

Carl Christian Tönnies

Datenbasierte Informationsmodelle zur explorativen Analyse von Anlagenkonfigurationen



Datenbasierte Informationsmodelle zur explorativen Analyse von Anlagenkonfigurationen

Data-Based Information Models for Explorative Analyses of Plant Configurations

Von der Fakultät für Maschinenwesen
der Rheinisch-Westfälischen Technischen Hochschule Aachen
zur Erlangung des akademischen Grades eines
Doktors der Ingenieurwissenschaften
genehmigte Dissertation

vorgelegt von

Carl Christian Tönnies

Berichter:

Univ.-Prof. Dr.-Ing. Dipl.-Wirt. Ing. Günther Schuh
apl. Prof Dr.-Ing. Wolfgang Boos

Tag der mündlichen Prüfung: 27. November 2020

ERGEBNISSE AUS DER PRODUKTIONSTECHNIK

Carl Christian Tönnes

Datenbasierte Informationsmodelle zur
explorativen Analyse von Anlagenkonfigurationen

Herausgeber:

Prof. Dr.-Ing. T. Bergs
Prof. Dr.-Ing. Dipl.-Wirt. Ing. G. Schuh
Prof. Dr.-Ing. C. Brecher
Prof. Dr.-Ing. R. H. Schmitt

Band 1/2021



Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <https://portal.dnb.de> abrufbar.

Carl Christian Tönnies:

Datenbasierte Informationsmodelle zur explorativen Analyse von Anlagenkonfigurationen

1. Auflage, 2021

Apprimus Verlag, Aachen, 2021

Wissenschaftsverlag des Instituts für Industriekommunikation und Fachmedien
an der RWTH Aachen

Steinbachstr. 25, 52074 Aachen

Internet: www.apprimus-verlag.de, E-Mail: info@apprimus-verlag.de

ISBN 978-3-86359-948-5

D 82 (Diss. RWTH Aachen University, 2020)

Vorwort

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Lehrstuhl für Produktionssystematik des Werkzeugmaschinenlabors WZL der RWTH Aachen. Die Erstellung dieser Dissertation wäre ohne die Unterstützung vieler Personen nicht möglich gewesen, denen ich an dieser Stelle danken möchte.

Mein Dank gilt insbesondere Herrn Professor Günther Schuh für die Möglichkeit zur Anfertigung meiner Dissertation. Durch seine inspirierende Art, sein Vertrauen in meine Person und die mir im Rahmen von vielen Projekten übertragene Verantwortung hat er sowohl meine Promotion als auch meine persönliche Entwicklung maßgeblich geprägt. An seinem Lehrstuhl für Produktionssystematik habe ich ein zwar herausforderndes, aber auch innovatives Umfeld vorgefunden, welches diese Arbeit erst ermöglicht hat.

Für die Übernahme des Koreferats danke ich Herrn Professor Wolfgang Boos. Herrn Professor (Osaka) Jochen Büchs danke ich für die Übernahme des Prüfungsvorsitzes.

An dieser Stelle möchte ich mich auch bei den Oberingenieuren, die meinen Weg am Werkzeugmaschinenlabor WZL begleiteten und maßgeblich prägten, bedanken. Insbesondere gebührt mein Dank Dr. Michael Riesener, der vor allem die nicht einfache Phase der Themenfindung und -schärfung mit seiner erfrischenden, aber auch kritischen Diskussionsfreude begleitet und unterstützt hat. Auch die Gruppenleiter und hier vor allem Dr. Christian Mattern, möchte ich nicht unerwähnt lassen. Mit seiner Akribie und konstruktiven Anmerkungen stand er kontinuierlich als Diskussionspartner zur Verfügung.

Aus meiner Anfangszeit in der Abteilung Innovationsmanagement möchte ich mich vor allem bei meinen ersten Bürokollegen Dr. Frederic Diels, Dr. Casimir Ortlieb und Dr. Stefan Breunig bedanken, die mir den Einstieg leichtgemacht haben sowie immer viele wertvolle Ratschläge für mich hatten. Dr. Sebastian Schloesser und Jan Koch möchte ich für die unzähligen gemeinsam am Institut verbrachten Wochenenden danken, die immer wieder aufs Neue motiviert haben und so zum Gelingen der Arbeit beitragen.

Aber erst durch die vielen anderen Kollegen wurde die Promotion in der Abteilung Innovationsmanagement zu einer unvergesslichen Zeit. Bedanken möchte ich mich bei Michael Mendl-Heinisch, mit dem ich die unterhaltsamsten Dienstreisen verbrachte, bei Johanna Ays, für die immer offene Bürotür, und bei Manuel Ebi, für die vielen Gespräche auf der Autobahn zwischen Düsseldorf und Aachen. Aber auch die jüngeren Kollegen und da vor allem Jonas Tittel und Julian Kress haben zu dem erfolgreichen Abschluss meiner Promotion beigetragen. Durch ihre Wissbegierde haben sie mich immer wieder dazu gebracht, Dinge kritisch zu hinterfragen.

