

# Ivan Smirnov

computational social scientist

updated on 05.06.2024

Research Fellow, University of Technology Sydney  
2023– Graduate Research School

Assistant Professor, University of Mannheim  
2022–2023 Business School

---

I was on six months of parental leave, followed by six months of reduced professional activity due to caring responsibilities

---

Research Fellow, RWTH Aachen University  
2020–2021 Chair for Computational Social Sciences and Humanities

Group Leader, Higher School of Economics, Moscow  
2018–2020 Computational Social Science Group

Research Fellow, Higher School of Economics, Moscow  
2014– 2018 Center for Modern Childhood

## Education

2018 Ph.D. (cum laude<sup>1</sup>), Higher School of Economics, Moscow  
Thesis: Students' differentiation by academic performance on a social networking site

2014 M.Sc. Interdisciplinary Approaches to Life Science, Université Paris Cité

2009 Applied Mathematics and Computer Science, St. Petersburg State University  
Specialist degree<sup>2</sup> with a specialization in mathematical statistics

## Teaching Experience

### University courses

2022/23 WS University of Mannheim  
Scientific Programming with Python

2021/22 WS RWTH Aachen  
Social Data Science  
Master Project (Social Data Science Major)

---

<sup>1</sup> In Russia, only one class of honors is used

<sup>2</sup> It is a 5-year degree equivalent to Master

2021 SS RWTH Aachen

Advanced Scientific Programming with Python  
Social Networks  
Seminar Data, Algorithms & Society  
Master Project (Social Data Science Major)

2020/21 WS RWTH Aachen

Introduction to Scientific Programming with Python  
Seminar Data, Algorithms & Society

2019/20 Higher School of Economics

Big Data and Machine Learning in Educational Research  
Introduction to Computational Social Science

Best Course Award based on student evaluations (nomination "Expanding horizons")

2018/19 Higher School of Economics

Introduction to Educational Data Science  
Introduction to Computational Social Science

2014 University of Paris

Introduction to Statistics

## Thesis Supervision

2024 E. Anderlohr (Master, Mannheim) Interpretable detection of political framing  
ongoing, co-supervising with M. Lutz

2024 I. Shakhova (Master, Mannheim) Examining gendered word embeddings and their effects on  
downstream tasks  
co-supervised with M. Lutz

2024 S. Guliyev (Master, Mannheim) Measuring perceptions of veganism on social media: Diachronic  
analysis using word embeddings  
co-supervised with M. Lutz

2023 J. Stritzky (Master, Mannheim) Corporate gender biases in contextual neural word embeddings  
co-supervised with M. Lutz

2022 C. Oprea (Master, RWTH) Determining the impact of toxicity on Wikipedia's talk pages

2021 M. Zubkova (Bachelor, HSE) Using NLP to understand the gender pay gap in the online  
freelance marketplace

2019 M. Gizatullin (Master, HSE) Professional identity development of university students

## Other teaching activities

2022 Summer Institute in Computational Social Science

What is computational social science (CSS) really?  
How to write a (good!) CSS paper  
Potential pitfalls and problems for awesome CSS projects  
CSS Research Project

2020–2021 RWTH Aachen

Student advisor for Master program "Computational Social Systems"

## Work Experience

Prior to academia, I worked as a web developer and a technical leader at several startups.

I also worked on various non-profit projects, for example, I was part of the Teach for Russia founding team (6 people) who raised \$2M to launch the project, and I co-founded Sci.STEPS – a mentorship program for early career researchers.

