

Knut Hildebrand
Marcus Gebauer
Holger Hinrichs
Michael Mielke *Hrsg.*

Daten- und Informationsqualität

Auf dem Weg zur Information Excellence

3. Auflage

 Springer Vieweg



Daten- und Informationsqualität

Knut Hildebrand • Marcus Gebauer
Holger Hinrichs • Michael Mielke
(Hrsg.)

Daten- und Informationsqualität

Auf dem Weg zur Information Excellence

3., erweiterte Auflage

Herausgeber
Knut Hildebrand
Fakultät WF
Hochschule Weihenstephan-Triesdorf
Freising, Deutschland

Marcus Gebauer
Haan, Deutschland

Holger Hinrichs
Lübeck, Deutschland

Michael Mielke
Helferskirchen, Deutschland

ISBN 978-3-658-09213-9
DOI 10.1007/978-3-658-09214-6

ISBN 978-3-658-09214-6 (eBook)

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

Springer Vieweg

© Springer Fachmedien Wiesbaden 2008, 2011, 2015

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften. Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen.

Gedruckt auf säurefreiem und chlorfrei gebleichtem Papier

Springer Fachmedien Wiesbaden ist Teil der Fachverlagsgruppe Springer Science+Business Media
(www.springer.com)

Grußwort Prof. Richard Wang

It is a great honor and privilege to be part of the German monograph „Daten- und Informationsqualität“. This compendium for German readers is a rigorous introduction to information quality (IQ). When Michael Mielke invited me to write this preface, I accepted without reservation.

The field of Information Quality has witnessed significant advances over the past two decades. In 1988, Professors Stuart Madnick and Richard Wang at the Massachusetts Institute of Technology pioneered the Total Data Quality Management (TDQM) program, beginning a journey of research publications with key TDQM members such as Professors Yang Lee, Leo Pipino, and Diane Strong. An applied, multi-disciplinary field such as Information Quality demands interaction and collaboration between practitioners and researchers. As such, in 1996, the MIT TDQM program organized the first International Conference on Information Quality (ICIQ) to encourage the exchange of research ideas and results between researchers and practitioners. In 2002, the MIT Information Quality Program was established by the Center for Technology, Policy, and Industrial Development to conduct research on all aspects of Information Quality, such as managing information as a product, developing information product maps, and adopting information quality practices in organizations. In 2007, the MIT Information Quality Program launched the first Industry Symposium to further interactions and collaborations among practitioners, vendors, and academicians. In addition to presentations and workshops, the Symposium also includes vendor presentations, product announcements, and consultancy methods to complement the annual ICIQ conference.

Over the years, the MIT efforts have spawned many conferences, workshops and communities, such as the SIGMOD workshops on Information Quality in Information Systems, the CAiSE workshop on Data and Information Quality, and the German Society for Information and Data Quality that organizes regular conferences, workshops and roundtable meetings. Today, researchers and practitioners have moved beyond establishing Information Quality as a field to resolving IQ problems, which range from defining, measuring, analyzing, and improving IQ to developing tools, methods, and processes for improving the quality of information. As a result, numerous IQ resources are now available for the reader to use. In the industry, vendors such as Acxiom, A.I.D. (France), Deloitte Consulting, EDS, FAST, Firstlogic, FUZZY! Informatik AG (Germany), IBM, Informatica, SAS

and Serasa S.A. (Brazil) are actively promoting information quality. We as a community can be proud of what we have accomplished.

I want to commend Michael Mielke and Marcus Gebauer for outstanding contributions to the Information Quality Community over the last decade. This first German book that includes „Information Quality“ in its title will attract German MIS and IT academics, students, and industry people to information quality.

Richard Y. Wang
Cambridge, Massachusetts, USA
rwang@mit.edu
<http://mitiq.mit.edu>

Editorial

Daten- und Informationsqualität ist im Zeitalter des Internet und der Informationsgesellschaft zunehmend ein entscheidender Faktor für den Erfolg eines Unternehmens. Dabei ist zu beobachten, dass Unternehmen trotz steigender Verfügbarkeit von Informationen nicht unbedingt bessere Entscheidungen treffen. Plakativ ausgedrückt: „Wenn unsere Autos die gleiche Qualität wie unsere Daten hätten, kämen die meisten nicht einmal aus der Garage.“ Die Beobachtung erfolgreicher Unternehmen hat gezeigt, dass der Erfolg maßgeblich vom Vertrauen der Führungskräfte in die Glaubwürdigkeit der zur Verfügung stehenden Informationen abhängt. Erfolgreiche Unternehmen entwickeln daher ihre Business Excellence zunehmend zu einer Business Information Excellence weiter.

Obwohl seit Anfang der 1990er Jahre intensiv am Thema Management der Daten- und Informationsqualität – IQM (Information Quality Management) – geforscht wird, ist es immer noch eine junge Disziplin. Eine Vielzahl von Publikationen, auch zu angrenzenden Themengebieten, ist bereits erschienen, vornehmlich in englischer Sprache. Erst Anfang des 21. Jahrhunderts findet das Thema mehr Beachtung im deutschen Sprachraum. Den Auftakt bildete 2002 die Entwicklung der IQ-Zertifikatskurse von Mielke und Wang, dann 2003 die erste deutsche IQM-Konferenz unter Mitwirkung von Richard Wang gefolgt von offenen Round-Table- Meetings, IQ-Best Practice Day, IQ-Contest, IQ-Challenge u. v. a. m. bis zur Geburtshilfe von IQM-Communities in Europa, Nord- und Südamerika und Australien/ Ozeanien. Darin erkennen Sie die Bedeutung, die die deutsche IQM-Gemeinschaft – organisiert in der DGIQ (Deutsche Gesellschaft für Informations- und Datenqualität e. V.) – für die Entwicklung des Themas Informationsqualität mittlerweile international hat.