Einen herzlichen Dank möchte ich an meine studentischen Mitarbeiter richten. Insbesondere hervorzuheben sind hier Alexandra Mierke, Andrej Naumenko und Thorben Ipers, die auch nach Auslandsaufenthalten und Praktika immer wieder an den Lehrstuhl zurückgekehrt sind und so meinen Dissertationsprozess wortwörtlich von der ersten Idee bis zu den letzten Herausforderungen mitbegleitet haben.

Nicht unerwähnt bleiben darf die Düsseldorfcrew – Lisa, Henning, Julie, Jacky, Peter, Vanessa und Annika – für die regelmäßige Resozialisierung nach langen Arbeitstagen. Julian, Tobias und Kai danke ich für die langjährige verlässliche Freundschaft auch nach Phasen des Abtauchens.

Mein größter Dank gebührt meiner Familie, allen voran meinen Eltern und meiner Schwester, die mir auf meinem Weg gerade auch in schweren Zeiten immer zur Seite gestanden haben und auf die ich mich immer bedingungslos verlassen konnte. Ohne diesen Rückhalt wäre ich heute nicht da, wo ich bin und wäre nicht der, der ich bin. Mein ebenso großer Dank gilt meiner Freundin. Die fortwährende Ermunterung und das kritische Lektorat waren insbesondere in der heißen Phase die zentrale Motivation und die notwendige Unterstützung, die zum erfolgreichen Abschluss dieser Arbeit geführt haben.

Euch widme ich diese Arbeit.

Stuttgart, im Januar 2021

Carl Christian Tönnies

Inhaltsverzeichnis

Verzeichnis der Abbildungen.....	VII
Verzeichnis der Tabellen.....	XIII
Verzeichnis der Abkürzungen.....	XV
Verzeichnis der Formelzeichen.....	XVII
Zusammenfassung.....	XXI
Summary.....	XXIII
1 Exposition	1
1.1 Einleitung und Motivation des Themas	1
1.1.1 Relevanz.....	2
1.1.2 Problemidentifikation	4
1.2 Zielsetzung der Arbeit	9
1.3 Forschungskonzeption.....	11
1.3.1 Einordnung des Forschungsobjekts in die Wissenschaftssystematik	11
1.3.2 Design Science Research als forschungsmethodisches Vorgehen	13

1.4	Aufbau der Arbeit	17
2	Grundlagen und Definitionen im Kontext der Arbeit.....	19
2.1	Grundlagen der Informationsmodellierung.....	19
2.1.1	Definition relevanter Begrifflichkeiten	20
2.1.2	Vorgehensweise der konzeptionellen Informationsmodellierung	28
2.1.3	Methoden zur Modellierung von Informationsstrukturen.....	35
2.2	Produktinformationen im Anlagenbau	39
2.2.1	Definition relevanter Begrifflichkeiten	39
2.2.2	Verwendung von Produktmodellen	46
2.2.3	Produktinformationen im Produktlebenszyklus.....	50
2.3	Systembasierte Informationsbereitstellung.....	55
2.3.1	Definition relevanter Begriffe	55
2.3.2	Konzepte zur Bereitstellung von Informationen im Unternehmen.....	62
2.3.3	Bereitstellung von Produktinformationen im Kontext von Industrie 4.0.....	67
2.4	Datenanalyse zur Informationsgenerierung	71
2.5	Zwischenfazit	76
3	Bestehende Ansätze zum Aufbau datenbasierter Informationsmodelle.....	79
3.1	Kriterien zur Bewertung bestehender Ansätze	79
3.2	Bestehende Ansätze zur Gestaltung datenbasierter Informationsmodelle	82
3.2.1	Darstellungen von Normen und Richtlinien.....	82
3.2.2	Darstellung von Ansätzen des Informationsmanagements.....	84

3.2.3	Darstellung von Ansätzen des Datenmanagements	90
3.2.4	Darstellung von Ansätzen zur Gestaltung von Produktmodellen.....	94
3.3	Bewertung bestehender Ansätze und Positionierung der Arbeit	100
3.4	Zwischenfazit	104
4	Konzeption der Methodik zum Aufbau eines datenbasierten Informationsmodells	105
4.1	Ordnungsrahmen der Methodik	105
4.2	Anforderungen an die Methodik zur datenbasierten Informationsmodellierung.....	107
4.2.1	Inhaltliche Anforderungen	107
4.2.2	Formale Anforderungen.....	109
4.3	Formale Konstruktion und modelltheoretische Konzeptionierung der Methodik.....	110
4.3.1	Definition der Elemente einer Methodik	110
4.3.2	Konstruktionsorientiertes Modellverständnis.....	113
4.4	Konzeptionierung der Methodik zur Gestaltung datenbasierter Informationsmodelle.....	116
4.4.1	Phase I: Beschreibung des Zielsystems	117
4.4.2	Phase II: Ableitung des Informationsbedarfs	118
4.4.3	Phase III: Modellierung der Informationsstruktur	119
4.4.4	Phase IV: Beschreibung von Datenquellen	119
4.4.5	Phase V: Verknüpfung von Informationsmodell und Datenquellen.....	121
4.5	Ableitung der zu entwickelnden Artefakte	122
4.6	Zwischenfazit	123

