Aviv Orly

Senior Software Engineer · Applied Machine Learning & Scientific Computing

12+ years building software · M.Sc. in theoretical physics (dark-matter phenomenology) · LLMs, GPU inference, large-scale data analysis

Tel Aviv, Israel · avivbinyaminorly@gmail.com · Google Scholar

SUMMARY

Senior software engineer and machine-learning researcher with 12+ years of hands-on development — from full-stack systems and workflow automation in the IDF Intelligence Corps, through large-scale scientific data analysis for an international physics collaboration, to LLM/GPU-based products. Completing an M.Sc. in theoretical physics at Tel Aviv University; in parallel, led the statistical analysis behind a 2026 *Physical Review Letters* publication.

SKILLS

- **Programming languages:** Python (primary), SQL, Ruby / Ruby on Rails, JavaScript (ExtJS), C#
- **Machine learning & AI:** deep learning, LLMs, PyTorch, scikit-learn, vLLM, GPU inference, statistical modeling and analysis, retrieval/search algorithms, token-efficiency optimization
- **Data & imaging:** scientific computer vision (CCD sensor images), image masking and filtering, rare-event/low-signal detection, large-scale data pipelines, parallelized processing
- **Databases & cloud:** PostgreSQL, cloud deployment, version control and reproducibility
- **Scientific computing:** CERN ROOT, Wolfram Mathematica, Qt, numerical methods

EXPERIENCE

Lead Engineer — AI Legal-Assistant Product · 2026–Present

Building an AI legal-assistant product end to end; first deployed at a law firm and designed to expand to additional clients.

- Architected and built the full product: backend, data layer, user-facing workflows, and deployment.
- Deployed and orchestrated LLMs locally on GPU and in the cloud.
- Designed a custom retrieval/search algorithm for legal knowledge, going beyond off-the-shelf keyword search.
- Implemented token and context optimizations to cut inference cost and latency; built the PostgreSQL knowledge store and cloud deployment.

Research Software Lead & Data Analyst — Tel Aviv University / SENSEI Collaboration · 2019–Present

- Headed development of **Eureka**, a shared data-analysis and visualization platform used by the international SENSEI dark-matter collaboration; featured in the national press (*Maariv*).
- Led the statistical analysis for *SENSEI: A Search for Diurnal Modulation in Sub-GeV Dark Matter Scattering* (*Physical Review Letters*, 2026) — defined the methodology and guided the programming team that implemented it.
- Built the computer-vision-style analysis of silicon skipper-CCD sensor images: image masking, artifact and defective-pixel rejection, and isolation of single-electron signals from millions of pixel values per image.
- Ran large-scale LLM inference experiments on GPUs (vLLM) for *One Temperature to Rule Them All?* (inference-time temperature sampling; under review at NeurIPS 2026).
- Tools: Python, Qt, CERN ROOT, Wolfram Mathematica, SQL.

Full-Stack Developer — Technological Units, IDF Intelligence Corps · 2013–2017 · Captain (res.)

- Developed full-stack operational systems over four-plus years of service.
- Backend: Ruby on Rails — REST APIs, SQL, and database logic. Frontend: JavaScript (ExtJS).
- Automated workflows in Python, Qt, Ruby, VBA, and AutoIt.

SELECTED PUBLICATIONS

Full publication list and citation metrics: Google Scholar.

- SENSEI: A Search for Diurnal Modulation in Sub-GeV Dark Matter Scattering — *Phys. Rev. Lett.* **136**, 211001 (2026). *Led the statistical analysis.*
- First Direct-Detection Results on sub-GeV Dark Matter Using the SENSEI Detector at SNOLAB — *Phys. Rev. Lett.* **134**, 011804 (2025).
- Search by the SENSEI Experiment for Millicharged Particles Produced in the NuMI Beam — *Phys. Rev. Lett.* **133**, 071801 (2024).
- SENSEI: Direct-Detection Results on sub-GeV Dark Matter from a New Skipper CCD — *Phys. Rev. Lett.* **125**, 171802 (2020).
- Analytical Approximations as Close as Desired to Special Functions — *Axioms* **14**(8), 566 (2025). *Sole author.*
- One Temperature to Rule Them All? — inference-time temperature sampling in LLMs; *under review at NeurIPS 2026.*

EDUCATION

M.Sc. in Physics — Tel Aviv University · Expected 2026
Theoretical and phenomenological high-energy physics. Thesis: a novel mechanism for producing heavy dark matter via the collapse of subcritical vacuum bubbles (manuscript in preparation).

B.Sc. in Physics — Tel Aviv University · 2022 · Magna Cum Laude

"Young Researchers" Program — Weizmann Institute of Science · 2009–2012
Included a year of hands-on programming; final simulation project on hypervelocity stars emerging from binary-star separation; selected for a youth physics delegation to Seoul.

LANGUAGES

Hebrew (native) · English (fluent)