Die Herausgeber und Autoren des vorliegenden Buches befassen sich seit vielen Jahren mit dem Thema und tragen mit ihrem Engagement zu einer lebendigen IQM-Gemeinschaft bei, die spätestens mit der ersten deutschen IQM-Konferenz im Jahr 2003 ihren Start hatte. Seitdem sind viele unserer Kollegen mit dem Wunsch nach einem deutschsprachigen Buch an uns herangetreten. Hierbei sollte der Schwerpunkt vor allem auf ‚What Works‘, also dem Machbaren liegen, um allen Interessierten den Einstieg in das Thema Daten- und Informationsqualität so leicht wie möglich zu machen. Allerdings wollten wir dabei nicht unsere Wurzeln vernachlässigen, die bei den Forschungsarbeiten von Prof. Richard Wang vom Massachusetts Institute of Technology (Cambridge, USA) liegen. So

finden sich neben den Praxis-Kapiteln auch Beiträge, in denen die Grundlagenforschung im Mittelpunkt steht. Das Ergebnis unserer Bemühungen halten Sie in Ihren Händen.

Gehen Sie mit uns auf eine spannende Reise durch Grundlagen, Methoden und Praxisbeispiele aus dem Themenfeld der Daten- und Informationsqualität.

Die Herausgeber
Knut Hildebrand
Marcus Gebauer
Holger Hinrichs
Michael Mielke

Editorial zur 3. Auflage

Die erste Auflage der „Daten- und Informationsqualität“ erschien in 2008 und traf einen aufnahmebereiten Markt. Das Feedback war sehr positiv, so dass schon in 2011 die zweite Auflage erfolgte. Seitdem hat das Thema weitere Verbreitung gefunden, sowohl in der Literatur, aber vor allem in der betrieblichen Praxis. Wir haben dies in der jetzt vorliegenden dritten Auflage berücksichtigt, sie wurde überarbeitet und erweitert. Unser Dank geht an die fleißigen Autoren und die Leser, denen dieser Erfolg zu verdanken ist.

Wir wünschen Ihnen neue Erkenntnisse, Freude beim Lesen und viel Erfolg in den DQ-Projekten.

Die Herausgeber
Knut Hildebrand
Marcus Gebauer
Holger Hinrichs
Michael Mielke

Danksagung

Dank gebührt natürlich unseren vielen Co-Autoren, die mit ihrer weitreichenden Erfahrung nicht nur die IQM-Gemeinschaft bereichern, sondern nun auch dieses Buch ermöglicht haben. Allerdings wäre nichts möglich gewesen, wenn nicht unsere Familien geduldig auch dieses Projekt mitgetragen hätten. Sie müssen uns ohnehin schon mit häufig einnehmenden Arbeitgebern teilen. Die Leidenschaft für das Thema Daten- und Informationsqualität erfordert dann nur noch um so mehr Verständnis. Danke an Euch alle.

Knut Hildebrand sei an dieser Stelle ein besonderer Dank ausgesprochen. Ohne sein Antreiben wäre dieses Buch nicht möglich gewesen.

Inhaltsverzeichnis

Teil I Informationsqualität - Grundlagen

1 Was wissen wir über Information?	3
Florian Engelmann und Christoph Großmann	
1.1 Einleitung	3
1.2 Grundlegung	5
1.3 Information im Wissens- und Informationsmanagement	7
1.4 SHANNONsche Informationstheorie	11
1.5 STEINMÜLLERs Informationsmodell	13
1.5.1 STEINMÜLLERs System- und Prozessverständnis	14
1.5.2 Information als allgemeines Modell	14
1.5.3 Modell eines Informationssystems unter Einbezug der Semiotik ..	15
1.5.4 Fazit	16
1.6 Information als Produktionsfaktor	17
1.6.1 Perspektive der Produktionstheorie	17
1.6.2 Produktionsfaktor Information	20
1.6.3 Fazit	21
1.7 Zusammenfassung des Beitrages	21
Literatur	21
2 Informationsqualität – Definitionen, Dimensionen und Begriffe	25
Jan P. Rohweder, Gerhard Kasten, Dirk Malzahn, Andrea Piro und Joachim Schmid	
2.1 Einleitung	25
2.2 IQ-Dimensionen und Definitionen	27
2.2.1 Die 15 IQ-Dimensionen im Überblick	28
2.2.2 Graphische Darstellung der 15 IQ-Dimensionen und 4 IQ-Kategorien	29
2.2.3 Die 15 IQ-Dimensionen: Definitionen und Beispiele	31
2.3 Zusammenfassung und Ausblick	44
Literatur	44

Teil II Methoden – Techniken – Tools – Regelwerke/Standards

3 Datenqualitätsmetriken für ein ökonomisch orientiertes Qualitätsmanagement	49
Bernd Heinrich und Mathias Klier	
3.1 Einleitung	49
3.2 Anforderungen an Datenqualitätsmetriken	51
3.3 Bisherige Beiträge zur Messung von Datenqualität	52
3.4 Metriken und Messverfahren für DQ	54
3.4.1 Metrik für die DQ-Dimension Vollständigkeit	54
3.4.2 Metrik für die DQ-Dimension Fehlerfreiheit	57
3.4.3 Metrik für die DQ-Dimension Konsistenz	60
3.4.4 Metrik für die DQ-Dimension Aktualität	62
3.5 Praktische Anwendung der Metrik für Aktualität	63
3.6 Zusammenfassung und Ausblick	65
Literatur	65
4 Datenqualitätsmanagement – Steigerung der Datenqualität mit Methode ...	69
Niels Weigel	
4.1 Die Bedeutung des Total Data Quality Management	70
4.1.1 Vorgehensmodelle	70
4.1.2 Datenqualitätsmanagement sichert Ihren Unternehmenserfolg	71
4.2 Phasen eines ganzheitlichen Datenqualitätsmanagements	71
4.2.1 Initiierung des Datenqualitätsprojekts	72
4.2.2 Definition der Datenqualitätsanforderungen	75
4.2.3 Messung der vorhandenen Datenqualität	77
4.2.4 Analyse der Fehlerursachen	81
4.2.5 Verbesserung der Datenqualität	82
4.2.6 Permanente Überwachung der Datenqualität	84
4.3 Anreize für ein Datenqualitätsmanagement	84
Literatur	85
5 Strukturierte Datenanalyse, Profiling und Geschäftsregeln	87
Marcus Gebauer und Ulrich Windheuser	
5.1 Datenqualität	87
5.2 Merkmale der Datenqualität	89
5.3 Geschäftsregeln	92
5.4 Methoden der Datenanalyse	93
5.5 Metriken im Detail	95
5.6 Datenqualität in der Anwendung	96
Literatur	99

