DATA STRATEGY: PART 3 DATA DRIVEN DECISION MAKING

Dr. Eva Savelsberg, Senior Vice President, INFORM;
Dr. Boris Michel, Head of Sales & Technical Sales INFORM DataLab;
Matthew Wittemeier, Senior Manager International Marketing and
Customer Relations, INFORM's Logistics Division

"IN CONTRAST TO SIMPLER BUSINESS INTELLIGENCE, DATA SCIENCE IS ABLE TO SIMULTANEOUSLY PROCESS A MORE CONSIDERABLE AMOUNT OF DATA TO ANSWER MORE COMPLEX QUESTIONS."

In 2020, we kicked off a three-part series on data strategy, starting with Part 1, Data Management is Fundamental, in Edition 100 of Port Technology. If you haven't started there, this would be the ideal place to begin your reading. Part 1 was followed up with Part 2, Decision Democratization Using Modern Data Analytics, in Edition 106 of Port Technology. Over the first two parts, we've built steadily from the basics of a data strategy through to analytics. Now, over a year after we started, we're concluding the series with Part 3, the third and final installment looking at data science applications.

The good news is these data science applications are hiding in plain sight all around us. In many areas, these so-called 'AI solutions' have started to make regular appearances and, in most cases, add real value to our everyday lives. Real-time weather and traffic predictions, fraud detection, process automation, and recommendation engines – think Amazon knowing what you're looking for before you do – are just a few popular examples from our everyday life. Even in the fields of maritime and intermodal terminal operations, many of these technologies have emerged over the past few years to become game-changers for our sector as well. This is why we wind up our trilogy on data value generation with this article on data science.

DATA SCIENCE 101

First things first – some vocabulary. When we talk about data science, we often see the term used interchangeably with AI, machine learning, and many others. To clarify:

- **Data science** is a discipline that uses scientific methods and other techniques to extract information from data, often huge amounts.
- **Artificial Intelligence (or AI)** is the emulation of human-like or better-than-human intelligence by machines.
- **Machine learning (or ML)** is a discipline within the field of AI whereby data scientists use data to train algorithms to perform a multitude of different tasks.

In this article, we will be mainly talking about data science, its benefits, its limitations, its characteristics, and how we can put it to good use within the terminal operations industry.

One of the main misconceptions about data science is that it is 'software' that one develops or purchases to automatically make decisions. Actually, data science is rather a continuous activity creating models, using statistical analysis, deploying algorithms, and much more in combination with often vast amounts of data collected within an organisation and/or bought from outside sources. The ultimate goal, usually, is to make better decisions. In contrast to simpler business intelligence (see Part 2 of this series), data science is able to simultaneously process a more considerable amount of data to answer more complex questions. One could interpret these models and algorithms as 'virtual analysts' that support 'human analysts' and the multitude of business decision-makers to better understand context allowing them to make the best possible decisions.

That is one of the misunderstandings we come across the most: companies are looking for a data science solution. But in our humble opinion, practicing data science is much more about the process, its flexibility, and its results and less about having something that consumes inputs and produces an output or answer- AKA a program.

OK, THANKS, BUT WHAT ABOUT MACHINE LEARNING SOLUTIONS?

We hear you. The next logical question is, "What about these ready-to-buy machine learning solutions that I get 15 LinkedIn messages about daily?" And that's a good question. Heck, we at INFORM also have a machine learning add-on module with a proven solution for reducing yard rehandles by 10% and improving the productivity of yard-lift equipment by 15%.

Having recurring patterns and business questions allows for standardisation – to some extent, at least. Frequently vendors, INFORM included, with deep knowledge in a specific vertical produce great solutions that can be easily deployed and integrated into terminal operations. Given that the right data is available in appropriate quantity and quality, these solutions can quickly be deployed with immediate results. Off-the-shelf solutions offer great time-to-value solutions as a shortcut.

On the flip side, while they may be able to provide insights for, or even solve, a single, short-term challenge, they will not deliver the long-term opportunities from "active" data science, nor will they be able to evolve over time to assist you in answering your questions as they evolve. In business, as it is with real-life, we

