

# Andrei (Andreas) Shcherbakov

[andreas@softwareengineer.pro](mailto:andreas@softwareengineer.pro)  
<http://softwareengineer.pro>

**Research Scientist, Software Expert**

PhD candidate (*viva passed in Oct 2021*), ex-Intel research scientist

- **Natural Language Processing**
- **Machine Learning**
- **Information Retrieval**
- **Formal Logic, Reasoning, Verification**
- **Algorithms & Optimization**
- **Automata and graphs**

## TECHNICAL EXPERTISE

- *Computer Science & Technology – research, development, teaching:*
  - Natural and Formal Language processing, compiling, parsing, automata/transducers, regular expressions, structured information extraction etc.
  - Machine learning (including “deep learning”; mostly natural language related applications). Building standard and custom model architectures, both neural and non-neural. Experience with high performance computational graph frameworks (TensorFlow, DyNet).
  - Logic/combinatorial problems: Formal and dynamic verification, reasoning, proving, simulation and modeling, equation [SAT/SMT] solver based techniques.
  - Full stack development of low-scale services.
  - *Programming:* Python, C++/C, Java, functional languages (OCaml), PHP, JavaScript, Linux, Perl, Shell, data processing languages (Octave, R), visualization (Tcl/Tk, Qt). Rapid familiarization to new languages.

## EXPERIENCE



### The University of Melbourne

04/2015 – present

- Academic teaching:
  - “Knowledge Technology” (Information Retrieval, Probability, Machine Learning)
  - Web Search and Text Analysis (Natural Language Processing, Information Retrieval)
  - Stream Computing Applications (Statistical sketching of big data)
  - Database Modeling (Mostly about SQL databases)
  - Foundations of computing (Python programming – an undergraduate summer course)
  - Tutor in Academic Excellence Program for Indigenous students.
- Research:
  - Multi-word expression detection;
  - Prevention of looping effects in neural decoders;
  - Multiple string alignment; producing automata from string sample sets;
  - Modeling morphology;
  - Phonetics modeling.
- Software development and research assistance
  - (Joint research project with Austin Health) -- Automated annotation/extraction of sensitive data in clinical records ~ full-stack development, administration.
  - Word prediction application for studying under-resourced languages -- “*Word Generator*” ( <http://paradisec.org.au/wordgen/wg.php> ) ~ algorithm research, web service development.

**Neo AI Systems P/L ~ Project Lead**  
01/2017-10/2017

- R&D in Natural Language Understanding:
  - (Elastic Search based) semantic search engine for news.
  - Anaphora/coreference resolution flow for Web texts.
  - Annotation and visualization applications.



**Intel Corporation ~ Research Scientist**  
09/2008 – 12/2013  
2009 - 2012 - Logic Verification Group (Moscow) tech lead

- R&D Software Verification algorithms and tools
  - Proposed and implemented multiple efficient heuristics for Software testing that led to 6+ times speed-up of testing without loss of coverage.
  - Developed an event-based flow for the model equivalence checking at different abstraction levels.
  - Explored applicability of equation solvers in layout design problems. Prototyped a pioneering flow for rule-compliant layout elements generation.
  - Developed customized data mining applications for diagrams and text specifications.
  - Developed a power consumption estimation flow for early stages of a microarchitecture design.



**Intel Corporation ~ Senior Software Engineer**  
08/2004 – 12/2013

- Temporal logic compilers for formal and dynamic verification
  - Actively drove adoption of a modern property specification language (SVA) at Intel:
  - Developed a set of key features within the Intel's native back-end for property verification ("Property compiler")
  - Developed a WYSIWYG advisory tool for composing property specifications
  - Advised an internship project



**Moscow Center for SPARK technology (MCST) ~ Software R&D Engineer**  
03/1997 - 08/2004

- Design automation software development [outsourcing for Avant!, Synopsys]
  - Rule checking and what-if advisor tools, framework, graphics, configuration

Earlier:

**Lebedev Institute of Precision Mechanics and Computer Engineering (IPMCE) ~ Intern Engineer**

- Developed a customized array logic synthesis tool.

**Sinq Ltd (small enterprise) ~ Co-founder, Electronics designer**

## MISC PERSONAL PROJECTS

- A Regular Expression extraction service – <http://regexus.com>
- A thermo-hydraulic service site with liquid flow visualizations – <http://simmixflow.com>
- A translator of semantic relations into SQL database

## LINKS

Google Scholar: <https://scholar.google.com/citations?user=h7Skt54AAAAJ>

GitHub: <https://github.com/andreas-softwareengineer-pro>

Stack Overflow: <https://stackoverflow.com/users/4819357/andreys-scherbakov>

## EDUCATION

- *Ph.D. in Engineering Sciences [Computer science]*, **Moscow Technological University (MIREA)**, Moscow, 2021 (awaiting grant approval).
- *M.S. in Computers & Networks*, **Moscow Institute for Physics & Technology (Technical University) (MIPT)**, Moscow, 1993.