6	Datenbereinigung zielgerichtet eingesetzt zur permanenten Datenqualitätssteigerung	101
	Marcus Zwirner	
6.1	Definition „Datenbereinigung“	102
6.2	Ursachenanalyse	103
6.3	Bewertungskriterien für Datenfehler und Korrekturmaßnahmen	103
6.4	Methoden des Datenqualitätsmanagements	107
6.5	Datenqualitätsmaßnahmen im Detail	109
6.6	Zusammenfassung	120
7	Datenintegration und Deduplizierung	121
	Jens Bleiholder und Joachim Schmid	
7.1	Schritt 1: Schema Matching	124
7.2	Schritt 2: Dublettenerkennung	127
7.2.1	Auswirkungen von Dubletten	129
7.2.2	Entstehung von Dubletten	130
7.2.3	Erkennen von Dubletten	130
7.2.4	Durchführung der Dublettenerkennung	132
7.3	Schritt 3: Datenfusion	133
7.3.1	Konflikte ignorieren	135
7.3.2	Konflikte vermeiden	135
7.3.3	Konflikte auflösen	136
7.4	Erweiterungen	137
7.4.1	Strukturierung	137
7.4.2	Standardisierung	138
7.5	Zusammenfassung	139
	Literatur	139
8	Definition von Datenarten zur konsistenten Kommunikation im Unternehmen	141
	Andrea Piro und Marcus Gebauer	
8.1	Einleitung und Zielsetzung	141
8.1.1	Informationsqualität und Datenarten	142
8.2	Datenarten in der Informationslandschaft	143
8.3	Beschreibungskriterien	143
8.3.1	Beschreibung der Eigenschaften	144
8.3.2	Beschreibung des Kontextes	147
8.4	Beispiele für den Praxiseinsatz	148
8.4.1	Analyseebenen der Informationsqualität	148
8.4.2	Visualisierung des IQ-Status	149
8.5	Zusammenfassung	153
	Literatur	153

9 Suchmaschinen und Informationsqualität: Status quo, Problemfelder, Entwicklungstendenzen	155
Christian Maaß und Gernot Gräfe	
9.1 Ausgangssituation	155
9.2 Charakterisierung algorithmenbasierter Suchmaschinen	157
9.2.1 Funktionsweise algorithmenbasierter Suchmaschinen	157
9.2.2 Anfrageabhängige Ranking-Faktoren	157
9.2.3 Anfrageunabhängige Ranking-Kriterien	159
9.3 Semantisches Web und semantische Suchmaschinen	161
9.3.1 Vision und Grundlagen des semantischen Webs	161
9.3.2 Technische Grundlagen des semantischen Web	164
9.3.3 Problemfelder und Herausforderungen im Bereich der semantischen Suche	164
9.4 Fazit und Ausblick	166
Literatur	167
10 Bedeutung der Informationsqualität bei Kaufentscheidungen im Internet ...	169
Gernot Gräfe und Christian Maaß	
10.1 Einleitung	169
10.2 Informationsqualität in Entscheidungsprozessen	170
10.2.1 Informationen und Kaufentscheidungen	170
10.2.2 Informationsqualitätskriterien	172
10.3 Ursachen mangelnder Informationsqualität im Internet	174
10.3.1 Opportunistische Verhaltensspielräume der Anbieter	174
10.3.2 Informationsqualität aus der Nachfragerperspektive	177
10.3.3 Gründe für Opportunismus im Internet	180
10.4 Fazit und Handlungsempfehlungen	185
Literatur	188
11 Datenqualitäts-Audits in Projekten	193
Marcus Gebauer und Michael Mielke	
11.1 Einleitung	193
11.2 Abstimmung mit anderen Regelwerken	195
11.3 Glossar	195
11.4 Gebrauch der Generischen Checkliste	196
11.5 Datenqualitätsbewertung einer Datensammlung	198
11.5.1 Anforderungen an das Management	198
11.5.2 Service Level Agreements	199
11.5.3 Organisatorische Spezifikationen	199
11.5.4 Prozess-Definitionen	200

11.5.5 Datensammlung, Datenverarbeitung und Datennutzung	202
11.5.6 Messung, Maßnahmen und Überwachung	203
11.5.7 Technische Anforderungen	204
11.5.8 Dokumentation	206
11.6 Zusammenfassung	206
12 Bewertung der Informationsqualität im Enterprise 2.0	209
Sven Ahlheid, Gernot Gräfe, Alexander Krebs und Dirk Schuster	
12.1 Einführung	209
12.2 Beurteilung der Informationsqualität einer Enterprise 2.0	
Wissensplattform mittels eines hybriden Ansatzes	210
12.2.1 Automatische Beurteilung der Informationsqualität	211
12.2.2 Implizites Nutzer-Feedback	213
12.2.3 Explizites Nutzer-Feedback	214
12.2.4 Zusammenwirken der drei Ansätze und Fazit	216
Literatur	216
 Teil III Organisation	
13 Organisatorische Ansiedlung eines Datenqualitätsmanagements	223
Jens Lüssem	
13.1 Einführung	223
13.1.1 Motivation	223
13.1.2 Gliederung des Kapitels	224
13.2 Datenqualitätsmanagement – Entwicklungsstufen und Aufgaben	225
13.2.1 Sicherung der Datenqualität	226
13.2.2 Management der Datenqualität	226
13.3 Datenqualitätsmanagement – Ansiedlung im Unternehmen	227
13.3.1 Kopplung von Datenqualitätsmanagement mit anderen Unternehmensbereichen	227
13.3.2 Folgerungen für die Ansiedlung eines Datenqualitätsmanagements	229
13.4 Datenqualitätsmanagement in Projekten	230
13.4.1 Aufgaben des Datenqualitätsmanagements in Projekten	231
13.4.2 Organisatorische Verankerung des Datenqualitätsmanagements in Projekten	231
13.5 Zusammenfassung und Ausblick	232
13.5.1 Zusammenfassung	232
13.5.2 Ausblick	233
Literatur	233

