

# Roger R. Thomas, B.Sc. (Hons), M.Sc.

Document Intelligence Architect

📍 Dublin, Ireland    🌐 [rt.two4u.fun](https://rt.two4u.fun)    [in](#) LinkedIn    [GitHub](#)

## Professional Summary

Software engineer with a background in astrophysics and scientific computing. Early work using C/C++ to develop scientific models led to more than a decade of building backend services, APIs, data platforms, and high-performance systems. My experience spans Python web frameworks including FastAPI, Django/DRF, and Flask; relational and NoSQL databases including MySQL, PostgreSQL, DynamoDB, and MongoDB; and the PyData ecosystem, including Pandas, Dask, NumPy, and Numba.

Currently a Document Intelligence Architect (Principal Engineer) at Nitro, building scalable document-processing systems using machine learning and generative AI, including document indexing, semantic retrieval, and retrieval-augmented generation (RAG).

## Experience

**Principal Engineer, Nitro, Inc.** – Dublin, Ireland May 2023 – present

- Design and build scalable document-intelligence services, combining hands-on backend engineering with system design across machine learning, generative AI, document processing, indexing, semantic retrieval, and retrieval-augmented generation.
- **Technology**  
**High level:** Python, FastAPI  
**AWS:** EKS, Lambda, DynamoDB, Textract, Step Functions, Bedrock  
**Azure:** OpenAI, AI Search  
**Infrastructure and delivery:** Terraform, GitHub Actions

**Senior Lead Software Engineer, Oliver Wyman** – Dublin, Ireland Jan 2022 – May 2023

- Led platform architecture and engineering delivery, setting technical direction across API design, infrastructure modernisation, reliability, and feature development.
- Partnered with product owners, clients, and stakeholders to shape priorities, plan delivery, mentor engineers, and support the platform's long-term operation.
- Technology: Python, Django/DRF, Airflow, Argo Workflows, MongoDB, PostgreSQL, SQL Server, Docker, Kubernetes, and Azure.

**Lead Software Engineer, Oliver Wyman** – Dublin, Ireland Jan 2019 – Jan 2022

- Led the technical design of the platform's analytics architecture, spanning Python, Airflow, Django/DRF, MongoDB, and the PyData ecosystem.
- Designed and delivered a general-purpose asynchronous job-execution system for distributed data processing using Python, Dask, and Azure.

**Senior Software Developer, Oliver Wyman** – Dublin, Ireland June 2017 – Jan 2019

- Built and delivered core retail SaaS features across backend services, data processing, and computational workloads using Python, SQL Server, MongoDB, and the PyData ecosystem.

**Technical Lead / Software Developer, Creme Global** – Dublin, Ireland Jan 2014 – June 2017

- Led and modernised the Python and Django backend for a computational risk-exposure platform, including REST APIs and automated testing.

- |                                                                                                                                                                           |                     |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|
| <b>Python Developer</b> , AGOGO Amalgamated, Inc. – San Francisco, USA                                                                                                    | Feb 2013 – Dec 2013 |
| <ul style="list-style-type: none"> <li>• Developed and maintained a large-scale Python web-scraping framework.</li> </ul>                                                 |                     |
| <b>Mathematical Software Developer</b> , <a href="#">Creme Global</a> – Dublin, Ireland                                                                                   | Oct 2011 – Dec 2012 |
| <ul style="list-style-type: none"> <li>• Developed scientific models and distributed Python and AWS systems for consumer risk and chemical-migration analysis.</li> </ul> |                     |

## Selected Projects

---

### TaglyDB

A simple REST database for tagged JSON and blob storage – no schema, just tag it and find it.

- Store JSON or blobs, tag them, and query by tag combinations – no tables to design.
- REST API with per-integration API keys and signed, shareable links.
- **FE:** Next.js(TS), Tailwind CSS, daisyUI.
- **BE:** Hono(TS), Cloudflare Workers, D1, Typebox.

### jero

A high-performance, msgspec-first ASGI framework for modern Python.

- Class-based routing through Resource and Endpoint, with type hints as the API contract.
- Strict startup validation, typed request binding, typed responses, and explicit application wiring.

### yeetr

A tiny, typed, signature-driven Python CLI runner.

- Runs ordinary synchronous or asynchronous Python functions as command-line tools.
- Uses function signatures for arguments and options, with rich help and support for structured types.

### lothc

A typed HTTP client for Python, built on Rust-backed pyreqwest.

- Typed request and response decoding into pydantic models or msgspec structs, with sync and async clients sharing one API.
- Published benchmarks showing validated msgspec decoding outperforming unvalidated dict parsing.

### two4u

A mobile-first sports prediction application built for a long-running match-day game with friends.

- **FE:** SvelteKit(TS), Tailwind CSS, Skeleton.
- **BE:** Python, FastAPI, DynamoDB.
- **Infra:** AWS (API Gateway, Lambda, S3), Cloudflare, Terraform.

## Education

---

- |                                                                                            |             |
|--------------------------------------------------------------------------------------------|-------------|
| <b>Trinity College Dublin</b> , M.Sc. (Hons), High Performance Computing – Dublin, Ireland | 2010 – 2011 |
| <b>University College Dublin</b> , B.Sc. (Hons), Astrophysics – Dublin, Ireland            | 2006 – 2010 |

## Technologies

---

**Languages:** Python (primary); TypeScript (working knowledge); C, C++, Go, Java (familiar)

**Python:** FastAPI, Django/DRF, Flask, ASGI, Pandas, NumPy, Numba, Dask

**Cloud & Infrastructure:** AWS, Azure, Kubernetes, Docker, Terraform, Argo Workflows, Airflow, GitHub Actions

**Data:** DynamoDB, MongoDB, PostgreSQL, SQL Server, MySQL, distributed data processing

**Engineering:** Backend engineering, API design, observability, CI/CD, technical leadership and mentoring