Anton Wolkov

Software Engineer specializing in advanced architectures, data infrastructure for large scale, reliable development and deployment of AI and data lake management tasks (**DevOps**, **MLOps**)

🔖 Experience Summary

Docker 8 years	**Kubernetes** 5 years	**AWS** 9 years	**Jenkins** 5 years	**Linux** 12 years
DevOps 5 years	**MLOps** 7 years	**Java** 6 years	**Python** 8 years	**Go** 2 years

📍 **Available in EU Time**

Contact me at

Email: anton@l6.ee

Schedule a call: https://l6.ee

linkedin.com/in/antonwolkov

⊞ Employment

- ### Lead DevOps, MLOps, Software Engineer
  Proofpoint — 2020 - PRESENT

  I'm a member of a small team consisting of backend, full stack developers and data scientists. I am responsible for all production and deployment infrastructure of our web based platform (DevOps), our data platform (MLOps) and integrations with other teams.

  Notable achievements I accomplished at this position:
  - Designed a microservice architecture using python, docker, helm 3, AWS operator
  - Built jenkins based CI/CD kubernetes deployment pipelines for frontend and backend
  - Merged prometheus and grafana dashboards from multiple EKS clusters using thanos
  - Created ETLs for data exploration and integration tests on anonymized data
  - Automated pagerduty incident management with playbooks and CI/CD deployments
  - Integrated multiple teams' data into an NLP oriented batch data pipeline

  Technologies I used extensively at this position - grouped by field of responsibilities:
  MLOps ETLs: athena, glu. Scheduling: airflow, kubernetes. KPIs: mlflow, prometheus. Infra-as-code: aws operators, helm 3, terraform. Dashboards: grafana. NLP: tensorflow.
  DevOps CI/CD: jenkins, docker, helm, minio, elasticsearch. Environment: docker, pypi, chartmuseum, os agnostic deployment and management tools. Cloud: aws, gcp. SSO Integration: onelogin, oauth2-proxy, cognito. CDN: cloudfront. DNS: route53.

- ### MLOps, Principal Software Engineer
  Oracle **Data Cloud** — 2018 - 2020

  I managed the data lakes and infrastructure-as-code pipelines for multiple teams of data scientists and data analysts. I've also worked on real time audience tagging microservices.

  Notable achievements I accomplished at this position:
  - Created a self service process for data scientists to productize their data POCs
  - Integrated metrics collection and reporting into all parts of the pipeline and git PRs
  - Migrated AWS EMR workloads into spark-on-kubernetes running on OCI infrastructure
  - Developed lightweight microservices to handle real time pixel requests with strict SLAs
  - Extended python s3 library to support OCI and optimized for high latency operations

  Technologies I used extensively at this position - grouped by teams I was a part of:
  Batch Data Pipeline ETLs: luigi, spark, presto, pig. Spark UDFs: scala, java, python pandas. Scheduling: jenkins, luigi, airflow. Monitoring: luigi, airflow, mlflow, kibana, s3. Environment Orchestration: emr, qubole, kubernetes. Data Lake: luigi, hive metastore.
  Real Time Pixel Software: go, nginx ingress. DBs: aerospike. Queue: kafka CI/CD: jenkins. Monitoring: thousandeyes, grafana, prometheus. Cloud: aws, oci. CDN: akamai.

Software Engineer for hire with over **12 years** of experience acquired in small startups, large corporations and fortune 500 companies including industry leaders in A.I.

B.Sc. in Computer Science Graduated from a top ranked university

I am a Data platform architect with experience in infrastructure (**DevOps**) and data science workflows (**MLOps**)

At Proofpoint, I was able to provide multiple teams of security researchers and product engineers with tools to discover new vulnerabilities and protect our customers from previously unknown types of attacks

At Oracle, I supported 3 teams of data scientists and 3 teams of data engineers across multiple continents, providing a unified data lake with common self-service tools for exploration, ETLs, productization and code lifecycle (tests, KPIs)

● MLOps, Software Engineer II

Amazon Lab126 — 2017 - 2018

I worked with a dozen teams of data scientists across multiple sites. My team consisted of 3 Israeli and 8 American developers. I built tools for secure data exploration, experiment management and POCs in computer vision, NLP and discrete recommendations.

Notable achievements I accomplished at this position:
- Onboarded a new realtime database to sync annotators' inputs (javascript), ETLs (java) and report generators / data exploration tools for A.I. experiments / POCs (python)
- Repurposed an internal voice annotation platform to be used for computer vision
- Automated status reporting from an experiment management platform to confluence
- Created a jira ticket templating system to simplify operational process status tracking

Technologies I used extensively at this position - grouped by field of responsibilities:
MLOps ETLs: spark, python, java. DBs: rethinkdb, postgresql, s3. Alerting: pong, chime, email. Notebooks/Reports: zeppelin, jupyter. Metrics: spark counters, rethinkdb.
Software Engineering Integrations: sso, software vpn, security scanners. Ticketing: jira, sim.

