

# Diego Garcia-Olano

4905 Caswell Avenue, Austin, TX 78751 | 830-421-1805 | diegoolano@gmail.com

[www.diegoolano.com](http://www.diegoolano.com) | [github.com/diegoolano](https://github.com/diegoolano) | [linkedin.com/in/diego-garcia-olano/](https://linkedin.com/in/diego-garcia-olano/) | [twitter.com/dgolano](https://twitter.com/dgolano)

## EDUCATION

**The University of Texas at Austin**, Austin, TX, USA.

9/2017-7/2022

### Ph.D Candidate in Electrical and Computer Engineering

Decision, Information, and Communications Engineering (DICE) Track

**Advisor:** Dr. Joydeep Ghosh. Intelligent Data Exploration & Analysis Lab. [ideal.ece.utexas.edu](http://ideal.ece.utexas.edu)

**Dissertation:** "In-process Diagnostic methods for Entity Representation Learning on Sequential data at Scale"  
Committee: Alex Dimakis, Harris Vikalo, Atlas Wang & Byron Wallace (external - Northeastern University)

**Universitat Politècnica de Catalunya** - Barcelona, Spain.

Graduated: 7/2015

### Masters in Computer Science - Data Science

**Advisors:** Marta Arias, Josep Lluís Larriba Pey

**Research Groups:** Laboratory for Relational Algorithmics, Complexity and Learning ( LARCA ) / DAMA

**Thesis:** "Automated Construction & Analysis of Political Networks". [whoyouelect.com/texas](http://whoyouelect.com/texas)

**The University of Texas at Austin** Austin, TX, USA.

Graduated: 5/2004

**Computer Science** (B.A., 05/2004) (3.86 GPA of 4) Graduated with Honors

**Government** (B.A., 05/2005) (4.0 GPA) with minor in: Business

**Hispanic Studies** (B.A., 05/2004) Graduated with Honors

**Universidad de Sevilla**, College of Arts & Letters (Filología) Sevilla, Spain;

Attended: Fall 2002

**Institut Catholique de Toulouse**, Institut de Langue, Toulouse, France;

Attended: 03/2012 - 06/2013

## RESEARCH INTERESTS

Explainable ML/NLP with a focus on multimodal entity learning for use in health, policy and knowledge based tasks.

## PUBLICATIONS

- Fulton Wang, Julius Adebayo, Sarah Tan, **Diego Garcia-Olano**, Narine Kokhlikyan. "Error Discovery by Clustering Influence Embeddings". ICLR 2023. Oral paper in Trustworthy ML workshop.
- **Diego Garcia-Olano**, Yasumasa Onoe, Joydeep Ghosh, Byron C. Wallace "Intermediate Entity-based Sparse Interpretable Representation Learning". EMNLP 2022. Blackbox NLP workshop
- **Diego Garcia-Olano**, Yasumasa Onoe and Joydeep Ghosh. "Improving and Diagnosing Knowledge-Based Visual Question Answering via Entity Enhanced Knowledge Injection". ACM WWW 2022. Workshop on Multimodal Understanding for the Web and Social Media. [arxiv.org/abs/2112.06888](https://arxiv.org/abs/2112.06888)
- **Diego Garcia-Olano**, Yasumasa Onoe, Ioana Baldini, Joydeep Ghosh, Byron Wallace and Kush Varzney. "Biomedical Interpretable Entity Representations". ACL-IJCNLP 2021.
- Daniel Gillick, Sayali Kulkarni, Alessandro Presta, Jason Baldrige, Eugene Le and **Diego Garcia-Olano**. "Learning Dense Representations for Entity Retrieval". CoNLL 2019
- **Diego Garcia-Olano**, Alan Gee, Joydeep Ghosh and David Paydarfar. "Deep Classification of Time-Series Data with Learned Prototype Explanations". ICML 2019 Time Series Workshop ( 4 pages ) [bit.ly/2lIuEK](https://bit.ly/2lIuEK) & IJCAI 2019 Workshop on Knowledge Discovery in Healthcare Data ( 8 page ) [arxiv.org/abs/1904.08935](https://arxiv.org/abs/1904.08935)
- Kris Sankaran, **Diego Garcia-Olano**, et.al, "Applying Machine Learning methods to Enhance the Distribution of Social Services in Mexico", Data Science For Social Good Datafest, 08/2016 and arxiv update 2017, Chicago. Paper: [arxiv.org/abs/1709.05551](https://arxiv.org/abs/1709.05551) , talk at UChicago DataFest: [goo.gl/AG66bg](https://goo.gl/AG66bg) , poster: [goo.gl/Bc4HAF](https://goo.gl/Bc4HAF)
- **Diego Garcia-Olano**, Marta Arias and Josep L Larriba-Pey: "Automated Construction and Analysis of Political Networks via open government and media sources", European Conference on Machine Learning. ECML SoGood workshop, 09/2016, Riva del Garda, Italy. paper: [goo.gl/xgCU81](https://goo.gl/xgCU81)
- Ivan Paz-Ortiz, **Diego Garcia-Olano**, Carlos Gay-García. "Term-frequency Inverse Document Frequency for the Assessment of Similarity in Central and State Climate Change Programs: An Example for Mexico", Simultech 2015. Special Session on Applications of Modeling and Simulation to Climatic Change and Environmental