5	Detaillierung der Methodik zum Aufbau datenbasierter Informationsmodelle	125
5.1	Beschreibung des Zielsystems	126
5.1.1	Definition des Anwendungsfalls.....	127
5.1.2	Beschreibung des Informationssystems	134
5.1.3	Bestimmung der informatorischen Anforderungen.....	140
5.2	Ableitung des Informationsbedarfs	147
5.2.1	Identifizierung des Informationsbedarfs	148
5.2.2	Homogenisierung der Informationen	156
5.2.3	Charakterisierung der Informationen.....	160
5.3	Modellierung der Informationsstruktur.....	168
5.3.1	Aktivitäten zur Konstruktion der Dimensionen von Attributen	170
5.3.2	Aktivitäten zur Definition der Aggregationspfade von Fakten	175
5.3.3	Aktivitäten zur Modellierung der Beziehungen	178
5.4	Beschreibung der Datenquellen	180
5.4.1	Identifizierung von Datenquellen.....	181
5.4.2	Katalogisierung von Datenquellen.....	185
5.4.3	Bewertung der Datenquellen	189
5.5	Verknüpfung von Informationsmodell und Datenquellen	199
5.5.1	Definition von Datenquellenoptionen	200
5.5.2	Berechnung des Data-Information-Fit-Indicator.....	202
5.5.3	Auswahl geeigneter Datenquellen	208
5.6	Zwischenfazit	211

6	Evaluierung und kritische Reflexion	215
6.1	Herleitung der Evaluierungsstrategie.....	215
6.2	Evaluierung des Vorgehens am Beispiel eines Herstellers für Windenergieanlagen.....	217
6.2.1	Ausgangssituation der Windkraft AG	217
6.2.2	Demonstration der Methodik	217
6.3	Kritische Reflexion der Anwendungserfahrung.....	229
7	Zusammenfassung und Ausblick	231
	Literaturverzeichnis	235
A	Anhang	269
A.1	Übersicht möglicher Datengliederungsebenen	269
A.2	Dimensionen zur Beschreibung von Daten- und Informationsqualität	270
A.3	Übersicht bestehender Literatur zur Abgrenzung des digitalen Schattens ...	276
A.4	Darstellung weiterer Anwendungsfälle	279
A.5	Herleitung des Modells zur Formalerfassung von Datenquellen.....	281
A.6	Berechnungsvorschriften zur Ermittlung der Datenqualität	284
A.7	Evaluierung am Beispiel der Windkraft AG.....	291

Verzeichnis der Abbildungen

Abbildung 1-1:	Potenziale datenbasierter Entscheidungen nach JANK.....	3
Abbildung 1-2:	Datenführende Systeme für Produktinformationen nach RIESENER.....	5
Abbildung 1-3:	Wissenschaftssystematik nach ULRICH UND HILL	11
Abbildung 1-4:	Ordnungsrahmen Design Science Research Zyklen	14
Abbildung 1-5:	Prozessmodell der Design Science Research	15
Abbildung 1-6:	Umsetzung des Forschungsprozesses nach PEFFERS ET AL.	16
Abbildung 1-7:	Aufbau der Arbeit	17
Abbildung 2-1:	Data-Information-Knowledge-Wisdom Hierarchy nach ROWLEY	20
Abbildung 2-2:	Aufgaben des Informationsmanagements nach FERSTL UND SINZ	24
Abbildung 2-3:	Abgrenzung von Informationssystemen, -modellen und - strukturen von Datenmodellen, und -architekturen.....	26
Abbildung 2-4:	Teilmengen des Informationsstands.....	30
Abbildung 2-5:	Beschreibungselemente der DFM nach GOLFARELLI ET AL.....	37
Abbildung 2-6:	Darstellung des Kubenstrukturmodell nach SCHELP	38
Abbildung 2-7:	Ausschnitt der idealtypischen Charakterisierung eines Auftragsfertigers nach SCHUH	40
Abbildung 2-8:	Einordnung des Maschinen- und Anlagenbaus nach BELZ ET AL.	41
Abbildung 2-9:	Differenzierung technischer Systeme nach KOLLER	43

