Adriana Meza Soria

PhD in Software Engineering | University of California, Irvine | She, her, hers

🎓 Google Scholar   🖥 adriana.mezasoria.com   📬 adriana.meza.soria@ibm.com   🕸 LinkedIn

EDUCATION

- **Ph.D. in Software Engineering**   2017–2022
  University of California, Irvine
  GPA: 3.98 (0-4 scale)
- **M.S in Engineering (Summa Cum Laude)**   2016
  CETYS University, Tijuana, Mexico
  GPA: 100 (0-100 scale)
- **B.S. in Computational Systems Engineering**   2013
  Technological Institute of Tijuana (ITT), Tijuana, Mexico
  GPA: 96.39 (0-100 scale)

PROFESSIONAL EXPERIENCE

- **MIT-IBM Watson AI Lab, AI Models Engineering team,** *Research Engineer*   2023-present
  Synthetic Data Generation, Data Engineering, AI for Code
- **MIT-IBM Watson AI Lab, AI Models Software Group (APT),** *Research Engineer*   2022-2023
  Generative AI and Software Engineering
- **MIT-IBM Watson AI Lab, APT** *Research Intern*   Summer 2021
  Design and prototyping
- **Grupo Tress Internacional (GTI)***, Senior Software Engineer*   2013–2017
  End-user application design, development and modernization
- **IWAI Metal Mexico,** *IT Assistant*   2012–2013
  Internal software development and IT management
- **TELNOR,** *Intern*   2011–2012
  Design and prototyping

TEACHING EXPERIENCE

- **Professor**
  UC Irvine CA, U.S.A   Summer 2020
  Programming in Java as a second language (undergraduate)
- **Teaching Assistant**   Fall 2018–Winter 2021
  UC Irvine CA, U.S.A
- **Professor**   2016–2017
  CETYS University Tijuana, Mexico
- **Professor**   2014–2017
  Autonomous University of Baja California (UABC) Tijuana, Mexico

ACADEMIC SERVICE

- **Program Committee Member**
  International Conference in Software Engineering (ICSE)   2026
  [Tracks: Research]
  International Conference on the Foundations of Software Engineering (FSE)   2025
  [Tracks: Industry track] [Workshops: AI IDE Workshop]
  International Conference in Learning Representations (ICLR)   2025
  [Tracks: Datasets & Benchmarks]
  International Conference in Software Engineering (ICSE)   2025
  [Tracks: Research(shadow), SE in Practice (SEIP), Artifact Evaluation] [Workshops: Designing]
  International Conference on Cooperative and Human Aspects of Software Engineering (CHASE)   2025
  [Tracks: Research]

International Conference in Software Engineering (ICSE)   2024
[Workshops: Designing]
IEEE Software Special Issues   2024
[Creativity in Software Engineering]
International Conference on Cooperative and Human Aspects of Software Engineering (CHASE)   2023
[Tracks: Research]
Mining Software Repositories   2021
[Tracks: Research (shadow)]

- **Organizing Committee Member**
International Conference on Cooperative and Human Aspects of Software Engineering (CHASE)   2024
[Program co-chair of short-paper track]
International Conference on Cooperative and Human Aspects of Software Engineering (CHASE)   2023
[Proceedings chair]

HONORS AND AWARDS

- IBM Tech Recognition Award   2025
- Recipient of Miguel Velez Scholarship (3rd)   2022
- Latino Excellence and Achievement Award   2021
- Grace Hopper Celebration Scholar   2020
- Best product idea & CodePath favorite   2020
- Recipient of Miguel Velez Scholarship (2nd)   2019
- Recipient of Rosalva Gallardo Valencia Graduate Award   2019
- Second place at AMIA Design Challenge   2018
- Recipient of Miguel Velez Scholarship (1st)   2017
- Summa Cum Laude of class 2016 for the M.S in Engineering program   2016