"IF WE WANT TO LOOK INTO OPERATIONAL AND/OR BUSINESS PHENOMENA, WE NEED DATA THAT DESCRIBE THESE AREAS."
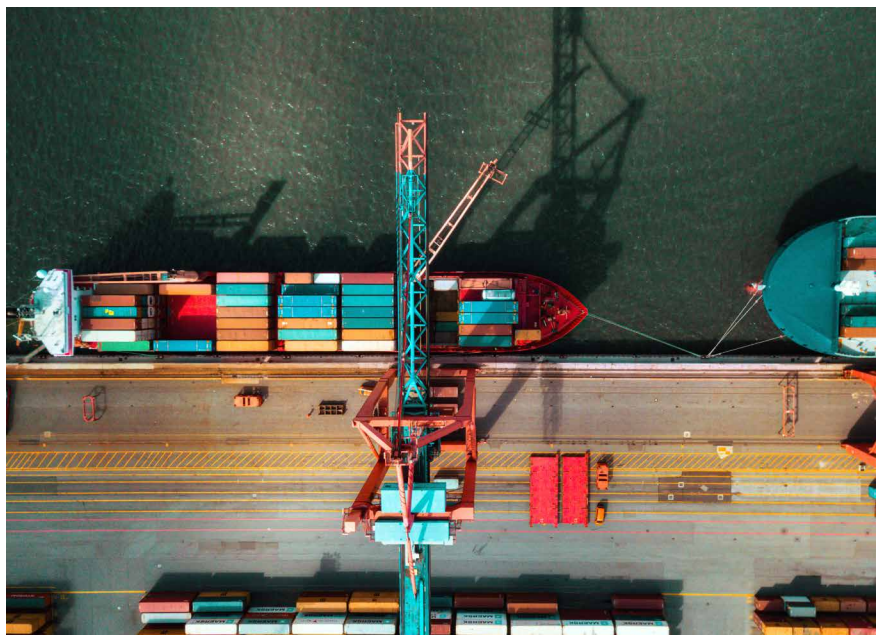
"AS WITH ANY LONG-TERM PROJECT, CRAWL BEFORE YOU WALK; WALK BEFORE YOU RUN."

know that once we answer one question, another will inevitably replace it. To draw another parallel to traditional business analytics, it is like comparing purchasing a pre-built template for, let's say, a sales dashboard (market solution) versus having the capability to create these dashboards internally to answer your questions now as well as subsequent questions. This requires data science operations.

IS THIS SOMETHING WE SHOULD HAVE IN-HOUSE?

If you're asking this question, the answer is probably "no," or at least "not yet." A data science team is very valuable and will create immeasurable value and an immense advantage to an organisation when deployed in the right context. Why "no" then? The issue will most likely be that if you are only starting out, the above-mentioned "context" is likely lacking. As with any scientific project, science never happens for its own sake. It is used to answer a potent question discerned from an observation. You must first ask yourself, are the project scope, the question, and the expectation of the project clear?

The next and equally important element is resources. We are not talking about laptops and computing power; we are talking about data. If we want to look into operational and/or business phenomena, we need data that describe these areas. Finding correlations between crane movements, vessel names, and delays is next to impossible if only half of these records exist in a semi-structured, electronic form. Only starting to measure these parameters and still hiring a data scientist for the project will lead to great frustration on all sides.

Another crucially important aspect is the impact that these results will have. Is the organisational culture ready to accept insights and integrate them into their everyday decision-making processes? In our experience, organisations just getting started with data science have a long way to go towards building data insights into their day-to-day decision-making strategies.

As with any long-term project, crawl before you walk; walk before you run. Don't be discouraged, though. The first step is always the hardest. A specialised data science consultant can help your organisation see the benefits of leveraging

data to improve decision-making without the long-term cost risks or the frustration.

ENOUGH WITH THE THEORY, WHAT'S IN IT FOR US?

Data science offers broad utility across the entire enterprise. Common, classical use cases can be found in sales or finance, such as forecasting, for example. Statistical extrapolation is one of the oldest tools used to predict the future. Modern regression algorithms reach far beyond polynomials of the third degree but allow fitting of functions of very high complexity and, therefore, high precision. Be careful; unforeseen large-scale events like the COVID-19 pandemic can disrupt these patterns. And unfortunately, the short, medium, and long-term disruptions to our industry are increasing both in terms of frequency and duration. Climate change, cybersecurity threats, trade wars, and global transnational events are impacting the global supply chain unlike any time before.

Apart from these, maybe mundane, forecasting examples, other examples that beat closer to the heart of operations can be found. A use case we'd all recognise is real-

time image processing combined with optical character recognition (OCR). This technology stack, along with the AI working behind the scenes, allows for automated registration of equipment, trucks, or containers. And when combined in a network, these could be used for location purposes or usage quotas. What's more, we haven't even begun to scratch the surface of what insights this data source could provide when combined with other operational data. Imagine a network of cameras across your terminal able to accurately predict operational bottlenecks five, 10, or even 15 minutes before they happen, giving you time to preemptively address these issues.

