

Cloud Computing Technology — New and Innovative Approaches to Big Data

Khalid Lemzouji, Peter Solymos and Brandon Smith, Analythium Solutions Inc

Data has quickly become one of the most valuable resources for any organization within the environmental industry. New data collection tools such as remote sensors, drones and private satellite imagery has made it possible to gather massive sets of data—dubbed "Big Data" at lower costs than ever before. One area being reshaped by low-cost Big Data is environmental management and research. New and innovative technologies are providing researchers, businesses, and governments with higher quality data than ever before and opening new doors for informed decisions with respect to optimized resource use and conservation. At the same time, Big Data presents challenges; as the volume of the data increases larger infrastructure is required. The infrastructure includes more servers, larger database capacity, more powerful software, more database administrators for data management and security, and more data scientists to convert the data to useful information. This capital cost is beyond the reach of most small and midsize environmental businesses, government agencies, and research centers.

Cloud computing coupled with customized data application addresses most of these challenges. First, cloud-based computing allows users to scale based on need and you pay only for what you use. Second, data-applications deployed within a cloud-based network automatically convert data to information in a visual and reactive way without the need of data scientists, expensive analytical or GIS software. Third, cloud solutions can be combined with inhouse servers (hybrid architecture). Together, this approach ensures higher security for sensitive data, intractability between remote users all at lower costs than traditionally thought possible. To reinforce this innovative approach, we will explain the opportunities and challenges of Big Data by showcasing a cloud computing application running live simulation for approximately 200 million data points. The application, which Analythium provided the infrastructure expertise to run at scale, is the British Columbia Ministry of Forests, Land, Natural Resource Operations, and Rural Development prize winner out of 179 worldwide apps submitted in the 3rd Annual RStudio Shiny Content for analytics and machine learning visual applications.

Peter Solymos

Peter Solymos, Ph.D., is a technologist and biologist. He has been developing and applying statistics and machine learning in natural resource management, environment, and nature conservation for over 20 years. His experience and focus areas are in biology, environmental sciences, programming, web-based and cloud-based infrastructure development.

Khalid Lemzouji

Khalid Lemzouji, B.Sc., B.Eng., is a senior statistician and data scientist. He has 15 years of experience in environmental, public health, and pipeline reliability risk decision. His skills set focuses on the integration of statistical and machine learning tools to transform data to knowledge. As a professional statistician with a double bachelor's in chemical engineering and statistics he specializes in sampling design and statistical modeling for environmental data for regulatory compliance in Canada and USA.

Brandon Smith

Brandon Smith, M.Sc. is a professional biologist and has over a decade of experience working as both a human and ecological health risk assessor. He brings with him a strong background in both human and environmental toxicology. His work focuses on the unison of sound statistics and innovative technologies with environmental liability reduction and alternative remediation strategy design.