PUBLICATIONS

- API Pack: A Massive Multi-Programming Language Dataset for API Call Generation. Zhen Guo, **Adriana Meza Soria**, Wei Sun, Yikang Shen, Rameswar Panda. *[ICLR 2025, PFD]*

- 2024. Power Scheduler: A Batch Size and Token Number Agnostic Learning Rate Scheduler. Yikang Shen, Matthew Stallone, Mayank Mishra, Gaoyuan Zhang, Shawn Tan, Aditya Prasad, **Adriana Meza Soria**, David D. Cox, Rameswar Panda. *[under submission 2024, PDF]*

- Granite-Function Calling Model: Introducing Function Calling Abilities via Multi-task Learning of Granular Tasks. Ibrahim Abdelaziz, Kinjal Basu, Mayank Agarwal, Sadhana Kumaravel, Matthew Stallone, Rameswar Panda, Yara Rizk and GP Bhargav, Maxwell Crouse, Chulaka Gunasekara, Shajith Ikbal, Sachin Joshi, Hima Karanam, Vineet Kumar, Asim Munawar, Sumit Neelam, Dinesh Raghu, Udit Sharma, **Adriana Meza Soria**, Dheeraj Sreedhar, Praveen Venkateswaran, Merve Unuvar, David Cox, Salim Roukos, Luis Lastras, Pavan Kapanipathi. *[EMNLP 2024, PDF, HF]*

- Granite Code Models: A Family of Open Foundation Models for Code Intelligence. Mayank Mishra, Matt Stallone, Gaoyuan Zhang, Yikang Shen, Aditya Prasad, **Adriana Meza Soria**, Michele Merler, Parameswaran Selvam, Saptha Surendran, Shivdeep Singh, Manish Sethi, Xuan-Hong Dang, Pengyuan Li, Kun-Lung Wu, Syed Zawad, Andrew Coleman, Matthew White, Mark Lewis, Raju Pavuluri, Yan Koyfman, Boris Lublinsky, Maximilien de Bayser, Ibrahim Abdelaziz, Kinjal Basu, Mayank Agarwal, Yi Zhou, Chris Johnson, Aanchal Goyal, Hima Patel, Yousaf Shah, Petros Zerfos, Heiko Ludwig, Asim Munawar, Maxwell Crouse, Pavan Kapanipathi, Shweta Salaria, Bob Calio, Sophia Wen, Seetharami Seelam, Brian Belgodere, Carlos Fonseca, Amith Singhee, Nirmit Desai, David D. Cox, Ruchir Puri, Rameswar Panda. *[Technical report 2024, PDF, HF]*

- Daniel Graziotin, Alexander Nolte, Birgit Penzenstadler, Klaas-Jan Stol, Giuseppe Destefanis, **Adriana Meza Soria**, Silvia Abrahão: Proceedings of the 2024 IEEE/ACM 17th International Conference on Cooperative and Human Aspects of Software Engineering, CHASE 2024, Lisbon, Portugal, April 14-15, 2024. ACM. *[Proceedings edition 2024, PDF]*

- Characterizing Software Maintenance Meetings: Information Shared, Discussion Outcomes, and Information Captured. **Adriana Meza Soria**, Taylor Lopez, Elizabeth Seero, Negin Mashhadi, Emily Evans, Janet Burge, and André Van der Hoek. 2024. *[ICSE 2024, PDF]*

- Exploring a Research Agenda for Design Knowledge Capture in Meetings. L. Seero, J. Burge, **A. M. Soria** and A. Van Der Hoek *[CHASE 2023, PDF]*

- **Adriana Meza Soria**. Understanding How Information Flows In and Out of Regularly Scheduled Software Maintenance Design Meetings: A Case Study. *[Dissertation 2022, PDF]*

- Recurring distributed software maintenance meetings: toward an initial understanding. **Adriana Meza Soria**, André van der Hoek, and Janet Burge. *[CHASE 2022, PDF]*