At Amazon, I got the opportunity to bring multiple products from a data science POC to a fully fledged hardware and software solution available to Amazon shoppers and home automation users

My modified annotations collector for images was the basis of a secure, hybrid pipeline based on neural networks and human annotators

● MLOps, Software Engineer II

Microsoft — 2014 - 2017

I worked on a cloud app security broker product, integrating major cloud vendors' APIs and providing the data science team with anomaly detection tools and infrastructure.

Notable achievements I accomplished at this position:
- Integrated users and files APIs from office 365, google, servicenow, salesforce, okta…
- Created a stress test CI/CD for data science POCs from anonymised production data
- Developed a distributed rate limiter logic for integrations' SDKs running async.
- Created a wrapper to remotely sort and filter deleted files on generic files APIs
- Responsible for integrating all services to a customer facing availability status page
- Optimized mongodb and elasticsearch to scale for all of Microsoft's customers' data

Technologies I used extensively at this position - grouped by team focus:
MLOps ETLs: python, flink. Monitoring: splunk, kibana. Scheduler: jenkins, azure.
Development Lang: java, python. DBs: elasticsearch, mongodb, apache ignite, cassandra, s3.

During my last year I joined 2 other data scientists to create a new user behavior anomaly detection engine and a soft compliance enforcement sold under the Office 365 E5 license

My hackathon project paperknife.me downloads attachments linked in emails (gmail, outlook) and attaches them to the original email thread

● IT support (student)

Technion (University) — 2012 - 2014

I provided IT support to students, mainly on linux desktops at the CS faculty.

Notable achievements I accomplished at this position:
- Created a GUI printer manager to auto-assign printers based on location (IP subnets)
- Crafted a large software bundle for developers to support all CS courses' requirements

Technologies I used extensively at this position:
Linux OS: ubuntu, Software: apt packages, qt creator. Tickets: jira.

At Proofpoint, Oracle, Amazon, Microsoft and the Technion: I've created scripts in python and ruby to orchestrate laptops and servers

I was also responsible for onboarding new data scientists and developers

● Lead Backend Software Engineer

S'No Control — 2008 - 2009

I created a fully custom realtime web based chat backend on a MySQL cluster sending messages using http comet (supporting IE6) handling a reasonably large user base.

Notable achievements I accomplished at this position:
- Creating a new chat backend from scratch supporting using very limited hardware

Technologies I used extensively at this position:
Development Lang: php. DBs: mysql.

Before university, I got an opportunity to create a chat platform for a popular Israeli website, which I've created the from scratch, using bleeding edge backend technologies

- ## Cyber Security Expert

BugSec — 2007 - 2008

I provided security consultations, penetration testing and code hardening for over 50 clients, on-prem, including startups and industry leaders in banking and health care.

Notable achievements I accomplished at this position:
- Developed a knowledge base product for on-prem clients, allowing their developers to quickly lookup code hardening guidelines for their specific tech-stack and situation

Technologies I used extensively at this position:
On-Prem KB Product OS: ubuntu. Lang: php. DBs: mysql. Web Server: lighttpd.

At BugSec I did a lot of short-medium term projects with different clients. My reports always included practical remediation suggestions tailored and adapted to the technologies and programming languages used by the client

🎓 Education

- ## B.Sc. in Computer Science

Technion Israel Institute of Technology — Graduated 2014

Notable grades: OS: 90, Intro to Sys. Programming: 93, Project in Parallel and Distr. Comp.: 100

At university, my seminar was about automatic annotation inference, static validation of concurrency runtime safety contacts for Java structures and deadlock detection

📑 Portfolio

- ## kix.co.il chat platform

I've created a slack-like, privacy oriented, fully self-hosted chat and community suite of tools, using cloudflare, digitalocean's kubernetes, mattermost, and lots of custom addons and bots.

You can check out all of the source code, including full install helm charts, bots, other tools linked via SSO, backend and frontend mods at: github.com/phntom

My chat enjoys an active core of loyal users. The new platform is up since 2020 superseding the IRC network which is up for over 20 years

- ## Condi android automation if-this-then-that app

My university project was to create a robust, free and intuitive automation app for android similar to tasker, with natural language descriptions and sharing capabilities. The app was downloaded over 100,000 times and was translated to 8 languages by volunteers.

My android project is currently delisted due to a lack of updates, but it's still used by a small group of people, since the app is fully self contained on the device.

📁 Skills

- Data lake management
- CI/CD automation
- Migrating into Kubernetes
- Helm 3 templating
- MLOps workflow incorporation
- Security Auditing