LAZAR KOVACEVIC   lazar.kovacevic@yahoo.com | linkedin.com/in/lazark/

DATA SCIENTIST, SOFTWARE DEVELOPER, ML / NLP, SEO

Lead role in full life-cycle of ML web apps, *ideation to deployment*; *100M+* aggregate views.

Worked solo & in teams in a variaty of domains: news, sci, med, food, gov, law, e-comm, etc.

Attended & sponsored conferences & workshops on ecology, *collective IQ & collaboration.*[1]

- **Programming:** Python · Java · PHP · Mathematica · Multi-Agent Systems

- **DB / Storage:** MySQL · Postgres · MongoDB · SOLR · Elasticsearch · Neo4j / Faiss · Gluster · Redis

- **Web Development:** LAMP · DevOps · Docker Swarm · Flask · API · SaaS · Bot Prevention · AWS

- **Data Science:** NumPy · Pandas · sklearn · Matplotlib · D3.js · Jupyter · LightGBM · RF · SVM · PCA
  · LDA · KMeans · ANN · PyTorch · SpaCy · AllenNLP · Gensim · SBERT · NLTK · FastText
  · Splink · WordNet · HuggingFace · IBM Watson · OpenAI · LLM · Langchain · Data Collection

- **Algorithm Design:** Recommender Systems · Novelty Search · Data Modeling · Optimization

- **Information Retrieval:** Scraping · ETL · NLP · NEL · ML · AI · LTR · KG · RAG · LSH · SQL · Record Linkage

POSITION	COMPANY	FROM – TO
Data Scientist, Affiliate Faculty	**George Mason University**	**2022 – Current**

- Developed a platform easily extensible to new sources and **ad hoc faceted searches**
- Deployed end-to-end data science workflow for *Detecting National Academies of Sciences, Engineering, and Medicine [scientific] evidence citation in US Congress* [2]

Software Developer & SEO, Consultant	**Inverudio**	**2005 – Current**

- Developed *video transcript **summarization*** for *WelcomeTerra*; acquired by Pictory.ai
- Developed a legal contract **red flag issue detector** for *ContractReader.us* (team)
- Built nontrivial content & connections **recommender system** for *Macrovo.com* (team)
- ***Opinion mining*,** condensed 300k Net Neutrality comments into 300 arguments
- ***DevOps*** on many projects including Docker Swarm based **microservices** behind HAProxy
- ***Location prediction*** w/ ML & NLP in unstructured non-lethal violence reports (team)
- LSTM based ***intrusion detection system (IDS)*** through network packets analysis (team)
- Used controlled ***medical*** and other ***ontologies*** for ***data augmentation*** and retrieval
- Wrote an algorithm for finding optimal layover paths for ***flight route optimization***
- ***Fuzzy matching*** & ***entity linking*** for product / company mention disambiguation
- Analyzed *Machine learning **trends*** in SSRN finance & economy networks
- Built GUI based ***multi-agent simulation*** of infectious diseases, i.e. COVID-19
- Extracted ***structured data from unstructured texts*** in numerous domains

100+ Projects (team work & solo)
- LLM integration, **prompt engineering**, auto-labeling, **sentiment analysis**, medical KG, ML test battery, topic modeling, **clustering**, job-cv match, A/B tests, ETL,...

50+ SEO Projects (solo)
- *On-page SEO, **keyword research, (tech) audits**, competition & niche analysis for:*
    UnifiedPatents, Recruit, EMusic, NY-Engineers, WildHorseTour, Altijdzon, Links...

Stealth startup (2+ years; solo)
- *1ms **sentence alignment** for in-text paraphrase citation detection **in nonparallel texts***
- *Mapped evolution of millions of recurring ideas* from 500,000 multi-lingual books
- Created a Blekko-inspired search engine: selective indexing, slashtag filters, SEO queries
- LSM+BTree **bounded index** minimizes space / read amplification on updates / reads
- Invented data structures to *estimate multidimensional maxima* for concept/url selection

EZGlot.com (2 years; team work)
- Designed algorithms for sound, orthography, and graph based *similarity measures*
- Mined and bootstraped thousands of dictionaries for **semantic similarities** [3]
- *End-to-end data science workflows*: model development to testing & implementation
- Implemented **RAG for multilingual content** based on large data-set amassed over time

SemanticJuice.com (3+ years; team work)
- Built in team ML driven *topical crawling SaaS*, a distributed **search engine generator**
- Created *Semantic flow* measure for rapid identification of top niche resources
- Practiced **pair programming** w/ a university professor during core development phase

---

Programmer, Research Associate Chicago Center for Creative Dev.   2003 – 2006

- ***Interdisciplinary research*** team member, programmer, *co-author of academic papers* [4]
- Developed GUI desktop applications for *recurrence plots & time series analysis (TSA)*
- Collaborated on RRI-based **mental health classification**
- Developed Schrodinger equation simulation applet in Java

SELECTED PUBLICATIONS [5]
- Bios data analyzer. H Sabelli, A Sugerman, L Kovacevic, L Kauffman, L Carlson-Sabelli, M. Patel, and J. Konecki. *Nonlinear dynamics, psychology, and life sciences* 9 (4), 505-538, 2005.
- Quantum bios and biotic complexity in the distribution of galaxies H Sabelli, L Kovacevic. *Complexity* 11 (4), 14-25, 2006.
- The biotic pattern of heartbeat intervals H Sabelli, J Messer, L Kovacevic, K Walthall. *International journal of cardiology* 145 (2), 303-304, 2010.
- Augmenting and Informing the Translation Process through Workflow-Enabled CALT Tools SM Rudan, E Kelbert, L Kovacevic, M Reynolds, S Rudan *Computer-Assisted Literary Translation*, 258-281

EDUCATION: Dipl. Ing. (M.Sc.E.E.) **University of Belgrade**, Yugoslavia   **1997 – 2002**