

# Business Intelligence in KMU – Realisierung in einer heterogenen IT-Systemlandschaft

Sebastian Kollmar, B.Sc., Industrial Analytics Lab GmbH, Bochum, Deutschland

Prof. Dr.-Ing. Dipl.-Wirt. Ing. Andreas Merchiers und Prof. Dr. rer. nat. Henrik Blunck, Bochum University of Applied Sciences, Deutschland

## 1 Einleitung

### 1.1 Aktueller Stand von Business Intelligence in KMU's

Im Rahmen der Digitalisierung der klassischen Geschäftsprozesse vergrößert sich die zugängliche betriebliche Datenmenge mit jeder Transaktion und jedem Werktag. Moderne Business Intelligence (BI) nutzt dieses Datenvolumen, um das bisherige Reportwesen zu ersetzen und neue Zusammenhänge sichtbar zu machen. Von der operativen Überwachung einer Werkhalle bis zur abteilungsübergreifenden Kennzahlenanalyse kann BI zahlreiche Perspektiven eröffnen und Optimierungspotenzial aufzeigen. Was früher abhängig vom Bauchgefühl der Führungskräfte war, kann heute bereits für jedermann sichtbar auf's Papier oder den Bildschirm gebracht werden – sofort und auf Knopfdruck.

Obwohl kaum jemand im digitalen Zeitalter die Relevanz von Datenanalyse infrage stellt, wird das Potenzial von Business Intelligence bei mittelständischen deutschen Industriebetrieben noch immer bei weitem nicht vollständig ausgeschöpft. Wertvolle Unternehmensdaten entstehen ungeordnet auf getrennten Dateisystemen und werden daher niemals ausgewertet. Die Gründe dafür liegen nicht im fehlenden Bewusstsein der Unternehmerinnen und Unternehmer, sondern vielmehr in den Hindernissen der eigenen IT-Systemlandschaft. Bereits verbaute Software ist oft schon etwas älter, verfügt über keine geeigneten Datenschnittstellen oder Exportfunktionen und ist daher schwierig in eine BI-Infrastruktur zu integrieren. Ein Upgrade der zahlreichen Unterstützungssysteme ist kostspielig und mit Produktivitätsausfall verbunden. Und nur zu gern wird das Problem daher nach dem Prinzip „never change a running system“ nach hinten verschoben.

Für eine stabile Kennzahlensteuerung ist aber keineswegs eine Generalüberholung der IT-Infrastruktur notwendig. Auch wenn die gewünschten Bereiche, wie Buchhaltung, Personalzeiterfassung, Maschinendaten oder Logistik, eigene Software zur Dateneingabe nutzen, können die Anwendersysteme nacheinander individuell in eine gemeinsame Datenpipeline integriert werden. Fehlende Datenschnittstellen können durch Methoden wie „Robotic Process Automation“ (RPA) umgangen werden, indem Datensätze automatisiert über die Benutzeroberfläche exportiert werden. Dieser Beitrag zeigt daher eine praktikable Lösung zum Aufbau einer BI-Struktur auf, in der RPA genutzt wird, um Anwendersysteme ohne Datenschnittstellen zu integrieren.

Business Intelligence kann auch in kleinen und mittleren Unternehmen (KMU) eingeführt werden und zur Prozessoptimierung und Effizienzsteigerung dienen.

**KMU • Digitalisierung • Business Intelligence • Prozessoptimierung • Effizienz**

### 1.2 Automatisierung des Berichtswesens

In einer traditionellen Firmenstruktur, in der die Geschäftsführung monatliche Berichte als Entscheidungsgrundlage erwartet, sind Managerinnen und Manager für die Präsentation von Unternehmensdaten verantwortlich. In einem häufig manuellen Prozess sammeln, aggregieren und selektieren Führungskräfte die relevanten Informationen, welche dann für Managemententscheidungen genutzt werden sollen [1, S. 49]. Operative Kennzahlen werden in mühsamer Kleinstarbeit in Excel-Tabellen zusammengetragen, vielfach überschrieben und umformatiert, bevor sie dem Management in höheren Unternehmensebenen Wochen später präsentiert werden können. Unterstützende Softwaresysteme liefern zwar teils die gewünschten Rohdaten, diese müssen aber noch bereinigt und aggregiert werden. Die Gefahr menschlicher Fehler ist groß, Daten sind zum Zeitpunkt der Betrachtung der Auswertung bereits nicht mehr aktuell, aufgrund ungelöster technischer Probleme nicht verfügbar oder hinsichtlich der Entscheidungsfindung irrelevant [2, S. 46].

Computergestütztes Reporting hat das Ziel, diesen Prozess zu beschleunigen und zu systematisieren. Umsatzzahlen, Maschinenauslastungen und Lagerbestände können nicht nur monatlich, sondern täglich oder in Echtzeit und auf Knopfdruck zur Verfügung gestellt werden. In vielen Fällen lässt sich das interne Berichtswesen durch automatisierbare Prozesse vollständig digital ausführen. Ist die entsprechende Infrastruktur einmalig aufgebaut, lässt sie sich außerdem leicht skalieren, da die Beschaffung und Erzeugung von zusätzlichen Informationen lediglich etwas mehr Rechenleistung erfordert.