Abbildung 2-10:	Abgrenzung zwischen Konfiguration und Konstruktion in Abhängigkeit der Produktdetaillierungsebene nach BONGULIELMI	45
Abbildung 2-11:	Aspekte eines Produkts nach DIN 81346-1	48
Abbildung 2-12:	Informationsströme nach WUEST ET AL.	55
Abbildung 2-13:	Datenhierarchie.....	56
Abbildung 2-14:	Datenbestandstypen nach GEBAUER UND WINDHEUSER	57
Abbildung 2-15:	Ordnungsrahmen zur Beschreibung betrieblicher Daten	58
Abbildung 2-16:	Abgrenzung verschiedener Typen von Informationssystemen	63
Abbildung 2-17:	Einbindung eines data warehouse in den Entscheidungsunterstützungsprozess nach HINRICHS.....	67
Abbildung 2-18:	Entwicklung von statistischen Methoden zum Data Mining	72
Abbildung 2-19:	Darstellung der statistischen Arbeitsweise nach BECKER ET AL.	73
Abbildung 3-1:	Literaturanalyseprozess in Anlehnung an VOM BROCKE ET AL.	80
Abbildung 3-2:	Ausschnitt des Informationsmodells und den Fähigkeiten nach ISO 10303:242.....	83
Abbildung 3-3:	Ordnungsrahmen der Diagrammtypen nach DETERMANN	86
Abbildung 3-4:	Multidimensionale Datenstruktur nach NUSSELEIN	87
Abbildung 3-5:	Funktionen des digitalen Schattens nach BAUERNHANSL ET AL.	89
Abbildung 3-6:	Migrationspfad des digitalen Schattens	91
Abbildung 3-7:	Auszug des Referenzproduktmodells nach NURCAHYA.....	95
Abbildung 3-8:	Ausschnitt des Produktdatenmodells nach DEMOLY ET AL.	96
Abbildung 3-9:	Zusammenfassende Bewertung der bestehenden Ansätze	102
Abbildung 4-1:	Ordnungsrahmen zur Strukturierung der Methodik.....	106

Abbildung 4-2:	Elemente des Methoden-Engineering	112
Abbildung 4-3:	Modellbildung und Modellarten nach RESCHKE	115
Abbildung 4-4:	Ableitung der zu entwickelnden Artefakte	123
Abbildung 5-1:	Notation der Methodikbeschreibung	125
Abbildung 5-2:	Struktur zur Beschreibung des Zielsystems	127
Abbildung 5-3:	Phasen des Produktlebenszyklus	128
Abbildung 5-4:	Exemplarische Anwendungsfälle eines digitalen Schattens	131
Abbildung 5-5:	Anwendung der Kriterien zur Definition des Anwendungsfalls	133
Abbildung 5-6:	Herleitung des Modells zur Beschreibung von Informationssystemen	135
Abbildung 5-7:	Modell zur Beschreibung des Informationssystems.....	135
Abbildung 5-8:	Beispielhaftes Ergebnis der Beschreibung des Informationssystems für den Anwendungsfall <i>Konstruktion</i>	140
Abbildung 5-9:	Dimensionen zur Beschreibung von Informationsqualität	144
Abbildung 5-10:	Matrix für den paarweisen Vergleich	145
Abbildung 5-11:	Struktur zur Ableitung des Informationsbedarfs	147
Abbildung 5-12:	Strukturierungsebenen des Modells zur Ableitung von Produktperspektiven.....	149
Abbildung 5-13:	Modell zur Ableitung von Beschreibungsperspektiven für Erzeugnisse des Maschinen- und Anlagenbaus.....	151
Abbildung 5-14:	Ableitung des Informationsbedarfs mittels Geschäftsfragen	155
Abbildung 5-15:	Definition von Basiswörtern mittels Bezeichnungszersetzung	157
Abbildung 5-16:	Einordnung sprachlicher Defekte zur Homogenisierung	159
Abbildung 5-17:	Übersicht der Parameter zur Charakterisierung von Informationen	161
Abbildung 5-18:	Aufteilung des Informationsrasters in Fakten und Attribute	162

Abbildung 5-19:	Parameter zur Beschreibung von Fakten	163
Abbildung 5-20:	Parameter zur Beschreibung von Attributen	164
Abbildung 5-21:	Aktivitäten zur Charakterisierung der Informationen	165
Abbildung 5-22:	Fakten-Attribute-Tableau als Ergebnis der Phase II	168
Abbildung 5-23:	Vereinfachte Darstellung des Informationsraums als Würfel	169
Abbildung 5-24:	Struktur zum Aufbau des Informationsmodells	169
Abbildung 5-25:	Vorgehen zur Definition der Hierarchie von Dimensionen	171
Abbildung 5-26:	Unterschiedliche Hierarchiestrukturtypen nach SCHELP	172
Abbildung 5-27:	Typisierte Darstellung einer einfachen Hierarchie	173
Abbildung 5-28:	Typisierte Darstellung von Attributhierarchien	173
Abbildung 5-29:	Visualisierung eines Fakts im Kubenstrukturmodell	175
Abbildung 5-30:	Berechnungsvorschrift für Fakten	176
Abbildung 5-31:	Aggregation von Faktenwerten entlang der Vertriebsstruktur	177
Abbildung 5-32:	Vorgehen zur Identifizierung von Beziehungen zwischen Fakten	179
Abbildung 5-33:	Fakten-Relationsdiagramm mit typisierten Dimensionshierarchien	180
Abbildung 5-34:	Struktur zur Beschreibung von Datenquellen	181
Abbildung 5-35:	Beispiele typischer Datenquellen entlang des Produktlebenszyklus	182
Abbildung 5-36:	Aktivitätskarte zur Identifikation von Datenquellen	184
Abbildung 5-37:	Inhaltliche Beschreibung von Datenquellen	186
Abbildung 5-38:	Kriterien zur Beschreibung von Datenquellen	189
Abbildung 5-39:	Zusammenhang zwischen Daten- und Informationsqualität	190
Abbildung 5-40:	Mögliche Abstandsfunktionen zur Ermittlung der Richtigkeit	193

