Bioremediation and Sustainable Environmental Technologies—2017. Fourth International Symposium on Bioremediation and Sustainable Environmental Technologies (Miami, FL; May 2017). ©2017 Battelle Memorial Institute, Columbus, OH, www.battelle.org/biosymp. A. Barton and S. Rosansky (Chairs)

Leveraging Big Data and Cognitive Computing for Remediation Selection, Benchmarking, and Environmental Portfolio Optimization

Craig E. Divine (craig.divine@arcadis.com) (Arcadis, Washington DC, USA)
Marc Paquet, Daniel Fortin, Amelie Rolland (WikiNet, Levis, QC, Canada)
Martin Beaudoin (Arcadis, Longueuil, QC, Canada)

Background/Objectives. Environmental experts today are challenged with the extremely rapid growth of information across the planet (Big Data), making it impossible for an individual to assimilate knowledge available and quickly integrate it to his or her expertise. While traditional information systems allow experts to have access to a wide variety of information sources, a detailed review of the scientific literature is a complex and laborious task. In addition, the information is generally complex and available in different formats; both structured (e.g., data tables) and unstructured (e.g., text reports). Therefore, integration, synthesis and use of all information remains very difficult and requires a significant amount of time. Obtaining and fully analyzing all appropriate information to identify all variables required to select the most efficient solution with high level of confidence is key to an economic and efficient remediation of contaminated sites. Furthermore, environmental portfolio managers are challenged with the optimization problem – how to make use of all of their project data, assess the performance of their current remediation projects, and maximize the return on the resources applied to restoration.

Approach/Activities. A new evolutive technological tool is underdevelopment to support site remediation selection and design by integrating and leveraging all available technical, regulatory and financial information. Specifically, this application incorporates the new machine learning and cognitive processing technology created by IBM/Watson. This technology allows experts to exploit very quickly and efficiently environment “Big Data”. Current cognitive tools are used to develop an access to different sources of structured and unstructured massive data in order to make their assimilation and treatment very quickly. In addition, this tool has the ability to enrich its knowledge exponentially by integrating the information, as soon as it is available. Consequently, the uncertainty related to the selection of the remediation technology is continually improving and is determined dynamically. Additionally, the use of this tool allows the professional to work in natural language environment. The application’s outcome is based on the search of benefits at different level: user, business, financial, environment and society by increasing and sharing the knowledge and the information related to development of remediation technologies for contaminated sites.

Results/Lessons Learned. The application has been trained, with a current focus on petroleum hydrocarbon site, by “reading” hundreds of site reports representing several hundred mature remediation sites. Based on this training, the application can extract most relevant site information from a report, such as contaminant type and levels and geologic conditions in order to develop recommend site-specific remedial approaches with anticipated performance and cost range estimates. This capability is currently being used to support remedy selection and performance benchmark for current remediation projects for clients in the oil and gas and rail sectors. The application will be further expanded to address other contaminant types, including chlorinated solvents, in the next development phases.