## Languages

English (IELTS 8), French (DELFI B2), Russian (native), German (beginner)

## Programming Languages

Actively using: Python; used in the past: R, MATLAB, JavaScript, PHP, C#

## Grants, Scholarships, and Awards

2022 Co-organizer, SICSS TU Graz-RWTH from TU GRAZ (€15,000) w/ J. Lasser

2021 Co-organizer, SICSS HSE from SICSS Core Sponsors (\$6,450) w/ E. Sivak and S. Dokuka

2020 ~~Co-organizer, SICSS HSE from SICSS Core Sponsors (\$10,000)~~ (canceled due to COVID-19)

2020 Russian Foundation for Basic Research grant №20-311-90056 w/ M. Bogdanov (~ \$16,000)

2019 PI, Russian Science Foundation grant №19-18-00271 (~ \$261,000)

2018 Co-PI, HSE University Basic Research Program grant №31 (~ \$47,000)

2016 Key Personnel, Russian Foundation for Basic Research grant №16-06-00916 (~ \$26,000)

2016 Institute of Education The Pioneers Award for the Digital Trace project

2014 Advanced Doctoral Program Scholarship, HSE University

2013 SCIRE Foundation Excellence Scholarship

2012 French Government Scholarship

## Service

Program Committee

IC2S2 2024, Wiki Workshop 2024, ICA regional conference 2022

Reviewer

Science; The Web Conference / WWW; ICWSM; Demographic Research; ACM Transactions of Social Computing; PLOS One; Sage Open; Frontiers in Psychology

## Selected Invited Talks

2023 King's College London

Neural word embeddings and gender biases in large corpora

2022 Stanford University

Estimating Educational Outcomes from Students' Short Texts on Social Media

2022 University of Maastricht

Using social media data to understand inequality and segregation in the digital age

2021 TU Graz

Understanding emotional well-being in the digital age

2020 Complexity Science Hub, Vienna

Inequality and well-being in the digital age

## Conference Talks

IC2S2 – International Conference on Computational Social Science

2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024

ICWSM – International AAAI Conference on Web and Social Media

2018, 2019

NetSci-X – International Conference on Network Science

2017, 2020, 2021

SocInfo – International Conference on Social Informatics

2017, 2019, 2022

AERA Conference on Educational Data Science

2020

Wiki Workshop

2023

2023 Think Tank ENGAGE EU

A week long event focusing on tackling gender inequalities to build a more sustainable 'new normal' in Europe

## Publications

Smirnov I. (2024) The Echoes of the 'I': Tracing Identity with Demographically Enhanced Word Embeddings, *arXiv* <https://doi.org/10.48550/arXiv.2407.00340>

Urman A., Smirnov I., Lasser J. (2024) The right to audit and power asymmetries in algorithm auditing, *EPJ Data Science* <https://doi.org/10.1140/epids/s13688-024-00454-5>

Smirnov I., Oprea C., Strohmaier M. (2023) Toxic comments are associated with reduced the activity of volunteer editors on Wikipedia, *PNAS Nexus* <https://doi.org/10.1093/pnasnexus/pgad385>

Bessudnov A., Tarasov D., Panasovets V., Kostenko V., Smirnov I., Uspenskiy V. (2023) Predicting perceived ethnicity with data on personal names in Russia, *Journal of Computational Social Science* <https://doi.org/10.1007/s42001-023-00205-y>

Sivak E., Smirnov I., Dementeva Y. (2022) Online Social Integration and Depressive Symptoms in Adolescents In *International Conference on Social Informatics* [https://doi.org/10.1007/978-3-031-19097-1\\_21](https://doi.org/10.1007/978-3-031-19097-1_21)

Dokuka S., Sivak E., Smirnov, I. (2022) Core but not peripheral online social ties is a protective factor against depression: evidence from a nationally representative sample of young adults, In *Proceedings of NetSci-X Conference* [https://doi.org/10.1007/978-3-030-97240-0\\_4](https://doi.org/10.1007/978-3-030-97240-0_4)

Ahnert G., Smirnov I., Lemmerich F., Wagner C., Strohmaier M. (2021) The FairCeptron: A framework for measuring human perceptions of algorithmic fairness, *Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization* <https://doi.org/10.1145/3450614.3463291>

