

# Tristan Deleu

6666 St-Urbain

H2S 3H1 Montréal, QC – Canada

<https://tristandeleu.github.io>

## Education

2017 - present	<b>Université de Montréal / Mila</b>	Montréal, QC, Canada
2018 - pres.	<b>Ph.D. Computer Science - Artificial Intelligence</b>	
Research	<i>Probabilistic models, Structure learning, Causality, Reinforcement learning, Meta-learning, Few-shot learning</i>	
2017 - 2018	<b>M.Sc Computer Science - Artificial Intelligence</b>	
2011 - 2015	<b>Ecole Normale Supérieure de Cachan</b>	Cachan, France
2013 - 2014	<b>Master 2 - Applied Mathematics (M.Sc)</b> , with highest honors	
	Master <i>Mathematics, Computer Vision and Machine Learning (MVA, ENS Cachan)</i>	
2012 - 2013	<b>Master 1 - Mathematics</b> , with high honors	
	ENS Cachan & Université Paris Diderot VII	
2011 - 2012	<b>Licence 3 - Mathematics (B.Sc)</b> , with highest honors	
2011 - 2012	<b>Licence 3 - Computer Science (B.Sc)</b> , with highest honors	

## Professional Experience

2024	<b>Valence Labs, Recursion</b> , Senior Research Scientist	Montreal, QC, Canada
Jan 24 - present		
2022	<b>Google Research</b> , Student Researcher	Mountain View, CA, USA
Jun 22 - Aug 22	Project: <i>Combining Zero-shot &amp; Few-shot models</i>	
	Description: <i>Evaluation of zero-shot &amp; few-shot models based on large-scale pre-trained encoders on the Meta-Dataset benchmark. Creation of a more challenging episodic few-shot classification benchmark, based on the Open Images dataset. Study of multiple strategies to combine zero-shot &amp; few-shot models, including an empirically effective regularization method. Design of a bi-level optimization problem to automatically meta-learn the regularization constants.</i>	
2015 - 2017	<b>Snips</b> , Researcher	Paris, France
Oct 16 - Jul 17	Project: <i>Query parsing with Convolutional Neural Networks</i>	
	Description: <i>Design of a Convolutional Neural Network (CNN) and Conditional Random Field-based (CRF) model to parse queries in natural language (intent classification &amp; slot filling) on 10 preset domains (including “Place details”, “Restaurant reservation”, “Directions”, etc...). Collaboration with engineering teams to run inference on mobile devices (see <a href="#">Snips SDK</a>).</i>	
Sep 15 - Oct 16	Project: <i>Memory-augmented Neural Networks applied to Reasoning</i>	
	Description: <i>Literature review on memory-augmented networks. Implementation of a Neural Turing Machine (NTM) library in Theano/Lasagne. Reproduction of most of the results on the algorithmic tasks from the <a href="#">original paper</a>, with similar generalization performances. Application of NTMs to the problem of learning matching parentheses (Dyck words). Study of the bAbI dataset and generation of similar unbiased reasoning tasks based on rejection sampling &amp; application of NTMs on these Q&amp;A tasks (graph data).</i>	

2014	<b>Snips</b> , <i>Research Internship</i>	Paris, France
Apr 14 - Jan 15	Project: <i>Urban context modeling with Dynamic Bayesian Networks</i> Description: <i>Theory and design of a hierarchical graphical model to gather contextual data from multiple sources, and infer informations about a user. Design of a 2-stage classifier based on a Random Forest and a Hidden Markov Model to detect the mode of transportation online and on-device from raw accelerometer data.</i>	
2013	<b>University College London</b> , <i>Research Internship</i>	London, UK
Jun 13 - Aug 13	Project: <i>Hierarchical Representations using Deep Belief Networks</i> Supervisor: <i>Dr. David Barber (CSML/UCL)</i> Description: <i>Theory of Deep Belief Networks, Gibbs sampling-based learning, and inference algorithms. Design of two autoencoder-based learning algorithms, approaching the state of the art in terms of data reconstruction: the Neural Principal Component Analysis and the Implicit Neural Network. Evaluation of these algorithms on shallow neural networks on the MNIST dataset.</i>	
2012	<b>ENS Cachan &amp; Université Paris Sud XI</b> , <i>Research Internship</i>	Orsay, France
Feb 12 - Jun 12	Project: <i>Optimal Transport: Theory and Continuity Optimization</i> Supervisor: <i>Pr. Filippo Santambrogio (P11)</i> Description: <i>Theory of distributions and optimal transport. Design of an evolutionary algorithm for minimizing the oscillation in transport problems, with applications to privacy addressed by the uniform encoding of geolocation data.</i>	

