Data management — enabling a timely, effective and cost-efficient decision making by track machine and railway infrastructure operators

The experts at Plasser & Theurer Connected GmbH, a subsidiary of Plasser & Theurer, who come from various backgrounds, such as software engineering, physics, mathematics and statistics, engage in the collection, networking, integration and evaluation of track construction and maintenance machine data, with the aim to obtain concrete maintenance guidelines, in order to reduce track and machine maintenance costs. Should an unexpected problem occur, the information can assist in finding the best possible solution that is tailored to the needs of the client – having the right data available at the right time can assist both track machine and railway infrastructure operators to make timely, effective and cost-efficient decisions.

Worldwide, there is an increase in demand for mobility. The use and analysis of data provides information that could positively impact mobility.

Bernhard Maier
Chief Operating Officer – P&T Connected
Plasser & Theurer Connected GmbH

Current studies predict a worldwide increase in data volume generation to 163 zettabytes (1 zettabyte = 1,000,000,000 terabytes = 10^{21} bytes) by the year 2025 (Fig. 1) [1]. Storage and processing of this data could create an added value for companies in general, and railways in particular.

IoT – "Internet of Track"

IoT (Internet of Things), the networking of objects, will have just as much of an impact on data management and analysis as the increase in data volume will have.

It is expected that IoT will significantly improve processes and operational procedures, in combination with new business models. Gartner, in [2], predicts that there will be 20 billion IoT devices by the year 2020 (see table below); Cisco even forecasts that there will be up to 50 billion interconnected objects. This brings with it the opportunity that IoT may also become the "Internet of Track".

Category	2016	2017	2018	2020
Consumer	3.9630	5.2443	7.0363	12.8630
Business: Cross-Industry	1.1021	1.5010	2.1326	4.3814
Business: Vertical-Specific	1.3166	1.6354	2.0277	3.1710
Total	**6.3817**	**8.3807**	**11.1966**	**20.4154**

Installed IoT devices by category (billion units) [2]

If the analysts at Gartner are to be believed, more than half of all major business processes will, in some form, be integrated with IoT. It is expected that the required technology will become considerably more practical and its integration easier [3]. The potential of IoT as regards track construction and maintenance lies in the continuous collection of data, for instance by linking up sleepers that daily report their status. In this way, IoT could become a valuable data source for analysis purposes, as well as for planning track inspections and maintenance.

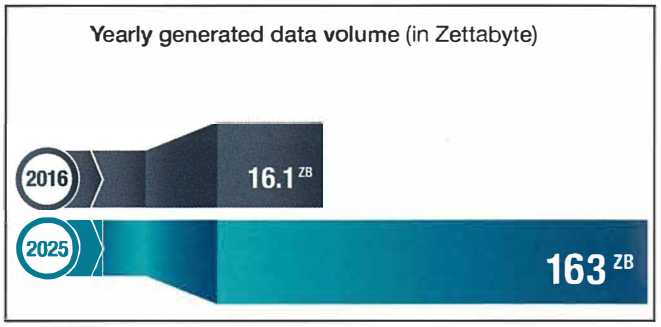

Fig. 1: Increase in data volume [1]

Valuable machine condition data

Regular condition inspection of track construction and maintenance machines, either on-site or in a depot, is regarded as part of their servicing and maintenance. There are up to 250 sensors integrated in these machines that supply a wealth of information, which can be evaluated for control purposes, and also allows the tracking of condition and processes at any given time.

When also taking into consideration additional data as regards:
— spare parts orders;
— maintenance contracts;
— error messages;
— log files; and
— weather and geo data;
a wide range of information about the condition of a machine becomes available. The analysis can be further enhanced by linking historical with current information – this combination of data sources allows a completely new level of analysis.

A lot of data, but where can the key information be found?

When analysing data, the primary objective of "data scientists" is to transform big data into smart data. Whether it concerns infrastructure management or machine condition, there is an increasing demand to provide system-relevant information for action recommendations.

By all means, the demand for flexible evaluations and analyses is growing. This poses new requirements on the "digital" strategy of companies. System solutions are in greater demand than ever before. Data-driven wear models and specific action recommendations will help reduce maintenance costs. Each company's strategic decision on how it deals with data, data management and big data analyses will have a direct impact on both product life-cycle cost and economic success.