Abbildung 5-41:	Struktur zur Auswahl von Datenquellen	200
Abbildung 5-42:	Definition von Varianten zur Erzeugung von Informationen aus unterschiedlichen Datenquellen	201
Abbildung 5-43:	Wirkzusammenhang von Informations- und Datenqualität	203
Abbildung 5-44:	1:1-Zuordnung von Daten- und Informationsqualität	204
Abbildung 5-45:	2:1-Zuordnung von Daten- und Informationsqualität	204
Abbildung 5-46:	3:1-Zuordnung von Daten- und Informationsqualität	205
Abbildung 5-47:	Paarweiser Vergleich zur Ermittlung der Transformationsfaktoren	205
Abbildung 5-48:	Vorgehen zur Auswahl geeigneter Datenquellen	209
Abbildung 5-49:	Darstellung der entwickelten Methodik	213
Abbildung 6-1:	Effektivität und Effizienz einer Methodik	216
Abbildung 6-2:	Formalisierte Beschreibung des Zielsystems	218
Abbildung 6-3:	Informatorische Anforderungen des Anwendungsfalls	219
Abbildung 6-4:	Ausschnitt des Fakten-Attribute-Tableaus des Evaluierungsbeispiels	221
Abbildung 6-5:	Konstruktion der Dimension Parkstruktur	222
Abbildung 6-6:	Berechnungsvorschrift für den Fakt Windbedingungen	222
Abbildung 6-7:	Kubenstrukturmodell für den Fakt Windbedingungen	223
Abbildung 6-8:	Ausschnitt des Fakten-Relationsdiagramms	223
Abbildung 6-9:	Auswahl relevanter Datenquellen	224
Abbildung 6-10:	Datenquellenoptionen für Informationen zum Netzanschlusspunkt sowie zur Temperatur und zu den Windbedingungen	226
Abbildung A-7-1:	Mögliche Anwendungsfälle eines digitalen Schattens	279
Abbildung A-7-2:	Vollständiges Fakten-Relationsdiagramm	291

Verzeichnis der Tabellen

Tabelle 2-1:	Funktionalitäten von PDM-Systemen.....	52
Tabelle 6-1:	Ausschnitt der definierten Geschäftsfragen	220
Tabelle 6-2:	Datenquellspezifische Informationsqualität.....	228
Tabelle 6-3:	Data-Information-Fit-Indicator für ausgewählte Datenquellen	228
Tabelle A-1:	Übersicht der Datenbankstrukturebenen im Vergleich	269
Tabelle A-2:	Ansätze zur Beschreibung von Daten- und Informationsqualität.....	270
Tabelle A-3:	Übersicht relevanter Veröffentlichungen zur Produktinformationsbereitstellung im Kontext von Industrie 4.0.....	276
Tabelle A-4:	Jung (2006) Architekturen zur Datenintegration	281
Tabelle A-5:	Bittel (2014) Semantische Informationsintegration.....	282
Tabelle A-6:	Ebner (2014) Entwicklung einer Methode zum Entwurf einer Unternehmensdatenarchitektur.....	283
Tabelle A-7:	Ansätze zur Ermittlung des Alters.....	284
Tabelle A-8:	Ansätze zur Ermittlung der Granularität	284
Tabelle A-9:	Ansätze zur Ermittlung der Richtigkeit	285
Tabelle A-10:	Ansätze zur Ermittlung der Nachvollziehbarkeit	286
Tabelle A-11:	Ansätze zur Ermittlung der Verarbeitungsgeschwindigkeit	286
Tabelle A-12:	Ansätze zur Ermittlung der Redundanz.....	286
Tabelle A-13:	Ansätze zur Ermittlung der Genauigkeit	287

Tabelle A-14:	Ansätze zur Ermittlung der Abrufgeschwindigkeit	287
Tabelle A-15:	Ansätze zur Ermittlung der strukturellen Konsistenz	288
Tabelle A-16:	Ansätze zur Ermittlung des Umfangs	288
Tabelle A-17:	Ansätze zur Ermittlung der semantischen Konsistenz	289
Tabelle A-18:	Ansätze zur Ermittlung der Vollständigkeit	290
Tabelle A-19:	Bewertung der Datenqualität Dimensionen 1-6	292
Tabelle A-20:	Bewertung der Datenqualität Dimensionen 7-12	293

