

Datenmanagement im Gleisbau

Richtige Daten zur richtigen Zeit helfen den Betreibern von Gleisbaumaschinen und Infrastruktur, Entscheidungen früher, effektiver und kostengünstiger zu treffen.

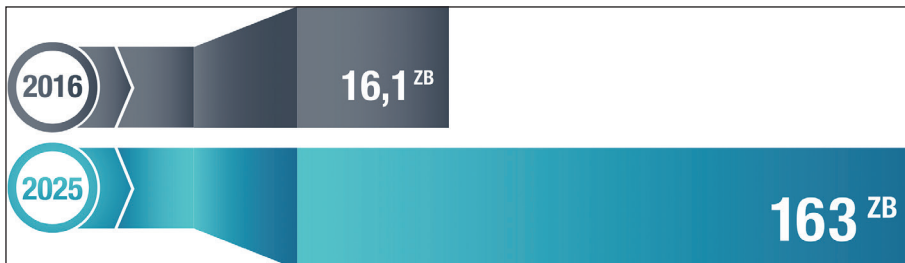


Abb. 1: Steigerung Datenvolumen [1]

BERNHARD MAIER

Plasser & Theurer Connected, eine Tochtergesellschaft von Plasser & Theurer, beschäftigt sich mit ihren SpezialistInnen aus verschiedenen Bereichen wie Software Engineering, Physik, Mathematik und Statistik mit der Erfassung, Vernetzung, Integration und Auswertung von Gleisbaumaschinen- und Maschinendaten. Ziel ist es, auf dieser Basis konkrete Handlungsanweisungen für die Wartung zu erhalten, um so die Wartungskosten für Maschinen zu senken. Droht also an Maschine oder Fahrweg ein unerwartetes Problem, kann anhand der bereitgestellten Informationen bestmöglich reagiert werden; ganz nach den Bedürfnissen der Kunden.

Weltumspannend steigt der Bedarf an Mobilität. Die Verwendung von Daten und deren Analyse liefert hier Informationen, um die Mobilität positiv beeinflussen zu können. Daten sind daher der Kraftstoff des Motors, der uns in die Zukunft bringt. Die digitale Transformation hat den Bahnsektor erfasst, und sie wird diesen nachhaltig verändern. Aktuelle Studien gehen von einer weltweiten Verzehnfachung der pro Jahr generierten Datenmenge auf 163 Zettabyte (1 Zettabyte = 1 000 000 000 Terabyte = 10^{21} Byte) bis 2025 aus, wie Abb. 1 zeigt. Die Speicherung und Verarbeitung wird hier zu einem essenziellen Mehrwert für Unternehmen.

IoT oder Internet of Trains

IoT („Internet of Things“), die Vernetzung von Gegenständen, wird das Datenmanagement und die Analyse im Gleisbau ebenso beeinflussen wie die steigende Datenmenge. Die Erwartung an IoT zur Verbesserung von Prozessen und Betriebsabläufen verbunden mit neuen Geschäftsmodellen ist groß. Gartner [2] prognostiziert derzeit 20 Milliarden IoT-Geräte bis 2020 (siehe Tab. 1), Cisco geht sogar von bis zu 50 Milliarden vernetzten Gegenständen aus. Dabei besteht die Chance, dass IoT auch zum „Internet of Trains“ bzw. „Internet of Rails“ wird.

Glaukt man den Analysten von Gartner, wird mehr als die Hälfte aller großen Geschäftsprozesse in irgendeiner Form mit dem IoT vernetzt sein. Es ist zu erwarten, dass die Technologie wesentlich praktikabler wird und die Integration einfacher [3]. Das Potenzial für den Gleisbau liegt in der kontinuierlichen Sammlung von Daten. Eine vernetzte Eisenbahnschwelle, die ihren Status einmal pro Tag meldet, ist erst der Anfang. IoT wird so zu einer weiteren wertvollen Datenquelle für Analyse Zwecke und zur Planung von Wartungs- und Inspektionsmaßnahmen.

