

# Andrew Kurochkin



Data Scientist

- [andrewkurochkin.com](http://andrewkurochkin.com)
- [me@andrewkurochkin.com](mailto:me@andrewkurochkin.com)
- Lviv, Ukraine

---

## KEY SKILLS

Computational Social Science  
Behavior analysis Research  
Negotiation  
Communication  
Presentation  
Data analytics  
Programming  
Problem solving  
Database design  
Data visualization  
Team organizing

---

## TECHNICAL SKILLS

Python & R & PHP  
JS, AngularJS, CSS  
UNIX / Linux  
BI (Looker, etc)  
GCP, AWS, S3  
GIT, CI/CD  
SQL, BQ, ETL  
MTurk, A/B testing

## Education

**Data Science**, Master degree 2018 - 2020  
Ukrainian Catholic University  
Thesis: Meme Generation for Social Media Audience Engagement

**Software engineer**, Bachelor degree 2013 - 2017  
National University "Lviv Polytechnics"  
Thesis: Hiring automation for IT company using ML

## Experience

**Data scientist**, R&D, Palo Alto Insight 2018

- Researched generative adversarial networks
- Created an automated pipeline for effective anime faces recognition, capturing and processing
- Developed NNs to produce new anime faces.

**Data engineer**, Digital Media Agency, Lohika 2017 - pres

- Built BI infrastructure for an advertising data
- Implemented BI semantic layer
- Integrated data warehouse architecture with big data technologies
- Developed high-performance data processing procedures
- Reviewed, analyzed, and ensured the quality of data loaded into the database systems from hundreds of sources.

**Principal developer**, Headliner 2015 - 2017

- Cooperated with client's company executives to set up priorities and establish requirements
- Collected and analyzed pre-compiled data to generate reports
- Designed and implemented highly efficient data collection procedures to boost overall performance.

**Full-stack web developer**, Freelance 2011 - 2015

- Managed business processes, created proposals and project requirements
- Designed and implemented various functionalities to tens of projects
- Cooperated efficiently with team of freelancers.

## Projects

### **Meme Generation for Social Media Audience Engagement**

A memes dataset of ~650K meme instances was collected. Applied SOTA technique towards computational creativity problems. Used GPT-2 in the combination of the other methods for meme generation. Built pipeline to compare machine-generated memes with human-created. Justified that MTurk workers can be involved in the approximate estimating of users' behavior in a social network, to measure engagement. Generated memes cause the same engagement as human memes, which didn't collect engagement in the social network (historically).

**Tools:** python, GPT-2, GCP, MTruk, tensorflow, Azure, S3.

### **Anime face generation by GAN, Silicon Valley**

R&D project for the San Francisco based company.

Solution for collecting anime faces dataset from videos was created. Trained generative adversarial networks (GAN) to generate new faces for cartoon and anime characters.

**Tools:** python, tensorflow, AWS, S3.

### **Data engineering for Digital Media Agency, Worldwide**

The agency manages over \$600MM in media and deploys campaigns across world markets. Our solution delivers tons of data from advertising platforms to customers. As a part of the team, I'm responsible for BI infrastructure for the company, data quality management (DQM), creating reports and dashboards for the system and business processes monitoring, integration of different data warehouses in one architectural solution.

**Tools:** Looker, python, hive, presto, PostgreSQL, SQL.

### **Panacea. CRM for city Ambulance**

Software for the Lviv Central Ambulance Station. It allows to gather statistics, search for lost people, generate reports, and create dashboards for the Ministry of Health of Ukraine. The solution has been in effect at the city Ambulance Station from 2015 till 2018.

**Tools:** PHP, PostgreSQL, Ubuntu, JS, AngularJS.

### **Recommendation system for recruiters**

The software evaluates how new candidates suit unfilled positions based on social network data and feedback from interviewers. The developed machine learning model showed ~70% accuracy.

**Tools:** python, pandas, scikit-learn, seaborn, casperJS, phantomJS, PostgreSQL, PHP, JS, AngularJS.

### **Optimizing ambulance service with AI instruments**

Research was based on a dataset of 2 years of ambulance exits. We investigated data and created a model for prediction of emergency call result from primary data.

**Tools:** R, python, ggplot2, pandas, scikit-learn, numpy.

## Other activities

1. 2020, Poster at the **Stanford 2020 Conference on Computational Sociology**  
[iriss.stanford.edu/css/conferences/2020-conference-computational-sociology/agenda](http://iriss.stanford.edu/css/conferences/2020-conference-computational-sociology/agenda)
2. 2020, Talk at **UCU webinar** – “Meme Generation for Social Media Audience Engagement [Part #1]”
3. 2019, **Masters Symposium** on Advances in Data Mining, Machine Learning, and Computer Vision (MS-AMLV 2019)  
[apps.ucu.edu.ua/en/masters-symposium-amlv-2019](http://apps.ucu.edu.ua/en/masters-symposium-amlv-2019)
4. 2018-2019, **Program committee member** in the “Morning at Lohika”, monthly local tech talks.  
Talks - [youtube.com/channel/UCRGTTEhkjGTfpypmHFtT\\_Ww](https://youtube.com/channel/UCRGTTEhkjGTfpypmHFtT_Ww)  
FB community - [facebook.com/groups/morning.lohika](https://facebook.com/groups/morning.lohika)