14 Organisatorische Maßnahmen für gute Datenqualität	235
Jürg Wolf	
14.1 Messungen, Ursachen und generische Ansätze	235
14.1.1 Möglichen Arten von Datenqualitätsmängeln	235
14.1.2 Datenqualitätsmängel – Entstehung und Bekämpfung	236
14.1.3 Vier Generische Ansätze	237
14.1.4 Aus den generischen Ansätzen abgeleitete Strategien	238
14.2 Strategie A: Transparenz schafft Vertrauen	239
14.2.1 Ansatzpunkt dieser Strategie	239
14.2.2 Nutzen dieser Strategie	240
14.2.3 Nachteile und Risiken dieser Strategie	240
14.3 Strategie B: Definition von Verantwortlichkeiten	240
14.3.1 Ansatzpunkt dieser Strategie	240
14.3.2 Positionierung dieser Businessrollen im Modell	240
14.3.3 Nutzen dieser Strategie	242
14.3.4 Nachteile und Risiken dieser Strategie	242
14.4 Strategie C: gezielt Abhängigkeiten suchen	243
14.4.1 Ansatzpunkt dieser Strategie	243
14.4.2 Gezielte Definition von Master und Slave	243
14.4.3 Nutzen dieser Strategie	244
14.4.4 Nachteile und Risiken dieser Strategie	244
14.5 Strategie D: Daten-Lifecycle auf Basis des Prozesses	245
14.5.1 Ansatzpunkt dieser Strategie	245
14.5.2 Der Prozess und Lebenszyklus	245
14.5.3 Nutzen dieser Strategie	246
14.5.4 Nachteile und Risiken dieser Strategie	246
14.6 Strategie E: Niederschwellige Verbesserungs-Werkzeuge	247
14.6.1 Ansatzpunkt dieser Strategie	247
14.6.2 Beispiel eines niederschwelligen Verbesserungs-Werkzeuges	247
14.6.3 Die Infrastruktur dieses Werkzeuges	248
14.6.4 Nutzen dieser Strategie	249
14.6.5 Nachteile und Risiken dieser Strategie	249
14.7 Vor- und Nachteile aller erwähnter Strategien	249
14.7.1 Der Prozess ist die Vorgabe	249
14.7.2 Das Saatkorn ist der Beginn	249
14.7.3 Komplexität des Systems und Datenvolumen	250
14.8 Vorgehen bei der Umsetzung dieser Strategien	250
14.8.1 Kontakt zwischen den Parteien	250
14.8.2 Management-Unterstützung	250
14.9 Schlussfolgerungen und Ausblick	250
Literatur	251

15 Informationsmanagementprozesse im Unternehmen	253
Klaus Schwinn	
15.1 Motivation	253
15.2 Ausgangslage	254
15.3 Bewertung	255
15.4 Informationsmanagementprozess	257
15.5 Schema einer Informationsplanung	259
15.6 Datenlandkarte und Datenarchitektur	260
15.7 Geschäftsprozesse und Informationsmanagementprozess	261
15.8 Qualitätsaspekte	263
15.9 Ökonomische Aspekte	266
15.10 Zusammenfassung	267
Literatur	268
16 Data Governance	269
Boris Otto und Kristin Weber	
16.1 Einführung	269
16.2 Stand der Wissenschaft und Praxis	271
16.2.1 Abgrenzung des DQM	271
16.2.2 Bedeutung des Governance-Begriffs	271
16.2.3 Data Governance	272
16.2.4 Grundmuster für IT Governance	273
16.2.5 Einflussfaktoren auf IT Governance	274
16.3 Ein Modell für Data Governance	276
16.3.1 Rollen	276
16.3.2 Aufgaben	277
16.3.3 Zuständigkeiten	280
16.3.4 Einflussfaktoren und Gestaltungsparameter	281
16.4 Zusammenfassung	284
Literatur	285
17 IQM-Reifegradmodell für die Bewertung und Verbesserung des Information Lifecycle Management Prozesses	287
Saša Baškarada, Marcus Gebauer, Andy Koronios und Jing Gao	
17.1 Einleitung	287
17.2 Hintergrund	288
17.2.1 Total Quality Management	288
17.2.2 QM-Reifegrad	289
17.2.3 Information Quality Management	290
17.2.4 Existierende IQM-Reifegradmodelle	291

