

Senior Big Data Engineer

If you are an analytical problem solver with a vast knowledge of Java, we have the perfect job for you. We are seeking a **Big Data Developer** to join our friendly team of experts. Your mission will be to improve the overall quality, width and depth of the data collected on companies worldwide.

About Our Customer:

The Customer provides business data enrichment with unprecedented coverage of private companies, accurate classification and in-depth insights based on real-time updates, for procurement, insurance, market intelligence and more.

Project Team:

When you join our team, you'll immerse in a culture where teammates support each other to achieve better results. We believe that together we are better and can find brilliant solutions by sharing ideas.

Requirements:

- Experience using statistical computer languages, preferably Scala (or R, Python, SQL, etc.) to manipulate data and draw insights from (very!) large data sets
- Expertise working with and creating data architectures (Spark, Cassandra)
- Knowledge of advanced statistical techniques and concepts (regression, properties of distributions, statistical tests and proper usage, etc.) and experience with applications
- High speed and uncompromising quality in your work
- A growth mindset, able to capitalize on unprecedented contexts through your skills and abilities
- An appetite to grapple with a variety of technical challenges
- The ability to quickly and effectively evaluate technical tradeoffs and translate them into relevant scenarios

English level:

Intermediate+

Nice to Have:

Any knowledge of machine learning techniques (clustering, decision tree learning, artificial neural networks etc.) and their real-world advantages/drawbacks

Responsibilities:

- Mine and analyze data from across the web
- Assess the effectiveness and accuracy of new data sources and data gathering techniques
- Develop custom data models and algorithms to apply to text datasets
- Use predictive modeling to increase and optimize data extraction and data quality at ingestion and post-processing
- Develop the company's A/B data testing framework and test the quality of the data continuously
- Develop processes and tools to monitor and analyze model performance and data accuracy.
- Prototype quickly to solve thorny use cases, without getting stuck in theory, as we're prone to shipping early and often
- Write well-designed, testable, efficient code
- Identify areas of opportunity and improvement