Sciences. 06/2015, Colmar, France. [diegoolano.com/ivanpaz](http://diegoolano.com/ivanpaz)

## COMMUNITY INVOLVEMENT

- NeurIPS 2023, 2022, 2021 reviewer,
- ACL 2023 reviewer, EMNLP 2023, ACL-IJCNLP 2021 reviewer, ACL Rolling Reviewer
- ICLR 2023, 2022, 2021, 2020 reviewer ( Top Reviewer Award ), ICLR AI for Social Good 2019 reviewer
- ICML 2021, 2020 reviewer, Texas Linguistics Summit Conference 2019 reviewer
- Data Science For Social Good - Project scoping and Candidate interviewer: 2017, 2018

## SELECTED PROJECTS & PRESENTATIONS

- Leveraging State-of-the-Art Unsupervised Style Transfer Models for Counterfactual Text Generation. 5/2020. RL, Retrieval & Distengagled learning methods, Computational Linguistics Seminar project. [bit.ly/3DLZFZx](http://bit.ly/3DLZFZx)
- Explainability methods for NLP talk at Cognitive Scale 2/2020. [bit.ly/3FYbeh5](http://bit.ly/3FYbeh5) , [bit.ly/3l31Al6](http://bit.ly/3l31Al6)
- Link Detection in Political Networks using GCNs, 12/2018. NLP Final Project [bit.ly/3CQHvnX](http://bit.ly/3CQHvnX)
- Lecture on Modern Visualization, UT Austin, 11/2020. slides: [goo.gl/R6PVzF](http://goo.gl/R6PVzF)
- Network Science Institute at Northeastern University, 10/2016: "Whoyouelect.com" tech talk. [goo.gl/F1dNcv](http://goo.gl/F1dNcv)
- Aquavalens Project meeting, Alicante, Spain. 11/2014. "WP5: Preliminary analysis of the Aquavalens matrix with ICHNAEA® software for Microbial Source Tracking". <http://bit.ly/2eDp9ga>
- "Understanding US counties: Clustering and Classification over voting, socio-economic and public health", 4/2014 Predicting Blue Islands in the US. Paper/Code/Visualization: [diegoolano.com/electionmap](http://diegoolano.com/electionmap)

## PROGRAMMING SKILLS

Highly skilled with Python, PyTorch, Tensorflow, Huggingface, unix, d3, R, mysql/postgres, git, js.

Experience with: keras, postgis/qgis, spark, java, c++, nodejs, matlab, apache/nginx [www.github.com/diegoolano](http://www.github.com/diegoolano)

## EMPLOYMENT EXPERIENCE

**MetaAI - Responsible AI** Research Scientist **9/2022 - Present**  
Research on fair, transparent and guidable AI systems for social impact. Currently: label noise and influence functions

**Fulbright - Azul, Argentina** Fulbright Specialist **3/2022 - 4/2022**  
Lead development and research on a tool for predicting gender bias in judicial proceedings using NLP. Hosted and taught a workshop (in spanish) to faculty/researchers in the Law & Computer Science departments of UNICEN.

**IBM Research- NY** IBM Science for Social Good PhD Fellow **5/2020 - 8/2020**  
Applied research on Biomedical Interpretable Entity Representations For Drug Repurposing using PubMed articles and links between UMLS and Wiki for learning word embeddings for tasks. Host: Kush Varshney, Co-host: Ioana Baldini

**Google AI - Seattle** Software Engineering Research Intern **5/2019 - 8/2019**  
Conducted applied research on explaining seq2seq models using feature attribution methods on Transformer and LSTM based architectures with attention for machine translation. Host: Besim Avci, Co-host: Frederick Liu

**Google AI - Mountain View** Software Engineering Research Intern **5/2018 - 8/2018**  
Conducted applied research on Entity Linking task, built upon internal framework and worked on reproducing findings for OpenAI Deep Type paper. Work from project published in CoNLL 19. Host: Jason Baldridge, Co-host: Daniel Gillick

**University of Texas at Austin - Intelligent Data Exploration & Analysis Lab** Teaching Assistant **9/2017 -7/2022**  
TA for Fall 17 Advanced Predictive Modeling course on Machine Learning with Python.  
TA for Spring 19 Responsible AI Graduate Engineering seminar course. Syllabus: [www.github.com/ideal-ut/RAI-course/](http://www.github.com/ideal-ut/RAI-course/)  
TA for Spring 19 Manoj Saxena's Ethics in AI Design short course

**Eric & Wendy Schmidt Data Science For Social Good at the University of Chicago** DSSG Fellow **5/2016 - 9/2016**  
Worked with SEDESOL of Mexico to improve its allocation of social service benefits for 80 million people. Developed predictive models to estimate incomes, needs, and likelihood of eligibility during recertification periods. Received training in machine learning for social projects by the Chief Scientist of Barack Obama's 2012 election campaign.