Gleisbaumaschinen liefern wertvolle Daten

Bei Gleisbaumaschinen ist die regelmäßige, intervallbasierte Inspektion vor Ort oder in einem Depot Bestandteil der Überprüfung und Wartung. In Gleisbaumaschinen sind bis zu 250 Sensoren integriert, die den Verantwortlichen eine Fülle von Informationen zur Verfügung stellen.

Diese Sensordaten wurden und werden in festen Intervallen für Steuerungszwecke ausgewertet und erlauben zu jeder Zeit die Verfolgung von Zuständen und Prozessen.

Die Berücksichtigung zusätzlicher Daten von:

- Ersatzteilbestellungen,
- Instandhaltungsaufträgen,
- Fehlermeldungen,
- Logdateien sowie
- Wetter- und Geodaten.

liefert weitreichende Mehrinformation über den Maschinenzustand.

Die Analyse wird durch Vernetzung von historischen und aktuellen Informationen verbessert. Diese Kombination vielfältiger Datenquellen erlaubt ein gänzlich neues Niveau in der Analyse.

Viele Daten, doch wo steckt die Schlüsselinformation?

In der Analyse der Daten geht es für die Experten („Data Scientists“) vor allem darum, aus Big Data schlussendlich Smart Data zu machen. Ob Infrastrukturmanagement oder Gleisbaumaschine, es steigt die Notwendigkeit, systemnahe Information für Handlungsempfehlungen bereitzustellen. Der Bedarf an flexiblen Auswertungen und Analysen wächst. Das stellt neue Anforderungen an die „digitale“ Strategie von Unternehmen. Systemlösungen sind gefragter denn je. Datengetriebene Verschleißmodelle und spezifische Handlungsempfehlungen helfen bei der Reduktion von Instandhaltungskosten. Die strategische Entscheidung jedes Unternehmens, wie es mit Daten, Datenmanagement und Big Data-Analysen umgeht, wird direkten Einfluss auf die Product Life Cycle Costs und den wirtschaftlichen Erfolg haben.

Ein zentrales Element im Datenmanagement ist die Kategorisierung von Daten hinsichtlich:

- Struktur,
- Volumen und Geschwindigkeit sowie
- Datentemperatur.

Daten sind nicht gleich Daten

Daten können von ihrem Strukturierungsgrad in drei Kategorien unterteilt werden:

| 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 |
| Grand Total | 6.3818 | 8.3806 | 11.1966 | 20.4154 |

Tab. 1: Installierte IOT-Einheiten nach Kategorie (Milliarden Einheiten) [2]

- strukturierte Daten,
- semistrukturierte Daten und
- unstrukturierte Daten.

Strukturierte Daten können normalisiert bzw. in ein Schema (Tabellenform) gebracht werden und haben ein vorgegebenes Format. Unstrukturierte Daten wie Bilder, Text- oder Audiodaten haben keine formale Struktur und passen in kein Schema. Semistrukturierte Daten zeigen zwar keinen festen Typ, haben aber eine gewisse „versteckte“ Struktur.

Das ständig steigende Datenvolumen ist klar erkennbar (Abb. 1). Die Herausforderung ist die rasant zunehmende Geschwindigkeit, mit der Daten generiert werden. Daraus resultieren ständig steigende Datenraten (Daten pro Zeiteinheit). Hohe Datenraten sind vor allem bei „Real-Time“-Anwendungen von enormer Bedeutung, denn nichts ist ärgerlicher als lange Wartezeiten.

