysis of multilingual social media data
- Coordinated online data collection across an international research consortium
- Authored research papers analysing topic and sentiment dynamics from text and survey data

## PUBLICATIONS (SELECTED)

---

**Rocca, R.**, Pistilli, G., Maheshwari, K., Fusaroli, R. (forthcoming), Introducing ELLIPS: an ethics-centered approach to research on LLM-based inference of psychiatric conditions, *Proceedings of the 2024 AAAI/ACM Conference on AI, Ethics, and Society*

Hansen, L., **Rocca, R.**, Simonsen, A., Olsen, L., Parola, A., Bliksted, V., ... & Fusaroli, R. (2023), Speech-and text-based classification of neuropsychiatric conditions in a multidagnostic setting. *Nature Mental Health*, 1-11.

**Rocca, R.**, Tamagnone, N., Contla, X., Bove, J.B., Rekabsaz, N. (2023), Natural language processing for humanitarian action: challenges, opportunities, and the path towards a humanitarian NLP community, *Frontiers in Big Data*

**Rocca, R.**, Yarkoni, T. (2022), Language as a fingerprint: Self-supervised learning of user encodings using transformers, *Findings of the Association for Computational Linguistics: EMNLP 2022*

de la Vega, A.\*, **Rocca, R.\***, Blair, R., Mentch, J., Markiewicz, C., Ghosh, S., Poldrack, R., Yarkoni, T. (2022), Neuroscout: a unified platform for generalized and reproducible fMRI research, *eLife*

**Rocca, R.**, Tylén, K. (2022), Cognitive diversity promotes collective creativity: an agent-based simulation, *Proceedings of the 44th Annual Meeting of the Cognitive Science Society*

- Rocca, R.**, de la Vega, A. (2022), Evaluating the role of non-lexical markers in GPT-2’s language modeling behavior, *Proceedings of the 3rd Workshop on Evaluation and Comparison of NLP Systems*
- Rocca, R.**, Yarkoni, T. (2021). Putting psychology to the test: rethinking model evaluation through benchmarking and prediction, *Advances in Methods and Practices in Psychological Science*
- Rocca, R.**, Coventry, K. R., Tylén, K., Staib, M., Lund, T. E., Wallentin, M. (2020). Language beyond the language system: dorsal visuospatial pathways support processing of demonstratives and spatial language during naturalistic fast fMRI, *NeuroImage*
- Wallentin, M., **Rocca, R.** (2020). The semantics of spatial demonstratives. *Proceedings of the 42nd Annual Meeting of the Cognitive Science Society*
- Rocca, R.**, Wallentin, M., Vesper, C., Tylén, K. (2018). This and that back in context: grounding demonstrative reference in manual and social affordances, *Proceedings of the 40th Annual Meeting of the Cognitive Science Society*

---

## PREPRINTS

- Kardós, M., Kostkan, J., Vermillet, A.Q., Nielbo, K., Enevoldsen, K., **Rocca, R.** (2024), S<sup>3</sup> – Semantic Signal Separation, [arXiv:2406.09556](https://arxiv.org/abs/2406.09556)
- Rocca, R.**, Tylén, K. (2024). The effect of diversity and social interaction on cognitive search: An agent-based simulation, PsyArXiv, <https://doi.org/10.31234/osf.io/n3t6j>
- Luo, X., Rechardt, A., Sun, . . . , **Rocca, R.**, . . . , & Love, B. C. (2024). Large language models surpass human experts in predicting neuroscience results, arXiv, [arXiv:2403.03230](https://arxiv.org/abs/2403.03230)

---

## OPEN-SOURCE CONTRIBUTIONS

- Co-developer** of *Turftopic*, a Python library for scikit-learn compatible transformer-based topic modeling: <https://github.com/x-tabdeveloping/turftopic>
- Co-developer** and **maintainer** of *pliers*, Python library for automated multimodal feature extraction using machine learning: <https://github.com/PsychinformaticsLab/pliers>
- Co-developer** and **maintainer** of the *Neuroscout* ecosystem, a suite of tools for fast and reproducible analysis of naturalistic fMRI data: <https://github.com/neuroscout>
- Contributor** to the Massive Multilingual Text Embedding Benchmark (MMTEB), a big science initiative for multilingual and multi-task benchmarking of LLM embeddings: <https://github.com/embeddings-benchmark/mteb>
- Code to **reproduce my research papers** is available on GitHub, see for example:
- Rocca & Tylén, 2024: <https://github.com/rbroc/simcat>
  - Rocca & Yarkoni, 2022: <https://github.com/rbroc/contrastive-user-encoders>
  - Rocca & Yarkoni, 2021: [https://github.com/rbroc/benchmarks\\_paper](https://github.com/rbroc/benchmarks_paper)

---

## BLOG POSTS AND TECHNICAL REPORTS

- Rocca, R.** (2024), [Interpreting R<sup>2</sup>: a Narrative Guide for the Perplexed](#), *Towards Data Science*
- Rocca, R.** (2021) [Complex Systems Modeling for Humanitarian Action: Methods and Opportunities](#), *Research Report for the United Nations’ Centre for Humanitarian Data*

---

## INVITED TALKS, GRANTS AND AWARDS (SELECTED)

Invited Talk at Centre for Mind/Brain Sciences, University of Trento	June 2024
Recruitment Grant by Aarhus University Research Foundation: 1.949.000 DKK	December 2023
Invited Talk at Humanitarian Networks and Partnerships Week (HNPW)	April 2023
Invited Talk at MIT/Harvard Speech and Language Biomarkers Interest Group	December 2022
Invited Talk at Joint seminar series – University of Toronto, McGill University, Karolinska Institute	May 2021
University of Trento Merit Award: 3.000 EUR	July 2016