



Projekt: DaFuER

Anwendung der Datenfusion bei der Erfassung und Speicherung betrieblicher Rückmeldedaten

Steigerung der Datenqualität betrieblicher Rückmeldedaten durch Methoden der Datenfusion

Dieser Artikel gibt einen einführenden Überblick über die Ziele des Forschungsprojekts ‚DaFuER‘ und das Vorgehen in demselben. Ziel des Forschungsprojekts ist es, Methoden der Datenfusion zur Steigerung der Datenqualität im Kontext betrieblicher Rückmeldedaten anzuwenden. Hierzu soll im Vorhaben die Frage beantwortet werden, wie durch Ansätze der Datenfusion eine hinreichende Qualität von Rückmeldedaten in der Produktion kosteneffizient sichergestellt werden kann. Das IGF-Vorhaben 20579 N der Forschungsvereinigung FIR e. V. an der RWTH Aachen, Campus-Boulevard 55, 52074 Aachen wird über die AiF im Rahmen des Programms zur Förderung der industriellen Gemeinschaftsforschung (IGF) vom Bundesministerium für Wirtschaft und Energie (BMWi) aufgrund eines Beschlusses des Deutschen Bundestages gefördert.

Durch die zunehmende Globalisierung stehen produzierende Unternehmen vor der Herausforderung, hochindividuelle Produkte zu niedrigen Kosten effizient herzustellen. Besonders in Hochlohnländern, in denen hauptsächlich hochwertige Produktionsgüter hergestellt werden, muss dazu eine immer flexiblere Fertigung ermöglicht werden.¹ Dies wird erst durch Fortschritte in digitalen Technologien umsetzbar: Durch eine Darstellung des realen Produktionsfortschritts, nahezu in Echtzeit, gelingt es, schnell auf Änderungen des Produktionsablaufs zu reagieren. Damit solche Entscheidungen jederzeit richtig und schnell getroffen werden können, ist eine solide Datengrundlage unabdingbar.²

Die Basis dafür bilden betriebliche Rückmeldedaten. Diese beinhalten unter anderem logistische Daten, Personalzeitdaten, Maschinendaten, Betriebsdaten, Qualitätsdaten, aber auch Leistungsdaten.³ In der Praxis reicht die Art der Datenerfassung von händischen Eintragungen über gefertigte Mengen und benötigte Zeiten in das IT-System bis

hin zu einer Anbindung an vorhandene Maschinen und Transportsysteme, die Mengen mit einem Zeitstempel versehen in das IT-System übertragen.⁴

Um gute Entscheidungen im Rahmen der Produktionssteuerung sowie Kalkulationen des operativen Produktionscontrollings treffen zu können, sind vor allem eine hohe Datenkonsistenz und deren Widerspruchsfreiheit notwendig.⁵ Des Weiteren bildet eine verlässliche Datenbasis die Grundlage für Prozessoptimierung und die anschließende Verarbeitung der Daten im Rahmen von Big-Data-Analysen. Da die Datenaufbereitung für spezifische Analysen bis zu 80 Prozent des Arbeitsaufwands ausmacht, kann eine korrekte und präzise

Aufnahme betrieblicher Rückmeldedaten Kosten sparen und schnellere Prozessverbesserungen ermöglichen. Der direkte Einfluss mangelhafter Daten lässt sich auch anhand des Unternehmensumsatzes bemessen. So wurde bereits 1998 in einer Studie von REDMAN nachgewiesen, dass durch mangelhafte Datenqualität bis zu 12 Prozent weniger Unternehmensumsätze erzielt werden konnten.⁶

Entgegen der Bedeutung für den wirtschaftlichen Erfolg eines Unternehmens wird die Datenqualität bei KMU zurzeit oft als unzureichend eingeschätzt.⁷

Zudem haben Erfahrungen aus vorherigen Forschungsprojekten des FIR (iProd, bigPro)

