

## CURRICULUM VITAE

### Mateusz Konopelski

ml.konopelski@gmail.com

linkedin.com/in/mateusz-konopelski/

+48 576 716 000

### Computer Vision projects: [github.com/mlkonopelski](https://github.com/mlkonopelski)

Goal: learn CV by building models from scratch in PyTorch or fine-tuning sota architectures to solve problems like: cancer detection on ct scans, wood knot counting on production lines, construction blueprint analyser, plant identification, etc.

## Experience

### COIG SA ML Engineer (07.2022 - 12.2022: 8 mo)

Lead programmer and architect of AIDA (AI Document Assistant) platform with following functionalities:

- o Text Categorization
- o Text Summarization
- o Named-entity recognition

delivered as MVP and tested with first customer (City Hall of Lublin).

Technology:

- o Text Cleaning: spaCy + own algorithms
- o ML (annotations: LabelStudio):
  - Classification tasks with optimized xgboost
  - Classification tasks with tuned HerBERT transformer
  - Named Entity Recognition with LayoutLM
  - Summarization with heuristics or T5 transformer
- o MLOps: MLFlow
- o DB: Postgres & Oracle & MongoDB
- o API: FastAPI + Gunicorn
- o DevOps: GitLab CI/CD with auto build and deployment
- o Containerization: Docker & Portainer
- o Metrics & Storage Viz: Lok & Promtail & Prometheus & Grafana Dashboards

### BasicLab Dermocosmetics • Data Engineer (02.2022 - 12.2022) weekend contract

- Building BI platform on AWS Cloud:
  - o SQL DB: Postgresql on AWS RDS
  - o noSQL on AWS DynamoDB
  - o ETL + RPA (Python / Docker) on AWS ECS (scheduled tasks)
  - o Logging on AWS WatchCloud
  - o Dashboards on Tableau

### T-Mobile Polska Senior Data Scientist (04.2020 - 06.2022: 2 yrs 2 mo) Remote

- Geopositioning analytics platform:
  - o Big Data (300M rows daily) processing with Hive sql, Hivemall, pySpark
  - o Time Series modelling with Statsmodels (sarimax), Tensorflow (LSTM – own architecture), pyTorch (ESRNN – implemented from github); Scikit-learn, Scikit-optimize
  - o Database Management with PostgreSQL + sqlalchemy
  - o ETL with Airflow
- Age & Gender Classification (stopped at POC)

- o Data processing with Hive sql
- o Data visualization with Tableau
- o Classification and Regression with boosting trees (XGBoost, LightGBM, CatBoost) optimized (Scikit-optimize, Hyperopt).
- CluE - Unsupervised clustering of service desk incidents:
  - o Text-mining with Spacy and clustering algorithms with DBScan, K-means
  - o Backend with: FastAPI, Joblib
  - o Frontend with: Dash
  - o Deployment: Docker, Kubernetes
- Service Desk Robots - solving SD Incidents automatically:
  - o Data processing: Pandas, Numpy etc
  - o Database Management with Oracle + sqlalchemy
  - o Deployment: Docker, Kubernetes
  - o CI/CD: Gitlab, Jenkins

**Innogy Polska** Data Scientist (08.2018 – 02.2020: 1 yr 6 mos) Warsaw

- Churn modelling:
  - o ETL: SAS Data Integration Studio
  - o Data processing: Python Scikit-learn on AWS
  - o Classification: XGBoost
  - o Deployment and Experiments: AWS SageMaker
- Workforce prediction – scheduling Call Center agents:
  - o Data processing: SAS 4GL
  - o Time Series prediction: SAS Miner (SARIMA & DNN)
  - o Data Exploration: SAS Viya

**Hilton Worldwide HQ** Data Analyst (09.2014 – 03.2017: 2 yrs 7 mos) London

- Revenue Management System for Group Reservations. Deployed to 300 hotels:
  - o ETL: Microsoft Data Integrated Studio
  - o Application: Microsoft Excel + VBA
  - o Databases: Microsoft SQL Server
  - o Reporting: Tableau

**IKEA Retail** Business Analyst part-time (1 yr) Remote

- Sales tracking tools and methods in cooperation with Customer Service and IT.

**IKEA Retail** Junior Treasury Specialist (06.2011 – 08.2012: 1 yr 3 mos) Warsaw

- Daily liquidity management activities for IKEA group. In cooperation with my manager: working on capital financing, investment financing, managing derivatives

## Education

**Warsaw School of Economics** (10.2018 – 06.2019) Analytics Academy with SAS, R and Python. Thesis: “Artificial Neural Networks with predefined convolutional layers for image recognition. Implementation in Numpy.”

**Kozminsky University, University of Bradford** (09.2012 – 06.2014) IB

Master Thesis: Combined vs. Individual Forecasting Methods in Revenue Management

**University of Victoria** (09.2013 – 12.2013) Student Exchange

**Warsaw School of Economics** (10.2008 – 06.2011) Quantitative Methods in Economy and Information Systems. Specialization: Decision Analysis Methods, Econometrics.

Thesis: Credit risk models for households.