

# Maria De-Arteaga

---

## CONTACT INFORMATION

Pittsburgh, PA

mdeartea@andrew.cmu.edu  
mariadearteaga.com

## RESEARCH OVERVIEW

My research focuses on issues of fairness, accountability and sustainability when using machine learning (ML) to support human decision-making. I characterize the limits and risks of using ML in these contexts, and develop human-centered ML that can improve expert decision-making.

## EDUCATION

**Carnegie Mellon University** Pittsburgh, PA, USA

*Joint Ph.D. Candidate in Machine Learning & Public Policy*

**Aug 2014 – Present**

*Machine Learning Department and Heinz College of Information Systems and Public Policy*

**Carnegie Mellon University** Pittsburgh, PA, USA

*Master of Science, Machine Learning*

**Aug 2014 – May 2017**

*Machine Learning Department*

**Universidad Nacional de Colombia** Bogota, Colombia

*Bachelor of Science, Mathematics*

**Aug 2008 – Aug 2013**

**Universidad Autónoma de Madrid** Madrid, Spain

*Exchange student*

**Jan 2012 – Dec 2012**

## RESEARCH EXPERIENCE

**Auton Lab, Carnegie Mellon University** Pittsburgh, PA, USA

*PhD Student*

**Aug 2014 – Present**

- Co-advisors: Artur Dubrawski and Alexandra Chouldechova.
- Characterize risks of algorithmic bias, and develop machine learning methodology to mitigate biases and to assist expert decision-making.

**Microsoft Research** Cambridge, MA, USA

*Research intern*

**May 2018 – August 2018**

- Advisors: Adam Kalai (primary advisor), Hanna Wallach, Jennifer Chayes and Christian Borgs.
- Conducted a large-scale study of gender bias in automated recruiting, and developed an algorithm for unsupervised bias enumeration in word embeddings.

**Microsoft Research** Redmond, WA, USA

*Research intern*

**May 2017 – August 2017**

- Healthcare NeXT's Clinical Sensing and Analytics Group. Advisor: Sumit Basu.
- Proposed an approach to gain epidemiological insights from unstructured medical records using distance metric learning.

**University of Applied Sciences of Western Switzerland HES-SO** Sierre, Switzerland

*Research intern*

**Jun 2014 – Aug 2014**

- MedGIFT research group. Advisor: Henning Müller.
- Analyzed search behavior of radiologists using specialized search engines and applied supervised machine learning algorithms to improve search results.

**Quantil S.A.S.** Bogota, Colombia

*Data mining researcher*

**Aug 2013 – Jun 2014**

- Applied machine learning and data mining methodologies to assist high-stakes decision-making and oversight in the Colombian government, including the Finance and Health Ministries.

**Computational Linguistics Lab, Universidad Nacional de Colombia** Bogota, Colombia

*Research group member*

**Jan 2013 – Aug 2013**

- Analyzed the predictability of authors' demographics using corpus statistics, lexicons and stylistic features across languages.

SELECTED  
PUBLICATIONS

**M. De-Arteaga\***, R. Fogliato\*, and A. Chouldechova. *A case for humans-in-the-loop: decisions in the presence of erroneous algorithmic scores*. In Proceedings of ACM CHI Conference on Human Factors in Computing Systems (CHI), 2020. [To Appear] (\*Indicates equal contribution)

A. Romanov, **M. De-Arteaga**, H. Wallach, J. Chayes, C. Borgs, A. Chouldechova, S. Geyik, K. Kenthapadi, A. Rumshisky, A. Kalai, *Whats in a name? Reducing bias in bios without access to protected attributes*, In Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2019. **Best Thematic Paper**.

**M. De-Arteaga**, A. Romanov, H. Wallach, J. Chayes, C. Borgs, A. Chouldechova, S. Geyik, K. Kenthapadi, A. Kalai, *Bias in Bios: A Case Study of Semantic Representation Bias in a High-Stakes Setting*, In Proceedings of the ACM Conference on Fairness, Accountability, and Transparency (FAT\*), 2019.

N. Swinger\*, **M. De-Arteaga\***, N. Heffernan IV, M. Leiserson, A. Kalai, *What are the biases in my word embedding?*, In Proceedings of the AAAI/ACM Conference on Artificial Intelligence, Ethics, and Society (AIES), 2019. (\*Indicates equal contribution)

**M. De-Arteaga**, J. Chen, P. Huggins, J. Elmer, G. Clermont, A. Dubrawski, *Predicting Neurological Recovery with Canonical Autocorrelation Embeddings*, PLoS ONE, 2019.

**M. De-Arteaga**, A. Dubrawski, A. Chouldechova, *Learning under selective labels in the presence of expert consistency*, Workshop on Fairness, Accountability, and Transparency in Machine Learning (FAT/ML), ICML, 2018.

**M. De-Arteaga**, W. Herlands, D. Neill, A. Dubrawski, *Machine Learning for the Developing World*, ACM Transactions on Management Information Systems, 2018.

**M. De-Arteaga**, A. Dubrawski, *Discovery of Complex Anomalous Patterns of Sexual Violence in El Salvador*, Data for Policy, 2016. **1st Place Innovation Award on Data Science**.

W. Herlands, **M. De-Arteaga**, D. Neill, A. Dubrawski, *Lass0: Sparse Non-Convex Regression by Local Search*, Workshop on Optimization for Machine Learning, NIPS, 2015.

