

Saahil Mehta

London, UK · self@saahil.co.uk · +44 7537 830191 · in/saahil-mehta · github.com/saahil-mehta · saahil.co.uk

SUMMARY

Associate Director leading a six-person data and insights team at South Pole, a 501-person climate solutions company. Built Orca, the company's enterprise AI assistant, and the governed platform underneath it: a semantic layer on OpenMetadata, self-hosted Dagster orchestration, automated access governance and FinOps across 195 GCP projects. Author of the data strategy adopted by ExCo, with ~295k EUR a year of direct savings in its plan. Hands-on across GCP, BigQuery, Terraform, dbt, Dagster, FastAPI and LLM systems.

EXPERIENCE

South Pole

London / Remote

Associate Director, Data & Insights

Jul 2026 to Present

- Own the company data strategy: defined, costed and secured ExCo endorsement for the Data Foundation Plan, a roadmap to 2028 across seven workstreams (semantic layer, orchestration, visualisation, AI platform, data quality at scale, access governance, SLAs). Lead a team of six serving 501 employees.
- Built Orca, South Pole's enterprise AI assistant: deep research grounded in company data, South Pole-format documents, decks and sales proposals, position-aware access. Within three weeks of the 2.0 release, 131 employees, over a quarter of the company, were querying governed data through it; 560,000+ requests served at 99.99% success.
- Designed Orca Semantics, a governed semantic layer on self-hosted OpenMetadata: version-controlled definitions, named domain owners, streamed into Orca as context and underpinning the visualisation layer.
- Replaced dbt Cloud with self-hosted Dagster on GKE in three weeks, validated by shadow runs across 1,800+ models matching production output, avoiding a three-year renewal at ~27k EUR a year (81k EUR over the term). Deploys run without a human step.
- Retiring Power BI (98 reports, 410 users) for Orca Visuals, removing ~200k EUR a year in licences; ~295k EUR a year of direct savings across the ecosystem plan.
- Took platform incidents from one every two days to roughly one a month with proactive alerting and self-healing pipelines.
- Automated data access governance: a self-service portal opens IAM grant pull requests through a GitHub App, returning 40+ hours a month to the team.
- FinOps across 195 GCP projects: anomaly dashboard with P99 alerting caught a 6x BigQuery spend spike in March 2026 and returned it to baseline; marts query spend down 46% year on year.
- Codified five rotating admin roles (support, sprint, access, reliability, tech debt) as runbooks and AI-assisted skills; set the team's AI harness standards and moved delivery to product-based thinking.

South Pole

London / Remote

Lead Specialist, Data Engineering

Mar 2024 to Jul 2026

- Migrated 18,000+ cloud resources from Pulumi to modular Terraform; standardised IAM, networking and policy-as-code so five business units share one platform.
- Shipped governed FastAPI/FastMCP services on Cloud Run that let LLM and RAG workloads reach sensitive carbon-credit transaction data under enforced access controls, with eval suites guarding accuracy.
- Built and launched Orca 1.0 to the company in June 2026, the foundation of the enterprise AI assistant now used across South Pole.
- Cut data marts BigQuery query spend 46% year on year through FinOps and DataOps practice; the platform ran governed, observable and within budget.
- Ran the data estate (95 datasets, 8,000+ tables, 21 TB) with strict standards, real backups, regular drills and full observability.

Publicis Groupe, Digitas UK

London

Senior Data Engineer II

Jul 2023 to Mar 2024

- Cut flagship GCP + Airflow workloads from four hours to two.
- Built GitHub Actions CI/CD that replaced manual deployments across client pipelines.

Publicis Groupe, Digitas UK

London

Senior Data Engineer I

Oct 2022 to Jul 2023

- Normalised cross-channel marketing data for clients including Samsung and Formula 1, on tight campaign deadlines.

Publicis Groupe, Digitas Mumbai

Mumbai

Senior Associate, Data Analytics Engineering

Aug 2021 to Sep 2022

- Cut a client's monthly cloud spend by 98% by rationalising the compute layer of the pipelines.
- Scaled Facebook Ads ingestion from 3 to 100+ weekly loads, taking deployment from hours to minutes.
- Led a team of 3 on automated pipelines and dashboards for near real-time campaign analytics.