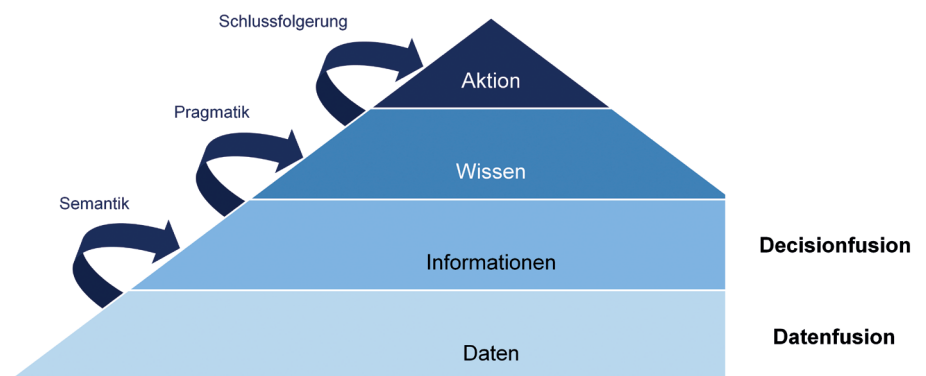


Bild 1: Verortung der Daten- und Decisionfusion in der Wissenspyramide (eigene Darstellung i. A. a. WOLF ET AL. 1999, S. 748)

¹ S. NIEHUES ET AL. 2017, S. 169
² S. MÜLLER 2000, S. 14; NAYAR 1993, S. 51 – 58
³ S. VDMA 2009, S. 3
⁴ S. KLETTI 2015, S. 19 – 21
⁵ S. BAUER 2003, S. 15
⁶ S. REDMAN 1998, S. 82
⁷ S. SPATH 2013, 92 ff.

zeigt, dass reale Rückmeldeprozesse eine exakte Rückmeldung oft verhindern. Nach APEL ET AL. sind Gründe hierfür ein mangelndes Problembewusstsein, mangelhaftes Design der Eingabemasken, die Heterogenität der Datenquellen sowie fehlende Ressourcen für die Erfassung der Daten. Diese ungünstige Prozessgestaltung führt zu einer Verzerrung der Rückmeldedaten und daher zu einer schlechten Datenqualität.⁸

Neben dem Einsatz zusätzlicher Sensorik stellt die Verknüpfung vorhandener Datenquellen eine kostengünstigere Alternative zur Steigerung der Datenqualität dar. Damit eignet sich diese insbesondere zur Implementierung bei KMU. Das Forschungsprojekt ‚DaFuER‘ zielt zu diesem Zweck auf eine Steigerung der Datenqualität von Rückmeldedaten durch systemorientierte, technische Ansätze ab. Hierbei werden insbesondere die relevanten Rückmeldedaten für die Produktionssteuerung und das operative Produktionscontrolling fokussiert.

Bei der Betrachtung der Datenqualitätssteigerung sollen insbesondere Methoden der Datenfusion und der Decisionfusion

eingesetzt werden. Unter dem Begriff der Datenfusion wird die „Zusammenführung mehrerer verschiedener Datensätze, die alle dasselbe Objekt einer realen Welt beschreiben“⁹, verstanden. Dabei folgt auf eine Prüfung der Schemen der einzelnen Datensätze (beispielsweise der Spalteneinträge) das Schema-Mapping, eine Duplikatsprüfung. Zuletzt werden auftretende Datenkonflikte bereinigt. Insbesondere im letzten Schritt lassen sich nach BLEIHOLDER U. NEUMANN verschiedene Methoden beschreiben, welche sich in ihrer Komplexität bei der Umsetzung unterscheiden. Hierzu zählen beispielsweise einfache Ansätze wie „Take the Information“, der einen vorhandenen Wert einem leeren Feld vorzieht, oder aber auch ein Vergleich mit großen Datenbanken („Trust your Friends“) sowie der Abgleich mit häufig auftretenden Werten („Cry with the wolves“).¹⁰

