

# Jérémy Perret

AI alignment outreach, AI research & engineering

✉ me@jrmyp.net

📄 jrmyp.net

French nationality

## Full-time work experience

- 2023–2024 **AI alignment outreach**, *Suboptimal IA*, Lyon, France.  
(ongoing) Production of video, audio and text content on AI alignment, safety and governance, for a French-speaking audience. Funded by the Long-term Future Fund until February 2024, full-time since September 2023.
- 2023 **Head of Data**, *Together.do*, Lyon, France (remote).  
Design, implementation and maintenance of the company's AI infrastructure. Deployment of fast matching models for brand co-marketing; of large language model-based annotation and classification of text; of interactive dashboards for hyperparameter tuning. Use of Python, scikit-learn, spaCy, Dash, OpenAI API, Docker, Google Cloud Platform. Internal & external AI communication.
- 2021–2022 **Data Analysis Instructor**, *Wild Code School*, Lyon, France.  
Teaching, management and support of 24 students over two cohorts in a bootcamp-like environment, studying Python, SQL, and a wide variety of data analysis skills. Design and improvement of course materials, lectures and workshops.
- 2017–2020 **AI R&D Lead**, *Hubware*, Toulouse, France.  
Design, implementation and maintenance of the company's AI infrastructure. Fast creation and deployment of predictive models to accelerate and optimize customer service operations. Applied research in natural language processing, design of novel unsupervised learning algorithms. Management of data quality, scalability, annotation campaigns. Use of Python, TensorFlow, Keras, SciPy, GCP. Management of the 3-people R&D team, internal & external AI communication.
- 2013–2016 **Ph.D. in Artificial Intelligence**, *Toulouse Institute of Computer Science Research (IRIT)*.  
Parsing dialogue and argumentative structures (see dedicated section).
- 2012–2013 **Masters Degree in Computer Science, Artificial and Collective Intelligence**, *University of Toulouse*.  
Study of semantic compositionality of adjectives. Development of a word embedding Python library.
- 2009–2013 **Engineering Degree in Computer Science and Applied Mathematics**, *ENSEEIH*, Toulouse.

## Part-time work & volunteer experience

- 2018–2024 **Effective altruism community-building**, *Altruisme Efficace France*.  
(ongoing) Organization of regular meetups, physical and online. Writing online primers about AI risk and longtermism for a French-speaking audience. Board member since 2021, current treasurer.
- 2022–2024 **Public speaking on AI alignment**, *various outlets*.  
Appearances in three French-speaking podcasts, questions on the case for AI alignment research and on extreme risk scenarios. Academic seminar presentation at Grenoble University.
- 2021–2023 **Cohort facilitation**, *AGI Safety Fundamentals*.  
Coordination and facilitation of five cohorts for three iterations of the programme, covering a wide range of AI alignment and governance topics.
- 2020–2021 **Modeling Transformative AI Risk**, *Johns Hopkins University APL, then independent group*.  
Forecasting AI risk through probabilistic modeling and expert elicitation. Detailed breakdown of medium- and long-term scenarios, identification of critical risk factors, integration of expert opinion into a unified framework of risk assessment. Funded by the Long-term Future Fund under David Manheim's management.
- 2019–2021 **Lectures on data science and management**, *Digital Campus*, Toulouse.  
How to manage data-related projects, data pipelines, storage, visualization, quality and automation.
- 2017–2018 **Public speaking on AI**, Toulouse.  
Primers on general concepts on AI and automation to a general audience.

---

## Languages

French native

English fluent

---

## Ph.D. thesis

Title *Parsing dialogue and argumentative structures.*

Advisors Professor Nicholas Asher & Lecturer Stergos Afantenos, MELODI team

Description Automated extraction of dialogue and argumentative structures from natural language dialogues. Use of combined techniques from machine learning and integer linear programming to create constrained graph structures, following formalized linguistic norms and cues.

Open access <http://thesesups.ups-tlse.fr/3585/>

---

## Publications

- 2016 **Integer Linear Programming for Discourse Parsing**, *Jérémy Perret, Stergos Afantenos, Nicholas Asher, Mathieu Morey*, North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2016.  
A new approach for extracting discourse structure from dialogue, combining machine learning and linear optimization.
- 2016 **Parallel discourse annotations on a corpus of short texts**, *Manfred Stede, Stergos Afantenos, Andreas Peldszus, Nicholas Asher, Jérémy Perret*, Tenth International Conference on Language Resources and Evaluation, 2016.  
Annotation of a corpus of argumentative texts with three distinct frameworks.
- 2015 **Defining the Right Frontier in Multi-Party Dialogue**, *Julie Hunter, Nicholas Asher, Eric Kow, Jérémy Perret, Stergos Afantenos*, Workshop on the semantics and pragmatics of dialogue (Semdial), 2015.  
Study of valid discourse structures in multi-party chat.
- 2015 **Discourse parsing for multi-party chat dialogues**, *Stergos Afantenos, Eric Kow, Nicholas Asher, Jérémy Perret*, Empirical Methods on Natural Language Processing, 2015.  
Building discourse structure automatically from dialogue with machine learning and graph algorithms.
- 2014 **Revealing Resources in Strategic Contexts**, *Jérémy Perret, Stergos Afantenos, Nicholas Asher, Alex Lascarides*, Workshop on the semantics and pragmatics of dialogue (Semdial), 2014.  
Extracting useful information from negotiations between humans, in order to participate in them.