A key element of data management is the categorisation of data as regards its:

— *structure:* in general, data can be categorised as:
  — structured: this concerns data that can be standardised or put into a scheme (table format) – it has a predefined format;
  — semi-structured: this concerns data that does not have any specific characteristics, but has a certain "hidden" structure;
  — unstructured: this concerns data that has no formal structure (e.g. images, text or audio data) and, therefore, does not fit into any scheme;
— *volume and speed:* as noted earlier, the volume of data that is becoming available is ever-increasing. The challenge lies in the rate at which all this data can be processed. High data rates (data per unit of time) are immensely important when it comes to "real-time" applications;
— *temperature:* whether data is discerned as either "hot" or "cold" is of influence in decision-making. Initially, new data has a high relevance ("hot") – the older a data set becomes, the more it tends to lose its relevance, i.e. it gets "colder". However, it should be noted that, in aggregate, a large set of cold data could be very valuable for analysis purposes, particularly in the case of products that have a long life cycle (see also (Fig. 2)).

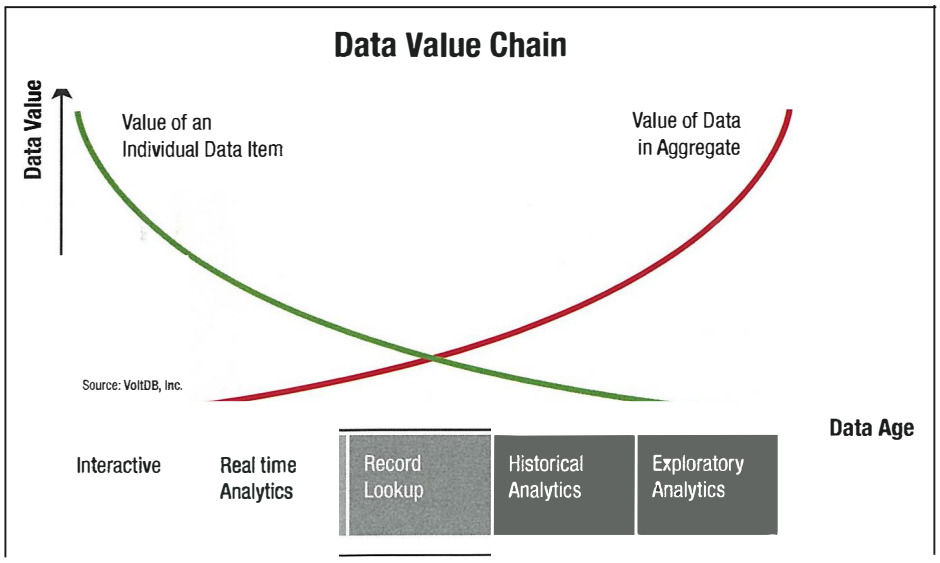

Fig. 2: Data value chain [4]

The aforementioned requires an efficient data management technology, the selection of which should be based on the intended use, as well as structural characteristics and access pattern – the selection of the right technology results in optimised costs and an efficient use of resources.

By adopting the right data management strategy, track construction and maintenance companies can save a lot of time and money. In Fig. 3, a selection of data management technologies based on Amazon AWS is shown.

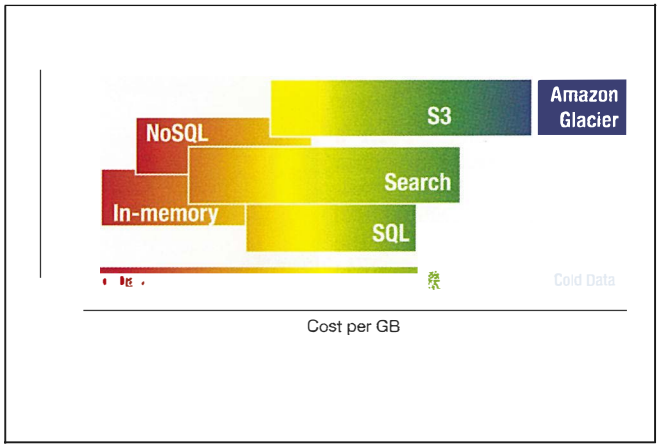

Fig. 3: Technology selection for data management based on Amazon AWS [5]

Should data go in the cloud?

Cloud services come with risks and opportunities! In cloud computing, clients rent storage and/or processing capacity instead of setting up their own physical infrastructure, such as a server or the like. Cloud service providers offer three abstraction levels:

— Infrastructure as a Service (IaaS) – infrastructure level;
— Platform as a Service (PaaS) – developer level;
— Software as a Service (SaaS) – user level.

Of course, cloud services have both advantages and disadvantages. On the one hand, they require no costly investment in one's own infrastructure; the services used are charged on a "pay per use" basis. This allows free scaling, depending on requirements and data volume. In addition, they ensure high availability and redundancy, and make administration easier. Further, a global use of services can easily be implemented.

On the other hand, cloud computing will not function without internet access and can be slow when there is insufficient bandwidth – here, the size of the selected cloud service provider is essential. Also, the matter of data security and data protection needs to be taken into account. For instance, in the case of trans-national data processing, various regulations would have to be observed.