17.3	Methodologie	291
17.3.1	Die Delphi-Methode	292
17.4	IQM-Reifegradmodell	292
17.4.1	Chaotisch	293
17.4.2	Reaktiv	294
17.4.3	Messend	294
17.4.4	Steuernd	295
17.4.5	Optimierend	296
17.5	Zusammenfassung und Ausblick	296
	Literatur	296
18	Master Data Life Cycle – Stammdatenprozesse am Beispiel	
	Materialstamm in SAP ERP	301
	Knut Hildebrand	
18.1	Stammdaten – die Basis der Informationssysteme	301
18.2	Stammdatenqualität führt zu Prozessqualität	303
18.2.1	Qualitätseigenschaften	303
18.2.2	Probleme der Datenqualität und ihre Auswirkungen	303
18.3	Master Data Life Cycle (MDLC) – der Stammdatenprozess	304
18.3.1	Statuskonzept	304
18.3.2	Hindernisse und Problemfälle	308
18.3.3	Tools der SAP	309
18.4	Implementierung des MDLC	309
	Literatur	311
19	Prinzipien erfolgreichen Informationsqualitätsmanagements im Lichte	
	von Industrie 4.0	313
	Michael Mielke	
19.1	Big Data = Big Data Quality?	313
19.2	Datenqualität und Industrie 4.0	315
19.3	Übergeordnete Grundsätze und Einordnung der IQM-Grundsätze	317
19.4	Verantwortung für die Daten übernehmen	320
19.5	Agile DQ-Entwicklung	321
Teil IV Praxisbeispiele		
20	Ein Entscheidungsmodell zur Weitergabe persönlicher Daten im Internet	327
	Horst Treiblmaier	
20.1	Einleitung	327
20.2	Entscheidungsmodell	329
20.2.1	Intention	331
20.2.2	Nutzen	331

20.2.3	Vertrauen	333
20.2.4	Datenarten	336
20.2.5	Eingabefehler	338
20.3	Ausblick	338
	Literatur	339
21	Einführung eines proaktiven DQ-Managements	341
	Volker Landt	
21.1	Die Bremer Landesbank	341
21.1.1	Der Auftrag	342
21.2	Proaktives Datenqualitätsmanagement	342
21.3	Datenqualitätsorganisation	344
21.4	Eskalationsinstanz	345
21.5	Reporting	345
21.6	Messung von Datenqualität	346
21.6.1	Die Themen	346
21.6.2	Messsysteme	350
21.6.3	Messung bankfachlicher Datenzusammenhänge	351
21.7	Visualisierung der Messergebnisse	352
21.8	Messergebnisse und Fazit	356
22	Informationsqualität für das Management mit TOPAS®	359
	Otto-Ernst Brust, Frank Möller und Werner Skrablies	
22.1	Informationsmanagement für Führungskräfte	359
22.1.1	Qualitätskriterien von Informationen für das Management	360
22.1.2	Absicherung der Informationsqualitätskriterien: TOPAS®-Methodik	361
22.2	TOPAS®: Methode und Modell	361
22.2.1	Geschäftsprozessmanagement (GPM) mit der TOPAS®-Methode ...	361
22.2.2	Regelkreis für das Geschäftsprozessmanagement	362
22.2.3	4-Ebenen-Modell (Business Excellence Kriterien)	363
22.3	Anwendung: Management von Informationen und Daten	366
22.3.1	Informationen über Strukturen und Abläufe	366
22.3.2	Informationen und Daten für die strategische und operative Planung	367
22.4	Informationserfassung	369
22.4.1	Quellen zur Informationserfassung	370
22.4.2	Anforderungen zur Sicherung der Informations- und Datenqualität	370
22.5	Informationsverarbeitung	370
22.5.1	Operative Planung, Ziele	371

22.5.2	Festlegung von Kennzahlenstrukturen (KPI) in der Balanced Scorecard	371
22.5.3	Sicherung der Informations- und Datenqualität: Kennzahlensteckbrief	372
22.5.4	Erfolgskontrolle via KPIs: Reviews und Audits	373
22.6	Informationsdarstellung	373
22.6.1	Standard-Reporting mit dem TOPAS®-Prozessmodell	373
22.6.2	IT-Portal für das Management: Business Cockpit	374
22.6.3	Effizienzsicherung der Prozess-Architektur	374
22.6.4	Identifizierung von Korrekturmaßnahmen und Kontinuierlicher Verbesserungsprozess (KVP)	375
22.7	Wirkungsspektrum von TOPAS®	376
	Literatur	377
23	Datenqualitäts-Modell der Volkswagen Financial Services AG	379
	Helena Moser	
23.1	Einleitung	379
23.2	Das Projekt „Datenqualität Strukturen/Standards und Drittmarktfähigkeit“	381
23.2.1	Warum ist Datenqualität nötig?	381
23.2.2	Projektauftrag	381
23.2.3	Projektziel	382
23.2.4	Ist-Analyse	382
23.2.5	Prozessanalyse	383
23.2.6	Sollkonzept	384
23.2.7	Das Datenqualitäts-Modell und deren Zuständigkeiten (Abb. 23.1)	386
23.2.8	Das Datenqualitäts-Modell und sein Regelwerk	390
23.2.9	Monitoring/Reports	391
23.2.10	Realisierungs- und Einführungsphase	392
23.3	Fazit	393
24	Verknüpfung von DQ-Indikatoren mit KPIs und Auswirkungen auf das Return on Investment	395
	Frank Block	
24.1	Beispiele zur Illustration von DQ-Problemen	396
24.2	Wie wirken sich DQ-Probleme auf Unternehmen aus – Der Zusammenhang zwischen Daten- und Prozessqualität	397
24.2.1	Beispiel – Call Center	398
24.2.2	Beispiel – Kundenbeziehungsmanagement (CRM)	398
24.2.3	Beispiel – Data Mining Prozess im Marketingumfeld	399
24.2.4	Beispiel – Direktmailprozess	400

24.3	Wie viel kosten schlechte Daten den Unternehmer?	402
24.4	Der Einfluss von DQ-Indikatoren auf KP-Indikatoren – wie beeinflusst Datenqualität den Unternehmenserfolg?	403
24.5	Beschreibung eines KPI orientierten DQ-Managementprozesses	405
24.5.1	Phase 1 – Selektiere zu untersuchende Komponenten	406
24.5.2	Phase 2 – Mitarbeiterbefragung	407
24.5.3	Phase 3 – DQ-Assessment	409
24.5.4	Phase 4 – Validieren und Quantifizieren	410
24.5.5	Phase 5 – DQ-Projekte definieren, Korrekturmaßnahmen durchführen	411
24.6	Fallstudie – Banque Cantonale Vaudoise (BCV)	413
	Literatur	415
	Stichwortverzeichnis	417

Teil I
Informationsqualität - Grundlagen

Florian Engelmann und Christoph Großmann

1.1 Einleitung

Der Begriff Information stammt aus dem Lateinischen. Von der Antike bis in das 20. Jahrhundert war Information ein kaum verwendetes und unspektakuläres Wort. Zemanek belegt diese Hypothese mit dem Verweis auf ein Häufigkeitswörterbuch der deutschen Sprache aus dem Jahre 1897. Dort kommt in ca. 11 Mio. Wörtern nur 55 mal das Wort Information vor (Zemanek 1986, S. 19).