Durch die Fusionierung von Rohdaten verschiedener Unternehmensbereiche in Kombination mit explorativer Datenanalyse können außerdem Zusammenhänge sichtbar gemacht werden, welche weit über das traditionelle Reporting hinausgehen. Führungskräfte können die bisher benötigte Zeit zur Erstellung dieser Kennzahlen stattdessen in die Dateninterpretation stecken, um die Entscheidungsqualität zu erhöhen.

### 1.3 Entdeckung und Ausschöpfung verborgener Potenziale

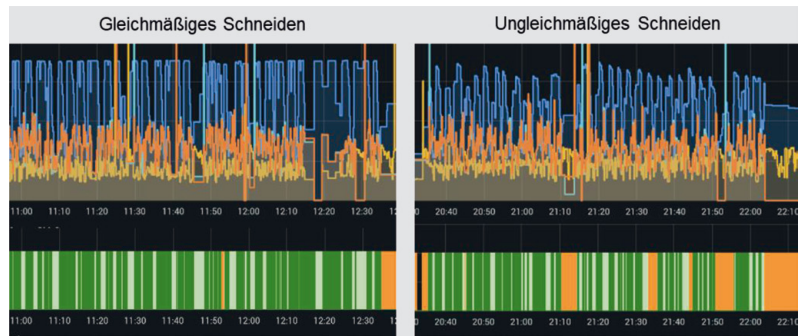
Computergestützte Datenauswertung kann nicht nur das klassische Reporting ersetzen, sondern auch neue Verbesserungspotenziale aufdecken, die vorher nicht oder nur schwer erkennbar waren.

Die schnelle Visualisierung historischer Daten kann zur frühzeitigen Erkennung von Trends, zum Beispiel in der Produktion, beitragen. Im Laufe einer Periode können daher schon frühzeitig Ausfälle, Kapazitätsspitzen oder Veränderungen in der Durchlaufzeit prognostiziert werden. Negative Trends können frühzeitig ermittelt und bekämpft werden. Natürlich hat eine Produktionsleitung mit jahrelanger Erfahrung ein gutes Gefühl dafür, wie viele Tonnen Material im Verlauf der Woche verarbeitet werden, aber noch besser ist es, wenn eine mitlaufende Kennzahl dieses Gefühl stützen kann. Dann ist eine Abweichung von der Norm frühzeitig zu erkennen und das Problem zu ermitteln und zu lösen. Gute BI-Systeme liefern dabei detaillierte Daten zu den Produktionsfaktoren, sodass zusätzlich zur Trendermittlung auch eine Ursachenanalyse möglich ist. Sinkt die produzierte Stückzahl einer Station, gibt z. B. eine Visualisierung der Materialeingangslager, der Arbeitszeit und der Laufzeit der Maschinen über diesen Zeitraum bereits Hinweise darauf, wo das Problem liegt.

Die stetige Beobachtung von Echtzeitdaten hat außerdem den Vorteil, dass akute Probleme zeitnah erkannt werden. Durch häufige Anwendung bekommen die Nutzerinnen und Nutzer Erfahrungswerte, sodass Abweichungen schnell identifiziert und die Ursachenanalyse eingeleitet werden können. Werden auch Daten aus angrenzenden Arbeitsbereichen zur Verfügung gestellt, kann die Ursachenanalyse sogar ausgeweitet werden und ein besseres Verständnis für die Zusammenhänge der eigenen Abteilungspersormance generieren. Da all diese Ergebnisse objektiv mit Zahlen abgebildet werden, können auch neue Mitarbeitende mit weniger Erfahrung wirksame Entscheidungen treffen.

Ein weiterer Vorteil einer Echtzeitkennzahlensteuerung ist ein objektiver Vorher-Nachher-Vergleich einer Maßnahme oder Änderung im Prozessablauf. Wie wirkt sich die Einführung einer viertelstündigen Frühstückspause auf die Produktion aus? Mit einer guten Datenvisualisierung kann die Auswirkung bereits nach wenigen Tagen quantifiziert werden. Eine kontinuierliche Optimierung über Monate oder Jahre hinweg kann langfristig einen sehr großen Unterschied ausmachen.

Zu guter Letzt ermöglicht eine gezielte Sammlung von Unternehmensdaten, Muster und Korrelationen in den Daten zu analysieren. Sogenanntes Data Mining beschreibt die Extraktion von Wissen aus großen Datenmengen [3]. Durch exploratives Data Mining können sonst verborgene Optimierungspotenziale entdeckt und genutzt werden. Einige Beispiele für potenzielle oder tatsächliche Anwendungen sind:



**Bild 1:** Unterschiedliche Maschinenbedienung einer Bergbaumaschine

- ▶ Identifikation des Anwenderverhaltens bei der Maschinenbedienung (**Bild 1**) zur Bestimmung von Best-Practices
- ▶ Qualitätskontrolle von Produktionsprozessen mit Statistical Process Control (SPC)
- ▶ Implizites Prozessmonitoring statt expliziter Überwachung
- ▶ Analyse von Kausalketten im Unternehmen
- ▶ Hypothesentests von Ablaufänderungen

Der Einführung von Business Intelligence lässt sich leider keine direkte Investitionsrendite zuordnen, da die Erkenntnisse vorher ungewiss sind. Ist die Infrastruktur allerdings einmal aufgebaut, können die Auswirkungen daraus resultierender Ablaufänderungen im Nachgang quantifiziert werden. Im Laufe der Zeit kann Business Intelligence auch bei produzierenden Unternehmen einen messbaren Anteil der Wertschöpfung ausmachen und klassischen IT-Abteilungen zu mehr Nutzen verhelfen.