Ob ein Datensatz „heiß“ oder „kalt“ ist, macht einen Unterschied. Der einzelne aktuelle Datensatz hilft uns bei täglichen Entscheidungen. Am Beginn ist die Relevanz eines neuen Datensatzes groß – je älter ein Datensatz, desto mehr verliert er an Relevanz und wird sozusagen „kälter“. Prinzipiell sollten alle wertvollen Daten gespeichert werden. Die richtige Auswahl fällt oft schwer, da sich der Wert von Daten im Vorfeld schlecht bestimmen lässt. Durch die Aggregation vieler kalter Daten steigt deren Wert für Analysezwecke (Abb. 2). Dieser ist besonders bei Produkten mit langen Lebenszyklen von Bedeutung.

Mit Struktur, Temperatur und Volumen „vernetzt“ sind auch die Kosten, die für Datenmanagement entstehen. Anhand der geplanten Nutzung in Zusammenhang mit den strukturellen Eigenschaften und dem Zugriffsmuster ist die geeignete kostenoptimierte Technologie auszuwählen. Die richtige Wahl führt zu Kostenoptimierung und effizientem Einsatz der Ressourcen. Eine richtige Datenmanagementstrategie wird Unternehmen im Gleisbau viel Zeit und Geld sparen. Ob SQL oder NoSQL, Objektspeicherung oder historische Daten. Abb. 3 zeigt eine Technologieauswahl auf Basis Amazon AWS.

Sollen die Daten in die Cloud?

Cloud-Dienste haben Risiken und Chancen! Bei Cloud Computing mietet der Kunde Speicher bzw. Rechnerleistungen, anstatt eigene physikalische Infrastruktur in Form eines Servers etc. aufzubauen. Drei Abstraktionsschichten werden von Cloud-Anbietern zur Verfügung gestellt:

- Infrastructure as a Service (IaaS) – Infrastrukturerbene,
- Platform as a Service (PaaS) – Entwickler-Ebene und
- Software as a Service (SaaS) – User-Ebene.

Cloud-Dienste haben natürlich Vor- und Nachteile. Zum einen müssen keine teuren Investitionen in eigene Infrastruktur vorgenommen werden, da die genutzten Services nach dem

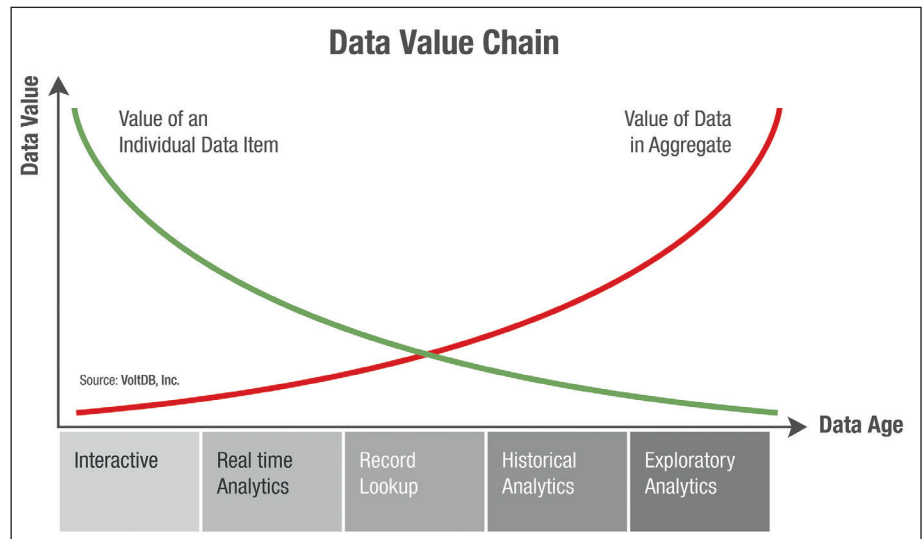


Abb. 2: Datenwertschöpfungskette [4]

Prinzip „Pay per use“ abgerechnet werden. Das erlaubt eine freie Skalierung je nach Anforderungen und Datenaufkommen. Ebenfalls werden hohe Verfügbarkeit und Redundanz gewährleistet und die Administration erleichtert. Auch eine weltweite Nutzung von Diensten kann einfach umgesetzt werden.

