# Curriculum Vitae - Octavian-Eugen Ganea

Postdoctoral AI researcher at MIT


## CONTACT:

**Email:** oct@mit.edu **Webpage:** https://people.csail.mit.edu/oct


## ACADEMIC HISTORY

*Sep 2019 - present: Postdoctoral researcher*, **Massachusetts Institute of Technology**

- Advisors: Tommi Jaakkola, Regina Barzilay
- Geometric machine learning for accelerating drug and material discovery.
- I contribute to the Machine Learning for Pharmaceutical Discovery and Synthesis MLPDS Consortium, AI Cures initiative, DARPA AMD project. I am a member of the ELLIS society.

*2014 - 2019: PhD in Computer Science*, **ETH Zurich**, Switzerland

- Advisor: Thomas Hofmann
- Thesis: Non-Euclidean neural representation learning of words, entities and hierarchies

*2013: Research intern*, Algorithms & Data Structures lab, **ETH Zurich**

- Algorithms for de novo peptide sequencing. Advisor: Peter Widmayer

*2010 - 2012: MSc in Computer Science*, **Ecole Polytechnique Fédérale de Lausanne**

- MSc thesis at Google Zurich: Bayesian networks for ranking entities in a geographic search engine. Thesis advisors: Bernard Moret (EPFL), Radu Jurca (Google)

*2006 - 2010: BSc in Computer Science*, **University Politehnica of Bucharest (UPB)**

## EMPLOYMENT HISTORY

Jun - Oct 2018: **Google Brain Research**, Zurich, CH - *Research intern*

- Generative models for text.

May - Aug 2017: **Google AI Research**, Mt. View, CA - *Research intern*

- Representation learning for hierarchical structures and knowledge graphs.

Feb - Oct 2012: **Google Maps Search Quality**, Zurich, CH - *Research intern, MSc thesis*

Jun - Sep 2011: **Google**, Mt. View, CA - *Software engineer intern*

## RESEARCH PUBLICATIONS

19 peer-reviewed publications including 10 (co-)first author. — Google Scholar Profile

### Featured Publications on Euclidean 3D Models for Chemistry/Biology

- *GeoMol: Torsional Geometric Generation of Molecular 3D Conformer Ensembles*
  O-E. Ganea*, L. Pattanaik*, C. Coley, R. Barzilay, K. Jensen, W. Green, T. Jaakkola
  **Spotlight (top 3% submissions)** at **NeurIPS 2021**: Conference on Neural Information Processing Systems
  **Covered by MIT News** [link]
- *Independent SE(3)-Equivariant Models for End-to-End Rigid Protein Docking*
  O-E. Ganea*, X. Huang*, C. Bunne, Y. Bian, R. Barzilay, T. Jaakkola, A. Krause.
  **Spotlight** at **ICLR 2022** (International Conference on Learning Representations)
  Top 13 of 3326 ICLR submissions (**top 0.4%**), ranked by average review score (source).
  **Covered by MIT News** [link]

Contributed Talk at ELLIS Machine Learning for Molecule Discovery Workshop, 2021

- EquiBind: *Geometric Deep Learning for Drug Binding Structure Prediction*
  H. Stark*, O-E. Ganea* (equal first author), L. Pattanaik, R. Barzilay, T. Jaakkola
  Pre-print, under submission.
- *Crystal Diffusion Variational Autoencoder for Periodic Material Generation*
  T. Xie*, X. Fu*,O-E. Ganea* (equal first author), R. Barzilay, T. Jaakkola.
  **Contributed Talk** at NeurIPS 2021 Workshop on Machine Learning and the Physical Sciences.
  Accepted to **ICLR 2022** (International Conference on Learning Representations)

## Featured Publications on Foundations of Non-Euclidean Machine Learning

- *Hyperbolic Neural Networks*
  O-E. Ganea*, G. Bécigneul*, T. Hofmann
  **Spotlight (top 4% submissions)** at **NeurIPS 2018**: Conference on Neural Information Processing Systems
- *Hyperbolic Entailment Cones for Learning Hierarchical Embeddings*
  O-E. Ganea, G. Bécigneul, T. Hofmann
  **Oral talk**, full paper at **ICML 2018**: International Conference on Machine Learning.
- *Poincaré GloVe: Hyperbolic Word Embeddings*
  A. Tifrea*, G. Bécigneul*, O-E. Ganea*
  **ICLR 2019**, full paper: International Conference on Learning Representations.
- *Constant Curvature Graph Convolutional Networks*
  G. Bachmann, G. Bécigneul, O-E. Ganea
  **ICML 2020**, full paper: International Conference on Machine Learning.

