

Claudio Aracena

claudio.aracena@uchile.cl – github.com/caracena – claudioaracena.com

Education

- **University of Chile** **Santiago, Chile**
PhD candidate in Engineering Systems 2021 - Present
 - **University of Sydney** **Sydney, Australia**
Master of Information Technology and Master of Information Technology Management 2015 - 2017
 - **University of Chile** **Santiago, Chile**
Industrial Engineering with Major in Information Technology 2008 - 2013
Graduated with highest distinction
-

Experience

- **GobLab UAI - Adolfo Ibáñez University** **Santiago, Chile**
Associate Researcher January 2022 - Present
Data Scientist October 2019 - December 2021
 - Data analysis for GobLab UAI projects and grant applications. Main projects: analysis of citizens' comments about Congress bills using NLP and air pollution forecasting in an industrial zone.
- **Chilean Safety Association** **Santiago, Chile**
Data Science Consultant October 2022 - Present
 - Pre-training of clinical language models and applications.
- **Center for Mathematical Modeling - University of Chile** **Santiago, Chile**
Project Engineer October 2018 - September 2019
 - No-show prediction for medical appointments at the Luis Calvo Mackenna Hospital using machine learning.
- **Chatbot Chile** **Santiago, Chile**
Co-founder July 2017 - Present
 - Development of chatbots for several clients and dashboard for chatbot administration (chatbotchile.cl).
- **Plannertrip** **Sydney, Australia**
Co-founder December 2016 - September 2017
 - Development of Plannertrip platform (plannertrip.com, travel startup), apply for funding, and manage Plannertrip team. Plannertrip was part of the INCUBATE program from the University of Sydney.
- **Data61, CSIRO** **Sydney, Australia**
Associate Software Developer October 2015 - December 2016
 - Software development for 64knots team (freight and shipping optimization and machine learning group).
- **VSB - Technical University of Ostrava** **Ostrava, Czech Republic**
Research Intern October 2014 - January 2015
 - Study of relation between pupil dilation and emotions based on neural network analysis.
 - * Identifying salient website object: Comparative study between neurodata, pupil dilation and eye fixations.
 - * Survey on Big Data and Internet of Things.
- **Web Intelligence Centre - University of Chile** **Santiago, Chile**
Project Engineer 2014
 - Study of relation between EEG, pupil dilation, and emotions based on machine learning methods. Application of results to web object saliency.
 - Keep track of the AKORI project and write progress reports for funding institution.

- **Infosys Limited**

Intern

Bangalore, India

January - March 2013

- Development of a text classification project using tweets. Data modeling and implementing a database for tweets storage and posterior analysis.
-

Teaching Experience

- **Adolfo Ibáñez University, School of Government**

Academic Director

Lecturer

Santiago, Chile

2023 - Present

2020 - Present

- Machine Learning and Text Mining courses of the Diploma of Data Science for Public Policy, and Data Science Project Scoping course.

- **University of Development, Faculty of Engineering**

Lecturer

Santiago, Chile

2018 - 2022

- Course of storage and data capture of Master of Data Science program.

- **Intersect Australia**

Trainer

Sydney, Australia

February 2018 - August 2018

- Teaching using live coding methodologies (Python, R, SQL, MATLAB, BASH, GIT, among others).

- **University of Chile, Department of Industrial Engineering**

Lecturer

Santiago, Chile

March 2014 - July 2014

- Course of Data Warehousing.

- **University of Chile, Department of Industrial Engineering**

Teaching Assistant

Santiago, Chile

2012 - 2013

- Information Technologies: 2 semesters; Web Mining: 2 semesters
 - Introduction to Knowledge Management and Web and Business: 1 semester each course
-

Publications

Journal papers

Dunstan, J., Vakili, T., Miranda, L., Villena, F., **Aracena, C.**, Quiroga, T., Vera, P., Viteri, S. Rocco, V. A pseudonymized corpus of occupational health narratives for clinical entity recognition in Spanish (2024). *BMC Medical Informatics and Decision Making*, 24(1), 204.