In Analogie zur Wissenspyramide wird bei der Vereinigung von Informationen (verarbeiteten Daten und Mustern) von Decisionfusion gesprochen (s. Bild 1, S. 46). Ziel ist es hierbei, eine Schlussfolgerung aus den verschiedenen vorhandenen Informationen abzuleiten, wobei aufgrund von Datenqualitätsmängeln auch

widersprüchliche Informationen vorliegen können.¹¹

In der Praxis lassen sich diverse Anwendungsfälle im Bereich der Produktionssteuerung und des operativen Produktionscontrollings finden. So ist beispielsweise eine häufig auftretende Problematik die genaue Bestimmung der Durchlaufzeiten und Rüstzeiten. Dabei werden Zeiten einzelner Bearbeitungsschritte oder möglicher Störungen oft nicht durch geeignete Rückmeldeprozesse oder Rückmeldepunkte erfasst. Dies führt dazu, dass die Datenqualität in Bezug auf verschiedene Beschreibungsdimensionen nicht korrekt ist.

Bei der Bestimmung der Datenqualität werden nach WANG U. STRONG vier Dimensionen und entsprechende Kriterien unterschieden. Die Dimension der intrinsischen Datenqualität untersucht den Inhalt eines Datensatzes anhand der Kriterien der *Glaubwürdigkeit*, der *Genauigkeit*, der