Verzeichnis der Abkürzungen

acatech:	Deutsche Akademie der Technikwissenschaften
AtO:	Assemble to Order (Auftragsabwicklungsart)
BOM:	Bill of materials (dt. Stückliste)
BPMN:	Business Process Modell and Notation
bspw.:	beispielsweise
bzw.:	beziehungsweise
CAD:	Computer-Aided Design
CAX:	Computer-Aided X
CPS:	cyber-physische Systeme
CRM:	Customer Relationship Management
d. h.:	das heißt
DIFI:	Data-Information-Fit-Indicator
DIN:	Deutsches Institut für Normung
DSR:	Design Science Research
DSS:	Decision Support System
EDSS:	Engineering Decision Support System
EERM:	erweiterte Entity-Relationship-Modell
ERM:	Entity-Relationship-Modell
ERP:	Enterprise Resource Planning
Geschw.:	Geschwindigkeit

ggf.:	gegebenenfalls
i. S.:	im Sinne
IoT:	Internet of Things (Internet der Dinge)
IT:	Informationstechnik
KSM:	Kubenstrukturmodell
KW:	Kalenderwoche
MPM:	Manufacturing Process Planning
NASA:	National Aeronautics and Space Administration
o. ä.:	oder ähnlich
o. Ä.:	oder Ähnliches
OLAP:	Online Analytical Processing
OLTP:	Online Transaction Processing
PDM:	Product Data Management
PLM:	Product Lifecycle Management
rel.:	relative
sog.:	sogenannte
teilw.:	teilweise
u. a.:	unter anderem
u. U.:	unter Umständen
VDI:	Verein deutscher Ingenieure
VDMA:	Verband Deutscher Maschinen- und Anlagenbau
vgl.:	vergleiche
WEA:	Windenergieanlage
WGP:	Wissenschaftliche Gesellschaft für Produktionstechnik
z. B.:	zum Beispiel

Verzeichnis der Formelzeichen

A :	Attribut
\tilde{A} :	Menge der Attributwerte, zu denen A gehört
$d(w_I, w_R)$:	Abstandsfunktion
DQ :	Informationsqualitätsvektor
d_i :	Element i des Vektors der Datenqualität
$Dup(N, E)$:	Anzahl der Duplikate von E in N
E :	Ausprägung eines Entitätstyps
I :	Informationssystem
$i_{sp,j}$:	Element j des spezifischen Informationsqualitätsvektors
i_j :	Element j des Vektors der Informationsqualität
IQ :	Informationsqualitätsvektor
IQ_{max} :	maximal mögliche Bewertung der Informationsqualität
IQ_{min} :	minimal mögliche Bewertung der Informationsqualität
IQ_{sp} :	datenquellenspezifische Informationsqualität
N :	Gesamtzahl der Entitäten zur Beschreibung von E
n_{Herk} :	Anzahl von Datensätzen mit Angaben zur Herkunft
$n_{S,ges}$:	Anzahl der gesamten Datensätze einer Datenquelle
$n_{S,benöt}$:	Anzahl benötigter Datensätze
$n_{S,bereit}$:	Anzahl bereitgestellter Datensätze
$n_{S,str.kons}$:	Anzahl konsistenter Datensätze

$n_{w,inak}$:	Anzahl inaktueller Attributwerte
$n_{w,ges}$:	Anzahl gesamter Attributwerte
$p(w)$:	Position des Werts w in einer Hierarchie
$p_{opt}(A)$:	Optimale Position des Attributs A
Q_{Abruf} :	Datenqualität für Abrufgeschwindigkeit
Q_{Alter} :	Datenqualität der Dimension Alter
Q_{Gen} :	Datenqualität der Dimension Genauigkeit
Q_{Gran} :	Datenqualität der Dimension Granularität
Q_{Nach} :	Datenqualität der Dimension Nachvollziehbarkeit
Q_{Red} :	Datenqualität der Dimension Redundanz
Q_{Rich} :	Datenqualität der Dimension Richtigkeit
$Q_{sem.Kons}$:	Datenqualität der Dimension semantische Konsistenz
$Q_{str.Kons}$:	Datenqualität der Dimension strukturelle Konsistenz
Q_{Umf} :	Datenqualität der Dimension Umfang
Q_{Ver} :	Datenqualität der Dimension Verarbeitungsgeschwindigkeit
Q_{Voll} :	Datenqualität der Dimension Vollständigkeit
$Rep(N, E)$:	Anzahl der Elemente von E für die ein korrespondierendes Element in N existiert
r_g :	Konsistenzregel g
s :	Sekunde
$s(A)$:	Stelligkeit des Attributs A
$s_{opt}(A)$:	Optimale Anzahl von (Nachkomma)-Stellen
S :	Datensatz einer Datenquelle
T :	Transformationsmatrix
$t_{i,j}$:	Transformationsfaktor

t_E :	Erfassungszeitpunkt
t_V :	Verarbeitungszeitpunkt
w_I :	Attributwert im Informationssystem
w_R :	Attributwert in der Realität

