

# MYKOLA TROKHYMOVYCH

Barcelona, Spain | [trokhymovych.com](https://trokhymovych.com) | [www.linkedin.com/in/trokhymovych](https://www.linkedin.com/in/trokhymovych)

## SUMMARY

---

- Research experience in AI, NLP, data mining, and fact-checking. Three top-tier conference publications.
- 6+ years of industry experience in a mix of data science and engineering roles
- Experience as a university teacher, thesis mentor, and project leader

## EDUCATION

---

- **Universitat Pompeu Fabra (UPF), PhD student, Information Technology** Nov 2022 - Now
  - Working on Knowledge Integrity, focusing on identifying and mitigating information manipulation and vandalism in digital ecosystems and AI-generated content.
  - Teaching: Information Retrieval and Web Analytics, Data Mining, Computer Organization.
- **Ukrainian Catholic University, MSc. Data Science** Sep 2019 - Jun 2021
  - Average grade 96 (out of 100) points, diploma with magna cum laude honors
  - Master thesis: "Natural Language Inference for Fact-checking in Wikipedia."
  - Teacher assistant in Linear Algebra, Corporate Finance.
- **NTUU "Kyiv Polytechnic Institute," BSc, Institute for Applied Systems Analysis** Sep 2015 - Jun 2019
  - Bachelor thesis: "Geodata analysis methods for venues establishment recommendation in Kyiv."

## SELECTED PUBLICATIONS

---

- Author: Trokhymovych M., Sen I., et al. **An Open Multilingual System for Scoring Readability of Wikipedia.** *ACL'24 Main Track*, 6296–6311. [Publication link](#)  
*Summary:* Present a new multilingual dataset and system to score Wikipedia article readability, along with the first systematic overview of Wikipedia readability beyond English.
- Author: Trokhymovych M., Aslam M., et al. **Fair Multilingual Vandalism Detection System for Wikipedia.** *KDD'23 Applied Track*, 4981–4990. [Publication link](#)  
*Summary:* Introduce a new generation of systems designed to help the Wikipedia community deal with vandalism on the platform, improving performance, language coverage, and fairness.
- Author: Trokhymovych M. and Saez-Trumper D. **WikiCheck: An end-to-end open-source Automatic Fact-Checking API based on Wikipedia.** *CIKM'21 Applied Track*, 4155–4164. [Publication link](#)  
*Summary:* Present a new fact-checking system based on the Wikipedia knowledge base. It is comparable to SOTA solutions in terms of accuracy and can be used on low-memory CPU instances

## EXPERIENCE

---

- **Wikimedia Foundation, Research team, Remote** – Research contractor May 2021 - Jun 2021, Aug 2022 - Now
  - Implemented a new generation of the multilingual vandalism detection system for Wikipedia ([link](#))
  - Created a prototype for Automatic fact-checking based on Wikipedia, using the NLI model ([link](#))
  - Built a new model for multilingual text readability evaluation ([link](#))
- **Surprise.com, Data Science team, Kyiv** – Data Scientist/ML engineer, Jan 2022 - Nov 2022
  - Designed pipeline for custom knowledge-aware content generation
- **Joble, Search team, Kyiv** – Data Scientist/ML engineer, Sep 2020 - Jan 2022
  - Added advanced semantic textual similarity features for high-load search inference
  - Implemented cold-start recommender system, adding personalization for search
- **Ciklum, Big Data & Analytics, Kyiv** – Junior Data Scientist, Jan 2018 - Sep 2020
  - Developed an NLP-based search system automating 85% of client's procurement processes

## EXTRACURRICULAR EXPERIENCE

---

- **Hackathons and competitions**
  - **Foursquare - Location Matching Kaggle (Geospatial, NLP)**, 37 out of 1079 teams
  - **SoftServe Intelligent Search Hackathon 2021 (Search)**, 2nd out of 28 teams
  - **Deloitte Data Science Hackathon 2020 (NLP), Team leader**, 2nd out of 37 teams
- **Languages:** Ukrainian (native), Russian (fluent), English (advanced), Spanish (elementary)