

# Senior Data Engineer

Data Science India

This is a great opportunity to assume a technical position in a fast-growing and ambitious company. You will get a chance to address some of the most challenging real-world problems in the field of ML and data systems development and operations. You will work closely with Machine Learning engineering teams to anticipate company needs for creating data pipelines & datasets required to support ML solutions development. You will help create, improve and extend the underlying data infrastructure that powers our ML teams, thus simplifying the development and deployment cycles of our ML solutions.

## Responsibilities:

- Design and build effective, user-friendly data pipelines, tooling, and automation to accelerate Machine Learning at BlackSwan Technologies
- Collaborate with ML teams to drive the data infrastructure roadmap
- Support company's internal ML teams with data engineering best practices
- Build complex automated data pipelines for creating datasets for ML products, facilitate efficient data management
- Build data extraction & scraping pipelines for creating datasets from open web data
- Build basic data analytics pipelines
- Assist in the search for relevant datasets for various ML, NLP tasks
- Help establish standards, practices for managing ML teams' data infrastructure
- Write clean and tested code that can be maintained and extended by fellow engineers

## Requirements:

- A strong predilection for good software and the processes that make it
- 3+ years of experience as a Data Engineer working in Python
- Good working SQL knowledge and experience working with relational databases, as well as working familiarity with a variety of databases
- A successful history of manipulating, processing, and extracting value from large datasets
- Experience building and optimizing data pipelines & architectures
- Strong analytical skills related to working with structured datasets
- 4+ years experience with AWS or other public cloud platforms (GCP, Azure, etc.) and corresponding data tools (e.g. AWS Athena, etc.)
- Excellent verbal and written communication skills for effective communication in a multicultural environment with teams spread throughout the world

## Nice To Have:

- Experience managing and building data pipelines with private dumps of large knowledge bases such as Wikidata
- Experience with Big Data processing tools
- Experience with NLP and Machine Learning product development
- Familiarity with Knowledge Engineering, Graph technologies & Graph databases, Knowledge Graphs

## Who We Are:

BlackSwan Technologies is reinventing enterprise software through Agile Intelligence for the Enterprise – a fusion of data, artificial intelligence, and cloud technologies that provides unparalleled business value. Our multi-tiered enterprise offerings include the award-winning platform-as-a-service, ELEMENT, which enables organizations to build enterprise AI applications at scale for any domain quickly and at a fraction of the cost of alternatives. BlackSwan and its global partners also provide industry-proven applications that are ready-made and fully customisable for rapid ROI. These offerings are generating billions of dollars in economic value through digital transformation at renowned global brands. The private company maintains gravity centers in the UK, Europe, Israel, the US, and Sri Lanka.  
[www.blackswantechnologies.ai](http://www.blackswantechnologies.ai)

#### Equal Opportunities:

At BlackSwan Technologies, we prioritize diversity. We celebrate difference and embed it into every aspect of our workplace, from our community to our product. BlackSwan Technologies is proud and committed to providing equal opportunity employment to all individuals regardless of race, color, religion, sex, sexual orientation, citizenship, national origin, disability, Veteran status, or any other characteristic protected by law. In addition, BlackSwan Technologies will accommodate individuals with disabilities or a special need.

Apply