

Tathagata Ghosh

Research data · FAIR stewardship · Knowledge graphs · Data engineering

Magdeburg, Germany · tghosh12101997@gmail.com · gh.tathagata@gmail.com · +49 152 0825 8194 · linkedin.com/in/t-ghosh

SUMMARY

Research-oriented data professional with hands-on experience building **FAIR-aligned research data infrastructure**, **ontology-driven** modeling, and **knowledge graphs** (RDF/OWL) in a Helmholtz lab setting. Strong grounding in **research data management** at university level (DMPs, FAIR/CARE, tooling evaluation) and **health data integration** (FHIR, OMOP). Comfortable shipping **Python/SQL ETL**, **Docker**-based services, and **cloud** pipelines alongside scientists and RDM stakeholders. Open to RDM, data stewardship, research software, data engineering, and doctoral paths in Germany/EU.

CORE COMPETENCIES

RDM / stewardship: FAIR · CARE · DMPs · metadata · documentation

Semantics: RDF · OWL · Protégé · RDFLib · JSON-LD · knowledge graphs

Engineering: Python · SQL · ETL · MongoDB · PostgreSQL · MinIO · Docker · GCP **Analytics / viz:** Streamlit · PyVis · Plotly · Grafana

EXPERIENCE

Full-stack developer (freelance) OWIPL

Oct 2024 – present · Remote

- Multi-tenant SaaS (workforce/operations): microservices (React, Node.js, FastAPI, Flask, MongoDB), Docker Compose, auth/RBAC, audit logging; document pipelines (OCR, local LLM). Complements research-side work with production software and data-layer delivery.

Student assistant — data science & visualization Helmholtz-Zentrum Dresden-Rossendorf (HZDR)

Apr 2024 – Mar 2025 · Dresden, Germany

- Built **DRACO DataMaster**: FAIR-oriented research data platform for the DRACO laser experiment — ontology-driven schema, RDF/OWL in Protégé, Python ETL, MongoDB, CEPH/MinIO at scale, monitoring (InfluxDB, Grafana).
- Presented RDM work at OUTPUT2024 (TU Dresden).

Student RPA developer (GCP focus) Otto-von-Guericke-Universität Magdeburg

Feb – Nov 2024 · Magdeburg, Germany

- University HR proof-of-concept: document intelligence (Document AI, Vertex AI), structured extraction, BigQuery, orchestration with Cloud Composer (Airflow), Cloud Run, Terraform, Docker — project documentation reported a large reduction in initial screening time for the evaluated workflow.

Wissenschaftlicher Hilfskraft — Research Data Management Otto-von-Guericke-Universität Magdeburg

Mar – Sep 2023 · Magdeburg, Germany

- Data Management Plans; FAIR/CARE alignment; evaluation of RDMO and RADAR; bilingual RDM web content; support on legal/ethical data handling for scientific projects.

Technischer Mitarbeiter IT — Datenintegrationszentrum (MIRACUM) Universitätsmedizin Magdeburg

May – Dec 2022 · Magdeburg, Germany

- Clinical research data integration: FHIR server work, OMOP, Python ETL, PostgreSQL; deployment in local/intranet environments.

Earlier (India, 2018–2021): Operations, quality, and customer operations roles (e.g. apparel operations trainee, workforce/fraud operations) — foundation in process discipline, metrics, and coordination before M.Sc. and research data focus.

SELECTED PROJECTS

DRACO DataMaster (M.Sc. thesis, Feb 2025)

Metadata-driven pipelines, experiment ontologies, knowledge graphs, FAIR alignment — OVGU & HZDR supervision. Full thesis PDF on request / portfolio site.

OttoBot — educational RAG chatbot (OVGU)

Institutional knowledge access: LangChain, Llama-2, embeddings, FAISS, Streamlit.

OCR & slide recommendation (OVGU)

Information retrieval from lecture slides: Tesseract, TF-IDF, similarity ranking.

Project partner recommendation

Collaborative filtering for course teams — co-authored IEEE WCCCT 2023 paper.

EDUCATION

M.Sc. Digital Engineering — Otto-von-Guericke Universität Magdeburg · Apr 2021 – Mar 2025 · Thesis: DRACO DataMaster (HZDR)

B.Tech Mechanical Engineering — GITAM University, India · Sep 2015 – Apr 2019

RESEARCH & DATA INTERESTS

- FAIR data stewardship and metadata-driven research infrastructures
- Semantic interoperability and ontology-based integration
- Knowledge graphs for discovery, validation, and reuse in science
- Reproducible pipelines and maintainable research software