

DR. NIKOLA KONSTANTINOV

INSAIT, Synergy Tower, Tsarigradsko shose 111P, 1784 Sofia, Bulgaria
Website: <https://insait.ai/dr-nikola-konstantinov>, Email: nikola.konstantinov@insait.ai

EMPLOYMENT

- | | |
|-------------------|---|
| 06.2023 - | INSAIT, Sofia University: Tenure-Track Faculty <ul style="list-style-type: none">- Leading the machine learning group- Research focus: mathematical guarantees for trustworthy ML |
| 04.2022 – 04.2023 | ETH Zurich AI Center: Postdoctoral fellow <ul style="list-style-type: none">- Mentors: Prof. Martin Vechev and Prof. Fanny Yang |

EDUCATION

- | | |
|----------------|--|
| 2017 – 03.2022 | IST Austria (ISTA): PhD Student; Member the ELLIS PhD Program <ul style="list-style-type: none">- Supervisor: Prof. Christoph H. Lampert- Thesis: <i>Robustness and Fairness in Machine Learning</i>- Nominated for the best ISTA PhD thesis award of 2022 |
| 2013 – 2017 | University of Oxford: Master in Mathematics and Statistics; Grade: 1.1 <ul style="list-style-type: none">- Master Thesis: “<i>Kernel Dependence Measures for Unsupervised Learning</i>”, supervised by Prof. Dino Sejdinovic- Focus on Probability Theory, Stochastic Analysis, Statistics |
| 2005 – 2013 | Sofia High School of Mathematics <ul style="list-style-type: none">- Bulgarian High School Diploma, grade 5.99 out of 6- Extracurricular classes in Mathematics, Informatics and English |

INTERNSHIPS AND ACADEMIC VISITS

- | | |
|----------------|--|
| Spring of 2021 | University of Milan, Italy: ELLIS PhD Exchange Visit (3 months) <ul style="list-style-type: none">- Host: Prof. Nicolò Cesa-Bianchi- Studying models for online learning from multiple unreliable data sources |
| Summer of 2019 | Amazon.com, Berlin, Germany: Applied Research Internship (3 months) <ul style="list-style-type: none">- Group: Intelligent Cloud Control; Mentor: Paul Vazquez- Evaluation and model comparison of predictive models for cloud computing |
| Summer of 2016 | ATASS Sports, Exeter, UK: Research Internship (7 weeks) <ul style="list-style-type: none">- Statistical modelling for sports prediction- Experience in statistical programming with R and manipulating large datasets |
| Summer of 2015 | IST Austria (ISTA): Research Internship (11 weeks) <ul style="list-style-type: none">- Supervisor: Prof. Nick Barton- Mathematical models of evolution: quantifying the effects of various genetic phenomena on spatially extended populations |

Supervision and group

PhD Students (current): Nikita Tsoy (joint with Martin Jaggi), Ivan Kirev (joint with Andreas Krause), Kristian Minchev

M.Sc. Students: Andor Vari-Kakas (ETH, 2023), Kristof Szabo (ETH, 2023-2024), Florian Dorner (ETH, 2022-2023)

Interns: Anna Mihalkova (INSAIT, 2023), Anna-Maria Gueorguieva (INSAIT, 2023), Dimitar Chakarov (INSAIT, 2024), Lyuben Baltadzhiev (INSAIT, 2025), Kalina Nikolova (INSAIT, 2025)