- Smirnov, I., Lemmerich, F., Strohmaier M. (2021) Quota-based debiasing can decrease representation of already underrepresented groups, *Royal Society Open Science* <https://doi.org/10.1098/rsos.210821>
- Bogdanov M., Smirnov I. (2021) Opportunities and limitations of digital traces and machine learning methods in sociology, *Monitoring of Public Opinion: Economic and Social Changes* <https://doi.org/10.14515/monitoring.2021.1.1760>
- Bekova S., Dementeva J., Smirnov, I. (2021) Factors associated with depression among university students: the role of students satisfaction <https://osf.io/preprints/socarxiv/9fg8t/>
- Dokuka, S., Smirnov, I. (2020) High academic performance is associated with shorter sleep and later bedtimes for young adults <https://arxiv.org/abs/2005.07806>
- Smirnov, I. (2020) Estimating educational outcomes from students' short texts on social media, *EPJ Data Science* <https://doi.org/10.1140/epjds/s13688-020-00245-8>
- Sivak, E. & Smirnov, I. (2020) Measuring Adolescents' Well-Being: Correspondence of Naïve Digital Traces to Survey Data In *International Conference on Social Informatics* [https://doi.org/10.1007/978-3-030-60975-7\\_26](https://doi.org/10.1007/978-3-030-60975-7_26)
- Bekova, S. & Smirnov, I. (2020) Determinants of doctoral degree aspirations: insights from a nationally representative panel study, *Higher Education in Russia and Beyond* <https://publications.hse.ru/en/articles/429127781>
- Sivak, E. & Smirnov, I. (2019) Parents mention sons more often than daughters on social media, *Proceedings of the National Academy of Sciences* <https://doi.org/10.1073/pnas.1804996116>
- Bochaver, A., Dokuka, S., Novikova, M., Sivak, E., Smirnov, I. (2019) Report on Well-being of children in the digital age (*non peer-reviewed*)
- Bochaver, A., Dokuka, S., Sivak, E., Smirnov, I. (2019) Internet Use and Depressive Symptoms in Adolescents: a Review. *Clinical Psychology and Special Education* <https://doi.org/10.17759/cpse.2019080301>
- Smirnov, I. (2019) Schools are segregated by educational outcomes in the digital space, *PLOS One* <https://doi.org/10.1371/journal.pone.0217142>
- Smirnov, I. (2018) Predicting PISA scores from students' digital traces. In *International AAAI Conference on Web and Social Media* <https://ojs.aaai.org/index.php/ICWSM/article/view/14996>
- Smirnov, I., & Thurner, S. (2017) Formation of homophily in academic performance: students change their friends rather than performance, *PLOS One* <https://doi.org/10.1371/journal.pone.0183473>
- Smirnov, I. (2017) The digital Flynn effect: complexity of posts on social media increases over time. In *International Conference on Social Informatics* [https://doi.org/10.1007/978-3-319-67256-4\\_3](https://doi.org/10.1007/978-3-319-67256-4_3)
- Polivanova K., & Smirnov, I. (2017) What is in my profile? Understanding the interests of modern adolescents through VKontakte data, *Educational Studies* <https://doi.org/10.17323/1814-9545-2017-2-134-152>
- Smirnov, I. (2017) Identifying factors associated with the survival and success of grassroots educational innovations, In *Reforms and innovation in education* [https://doi.org/10.1007/978-3-319-60246-2\\_6](https://doi.org/10.1007/978-3-319-60246-2_6)
- Smirnov, I., Sivak E., Kozmina Y. (2016) In search of lost profiles: the reliability of VKontakte data and its importance for educational research, *Educational Studies* <http://doi.org/10.17323/1814-9545-2016-4-106-122>

## Selected Media Coverage

ABC ([The Drum](#), [Triple J Hack](#)), [The Australian](#), [Radio New Zealand](#), [The Times](#), [MIT Technology Review](#), [la Repubblica](#) (it), [CNN](#) (el), [Business Insider](#) (de), [Paris Match](#) (fr), [Nature](#), [Daily Mail](#), [The Times of India](#), [Le Progrès](#) (fr)