

VEREIN
DEUTSCHER
INGENIEURE

VERBAND DER
ELEKTROTECHNIK
ELEKTRONIK
INFORMATIONSTECHNIK

Implementierung und Betrieb von
Big-Data-Anwendungen
in der produzierenden Industrie
Durchführung von Big-Data-Projekten

Implementation and operation of Big Data
application in the manufacturing industry

Implementation of Big Data projects

VDI/VDE 3714
Blatt 1 / Part 1

Ausgabe deutsch/englisch
Issue German/English

Die deutsche Version dieser Richtlinie ist verbindlich.

The German version of this standard shall be taken as authoritative. No guarantee can be given with respect to the English translation.

Inhalt	Seite	Contents	Page
Vorbemerkung	2	Preliminary note.....	2
Einleitung.....	2	Introduction.....	2
1 Anwendungsbereich	3	1 Scope	3
2 Begriffe	4	2 Terms and definitions	4
3 Durchführung von Big-Data-Projekten	12	3 Implementation of Big Data projects	12
3.1 Voraussetzungen und Rahmen.....	13	3.1 Prerequisites and framework	13
3.2 Definitionsphase	14	3.2 Definition phase.....	14
3.3 Sondierung der Datenlage.....	17	3.3 Exploring the data situation	17
3.4 Datenbewirtschaftung	21	3.4 Data management	21
3.5 Modellierung.....	25	3.5 Modelling.....	25
3.6 Bewertung der Datenanalyse- Ergebnisse	31	3.6 Evaluation of the data analysis results.....	31
3.7 Umsetzung und Einführung	35	3.7 Implementation and introduction.....	35
3.8 Nachhaltigkeit	39	3.8 Sustainability	39
Anhang Vorstellung etablierter organisatorischer Prozessmodelle	43	Annex Presentation of established organizational process models	43
Schrifttum	51	Bibliography	51

VDI/VDE-Gesellschaft Mess- und Automatisierungstechnik (GMA)

Fachbereich Digitale Transformation

VDI/VDE-Handbuch Automatisierungstechnik
VDI-Handbuch Informationstechnik, Band 1: Angewandte Informationstechnik

Vorbemerkung

Der Inhalt dieser Richtlinie ist entstanden unter Beachtung der Vorgaben und Empfehlungen der Richtlinie VDI 1000.

Alle Rechte, insbesondere die des Nachdrucks, der Fotokopie, der elektronischen Verwendung und der Übersetzung, jeweils auszugsweise oder vollständig, sind vorbehalten.

Die Nutzung dieser Richtlinie ist unter Wahrung des Urheberrechts und unter Beachtung der Lizenzbedingungen (www.vdi.de/richtlinien), die in den VDI-Merkblättern geregelt sind, möglich.

Allen, die ehrenamtlich an der Erarbeitung dieser Richtlinie mitgewirkt haben, sei gedankt.

Eine Liste der aktuell verfügbaren und in Bearbeitung befindlichen Blätter dieser Richtlinienreihe sowie gegebenenfalls zusätzliche Informationen sind im Internet abrufbar unter www.vdi.de/3714.

Einleitung

Der Fachausschuss „Big Data“ der VDI/VDE-Gesellschaft für Mess- und Automatisierungstechnik hat sich mit der Erstellung dieser Richtlinie der Aufgabe angenommen, den ökonomischen und ökologischen Nutzen von Big Data aufzuzeigen, den Wissenstransfer über verschiedene Industrien und Branchen hinweg zu verbessern und die Implementierung und den Betrieb von Big-Data-Anwendungen in der produzierenden Industrie voranzutreiben und zu vereinheitlichen.

Die Richtlinienreihe soll eine Orientierung über erforderliche Maßnahmen zur Big-Data-Analyse geben und aufzeigen, welche Methoden für eine zielführende Arbeit geeignet sind und welche Einschränkungen und Hindernisse bestehen. Dem Praktiker sollen Hinweise gegeben werden, welche Methoden und Betrachtungen für den Erfolg eines Big-Data-Projekts hinsichtlich Einsatzes und nachhaltigen Betriebs notwendig sind.

Die Richtlinienreihe VDI/VDE 3714 umfasst die Blätter:

Blatt 1 Durchführung von Big-Data-Projekten

Blatt 2 Datenqualität

Blatt 3 Datenbewirtschaftung

Blatt 4 Analyseverfahrensklassen

Blatt 5 Modellierungsverfahren

Blatt 6 Validierung von Modellen

Blatt 7 Online-Anwendung von datengetriebenen Modellen

Die Richtlinienreihe VDI/VDE 3714 ist im Fachausschuss 7.24 „Big Data“ des Fachbereichs 7

Preliminary note

The content of this standard has been developed in strict accordance with the requirements and recommendations of the standard VDI 1000.

