

Maria De-Arteaga

CONTACT INFORMATION	Austin, TX	dearteaga@mcombs.utexas.edu mariadearteaga.com
CURRENT APPOINTMENT	University of Texas at Austin Austin, TX, USA <i>Assistant Professor</i> Information, Risk, and Operations Management Department McCombs School of Business <ul style="list-style-type: none">• Other appointments: Machine Learning Laboratory (core faculty), Good Systems (researcher), Center for Global Business (affiliated faculty).	July 2020 – Present
EDUCATION	Carnegie Mellon University Pittsburgh, PA, USA <i>Ph.D., Machine Learning & Public Policy</i> <ul style="list-style-type: none">• Doctoral Dissertation: <i>Machine Learning in High-Stakes Settings: Risks and Opportunities</i>• Advisors: Alexandra Chouldechova & Artur Dubrawski Carnegie Mellon University Pittsburgh, PA, USA <i>Master of Science, Machine Learning</i> Universidad Nacional de Colombia Bogota, Colombia <i>Bachelor of Science, Mathematics</i> Universidad Autónoma de Madrid Madrid, Spain <i>Exchange student</i>	Aug 2014 – May 2020 Aug 2014 – May 2017 Aug 2008 – Aug 2013 Jan 2012 – Dec 2012
PUBLICATIONS	h-index: 18. i-10 index: 23. Citations: 1700+. <i>Note on authorship: As faculty, I tend to follow Business School convention, which does not give special meaning to the “last author”, and orders according to contribution.</i> <i>* indicates equal contribution, ° indicates student author.</i>	
JOURNAL ARTICLES	<ul style="list-style-type: none">[1] M. De-Arteaga, V. Jeanselme°, A. Dubrawski, A. Chouldechova. <i>Leveraging Expert Consistency to Improve Algorithmic Decision Support</i>. Management Science, forthcoming.[2] S. Gupta°, V.Kovatchev, A.Das°, M. De-Arteaga, M. Lease. Finding Pareto trade-offs in fair and accurate detection of toxic speech. Information Research an International Electronic Journal, (iConf), 2025.[3] Y. Li°, M. De-Arteaga, M Saar-Tsechansky. <i>Label Bias: A Pervasive and Invisibilized Problem</i>. Notices of the American Mathematical Society, 2024.[4] M. Cheng°, M. De-Arteaga, L. Mackey, A. Kalai. <i>Social Norm Bias: Residual Harms of Fairness-Aware Algorithms</i>. Data Mining & Knowledge Discovery, 2023.[5] J. Elmer, M. Kurz, P. Coppler, A. Steinberg, S. DeMasi, M. De-Arteaga, N. Simon, V. Zadorozhny, K. Flickinger, C. Callaway. <i>Time to Awakening and Self-Fulfilling Prophecies After Cardiac Arrest</i>. Critical Care Medicine, 2023.[6] M. De-Arteaga, S. Feuerriegel, M. Saar-Tsechansky. <i>Algorithmic Fairness in Business Analytics: Directions for Research and Practice</i>. Production & Operations Management, 2022.[7] M. De-Arteaga, J. Elmer. <i>Self-fulfilling prophecies and machine learning in resuscitation science</i>. Resuscitation, 2022.[8] S. Fazelpour, M. De-Arteaga. <i>Diversity in Sociotechnical Machine Learning Systems</i>. Big Data & Society, 2022.[9] V. Jeanselme°, M. De-Arteaga, J. Elmer, S. M.Perman, A. W. Dubrawski. <i>Sex differences in post cardiac arrest discharge locations</i>, Resuscitation Plus, 2021.[10] M. De-Arteaga, J.Chen, P. Huggins, J. Elmer, G. Clermont, A. Dubrawski, <i>Predicting Neurological Recovery with Canonical Autocorrelation Embeddings</i>, PLoS ONE, 2019.	