### 1.4 Zielsetzung: Nutzerspezifische Business Intelligence für alle Führungsebenen

Ziel der Business Intelligence soll es sein, Wissen aus rohen externen und internen Unternehmensdaten zu gewinnen. Dieser Wissensgewinn kommt keinesfalls nur den oberen Führungskräften zugute, sondern kann für alle Managementebenen realisiert werden.

Die dargestellten Informationen unterscheiden sich nutzerspezifisch und bieten individuelle Entscheidungsgrundlagen. Jede Instanz bildet ihre eigene „Sicht“ auf den gemeinsamen Datenwürfel und arbeitet die für den jeweiligen Zuständigkeitsbereich benötigten Datensätze heraus (**Bild 2**). Die Dashboards der Führungskräfte unterscheiden sich damit hinsichtlich der Detailtiefe, des Geltungsbereichs und des Zwecks.

Historische Daten können Zusammenhänge erläutern, zur Trendanalyse dienen oder für Vorjahresvergleiche verwendet werden. Tagesaktuelle Daten dienen der kurzfristigen Überwachung oder unterstützen sogar die Personaleinteilung. Die Geschäftsführung erhält grobe unternehmensweite Informationen, während die Abteilungsleitungen spezifische Details zu ihrer Abteilung



**Bild 2:** OLAP Datenwürfel

erhalten. Dies können zum Beispiel die Maschinenauslastung der zugewiesenen Maschinen sein, die aktuelle Krankheitsquote, die Anzahl der Überstunden oder schlicht eine Liste der offenen Aufträge sein.

Auch eine Verteilung an Mitarbeitende ohne Führungsfunktion ist denkbar, zum Beispiel indem ein Bildschirm mit abteilungsspezifischen Kennzahlen innerhalb der öffentlichen Arbeitsumgebung platziert wird. Mitarbeitende haben damit auch die Möglichkeit, die eigene Performance gegenüber ihren Vorgesetzten zu rechtfertigen oder ihren eigenen Beitrag zur Abteilung zu beobachten.

## 2 Einführung von BI in KMU's

### 2.1 Ausgangssituation der IT-Systemlandschaft

Bei der Auswahl von Anwendersoftware tendieren KMU's häufiger zu branchenspezifischen Lösungen, während größere Firmen auch größere Softwareanbieter mit entsprechenden Funktionsumfängen auswählen. Branchenapplikationen spezialisieren sich auf ein Marktsegment, weshalb ihr Funktionsumfang eher in die Tiefe als in die Breite geht. Für Nischenanwendungen kann diese Software daher wesentlich besser als Standardsoftware auf die Anforderungen des Unternehmens abgestimmt werden. Zudem ist Branchensoftware günstiger und kurzfristiger verfügbar. In einem wachsenden Unternehmen kann eine akute EDV-Lücke schneller mit spezialisierter Software von kleineren Anbietern gefüllt werden.

Im Laufe der Zeit entsteht durch diese Vorgehensweise eine heterogene Softwarelandschaft vieler verschiedener Anbieter. Beispielsweise kann sich das ERP-System auf branchenspezifische Auftragsflusskontrolle konzentrieren, während die Buchhaltung von einem separaten Buchhaltungsprogramm und die Arbeitszeiterfassung von einer Personalsoftware übernommen wird. Im operativen Bereich werden Anbieter für Maschinendatenerfassung, Materialflusskontrolle und Energiemanagement genutzt.

Sicherlich sind diese Anwendungen durch ihre Detailtiefe effektiv, die Vielfalt unterschiedlicher Software

verkompliziert allerdings den Zugriff auf alle Unternehmensinformationen für eine umfangreiche Business Intelligence.

Bei der Beschaffung von neuer Software muss daher darauf geachtet werden, dass eine Datenschnittstelle passend zur BI-Infrastruktur oder zumindest eine automatisierbare Exportfunktion in ein gemeinsames Datenformat vorliegt. Steht bei einem System von Haus aus keine Schnittstelle zur Verfügung, kann diese möglicherweise nachträglich hinzugefügt werden. Dies erfordert Programmierarbeit, die kostenaufwendig sein kann. Aber auch ohne direkte Datenschnittstelle ist eine Integration möglich, wie später noch an einem Beispiel erläutert wird.