Big data and data science enable the move from time-dependent to condition-based maintenance

Big data and cloud computing will often go hand in hand. The excellent scalability and distributability of cloud services lend themselves to big data applications, such as identifying patterns from the various data. Everything revolves around analysing, forecasting and data-driven modelling and, in particular, around providing the tools for improved and speedy decision-making.

Data-driven analyses should make the railway system more competitive. Big data and data science allow the move from time-dependent to condition-based maintenance.

Condition-Based Maintenance (what is happening?)

Condition-based maintenance (CBM) is already adopted by many companies. By means of sensors, the current condition of an object and its components is established (condition monitoring). Maintenance is carried out as soon as it is deemed necessary, i.e. when one or more indicators show signs of deterioration or exceed certain threshold values.

Predictive Maintenance (what will happen?)

Predictive maintenance takes CBM one step further. Rather than just analysing the existing situation, potential future events are forecasted, which can be generated through the use of a number of approaches, such as mathematical modelling in combination with machine learning and neuronal networks. Model-based and data-driven knowledge results in hybrid methods for generating new knowledge [6].

In Fig. 4, a hybrid data analysis (physical modelling combined with machine learning) of a Plasser & Theurer 09-series tamping machine is shown. Various impacts from working the tamping unit have been examined, in order to produce a forecast model as regards its wear condition. The data volume analysed amounts currently to around 1.1 terabytes and is continuously increasing. The blue and green dots mark correct forecasts for occurring wear conditions, as compared to real deployment data. The forecast accuracy of the model used is currently around 90%. This means that a worn tamping unit can be identified much earlier, thus allowing a better planning.

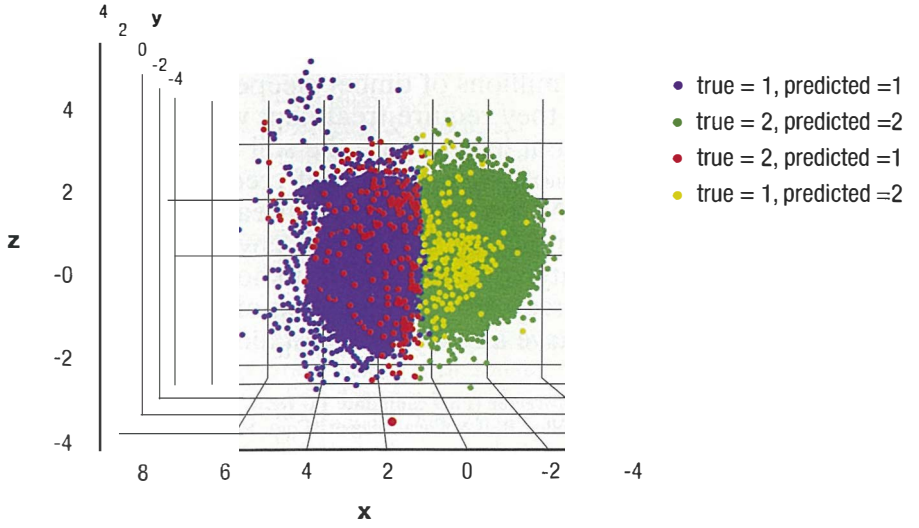

Fig. 4: Hybrid data analysis of a Plasser & Theurer 09-series tamping machine

Prescriptive Maintenance (what should happen?)

Prescriptive maintenance is about the future. Rather than analysing what will happen, actions to be undertaken and decisions to be made are immediately proposed. The decisions and interaction between decisions, as well as their influences and their relationship to the starting position are hereby analysed, in order to determine the optimum process [7].

Final remarks

Whether the linked-up refrigerator will soon do our shopping or a linked-up pedometer will make us healthier remains to be seen. Data and data analysis have an impact on our lives, as well as on track construction and maintenance. Correct data management in track construction and maintenance will create a considerable added value, as the availability of the right data will assist both track machine and railway infrastructure operators to make timely, effective and cost-efficient decisions.

References

[1] https://de.statista.com/statistik/daten/studie/267974/umfrage/prognose-zum-weltweit-generierten-datenvolumen/
[2] https://www.gartner.com/newsroom/id/3598917
[3] https://www.bigdata-insider.de/internet-der-dinge-gewinnt-stark-an-einfluss-a-518990/
[4] https://gotocon.com/aarhus-2012/thebigdatavaluecontinuum
[5] Own diagram based on Big Data Architectural Patterns & Best Practices on AWS, Siva Raghupathy, https://www.slideshare.net/AmazonWebServices/aws-reinvent-2016-big-data-architectural-patterns-and-best-practices-on-aws-bdm201
[6] https://www.cancom.info/2016/05/prescriptive-analytics-ist-die-zukunft/
[7] https://www.searchenterprisesoftware.de/definition/Prescriptive-Analytics