- [11] **M. De-Arteaga**, W. Herlands, D. Neill, A. Dubrawski, *Machine Learning for the Developing World*, ACM Transactions on Management Information Systems, 2018.
- [12] **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.
- [13] **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.
- [14] L. Deck^o, J. Schoeffer, **M. De-Arteaga**, N. Kuehl. *A Critical Survey on Fairness Benefits of Explainable AI*. In Proceedings of the ACM Conference on Fairness, Accountability, and Transparency (FAccT), 2024.
- [15] J. Schoeffer^o, **M. De-Arteaga***, N. Kuehl*. *On Explanations, Fairness, and Appropriate Reliance in Human-AI Decision-Making*. Accepted to ACM CHI Conference on Human Factors in Computing Systems (CHI), 2024. **Honorable Mention Award (top 5% submissions)**.
- [16] S. Gupta^o, S. Lee^o, **M. De-Arteaga**, M. Lease. *Same Same, But Different: Conditional Multi-Task Learning for Demographic-Specific Toxicity Detection*. In Proceedings of the ACM Web Conference (WWW), 2023.
- [17] **M. De-Arteaga***, K. Holstein*, Lakshmi Tumati^o, Yanghui Cheng^o. *Toward Supporting Perceptual Complementarity in Human-AI Collaboration via Reflection on Unobservables*. In Proceedings of the ACM on Human-Computer Interaction (CSCW), 2023.
- [18] V. Jeanselme^o, **M. De-Arteaga***, Z. Zhang*, J. Barrett, B. Tom. *Imputation Strategies Under Clinical Presence: Impact on Algorithmic Fairness*. In Proceedings of the 2nd Machine Learning for Health symposium, PMLR, (ML4H) 2022.
- [19] Y. Li^o, **M. De-Arteaga**, M. Saar-Tsechansky. *More Data Can Lead Us Astray: Active Data Acquisition in the Presence of Label Bias*. In Proceedings of the AAAI Conference on Human Computation and Crowdsourcing (HCOMP), 2022.
- [20] T. Neumann^o, **M. De-Arteaga**, S. Fazelpour. *Justice in Misinformation Detection Systems: An Analysis of Algorithms, Stakeholders, and Potential Harms*, In Proceedings of the ACM Conference on Fairness, Accountability, and Transparency (FAccT), 2022.
- [21] R. Gao^o, M. Saar-Tsechansky, **M. De-Arteaga**, L. Han^o, M. Kyung Lee, M. Lease. *Human-AI Collaboration with Bandit Feedback*. In Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), 2021.
- [22] N. Akpınar^o, **M. De-Arteaga**, A. Chouldechova. *The effect of differential victim crime reporting on predictive policing systems* In Proceedings of the ACM Conference on Fairness, Accountability, and Transparency (FAccT), 2021.
- [23] **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.
- [24] A. Romanov, **M. De-Arteaga**, H. Wallach, J. Chayes, C. Borgs, A. Chouldechova, S. Geyik, K. Kenthapadi, A. Rumshisky, A. Kalai, *What's 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 Award**.
- [25] **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 (FAccT), 2019.
- [26] 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.
- [27] **M. De-Arteaga***, R. Fogliato*, and A. Chouldechova. *Human Discernment of Algorithmic Errors: A Case Study in Child Welfare*. Under 2nd round review at Management Information Systems Quarterly (MISQ).

- [28] V. Jeanselme[°], **M. De-Arteaga**^{*}, Z. Zhang^{*}, J. Barrett, B. Tom. *Imputation Strategies Under Clinical Presence: Impact on Algorithmic Fairness*. Under 2nd round review at Management Science.
- [29] R. Gao[°], M Saar-Tsechansky, **M. De-Arteaga**. *Learning Complementary Policies for Human-AI Teams*. Under 2nd round review at Management Science.
- [30] **M. De-Arteaga**^{*}, K. Holstein^{*}. *Supporting Information Integration in Human-AI Collaboration via Reflection on Unobservables*. Major revision at Management Information Systems Quarterly (MISQ).
- [31] W. Dong[°], **M. De-Arteaga**, M. Saar-Tsechansky. *Using Machine Bias To Measure Human Bias*. Under review at Management Information Systems Quarterly (MISQ).
- [32] J. Schoeffer[°], **M. De-Arteaga**, Niklas Kühl. *Looks Can Mislead: Feature-Based Explanations Do Not Reliably Improve Fairness of AI-Assisted Decisions*. Under review at Management Science.

