

Dr. Juan Carlos Roldán

Co-founder & CTO @ Stargazr | Data extraction researcher

[linkedin.com/in/jcarlosroldan](https://www.linkedin.com/in/jcarlosroldan)

juancarlos@sevilla.es

May 12, 2023

I started my research career in 2014, focused on enterprise data integration. I prefer hybrid research approaches to incorporate R&D into industry development cycles. In 2020, I launched an AI-based fintech company that leverages the knowledge generated during my PhD. My personal interests include climbing, making music and space exploration.

Work experience

- 2020 – now** **Co-founder & CTO**
Seville, Spain *Stargazr Inc.*
Los Angeles, USA
Hamburg, Germany Founded a fintech startup to improve profitability in manufacturing companies. Currently in pre-seed stage, it's already present in three continents, have 10+ FTEs and raised \$2.5M total funding.
- 2019 – now** **Substitute Teacher**
Seville, Spain *University of Seville*
Researching new approaches to information extraction focused on reinforcement learning, knowledge graphs and quantum annealing. Lecturing related subjects such as Intelligent Information Systems.
- 2018 – 2019** **Visiting Researcher**
Los Angeles, USA *USC Information Sciences Institute, sponsored by Fulbright*
I researched on unsupervised web table extraction. The results are deployed into a IARPA geo-political forecasting tool called SAGE and a search engine for datasets and APIs called DataMart.
- 2015 – 2021** **Researcher**
Seville, Spain *University of Seville*
Worked on the creation of automated machine learning techniques for web information extraction, which led to several international collaborations and industrial applications.
- 2014 – 2015** **Research Intern**
Seville, Spain *University of Seville*
Researched on detecting keywords in the Spanish Twitter and cluster them into topics. The result has been cited several times and was used to predict the evolution of COVID-19 in LATAM countries.

Formal education

- 2016 – 2020** **Ph.D. in Computer Science**
Seville, Spain University of Seville – Honors: Cum Laude
- 2015 – 2016** **M.S. in Computer Science**
Seville, Spain University of Seville – GPA: 9.28 / 10
- 2011 – 2015** **B.S. in Software Engineering**
Seville, Spain University of Seville – GPA: 9.02 / 10

Journal articles

- January 2022** **Information Fusion**
Under review TOMATERA: A tool to support table extractors.
- January 2022** **Knowledge and Information Systems**
Published A hybrid quantum approach to leveraging data from HTML tables.
- January 2022** **Applied Soft Computing**
Published A coral-reef approach to extract information from HTML tables.
- December 2022** **Information Sciences**
Published On Exploring Data Lakes by Finding Compact, Isolated Clusters.
- November 2021** **Information Processing & Management**
Published A genetic clustering approach to extracting data from HTML tables.
- June 2020** **Software, Practice and Experience**
Published On the synthesis of meta-data tags for HTML files.
- May 2020** **Enterprise Applications of Artificial Intelligence**
Under review A clustering approach to extract data from HTML tables.
- May 2020** **Information Sciences**
Published TOMATE: Online Multi-document Automated Table Extraction.
- November 2019** **Knowledge-based Systems**
Published On extracting data from tables that are encoded using HTML.
- June 2015** **Advances in Distributed Computing and Artificial Intelligence**
Published An approach to discover keywords from tweets using Wikipedia.

Conferences and public talks

- November 2019** **TechCrunch Disrupt**
Manila, Philippines Invited to the Startup Battlefield 200.
- November 2019** **Google DGPH**
Manila, Philippines Talk at the keynote track "Turning web data into knowledge".
- June 2019** **Business Information Systems**
Seville, Spain Part of the organising committee of the conference.
- October 2017** **ACM/IEEE Symposium on Edge Computing**
San José, USA Presenting "Extracting web data using representation patterns".
- June 2016** **Practical Applications of Agents and Multi-Agent Systems**
Seville, Spain Presenting "Kizomba: unsupervised heuristic-based extraction".

I+D projects

- 2020 – 2022** **DESK**
39,446 € Ministry of Economic Affairs and Digital Transformation
- 2016 – 2020** **VORTEX**
117,666 € Ministry of Economy and Business
- 2014 – 2017** **ISIDORO**
142,375 € Ministry of Economy and Business
- 2013 – 2015** **PISCIS**
239,342 € 04INNOVA24H, Indra