## Featured Publications on Natural Language Processing

- *Breaking the Softmax Bottleneck via Learnable Monotonic Pointwise Non-linearities*
  O-E. Ganea, S. Gelly, G. Bécigneul, A. Severyn
  **Oral talk**, full paper at **ICML 2019**: International Conference on Machine Learning.
- *Deep Joint Entity Disambiguation with Local Neural Attention*
  O-E. Ganea, T. Hofmann
  **EMNLP 2017**, full paper: Empirical Methods in Natural Language Processing
- *End-to-end Neural Entity Linking*
  N. Kolitsas*, O-E. Ganea*, T. Hofmann
  **CoNLL 2018**, full paper: Conference on Natural Language Learning.

## Other Publications

- *Computationally Tractable Riemannian Manifolds for Graph Embeddings*
  C. Cruceru, G. Bécigneul, O-E. Ganea
  Full paper, **AAAI 2021**: The Thirty-Fifth AAAI Conference on Artificial Intelligence.
- *Mixed-curvature Variational Autoencoders*
  O. Skopek, O-E. Ganea, G. Bécigneul
  Full paper, **ICLR 2020**: International Conference on Learning Representations.
- *Message Passing Networks for Molecules with Tetrahedral Chirality*
  L. Pattanaik, O-E. Ganea, I. Coley, K. Jensen, W. Green, I. Coley
  Paper at **NeurIPS 2020** Machine Learning for Molecules **Workshop**.

- *Hierarchical Image Classification using Entailment Cone Embeddings*
  A. Dhall, A. Makarova, O-E. Ganea, D. Pavllo, M. Greeff, A. Krause
  **CVPR** 2020 Intl. Workshop on Differential Geometry in Computer Vision and ML
- *Riemannian Adaptive Optimization Methods*
  G. Bécigneul, O-E. Ganea
  Full paper, **ICLR 2019**: International Conference on Learning Representations.
- *Learning and Evaluating Sparse Interpretable Sentence Embeddings*
  V. Trifonov, O-E. Ganea, A. Potapenko, T. Hofmann
  **EMNLP 2018 Workshop** on analysis and interpretation of neural networks for NLP
- *Neural Multi-Step Reasoning for Question Answering on Semi-Structured Tables*
  T. Haug, O-E. Ganea, P. Grnarova
  **ECIR 2018**: European Conference on Information Retrieval .
- *Web2Text: Deep Structured Boilerplate Removal*
  T. Vogels, O-E. Ganea, C. Eickhoff
  **ECIR 2018**: European Conference on Information Retrieval .
- *Probabilistic Bag-Of-Hyperlinks Model for Entity Linking*
  O-E. Ganea, M. Ganea, A. Lucchi, C. Eickhoff, T. Hofmann
  **WWW 2016**, full paper: International World Wide Web Conference
- *Optimal Transport Graph Neural Networks*
  B Chen*, G Bécigneul*, O-E. Ganea*, R Barzilay, T Jaakkola
  Under review at AAAI 2022

## HONORS AND AWARDS

- 2021 - **Outstanding Reviewer Award - NeurIPS 2021**, Thirty-fifth Conference on Neural Information Processing Systems
- 2021 - **Top 3% of all submitted papers - NeurIPS 2021**, Spotlight paper at Conference on Neural Information Processing Systems
- 2019 - **Fellowship Grant - Institute for Advanced Study**, special-year program 2019 - 2020 in Machine Learning led by Sanjeev Arora (declined)
- 2018 - **Top 4% of all submitted papers - NeurIPS 2018**, Spotlight paper at Conference on Neural Information Processing Systems
- 2013 - **Top 5**%, Google Code Jam algorithms contest
- 2010 - **Excellence Scholarship** for master studies at EPFL: 30,000 $ - Dinu Patriciu foundation (Romania)
- 2010, 2009, 2008 - **Excellence Teaching Diplomas**, awarded three times by the *prime-minister of Romania* or *Ministry of Education and Research in Romania* for my teaching activity for international mathematical contests and olympiads, e.g., my students won 5 gold and silver medals at IMO.
- 2009, 2008, 2007 - $1^{st}$ **and** $2^{nd}$ **prizes** - International Mathematical Contest for University Students IMC (www.imc-math.org)
- 2008, 2007 - **Gold Medal** (2008) and **Silver Medal** (2007) - South Eastern Mathematical Olympiad for University Students - SEEMOUS, Greece / Cyprus
- 2005 - **Silver Medal (**$3^{rd}$ **place)** - Tuymada International Olympiad in Mathematics, Yakutsk, Russia
- 2005 - **Silver Medal** - Balkan Mathematical Olympiad, Iasi, Romania
- 2006, 2005, 2004, 2003, 2002 - **Gold Medal** (top 7 places), Romanian National Mathematical Olympiad
- 2009, 2008 - Honorable Mention - South-Eastern Europe ACM algorithms contest