All rights are reserved, including those of reprinting, reproduction (photocopying, micro copying), storage in data processing systems and translation, either of the full text or of extracts.

The use of this standard without infringement of copyright is permitted subject to the licensing conditions (www.vdi.de/richtlinien) specified in the VDI Notices.

We wish to express our gratitude to all honorary contributors to this standard.

A catalogue of all available parts of this series of standards and those in preparation as well as further information, if applicable, can be accessed on the internet at www.vdi.de/3714

Introduction

The technical committee “Big Data” of the VDI/VDE Society Measurement and Automatic Control has taken on the task of creating this standard to demonstrate the economic and ecological benefits of Big Data, to improve the transfer of knowledge across different industries and sectors, and to promote and standardize the implementation and operation of Big Data applications in the manufacturing industry.

The series of standards is intended to provide orientation on the measures required for Big Data analysis and to show which methods are suitable for target-oriented work and which limitations and obstacles exist. The practitioner should be given advice on which methods and considerations are necessary for the success of a Big Data project in terms of implementation and sustainable operation.

The series of standards VDI/VDE 3714 comprises the parts:

Part 1 Implementation of Big Data projects

Part 2 Data quality

Part 3 Data management

Part 4 Analysis process classes

Part 5 Modelling procedures

Part 6 Validation of models

Part 7 Online application of data-driven models

The series of standards VDI/VDE 3714 has been developed in the Technical Committee 7.24 “Big Data”

„Digitale Transformation“ der VDI/VDE-Gesellschaft für Mess- und Automatisierungstechnik (GMA) entstanden. Damit stellen die Produktion sowie die Mess- und Automatisierungstechnik die Schwerpunkte dar. In den Produktionsprozessen werden beispielsweise für Steuerungs- und Regelungsaufgaben oder für die Qualitätssicherung große Datenmengen erhoben, die mittels Datenanalyse für weitere Prozess- und Geschäftsverbesserungen genutzt werden können. Die Richtlinienreihe gibt eine generelle Orientierung sowie Hinweise auf potenzielle Schwierigkeiten und Hürden bei der Durchführung von Big-Data-Anwendungen – von der Entwicklung über die Inbetriebnahme bis zum nachhaltigen Betrieb.

Benachbart zu Big Data finden sich Themen wie das Internet der Dinge (IoT), Vernetzung von Geräten (Smart Devices) oder die zunehmende „Rechnerallgegenwart“ (Ubiquitous Computing) sowie Begriffe wie Business Intelligence, Data Analytics, Advanced Analytics, Data Mining, Smart Data und Data-Warehouse-Systeme, die generell die Nutzung von Daten adressieren.

Die Richtlinienreihe geht von einer generellen Verfügbarkeit aller benötigten Daten aus. Bezüglich Datenmenge, ihrer Struktur und Integrität wird keine Annahme getroffen. Zur Diskussion und Charakterisierung der Daten helfen die sogenannten „fünf Vs“, die die einzelnen Dimensionen von Big Data bezeichnen. Die Daten werden durch Umfang (*Volume*), Unterschiedlichkeit (*Variety*) und ihre Schnellebigkeit (*Velocity*) charakterisiert. Insbesondere bei industriellen Anwendungen sind die Qualität der Daten (*Validity*) und der unternehmerische Mehrwert (*Value*) relevant.

Auf weitere grundsätzliche technische Regeln sei hier hingewiesen, insbesondere im Umfeld von Industrie 4.0:

- VDI 2222 Blatt 1
- VDI 3517
- VDI/VDE 4000
- VDI 4010
- DIN EN ISO 900x
- ISO 13053-x

1 Anwendungsbereich

Mit dem Begriff „Big Data“ werden – obwohl er bereits seit einigen Jahren verwendet wird – unverändert sehr unterschiedliche Themen und Aspekte assoziiert und entsprechend in der gesellschaftlichen Diskussion differenziert diskutiert. Die immer weiter vorschreitende digitale Kommunikation, der in der Umsetzung befindliche Breitbandausbau und die

of the Technical Division 7 “Digital Transformation” of the VDI/VDE Society Measurement and Automatic Control (GMA). Thus, production as well as measurement and automation technology represent the focal points. In production processes, for example, large amounts of data are collected for control and regulation tasks or for quality assurance, which can be used for further process and business improvements by means of data analysis. The series of standards provides a general orientation as well as indications of potential difficulties and hurdles in the implementation of Big Data applications – from development and commissioning to sustainable operation.