GRANTS & FUNDING

McCombs Research Excellence Grant: Overcoming Clinical Nihilism via Human-AI Collaboration. \$21,450 USD. 2024. *Principal Investigator*.

NIH R01: Optimizing Recovery Prediction After Cardiac Arrest. \$3,156,222 USD. 2022-2027. *Co-Principal Investigator* (PI: Jonathan Elmer).

UT Austin Machine Learning Laboratory: Learning from Noisy and Biased Labelers towards Accuracy and Fairness goals. \$50,000 USD. 2021. *Co-Principal Investigator* (with Maytal Saar-Tsechansky).

UT Austin Good Systems Grant: Designing Responsible AI Technologies to Curb Disinformation. \$810,000 USD. 2021-2027. *Senior Personnel* (PI: Matt Lease).

Google Award for Inclusion Research: Who is Most Likely to Advocate for this Case? A Machine Learning Approach. \$60,000 USD. 2020. *Principal Investigator*

Microsoft Research Dissertation Grant: Quantifying and Mitigating Risks of Algorithmic Decision Support. \$25,000 USD. 2018. *Principal Investigator*

AWARDS

- *Honorable Mention Award* (top 5% submissions), ACM Conference on Human Factors in Computing Systems (CHI), 2024.
- 2023 Management Science Distinguished Service Award.
- *Best Student Paper Award*, Conference on Information Systems and Technology (CIST), 2022. Student author: IROM PhD student Ruijiang Gao.
- *Best Paper Runner-Up Award*, Workshop on Information Technologies and Systems (WITS), 2021.
- *EECS Rising Stars*, 2019.
- *Best Thematic Paper Award*, Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2019.
- 2nd place *Premio ¡Investiga! 2019*, an award for Colombian investigative journalism.
- *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.
- *Grado de Honor*, Universidad Nacional de Colombia, 2013.

ADVISING EXPERIENCE

Postdoctoral advisor Jakob Schoeffer.

PhD co-advisor Yunyi Li (IROM), Terrence Neumann (IROM), Soumyajit Gupta (CS).

PhD committee member John-Patrick Akinyemi (IROM), Nicholas Wolczynski (IROM), Eye Bayiz (ECE), Ruijiang Gao (IROM grad 2024, *Assistant Prof. at UT Dallas Bschoool*), Wanxue Dong (IROM grad 2023, *Assistant Prof. at CUHK Bschoool*).

Undergraduate thesis advisor Jennifer Mickel (UT Polymathic Scholars Honors Program), Riya Cyriac (Plan II Thesis, McCombs).

TEACHING
EXPERIENCE

University of Texas at Austin Austin, TX, USA **Spring'24**

Instructor, Ethics of AI (MIS 285N)

- Development of new course for Master of Science in Business Analytics.

University of Texas at Austin Austin, TX, USA **Spring'24**

Instructor, Ethics of AI (BAX 372 / MIS 373)

- Development of new course for undergraduate in Business Analytics and undergraduate in Management Information Systems.

University of Texas at Austin Austin, TX, USA **Spring'21, Fall'21, Fall'22**

Instructor, Introduction to Programming and Problem Solving (Undergraduate, MIS-304)

- Taught online (Spring'21), hybrid (Fall'21), and in-person (Fall'22).
- Development of flipped classroom strategy and material.

