

## YASIR KHALID

London, SE2 9QQ

Mobile: 07577798274 Email: yasirrkhalidd@gmail.com

LinkedIn: [www.linkedin.com/in/yasir-khalid](https://www.linkedin.com/in/yasir-khalid) | GitHub: [www.github.com/yasir-khalid](https://www.github.com/yasir-khalid)

**Summary:** I work in the data consulting industry, supporting tier-1 banking clients using large-scale data engineering, entity resolution, generative A.I., and modern cloud capabilities. I am a passionate learner, consistently filling my knowledge gaps across the software engineering and product development lifecycle – having worked in Agile/Scrum processes. I represent KPMG in badminton, have a passion for skydiving, and have led an investment pitch for my startup in the final stages of Y-Combinator.

### PROFESSIONAL EXPERIENCE

#### Assistant Manager – Data Engineering/Analytics, KPMG (London)

Jul 2022 - Ongoing

*Supporting clients with large scale entity resolution solutions for a 360-view on their customers using external data sources*

- Implementing Quantexa's entity resolution solution using Spark, Scala, Elasticsearch, AWS and Microservices, for financial crime detection in trade finance
- Responsible to delivering technical presentations, stakeholder management and liaising with cross-functional teams
- Experienced with robust ETL processes, workflow orchestrations using Apache Airflow and product release management
- Developing context-aware document comparisons using Python and LLMs, using frameworks like Pydantic, Langchain and Solidatus for data processing automation and data governance
- Enhancing team's capability around Google Cloud's Analytics and A.I. offerings, and working with AWS data engineering stack

#### Data Scientist, Ipsos (London)

Feb 2022 - Jun 2022

*Worked on an UKOM endorsed audience measurement tool to monitor apps and website usage across the U.K.*

- Developed production ETL jobs in Python involving data versioning, transformations and read/write to cloud storage
- Collaborated with a team of 20+ developers on a python package, and up skilled on the use of git flow and CI/CD

#### Co-Founder (Digital and Analytics), Afterschool (Pakistan)

Jan 2020 - May 2022

*Led the investment pitch for Afterschool in the final round of Y Combinator, a prestigious accelerator program*

- Guiding automation of student onboarding and other operational processes using Google App script, G-suite tools, and API calls
- Implemented user behavior tracking through Google Tag Manager and Hotjar, leading to data-driven improvements that significantly reduced bounce rate by 20%.
- Secured a \$5000 grant from the Retool startup program, enabling experimentation with low-code dashboards for internal use, fostering a culture of innovation.
- Led the analysis of customer CRM data from HubSpot, MongoDB, Postman, fostering data-driven decision making

#### Data Scientist (Placement), Spotlight Data (Nottingham)

Feb 2021 - Jun 2021

- Oversaw development of continuous data fetching and monitoring from social media sites using their commercial APIs
- Developed data pipeline that used *Facebook prophet* for time-series forecasting and presented reports using *Tableau*

### LEADERSHIP AND EXTRACURRICULARS

- Actively mentor junior developers and lead a thriving Python community of over 100+ developers at KPMG UK. Conducted engaging webinars on topics including Pydantic, Airflow, RAG implementation via Langchain, and best practices
- Enthusiastic member of the London Public Speakers (Toastmasters club), honing communication and leadership skills
- Represent KPMG's badminton team in inter-firm tournaments, fostering teamwork and sportsmanship
- Having participated in tandem skydiving, currently working towards achieving my British Skydiving A-license

### PERSONAL PROJECTS / INTERESTS

I love expanding my knowledge across different domains (especially backend) and improving my public speaking and executive presentation skills. My personal project, that I actively contribute towards currently is:

**Sports bookings app:** A webapp developed in Python that scraps sports centers' websites across London (asynchronously) to fetch available playing slots, near the user, based on their geolocation and advanced preferences. The app is deployed using a cloud-hosted Docker container, has a CI/CD pipeline (via GitHub Actions) that runs unit tests, packages in a docker container and automatically triggers cloud deployments via webhooks.

### SKILLS / CERTIFICATIONS

**Technical skills:** Python, Scala, Elasticsearch, Kibana, Apache Spark, Apache Airflow, Docker, Kubernetes, Microservices architecture, Continuous Integration (Github Actions, Teamcity, GitLab CI), AWS, GCP, Postman, FastAPI, Databricks, Git, Knowledge graphs, Network analytics, Neo4j

**Certifications:** Quantexa Certified Data Engineer, Databricks Certified Spark developer 3.0, Core consulting (Chartered Management Institute), Google Cloud digital leader, Generative AI for Developers Path (Google Cloud)

### EDUCATION

#### MSc Information Systems & Operations, University of Nottingham, (Distinction)

Sep 2020-Sep 2021

Dean's International Excellence Award (100% scholarship) – Relevant modules included *Data management and Manipulation*

#### B.Eng. Mechatronics, National University of Sciences & Technology (NUST), Pakistan, (1st Class)

Oct 2016-Jul 2020

Relevant Modules include *Object Oriented Programming, Data Structures and Algorithms, Robotics and Automation*