⁸ S. APEL ET AL. 2015, S. 37

⁹ LEIHOLDER U. NAUMANN 2011, S. 59

¹⁰ S. BLEIHOLDER U. NAUMANN 2006, S. 2

¹¹ S. CASTANEDO 2013, S. 15

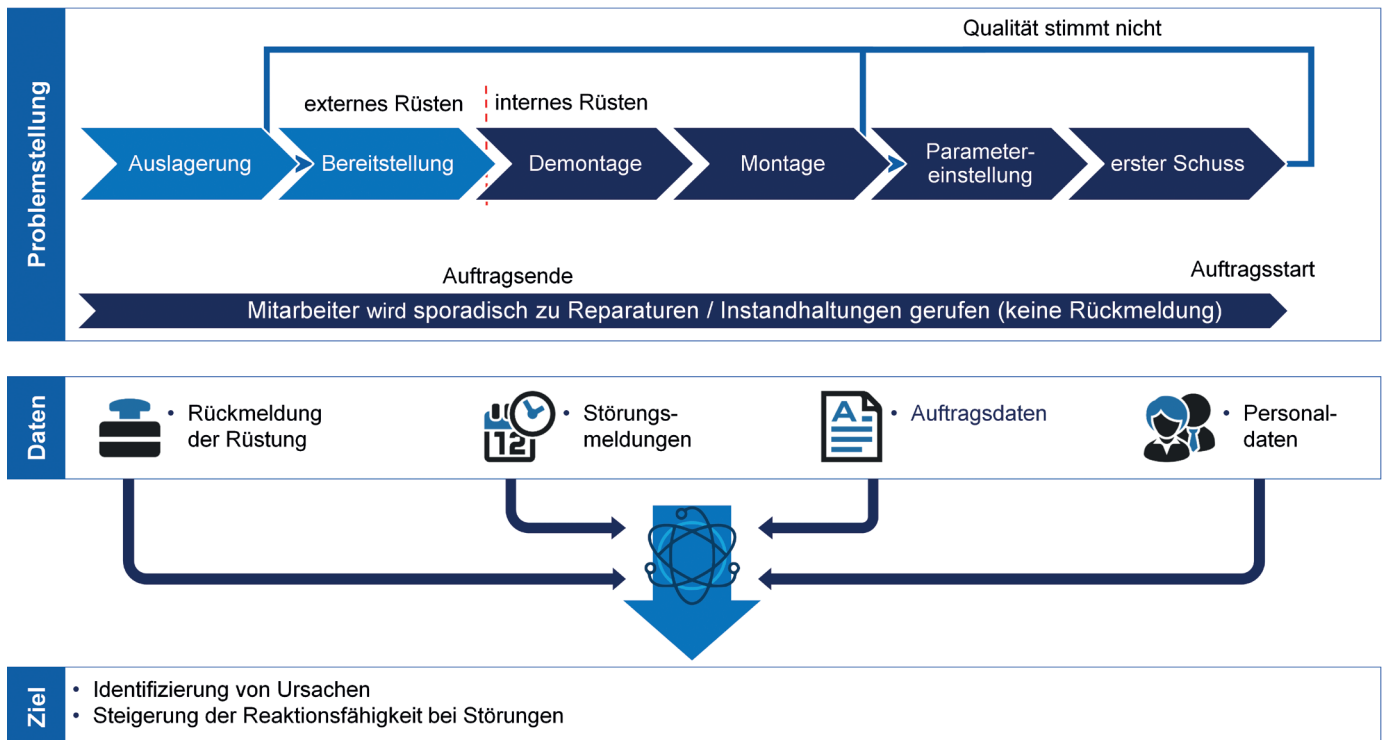


Bild 2: Datenfusion bei der Rückmeldungen eines Rüstprozesses (eigene Darstellung)

Objektivität und der Reputation. Bei der Dimension der kontextuellen Datenqualität wird ein Datensatz bezüglich seines Nutzens beurteilt (Kriterien: Mehrwert, Relevanz, Aktualität, Vollständigkeit und Datenmenge). Weiterhin lässt sich die repräsentative Datenqualität anhand der Darstellung einzelner Datensätze bewerten, sodass die Kriterien der Interpretierbarkeit, Verständlichkeit, Einheitlichkeit und Übersichtlichkeit relevant sind. Die letzte Dimension der Datenqualität ist die zugangsbezogene Datenqualität. Dabei wird eingeschätzt, wie ein Datensatz in Bezug auf die Kriterien der Erreichbarkeit und der Zugriffssicherheit im System vorliegt.¹²

Im Forschungsprojekt ‚DaFuER‘ werden unter anderem auch Anwendungsfälle für die genaue Bestimmung der Prozesszeiten untersucht. So variieren beispielsweise die systemseitig erhobenen Rüstzeiten bei einem Projektpartner sehr stark, wodurch eine effektive Produktionssteuerung wie auch eine Nachkalkulation der einzelnen Produkte nicht aussagekräftig sind (s. Bild 2, S. 45). Durch die Fusion bereits vorhandener Datenquellen, wie zum Beispiel der Rüst-Rückmeldungen, der Störungsmeldungen, der Auftrags- und Personaldaten, können Datenkonflikte aufgelöst und die gesamte Datenqualität gesteigert werden. Dabei wird nicht nur die inhaltliche Datenqualität verbessert, sondern es werden auch weitere Datenqualitätsdimensionen (z. B. die Nutzung oder Darstellung der Daten) angesprochen.

¹² S. WANG U. STRONG 1996, S. 20

Literatur

APEL, D.; BEHME, W.; EBERLEIN, R.; MERIGHI, C.: *Datenqualität erfolgreich steuern. Praxislösungen für Business-Intelligence-Projekte*. 3., überarb. und erw. Auflage. dpunkt.verl., Heidelberg 2015.

BAUER, J.: *Produktionscontrolling mit SAP®-Systemen. Effizientes Controlling, Logistik- und Kostenmanagement moderner Produktionssysteme*. 2., aktualis. u. erw. Auflage. Vieweg+Teubner Verlag, Wiesbaden 2003.

BLEIHOLDER, J.; NAUMANN, F.: *Conflict Handling Strategies in an Integrated Information System*. Humboldt-Universität, Berlin 2006. https://hpi.de/fileadmin/user_upload/fachgebiete/nau-

[mann/publications/IIWeb06.pdf](#) (Link zuletzt geprüft: 15.06.2020)

BLEIHOLDER, J.; NAUMANN, F.: KURZ ERKLÄRT. DATENFUSION. IN: DATENBANK-SPEKTRUM 11 (2011) 1, S. 59 – 61.