Mit Sicherheit kann man behaupten, dass der Informationsbegriff heute sowohl in der Alltagssprache als auch in der Wissenschaftssprache häufig Verwendung findet. Leben wir nicht im Informationszeitalter bzw. in einer Informationsgesellschaft? Könnte also der Information die Eigenschaft eines fächervereinenden, interdisziplinären Begriffs zukommen?

Während in der Umgangssprache Unschärfen der Normalfall sind, muss ein wissenschaftliches Begriffssystem möglichst präzise sein. In der betriebswirtschaftlichen Literatur spricht man von Information als Schlüsselgröße (Grochla und Wittmann 1993, S. 1725 ff.), Schlüsselbegriff (Lehner und Maier 1994, S. 1), kritischer Erfolgsfaktor im Wettbewerb (Pietsch et al. 2004, S. 17) oder als Produktions- und Wettbewerbsfaktor (Zie-

F. Engelmann (✉)

Ernst & Young GmbH Wirtschaftsprüfungsgesellschaft, Börsenplatz 1, 50677 Köln, Deutschland

E-Mail: florian.engelmann@email.de

C. Großmann

Oppenweiler, Deutschland

E-Mail: christophgrossmann@gmx.de

© Springer Fachmedien Wiesbaden 2015

K. Hildebrand et al., (Hrsg.), *Daten- und Informationsqualität*,

DOI 10.1007/978-3-658-09214-6_1

genbein 2004, S. 21). Je mehr wirtschaftswissenschaftliche Definitionen man aber überprüft, umso mehr Vorschläge erhält man. Häufig verwenden Autoren Theorien aus ganz anderen Wissenschaftszweigen, definieren Information für eigene Zwecke neu oder lassen Information einfach undefiniert.

In der Wirtschaftswissenschaft verfolgt man bestimmte Zielsetzungen mit Information, wie eine exemplarische Aufzählung wichtiger Perspektiven zeigt:

- produktionswirtschaftliche Sicht (Information als Produktionsfaktor, Zwischen- oder Endprodukt)
- entscheidungstheoretische Sicht (Information zur Vorbereitung von Entscheidungen und Handlungen)
- strategische Sicht (exklusiver Gebrauch von Information um einen Vorsprung zu erlangen)
- Sicht der neuen Institutionenökonomie (berücksichtigt u. a. asymmetrische Informationen, veränderbares Wissen, beschränkte Rationalität und Opportunismus).

Es gibt also einen latenten Widerspruch in den Anforderungen, wenn ein Begriff universell und präzise zu gleich sein müsste. Falls Information aber unbestimmt oder mehrdeutig sein sollte, welches Ziel wird dann im Rahmen eines Informationsqualitätsmanagements verfolgt? Dies bedarf einer Klärung.

Vor dem Hintergrund einer solchen Problematik gibt es mehrere Versuche, innerhalb der Betriebswirtschaftslehre den Begriff allgemeingültig zu fassen. Beispielsweise hat Bode eine Typologie entwickelt, um die unterschiedlichen Ansätze zu strukturieren. Hierzu betrachtet er fünf Dimensionen (Tab. 1.1):

Analysiert man Bodes Vorschlag, dann stellt sich die Frage, ob er mit seinen Dimensionen eine Grundlage für einen allgemeingültigen Informationsbegriff geschaffen hat. Das es wesentlich genauer geht, zeigen die 16 Dimensionen der Informationsqualität.

Ein wissenschaftliches oder praktisches Problem folgt aus unterschiedlichen Definitionen von Information zunächst nicht, obwohl überschneidende Definitionen zu Kommunikationsproblemen führen können. Eine interdisziplinäre Zusammenarbeit ist momentan also nur schwer möglich. Genauer betrachtet haben die Differenzen tiefergehende Ursachen und sind methodischer Natur, weil aktuelle Problemstellungen ein fächerübergreifendes Verständnis verlangen, kollidieren ganz unterschiedliche wissenschaftliche Methoden.

Tab. 1.1 Typologie der Informationsbegriffe in Anlehnung an Bode. (Bode 1997, S. 452)

Dimensionen	Ausprägungen		
Semiotik	Syntaktisch	Semantisch	Pragmatisch
Träger	Ungebunden		Menschengebunden
Neuheitsgrad	Subjektiv		Objektiv
Wahrheitsgehalt	Wahrheitsabhängig		Wahrheitsunabhängig
Zeitbezogenheit	Statisch		Prozessual