## Publications

Dec 2023	<b>Tristan Deleu</b> , Mizu Nishikawa-Toomey, Jithendaraa Subramanian, Nikolay Malkin, Laurent Charlin, Yoshua Bengio, <i>Joint Bayesian Inference of Graphical Structure and Parameters with a Single Generative Flow Network</i> , NeurIPS 2023 <a href="https://arxiv.org/abs/2305.19366">https://arxiv.org/abs/2305.19366</a>
Jul 2023	Salem Lahlou, <b>Tristan Deleu</b> , Pablo Lemos, Dinghuai Zhang, Alexandra Volokhova, Alex Hernández-García, Léna Néhale Ezzine, Yoshua Bengio, Nikolay Malkin, <i>A Theory of Continuous Generative Flow Networks</i> , ICML 2023 <a href="https://arxiv.org/abs/2301.12594">https://arxiv.org/abs/2301.12594</a>
Jul 2023	Sébastien Lachapelle*, <b>Tristan Deleu</b> *, Divyat Mahajan, Ioannis Mitliagkas, Yoshua Bengio, Simon Lacoste-Julien, Quentin Bertrand, <i>Synergies between Disentanglement and Sparsity: Generalization and Identifiability in Multi-Task Learning</i> , ICML 2023 <a href="https://arxiv.org/abs/2211.14666">https://arxiv.org/abs/2211.14666</a>
Jun 2023	Yoshua Bengio*, Salem Lahlou*, <b>Tristan Deleu</b> *, Edward J. Hu, Mo Tiwari, Emmanuel Bengio, <i>GFlowNet Foundations</i> , JMLR <a href="https://arxiv.org/abs/2111.09266">https://arxiv.org/abs/2111.09266</a>
May 2023	Nikolay Malkin*, Salem Lahlou*, <b>Tristan Deleu</b> *, Xu Ji, Edward Hu, Katie Everett, Dinghuai Zhang, Yoshua Bengio, <i>GFlowNets and variational inference</i> , ICLR 2023 <a href="https://arxiv.org/abs/2210.00580">https://arxiv.org/abs/2210.00580</a>
Apr 2023	Moksh Jain, <b>Tristan Deleu</b> , Jason Hartford, Cheng-Hao Liu, Alex Hernandez-Garcia, Yoshua Bengio, <i>GFlowNets for AI-Driven Scientific Discovery</i> , Digital Discovery, Royal Society of Chemistry <a href="https://arxiv.org/abs/2302.00615">https://arxiv.org/abs/2302.00615</a>
Feb 2023	Ramnath Kumar, <b>Tristan Deleu</b> , Yoshua Bengio, <i>The Effect of Diversity in Meta-Learning</i> , AAAI 2023 (Oral) <a href="https://arxiv.org/abs/2201.11775">https://arxiv.org/abs/2201.11775</a>
Aug 2022	<b>Tristan Deleu</b> , António Góis, Chris Emezue, Mansi Rankawat, Simon Lacoste-Julien, Stefan Bauer, Yoshua Bengio, <i>Bayesian Structure Learning with Generative Flow Networks</i> , UAI 2022 <a href="https://arxiv.org/abs/2202.13903">https://arxiv.org/abs/2202.13903</a>
Apr 2022	<b>Tristan Deleu</b> , David Kanaa, Leo Feng, Giancarlo Kerg, Yoshua Bengio, Guillaume Lajoie, Pierre-Luc Bacon, <i>Continuous-Time Meta-Learning with Forward Mode Differentiation</i> , ICLR 2022 (Spotlight) <a href="https://arxiv.org/abs/2203.01443">https://arxiv.org/abs/2203.01443</a>

Apr 2021	Yoshua Bengio, Prateek Gupta, Tegan Maharaj, Nasim Rahaman, Martin Weiss, <b>Tristan Deleu</b> , Eilif Muller, Meng Qu, Victor Schmidt, Pierre-Luc St-Charles, Hannah Alsdurf, et al., <i>Predicting Infectiousness for Proactive Contact Tracing</i> , ICLR 2021 ( <b>Spotlight</b> ) <a href="https://arxiv.org/abs/2010.12536">https://arxiv.org/abs/2010.12536</a>
Apr 2020	Sébastien Lachapelle, Philippe Brouillard, <b>Tristan Deleu</b> , Simon Lacoste-Julien, <i>Gradient-Based Neural DAG Learning</i> , ICLR 2020 <a href="https://arxiv.org/abs/1906.02226">https://arxiv.org/abs/1906.02226</a>
Apr 2020	Yoshua Bengio, <b>Tristan Deleu</b> , Nasim Rahaman, Rosemary Ke, Sébastien Lachapelle, Olexa Bilaniuk, Anirudh Goyal, Christopher Pal, <i>A Meta-Transfer Objective for Learning to Disentangle Causal Mechanisms</i> , ICLR 2020 <a href="https://arxiv.org/abs/1901.10912">https://arxiv.org/abs/1901.10912</a>

