

Data Engineer - Empower Data-Driven Recruitment

StartUp

Vast

75k

ABOUT THE COMPANY:

This organization is rewriting the rules of recruitment. While the industry has stayed stagnant for decades, they've built a platform that leverages data science and machine learning to empower businesses with actionable insights into their workforce. With over 50 paying clients, they're already making waves and preparing to scale globally with the backing of top-tier investors.

Join the data team and be part of building something transformative, where your expertise in data pipelines and databases will be key to their success.

YOUR RESPONSIBILITIES:

• 1. Building Data Pipelines from Scratch

- Design and implement ETL (Extract, Transform, Load) pipelines to process raw data into actionable insights.
- Ensure scalability, reliability, and efficiency in data flows, optimizing performance for large datasets.

• 2. Creating Data Infrastructure

- Set up and maintain data lakehouses, lakes, or warehouses to store, organize, and manage structured and unstructured data.
- Use cloud-based tools (AWS, Databricks, or equivalent) to enhance the platform's storage and retrieval capabilities.

• 3. Data Collection & Integration

- Develop advanced web crawlers and scrapers to gather diverse datasets, maintaining compliance with ethical and legal standards.
- Build robust integrations with third-party APIs to enhance and enrich data streams.

• 4. Supporting Machine Learning

- Create and maintain infrastructure to support machine learning models, enabling data extraction, classification, and prediction workflows.
- Collaborate with ML Engineers to deploy models and ensure smooth data-handling pipelines.

• 5. Cross-Functional Collaboration

- Work closely with a diverse team, including developers, ML Engineers, and analysts, to understand business needs and translate them into technical solutions.
 - Contribute to brainstorming sessions for innovative ways to automate and scale data collection and augmentation processes.
-

WHAT YOU BRING:

- Proven experience in designing and building ETL pipelines in the cloud.
- Expertise in SQL and NoSQL databases, with a deep understanding of database design and performance optimization.
- Proficiency in Python, with a focus on data engineering tasks.
- Familiarity with data-related cloud services, such as AWS or Databricks.
- Experience building crawlers and scrapers to gather diverse datasets.
- Creative, analytical problem-solving skills with a passion for innovation.
- A Bachelor's degree (or higher) in Computer Science, Data Engineering, or a related field.

Bonus Skills (Nice to Have):

- Experience with Docker, GitHub Actions, and Infrastructure as Code tools.
- Knowledge of machine learning and MLOps frameworks.
- Familiarity with distributed computing for large-scale data processing.

WHAT'S IN IT FOR YOU:

- Competitive salary: €45,000–€75,000 based on your experience.
- Equity options: Share in the success of a game-changing company.
- A hybrid work setup, with at least two days per week on-site in the Netherlands.
- A collaborative, low-ego culture that encourages personal and professional growth.
- A lively social calendar, including laser tag, board game nights, and pub quizzes.

Ready to Apply?

Send your CV to r.sadloe@haystackpeople.nl or reach out via +31 (0)6 83 93 19 68 for a faster response.

Contact & solliciteren

Rashied Sadloe

Recruitment Consultant

r.sadloe@haystackpeople.nl

0683931968

<https://www.linkedin.com/in/rashied-sadloe-b7a39845/>

Groningen

<https://haystackpeople.nl>