Zusammenfassung

Die Digitalisierung von Produkten und Prozessen im industriellen Umfeld und die resultierende Generierung von Daten bietet Unternehmen die Möglichkeit, mittels Analyseverfahren einen Wettbewerbsvorteil aufzubauen. Insbesondere im Maschinen- und Anlagenbau entsteht durch die kundenspezifischen Anforderungen eine hohe Vielzahl projektspezifischer Produktdaten. Jedoch zeigen Studien, dass diese produktbezogenen Daten nicht für Analysezwecke zur Verfügung stehen. Ursächlich dafür sind drei zentrale Hindernisse. Erstens geht mit der Diversität von datenführenden Systemen auch eine unterschiedliche Strukturierung der Datenbasis einher. Zweitens führt die dezentrale Pflege von Datenquellen zu einer semantischen Heterogenität und damit zu keinem einheitlichen Verständnis der Informationen. Drittens sind die IT-Systeme auf die Unterstützung von Prozessen ausgelegt, sodass historische Daten oftmals nicht unmittelbar abrufbar sind.

Das Ziel der Arbeit besteht in der Entwicklung einer Methodik zur Gestaltung datenbasierter Informationsmodelle für die explorative Analyse von Produktinformationen. Damit soll das Informationsmanagement von Unternehmen des Maschinen- und Anlagenbaus unterstützt werden, existierende produktbezogene Daten durch den Aufbau eines multidimensionalen Informationsmodells nutzbar zu machen.

Die Methodik gliedert sich in fünf Phasen. In der Ersten wird auf Basis eines unternehmensindividuell beschriebenen Anwendungsfalls die Ableitung der informatorischen Anforderungen an das Informationsmodell ermöglicht. Im Sinne einer Informationsbedarfsanalyse werden Anwender in der zweiten Phase über ein generisches Produktbeschreibungsmodell befähigt, die Informationen des Anwendungsfalls zu ermitteln und zu charakterisieren. In der dritten Phase werden die Informationen strukturiert in ein semantisches Informationsmodell überführt. Die vierte Phase dient der systematischen Identifizierung von Datenquellen und der Ermittlung der jeweiligen Datenqualität. Die fünfte Phase ermöglicht die Definition der Datenquellen, welche die benötigten Informationen in der angestrebten Qualität bereitstellen können, sodass ein datenbasiertes Informationsmodell entsteht.

Summary

The increasing digitalization of products and processes leads towards a tremendous generation of data in the industrial environment. These data offer the opportunity to build up a sustainable competitive advantage to companies by implementing suitable analysis methods. Data analysis results in a more precise and substantiated information base, which leads to an improvement in a decision-making process. Especially in plant engineering, customer-specific requirements and constraints result in a large number of project-specific product data. However, studies show that this product-related data is often not available for analytical use. There are three main barriers. Firstly, the diversity of data management systems in a company causes different structures of the underlying data. Secondly, the decentralized management of data sources leads to semantic heterogeneity and thus to a lack of consistent understanding. Thirdly, IT systems are designed for providing direct support for processes, so that historical data are often technically inaccessible.

The aim of this dissertation is the development of a methodology to design data-based information models for the explorative analysis of product information for supporting the information management of companies in the plant engineering sector. Existing product-related data are made usable for analyses by building a multidimensional information model.

The developed methodology is divided into five phases. The first phase is used to derive informational requirements for the information model based on a company-specific use case. In the second phase, the users are enabled to determine and characterize necessary information for the use case. A generic product description model supports the information requirements analysis. In the third phase, the identified information is structured and transferred into a semantic information model. The fourth phase aims at systematically identifying data sources and evaluating their data quality. Finally, in the fifth phase it is necessary to determine which data sources available in the company can provide the necessary information in the pursued quality. The result is a data-based information model.

1 Exposition

1.1 Einleitung und Motivation des Themas

Durch den technischen Fortschritt der Kommunikationstechnologien, die flächen-deckende Verbreitung des Breitbandinternets sowie die steigenden Kapazitäten zur Datenspeicherung wird eine wachsende Menge an Daten erzeugt.¹ Studien gehen davon aus, dass die heute existierende Datenmenge von 33 Zettabyte² auf etwa 175 Zettabyte im Jahr 2025 anwachsen wird.³ In der Folge verändert sich auch die Umgangsweise mit Daten im unternehmerischen Umfeld. Heutzutage sind Daten das Bindeglied zwischen den einzelnen Wertschöpfungsstufen eines Unternehmens, wie der Entwicklung, der Produktion und des anschließenden Servicegeschäfts.⁴ Im Rahmen dieser Vernetzung entsteht in Verbindung mit einer geeigneten Analytik ein „wertvoller Rohstoff“⁵ und ein nachhaltiger Wettbewerbsvorteil für produzierende Unternehmen.⁶ Diese Entwicklung der Wettbewerbsfähigkeit auf Basis einer gesteigerten Informationsverfügbarkeit und Transparenz, die auf der voranschreitenden Digitalisierung, dem zunehmenden Vernetzungsgrad von wertschöpfenden Prozessen sowie der angewendeten Datenanalyse beruht, wird als eine Kernerrungenschaft von *Industrie 4.0* gesehen.⁷

¹ Vgl. Ardagna et al. (2018), Data Quality Assessment for Big Data, S. 548; Schwaderer und Wieland (2018), Digital Index 2017/2018, S. 54f.; Taleb et al. (2018), Big Data Quality, S. 166; Schuh et al. (2017), Geschäftsmodell-Innovation, S. 6; Cai und Zhu (2015), The Challenges of Data Quality, S. 1.

