

ESTIMATING FLOOD RISK USING PREDICTIVE ANALYTICS

Trexin was chosen as the exclusive technology partner to design, develop, deploy, and manage an analytics platform to predict flood risk as far out as 30 years for individual parcels of land in U.S. coastal areas due to sea level rise associated with climate change.

BUSINESS DRIVER

Our Client, the founder and CEO of an innovative startup that developed an analytics model to accurately predict flood risk for a given parcel of land over a period of time, wanted to commercialize the technology to help consumers, businesses, and governments around the world better prepare and adapt to coast flooding, storm surges and sea level rise related to climate change. Their operational goal was to develop a completely automated online Internet-based system that could analyze multiple properties at once, but the current process needed extensive manual intervention to maintain accuracy and it took over two hours to analyze a single property.

The CEO asked Trexin to design and build a minimum viable product (MVP) that fully automated the predictive analytics model and complete a working prototype within 60 days so it could be presented to city leaders in Miami-Dade County (Florida) attending a local conference.

APPROACH

After our initial discovery sessions with our Client's team of data scientists who developed the analytics model, we quickly clarified the business problem but determined that a much deeper understanding of the metadata would be required for the various sources feeding the analytics model – data related to Light Detection and Ranging (LiDAR) measurements; Tidal Gauges; and Sea, Land and Overland Surges from Hurricanes (SLOSH) data. For example, the LiDAR data file generated for a single county can contain almost 2 billion data points. This volume of data required a big-data solution that scaled well and also minimized overhead costs.

To satisfy these scalability requirements at an incremental cost structure we based the solution architecture on an Infrastructure-as-a-Service platform and selected Amazon Web Services (AWS) as the provider. Then we selected the PostgreSQL open source database and extended it with PostGIS to arrive at a solution resembling Flickr's model for storing large amounts of data and geographic information. With these architectural pieces in place, we then built the application using Docker Containers to make the software easy to develop, manage, test, and deploy.

The key challenge we faced during the design and development process was emulating and automating a data scientist's judgement. We had the analytics models the data scientists developed, but the original algorithms didn't account for a human's ability to deal with details that fall outside of a formula's predefined parameters, making some tasks difficult to fully automate. We overcame this obstacle by having weekly video conferences with the data scientists to compare our automated results with their manual output. Each week we discovered inconsistencies, applied the new understanding to our algorithms, and returned the next week with a new iteration, repeating this process until we refined the computational model to produce a fully automated system.

RESULTS

In under 60 days, Trexin delivered a scalable prototype of the system that could take the address of a randomly selected parcel of land within a target geo and automatically calculate and graph its flood risk score without any user intervention. This prototype also succeeded in making it possible to perform this analysis on a portfolio of multiple addresses, and it reduced the calculation time from over two hours to less than five seconds per parcel. The MVP is now being extended and the product is publically accessible as a commercial product.

Analytics Platforms Assist with Flood Risk
Trexin was chosen as the exclusive technology partner to develop, build, and deploy an Internet-based analytics platform that predicted flood risk associated with climate change.

- Managing the Metadata**
Established a deep understanding of metadata generated by LiDAR, SLOSH, and Tidal Gauges and how manage it discharge databases
- Developing the Technology**
Selected Amazon Web Services and PostgreSQL to develop a system that stored data and geographic info
- Automating the Results**
Worked closely with Client's data scientists to close the gap between manual and automated results

✓ Completely automated
✓ Analyzes multiple addresses at once
✓ Reduced calculation time from 2 hours to 5 seconds

Completed in 60 days

CONTACT US

Other Practices
info@trexin.com
www.trexin.com