**Diego Garcia-Olano, www.diegoolano.com** Freelance Developer, Data Scientist and Data Visualizer **9/2011 - 9/2017**  
Principal architect/developer for clients including Glasstire, Spotify, Pitchfork, Boxee TV, Switch Energy Project, etc.  
Frontend/backend/analytics (mostly python/php, d3, javascript, unix, apache/nginx, mysql/postgres/mongodb)

**Freelancers Union** - Brooklyn, New York, Software Developer **4/2009 - 12/2009, 5/2010 - 9/2011**  
Principal developer of Political Action Committee Fundraising for State and Federal PAC's, Client Scorecard, and The  
Freelance Life applications used by over 120,000 members (django, jquery, postgres ).

**Sapling Systems - Science Technologies;** Austin, Texas, Developer & Adaptive Learning Researcher **2/2005 - 2/2009**  
Principal developer of the IBIS educational platform used by 35,000 students, saplinglearning.com, Provided research  
and working prototypes on Item Response Theory models for feedback functionality.

**Growth Accelerated Partners;** Buenos Aires, Argentina; Lead Software Developer **11/2007 - 11/2008**  
Principal developer of Label Designer for personalwine.com. Oversaw team of developers in Costa Rica

**Caritas of Austin;** Austin, Texas; Lead Developer **3/2006 - 3/2008**  
Developed app for a refugee employment program to digitize employment search data, allow better data sharing of job  
opportunities and facilitate job matches with potential employers. Developed time-tracking system for case workers to  
provide greater accountability and transparency of donors funds.

**Department of State,** US Embassy, Caracas, Venezuela; Political Affairs Intern **9/2004 - 1/2005**  
Maintained and updated the Embassy Caracas SIPRANET transferring sensitive cables  
Observed Regional Elections in twelve voting centers around the greater Caracas area

**Ford Motor Company** Global Manufacturing Supply Chain; Dearborn, Michigan; Developer **5/2003 - 8/2003**  
Developed project to simulate daily production levels for Ford's manufacturing plant in Belgium (VB, Java )  
Provided analysis, infrastructure support, software testing, and acceptance processes on vendor product.

Global Consumer Systems; Dearborn, Michigan; IT & eBusiness Intern **5/2002 - 8/2002**  
Lead various benchmark studies comparing Ford IT capabilities with OEM's.

**UT Learning Center;** Austin, Texas; Tutor, **1/2002 - 7/2004**  
Tutored students on a weekly basis in computer science, logic, math, spanish, accounting, and chemistry

#### **VOLUNTEERING AND SOCIAL PROJECTS**

- **Brandon Dudley campaign for Harris County Tax Assessor/Voter Registrar** **1/2016 - 4/2016**  
Created digital tools to assist the campaign's volunteers to get out the vote.
- **ACLU Immigration Policy Center** Criminal Alien Program Database and Data Visualization; **1/2008 - 1/2009**  
Developed system to parse, digitize and validate prisoner records from Texas counties between 2000 and 2008  
and built a visualization tool to explore the data. The following policy paper is based on it: <http://goo.gl/QItf8>
- **Relief Oversight** - Helped gather a team and develop a publicly generated, app to monitor the activities and  
effectiveness of organizations soliciting donations for relief after an earthquake in Haiti. **12/2009 - 2/2010**
- **Volunteer Healthcare Clinic;** Austin, Texas, Spanish Interpreter & Aide **5/2008 - 8/2008**
- **El Buen Samaritano;** Austin, Texas, ESL Teacher. **9/2003 - 5/2005**  
Taught English night courses for a non-profit provider of working-poor, Spanish-speaking families.

#### **LANGUAGE PROFICIENCIES**

Fluent Spanish (written & spoken)  
French ( Proficient - C1 level as defined by CEFR )

#### **AWARDS & AFFILIATIONS**

The Charles W. and Margaret A. Tolbert Endowed Scholarship in Electrical and Computer Engineering (2019 and 2021)  
Eric & Wendy Schmidt Data Science For Social Good Fellow  
HSF / Ford Corporate Scholarship Recipient & College of Natural Sciences College Scholar