- Reading to write code: an experience report of a reverse engineering and modeling course. Brooke Ryan, **Adriana Meza Soria**, Kaj Dreef, and André van der Hoek. 2022. *[SEET 2022, PDF]*

- The Design of a Study Concerning the Capture of Important Design Bits at the Whiteboard. **A. M. Soria** and A. Van Der Hoek. *[MODELS-C 2021, PDF]*

- KNOCAP: capturing and delivering important design bits in whiteboard design meetings. **Adriana Meza Soria**. *[ICSE 2020, PDF]*

- Collecting Design Knowledge through Voice Notes. **A. Meza Soria** and A. van der Hoek. *[CHASE 2019, PDF]*

- Toward Collecting and Delivering Knowledge for Software Design at the Whiteboard. **A. Meza Soria** and A. van der Hoek *[CHASE 2018, PDF]*

SELECT PROJECTS

- **Instruction Data Selection (IBM | MIT-IBM AI Watson Lab | AI Models Engineering)**   2025-present
  Synthetic Data Generation with Open Source LLMs and techniques
- **AI SE Agents (IBM | MIT-IBM AI Watson Lab | AI Models Engineering)**   2024-present
  LLM powered agents able to solve complex tasks end-to end (i.e., GitHub issues).
- **Synthetic Data (IBM | MIT-IBM AI Watson Lab | AI Models Engineering)**   2024-present
  Synthetic Data Generation with Open Source LLMs and techniques
- **API Pack (IBM | MIT-IBM AI Watson Lab | AI Models Engineering)**   2023-2025
  A code instruction dataset to improve LLMs ability to generate API calls.
- **Software Maintenance Meetings (UCI | SDCL)**   2020-2022
  Single case study of software development meetings.
- **Internship mini-project (IBM | MIT-IBM AI Watson Lab | APT)**   Summer 2021
  Architecture design and development of a service to leverage ML models for product demand forecasting.
- **KNOCAP (UCI)**   2018-2022
  A suite of tools to collect important design bits from developers' conversations during whiteboard design meetings.
- **Nana Stories (AMIA | Design Competition) – 2nd at AMIA student design competition.**   2018–2019
  An Alexa skill that offers in-home exercises for children who require speech and language therapy.

ACTIVE RESEARCH AREAS

- AI for Software Engineering: AI for Code, AI Software Agents
- Synthetic Data: Tool/API Instruction Data Generation, Agentic Trajectories Generation, Automatic Instruction Data Selection
- Empirical Studies in Software Engineering: Software Meetings

VOLUNTEER WORK

- Mexico Graduate Research Education Program, UC Irvine (member)   2018–present
- I-SURF summer program, UC Irvine (mentor)   2019
- APPCamp summer program, UC Irvine (speaker)   2019
- ExploreCSR workshop (Google sponsored workshop), CSULB and UC Irvine (mentor)   2019

SKILLS

Technologies
- Programming languages: Python, Java, C#, Delphi, JavaScript
- Database: MySQL, SQL Server, Oracle, PostgreSQL, Elastic Search

- IDEs: VS Code, Eclipse, RAD XE5, Android Studio, and XCode, PyCharm
- Data engineering: Jupyter Notebook, Pandas, Matplotlib, Pytorch, Datasets, LiteLLM, Transformers
- Sketching and modeling: Visio, StartUML, Moqups, Figma
- Project management: Trello and Target Process, Github Projects
- Code repositories: GIT, TSF (Microsoft), StarTeam
- Word editors: LATEX, Microsoft Word

Languages
- English (fluent), Spanish (native speaker)

WEB PROFILES

LinkedIn: https://www.linkedin.com/in/adriana-meza-soria-52799961
ResearchGate:https://www.researchgate.net/profile/Adriana-Meza-Soria
GoogleScholar: https://scholar.google.com/citations?user=BpMQCb4AAAAJ&hl=en
Personal Website: https://adriana.mezasoria.com/