Oleksandr NARYKOV

CONTACT INFORMATION

LOCATION:	Boston, MA	**LINKEDIN:**	www.linkedin.com/in/oleksandr-narykov	**WEBSITE:**	narykov.com
PHONE:	857-999-5317	**EMAIL:**	onarykov@wpi.edu		

SKILLS

Machine Learning
Supervised, Semi-Supervised
Deep Learning, Clustering
scikit-learn, TensorFlow
pyTorch, OpenCV

Data Analysis
Data Mining, Data visualization
Network Science
pandas, SQL, jupyter, Cytoscape
bokeh, seaborn, ggplot2

Bioinformatics & Comp Bio
RNA-Seq, Gene Expression Analysis
and Pathway Analysis, Proteomics
BioPython, Bioconductor, bcbio-nextgen
SMART-seq, Drop-seq, SomaScan

Dev-Ops
Docker, Linux, bash

High Performance Computing
Hadoop, pySpark, GROMACS, NAMD

Software Engineering
Python, C++, R, C, Java, git

Web Applications
C#, .Net Framework, flask, JS
NGINX, Apache

Biological Databases
NIH, EBI, TCGA, GTEx
UniProt, RCSB, KEGG

Image Processing
pyTorch, Keras
OpenCV, scipy

WORK EXPERIENCE

May 2019 — now

WORCESTER POLYTECHNIC INSTITUTE (WPI)
Research Assistant

My activity consists of building machine learning models for the biological data (**RNA-Seq**, **protein interactomes**, **protein structures**, **clinical trials**). For the last year I worked on **molecular dynamics** simulations for the SARS-CoV-2 proteins and envelope. I successfully finished 4 major research projects, currently preparing publications.

ALT-IN Tool - project to build machine learning model for prediction of gene isoforms interactions. It leverages existing interactome data and includes *alternative splicing-specific* information. I analyzed several machine learning approaches, including utilization of Learning Under Privileged Information (LUPI) paradigm. I handled machine learning-related tasks, a part of data extraction and case study analysis and produces deliverables publicly available on https://github.com/korkinlab/altintool and https://hub.docker.com/r/narykov/alt-in.

Molecular dynamics of SARS-CoV-2 envelope - Ongoing project close to the completion. Collaborative effort on bringing together accuate stoichiometry, geometry, and structural information on SARS-CoV-2 envelope. My responsibilities include part of the structural modeling of membrane (M) protein and system integration, as well as performing molecular dynamics simulations on **Frontera supercomputer** (https://frontera-portal.tacc.utexas.edu/) and network analysis on protein connectivity.

Clinical trials on Gulf War Illness treatment - I was responsible for identifying gene set that exhibit response to accupuncture treatment based on *SomaScan* proteomics data. First, I evaluated different strategies for intra- and interplate normalization. Then I employed a range of *statistical methods* and successfully identified relevant genes. This discovery was further supported by literature search. I also conducted a preliminary work on *biomarkers identification* for the successful treatment. This work was done as a part of a joint grant application.

Alternative splicing impact factor - The goal of this project is to distill quantitative measure of impact that *alternative splicing* has on protein function. My responsibilities include formulating a *mathematical description* of this concept and building *visualizations* that help people to extract information based on aforementioned measure. *Transcriptomics data* for case studies is based on GTEx and TCGA repositories.

scRNA-Seq Transcriptomics - Project dedicated to evaluation and development of *transcriptomics* analysis strategies for single-cell RNA-Seq (**scRNA-Seq**) for SMART-based and UMI-based protocols, including study of paired scRNA-Seq and bulk RNA-Seq samples.

DISPOT - **D**omain **I**nteraction **S**tatistical **POT**ential. Tool for computing domain-based statistical potentials from SUPERFAMILY annotation. Availability - GitHub: https://github.com/korkinlab/DISPOT, DockerHub: https://hub.docker.com/r/korkinlab/dispot.

SARS-CoV-2 structural genomics and interactomics - *homology modeling* of COVID-19 viral protein and analysis of their potential interactors. I was responsible for the homology modeling of intra- and host-viral protein complexes. http://korkinlab.org/wuhan

Aug 2015 May 2019	**Worcester Polytechnic Institute (WPI)** *Teaching Assistant*

I was teaching undergraduate students for the 4 years as CS department TA. My main responsibilities included **Teaching Assistantship** and **Machine Learning, Data Analysis, Statistics**.

DREAM Challenge 2016 - Identification of disease modules in multiomics molecular networks based solely on **network topology** and validated on GWAS data.

Hodgkins Lymphoma diagnosis using cross-lab data sets - Microarray study on the diagnosis of specific cancer type. Main challenge I encountered was a **batch effect** for the data obtained from different laboratories. It was solved by combination of applying existing normalization techniques and designing feature selection algorithm for selection of the most stable set of features. I constructed machine learning models for the heterogeneous dataset which included information on gene fusion events and functional annotation of genes for 4 different cancer types.