## INVITED TALKS

1. March 2022: New York University, host departments: CS, CDS, CSE, ECE
2. March 2022: Moderna Inc., host: Eric Ma, James Ross
3. March 2022: University of Waterloo, hosts: Lila Kari, Olga Veksler, Ian Goldberg
4. February 2022: Pfizer Inc., host: Vishnu Sresht
5. February 2022: Amazon Research UK, host: Tom Diethe
6. February 2022: University of Illinois at Urbana-Champaign, CSLS conference
7. February 2022: Sanofi S.A., host: Yu Qiu
8. January 2022: Microsoft Research UK (Cambridge) and Netherlands (Amsterdam), host: Max Welling, Rianne van den Berg
9. January 2022: Entos AI, host: Tom Miller
10. January 2022: Intel AI Lab, host: Mariano Phielipp
11. November 2021: Imperial College London, host: Michael Bronstein
12. November 2021: Mila - Quebec AI Institute, host: Jian Tang
13. June 2021: Microsoft Research New England, host: David Alvarez-Melis
14. April 2021: Machine Learning for Pharmaceutical Discovery and Synthesis Consortium
15. October 2020: MIT, guest lecture on generative models in the 6.867 Machine Learning lecture
16. July 2020: Georgia Institute of Technology, host: Le Song
17. July 2020: Northeastern University, hosts: T. Eliassi-Rad, D. Krioukov, R. Yu
18. December 2020: Bowdoin College, guest lecture, host: Jennifer Taback
19. April 2020: Machine Learning for Pharmaceutical Discovery and Synthesis Consortium
20. March 2020: IBM Global Business Services, host: Lucia Stavarache
21. February 2020: University of Massachusetts Amherst, host: Andrew McCallum
22. January 2020:Relational.ai, host: Nikos Vasilakis
23. October 2019: Aggregate Intellect, host: Amir Feizpour
24. October 2019: MIT, prof. Tommi Jaakkola's group
25. October 2018: Google Brain, Zurich
26. April 2018: ETH Zurich Machine Learning Seminar
27. May 2017: Google Research Mountain View
28. March 2017: ETH Zurich Machine Learning Seminar

## TEACHING

- Oct 2020: **6.867 Machine Learning lecture, MIT**: guest lecture on generative models – [**sole instructor**]
- Jun-Aug 2020: mathematics lectures for the candidates aiming to represent Romania at the International Mathematics Olympiad (website) – [**sole instructor**]
- Teaching assistant for the lectures: Deep Learning (2017, 2018 - ETHZ); Computational Intelligence Lab (2015, 2016, 2017, 2018 [**head TA**] - ETHZ); Information Retrieval (2014, 2015, 2016 - ETHZ); Advanced Algorithms (2011 - EPFL); Graph Theory (2011 - EPFL); Concurrency (2011 - EPFL); Numerical Methods (2008, 2009 - UPB)
- 2007 - 2010: *Lecturer for Mathematics Olympiads and Competitions* and Member of the Romanian national selection committee for the International Mathematical Olympiad (IMO). I taught mathematics lectures to the candidates aiming to represent Romania at the IMO. Several of my students won prizes and medals at international contests, including **5 gold and silver medals at IMO** (Stefan Ivanovici, Octav Dragoi, Ioana-Maria Tamas). – [**sole instructor**]

- 2007 - 2009: Member of the Romanian National Mathematical Olympiad committee: proposed original problems for various olympiad stages, graded students' papers, etc.