### 2.2 Frontend: Die richtige Wahl des BI-Tools

Aktuell gibt es eine ganze Reihe namhafter Anbieter zum Thema Business Intelligence, wie z. B. Microsoft Power BI, Tableau, Qlik Sense, SAP Analytics oder Grafana. Die Anbieter unterscheiden sich hinsichtlich der Preisgestaltung und der Funktionalitäten, weshalb vor dem Kauf eine genaue Bedarfsanalyse durchgeführt werden sollte. Folgende Kriterien können bei der Auswahl des Visualisierungstools helfen:

- ▶ Preisgestaltung
- ▶ Vielfalt der Datenschnittstellen
- ▶ Angebotener Support
- ▶ Plattformunabhängigkeit
- ▶ Echtzeitfähigkeit
- ▶ Nutzerfreundlichkeit
- ▶ Konfigurierbarkeit

Natürlich können auch die bereits bestehende Infrastruktur, die eigene Datenlandschaft oder spezielle Nutzungsvorlieben mit in die Auswahlkriterien einfließen.

Wird das ERP-System von SAP verwendet, muss deswegen nicht zwingend auf SAP Analytics als BI-Tool zurückgegriffen werden. Einheitliche Anbieter vereinfachen zwar die Integration, schließen aber andere Anbieter keinesfalls aus. Spezielle Daten erfordern auch spezielle Anbieter; Grafana eignet sich zum Beispiel hervorragend zur Visualisierung zeitbasierter Daten, beispielsweise ausgelesener Sensordaten einer speicherprogrammierbaren Steuerung (SPS). MS Power BI bietet sich zur Integration mit SQL-Servern an und ist in der Basisversion sogar in den meisten Office-Paketen enthalten.

Finden sich im Unternehmen einige begeisterte Programmierende, kann eine BI-Darstellung natürlich auch hauseigen entwickelt werden. Die bekannte Plotly-Bibliothek gibt zum Beispiel eine große Auswahl an Darstellungsmöglichkeiten in wenigen Codezeilen und erlaubt damit vollständige Unabhängigkeit von Fremdanbietern, mit der Möglichkeit Machine Learning und kompliziertere Statistikmodelle selbst zu konstruieren.

Die Visualisierungssoftware stellt den sogenannten Frontenteil des Systems dar, welcher von den Nutzerinnen und Nutzern bedient wird. Um Daten anzeigen zu können, müssen diese aber erst einmal im Hintergrund prozessiert werden. Ein einfacher Zugriff auf die Daten der Rohsysteme reicht nicht aus, um interpretierbare Auswertungen zu erstellen. Um zum Beispiel die Entwicklung der Anzahl der hergestellten Bauteile pro Schicht einer Maschine abzubilden, werden Rohdaten aus dem ERP-System mit Daten aus der Personalzeiterfassung kombiniert, nach Maschine und Zeitraum gefiltert, Wochenenden und Feiertage herausgerechnet sowie Pausen und Maschinenausfälle berücksichtigt. Diese Berechnungen werden im sogenannten Backend durchgeführt und abgespeichert.

Das Frontend greift zum Beispiel über das Internet oder das Unternehmensnetzwerk auf die Daten des Backends zu und bleibt damit austauschbar.

### 2.3 Backend: Datenpipeline und Data Warehouse

Das Backend besteht in diesem Falle aus einer Datenpipeline und dem Data Warehouse.

Im Backend werden die verschiedenen heterogenen Daten im Unternehmen zu einem kompakten Datenwürfel zusammengefasst. Unterschiedliche Datenformate, Zugriffsmechanismen und Speichertechniken verhindern, dass Mitarbeiterinnen und Mitarbeiter wissen, wo zu welchem Zeitpunkt Daten verfügbar sind. Aus gesammelten Daten sollen aber neue Erkenntnisse generiert und die Informationen aus verschiedenen Quellen in einer für die Entscheidungsfindung optimierten Datenbank, einem Data Warehouse, abgelegt werden. [4, S. 1–2]

Das Data Warehouse selbst beschreibt eine physische Datenbank, welche im Gegensatz zu den Datenbanken der Rohdaten nicht die Daten selbst, sondern Kennzahlen und Zusammenfassungen der Rohdaten enthält. Diese werden parallel zu den Rohdaten gehalten und periodisch aktualisiert (**Bild 3**). Damit stellt das Data Warehouse eine Momentaufnahme dar, während die operativen Datenbanken die Daten weiterhin in Echtzeit speichern. Die Dauer eines Aktualisierungszyklus hängt vom Bedarf ab und kann z. B. stündlich, täglich oder wöchentlich erfolgen. [4, S. 5–2]

Die Informationen aus den Rohdaten werden innerhalb einer „Datenpipeline“ von den Ursprungssystemen in das Data Warehouse transportiert. Die Pipeline stellt sicher, dass Informationen

- ▶ Systematisch
- ▶ Regelmäßig
- ▶ Vollautomatisch und
- ▶ Zuverlässig

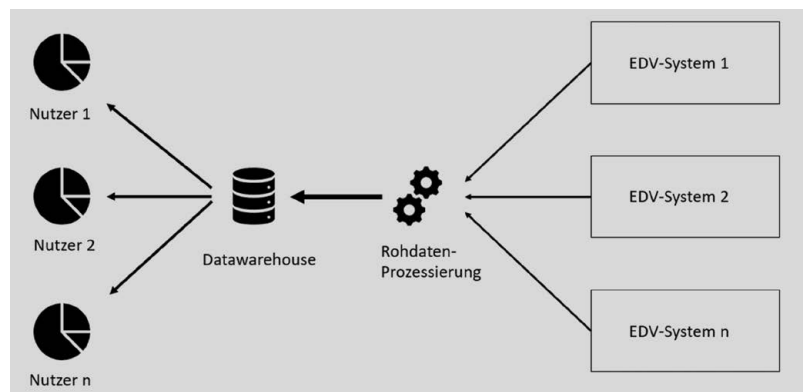
im Data Warehouse ankommen. Die Pipeline besteht aus drei Arbeitsschritten, die hinlänglich als „ETL-Pro-

zess“ (Extraction, Transformation und Load) bekannt sind [4, S. 8]:

- ▶ Export aus den originalen Systemen
- ▶ Datenprozessierung und Berechnung der Kennzahlen
- ▶ Import in das Data Warehouse

Die Umsetzung hängt von der bestehenden Infrastruktur des Unternehmens und den Quellsystemen ab. Idealerweise befinden sich diese im gleichen Netzwerk und der Datenzugriff lässt sich mit einer integrierten Schnittstelle realisieren. In diesem Fall reicht es bereits, einen zusätzlichen Rechner (physisch oder als virtuelle Maschine (VM)) mit Netzwerkzugriff für die Transformation bereitzustellen. Dieser Rechner empfängt die Rohdaten über das Netzwerk, verarbeitet sie und schickt die Auswertungen weiter an das Data Warehouse. Dabei kann sich das Data Warehouse entweder auf einem weiteren oder auf dem gleichen System befinden. Für größere Datenvolumen kann sich eine Lastverteilung auf mehrere Systeme anbieten.

Die Datentransformation übernimmt ein Programm, welches die benötigten Informationen aus den Rohdaten berechnet. Jede Kennzahl und Metrik muss vorher erst aus den Rohdaten berechnet werden. Dazu werden die Rohdaten gefiltert, selektiert, umformatiert und schließlich als Berechnungsgrundlage verwendet. Dies ist die „eigentliche Arbeit“, da sich bei stetig wechselnden Bedürfnissen der Nutzerinnen und Nutzer die gewünschten Kennzahlen und damit auch die Berechnungen immer wieder verändern werden. Die Berechnungen müssen daher für jeden nachvollziehbar und nachträglich anpassbar sein. Aus diesem Grund bietet es sich an, die notwendigen Anwendungen unternehmensintern zu erstellen oder zu beauftragen. Es wird auch vorgefertigte Software für ETL-Prozesse angeboten, deren Nutzung allerdings häufig mit Limitationen und geringerer Transparenz verbunden ist. Eigenprogrammierung ist mit Initialaufwand verbunden, kann aber sehr gut mit der individuellen Detailtiefe der Business Intelligence des Unternehmens mitwachsen. Im Jahr 2020 haben sich dafür die Programmiersprachen Python, Java und Ableger der C-Familie bewährt.



**Bild 3:** Datenpipeline und Data Warehouse

## 2.4 Anforderungen an IT-Systeme zur Einbindung in eine BI-Infrastruktur

Um eine Applikation in eine BI-Infrastruktur einzubinden, muss der Datentransport von der Applikation entlang der Datenpipeline zum Data Warehouse sichergestellt werden. Eine zentrale Herausforderung ist die Entnahme der Daten aus dem Quellsystem.

Der einfachste Weg ist eine integrierte On-Demand-Datenschnittstelle im Quellsystem, welche sich direkt mit der Software der Datenprozessierung verknüpfen lässt. Alternativ reicht auch eine Schnittstelle zu einer dritten Datenablagestruktur aus, welche ihrerseits dann eine Verbindung zur Datenprozessierung bereitstellt.

Ist keine direkte Schnittstelle vorhanden, bieten einige Systeme einen regelmäßigen Datenexport in einen Zielort an (zum Beispiel täglich oder wöchentlich). Von dort aus können die Daten dann abgeholt und final durch ein Drittprogramm zur Datenprozessierung gebracht werden. Ist auch dies nicht möglich, muss zumindest eine manuelle Exportfunktion über die grafische Benutzeroberfläche gegeben sein.

Moderne IT-Systeme verfügen meist von Haus aus über integrierte Datenschnittstellen oder Exportfunktionen, die sich automatisieren lassen. Wurde bei der Beschaffung der Software nicht auf derartige Funktionen geachtet, erschwert dies die Anbindung an die Datenpipeline. Die sauberste Lösung ist sicherlich ein Upgrade der Software, ist dies aber nicht vorgesehen, muss die grafische Benutzeroberfläche genutzt werden.

Bei handelsüblichen Datenverwaltungssystemen kann der aktuelle Datenbestand in der Regel mit wenigen Klicks heruntergeladen werden. Die Daten werden meistens im Textformat lokal abgespeichert und können nun weiterverarbeitet werden. Mithilfe von Robotic Process Automation (RPA) kann diese grafische Interaktion automatisiert und damit das Anwendersystem übergangsweise in die Datenpipeline integriert werden. Diese Einbindung ist nicht ideal, da bei einer grafischen Überholung des Anwendersystems auch die Integration angepasst werden muss. Sie ist aber schnell

umsetzbar und effektiv, sodass ein BI-Projekt trotz ungünstiger Startbedingungen begonnen werden kann.

