A nighttime cityscape with illuminated buildings and streets, viewed through a circular frame. The scene is dark with various lights from buildings and streetlights, creating a bokeh effect in some areas. The circular frame is defined by a thin blue line.

psl

**Geolocation Company  
Increases Efficiency  
in Database Queries  
by 87%**

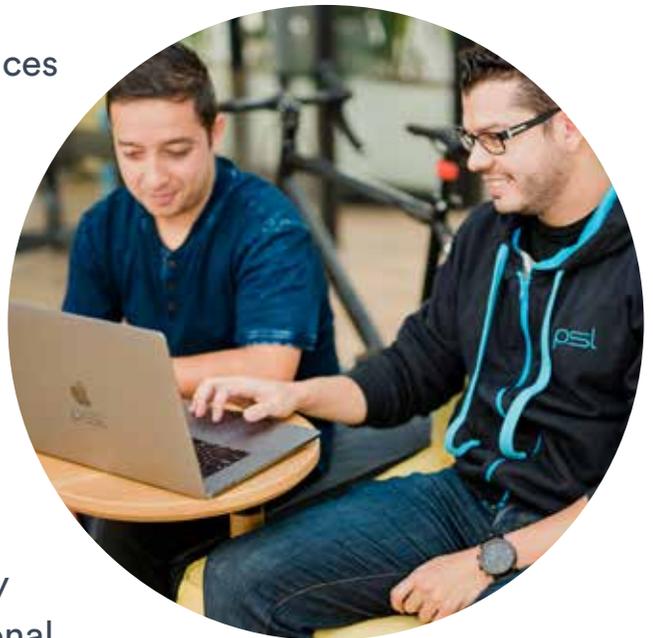
[www.pslcorp.com](http://www.pslcorp.com)

# Results

- Easy-to-use, web-based internal platform created to access specialized data in 1 query
- To be launched as a client-facing platform
- Platform evolved from only catering to Apache Zeppelin to supporting Zeppelin, Jupyter, Python, and more
- 87.7% reduction in process to pull relevant data from catalog
- Enabled easy access to information for relational analysis
- Completed migration from Amazon EMR to Kubernetes in 2/3 of the expected timeframe

## Executive Summary

The client, a global geolocation company with offices worldwide, receives millions of data points per second. With this amount of information at their disposal, they were interested in developing mechanisms to leverage data in order to scale tools capabilities and provide greater value. The client turned to PSL, a leading nearshore software development outsourcing provider, to build a team of highly skilled data scientists who could leverage machine learning algorithms, Big Data platforms, and data analytics to craft a highly interactive location services platform for professional and novice users, turning big data into valuable insights.



# Tech Stack



Apache Zeppelin

Jupyter

Python

## Challenge

PSL's client, a prominent geolocation company expecting to receive up to a billion data points per second by 2020, was faced with the question of how to best utilize the information. This wealth of intelligence offered a unique opportunity to expand their business tools and, in turn, provide added value to the end user, but also required highly specialized professionals to develop the strategies and mechanisms required to extract actionable insights.

Through a partnership with PSL, a leading software development outsourcing firm in Latin America, **the client assembled a multidisciplinary team which included highly skilled and experienced software engineers, data scientists, machine learning experts, architects, and DevOps specialists.**



# Solution

To devise ways for all type of users to consume and produce data from the client's ample data resources, the PSL team began by exploring a variety of Proof of Concepts (POC) for Big Data solutions. Because the project was based on a geographic relational database, they tested various Big Data platforms such as Spark, Hadoop, and others, to achieve geospatial analysis on a larger scale.

Because the client's core value resides in the value of its data, the objective was to allow the client's customers access, regardless of their technology of choice, and grant them the information needed to power their projects. The team initially built an environment that allowed customers to work easily with the client's data without having to refer to the code line or understand cluster details, so the solution, built in Scala Java with API libraries, mainly supported Apache Zeppelin. A later iteration would fully support Jupyter and Python to better facilitate projects for data scientists.

During the engagement, another application development project arose from the need to make the client's

database information more accessible for queries. To better leverage data for analytics development, the team focused on transforming previously difficult to pull information through an easy-to-access platform. Through a Scala-powered compiler and Spark pipelines, **they devised a system that cut query processing time by 87%, increasing efficiency and accelerating customers' development.** The product proved to be a huge success and is currently available internally to the client's workforce and is slated for launch as a client-facing platform as well.

**At the same time, the team completed a migration from Amazon EMR to Kubernetes in just 4 months—2 months faster than originally expected**—finalizing their DevOps implementation.

The products developed ensure the client's data can deliver the most value and power developments set to transform location services within the next decade. By leveraging the client's data, customers will be able to build highly customizable and rich user experiences and bring their services to the next level.