# Igor Shalyminov


E-mail : ishalyminov@gmail.com
Phone : +447514863061

Website: https://ishalyminov.github.io
Address: School of Mathematical and Computer Sciences, Heriot-Watt University
Edinburgh EH14 4AS
United Kingdom


## About me

I'm a PhD student at the Heriot-Watt University under the supervision of Prof. Oliver Lemon and Dr. Arash Eshghi.
My research interest is in the area of data-efficient methods for natural language processing. My PhD is focused on the techniques for creating practical neural dialogue models with minimum requirements on the amount of data and human supervision.

## Education

### PhD, Computer Science — Oct 2016 — Present

Heriot-Watt University

Research area: Data-efficient methods for neural dialogue systems
Supervisors: Prof. Oliver Lemon, Dr. Arash Eshghi

### Professional Retraining in Data Analysis (MSc equiv.) — Sep 2010 — Jun 2012

Yandex School of Data Analysis (Moscow Institute of Physics and Technology)

Major: Computer Science

GPA: 4,66

### Specialist (MSc equiv.) - with distinction — Sep 2005 — Jun 2010

Moscow State University of Instrument Engineering and Computer Science

Major: Computing machines, systems, and networks
GPA: 5,0

Defended thesis “A module for generating data for equipment reliability estimation in the 'Prognosis' software suite”

## Awards

Dialogue System Technology Challenge (DSTC) 8, Fast Domain Adaptation - 1st place, 2019
EMNLP 2019 - Student Travel Scholarship, 2019
SCAI workshop@EMNLP 2018 - travel grant from sponsors (Microsoft, Facebook, Google, and TextKernel), 2018
Amazon Alexa Prize 2018 (part of team Alana) - 3rd place prize + research grant, 2018
Amazon Alexa Prize 2017 (part of team Alana) - 3rd place prize + research grant, 2017
James Watt Scholarship from Heriot-Watt University, 2016

# Work experience

## Microsoft Research Montréal — Jul 2019 — Oct 2019

Research Intern

Mentor: Dr. Hannes Schulz
Project: Dialogue State Tracking Challenge (DSTC) 8 - Fast Domain Adaptation

## Microsoft Research AI — Jun 2018 — Sep 2018

Research Intern

Mentor: Dr. Sungjin Lee

Project title: "Improving robustness of goal-oriented dialog systems in an unsupervised way".

## Yandex — Aug 2013 — Sep 2016

Software Engineer

Projects I participated in:

- Voice interfaces for applications - dialogue management module
- Voice analytics for callcenters - search engine for voice data
- Russian National Corpus - search engine for linguistic research

## Yandex School of Data Analysis (Yandex / Moscow Institute of Physics and Technology) — Sep 2013 — Jun 2015

Teaching Assistant

Natural Language Processing course taught by Dr. Alexey Zobnin (Computer Science department, final year course).

My resposibilities for the course were:

- designing problems for the seminar classes;
- conducting the classes;
- helping the professor with designing home assignments and reviewing students' submissions.

## Microsoft Research — May 2014 — Aug 2014

Research Intern

Mentors: Dr. Yuxiong He, Dr. Sameh Elnikety

I worked on a query optimizer for a term distributed search engine. During my internship, I completed the following tasks:

- built combined analytical and statistical models of a search query for estimating its latency and amount of data for network communication;
- designed and implemented algorithms for producing optimal query plans in terms of minimizing either of the two objectives: latency or network communication.

The prototype of the query optimizer and distributed experimental platform was implemented in Java and Python.

## Yandex — May 2012 — Aug 2013

Junior Software Engineer

Participated in development of Russian National Corpus . I maintained the search engine at its existing architecture and explored the possibilities of improving it for better performance and reliability.

## Yandex — Jun 2011 — Dec 2011

Software Engineer Intern

Mentor: Dr. Alexey Zobnin

Built a prototype of an N-gram search service for Russian National Corpus. The service consisted of two components:

- N-grams storage - a compressed delta-encoded word trie written in C++;
- Search client - an http server in Python receiving user requests and fetching the results from the N-gram storage.