Another use case is the use of image processing to detect defects on a multitude of surfaces. A recent use case INFORM deployed was the aerial recognition of bird droppings on new vehicles using drones combined with a proprietary detection algorithm. This helped significantly lower damage incurred and operating costs.

The use of other integrated sensors allows the realisation of use cases like predictive maintenance. While this use case has been around for a while, it has gone through some evolution. It started by trying to predict whether or not equipment might soon fail based on historical data- i.e., when it failed in the past. What we're seeing today is the proliferation of IoT-bound sensors which allow for the detection of even the slightest changes in operations. As an example, an IoT sensor can detect the vibrations in a crane's motor to predict the potential for failure. This is often referred to as preventive maintenance, as well.

These are just a few examples of how data science and all its sub-disciplines can help companies become more efficient and more effective. In combination with comprehensive data analytics deployed on a solid data platform, following a data strategy that aligns with the overall organisation's strategy delivers a huge, positive impact on the administrative and operational body.

THAT'S A WRAP

In this three-part series, we started out with the basics of data strategy and have built up to some of the more advanced uses we're seeing in the industry today. But, by no means is this a comprehensive view of what is being achieved or what can be achieved in the industry. At INFORM, we combine a dedicated team of data scientists with our container terminal operational specialists to help our end customers derive insights all across their operations and organisations. Bringing over 25 years of operational experience deploying AI and OR-based optimisation into container terminal operations and combining it with a team of 35-plus data scientists ensures that our approach to data strategy is robust and considered.

So, whether you're just getting started on your journey to leverage data or are looking to tackle the next challenge, we're here to assist you on that path to data-driven, continuous improvement.

"FOLLOWING A DATA STRATEGY THAT ALIGNS WITH THE ORGANISATION'S STRATEGY DELIVERS A HUGE IMPACT."
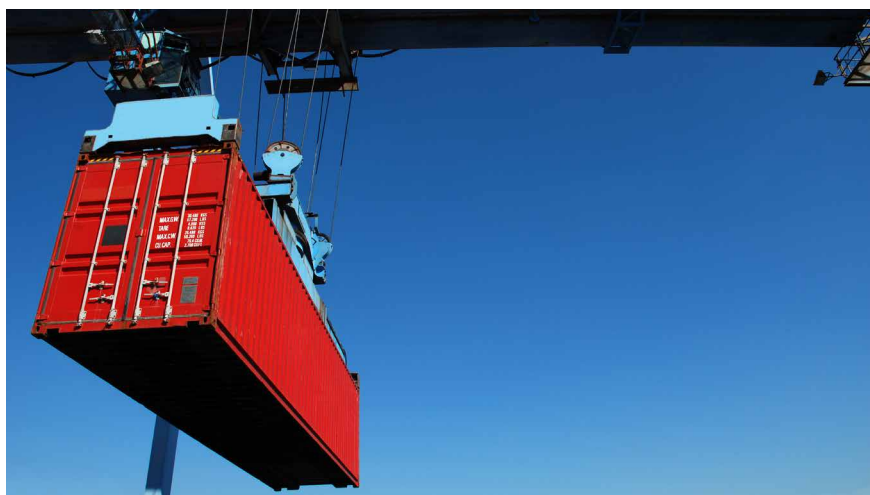
ABOUT THE AUTHOR

Boris Michel is the lead for Inform's DataLab Sales and Strategy effort. As the head of the Sales and Technical Sales department Boris and his team are working closely with customers and partners to drive the adoption of solutions and processes that enable people to make healthy, data driven decisions. These approaches cover the entire data value added chain starting from raw data all the way to sophisticated Business Intelligence Applications or AI solutions based on modern data science. Dr. Eva Savelsberg is Senior Vice President of Inform's Logistics Division. She specialises in optimisation software that renders a wide range of operational processes more productive, agile, and reliable. Eva is also lecturer at the University of Aachen (RWTH), where she received her PhD in Mechanical Engineering in 2002. Eva has published five books and over 40 papers on innovation in freight transportation. Matthew Wittemeier is Senior Manager International Marketing and Customer Relations at Inform's Logistics Division where he's become a thought-provoking contributor to many industry publications and conferences. He's co-author of the multi-award winning 2038: A Smart Port Story – a novella about the future of technology and the social challenges it may bring. In addition, he serves on the board of YoungShip Rotterdam.

ABOUT THE ORGANISATION

Inform specialises in AI and optimisation software to improve operational decision-making. Based in Aachen, Germany, the company has been in the optimisation business for 50 years and serves a wide span of logistics industries including maritime, intermodal, and inland terminals. With a broad range of standalone and add-on software modules, Inform's unique blend of algorithmic based software expertise, rich industry experience, and big world thinking delivers huge value for their customers.