

# Sergei Bogdanov, École Polytechnique

## CREATOR OF **NuNERZERO** (5K/MONTH) & **NuNER** (10K/MONTH)

## CO-CREATOR OF **NuEXTRACT** (15K/MONTH) AND **CLIPPINATOR**

📞 +33649985728 | ✉️ improvesergey@gmail.com | 🌐 <https://www.github.com/serega6678> | 🔗 <https://www.linkedin.com/in/svbogdanov/>

## Experience

### NuMind (YC 22), 8 people

Research Engineer (NLP)

Paris

04.2023-05.2024

- Created NuNERZero - Zero-Shot NER Model: **+4.5% over prev. SOTA + 4k context** + developed under MIT License: [HF link](#)
- Co-created NuExtact - **GPT-4-quality Structured Extraction model**, MIT License: [HF link](#)
- Created NuNER - SOTA (EN & Multi) NER Few-Shot Foundation Model: **+7-+20% increase over existing models**: [Hugging Face link](#); [Paper](#)

### Tinkoff, 1000+ people

Research Engineer (CV + NLP)

Moscow

10.2020 - 09.2022

- Developed **SOTA Masked Face Recognition** TPR=0.99 at FPR=0.001
- Gathered **200k images with 99.9% labeling accuracy** with novel automatic dataset gathering technique
- **Deployed 3 ML models** and implemented several gRPC endpoints
- **Doubled detector's speed** by leveraging **Model Compression techniques**: Pruning and Quantization
- **Recruited & mentored intern** Machine Learning Engineer. **Mentee was hired** and received positive feedback

### JetBrains, 1000+ people

Research Engineer intern (NLP)

Moscow

07.2020 - 09.2020

- Proposed model for **Method name generation** task which is comparable to **SOTA**
- Developed model for Code-SearchNet task: **Top-20% out of all competition participants**

## Education

### Ecole Polytechnique, QS top-21, Top-1 French Engineering School

M.S. in Artificial Intelligence, NLP, and Computer Vision

Paris

09.2022 - 10.2023

- Admitted to the 2nd year directly with 16.2k scholarship (highest among all Master's students of the University)

GPA: 3.95/4.00

### Yandex School of Data Analysis, Best National AI Masters

M.S. in Artificial Intelligence, NLP, and Computer Vision

Moscow

09.2020 - 08.2022

- **Courses**: Computer Vision (2d & 3d), Natural Language Processing, ML for Audio and Speech, Rec Sys, RL, Python, GO

GPA: 5.00/5.00

### Moscow Institute of Physics and Technology, Top-1 National Engineering School

B.S. in Computer Science and Applied Math

Moscow

4 years

- **Thesis**: Reliable and Efficient Face Identification during COVID-19

GPA: 3.82/4.00

## Skills

**General** Python, Scala, Golang, C++, SQL, NoSQL, AWS, Docker, GitHub (Git), Hadoop

**ML-related** Pytorch (Torch), Sklearn, Recommender Systems (Rec Sys), Computer Vision (CV), Natural Language Processing (NLP), Large Language Models (LLM)

## Achievements

- 2024 **NLP SOTA** - Co-Creator of the SOTA Open Source information extraction model, [link](#)
- 2024 **NLP SOTA** - My models are the best Zero-shot Token Classification models, [link](#)  
**Open Source commits** - [HuggingFace's Alignment](#)
- 2023 **HuggingFace #1 & #2 models** - My models are most popular Token Classification models, [photo](#)  
**NLP SOTAs** - Created the best Domain-Agnostic English & Multilingual NER Foundation Models, [Blog post](#)  
**230+ GitHub stars repo** - Co-created best Open-Source [AI Programming Assistant](#) and gave a [talk](#)  
**NLP SOTA** - Co-created the best Open-Source Sentiment Analysis Foundation Model, [Blog post](#)  
**Vice President** - VP of the most titled Toastmasters Public Speaking club in France
- 2022 **Apple, Amazon** - Got Research Scientist and Research Engineer offers  
**Absolute winners** - Agrocode Hackathon - Real-time aiming the harvester's dispenser, [video demo](#)  
**4th/150+ teams** - More Tech Hackathon - Rec Sys for news for businesses with trend and insight extraction

- 2021 **CV SOTA** - Created SOTA (at the moment) Masked Face Recognition model
- 2020 **NLP SOTA-like** - Created NLP SOTA-like model for Method name generation task
- 2018 **Top-3%, Winner** - ITMO Informatics Olympiad - Top-2 National Informatics Olympiad, awardee for 2 years  
**Top-3%, Winner** - MSU Mathematics Olympiad - Top-3 National Maths Olympiad  
**1st/1000+ participants** - BMSTU Informatics Olympiad

## Speaking

---

- 2024 **NuNER & NuSentiment - Creating efficient Foundation Models thanks to LLMs, [link](#)**  
Paris NLP Meetup (the biggest NLP meetup in Paris), ~ 80 ppl, in-person, 1 of 2 speakers
- 2023 **NuMind's SOTA NER Foundation Models, [post](#)**  
HuggingFace Conference ~ 2000 ppl, in-person, 1 of 20 booth speakers  
**NuMind's SOTA NER Foundation Models, [stream recording](#)**  
ODS.ai (the biggest Post-Soviet AI community)  
**Overview of our LLM-based Open-Source AI coding assistant, [video](#)**  
DataTalks.Club 30k subs, 500 views
- 2022 **Backbones in Computer Vision(Resnet, EfficientNet, MobileNet, ...)**  
lecture (20 students, online)  
**Paper Reading club**  
Gave several intra-corporate paper-review speeches (10-15 members, online)