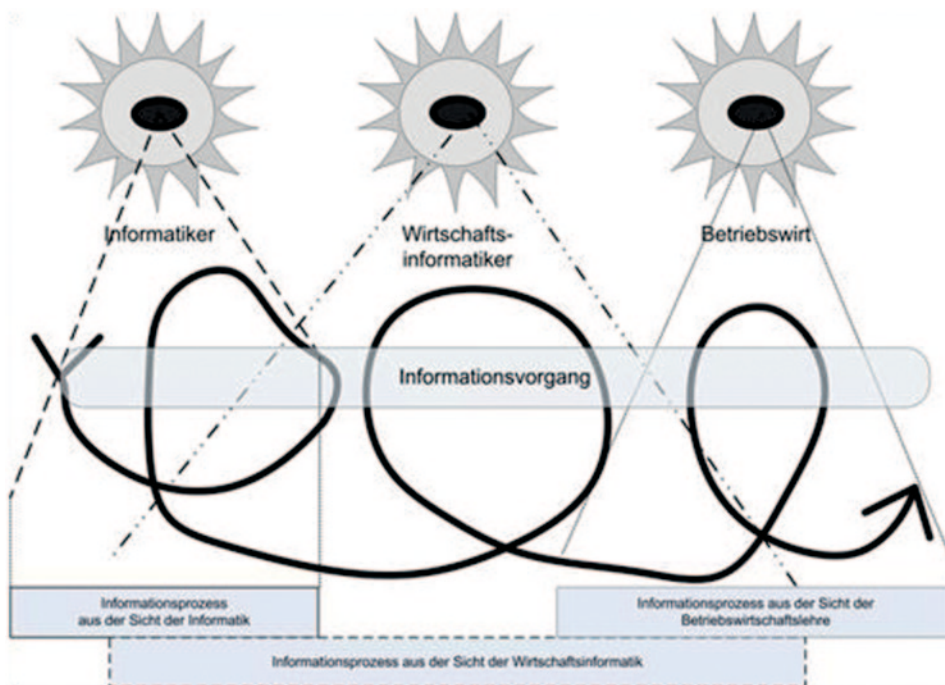


Abb. 1.1 Information interdisziplinär analysiert. (Steinmüller 1993, S. 224)

Nach einer Grundlegung über Information und Informationsmanagement sollen hier eine quantitative Methode (Kap. 4) aus der Nachrichtentechnik, eine modelltheoretische Betrachtung (Kap. 5) aus der Informatik und eine klassifikatorische Methode (Kap. 6) aus der Betriebswirtschaftslehre im Zusammenhang mit Information vorgestellt werden (Abb. 1.1).

1.2 Grundlegung

Eine weitverbreitete Vorstellung über den Aufbau von Information zeigt die Informations- oder Wissenspyramide. Dieses deskriptive Modell verwendet aufeinander aufbauende Ebenen. Information ist das mit Bedeutung versehene Datum. Nach dieser Auffassung lässt sich die höhere Ebene somit auf die Bestandteile der unteren Ebenen der Daten oder Zeichen reduzieren (Abb. 1.2).

Die Pyramide verschweigt aber ein Problem, das vor allem im Bereich der Kennzahleninterpretation bekannt ist: Wie kann einem Datum eine allgemeingültige Bedeutung von einem Sender zugewiesen werden? Und wenn das geschehen ist, wie kann dann diese Bedeutung von allen Empfängern gleichermaßen verstanden werden? Dies wäre nur dann eindeutig möglich, wenn die Information vorher allen bekannt ist. Das ist offensichtlich

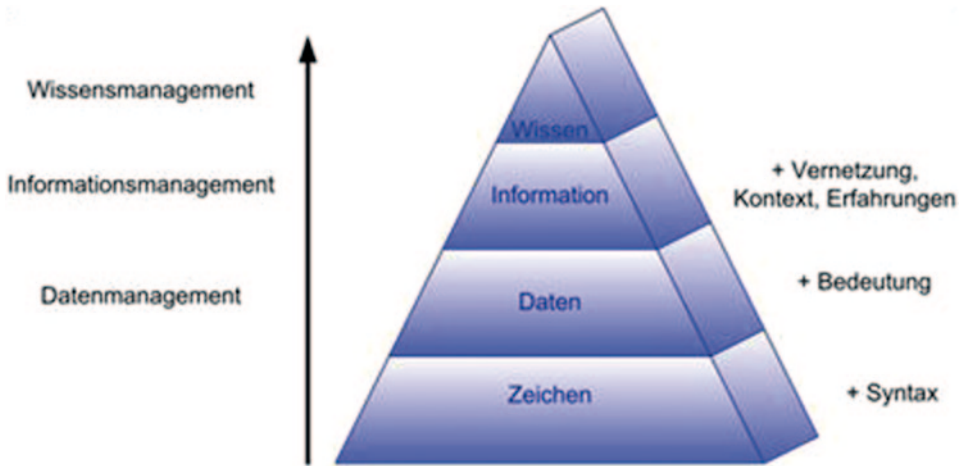


Abb. 1.2 Wissenspyramide. (Forst 1998, S. 1)

paradox, da der Zugang der Information in einen Fall keine Neuigkeit liefert und im anderen Fall nicht verständlich ist.

Weiter wird in der Pyramide der qualitative Unterschied zwischen speicherbaren Daten und verstandenen Informationen bzw. Wissen nicht deutlich. Während die Zeichen oder Daten auf einem materiellen Träger eindeutig verifiziert werden können, ist es zumindest in der Wissenschaft strittig, ob im Gehirn äquivalente Prozesse bei der „Speicherung“ ablaufen. Die neuere Gehirnforschung zeigt nämlich, dass das Gehirn einem Computer überhaupt nicht ähnlich ist (Brodbeck 2007). Folglich ist zumindest die höchste Ebene „Wissen“ nicht auf ein abgespeichertes Konglomerat von Zeichen, Daten oder Informationen reduzierbar.