Zum anderen funktioniert Cloud Computing nicht ohne Internet und bei nicht ausreichender Bandbreite ist mit Verzögerungen zu rechnen. Die Größe des gewählten Cloud-Anbieters ist hier wesentlich. Das Thema „Datensicherheit und Datenschutz“ ist ausreichend zu berücksichtigen. Besonders bei länderübergreifender Datenverarbeitung ist auf unterschiedliche Regulative zu achten.

Big Data & Data Science

Big Data und Cloud Computing gehen häufig Hand in Hand. Die gute Skalier- und Verteilbarkeit von Cloud-Diensten eignet sich hervorragend für Big Data-Anwendungen. Denn eines der primären Ziele von Big Data ist, aus ver-

schiedenen Daten Muster zu erkennen. Alles dreht sich hier um Analyse, Prognose und datengetriebene Modellbildung. Im Besonderen darum, Tools für bessere und vor allem frühere Entscheidungsfindung bereitzustellen. Durch datengetriebene Analysen soll das System Eisenbahn noch wettbewerbsfähiger gemacht werden. Der Weg von zeitabhängiger zu zustandsbasierter Wartung kann durch Big Data und die „Data Scientists“ vollzogen werden.

Condition-based Maintenance, kurz CBM (was passiert?), wird vielfach in Unternehmen bereits angewendet. Durch den Einsatz von Sensoren wird der aktuelle Zustand an und in Teilen festgestellt (Condition Monitoring). Die Instandhaltung wird durchgeführt, sobald sie notwendig ist bzw. eine Verschlechterung bzw. Überschreitung von Grenzwerten durch einen oder mehrere Indikatoren angezeigt wird.

Predictive Maintenance (was wird passieren?) bringt CBM einen Schritt weiter. Statt nur die bestehende Situation zu analysieren, werden

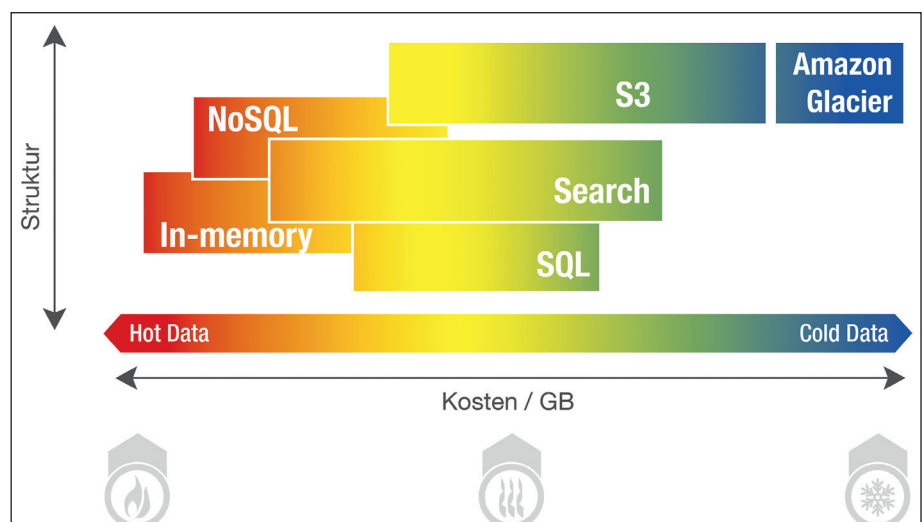


Abb. 3: Technologieauswahl für Datenmanagement auf Basis Amazon AWS, in Anlehnung [5]