Related to Big Data are topics such as the Internet of things (IoT), networking of devices (smart devices) or the increasing “ubiquitous computing” as well as terms such as business intelligence, data analytics, advanced analytics, data mining, smart data, and data warehouse systems, which generally address the use of data.

The series of standards assumes a general availability of all required data. No assumption is made regarding the amount of data, its structure and integrity. The so-called “five Vs”, which denote the individual dimensions of Big Data, help to discuss and characterize the data. The data is characterized by *volume*, *variety*, and *velocity*. The quality of the data (*validity*) and the corporate added value (*value*) are particularly relevant for industrial applications.

Further fundamental technical rules should be pointed out here, especially in the environment of Industry 4.0:

- VDI 2222 Part 1
- VDI 3517
- VDI/VDE 4000
- VDI 4010
- DIN EN ISO 900x
- ISO 13053-x

1 Scope

Although the term “Big Data” has been in use for several years, it continues to be associated with very different topics and aspects and is therefore the subject of differentiated discussion in society. The ever-increasing pace of digital communication, the expansion of broadband currently being implemented, and the fact that data can be processed anywhere are

überall verfügbare Verarbeitungsmöglichkeit von Daten beflügeln diese Diskussion sowohl in der Öffentlichkeit als auch in der Fachwelt. Die Themen reichen von Datenschutz und Datensicherheit bis hin zu generellen Strategien für die digitale Wertschöpfung bei kleinen und mittelständischen Unternehmen und auch bei Großunternehmen.

Im Kontext dieser Richtlinie geht es bei Big Data um Technologien zur Datenanalyse. Entsprechende Algorithmen und Werkzeuge können Erkenntnisse über betriebliche Abläufe liefern und zu deren Optimierung beitragen. Hierzu bedarf es der Umsetzung dieser Methoden und Werkzeuge zur Verarbeitung, Analyse und Interpretation von umfangreichen und komplexen Daten in Big-Data-Anwendungen. Die Richtlinienreihe unterstützt Erstellende und Nutzende bei der Vorbereitung, Entwicklung und Inbetriebnahme dieser Anwendungen sowie ihrem nachhaltigen Einsatz. Letztlich sollen diese Big-Data-Anwendungen verlässlichere Entscheidungsgrundlagen schaffen, um Produkte und Produktionsprozesse ökonomisch, ökologisch und technisch zu verbessern.

Die Richtlinienreihe soll dazu beitragen, die Vielfalt der in den letzten Jahren durch Forschungs-, Entwicklungs- und Praxisarbeiten entstandenen Erkenntnisse aufzubereiten, die Entwicklung und den Einsatz von Big-Data-Anwendungen in produzierenden Industrien sowie deren Nutzung im regulären Betrieb zu unterstützen.

Zur Zielgruppe gehören alle Stakeholder, von den Praktikern/Praktikerinnen bis zu den Entscheidern/Entscheiderinnen, von der Fertigungs- bis zur Prozessindustrie. Die Richtlinienreihe wendet sich dabei an die Nutzenden und die Erstellenden von Big-Data-Anwendungen in der produzierenden Industrie, unabhängig und übergreifend für alle Führungs- und Fachaufgaben.

Diese Richtlinie betrachtet anhand der generellen Phasen organisatorischer Prozessmodelle die Planung, Durchführung und die langfristige Sicherung der Nutzung der Big-Data-Anwendung.

fuelling this discussion both among the general public and among experts. The topics range from data protection and data security to general strategies for digital value creation for small and medium-sized enterprises as well as for large companies.

In the context of this standard, Big Data is about data analysis technologies. Corresponding algorithms and tools can provide insights into operational processes and contribute to their optimization. This requires the implementation of these methods and tools for processing, analysing, and interpreting extensive and complex data in Big Data applications. The series of standards supports creators and users in the preparation, development, and commissioning of these applications as well as their sustainable use. Ultimately, these Big Data applications should create a more reliable basis for decision-making in order to improve products and production processes economically, ecologically, and technically.

The series of standards is intended to help process the wide range of findings that have emerged in recent years through research, development, and practical work, and to support the development and use of Big Data applications in manufacturing industries as well as their use in regular operations.

The target audience includes all stakeholders, from practitioners to decision makers, from operations to process industries. The series of standards addresses the users and creators of Big Data applications in the manufacturing industry, independently and comprehensively for all management and specialist tasks.

This standard looks at the general phases of organisational process models to plan, implement, and ensure the long-term use of the Big Data application.