² 1 Zettabyte entspricht 10⁹ Terabytes oder 10²¹ Bytes.

³ Vgl. Reinsel et al. (2018), Data Age 2025, S. 3.

⁴ Vgl. Schuh und Riesener (2018), Produktkomplexität managen, S. 199; Hedberg et al. (2016), Testing the Digital Thread, S. 2.

⁵ Bauernhansl et al. (2016), WGP-Standpunkt Industrie 4.0, S. 34.

⁶ Vgl. Erevelles et al. (2016), Big Data Consumer Analytics, S. 897; Bauernhansl et al. (2016), WGP-Standpunkt Industrie 4.0, S. 34; Porter und Heppelmann (2014), How smart, connected products are, S. 77; Barton und Court (2012), Making Advanced Analytics Work, S. 80; Gabriel et al. (2009), Datawarehouse & Data Mining, S. 4

⁷ Vgl. Riesener und Dölle (2018) Marktspiegel Business Software PLM/PDM, S. 12; Reinhart und Zühlke (2017), Von CIM zu Industrie 4.0, S. XXXIV; Schuh et al. (2017), Geschäftsmodell-Innovation, S. 12.

Der Ausdruck *Industrie 4.0* wurde 2011 im Rahmen der Hannover Messe eingeführt, um in Anlehnung an die vorherigen drei industriellen Revolutionen eine Reihe von Technologien, welche die Produktionstechnik der Zukunft maßgeblich definieren, unter einem zentralen Mantel zusammenzufassen.⁸

1.1.1 Relevanz

Durch die dargelegten Errungenschaften der *Industrie 4.0*, stehen Unternehmen nun vor der Möglichkeit, Entscheidungen auf Basis einer fundierten Informationslage zu treffen und so die Produktivität von Prozessen zu steigern.⁹ Daten sind jedoch nicht per se wertschöpfend. Erst durch eine Interpretation durch Wissensträger im entsprechenden Kontext lassen sie sich als Information wertschöpfend einsetzen.¹⁰ Die Methoden der Datenanalyse können so Entscheidungen verbessern, indem historische Daten analysiert, Muster erkannt und auf dieser Basis Maßnahmen abgeleitet werden.¹¹ Eine von der Wirtschaftsprüfungsgesellschaft KPMG durchgeführte Studie hat ergeben, dass subjektive Entscheidungen auf der Grundlage von Erfahrungen immer häufiger durch objektive datenbasierte Erkenntnisse ersetzt oder mindestens unterstützt werden.¹² Dabei geben 41 % der befragten Unternehmen an, bereits heute durch den Einsatz von Datenanalysen Risiken im Unternehmen reduzieren zu können.¹³ Auch eine im Jahr 2016 von der Deutschen Akademie der Technikwissenschaften (acatech) durchgeführte Befragung von 67 Führungskräften, welche in Industrie 4.0-Bereichen aktiv sind, ergab, dass die Befähigung zur Echtzeit-Entscheidungsunterstützung eine wesentliche Anforderung an zukünftige Engineering-Methoden darstellt.¹⁴ Dabei zeichnet sich der Trend ab, dass Unternehmen zunehmend auf die Datenanalyse als Wettbewerbsvorteil vertrauen, je

⁸ Vgl. Jeschke et al. (2017), *Industrial Internet of Things*, S. VII; Reinhart und Zühlke (2017), *Von CIM zu Industrie 4.0*, S. XXXIV.

⁹ Vgl. Schuh et al. (2017), *Industrie 4.0 Maturity Index*, S. 25; Barton und Court (2012), *Making Advanced Analytics Work*, S. 80; Helu und Hedberg (2015), *Enabling Smart Manufacturing Research*, S. 86.

¹⁰ Vgl. Schuh et al. (2018), *Data-Based Determination of the Product Complexity*, S. 144; Krčmar (2015), *Einführung in das Informationsmanagement*, S. 42.

¹¹ Vgl. Schuh et al. (2018), *Derivation of a Digital Shadow*, S. 1; Schuh und Blum (2016), *Design of a Data Structure for Data Analytics*, S. 2164.

¹² Vgl. KPMG (2017), *Mit Daten Werte schaffen*, S. 11.

¹³ Vgl. KPMG (2017), *Mit Daten Werte schaffen*, S. 15.

¹⁴ Vgl. Abramovici und Herzog (2016), *Engineering im Umfeld von Industrie 4.0*, S. 25.