### Intel Corporation June 2010 — Aug 2010

Software Engineer Intern

Mentor: Dr. Denis Makoshenko

As a part of Intel Summer School 2010, I participated in Intel C/C++ Compiler development with a research project “Enhancing the methodology to access the bottom boundary of software pipelined cycle execution time as the application of an algorithm for finding a maximum cycle in a directed graph”.

### Moscow State University of Instrument Engineering and Computer Science 2006 — 2012

Software Engineer

Participated in the development of “Prognosis” - a software suite for reliability calculation of the equipment used in nuclear energetics.

## Additional Education

Summer Schools attended:

- Microsoft School on Algorithms for Massive Data (ALMADA). Attended lectures on the effective algorithms for External Memory computational model, algorithms and techniques for large-scale graph problems. Did theoretical assignments on graph algorithms and a Linq programming assignment in Linq (2013)
- Yandex Summer School On Distributed Computing. Attended lectures on Information Retrieval and MapReduce computing model. Did a practice assignment in Hadoop (2011)
- Intel Summer School. Attended lectures and webinars on parallel computing and processor architectures. Learned Intel IA64 architecture (Itanium I and II), participated in Intel C++ Compiler development. (2010)

University Courses Taken:

- Formal Semantics and Typology of Anaphora taught by Barbara H Partee. National Research University Higher School Of Economics (2014)

MOOCs Taken:

- Bayesian Statistics: Techniques and Models - Coursera (2019)
- Bayesian Statistics: From Concept to Data Analysis - Coursera (2019)
- Bayesian Methods for Machine Learning - Coursera (2018)
- Deep Learning Specialization - Coursera (2018)
- Neural Networks For Machine Learning - Coursera (2017)
- Artificial Intelligence for Robotics - Udacity (2016)
- Deep Learning - Udacity (2016)
- Machine Learning: Reinforcement Learning - Udacity (2015)
- Intro to Artificial Intelligence - Udacity (2013)
- Intro To Statistics - Udacity (2012)

Certifications:

- Microsoft Certified Professional - 70-536 Microsoft .NET Application Development Foundation
- Electronic office programmer - C++ and Pascal programming (Specialist Training Centre at the Bauman Moscow State Technical University)

# Tests Taken

- TOEFL iBT: 99 (Reading - 26, Listening - 24, Speaking - 24, Writing - 25)
- GRE: Verbal Reasoning - 152, Quantitative Reasoning - 161, Analytical Writing - 4.0
- IELTS Academic UKVI: 8.0 (Listening - 8.5, Reading - 8.5, Writing - 7.5, Speaking - 7.0)

# Language Ability

Russian - native, English - upper-intermediate (C1), Spanish - beginner (A1)

# Relevant Skills

**Programming Languages:** Python, C++, Java, Perl | **Frameworks:** Tensorflow, PyTorch, scikit-learn, numpy/scipy, pandas, Lucene | **Databases:** mongodb, MySQL | **Tools:** Jupyter notebook, matplotlib | **Core Technical Skills:** classic algorithms and data structures, mathematical foundations, machine learning algorithms, natural language processing

# Interests

Languages, photography, electronic music production.

# References

Oliver Lemon, PhD
Professor, School of Mathematical & Computer Sciences, Heriot-Watt University
Email: o.lemon@hw.ac.uk

Arash Eshghi, PhD
Assistant Professor, School of Mathematical & Computer Sciences, Heriot-Watt University
Email: a.eshghi@hw.ac.uk

Sungjin Lee, PhD
Principal Machine Learning Scientist, Amazon Alexa AI
Email: sungjinl@amazon.com

Hannes Schulz, PhD
Senior Researcher, Microsoft Research Montréal
Email: haschulz@microsoft.com

# Publications

**2019**

**Igor Shalyminov**, Sungjin Lee, Arash Eshghi, Oliver Lemon. Data-Efficient Goal-Oriented Conversation with Dialogue Knowledge Transfer Networks. EMNLP 2019