**MENTORING & SUPERVISED THESES / PROJECTS:**

I proposed and (co-)supervised **24 student research projects**, with **10** being later pushed into **research publications**. MSc thesis is 6 months long, BSc thesis is 4-6 months long:

1. Hannes Stark: *End-to-end Geometric Drug Binding*, research internship, MIT and TUMunich, 2021, [ongoing]
2. Julia Bala: *Generic Riemannian Embedding Spaces*, UROP project, MIT, 2021
3. Xinyuan Huang: *Modeling Protein Complexes*, MSc thesis, MIT and ETHZ, 2021, **research paper submitted**
4. Panayiotou Panayiotis: *Permutation Invariant Graph Generation via Optimal Transport*, MSc thesis, MIT and ETHZ, 2020 (now software engineer at Google)
5. Octav Dragoi: *Permutation Invariant Graph Optimization*, MSc thesis, MIT and TUMunich, 2020 (now quantitative researcher at Citadel)
6. Ondrej Skopek: *Mixed-curvature Variational Autoencoders*, MSc thesis, ETHZ, 2019, **research paper accepted @ ICLR** (now software engineer at Google)
7. Bachmann Gregor: *Riemannian Graph Neural Networks*, MSc thesis, ETHZ, 2019, **research paper accepted @ ICML** (now Ph.D. student at ETH Zurich)
8. Andreas Bloch: *Mixed-curvature Recommender Systems*, BSc thesis, ETHZ, 2019
9. Ankit Dhall: *Hierarchical Image Captioning*, MSc thesis, ETHZ, 2019, **research paper accepted @ CVPR Workshop** (now researcher at LatticeFlow.ai )
10. Philipp Wirth: *Hyperbolic Language Models*, MSc thesis, ETHZ, 2019 (now ML engineer at Lightly)
11. Jovan Andonov: *Neural ODEs for Language Modeling*, MSc thesis, ETHZ, 2019 (now AI software engineer at IBM Research)
12. Calin Cruceru: *Matrix Graph Embeddings*, MSc thesis, ETHZ, 2019, **research paper accepted @ AAAI** (now engineer at Daedalean AI)
13. Alexandru Tifrea: *Poincaré Glove: Hyperbolic Word Embeddings*, MSc thesis, ETHZ, 2018, **research paper accepted @ ICLR** (now Ph.D. student at ETH Zurich)
14. Kolitsas Nikolaos: *End-to-end Neural Entity Linking*, MSc thesis, ETHZ, 2018, **research paper accepted @ CoNLL** (now senior ML software engineer at Avaloq)
15. Valentin Trivonov: *Sparse and Interpretable Sentence Embeddings*, MSc thesis, ETHZ, 2018, **research paper accepted @ EMNLP Workshop** (now software engineer at Google)
16. Igor Petrovski: *Hyperbolic Sentence Embeddings*, MSc thesis, ETHZ, 2018 (now senior software engineer at Google)
17. Junlin Yao: *Detecting Medication and Adverse Drug Events from Electronic Health Records*, MSc thesis, ETHZ, 2018 (now data scientist at 4Paradigm)
18. Andreas Hess: *Reinforcement Learning for Question Answering with Semi-structured Tables*, MSc thesis, ETHZ, 2017
19. Yifan Su: *Deep Structured Prediction for Joint Entity Linking and Coreference Resolution*, MSc thesis, ETHZ, 2016
20. Till Haug: *Convolutional and Recursive Neural Networks for Question Answering on Semi-structured Tables*, BSc thesis, ETHZ, 2017, **paper accepted @ ECIR** (now Co-Founder and COO at Veezoo)
21. Severin Bahman: *Memory Networks for Entity Linking*, semester project, ETHZ, 2016 (now strategy consultant at Zurich Insurance)
22. Andreas Georgiadis: *Sentence Representations for QA*, semester project, ETHZ, 2016 (now quantitative developer at Squarepoint Capital)

23. Thijs Vogels: *Structured Prediction for Web Page Content Extraction*, semester project, ETHZ, 2016, **research paper accepted @ ECIR** (now Ph.D. student at EPFL)
24. Monteiro Joao Pedro: *Matrix Embeddings for Knowledge Base Completion*, BSc thesis, ETH 2016 (now CTO and Co-Founder of Veezoo)

**SERVICE**

- Reviewer for top tier research conferences in Artificial Intelligence:
  NeurIPS'21 **(Outstanding Reviewer Award)**, NeurIPS'20, ICML'20, AAAI'20, NeurIPS'19, EMNLP'19, ACL'18, EMNLP'18
- Co-founder of http://openconsulting.ai/, a free AI consulting platform for societal non-profit projects.
- Research mentor at London Geometry and Machine Learning Summer School, https://logml.ai, July 2021.

**MISC**

- Finished 6 mountain marathons and one road marathon
- Climbed 10 four-thousand meter peaks in the Alps [clip of my Zinalrothorn (4,221 m) climb]
- Completed the longest via ferrata in Switzerland (Leukerbad)
- Media coverage: Forbes "30 under 30", Romania's 2013 edition