**M. De-Arteaga**, I. Eggel, C. Kahn, H. Müller, *Analyzing Image Search Behaviour of Radiologists: Semantics and Prediction of Query Results*, Journal of Digital Imaging, 2015.

**M. De-Arteaga**, I. Eggel, B. Do, D. Rubin, C. Kahn Jr. H. Müller, *Comparing Image Search Behaviour in the ARRS GoldMiner Search Engine and a Clinical PACS/RIS*, Journal of Biomedical Informatics, 2015.

AWARDS

- *EECS Rising Stars, 2019.*
- *Best Thematic Paper Award, Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2019.*
- *Microsoft Research Dissertation Grant, 2018.*
- *1st Place Innovation Award on Data Science, Data for Policy, 2016.*
- *Best Student Presentation Award, Domestic Nuclear Detection Office (DNDO) Academic Research Initiative Grantees Conference, 2015.*
- *2nd place Premio ¡Investiga! 2019, an award for Colombian investigative journalism.*
- *Grado de Honor, Universidad Nacional de Colombia, 2013.*

TEACHING  
EXPERIENCE

**Intelligent Systems Architectures, Universidad de los Andes Bogota, Colombia**

*Visiting Instructor*

**Summer 2017**

**Data Mining (95-791), Carnegie Mellon University Pittsburgh, PA, United States**

*Teaching assistant*

**Fall 2016**

**Machine Learning (10-601), Carnegie Mellon University Pittsburgh, PA, United States**

*Teaching assistant*

**Spring 2016**

**Colombian Mathematical Olympiad (OCM) Bogota, Colombia**

*Instructor (number theory and combinatorics)*

**January 2010 – August 2010**

**Early SAS** Bogota, Colombia

*Co-founder*

**April 2017–Present**

- Social entrepreneurship focused on data-driven investigative journalism in Colombia.
- Example project: Data-driven journalistic investigation into killings of social leaders in Colombia. Joint work with CONNECTAS and El País de Cali (link).

**Connectas** Bogota, Colombia

*Investigative journalist*

**March 2012 – June 2012**

- In-depth research for the journalistic investigation “The Jungle Highway”.

**Publicaciones Semana - Especiales Regionales** Bogota, Colombia

*Journalist*

**August 2010 – December 2011**

- Reporting and writing mainly focused on agriculture and environment for *Semana*, a major opinion and analysis magazine in Colombia.

INVITED TALKS

*Will machine learning improve or bias expert decision-making?* Software Engineering Institute, 2019.

*Inteligencia Artificial Responsable*, Feria Digitech (Bogota, Colombia), 2019.

*What are the biases in my data and why do they matter?* IBM Research Africa, 2019.

*Bias in bios: fairness in a high-stakes machine-learning setting*, Computational Ethics in NLP course (CMU), 2019.

*Fairness In Prediction Models Used By Public Service Agencies*, Institute for Operations Research and the Management Sciences Annual Meeting (INFORMS), 2018.

*Guiding public service agencies with machine learning: Opportunities and risks*, Special Session on Social Change In and Through Mathematics and Education, AMS Sectional Meeting, 2018.

*Challenges of algorithmic decisions with humans in the loop*, Google Fairness in ML Workshop, 2018.

*Machine learning, sexual violence crimes and decision-support systems*, Office of the Attorney General of Colombia, 2018.

*Machine Learning for the Developing World*, Data science and machine learning for development and humanitarian action session, UNESCO Tech4Dev Conference, 2018. **Closing Keynote**.

*Machine Learning for the Developing World*, École polytechnique fédérale de Lausanne (EPFL), 2018.

*Using expert consensus to learn under selective labels*, Google Women in Tech Summit, 2018.

*Challenges of child maltreatment predictive models*, Allegheny County Department of Human Service, 2017.

SELECTED PRESS

*Consulted as expert*

- The New York Times - *Many Facial-Recognition Systems Are Biased, Says U.S. Study*.
- New Scientist - *All-Seeing Eyes: The Epic Plan to Track Almost Everything on Earth*.

*About my work*

- Motherboard (Vice) - *Algorithms Have Nearly Mastered Human Language. Why Cant They Stop Being Sexist?*
- VentureBeat - *Researchers Expose Biases in Datasets Used to Train AI Models*.
- Microsoft Research blog - *What are the Biases in my Data?*
- Microsoft Research blog - *Microsoft Research to Present Latest Findings on Fairness in Socio-technical Systems at FAT\* 2019*.
- SciDev - *Machine Learning Targets Development Problems*.
- ZettaBytes (EPFL) - *Principles for Effective Machine Learning for Developing Countries*.
- Quartz - *The most Impactful AI Research isnt Happening in Silicon Valley*. (Features ML4D Workshop)
- Connectas - *Patrones matemáticos que ayudan a hacer mejor periodismo*.

PROFESSIONAL  
SERVICE

*Organizer*

- Co-founder and co-organizer, Workshop series *Machine Learning for the Developing World (ML4D)* (NeurIPS 2017-2019).
- Co-organizer and program lead, CVPR Workshop *Computer Vision for Global Challenges* (2019).
- Co-founder and co-organizer FAT\* reading group at Carnegie Mellon University.
- Co-founder and co-organizer Machine Learning for the Developing World (ML4D) reading group at Carnegie Mellon University.

*Program committee*

AIES 2020, FAT\* 2019.

*Reviewer*

Journal of Digital Imaging (2016-2018), WiML NeurIPS 2017.

PROGRAMMING

Python, R, Matlab, L<sup>A</sup>T<sub>E</sub>X.

REFERENCES

Available upon request.