CASTANEDO, F.: A review of data fusion techniques. In: *The Scientific World Journal* (2013), 19 S. https://www.researchgate.net/publication/259003916_A_Review_of_Data_Fusion_Techniques/fulltext/54a1196f0cf267bdb9017265/A-Review-of-Data-Fusion-Techniques.pdf (Link zuletzt geprüft: 15.06.2020)

KLETTI, J. (HRSG.): *MES – Manufacturing Execution System. Moderne Informationstechnologie unterstützt die Wertschöpfung*. 2. Auflage. Springer Vieweg, Berlin [u. a.] 2015.

MÜLLER, J.: *Transformation operativer Daten zur Nutzung im Data Warehouse*. Deutscher Universitätsverlag, Wiesbaden 2000. – Zugl.: Bochum, Univ., Diss., 1999.

NAYAR, M.: *Achieving Information Integrity*. In: *Information Systems Management* 10 (1993) 2, S. 51 – 58.

NIEHUES, M.; REINHART, G.; SCHMITT, R.; SCHUH, G.; BRAMBRING, F.; ELLEREICH, M.; ELSER, H.; FRANK, D.; GROGGERT, S.; GÜTZLAFF, A.; HEINRICHS, V.; HEMPEL, T.; KOSTYSZYN, K.; NGO, H.; NIENDORF, L.; PERMIN, E.; PROTE, J.-P.; REUTER, C.; TÜRTMANN, R.: *Organisation,*

Qualität und IT-Systeme für Planung und Betrieb. In: *Handbuch Industrie 4.0. Geschäftsmodelle, Prozesse, Technik*. HRSG.: G. REINHART. Hanser, München [u. a.] 2017, S. 137 – 167.

REDMAN, T. C.: *The impact of poor data quality on the typical enterprise*. In: *Communications of the ACM* 41 (1998) 2, S. 79 – 82.

SPATH, D. (HRSG.): *[Studie] Produktionsarbeit der Zukunft – Industrie 4.0*. Fraunhofer-Verl., Stuttgart 2013. <https://www.iao.fraunhofer.de/images/iao-news/produktionsarbeit-der-zukunft.pdf> (Link zuletzt geprüft: 15.06.2020)

VDMA 66412-1: *Manufacturing Execution Systems (MES): Kennzahlen*. Verein Deutscher Maschinen- und Anlagenbau (VDMA) e. V.; ICS 03.100.50. Beuth, Berlin, Oktober 2009.

WANG, R. Y.; STRONG, D. M.: *Beyond Accuracy: What Data Quality Means to Data Consumers*. In: *Journal of Management Information Systems* 12 (1996) 4, S. 5 – 33.

WOLF, T.; DECKER, S.; ABECKER, A.: *Unterstützung des Wissensmanagements durch Informations- und Kommunikationstechnologie*. In: *[Proceedings] Electronic Business Engineering. 4. Internationale Tagung Wirtschaftsinformatik 1999*. Hrsg.: M. Nüttgens; A.-W. Scheer. Physica, Heidelberg 1999, S. 745 – 765.

Ansprechpartner:



Jokim Janßen, M.Sc.
FIR e. V. an der RWTH Aachen
Wissenschaftlicher Mitarbeiter
Bereich Produktionsmanagement
Tel.: +49 241 47705-413
E-Mail: Jokim.Janssen@fir.rwth-aachen.de

Gefördert durch:



aufgrund eines Beschlusses des Deutschen Bundestages



Projekttitel: DaFuER

Forschungs-/Projekträger: Bundesministerium für Wirtschaft und Energie (BMWi); Arbeitsgemeinschaft industrieller Forschungsvereinigungen "Otto von Guericke" e. V. (AiF)

Förderkennzeichen: 20579 N

Projektpartner: AUTO HEINEN GmbH, Berghof Systeme e. K., DFA Demonstrationsfabrik Aachen GmbH, INDUTRAX GmbH, Maschinenfabrik Möllers GmbH, Mattern Consult Gesellschaft für Produktionsregelung und Logistik mbH, mk Plast GmbH & Co. KG, NETRONIC Software GmbH, SICK AG, Ubisense AG, Westaflexwerk GmbH

Internet: dafuer.fir.de