Transforming water & wastewater utilities with Data Analytics

- Interview with Léon Patrick Zerbib on data analytic solutions from Mazars - zettafox

Today, we are lucky to have Léon Patrick Zerbib, US Partner at Mazars zettafox, with us to shed some light on zettafox’s solutions and technology, the services offered, and the value for our sectors and clients, particularly water utilities.

Q: Can you introduce yourself, zettafox, and its history and business?

A: Good morning, I’m a partner at Mazars zettafox. I co-founded zettafox, before joining the Mazars NY office about a year ago.

zettafox was founded in late 2014, and is now a business unit of Mazars Group, providing Advanced Analytics, generating significant measurable value for customers through its highly innovative, leading-edge technology. As businesses are faced with fast-growing volumes of complex data, they demand better predictive performance (what is likely to happen), and insight on strategic action. We build on traditional data mining and predictive techniques, augmenting them with prescriptive functions (proprietary Machine Learning algorithms). While our core expertise is data science, our team can also assist clients with the design and construction of next-generation data architectures, as well as the delivery of tailored web applications, decision-support models and data visualization tools.

Q: According to a study from Markets and Markets, “The big data market is expected to grow from USD 29 Billion in 2016 to USD 68 Billion by 2021, at a high Compound Annual Growth Rate (CAGR) of 18.45%.” We have all heard or read about Big Data, but as a data analytics specialist, can you give us your description of Big Data?

A: First of all Big Data and Data Science are two different concepts. Big Data refers to infrastructure, data storage, access, and collection. Data Science refers to manipulation and analysis of the collected data. “Big Data” is aptly named: It means “a lot of data” in terms of volume, variety and frequency, which increases the complexity of a problem. The key point here is that the digital transformation of businesses enables data processing in order to better understand routine transactions and to automate them.

Q: The Harvard Business Review described Data Scientist as “the Sexiest Job of the 21st Century.” What is a Data Scientist?

A: The job of a data scientist requires a good balance between business acumen, computer science, and mathematics. A data scientist has at his/her disposal a spectrum of analytics tools to retrieve data and reveal patterns that cannot be seen by the “naked eye”, but which could be used to attract new customers, better understand purchasing or investment (e.g. CAPEX) mechanisms, improve the quality of products, predict maintenance, and better manage risk. At zettafox, we created a software platform that allows us to do all of the above, and more.

Q: What are predictive and prescriptive analytics? How do they differ from traditional analytical procedures?

A: Predictive analytics have been around for a long time. Tools and models being used have evolved over time from traditional business intelligence systems to more sophisticated predictive modeling. But at the end of the day, it comes down to evaluating the probability of an event (fraud, purchase, mechanical breakdown) based on past observations. The methodology is mostly top-down, non-exhaustive (sampling is often required), hypotheses-based and linear. Outputs are usually presented in the form of “predictive scores” that provide limited insights on specific drivers of a phenomenon.

Prescriptive analysis, on the other hand, is bottom-up (no statistics), exhaustive (granular), does not require assumptions (no bias) and can detect non-linear relationships. Outputs are in the form of rules or scenarios that explicitly spell out the root causes of a phenomenon and suggest corrective actions (for risk, for example). Prescriptive analytics complement and enhance the performance of predictive analytics.

Our pragmatic, results-oriented approach complements conventional techniques (e.g. scores) by including a layer of next-gen rules-based algorithm that has proven to be highly effective for many applications.

Q: Now that we understand a few key terms, can you explain who your clients are and what services and value you provide?

A: We work with most industries, from banking and insurance to telecoms, from consumer packaged goods to the travel/hospitality sector, from cosmetics and luxury products to the public
A: Modern data science can handle voluminous sets of heterogeneous data flows from operational, financial, and other internal or external sources, to feed efficient, optimized, consolidated, and automatically refreshed reporting tools. The water and wastewater industries are facing many challenges today. We truly believe that the use of data analytics is a key tool to overcome perennial issues posed by:

Aging Infrastructure and Asset Management (Network): Data science tools will help water utilities to enhance the efficiency of the capital allocation process and improve return on investment through better preventive maintenance (CAPEX impact).

Demand Forecasting in order to better capture shifts in customer demand driven by economics and demographics.

Operational Efficiency by optimizing the deployment of resources and field forces (OPEX impact).

Segmentation and Offer Customization: Data analytics could be used to better understand customer needs, with the aim of customizing offers, improving plans and tariffs, extracting more value, and improving collection of receivables.

Anomaly Detection and Prevention (End Customer) by improving the utilities’ ability to detect, fix and reduce anomaly occurrences (e.g., meter error, unusual consumption patterns, fraud and non-revenue water issues).

Finally, in terms of Water Quality Management, data science technologies allow analysis of root causes and scenarios of potential water quality issues for early detection and prevention.

Q: Water and wastewater utilities generate information in ever-increasing volumes, including on consumption. How could data analytics help utilities maximize their use of this data to the benefit of their operations?

A: You are right, water utilities generate data in ever-increasing volumes, characterized by frequent updates and high heterogeneity. This makes it imperative to be able to extract, clean, blend and analyze data from all possible sources in order to get a more insightful, accurate picture. Operators have started to leverage predictive techniques to better understand the “who and when;” they also need to be able to make better-informed decisions based on the understanding of the “why and how,” which is only possible via more advanced prescriptive (machine learning) technologies. As mentioned in the previous question, one of the key challenges the water industry will need to resolve in the upcoming years is aging infrastructure. In its 2017 Infrastructure Report Card, the American Society of Civil Engineers grades the US water and wastewater systems a D and D+, respectively. The American Water and Wastewater Association’s 2017 State of the Industry Report finds that the number one concern of water utility workers in North America is updating aging assets. Indeed, today many water utilities across the United States are wrestling with the ability to maintain their infrastructure.

Utilities are under pressure to do more with less. Data analytics could help in addressing outdated systems at a time when traditional revenue streams are drying up. At Mazars zettafox we have developed tools enabling the transformation of “just data” into “intelligent Water Data.” Data generated by water networks is abundant: Flow, pressure, pumps and pipes activities, water supply metering, asset performance, non-revenue water, etc. External data like meteorological variables (soil moisture, precipitation and snowfall levels) could also be used. That data, once captured, aggregated and organized, will give operators powerful information to take efficient actions like:

- Identifying parts of the network that may be under stress.
- Predicting early failures.

In October 2017, Mazars Group announced the acquisition of Zettafox, a leader in data science and prescriptive data analysis. The acquisition of Zettafox is part of an ambitious plan for growth, innovation and transformation of all Mazars’ business lines on a global scale, including a growth target of 35%-40% of Group revenue in advisory services in 2018.

The integration of Zettafox’s technology and team will have three main objectives:

- The deployment of advanced analytics for the benefit of Mazars’ clients.
- The acceleration of a more data-focused approach in the areas of risk, fraud, and forensics.
- The digital transformation of Mazars and its auditing, consulting, and accounting services by integrating new know-how into all methodologies and approaches.

"Data scientist: the sexiest job of the 21st century"

- Harvard Business Review