**Igor Shalyminov**, Sungjin Lee, Arash Eshghi, Oliver Lemon. Few-Shot Dialogue Generation Without Annotated Data: A Transfer Learning Approach. SigDial 2019Sungjin Lee and Igor Shalyminov. Contextual Out of domain Utterance Handling with Counterfeit Data Augmentation. ICASSP 2019

**2018**

**Igor Shalyminov** and Sungjin Lee. Improving Robustness of Neural Dialog Systems in a Data-Efficient Way with Turn Dropout. NeurIPS 2018 workshop on Conversational AI

**Igor Shalyminov**, Arash Eshghi, and Oliver Lemon. Multi-Task Learning for Domain-General Spoken Disfluency Detection in Dialogue Systems. SemDial 2018

**Igor Shalyminov**, Ondrej Dusek, and Oliver Lemon. Neural Response Ranking for Social Conversation: A Data-Efficient Approach. Search-Oriented Conversational AI, an EMNLP 2018 Workshop

Amanda Cercas Curry, Ioannis Papaioannou, Alessandro Suglia, Shubham Agarwal, **Igor Shalyminov**, Xinnuo Xu, Ondřej Dušek, Arash Eshghi, Ioannis Konstas, Verena Rieser and Oliver Lemon. Alana v2: Entertaining and Informative Open-domain Social Dialogue using Ontologies and Entity Linking. Alexa Prize proceedings, 2018

**2017**

**Igor Shalyminov**, Arash Eshghi, and Oliver Lemon. Challenging Neural Dialogue Models with Natural Data: Memory Networks Fail on Incremental Phenomena. SemDial 2017

Arash Eshghi, **Igor Shalyminov**, and Oliver Lemon. Bootstrapping incremental dialogue systems from minimal data: the generalisation power of semantic grammars. EMNLP 2017

Arash Eshghi, **Igor Shalyminov**, and Oliver Lemon. Bootstrapping Dialogue Systems: Using a Semantic Model of Dialogue to Generalise from Minimal Data. LaML 2017

Arash Eshghi, **Igor Shalyminov**, and Oliver Lemon. Interactional Dynamics and the Emergence of Language Games. FADLI 2017

Ioannis Papaioannou, Amanda Cercas Curry, Jose L. Part, **Igor Shalyminov**, Xinnuo Xu, Yanchao Yu, Ondrej Dušek, Verena Rieser, Oliver Lemon. Alana: Social Dialogue using an Ensemble Model and a Ranker trained on User Feedback. 2017 Alexa Prize Proceedings

Ioannis Papaioannou, Amanda Cercas Curry, Jose L. Part, **Igor Shalyminov**, Xinnuo Xu, Yanchao Yu, Ondrej Dušek, Verena Rieser, Oliver Lemon. An Ensemble Model with Ranking for Social Dialogue. NIPS Workshop on Conversational AI 2017

**2010**

**Igor Shalyminov**. Using Policies in Designing of a Library for Cell Operations in ‘Prognoz’ Software Package. Software for PC-based systems for various purposes #13, pp. 101–107. ISBN 978-5-8068-0462-5

**2009**

**Igor Shalyminov**. Algorithms for calculation of constructions fault probability in ‘Prognoz’ software package. Applied Mathematics and Mathematical Modelling, pp. 70–77. ISBN 978-5-8122-1023-6

**2008**

Aleksandr Aleksandrov, Alexey Vostrikov, and **Igor Shalyminov**. An algorithm for generation geometric objects for crack

formation analysis. Software for PC-based systems for various purposes #11, pp. 145–148. ISBN 978-5-8068-0418-2

**2007**

Aleksandr Aleksandrov and **Igor Shalyminov**. An algorithm for definition of crack formation probability in statistical modeling. Software for PC-based systems for various purposes #10, pp. 11–15. ISBN 978-5-8068-0393-2

# Professional Service

Program committee: AAAI 2020, EMNLP 2019, SCAI@IJCAI 2019, ACL 2019, ACL 2018 (last-minute reviewer), COLING 2018, EMNLP 2018 (secondary reviewer), SCAI@EMNLP 2018