Aracena, C., Villena, F., Arias, F., Dunstan, J. Aplicaciones de aprendizaje automático en salud (2022). *Revista Médica Clínica Las Condes*, 33(6), 568-575.

Conference papers

Rodolfo Garcia-Flores, Soumya Banerjee, George Mathews, Blandine Vacher, Brian Thorne, Nazanin Borhan, **Claudio Aracena** and Yuriy Tyshetskiy. Analysis of Demand and Operations of Inter-modal Terminals (2017). National Conference of the Australian Society for Operations Research.

Oweis, N.E., **Aracena, C.**, George, W., Oweis, M., Soori, H., Snašel, V. Internet of Things: Overview, Sources, Applications and Challenges (2016). Proceedings of the Second International Afro-European Conference for Industrial Advancement AECIA.

Aracena, C., Basterrech, S., Snašel, V., and Velásquez, JD. Neural Networks for Emotion Recognition Based on Eye Tracking Data (2015). IEEE International Conference on Systems, Man, and Cybernetics (IEEE SMC).

Slanzi, G., **Aracena C.**, and Velásquez, JD. Eye tracking and EEG features for salient Web object identification (2015). International Conference on Brain Informatics & Health (BIH).

Workshop papers

Aracena, C., Miranda, L., Vakili, T., Villena, F., Quiroga, T., Núñez-Torres, F., Rocco, V., and Dunstan, J. A Privacy-Preserving Corpus for Occupational Health in Spanish: Evaluation for NER and Classification Tasks (2024). The 6th Clinical Natural Language Processing Workshop at NAACL.

Aracena, C., Rodríguez, N., Rocco, V., and Dunstan, J. Pre-trained language models in Spanish for health insurance coverage (2023). The 5th Clinical Natural Language Processing Workshop at ACL.

- Aracena, C.**, and Dunstan, J. Development of pre-trained language models for clinical NLP in Spanish (2023). Student Research Workshop (SRW) at EACL.
- Aracena, C.**, Villena, F., Rojas, M., and Dunstan, J. A Knowledge-Graph-Based Intrinsic Test for Benchmarking Medical Concept Embeddings and Pretrained Language Models (2022). International Workshop on Health Text Mining and Information Analysis (LOUHI) at EMNLP.
- Aracena, C.**, Loyola, P., Slanzi, G., Matsuo, Y., and Velásquez, JD. Towards An Unified Replication Framework for EEG-based Emotion Classification (2015). BrainKDD: The 2nd International Workshop on Data Mining for Brain Science at ACM SIGKDD.

Policy Papers

Gabriela Denis, María Paz Herмосilla, **Claudio Aracena**, Roberto Sánchez Ávalos, Natalia González Alarcón, and Cristina Pombo. Responsible use of AI for public policy: Project formulation manual (2021). Inter-American Development Bank Publications.

Core Technical Skills

Programming Languages: Python, R, Javascript, SQL, PHP, Java, MATLAB, \LaTeX .

Operative Systems: Ubuntu, Windows

Frameworks and technologies: Pytorch, HuggingFace, Django, Wordpress, Amazon Web Services, GIT, Docker, Microsoft Office (Excel, Word and Power Point).

Languages

Spanish: Native speaker

English: Advanced – TOEFL Score: 96 (2015)

Distinctions

- **ANID** **Santiago, Chile**
Scholarship for National PhD programs 2021 - Present
 - The scholarship is awarded to selected students who pursue a PhD program in a Chilean university.
- **Conicyt** **Santiago, Chile**
Becas Chile (Scholarship) 2015 - 2017
 - The scholarship is awarded to students who pursue a postgraduate degree in a remarkable foreign university and have shown excellent academic records.
- **University of Chile** **Santiago, Chile**
University of Chile Scholarship 2007 - 2013
 - It is awarded to the first 100 students that have been selected by the university for the first year.