ALISA ZHILA, PhD
Waltham, MA, USA ▪ +1-508-308-41-00

[alisa.zhila@gmail.com] ▪ [nlp.cic.ipn.mx/~alisa] ▪ [www.linkedin.com/in/alisazhila] ▪ [github.com/alisatl]

Objectives
Research & Development/Data Analyst/Software Engineer; availability **Fall 2017**

Summary
Currently Researcher/Engineer in Computer Vision group in IBM Watson. **Core Competencies:** Natural Language Processing: Information Extraction, Semantic Similarity, Topic Modeling, Knowledge Bases, Ontologies/Semantic Hierarchies, Text Analysis ▪ Computer Vision: Caffe, OpenCV ▪ Machine Learning ▪ Microservices/SaaS

Academic Education and Professional Training
PhD, 2014, Computer Science, GPA 10/10 (with Honors) ▪ Instituto Politécnico Nacional (CIC-IPN), Mexico
   Thesis: Open Information Extraction using Constraints over Part-of-Speech Sequences ▪ Microsoft Research Latin America Fellow
MS, 2008, Applied Physics and Mathematics, GPA 4.86/5 (with Honors) ▪ Moscow Institute of Physics and Technology, Russia
BS, 2006, Applied Physics and Mathematics, GPA 4.85/5 (with Honors) ▪ Moscow Institute of Physics and Technology, Russia
Translator in Professional Communications, 2006 (1 yr program), GPA 5.0/5 (with Honors)
   Department of Foreign Languages, Moscow Institute of Physics and Technology**,** Russia

Professional Experience
IBM, Littleton, MA   *Jan 2016 – present*
Watson Computer Vision (Oct 2016 - present)
   Semantic post-processing of Image Tagging output for multiple object detection ▪ Leading an intern's project on base tagging vocabulary expansion with a classifier ▪ Developing, deploying, maintaining cloud services for Visual recognition; DevOps
Watson Language Products (Jan-Oct 2016)
   Question answering pipeline, components for the microservice orchestration approach ▪ Performance analysis of Solr vs Elasticsearch ▪ DevOps and Provisioning as needed (Docker, Asgard)
IBM, Internship, Littleton, MA   *Feb – Mar 2015 (8 weeks)*
Watson Language Products
   Open Information Extraction based on Predicate-Argument Structures ▪ Designed and implemented a Relation Extraction system based on Open Information Extraction principles using predicate-argument structures
Yahoo, Internship (Tech Intern III), Sunnyvale, CA   *Jul – Sep 2014 (12 weeks)*
   Classification of Log Emergency Levels in Alert Monitoring Systems and Its Improvement ▪ Feature engineering; Compared simple threshold-based classifier (heuristics, no training) to LogReg classifier ▪ Data analysis in R
Oracle, Internship, Oracle MDC, Guadalajara, Mexico   *Mar – Jun 2014 (3 months)*
Product Group: Oracle Spatial and Graph
   Information Extraction System for Semantic Indexing Functionality Demo ▪ Information Extraction from text based on syntactic constraints (rules, heuristics) ▪ Converting extractions into RDF/XML format ▪ Oracle Semantic Technologies, SPARQL queries
Microsoft Research, Internship (Research Intern), Redmond, WA   *Jun – Sep 2012 (12 weeks)*
Project: Measuring Degrees of Relational Similarity between Word Pairs
   Feature engineering: relation-specific information, lexical patterns, vector space ▪ Development of a System for feature extraction and processing ▪ Classifier Training (LogReg) using an existing tool ▪ Publication in NAACL'2013
ABBYY Software Ltd., www.abbyy.com, Moscow, Russia   *Nov 2007 – Dec 2010 (3 years)*
Product: Machine Translation project Compreno based on semantic ontology
Position: Acting Head of IT Terminology Group, Sep 2009 – Dec 2010
   Syntactic-semantic rules for translation of words in context and collocations ▪ Managing semantic ontology manual population ▪ Managing working process of IT terminology group of 3 in-office and 6 remote employees
Position: Linguist, IT Terminology Group, Nov 2007 – Sep 2009
   Corpus analysis for word senses ▪ Formal description of word semantics including collocations and corresponding syntactic-semantic rules ▪ Semantic hierarchy manual population ▪ Lexicographic tasks

Skills
Programming Languages: Java, Python, Golang ▪ **Software Technologies**: Git, SQL, Bash, Bash, Unix/Mac ▪ **DevOps**: Docker, Kubernetes ▪ **Search Engines**: Solr, ElasticSearch ▪ **Computer Vision**: Caffe, OpenCV ▪ **Machine Learning**: NumPy, Scikit-learn, OpenCV, caffe, Octave ▪ **Data Analysis**: R, Python ▪ **Natural Language Processing**: NLP toolkits: GATE NLP, OpenNLP, NLTK, FreeLing, StanfordNLP — Methods and systems for Open Information Extraction — Vector Space Models (word2vec), feature engineering — Context Clustering, Word Sense Discrimination (SenseClusters tool) — Ontologies and Knowledge Graphs: WordNet, NELL, ProBase; RDF, SPARQL ▪ **Languages**: English (fluent), Spanish (fluent), Russian (native), German (reading)

Most Recent Publications & Patents (full list at *http://nlp.cic.ipn.mx/~alisa*)
Working with Deeply Nested Documents in SolrCloud @Lucene/SolrRevolution, 2016 ▪ Open Information Extraction for Spanish Language based on Syntactic Constraints @ACL SRW, 2014 ▪ Relational similarity measurement (Patent US 20140249799 A1, Publication Date: Sept. 4, 2014) ▪ Combining Heterogeneous Models for Measuring Relational Similarity @NAACL, 2013.