



Quality Assurance of Big Data Security product using automated query generation through R.

INDIUM SOFTWARE

Digital | QA | Gaming

*"Delivering exceptional customer-centric solutions across
Digital, QA & Gaming"*

Quality Assurance of Big Data Security product using automated query generation through R



Client Overview

- The client is based out of San Francisco Bay area.
- The client is a leading provider of data-centric security for next-gen data platforms.
- The client keeps enterprises in control of their data by allowing them to give users access to the data they need.



Business Requirements

- The client wanted to test how their product fared in various Big Data tools environments viz. Hive, Impala, SparkQL, Cassandra, HAWQ etc.
- The client was looking for an out of the box solution for query generation which would cover all the use-cases and generate queries in an automated fashion.

Quality Assurance of Big Data Security product using automated query generation through R



Key Challenges

- The challenge was that there were multitude of use-cases (join, subqueries, case when etc.) which needed to be tested.
- This further led to millions of queries for each tool.



Our Approach

- Our team of Data Scientists and Big Data Engineers devised combinatorial algorithms inspired from apriori algorithm to automate the query generation process.
- The algorithm was modified to generate meaningful queries for all the Big Data tools environment in which client's product worked.

Quality Assurance of Big Data Security product using automated query generation through R



Our Solutions

- The algorithm was written in R and packaged into modules to generate automated queries for all or one use-case.
- The algorithm could take any data model (set of tables) as input and generate meaningful queries (outputs non-trivial results) in an automated manner. The user just had to point the algorithm to the data model.



Results

Product QA became:


- 10X faster
- Replicable at each product deployment (client started shipping the algorithm code as-well-as sample query set to test the product deployment)
- Comprehensive (covered all the use-cases)
- Flexible with version changes (minimal user to generate new set of queries amenable to a new version)

THANK YOU




USA

Sunnyvale | Princeton


 Toll-free: 1-888-207-5969

INDIA


Chennai

 +91-44-6606 9100

Bengaluru


 +91-80-4645 7777

Mumbai

 +91-022-6215-4028


UK

London

 +44 773 653 9098

MALAYSIA

Kuala Lumpur

 +60 (3) 2298 8465



Sales Inquiries

sales@indiumsoftware.com



General Inquiries

info@indiumsoftware.com

www.indiumsoftware.com

We Are Also On