## OTHER

- Creative approach to task solving. Strong analytical skills. Prefers diverse, multi-aspect challenging tasks. Motivated by cognition.
- Extended engineering & natural science education and erudition.
- Willing to mentor, travel, advertise and communicate.

## PUBLICATIONS

- Shcherbakov, A., Muradoglu, S., & Vylomova, E. (2020, December). Exploring Looping Effects in RNN-based Architectures. In Proceedings of the The 18th Annual Workshop of the Australasian Language Technology Association (pp. 115-120).
- Shcherbakov, A., & Vylomova, E. (2019). A string-to-graph constructive alignment algorithm for discrete and probabilistic language modeling. In *Proceedings of the The 17th Annual Workshop of the Australasian Language Technology Association* (pp. 186-191).
- Scherbakov, A. (2018). A fast algorithm for finding accessible vertices in a control flow graph with path restrictions. *Problems of Perspective Micro- and Nanoelectronic Systems Development - 2018. Proceedings / edited by A. Stempkovsky, Moscow, IPPM RAS, 2016.* (pp. 92-96). [in Russian].
- Vylomova, E., Shcherbakov, A., Philippovich, Y., & Cherkasova, G. (2017). Men Are from Mars, Women Are from Venus: Evaluation and Modelling of Verbal Associations. In *International Conference on Analysis of Images, Social Networks and Texts* (pp. 106-115). Springer, Cham.
- Shcherbakov, A., Vylomova, E., & Thieberger, N. (2016) Phonotactic Modeling of Extremely Low Resource Languages. In *Australasian Language Technology Association Workshop 2016* (p. 84).
- Scherbakov, A. (2016). Branching Alignment based Synthesis of Regular Expressions. *Supplementary Proceedings of the Fifth International Conference on Analysis of Images, Social Networks and Texts (AIST 2016)*. (pp. 315-328).
- Shcherbakov A.S. (2016). A fast algorithm for data dependency tracking in Software and Firmware analysis and testing. *Problems of Perspective Micro- and Nanoelectronic Systems Development - 2016. Proceedings / edited by A. Stempkovsky, Moscow, IPPM RAS, 2016. Part II.* (pp. 76-83). [in Russian].
- Shcherbakov A.S. (2014). Powering Directed Automated Program Testing by Parallel Branch Alternation with Data Tracking. *Problems of Perspective Micro- and Nanoelectronic Systems Development - 2014. Proceedings / edited by A. Stempkovsky, Moscow, IPPM RAS, 2014. Vol. 2.* (pp. 15-21). [in Russian].
- Shcherbakov A.S. (2012). Conditional jump re-alternation Limiting based speed-up of Directed Automated program testing. *Problems of Perspective Micro- and Nanoelectronic Systems Development - 2012. Proceedings / edited by A. Stempkovsky, Moscow, IPPM RAS, 2012.* (pp. 89-94). [in Russian].

- Shcherbakov A.S. (2010). High level model based verification of digital circuits behavior. *Problems of Perspective Micro- and Nanoelectronic Systems Development - 2010. Proceedings / edited by A. Stempkovsky, Moscow, IPPM RAS, 2010.* (pp. 42-47). [in Russian].

### **Shared tasks**

- Pimentel, Tiago, et al. "SIGMORPHON 2021 Shared Task on Morphological Reinflection: Generalization Across Languages." *Proceedings of the 18th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology.* 2021.
- Scherbakov, A., Whittle, L., Kumar, R., Singh, S., Coleman, M., & Vylomova, E. (2021, June). Anlirika: an LSTM–CNN Flow Twister for Spoken Language Identification. In *Proceedings of the Third Workshop on Computational Typology and Multilingual NLP* (pp. 145-148).
- Scherbakov, A. (2020, July). The UniMelb submission to the SIGMORPHON 2020 shared task 0: Typologically diverse morphological inflection. In *Proceedings of the 17th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology* (pp. 177-183).
- Shcherbakov, A., Vylomova, E., Liu, F., & Baldwin, T. (2016). VectorWeavers at SemEval-2016 Task 10: From Incremental Meaning to Semantic Unit (phrase by phrase). *Proc. of SemEval. San Diego, California, USA.*

### **LECTURES & PRESENTATIONS**

- "Property Verification" as a part of "Math Basics of CAD" course, Moscow Institute for Physics & Technology (Technical University) (MIPT, 2010-2013)
  - Realization of solver based techniques for dynamic software verification. Microsoft Summer School in Software Engineering and Verification, Moscow, Russia, on July 17–27, 2011 (Invited lecturer)
-