## Awards & Scholarships

2021	<b>Sony Research Award Program</b> Project: <i>The learning dynamics of meta-learning</i>
2019	<b>Antidote scholarship 2019</b> Druide informatique
Oct 2019	<b>Best in Show award – Global PyTorch Summer Hackathon 2019</b> 1st place for the project <a href="#">Torchmeta</a>

## Teaching

Fall 2022	<b>Université de Montréal – Probabilistic Graphical Models</b> Teaching assistant
Fall 2021	<b>Université de Montréal – Probabilistic Graphical Models</b> Teaching assistant
Winter 2021	<b>Université de Montréal – Representation Learning</b> Teaching assistant

## Languages & Skills

French	Native
English	Fluent (Oral & Written)
Computer Skills	Jax, L <sup>A</sup> T <sub>E</sub> X, Python, PyTorch, Rust
Github	<a href="https://github.com/tristandeleu/">https://github.com/tristandeleu/</a>
Mila	<i>Meta-Learning Reading Group</i> organizer (2018 - 2020)

## Preprints & Workshop

- Jul 2023 **Tristan Deleu**, Yoshua Bengio, *Generative Flow Networks: a Markov Chain Perspective*  
<https://arxiv.org/abs/2307.01422>
- Feb 2023 Mizu Nishikawa-Toomey\*, **Tristan Deleu\***, Jithendaraa Subramanian, Yoshua Bengio, Laurent Charlin, *Bayesian learning of Causal Structure and Mechanisms with GFlowNets and Variational Bayes*, Workshop on Graphs and more Complex structures for Learning and Reasoning (GCLR), AAAI 2023  
<https://arxiv.org/abs/2211.02763>
- Feb 2021 **Tristan Deleu**, Yoshua Bengio, *Structured Sparsity Inducing Adaptive Optimizers for Deep Learning*  
<https://arxiv.org/abs/2102.03869>
- Sep 2019 **Tristan Deleu**, Tobias Würfl, Mandana Samiei, Joseph Paul Cohen, Yoshua Bengio, *Torchmeta: A Meta-Learning library for PyTorch*  
<https://arxiv.org/abs/1909.06576>
- Dec 2018 **Tristan Deleu**, Yoshua Bengio, *The effects of negative adaptation in Model-Agnostic Meta-Learning*, Workshop on Meta-Learning, NeurIPS 2018  
<https://arxiv.org/abs/1812.02159>
- Jul 2018 **Tristan Deleu**, Simon Guiroy, Seyedarian Hosseini, *On the reproducibility of gradient-based Meta-Reinforcement Learning baselines*, Reproducibility in Machine Learning Workshop, ICML 2018
- Dec 2016 **Tristan Deleu**, Joseph Dureau, *Learning Operations on a Stack with Neural Turing Machines*, 1st Workshop on Neural Abstract Machines & Program Induction (NAMPI), NIPS 2016  
<https://arxiv.org/abs/1612.00827>

## Projects

- Oct 2019 **Model-Agnostic Meta-Learning**  
Link: <https://github.com/tristandeleu/pytorch-maml>  
Description: *Implementation of the “Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks” (MAML) paper in PyTorch and Torchmeta. Applications to few-shot supervised learning tasks.*
- Sep 2019 **Torchmeta: A Meta-Learning library for PyTorch**  
Link: <https://github.com/tristandeleu/pytorch-meta>  
Description: *A collection of extensions and data-loaders for few-shot learning & meta-learning in PyTorch. It provides a unified for a wide range of standard benchmarks, including Omniglot, Mini-ImageNet, Tiered-ImageNet, CIFAR-FS, Fewshot-CIFAR100.*
- Apr 2018 **Meta-Learning for Reinforcement Learning**  
Link: <https://github.com/tristandeleu/pytorch-maml-rl>  
Description: *Implementation of the “Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks” (MAML) paper in PyTorch. Applications to Bandits, Tabular Markov Decision Processes and Continuous Control problems.*
- May 2016 **One-shot Learning with Memory-Augmented Neural Networks**  
Link: <https://github.com/tristandeleu/ntm-one-shot>  
Description: *Implementation of the “One-shot Learning with Memory-Augmented Neural Networks” paper in Theano, with experiments on the Omniglot dataset.*
- Feb 2016 **NTM-Lasagne: Neural Turing Machines in Theano/Lasagne**  
Link: <https://github.com/snipsco/ntm-lasagne>  
Description: *Implementation of the “Neural Turing Machines” paper in Lasagne/Theano as a Lasagne layer. Source and pre-trained models to reproduce the original experiments on algorithmic tasks (Copy, Repeat Copy, Associative Recall, Dyck words). Visualization of the heads & animations of the memory with Matplotlib.*