## 2.5 Integration schnittstellenloser Systeme mit Robotic Process Automation (RPA)

Robotic Process Automation (deutsch: Robotergesteuerte Prozessautomatisierung) beschreibt die Prozessautomatisierung durch Softwareroboter (Bots). Dabei werden nicht nur technische Systemschnittstellen verwendet, sondern auch die grafische Benutzeroberfläche, vergleichbar mit der Durchführung durch einen Menschen.

Vollautomatische Bots laufen auf Dauerschleife oder werden durch Events oder Zeitpläne aktiv und arbeiten eine vorgegebene Liste an Schritten ab. Dadurch können repetitive und monotone Arbeiten am Computer automatisiert werden. Ein Beispiel dafür ist: Ein RPA-Skript wird durch das Eintreffen einer E-Mail aus der Buchhaltung aktiviert. Steht im Betreff „Monatsabschluss“, so entnimmt das Skript automatisch die Excel-Datei aus dem Anhang, formatiert einige Spalten um, berechnet die Summe der Spalte H, trägt diese in ein Feld ein und legt die Datei in einem Ordner auf dem Netzwerk ab.

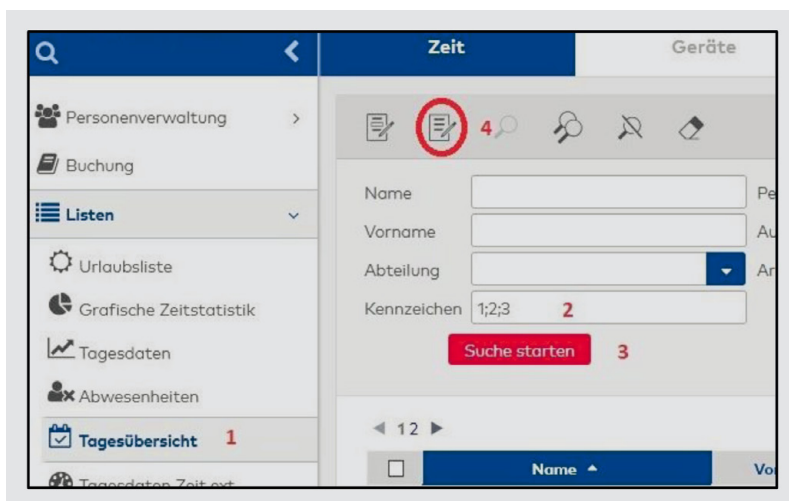
Halbautomatische Bots interagieren mit dem Nutzerinnen und Nutzern des Computers und automatisieren Teile ihrer Arbeit. Beispiele dafür können einfache Microsoft-Excel-Makros oder komplexere Bots sein, bei deren der Nutzung Eingaben zu tätigen sind.

Solche RPA-Skripte werden Schritt für Schritt vorprogrammiert und treffen regelbasierte Entscheidungen. In Spezialfällen können auch selbstlernende Roboter oder automatisiertes Auslesen von Texten („Screen Scraping“) zum Einsatz kommen.

Im Falle der Systemintegration übernimmt ein solches Skript den regelmäßigen Datenexport innerhalb der Datenpipeline. Mit RPA kann daher auch Software in eine Datenpipeline integriert werden, welche lediglich einen Datenexport über eine grafische Oberfläche zur Verfügung stellt (**Bild 4**).

In der Praxis bietet es sich für diese Methode an, auf virtuelle Maschinen zurückzugreifen, damit kein zusätzlicher Rechner mit Eingabegeräten benötigt wird. Auf der virtuellen Maschine läuft die zu exportierende Software und die RPA-Software, welche zu festgelegten Zeiten den Export automatisch durchführt und die Daten im Netzwerk ablegt. Das RPA-Skript wird z. B. täglich zur festgelegten Uhrzeit aktiviert, loggt sich in das Anwendersystem ein, navigiert selbstständig über die grafischen Benutzeroberflächen und führt den Export durch. Eine Überwachung durch Personal ist bei statischen Programmen in der Regel nicht nötig.

Der Vorteil von RPA-Skripten ist die Flexibilität in der Anwendung und die Skalierbarkeit. Diese Technik lässt sich im Grunde bei allen grafisch bedienbaren Applikationen anwenden und ist schneller als manuelle Bedienung durch eine Nutzerin oder einen Nutzer. Probleme entstehen allerdings dann, wenn sich die zu bedie-



**Bild 4:** Grafische Oberfläche zum Datenexport

nende Software verändert, z. B. durch ein Versionsupdate oder die Veränderung der Bildschirmauflösung. Auch andere Programme, die auf der virtuellen Maschine laufen, können ein RPA-Skript bei der Ausführung behindern.

### 3 Anwendung am Beispiel der Reese-Gruppe

#### 3.1 Implementiertes Kennzahlensystem

Die Firma Reese ist seit über 70 Jahren im Bereich Wärmebehandlung mit insgesamt vier Standorten in Deutschland aktiv. Das zunächst im größten Standort Bochum implementierte Kennzahlensystem hat das Ziel, bislang monatlich erstellte Kennzahlen automatisch und auf Knopfdruck bereitzustellen sowie eine Basis für digitales Shopfloor-Management bzw. ganzheitliches Management der Herstellung zu schaffen.

