Diego Rodríguez Atencia

+34 644 606 287 | dratencia@gmail.com | github.com/alektebel | alektebel.github.io | Madrid, Spain

Education

Universidad Autónoma de Madrid (UAM) — Madrid, Spain
B.Sc. Mathematics — *2018 – 2023*

- Thesis (9/10): general form of the fourth-order isotropic tensor in the Navier–Stokes constitutive relation. Coursework in probability, stochastic processes, functional analysis and numerical methods. Founded the university's AI student association (70+ members).

Universidad Internacional de Valencia — Valencia, Spain
M.Sc. Artificial Intelligence — GPA 8.67/10 — *2023 – 2024*

- Thesis: deep reinforcement learning for prosthetic-hand control (18-DoF Allegro Hand, MuJoCo); recurrent PPO with curriculum reward, converged on CPU in under 24h.

Experience

PwC Spain — Financial Risk — Madrid, Spain
Senior Associate, Data (promoted from Data Analyst, Jun 2025) — *Nov 2023 – Present*

- Own the LGD (Loss Given Default) pipeline feeding loan-loss provisions and regulatory capital at a top-4 Spanish bank, embedded on site since Apr 2024: schema and partitioning, load validation, join strategy, and reconciliation against the bank's official figures before submission. SQL, SAS and Oracle over 60M+ row datasets.
- Owned UAT sign-off on the migration of that database from SAS to Oracle. The new implementation (∼20k lines of business logic) was closed to inspection, so I re-derived it independently from the requirements and the legacy code and reconciled outputs against it to certify the migrated data could be trusted.
- Closed 45 of 50 outstanding audit findings on data quality and the default definition, some open since 2021: reconstructed intent from sparse sources across the bank, rewrote the classification rules in SQL, backfilled historical portfolios under the revised definition, and reconciled old against new to separate genuine reclassifications from migration artefacts.
- Built a FINREP/COREP reconciliation tool automating checks against stress-test report templates, directing two junior analysts. Used in the last three stress-test cycles.
- Replicated an ECB methodology paper on climate-risk PD stress testing into a working application, resolving specification questions directly with the NGFS.
- Built the transformation layer turning AI-generated audit aggregations held in DynamoDB into the entity's approved delivery template — parsing variable input, enforcing formats in three languages, and handling numeric tolerance and edge cases.
- Leading a refactor of the LGD calculation database, targeting a twofold throughput gain on the bottleneck steps by exploiting symmetries the data guarantees by construction.
- Built the client relationship directly and converted delivery quality into 8+ consecutive contract renewals with the bank, all under hard regulatory deadlines.

Álamo Consulting — Madrid, Spain
Data Analyst — regulatory reporting — *Oct 2022 – Nov 2023*

- Sole owner of SIRBE reporting to the Banco de España for an external financial entity: prepared, validated and submitted the periodic files on the regulator's fixed calendar. Extended the reporting datamart and reviewed other developers' work for conformance to regulatory requirements before release.

Projects & Open Source

Liger-Kernel (LinkedIn) — PR #1128, in review with maintainers — 2026

- Diagnosed why TiledMLP corrupts gradients under FSDP: per-tile `backward()` fires FSDP's post-backward reshard hooks mid-loop. Added a hook-based FP32 gradient accumulator (adapted from Axolotl) that sums per-shard parameter grads and writes `.grad` once, with a single backward after all tiles. Regression-tested against HuggingFace LlamaMLP at world sizes 2/4/8 on rented multi-GPU instances.

TorchRL (PyTorch) — design proposal, discussion #3538 — 2026

- Proposed an MC-PILCO implementation (model-based RL with moment-matching trajectory propagation): pen-and-paper derivation of marginal-likelihood gradients plus a NumPy reference implementation.

regllm — regulatory-validation agent (github.com/alektebel/regllm) — 2025 – 2026

- Rule-based verdict engine, retrieval over EBA documentation, QLoRA-fine-tuned 4B model for bounded tasks, LangGraph orchestration; deployed on AWS. Python, PostgreSQL/pgvector.

Systems programming, from scratch — 2025 – 2026

- SQLite clone in C (REPL, pager, B-tree); minimal container runtime (Linux namespaces, cgroups); grep; lexer and NFA/DFA work from compiler theory.

Technical Skills

Data: SQL (window functions, query planning, large-table optimisation), Oracle, SAS, DynamoDB, pandas, PySpark, ETL design and reconciliation
Languages: Python, SQL, C/C++ **Tools**: Git, Linux, AWS, PostgreSQL, PyTorch **Spoken**: Spanish, English