Florian J. B. Leitner

Data Scientist & Engineering Manager

C/ Juan Duque 26, 6C
E-28005 Madrid

✆ +34-XXX-XXX-XXX

❀ datacatalytics.com
❀ fnl.es

✉ leitner@datacatalytics.com
✉ me@fnl.es

- linkedin.com/in/florianleitner
- github.com/fnl
- stackexchange.com/users/fnl
- scholar.google.com/citations?user=EkXLlysAAAAJ

Profile

Scientific coach and leader specialized in machine learning + data and process analytics.
→ Coaching teams via the Toyota Kata to facilitate a mutual learning growth mindset
→ Teaching the principles of scientific management while leading through an agile strategy
→ Machine learning and data analytics specialist in biotech and finance (20+ years exp.)

Main career achievements: CTO of a financial news startup that successfully exited, Head of Machine Learning of an international financial data analytics entity, received a PhD in molecular biology (with distinction), published a first author paper in Nature Biotechnology, and founded a data science consultancy.

Recent Career

Head Of Machine Learning, ION Analytics, Madrid, ES & London, UK – 2021.05-Present
▸ Concurrently managing and coaching up to six different teams on disparate projects to develop a mutual learning mindset and apply the scientific method to their work.
▸ Developed the blue-print that the company uses to map its value streams and processes.
▸ Defining and promoting the company's data engineering, analysis, and governance doctrine.
▸ Tracking and orchestrating the company's LLM-related AI efforts.
▸ Coaching the product organization on ubiquitously applying augmented intelligence to automation, information extraction & retrieval, and recommendation.
▸ Teaching teams how to build supervised and semi-supervised human-in-the-loop systems to automate in-house workflows, increase structured data volumes, and reduce operational costs.
▸ Orchestrated early adaptation of LLMs to ensure long-term strategic positioning, making ION Analytics part of the exclusive set of multinationals with LLMs in production in 2023.
▸ Designed reusable natural language query interfaces to improve the UX and quality of content recommendations on the company's next-gen product line before LLMs were widely available.

Engineering Manager, ION Analytics, Remote (New York, USA & London, UK) – 2019.09-2021.04
▸ Led the development of a search interface for investor personas by extending the capabilities of Elasticsearch and improving the UX by allowing natural language-based data retrieval.
▸ Enhanced an in-house real-time content recommendation system by integrating Elasticsearch to enable the delivery of query-related context snippets ("highlights") and reduced the engine's overall memory footprint and cluster size, saving over $100k YOY in operational costs.
▸ Coached a team of analysts and engineers to scale up the reference data system by two orders of magnitude more entities.
▸ Coordinated a TechOps team that expanded our continuous integration capabilities and automated tedious tasks, improving developer experience and reducing maintenance overhead.

Chief Technology Officer, Selerity Inc., Remote (New York, USA) – 2018.05-2019.08

‣ Led a globally distributed team through a successful acquisition.

‣ Continually improved the quality of a proprietary recommendation system for public and private content with a team of data scientists by evolving the in-house NLP libraries that percolate financial events and information to multiple tier-1 banks.

‣ Overseeing the development of a credit trader chat analytics engine to predict trade partners, detect missed opportunities, and provide trade analytics that was sold to some of the biggest intra- and inter-bank Sales & Trading Platforms.

Faculty Member, IE University, Madrid, Spain – 2018.02-Present

‣ Teach core data science and machine learning concepts to MBA students while improving the class year-over-year using the students' evaluation feedback.

Senior Full-stack Data Scientist, Selerity Inc., Remote (New York, USA) – 2015.09-2018.04

‣ Evolved the in-house open information extraction and text classification systems for financial instruments, legal entities, and business concepts/topics by increasing systems performance against continuously annotated data generated from the QA process.

‣ Designed unsupervised text clustering algorithms for news and proprietary financial documents to detect correlations and trends by iteratively improving analyst reviews of cluster coherence.

‣ Domain-driven design of entities and aggregates needed for a trader chat analytics engine by collaborating with credit desk experts to establish all requirements and architect the data structures to encode the trader workflow.

‣ Tuning and maintaining Selerity's in-house message streaming platform (similar to NATS or Akka) by reducing end-to-end latency and resource use (CPU load & memory usage).

Full-stack Data Science Consultant, Data Catalytics, Madrid, Spain – 2015.09-2018.02

‣ Worked remotely as a part-time international consultant on various NLP-related projects, such as extracting structured data from CVs or patents via short, iterative deliverables defined at contract onset.

Instructor, Adv. Statistics and Data Mining Summer School, Madrid, Spain – 2014.02-2019.08

‣ Taught the NLP class component of a prize-winning summer school (5x3h lectures) to provide an in-person experience that is unavailable in online courses while improving class quality through participant feedback.

Juan de la Cierva fellow, Univ. Politechnica, Madrid, Spain – 2014.02-2015.08

‣ Researched probabilistic graphical models and developed text-mining systems in neurobiology to facilitate knowledge acquisition for neurobiology researchers and publish novel research in the field.

‣ Published a research article while teaching and mentoring PhD and master students.

‣ Invited speaker at research conferences and events.

Post-doctoral Fellow, Span. Nat. Cancer Research Centre, Madrid, Spain – 2012-2014

‣ Developed a tool to extract transcription factor-gene interactions with the experimental evidence for the interaction from scientific publications and delivered on the objective of providing systems biologists with an autogenerated mammalian transcriptome.

‣ Published various research articles and presented at conferences globally.

Education

PhD, Autonomous University of Madrid, 2012

Molecular Biology, Bioinformatics, BioNLP

Business Skills

‣ Toyota Kata coaching and facilitative leadership from a mutual learning mindset

‣ Implementing business strategy by maximizing alignment *and* decision making autonomy

‣ Building flow engineering teams that embrace agile principles

‣ Teaching and implementing Lean habits for business processes

‣ Evolving a generative company culture with a growth mindset

‣ Company policy design and implemention; quality assurance in business processes

‣ Public speaking and writing, remote-first working

Technical Skills

‣ Machine learning frameworks and libraries (TF, **Torch**, **SciKit-Learn**, Spacy, NumPy, *etc*.)

‣ Data science and statistics libraries (**Pandas**, **Seaborn**, SciPy, Plotly, statsmodels, matplotlib)

‣ Machine learning operations (MLOps) tooling (dvc/cml, eli5, Kibana, MLflow, SageMaker)

‣ DevOps & observability tooling (Ansible, Docker, Helm, Graphite, Grafana, Icinga, Jenkins)

‣ Pub/sub messaging & stream processing systems (in-house, **Spark**, Avro, Akka, Kafka, NATS)

‣ Search engines (in-house, Lucene/Solr/**ElasticSearch**, Sphinx, Vespa, Xaipan)

‣ SQL & NoSQL data-stores (Cassandra, **Databricks**, MongoDB, MySQL, **Postgres**)

‣ Web/API development (GraphQL, nginx, Node, Postman, Swagger/OpenAPI, Twisted, vert.x)

‣ Developer tooling ([C]Make, gdb, **Git**, Grunt, Maven, perf, psutil, sbt, Valgrind, Wireshark)

‣ **Ubuntu**, Debian, RedHat & Apple **MacOS** ops experience

‣ **AWS** Cloud computing (incl. SageMaker); **Azure** and some minor GCP exposure

‣ Agile development using XP, T/BDD, and DDD (paring, feedback, iterative, tidy first, etc.)

Languages

German, English, Spanish

All fluent.