Alle Zahlen werden aus bereits vorhandenen IT-Systemen entnommen und miteinander kombiniert, um daraus entscheidungsrelevante Informationen zu gewinnen. Die Berechnung wird täglich neu durchgeführt und ergänzt neue Daten in der Datenbank. Ferner werden wöchentlich die letzten sechs Monate der Datenbank für den Fall aktualisiert, dass ältere Daten nachträglich ergänzt wurden. Nutzerinnen und Nutzer können ihre Visualisierungssoftware jederzeit auf Knopfdruck aktualisieren und bekommen tagesaktuelle Daten.

Alle erhalten dabei Zugriff auf die für sie relevanten Datensätze. Die Geschäftsführung erhält Kennzahlen, die die allgemeine Unternehmenslage beschreiben, in diesem Fall zum Beispiel Umsätze und Auftragseingang in Gewicht oder Stückzahl – jeweils aufgeschlüsselt nach Abteilung und mit Vorjahresvergleich. Eine Abteilungsleitung erhält detailliertere Informationen zu ihrer Abteilung, z. B. eine Liste der offenen Aufträge und der Auslastung ihrer Maschinen sowie die aktuelle Krankheitsquote.

Die Firma Reese hat die Software Microsoft Power BI als Visualisierungssoftware ausgewählt. Diese Software ist in der Basisversion bereits im Office-Paket enthalten und verträgt sich gut mit der vorhandenen IT-Infrastruktur, da primär Microsoft Betriebssysteme und Microsoft Server verwendet werden. Jeder Nutzer besitzt eine lokale Installation von Power BI auf dem Rechner und kann sich seine Datensätze über das interne Netzwerk aktualisieren.

#### 3.2 Verwendete Datenquellen

Das eben genannte Kennzahlensystem erfordert Informationen aus verschiedenen Unternehmensbereichen, die zum Teil jeweils über eine eigene Software verarbeitet wurden.

Den größten Anteil an der Datenbasis macht das ERP-System aus. In diesem Fall wird die Software AMS verwendet, welche die Datenbank tagesaktuell mit Auftragsdaten, Stückzahlen, Gewichten und Preisen versorgt. Über dieses System werden auch die Rechnungen

erstellt, wodurch zur Umsatzermittlung kein Eingriff in die separate Buchhaltungssoftware notwendig ist.

Für mitarbeiterbezogene Kennzahlen werden Arbeitszeiten aus dem Personalmanagement-System Matrix exportiert. In Kombination mit Auftragsdaten können dadurch Performance-Kennzahlen im Verhältnis zur Abteilungsgröße oder den anwesenden Mitarbeitenden pro Tag adjustiert werden.

AMS und Matrix werden beide mit RPA über die grafische Benutzeroberfläche in die Datenpipeline integriert. In beiden Fällen wäre ein Software-Upgrade mit einer besseren Datenschnittstelle möglich, das aber bei AMS mit großem Aufwand und bei Matrix mit steigenden Lizenzkosten verbunden wäre. Aus diesem Grund hat sich stattdessen die Programmierung von RPA-Bots als günstigere Variante angeboten.

Auf der Kostenseite wird außerdem der Stromverbrauch der Maschinen erfasst und in die Datenbasis integriert. Ziel ist es, die Maschinen einzeln hinsichtlich der Profitabilität untersuchen zu können. Dazu werden auch die einzelnen Maschinendaten mitgeschrieben und ausgewertet. Über einen Feldbus-Adapter werden die SPS im Betrieb ausgelesen und die Daten in Echtzeit in eine Datenbank geschrieben. Ein Algorithmus bestimmt über maschinenspezifische Sensoren und Aktoren nachträglich den Zustand der Maschine und fasst diesen auf Tagesbasis zusammen. Dadurch können die Daten aus dem ERP-System mit echten Maschinendaten überlagert werden und vermitteln ein realistischeres Bild von der Auslastung jeder Maschine.

#### 3.3 Aufbau einer effektiven Business Intelligence Infrastruktur

Technisch wurde die Datenpipeline über eine dafür eingerichtete virtuelle Maschine (VM) aufgebaut (Bild 5). Dadurch musste keine Hardware beschafft werden, und der Wartungsaufwand hält sich in Grenzen. Diese virtuelle Maschine ist mit dem Unternehmensnetzwerk verbunden und stellt einen gemeinsamen Zielordner für die Rohdaten, aktiviert die RPA-Skripte, führt die Berechnungen durch und dient als Standort für das Data Warehouse.