WPP, GroupM · Mindshare

Mumbai

Manager, Data and Technology

Jun 2021 to Oct 2021

- Built a GCP data lake for Disney+ Hotstar that dropped reporting cycles from days to hours, unifying app, ad and CRM data for near real-time content decisions.

Publicis Groupe, Digitas Mumbai

Mumbai

Senior Associate, Data Science

Dec 2020 to Jun 2021

- Built and deployed multilingual sentiment analysis pipelines, doubling labelled-data coverage for planners.
- Created Power BI dashboards that spread across teams and cut manual update time by half.

Publicis Groupe, Digitas Mumbai

Mumbai

Associate, Data Science

Jul 2019 to Dec 2020

- Automated data collection and transformation workflows, turning months of manual effort into minutes.

PROJECTS

- Orca** (2026 - Present · AI assistant) Enterprise AI assistant for 501 employees: deep research over company data, South Pole-format documents, decks and proposals, position-aware access; 560,000+ requests at 99.99% success.
- Orca Semantics** (2026 - Present · Semantic layer) Governed semantic layer on self-hosted OpenMetadata: version-controlled definitions, named domain owners, streamed into Orca as context.
- Orca Visuals** (2026 - Present · Visualisation) The visualisation layer on the semantic layer: dashboards defined as code against certified metrics, replacing Power BI (98 reports, 410 users) and removing ~200k EUR a year in licences.
- Data Marts Core** (2026 · Orchestration) Orchestration for 1,800+ dbt models migrated from dbt Cloud to self-hosted Dagster on GKE in three weeks, avoiding a three-year renewal at ~27k EUR a year.
- Data Foundation Plan** (2026 - 2028 · Strategy) Seven-workstream data strategy to 2028 (semantic layer, orchestration, visualisation, AI platform, data quality, access governance, SLAs), costed at ~295k EUR/yr direct savings and adopted by ExCo.
- Data Access Governance** (2026 · Governance) Self-service access portal that opens Terraform IAM grant pull requests through a GitHub App; 40+ hours a month returned to the team.
- Cost Deep Dive** (2025 - Present · FinOps) FinOps dashboard across 195 GCP projects with P99 anomaly alerting; caught a 6x BigQuery spend spike and cut marts query spend 46% year on year.
- Enterprise Data Platform Migration** (2024 · Platform) 18,000+ cloud resources migrated from Pulumi to modular Terraform; IAM, networking and policy-as-code standardised for five business units.

VOLUNTEERING AND RECOGNITION

- Volunteer, Museum of English Rural Life (Jun 2025 to present): built an OCR + two-stage LLM pipeline that turned nine years of handwritten visitor cards into searchable data and a leadership report.
- Mentor, Data Science Festival (Mar 2026 to present): Spring 2026 cohort of early-career professionals moving into data and AI roles.
- Google Developer Specialist (2026): selected for Google's invite-only Developer Programme; technical content on Medium.

EDUCATION

Master of Science in Data Analytics

Glasgow, UK · 2022 - 2023

University of Strathclyde (Business School) · First-Class Honours / Distinction

Bachelor of Computer Engineering

Mumbai, India

University of Mumbai (Shah & Anchor Kutchhi Engineering College) · 3.25 GPA

Three-year Diploma in Computer Engineering

Mumbai, India

Shri Bhagubhai Mafatlal Polytechnic (SBMP) · 7.77 / 10

SKILLS

Platform: GCP, BigQuery, Terraform, Dagster, dbt, Airflow, Cloud Run, GKE, Cloud SQL, GitHub Actions, Python, SQL, TypeScript

AI: LLM applications, RAG, MCP, Agents, Evals, Vertex AI, Gemini, Claude, FastAPI, FastMCP

Governance: OpenMetadata, IAM, Policy tags and PII masking, Data quality, FinOps, Access governance

Leadership: Team leadership, Data strategy, ExCo communication, Product thinking, Mentoring, Hiring and development