

Valentin Malykh

Research Scientist

+7 (903) 966 8133

✉ valentin@maly.hk

📧 val.maly.hk

GitHub: madrugado

Google Scholar: <https://goo.gl/b6UPb5>



Education

2016–2019 **PhD**, *Moscow Institute of Physics and Technology*, Moscow.

(expected) Thesis: Noise Robustness in Various NLP Tasks

2007–2009 **M.Sc.**, *Moscow Institute of Physics and Technology*, Moscow.

Major: Computer Science

2003–2007 **B.Sc.**, *Moscow Institute of Physics and Technology*, Moscow.

Major: Computer Engineering

Experience

Oct 2018–p.t. **Applied Research Scientist**, *VK.com*, Saint-Petersburg.

Work on research problems related to social network, like text classification, summarization, etc.

2016–p.t. **Research Scientist**, *Laboratory of Neural Systems and Deep Learning, Moscow Institute of Physics and Technology*, Dolgoprudny, Moscow Region.

Work on iPavlov.ai project. Research on noise robustness in NLP tasks.

Detailed achievements:

- Published 5 papers,
- including 1 ACL Demo Paper with DeepPavlov library description
- & 2 papers in NIPS Proceedings on results from ConvAI Challenge.

2016–p.t. **Lecturer**, *Moscow Institute of Physics and Technology*, Dolgoprudny, Moscow Region.

I'm teaching Deep Learning in NLP course. The cumulative audience of that course is estimated to 3000 people.

2015–2016 **Developer**, *Yandex*, Moscow.

I was working in Yandex.News on the whole ML stack for ranking and also participated in clusterization development.

Detailed achievements:

- New ranking formula for news clusters.
- Improved news agency ranking.

2014–2015 **Research Engineer**, *Sputnik*, Moscow.

Sputnik is a Russian government-sponsored search engine, which was developed from scratch by a small team of engineers. I was in the Search Quality Department.

Detailed achievements:

- Web pages classifier.
- Malicious documents ranking.

2012–2014 **Research Engineer**, *Cognitive Technologies*, Moscow.

Cognitive Technologies is a company with long history of work in computer vision domain. It created second most common OCR solution in Russia. Nowadays company's main interest is self-driving cars.

Detailed achievements:

- Computer vision & control for a robot car.
- Research at Astarta project – high-load document classifier.

Miscellaneous

2017 **Certified Instructor**, *NVIDIA Deep Learning Institute*.

In addition to being a Certified Instructor, I have authored whole NLP Workshop in NVIDIA DLI and presented it at GTC EU in 2017.

2017 **Organizer**, *ConvAI Challenge*.

ConvAI Challenge is devoted to creation of human-level conversational intelligence. I was co-organizer in both scientific and technical spheres.

2018 **Organizer**, *ConvAI2 Challenge*.

ConvAI2 Challenge is devoted to creation of chat-bots with persona.

2015 **Winner**, *DeepHack.Game*.

DeepHack.Game was a hackathon on Reinforcement Learning applied to Atari games. URL: <http://game.deephack.me>. Our team is a winner of qualification round, and has overall 2nd place.

2016 **Participant**, *DeepHack.Q&A*.

DeepHack.Q&A is a hackathon on NLP, based on Kaggle Allen AI Challenge for school science tests. URL: <http://qa.deephack.me>. Our team has got 12th overall place at Kaggle among 1300 other participants.

Publications

- Generalizable Architecture for Robust Word Vectors Tested by Noisy Paraphrases, V. Malykh, The 6th International Conference On Analysis Of Images, Social Networks, And Texts (AIST), 2017
- The first conversational intelligence challenge. M. Burtsev et al. NIPS 2018 Proceedings Challenge Chapter.
- ConvAI dataset of topic-oriented human-to-bot dialogues. V. Logacheva, M. Burtsev, V. Malykh. NIPS 2018 Proceedings Challenge Chapter.
- Automatic quality evaluation of human-to-chatbot dialogues. V. Logacheva et al. NIPS 2018 Proceedings Challenge Chapter.
- DeepPavlov: Open-Source Library for Dialogue Systems. M. Burtsev et al. Proceedings of ACL 2018 Demo Track, Melbourne, Australia, 2018.
- Robust Word Vectors: Context-Informed Embeddings for Noisy Texts. V.Malykh et al. Proceedings of 4th Workshop on Noisy User Texts, Empirical Methods in Natural Language Processing, Brussels, Belgium, 2018.
- Alekseev, A., Malykh, V., Nikolenko, S., Shenbin, I., and Tutubalina, E. Aspera: Aspect-based rating prediction model. To appear in proceedings of European Conference on Information Retrieval, 2019.

Languages

English Fluent
Russian Native speaker

IELTS Academic 7.0