PAPERS

Publications

1. Nikita Tsoy, Ivan Kirev, Negin Rahimi, **N.K.**: On the Impact of Performative Risk Minimization for Binary Random Variables
In: *International Conference on Machine Learning (ICML)*, 2025
2. Nikita Tsoy, **N.K.**: Simplicity Bias of Two-Layer Networks beyond Linearly Separable Data
In: *International Conference on Machine Learning (ICML)*, 2024
3. Nikita Tsoy, Anna Mihalkova, Teodora Todorova, **N.K.**: Provable Mutual Benefits from Federated Learning in Privacy-Sensitive Domains
In: *Conference on Artificial Intelligence and Statistics (AISTATS)*, 2024
4. Nikita Tsoy, **N.K.**: Strategic Data Sharing between Competitors
In: *Conference on Neural Information Processing Systems (NeurIPS)*, 2023
5. Florian E. Dorner, **N.K.**, Giorgi Pashaliev, Martin Vechev: Incentivizing Honesty among Competitors in Collaborative Learning and Optimization
In: *Conference on Neural Information Processing Systems (NeurIPS)*, 2023
6. Florian E. Dorner, Momchil Peychev, **N.K.**, Naman Goel, Elliott Ash, Martin Vechev: Human-Guided Fair Classification for Natural Language Processing
In: *International Conference on Learning Representations (ICLR)*, **Spotlight**
Short version in: **TSRML@NeurIPS workshop**, 2022
7. Dimitar I. Dimitrov, Mislav Balunović, **N.K.**, Martin Vechev: *Data Leakage in Federated Averaging*
In: *Transactions of Machine Learning Research (TMLR)*, 2022
8. Eugenia Iofinova[#], **N. K.**[#], Christoph H. Lampert: *FLEA: Provably Fair Multisource Learning from Unreliable Training Data*
In: *Transactions of Machine Learning Research (TMLR)*, 2022; **# Equal contribution**
9. **N. K.**, Christoph H. Lampert: *Fairness-Aware PAC Learning from Corrupted Data*
In: *Journal of Machine Learning Research (JMLR)*, 2022
10. **N. K.**, Christoph H. Lampert: *On the Impossibility of Fairness-Aware Learning from Corrupted Data*
In: **AFRC@NeurIPS workshop, Contributed talk**, 2021
11. **N. K.**, Elias Frantar, Dan Alistarh, Christoph H. Lampert: *On the Sample Complexity of Adversarial Multi-Source PAC Learning*
In: *International Conference on Machine Learning (ICML)*, 2020
12. **N. K.**, Christoph H. Lampert: *Robust Learning from Untrusted Sources*
In: *International Conference on Machine Learning (ICML)*, **Long Talk**, 2019
13. Dan Alistarh, Torsten Hoefer, Mikael Johansson, **N. K.***, Sarit Khirirat, Cedric Renggli: *The Convergence of Sparsified Gradient Methods*
In: *Conference on Neural Information Processing Systems (NeurIPS)*, 2018; *** Alphabetic order**
14. Dan Alistarh, Chris De Sa, **N. K.***: *The Convergence of Stochastic Gradient Descent in Asynchronous Shared Memory*
In: *ACM Symposium of Principles of Distributed Computing (PODC)*, 2018; *** Alphabetic order**

Workshop Papers and Preprints

15. Nikita Tsoy, **N.K.**: Measuring the Layer-Wise Impact of Image Shortcuts on Deep Model Features
In: *SCSL Workshop @ ICLR*, 2025
16. Kristian Minchev, Dimitar I. Dimitrov, **N.K.**: LARP: Learner-Agnostic Robust Data Prefiltering
In: *DataWorld Workshop @ ICML*, 2025
17. Dimitar Chakarov, Nikita Tsoy, Kristian Minchev, **N.K.**: Incentivizing Truthful Collaboration in Heterogeneous Federated Learning
In: *OPT @ NeurIPS*, 2024

TALKS AND OUTREACH

Talks

1. July 2025: Invited talk at ISCPS 2025; Sofia, Bulgaria
Title: *Statistical Learning under Distribution Shift*
2. July 2025: ML seminar, University of Waterloo; Waterloo, Canada
Title: *Incentivizing Collaboration in Federated Learning*
3. October 2024: OxCSML seminar, University of Oxford; Oxford, UK
Title: *Incentivizing Collaboration in Federated Learning*
4. December 2023: Massachusetts Institute of Technology (MIT); New York, USA
Title: *Incentivizing Collaboration in Federated Learning*
5. December 2023: Microsoft Research; New York, USA
Title: *Incentivizing Collaboration in Federated Learning*
6. December 2023: Google Research; New York, USA
Title: *Incentivizing Collaboration in Federated Learning*
7. November 2023: INSAIT- RIKEN AIP Joint Workshop on Artificial Intelligence; Sofia Bulgaria
Title: *Incentivizing Collaboration in Federated Learning*
8. November 2023: New Horizons of Journalism Conference; Sofia Bulgaria
Title: *Generative AI: Risks, Responsibilities and Research*
9. July 2023: Third Annual Meeting of Young Bulgarian Mathematicians, IMI-BAN; Sofia, Bulgaria
Title: *Statistical Aspects of Trustworthy Machine Learning*
10. June 2023: Sofia Business School, NBU; Sofia, Bulgaria
Title: *Learning from Unreliable and Distributed Data*
11. April 2023: MLO group seminar; EPFL, Lausanne, Switzerland
Title: *Federated Learning in the Presence of Strategic Behavior*
12. March 2023: IML Seminar Series; ETH Zurich, Switzerland
Title: *Federated Learning in the Presence of Strategic Behavior*
13. October 2022: TrustML Young Scientists Seminars (online)
Title: *Statistical Aspects of Trustworthy Machine Learning*
14. October 2022: INSAIT Annual Conference, Tutorial Session; Sofia, Bulgaria
Title: *Statistical Machine Learning: Foundations and Present Challenges*