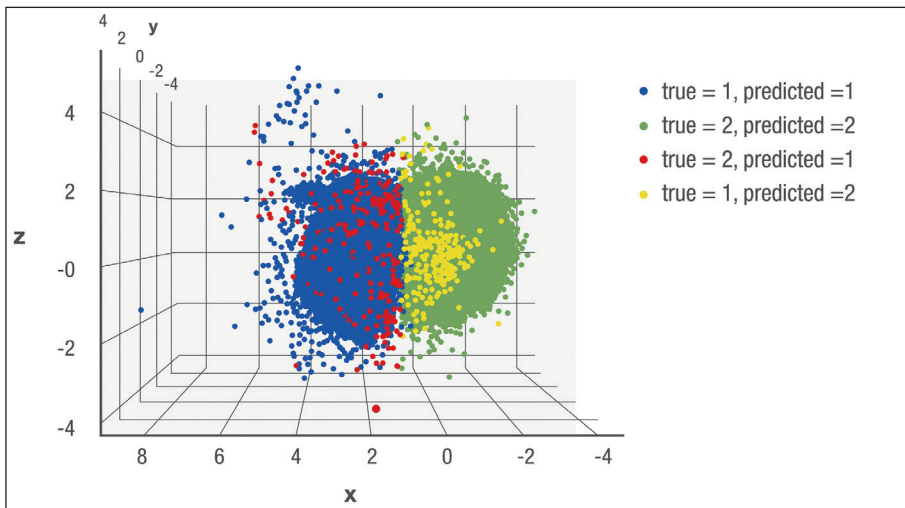


Abb. 4: Hybride Datenauswertung

mögliche künftige Ereignisse vorausgesagt. Durch verschiedene Ansätze wie mathematische Modellbildung in Kombination mit Machine Learning und neuronalen Netzen können Prognosen erstellt werden. Durch modellbasiertes und datengetriebenes Wissen entstehen hybride Methoden, um neues Wissen zu generieren. [6]

Abb. 4 zeigt eine hybride Datenauswertung (physikalische Modellbildung und Machine Learning kombiniert) einer Stopfmaschine der 09-Serie von Plasser & Theurer. Verschiedene Arbeitseinflüsse am Stopfaggregat wurden untersucht und so ein Prognosemodell zum Verschleißzustand erstellt. Die ausgewertete Datenmenge beträgt derzeit ca. 1,1 Terabyte

und steigt kontinuierlich. Blaue und grüne Punkte kennzeichnen richtige Prognosen zu auftretenden Verschleißzuständen im Vergleich mit echten Einsatzdaten. Die Prognosegenauigkeit des angewendeten Modells liegt derzeit bei ca. 90%. Ein auftretender Aggregatverschleiß kann somit wesentlich früher erkannt und die Einsätze dementsprechend geplant werden.

Prescriptive Maintenance (was sollte passieren?) ist die Zukunft. Anstatt zu analysieren, was passieren wird, werden direkt Handlungsanweisungen oder Entscheidungsoptionen vorgeschlagen. Es wird die Entscheidung und die Interaktion zwischen Entscheidungen, Einflüssen und deren Bezug zum Ausgang analy-

siert, um die optimale Vorgehensweise festzulegen. [7]

Resümee

Ob der vernetzte Kühlschrank bald unsere Einkäufe erledigt oder ein vernetzter Schrittzähler uns gesünder macht, bleibt abzuwarten. Daten und deren Analyse beeinflussen unser Leben wie auch den Gleisbau. Durch richtiges Datenmanagement lässt sich im Gleisbau ein erheblicher Mehrwert schaffen. Richtige Daten zur richtigen Zeit werden Gleisbaumaschinen- und Infrastrukturbetreibern helfen, Entscheidungen früher, effektiver und kostengünstiger zu treffen. ■

QUELLEN

- [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] <http://gotocon.com/aarhus-2012/thebigdatavaluecontinuum>
- [5] Eigene Darstellung in Anlehnung an 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] <http://www.searchenterprisesoftware.de/definition/Prescriptive-Analytics>



Bernhard Maier

Leiter P&T Connected
Plasser & Theurer Connected
Gesellschaft m.b.H., A-Hagenberg
office@ptconnected.com