

# Anton BELY

abel@jhu.edu · avbelyy.github.io

Scholar · GitHub · LinkedIn

## EDUCATION

---

- |                   |  |
|-------------------|--|
| AUG '19 — PRESENT | <b>PhD in Computer Science (NLP)</b> , Johns Hopkins University, U.S.A.<br>Advisors: Benjamin Van Durme (Primary), Vladimir Braverman (Secondary)  |
| SEP '13 — JUN '18 | <b>B.Sc. in Informatics and Applied Mathematics</b> , ITMO University, Russia<br>Thesis: <a href="#">Construction and Quality Evaluation of Heterogeneous Hierarchical Topic Models</a><br>Advisors: Andrey Filchenkov (Primary), Konstantin Vorontsov (Secondary) |

## RESEARCH INTERESTS

---

- Large-scale information retrieval
- Streaming and randomized algorithms
- Schema induction and structured text generation

## PUBLICATIONS

---

### 1. Script Induction as Association Rule Mining.

Belyy, A., & Van Durme, B. (2020). In *Proceedings of the 1st Joint Workshop on Narrative Understanding, Storylines, and Events (NUSE)*, pp. 55-62. [\[paper\]](#) [\[slides\]](#) [\[code\]](#)

### 2. Improved Evaluation Framework for Complex Plagiarism Detection.

Belyy, A., Dubova, M., & Nekrasov, D. (2018). In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (ACL)*, Vol. 2, pp. 157-162. [\[paper\]](#) [\[poster\]](#) [\[code\]](#)

### 3. Framework for Russian Plagiarism Detection Using Sentence Embedding Similarity and Negative Sampling.

Belyy, A., & Dubova, M. (2018). In *Proceedings of the 24th International Conference on Computational Linguistics and Intellectual Technologies (Dialogue)*, Issue 17, pp. 96-109. [\[paper\]](#) [\[slides\]](#) [\[code\]](#)

### 4. Quality Evaluation and Improvement for Hierarchical Topic Modeling.

Belyy, A., Seleznova, M., Sholokhov, A., & Vorontsov, K. (2018). In *Proceedings of the 24th International Conference on Computational Linguistics and Intellectual Technologies (Dialogue)*, Issue 17, pp. 110-123. [\[paper\]](#) [\[slides\]](#)

### 5. MEMOIR: Multi-class Extreme Classification with Inexact Margin.

Belyy, A., Sholokhov, A. (2018) [\[preprint\]](#)

## WORK EXPERIENCE

---

- |                   |   |
|-------------------|---|
| SEP '17 — AUG '19 | <b>Senior Data Scientist</b> , Tochka Bank, Russia <ul style="list-style-type: none"><li>• Anti-money laundering: prediction and scoring models</li><li>• Client communications: dialog intent recognition and exploratory analysis</li></ul> |
| SEP '15 — SEP '16 | <b>Software Developer</b> , VK.com, Russia <ul style="list-style-type: none"><li>• Fraud (anomaly) detection in ads' clicks</li><li>• Classification of trustworthy advertisers</li></ul>   |
| JUL '15 — SEP '15 | <b>Software Developer Intern</b> , JetBrains, Russia <ul style="list-style-type: none"><li>• Implementation of a garbage collection algorithm for <a href="#">dotMemory</a> profiler</li></ul>  |

## RESEARCH EXPERIENCE

---

MAR '18 — JUN '18	<b>Visiting Student</b> , Université Grenoble Alpes, France Supervisors: Massih-Reza Amini, Yury Maximov <ul style="list-style-type: none"><li>• Extreme multi-class classification with Pegasos and MIPS algorithms <a href="#">[report]</a></li></ul>
SEP '17 — FEB '18	<b>Research Assistant</b> , ITMO University, Russia Supervisor: Andrey Filchenkov <ul style="list-style-type: none"><li>• Novel evaluation metric for external plagiarism detection <a href="#">[paper]</a></li><li>• Framework for external plagiarism detection in Russian <a href="#">[paper]</a></li></ul>
MAR '17 — AUG '17	<b>Research Assistant</b> , Russian Academy of Sciences, Russia Supervisor: Konstantin Vorontsov <ul style="list-style-type: none"><li>• Hierarchical topic modeling for exploratory search over heterogeneous sources <a href="#">[paper]</a></li><li>• Topic-model driven exploratory search engine (full-stack implementation) <a href="#">[code]</a></li></ul>
JUL '14 — MAR '15	<b>R&amp;D Intern</b> , Synopsys, Russia Supervisor: Sergey Yakushkin <ul style="list-style-type: none"><li>• Static memory allocation for micro-controllers using graph <math>k</math>-coloring methods</li><li>• Implementation of three graph coloring algorithms, technical report with results</li></ul>

## TEACHING EXPERIENCE

---

SPRING 2021	<b>Intro Algorithms 601.433/633 (Head TA)</b> , JHU
SEP '17 — AUG '19	<b>Natural Language Processing (TA)</b> , Online <a href="https://coursera.org/learn/language-processing">https://coursera.org/learn/language-processing</a>

## REVIEWING

---

ICDT	2021 (secondary)
ICLR	2021 (secondary)
AKBC	2020 (secondary)
ACL	2020 (secondary)
AINL	2018, 2019, 2020

## COURSEWORK

---

AI / ML	Natural Language Processing, Machine Learning, Optimization Theory
ALGORITHMS	Discrete Mathematics, Algorithms & Data Structures, Randomized Algorithms, Computational Geometry, Complexity Theory, Sketching and Indexing, Information Retrieval
MATHEMATICS	Linear Algebra, Calculus, Probability Theory, Intro Stats, Statistical Theory, Information Theory, Mathematical Logic, Type Theory, Language & Automata Theory, Game Theory
PROGRAMMING	C++, Java, Haskell, Assembler, Web Development, Mobile Development (Android)
SYSTEMS	Computer Architecture, Parallel & Distributed, Operating Systems, Networks, Databases

## SKILLS AND TOOLS

---

SPOKEN LANGUAGES	Russian (Native Proficiency), English (Bilingual Proficiency), Chinese (Elementary)
PROGRAMMING	Python, Haskell, C / C++, Java, JavaScript, Assembler, *nix shell, *SQL, PHP, HTML
FRAMEWORKS	Research: NumPy, Pandas, PyTorch, fairseq, nmslib, BigARTM WebDev: Node.JS, Socket.IO, MongoDB, flask, memcache
TOOLS	Docker, git, LaTeX, (C)Make, JetBrains IDEs