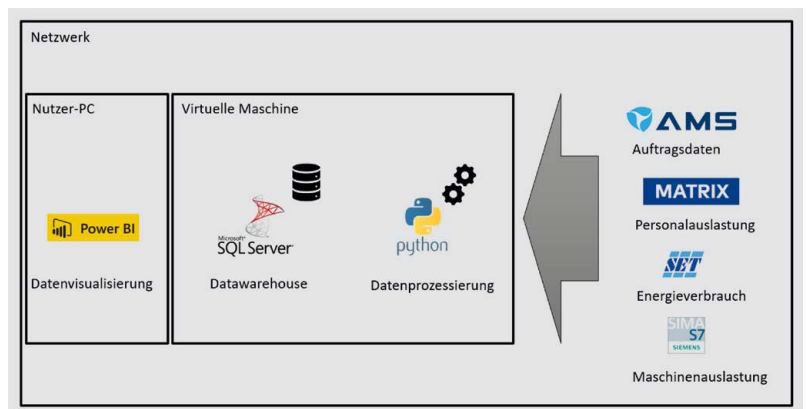


Bild 5: Beispiel einer BI-Infrastruktur

Die RPA-Skripte wurden mit der Programmiersprache Python selbst entwickelt. Über eine Schnittstelle zum Windows-Betriebssystem der VM werden die Nutzereingaben wie Mausklicks und Tastatureingaben simuliert, um die Rohdaten in einem festgelegten Zielordner abzuspeichern. Sollte sich die Software bei einem Versionsupgrade verändern, kann die Python-Software schnell angepasst werden.

Auch die Verarbeitung der Rohdaten wurde in Python umgesetzt und erfolgt auf der virtuellen Maschine. Die Rohdaten werden aus dem Zielordner entnommen und nach einem festgelegten Prozess bearbeitet, ausgewertet und schlussendlich im Data Warehouse gespeichert.

Das Data Warehouse wurde in einer Microsoft SQL-Server Datenbank umgesetzt, welche passend zum Betriebssystem und der Visualisierungssoftware ausgewählt wurde. Die Daten werden in einem Sternschema abgespeichert und können von MS Power BI direkt über das Netzwerk abgefragt werden.

#### 4 Fazit

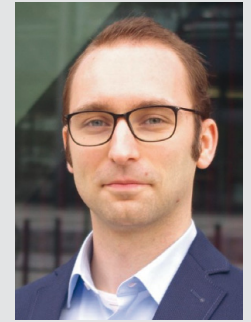
Business Intelligence ist auch für mittelständische Unternehmen umsetzbar und höchst relevant. Eine funktionsfähige BI-Infrastruktur kann in den meisten Fällen auch auf älteren Systemen aufgesetzt werden und passt sich flexibel an die Gegebenheiten in der Firma an. Ist eine solche Struktur einmal aufgebaut, kann die Detailtiefe außerdem beliebig und kostengünstig hochskaliert werden. Aus einer anfänglichen Umsatzvisualisierung kann so im Laufe der Zeit ein mächtiges Werkzeug zur umfassenden Wissensgenerierung für die komplette Firma entstehen. Auswirkungen von Änderungen im Betriebsablauf können direkt gemessen werden. Die Messergebnisse unterstützen daher den kontinuierlichen Verbesserungsprozess.

#### 5 Quellenverzeichnisse

- [1] Gluchowski, P.; Gabriel, R.; Dittmar, C. (2008): Management Support Systeme und Business Intelligence: Computergestützte Informationssysteme für Fach- und Führungskräfte, 2. Aufl., Berlin Heidelberg, 2008
- [2] Bullinger, H.-J.; Niemeier, J.; Koll, P. (1993): Führungsinformationssysteme (FIS: )Einführungskonzepte und Entwicklungspotentiale. In: Behme, W., u. a. (Hrsg.), Führungsinformationssysteme: Neue Entwicklungstendenzen im EDV-gestützten Berichtswesen, Wiesbaden, 1993, S. 44–62
- [3] Association for Computing Machinery (ACM) (2006): Data Mining Curriculum, 2006, 14.11.2020
- [4] Farkisch, K. (2011): Data-Warehouse-Systeme kompakt: Aufbau, Architektur, Grundfunktionen, Xpert.press, Berlin, Heidelberg, 2011

#### Sebastian Kollmar, B.Sc.

hat ein Studium zum Bachelor of Science der Wirtschaftsinformatik an der Hochschule Bochum absolviert. Aktuell ist er als Projektleiter im IT-Consulting für die Firma Industrial Analytics Lab GmbH, Bochum, tätig und berät Unternehmen im industriellen Umfeld. Sein interdisziplinäres Wissen hilft ihm dabei, moderne Methoden der Informatik betriebswirtschaftlichem Nutzen zuzuführen und Industriefirmen im digitalen Zeitalter bei der Erhebung, Archivierung und Auswertung ihrer betrieblichen Daten zu unterstützen.  
**Kontakt:** kollmar@ial.ruhr



#### Prof. Dr.-Ing. Dipl.-Wirt. Ing. Andreas Merchiers

hat Maschinenbau und zusätzlich Wirtschaftswissenschaften studiert. Er promovierte an der RWTH Aachen zum Dr.-Ing. und ist Professor an der Hochschule Bochum. Er lehrt und forscht zu Produktionsmanagement / Industrie 4.0 im Maschinen- und Bergbau sowie Technischer Investitionsplanung.  
**Kontakt:** andreas.merchiers@hs-bochum.de



#### Prof. Dr. rer. nat. Henrik Blunck

ist Professor für praktische Informatik an der Hochschule Bochum. Er ist Diplommathematiker und hat an der WWU Münster in Informatik promoviert. Seine Forschungsgebiete sind Data Science und Big Data Analytics, insbesondere in den Domänen Mobility Data, Context-aware Consumer Software und Industrie-4.0-Anwendungen.  
**Kontakt:** henrik.blunck@hs-bochum.de

