Teodor Yankov

📞 +447470030723 | ✉ teodor.yankov@outlook.com | 🔗 My LinkedIn | 🐙 My Portfolio

EDUCATION

University of Oxford   Oxford, UK
Master of Science, Social Data Science   *Oct. 2023 – Aug 2024*

- Grade (current): **Distinction**
- Courses in **Python**, analytical statistics, machine learning, network science, data analytics at scale, research methods.
- Thesis constructing **novel algorithm** for latent narrative extraction from unstructured text data, intersecting **SLMs, LLMs**, **NLP** and **network science**, evaluated on a novel benchmark I designed. Involved LLM fine-tuning, quantisation and performance evaluation.

University College London (UCL)   London, UK
Bachelor of Science, Political Science and International Relations (with Data Science)   *Sept. 2020 – May 2023*

- Grade: **First** Class Honours, **Dean's List**
- Courses in **R**, causal/statistical inference, measurement, quant. data analysis, migration, democracy, public policy, ethics.
- Dissertation exploring the predictive capacity of the **gravity model** in forecasting transnational refugee flows, as well as the effect of model ethno-cultural sensitivity on prediction accuracy.

EXPERIENCE

Associate Data Scientist   June 2022 – Present
World Data Lab   *Remote*

- **Mining**, **analysis**, **validation**, **modelling** & **visualisation**, using structured and unstructured data, for collaborative projects measuring and forecasting migration, greenhouse gas emissions, demography and consumer class dynamics, digital inequality and national post-crisis poverty.
- Data consulting public-sector clients (e.g. IOM, GIZ, Internet Society) and private-sector clients (e.g. L'Oreal, Colgate, Mastercard), building decks, datasets and constructing high-value insights.
- Key **insight mining** for data-driven webinars on consumer spending with **200+** global attendees
- Led COP27 **launch campaign**, insight-mining and video production for the World Emissions Clock.
- **Skills**: R, Python, APIs and pipeline automation, machine learning, data consulting, data tool building, QA

Co-founder   Jan. 2022 – Present
Youbee   *London, UK*

- One of two founders of an app-based personal growth and self-improvement business combining innovative technology and self-help tools to improve the lives of students and workers. My main contribution is in **data strategy**, **community outreach**, **website maintenance** and the creative process.
- **Skills**: app development (front-end, HTML, CSS), website management (WordPress), Google Analytics, teamwork, problem-solving, UX design, graphic design, media

Research Assistant   May. 2021 – Sept 2023
UCL Department of Political Science   *London, UK*

- Assisting Dr Samer Anabtawi, researching LGBTQ+ activist networks in the MENA region.
- Assisting Dr. Richard McMahon, exploring China-EU mutual representation.
- (As UCL Connect.Ed Mentee) assisting PhD project researching the Moria refugee camp and EU refugee policy response.
- **Skills**: multimodal (audio, text) data processing, reference automation, primary data acquisition, data visualisation, policy research, literature review, writing, QA

Summer Intern - Data Research   Jul 2021 – Sept 2021
Institute for Market Economics   *Sofia, Bulgaria*

- Synthesised reform plans submitted to the European Commission, helped **build database** of government & tax data for the assembly of regional profiles, used statistical software to filter and analyse **large-N EU funding data**.
- **Skills**: econometrics, data mining/analysis, databases, research, teamwork, communication, report writing

TECHNICAL SKILLS

Languages: Python, R, HTML, CSS
Methods: Machine Learning (multivar. linear/logit, regularised, tree-based and count regressions, neural networks), statistical analysis (hypothesis testing), NLP (NER, topic modelling), SLMs/LLMs (GPT-3.5-turbo, GPT-4, Llama 2/3, Mistral, BERT, BERTopic, SpacY), fine-tuning LLMs, quantised LLMs, network analysis (community detection, temporal network analysis, multiplex, network formation models, graph machine learning)
Tools: Git, HuggingFace, Docker, VS Code, R Studio, Overleaf, Zotero, Figma, Notion, WordPress, Microsoft Office, Google Suite
Libraries: pandas, NumPy, Matplotlib, scikit-learn, TensorFlow, PyTorch, networkX, nltk, langchain, tidyverse, data.table