15. July 2022: MLO group seminar; EPFL, Lausanne, Switzerland
Title: *Statistical Aspects of Trustworthy Machine Learning*
16. February 2022: ETH AI Center Postdoctoral Symposium (online)
Title: *Statistical Aspects of Trustworthy Machine Learning*
17. December 2021: Guest lecture at the ETH Zurich course on Trustworthy ML (online)
Title: *Statistical Aspects of Group Fairness-Aware Learning*
18. May 2021: GILT seminar series, University of Milan
Title: *On the Sample Complexity of Adversarial Multi-Source PAC Learning*

Interviews

1. [November 2023: Bulgarian National Radio \(BNR\), Program Horizont](#); Sofia, Bulgaria
Topic: AI safety and regulations
2. [October 2023: Bulgarian National Radio \(BNR\), Program Hristo Botev](#); Sofia, Bulgaria
Topic: Research abroad and in Bulgaria and career at INSAIT
3. [October 2023: Bulgarian National Television \(BNT\), Program “100% awake”](#); Sofia Bulgaria
Topic: Bulgaria’s breakthrough at NeurIPS; appearing together with INSAIT interns
4. June 2023: Fireside chat at VMWare; Sofia Bulgaria
Topic: *Trustworthy Machine Learning*

AWARDS AND SCHOLARSHIPS

Academic	<ul style="list-style-type: none"> - <i>ETH AI Center Postdoctoral Fellowship</i>, 2022 - <i>PODC student travel award</i>, SIGOPS, 2018 - <i>Davies Prize for MMath Part C</i>, Jesus College, Oxford, 2017 - <i>Department of Statistics Prize for MMath Part A</i>, Oxford, 2015 - <i>Open Scholarship</i>, Jesus College, Oxford – 3 awards (2014-2016) - <i>Bulgarian National Scholarship</i> for gifted students – 4 awards (2008 – 2011)
----------	--

SKILLS

IT skills	Experienced with: Python, Tensorflow, NumPy, R, LaTeX, Ubuntu Have worked with: JAX, C++, Matlab, HTML, Mathematica
Languages	Fluent: English, Bulgarian; Conversational: German, Russian

MISCELLANEOUS

Reviewing	- NeurIPS (2021-), ICML (2021-), AISTATS (2022-)
Teaching	ETH Zurich: Teaching Assistant - 09.2022 – 12.2022: Reliable and Trustworthy Artificial Intelligence (RTAI) IST Austria (ISTA): Teaching Assistant - 02.2020 – 04.2020: <i>Formal Methods for Learned Systems</i> - 11.2018 – 01.2019: <i>Deep Learning with Tensorflow</i> - 10.2018 – 11.2018: <i>Statistical Machine Learning</i>
Responsibilities	INSAIT: Internal seminar organizer - 2023 – present: organizing a weekly research seminar IST Austria (ISTA): ML meetup organizer - 2020 – 2021: Organizing a reading group for the ML students at ISTA

IST Austria (ISTA): Student representative in Data Science

- 2018 – 2020: Representing Data Science students at ISTA; providing feedback about related postgraduate courses; welcoming first-year Data Science students

Oxford University Bulgarian Society

- 2015 – 2016: *President* – organising weekly talks, film nights, joint meetings with other university societies; maintaining the society's website
- 2014 – 2015: *Secretary* – communicating with members and applicants