Dan Graur

Email: dan.ovidiu.graur@gmail.com
Personal Website: dangraur.github.io
LinkedIn: linkedin.com/in/graur-dan/

EXPERIENCE

ETH Zürich - Systems Group   Zürich, Switzerland
Research Assistant / PhD Student   *11/2019-2024*

- Supervised by Prof. Gustavo Alonso and Prof. Ana Klimovic.
- Researching Systems for Machine Learning: focusing on efficient, scalable ML data processing.
- Researching Databases: focusing on query languages and efficient means of processing semi-structured data.
- Published at top conferences such as VLDB, ATC, SoCC; often with industry collaborations (e.g. Google).
- Mentored 8 students for their MSc theses; most awarded maximal grades; some awarded prizes for their work

Google   Sunnyvale, California, USA
Research Intern   *07/2023-11/2023*

- Part of the SystemsResearch@Google group; mentored by Fatma Ozcan and Yannis Chronis.
- Researching novel adaptive indexing methods for improving performance and memory efficiency for distributed cloud databases. Working closely with AlloyDB.

Google   Mountain View, California, USA
Research Intern   *06/2022-09/2022*

- Part of the tf.Data team, where I researched novel ways to process data for large-scale ML jobs.
- I worked extensively with TensorFlow and ML preprocessing logic for production models.
- Published a paper at SoCC'23 around disaggregation for large-scale ML preprocessing.

Google   Zürich, Switzerland
Software Engineering Intern @ Google Brain   *06/2020-09/2020*

- Worked in the Flax team where I implemented RetinaNet. Mentored by Marvin Ritter.
- I worked exteinsively with Jax, Flax, TensorFlow, and the Google Cloud Platform.

TU Delft - Distributed Systems Lab   Delft, The Netherlands
Software Developer   *04/2018-10/2019*

- Software development in Python on the Trustchain and the Tribler P2P ecosystem.

Technical University of Cluj-Napoca   Cluj-Napoca, Romania
Student Researcher   *11/2016-07/2017*

- Worked in the Knowledge Engineering Group where I led and assisted research in fundamental methods for Multiple Instance Learning. Our work resulted in a paper published at an ECML-PKDD'18 workshop.

Telenav   Cluj-Napoca, Romania
Software Engineering Intern   *07/2016-08/2016*

- Worked in C++ on a geopositioning platform. I implemented two features (twice the workload expected from interns): (1) dynamically discovered POI and (2) circuit generation. Both features made it into production.
- Received full-time offer.

EDUCATION

ETH Zürich
PhD in Computer Science @ Systems Group   *2019-2024*

- 5th best university in Computer Science in the world
- Focus on Systems for Machine Learning and Databases

ETH Zürich & TU Delft
MSc in Computer Science   *2017-2019*

- MSc Thesis: Distribution Optimization of Large Scale Machine Learning
- Grades: thesis: 9 out of 10. GPA: 8.5 out of 10

Technical University of Cluj-Napoca
BSc in Computer Science   *2013-2017*

- BSc Thesis: Complex Localization in the Multiple Instance Learning Context
- Grades: thesis: 10 out of 10. GPA: 9.6 out of 10; Graduated top 1%.

PUBLICATIONS

- Audibert, A., Chen, Y., **Graur, D.**, Klimovic, A., Simsa, J. and Thekkath, C., *tf.data service: A Case for Disaggregating ML Input Data Processing*, 2023, 14th Symposium on Cloud Computing   2023
- ***Graur, D.**, Aymon, D., Kluser, D., Albrici, T., Thekkath, C. and Klimovic, A., Cachew: Machine Learning Input Data Processing as a Service, 2022, Proceedings of the USENIX ATC*   2022
- **Graur, D.**, Müller I., Proffitt M., Fourny G., Watts G. T., and Alonso G., *Evaluating Query Languages and Systems for High-Energy Physics Data*, 2022, VLDB Endowment   2022
- ***Graur, D.**, Bruno, R. and Alonso, G., Specializing Generic Java Data Structures, 2021, 18th ACM International Conference on Managed Programming Languages & Runtimes*   2021
- **Graur, D.**, Bruno, R., Bischoff, J., Rieser, M., Scherr, W., Hoefler, T. and Alonso, G., *Hermes: Enabling efficient large-scale simulation in MATSim*, 2021, Procedia Computer Science, 184, pp.635-641.   2021
- *Rellermeyer J. S., Khorasani S. O., **Graur D.** and Parthasarathy A., The Coming Age of Pervasive Data Processing, 2019, 18th International Symposium on Parallel and Distributed Computing,*   2019
- **Graur D.** , Maris R. A., Potolea R., Dinsoreanu M. and Lemnaru C., *Complex Localization in the Multiple Instance Learning Context*, 2018, ECML-PKDD (New Frontiers in Mining Complex Patterns)   2018

HONORS AND AWARDS

IDEA League
Research Grant   02.2019-07.2019
- Grant awarded for cutting-edge research projects developed within the IDEA League.

TU Delft
Excellence Scholarship   09.2017-08.2019
- Two-year scholarship awarded to academically excellent students.

Technical University of Cluj-Napoca
Merit Scholarship   2013-2017
- Best possible scholarship awarded based on academic excellence.

SKILLS

Languages:	PYTHON, C/C++, JAVA
Topics:	Distributed and Large Scale Systems, Databases, ML, Cloud Computing, Research
Frameworks:	TensorFlow, Spark, PyTorch, Jax

LANGUAGES

ENGLISH:	Proficient (C2) - 2012 Cambridge Certificate in Advanced English - Grade A
GERMAN:	Intermediate (B1)
FRENCH:	Intermediate (B1)
ROMANIAN:	Native