Universidad de los Andes Bogota, Colombia **Summer 2017**

Visiting Instructor, Intelligent Systems Architectures

Carnegie Mellon University Pittsburgh, PA, United States **Fall 2016**

Teaching Assistant, Data Mining (95-791)

Carnegie Mellon University Pittsburgh, PA, United States **Spring 2016**

Teaching Assistant, Machine Learning (10-601)

Colombian Mathematical Olympiad (OCM) Bogota, Colombia **Jan 2010 – Aug 2010**

Instructor (*number theory and combinatorics*)

WORK
EXPERIENCE

Early SAS Bogota, Colombia **Apr 2017 – 2020**

Co-founder & CTO

Social entrepreneurship focused on data-driven investigative journalism in Colombia.

Microsoft Research Cambridge, MA, USA **May 2018 – Aug 2018**

Research intern

Microsoft Research Redmond, WA, USA **May 2017 – Aug 2017**

Research intern

Haute École Spécialisée de Suisse Occidentale Sierre, Switzerland **Jun 2014 – Aug 2014**

Research intern

Quantil SAS Bogota, Colombia **Aug 2013 – Jun 2014**

Junior researcher

Connectas Bogota, Colombia **Mar 2012 – Jun 2012**

Investigative journalist

Publicaciones Semana - Especiales Regionales Bogota, Colombia **Aug 2010 – Dec 2011**

Journalist

INVITED TALKS
& PANELS

2025

University of Minnesota (Carlson School of Management, Information & Decision Sciences, seminar), Harvard Business School (Technology & Operations Management, seminar).

2024

International Symposium on Trustworthy Autonomous Systems (*Keynote*), U. Maryland (Values-Centered Artificial Intelligence (VCAI) colloquium), Annual Meeting of the Society for Business Ethics (Algorithmic Fairness and Business Ethics panel), Johns Hopkins AI Ethics and Governance Symposium.

2023

Stanford University (Graduate School of Business, OIT seminar), MIT (Sloan School of Management, OM seminar), Carnegie Mellon University (Tepper School of Business, seminar), University of California San Diego (Halicioğlu Data Science Institute, seminar), University of Sheffield (Speech and Language Technologies, seminar), UT Austin (Operations Research and Industrial Engineering, seminar), University of British Columbia (seminar), AI UK Conference (invited panelist), University of Chicago (CHAI research group, seminar), Good Systems Symposium (research to practice roundtable), Berkeley Simons Institute Toc4fairness series (invited talk), Berkeley Simons Institute workshop (invited talk), Colombian Mathematical Congress (*Keynote*).

2022

Harvard (*Keynote*, HDSI Annual Conference), LMU Munich (Keynote Series, Research in AI and Management), MIT (Guest lecturer, Ethical ML in Human Deployments course), Computing Community Consortium AI/OR Workshop, Penn State (AI for Social Impact Seminar Series), UW Madison (Statistics Dept. Seminar), McCombs Alumni (Knowledge to go Webinar), University of Memphis (Cognitive Science Seminar), Sociedad Colombiana de Matemáticas (*Keynote*). Texas State University (Math & Stats Awareness Series), Women in Data Science (WIDS) Charlottesville, CHI TRAIT Workshop Panel, CRA Snowbird Conference (From Fairness to Responsibility Panel).

2021

AI for People Conference (*Keynote*), Johns Hopkins University (MINDS & CIS Seminar Series), MSRI (Workshop on Mathematics and Racial Justice), University of Edinburgh (AI Ethics and Society seminar), University of Southern California (CAIS Symposium on Equity & AI), UT Austin (MLL & Good Systems seminar series on Ethics in AI), California State University Long Beach (Mathematics and Statistics colloquium), Rice University (ECE Data Science Seminar), UT Austin (McCombs Knowledge to go Webinar).

