

# Adrian Mirza

 GitHub |  LinkedIn |  Google Scholar |  adrianmirza.com |  mail@adrianmirza.com

## SUMMARY

---

PhD researcher specializing in the application of machine learning to chemistry problems. Proven track record of developing and deploying machine learning models for complex challenges in molecular chemistry, spectroscopy, and materials science. Experienced in leading projects from conception to completion, with practical expertise in molecular simulations and large-scale data management gained at Stability AI and Software for Chemistry and Materials.

## WORK EXPERIENCE

---

### Data Research Engineer – Part Time

August 2023 – July 2024

Stability AI

Contributed to Open Science and ChemNLP by collecting and curating chemistry datasets and performing data research.

### Intern in Computational Chemistry and Software Engineering

April 2023 – July 2023

Software for Chemistry and Materials

- Conducted literature reviews to identify relevant KPIs for conformer simulation tools.
- Applied graph theory to chemical problems related to conformers.
- Benchmarked internal conformer generation tools against literature.

### Teaching Assistant – Part Time

August 2022 – July 2023

Delft University of Technology

- Restructured a Master-level course on computational, numerical, and data science methods.
- Prepared tutorials for 140 graduate students.
- Prepared and graded assignments.

### Student Assistant in HR IT – Part Time

June 2020 – July 2021

TotalEnergies

- Analyzed and cleaned employee data.
- Built a Microsoft Office pipeline for data streamlining and analysis.
- Provided IT and HR support.

## EDUCATION

---

2023 – present **PhD in Digital Chemistry**, Helmholtz Zentrum Berlin

*Advisor: Kevin Jablonka*

2021 – 2023 **MSc Chemical Engineering**, Delft University of Technology

*Advisor: Evgeny Pidko (Joint thesis between TU Delft and Johnson & Johnson)*

2018 – 2021 **BSc Chemical Engineering & Biotechnology**, Aalborg University

*Advisor: Marco Maschietti*

## LANGUAGES

---

**Native** – Romanian, **Professional proficiency** – English (C2), Russian (C2), **Limited proficiency** – German (B1)

## SELECTED PUBLICATIONS

---

Mirza, Adrian, Nawaf Alampara, Anagha Aneesh, et al. (2025). “General purpose models for the chemical sciences”. In: *arXiv preprint arXiv:2507.07456*.

Mirza, Adrian, Nawaf Alampara, Sreekanth Kunchapu, et al. (2025). “A framework for evaluating the chemical knowledge and reasoning abilities of large language models against the expertise of chemists”. In: *Nature Chemistry*, pp. 1–8.

Mirza, Adrian, Nawaf Alampara, Martiño Ríos-García, et al. (2025). “ChemPile: A 250GB Diverse and Curated Dataset for Chemical Foundation Models”. In: *Advances in Neural Information Processing Systems 39: Datasets and Benchmarks Track*. URL: <https://openreview.net/forum?id=l8Mz3dcK1S>.

Kalikadien, Adarsh V et al. (2024). “Paving the road towards automated homogeneous catalyst design”. In: *ChemPlusChem* 89.7, e202300702.

Mirza, Adrian and Kevin Maik Jablonka (2024). “Elucidating structures from spectra using multimodal embeddings and discrete optimization”. In: *chemrxiv preprint*. DOI: [10.26434/chemrxiv-2024-f3b18-v2](https://doi.org/10.26434/chemrxiv-2024-f3b18-v2).

## SKILLS

---

<b>Programming Languages</b>	Python (primary), Bash, MatLab, Rust
<b>Scientific Computing and ML</b>	SciPy, NumPy, PyTorch, PyTorch Lightning, Modal, Amsterdam Modelling Suite, Gaussian16, Orca, HuggingFace
<b>Other Tools and Frameworks</b>	ChemDraw, Adobe Illustrator, LaTeX, Git, FastAPI

## CONFERENCES — SUMMITS — MEETINGS

---

2025 NeurIPS 2025: **Poster presentation**  
2025 AI4Mat ICLR Singapore: **Spotlight talk**  
2024 AI4Mat Vienna: **Spotlight talk**  
2024 L2M3 CECAM: **Invited talk**  
2024 HAICON24: **Poster presentation**  
2024 IWOMI: **Invited talk**  
2024 Llama Community Summit: Invited to Meta’s HQ for discussions about applications of AI  
2023 The Netherlands’ Catalysis and Chemistry Conference (NCCC): **Poster presentation**

## COMPETITIONS & AWARDS

---

2025 **Co-Winner** of German Research Foundation Best Environment contest in Germany  
2020 **Top-5** Technical University of Denmark (DTU) Countrywide Student Plastic Project Contest  
2018 **President’s Diploma** awarded to the best graduating high-school students in Moldova  
2018 **Bronze Medal** National Chemistry Olympiad of Moldova (Grade XII)  
2016 **Bronze Medal** National Chemistry Olympiad of Moldova (Grade X)