FuEL-Can - detecting **gene fusion** cancer drivers using tissue-specific learning classifiers.

Jun 2018 Aug 2018	**Tufts Medical Center CTSI** *Research Intern*

During my work in TMC **Clinical and Translational Science Institute** (CTSI) I foucsed on **pathway analysis** and **gene regulatory network reconstruction** algorithms in the context of **multi-omics** data integration. I reviewed existing methods and approaches, built Docker environment which included Hadoop, pySpark, TensorFlow, Anaconda and tested several computational methods.

Jun 2013 May 2015	**Samsung R&D Center, Kyiv** *Research Engineer*

Designing machine learning models for the wearable devices. My responsibilities consisted of **Machine Learning, Image segmentation, Statistics, Time Series Analysis** and **Signal Processing** Garamond

Gesture recognition project development of gesture recognition algorithms that use multiple smartphone sensors. Merging data from camera, proximity sensor, touch screen. Responsibilities included building Hidden Markov Models for multiple parameters and general software engineering problems in our gesture recognition engine. Main problem for my part of work was combining information from multiple sensors in one model. Another technical problem was bringing fast and optimized C++ core of our gesture recognition module to the android devices.

Medical imaging project - automatic diagnostics of the spine related disease based on the MRI medical images. I improved segmentation algorithm responsible for the spine detection and anomaly detection in the bones structure. Main challenge during my work was linking of the information about spine disc and facet joints. It was solved by using graph-based approach that integrated anomaly detection step into linking segmented parts of the image together.

Activity recognition project - consisted of the two main parts: sports activities recognition and vehicle recognition based on smartphone and smart watch sensors. Sports activity recognition included recognition of the predefined sets of exercises. Vehicle recognition part included distinguishing between riding a car, a bike and train. I performed signal processing and pattern recognition. One of the biggest challenges was necessity to reduce frequency of data retrieval from sensors in order to make a system energy efficient.

May 2011 Apr 2012	**Space Research Institute of the National Academy of Sciences of Ukraine and National Space Agency of Ukraine** *Software Engineer*

Development of geo-informational web-service adding maps to the web-service, coordinates systems. Leaf area index computation. Regression analysis of crop maps.

Education

2015-*Now*	**PhD in Computer Science** *Worcester Polytechnic Institute* GPA: 3.91/4.0
2013-2015	**Master Degree in Computer Science** *National Technical University of Ukraine Kyiv Polytechnic Institute, Institute of Physics and Technology*
2012-2014	**Data Analysis School** *Educational program featured by Yandex*

2009-2013	BACHELOR DEGREE IN COMPUTER SCIENCE
	National Technical University of Ukraine Kyiv Polytechnic Institute, Institute of Physics and Technology
	GPA: 3.86/4.0

PUBLICATIONS

O. Narykov, N. T. Johnson, D. Korkin "Predicting protein interaction network perturbation by alternative splicing with semi-supervised learning" *In revision in Cell Reports*

O. Narykov, S. Srinivasan, D. Korkin "Computational protein modeling and the next viral pandemic" *Bioinformatics*, 18, 444445 (2019) https://doi.org/10.1038/s41592-021-01144-0

S. Srinivasan, H. Cui, Z. Gao, M. Liu, S. Lu, W. Mkandawire, **O. Narykov**, M. Sun, D. Korkin "Structural Genomics of SARS-CoV-2 Indicates Evolutionary Conserved Functional Regions of Viral Proteins." *Viruses*, 12, 360. (2020) https://doi.org/10.3390/v12040360

O. Narykov, D. Bogatov, D. Korkin "DISPOT: A simple knowledge-based protein domain interaction statistical potential." *Bioinformatics*, 35, 53745378 (2019) https://doi.org/10.1093/bioinformatics/btz587

S. Choobdar, M.E. Ashen [et al, including **O. Narykov**] "Assessment of network module identification across complex diseases." *Nature Methods* 16, 843852 (2019) https://doi.org/10.1038/s41592-019-0509-5

Z. Gao, S. Lu, **O. Narykov**, S. Srinivasan, D. Korkin, Chapter 7 in Virus Bioinformatics (1st ed.) D. Frishman, M. Marz, (Eds.). (2021) *Chapman and Hall/CRC* https://doi.org/10.1201/9781003097679

PROJECTS

2014	GOODSINFO
	Team project in which I was in leading position. In its scope was created distributed application that allows user to get information about goods using their bar-codes. Has distributed architecture. GitHub link.

2013	KEYSTROKE AUTHENTICATION
	The goal of this project was to implement biometric authentication via keystroke dynamics and study its limitations. Result was a Windows application that had 89% user recognition accuracy (tested on 5 people). A method for coping with temporal variability during training stage was introduced. GitHub link.

INTERESTS AND ACTIVITIES

President for the Eastern-European cultural club at WPI during 2017-2019.