2020

NeurIPS (ML Retrospectives Workshop), UMass Amherst (Machine Learning & Friends Seminar), MIT Lincoln Laboratory (*Keynote*, Hispanic Heritage Month), Facebook (*Keynote*, Responsible NLP Summit), WIDS Bogota (*Keynote*), Universidad Javeriana (*Keynote*, MSc Data Science Inaugural Event), INFORMS en Español, X4D Seminar Series, Universidad del Rosario (Tic Tank seminar), UT Austin Conference on the Ethics of AI, Australian National University (Data, AI, and Society Seminar Series), FUDAPROMAT Panama, University of San Francisco (Data Ethics Seminar, cancelled due to COVID-19), Cornell Information Science (Department Colloquia), Cornell Policy Analysis & Management (Department Seminar), UC Santa Barbara (CS Department Seminar), Microsoft Research New England (MSR Seminar).

2019

UT Austin (IROM Department Seminar), University of Pittsburgh (Katz Department Seminar), CMU Software Engineering Institute (Lecture Series), Feria Digitech, IBM Research Africa (Africa Labs Deep Dive), Carnegie Mellon University (*Invited lecturer*, Computational Ethics in NLP course).

2018

INFORMS (Fairness and Transparency in Applied Operations Session), AMS Sectional Meeting (Special Session on Social Change In and Through Mathematics and Education), Google (Fairness in ML Workshop), Office of the Attorney General of Colombia (Lecture Series), UNESCO (*Closing Keynote*, Tech4Dev Conference, Data Science and Machine Learning for Development and Humanitarian Action Session), EPFL (CS Department Seminar), Google (Women in Tech Summit).

2017

Allegheny County Department of Human Service (DHS Speaker Series).

SELECTED PRESS *About my work*

- UT News - *Algorithms That Adjust for Worker Race, Gender Still Show Biases*
- McCombs Big Ideas - *Algorithms for Hiring: Bias In, Bias Out*
- MIT Technology Review - *Predictive policing is still racist—whatever data it uses*
- El Pais - *Algoritmos de predicción policial: para qué se usan y por qué se ensañan con los más pobres*

- Motherboard (Vice) - *Algorithms Have Nearly Mastered Human Language. Why Can't 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 isn't Happening in Silicon Valley.* (Features ML4D Workshop)
- Connectas - *Patrones matemáticos que ayudan a hacer mejor periodismo.*

Consulted as expert

- Fast Company - *Are you AI literate? Schools and jobs are insisting on it—and now it's EU law*
- The Wall Street Journal - *If AI Gives You Bad Answers, Tell It to Do Better*
- Quartz - *In a facial recognition bill backed by Microsoft, three words stand between citizens and their civil rights.*
- CNN - *Think your mask makes you invisible to facial recognition? Not so fast, AI companies say.*
- Quartz - *Amazon and Microsoft pause police facial recognition and demand regulation.*
- Quartz - *The influential project that sparked the end of IBM's facial recognition program.*
- 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.*

PROFESSIONAL SERVICE

Conference Leadership

- Executive Committee member, ACM FAccT 2022-2025.
- Diversity, Equity and Inclusion co-Chair, ACM FAccT 2021-2022.
- Organizer, Special Interest Group on Human-Centered Responsible AI, ACM CHI, 2023.
- Organizer, ICML Workshop on Human-Machine Collaboration and Teaming, 2022.
- Local Arrangements co-Chair, WITS 2021.
- Area Chair, ACM EAAMO 2021.
- Steering Committee Chair, ML4D 2020.
- 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, co-organizer: FEAT reading group and ML4D reading group at CMU.

Program committee

ACM COMPASS 2022, ICDM Workshop on Algorithms towards Ethical and Privacy challenges in Social Media Recommendation Systems 2022, MD4SG 2020, AIES 2020, FAT* 2019.

Reviewer

Management Science, Management Information Systems Quarterly, Information Systems Research, Nature Human Behavior, Journal of the Association for Information Systems, Journal of Digital Imaging, Big Data & Society, NeurIPS (Ethics reviewer), CHI, AIES, CSCW, FAccT, WiML NeurIPS, NeurIPS Workshop on Consequential Decision Making in Dynamic Environments, NeurIPS Workshop on Navigating the Broader Impacts of AI Research.

PROGRAMMING

Python, R, Matlab, L^AT_EX.

LANGUAGES

Spanish (native), English (fluent).