In der klassischen Zeichenlehre wird dieser qualitative Unterschied zwischen „Träger“ und „Getragenen“ klarer herausgearbeitet. An diesen beiden elementaren Kategorien jeder Information werden die Kernprobleme deutlich, die durch die Wissenspyramide nicht erfasst worden sind (Abb. 1.3):

Unter „Träger“ versteht man die materielle Gestalt der Information und das „Getragene“ ist etwas immaterielles, das mit den Gedanken eines Menschen zu tun hat. Dies bedeu-

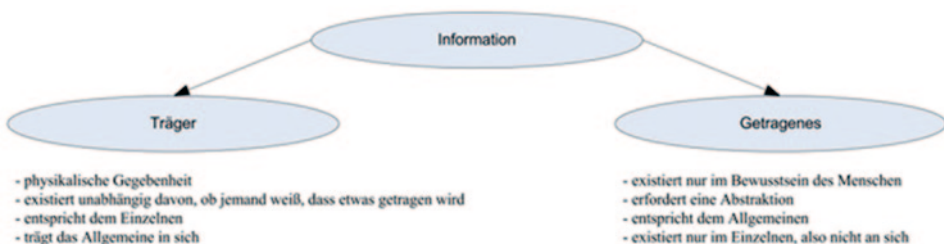


Abb. 1.3 Analyse von Information in Anlehnung an Völz. (Völz 1983, S. 228)

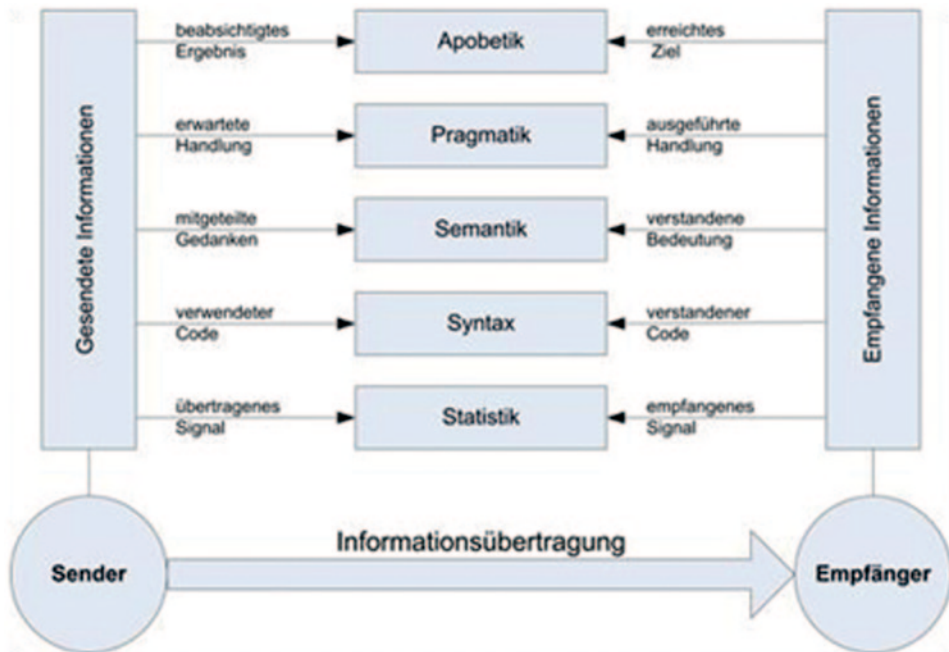


Abb. 1.4 Modell der Informationsübertragung. (Gitt 2002, S. 144)

tet jedoch, dass an einer Stelle der Wissenspyramide ein Übergang zu einem grundlegend anderen Verarbeitungssystem stattfinden muss.

Ein handlungsbezogenes Modell von Information muss Hintergrundinformation, vor allem über den Akteur und dessen Ziele mit einschließen. Den Betriebswirt interessieren primär die Zielsetzungen, während sich ein Informatiker wohl zuerst mit Codierungsaspekten beschäftigt. Ein Informationsmodell das beide Anforderungen integriert und den prozessualen Charakter von Information betont, sieht wie folgt aus (Abb. 1.4):

1.3 Information im Wissens- und Informationsmanagement

Information als zweckorientiertes Wissen hat in der Betriebswirtschaftslehre schon immer eine wichtige Rolle gespielt (Wittmann 1959, S. 14). Mit den gestiegenen technischen Möglichkeiten wächst aber auch der Koordinationsbedarf für die Beschaffung, Verarbeitung, Übertragung, Speicherung und Bereitstellung von Informationen. Kurz gesagt: Man benötigt ein unternehmensweites Informationsmanagement, weil das Wissen in den Köpfen mehr als die Summe der gespeicherten Daten ist, kann man das Wissensmanagement als eine Erweiterung des Informationsmanagements ansehen. Die Anforderungen einer wissensorientierten Unternehmensführung weichen aber konzeptionell nicht weit von denen eines informationsorientierten Managements ab. Nach Meinung von North handelt es sich bei Informationen um einen Rohstoff für Wissen. Wissen wird quasi aus Informationen

„generiert“ und „entsteht als Ergebnis der Verarbeitung von Informationen durch das Bewußtsein“ (North 2005, S. 33).

Die Erweiterung auf Wissen bedeutet zunächst nur eine Vergrößerung des Gegenstandsbereiches. Wenn der Kern erhalten bleibt, dann sind in beiden Fällen die Anforderungen ähnlich. Im Zentrum beider Konzepte steht zukünftig nicht mehr die Informationstechnologie, sondern das Benutzen der Technologien, um das Wissen der Mitarbeiter zu erweitern. Fundamental ist vor allem die Funktion von Information als verbindendes Kommunikationsinstrument in kollegialen Netzwerken. So verstanden geht es nicht mehr um einen exklusiven Zugriff auf Informationen, sondern um die Frage der effizienten und effektiven Kommunikation untereinander. Die strategischen Ziele fokussieren die praktischen Handlungen, die bei gekonnter Umsetzung der Ressource Wissen zu strategischen Wettbewerbsvorteilen führen (Abb. 1.5).

Idealtypische Konzepte des Informationsmanagements arbeiten unter anderem mit der Annahme, dass es optimale und berechenbare Lösungen für Informationsprobleme gibt. Unter diesen Bedingungen lässt sich ein objektiver Informationsbedarf a priori formulieren. Dieser wird jedoch niemals vom Benutzer (subjektiver Informationsnachfrage) genau nachgefragt, noch vom Betreiber des Informationsangebots exakt zur Verfügung gestellt. Der objektive Informationsstand vor Beginn einer Handlung ergibt sich als Schnittmenge aus allen drei Mengen. Diesen Stand zu optimieren, ist die Aufgabe des Informationsmanagements (Abb. 1.6).

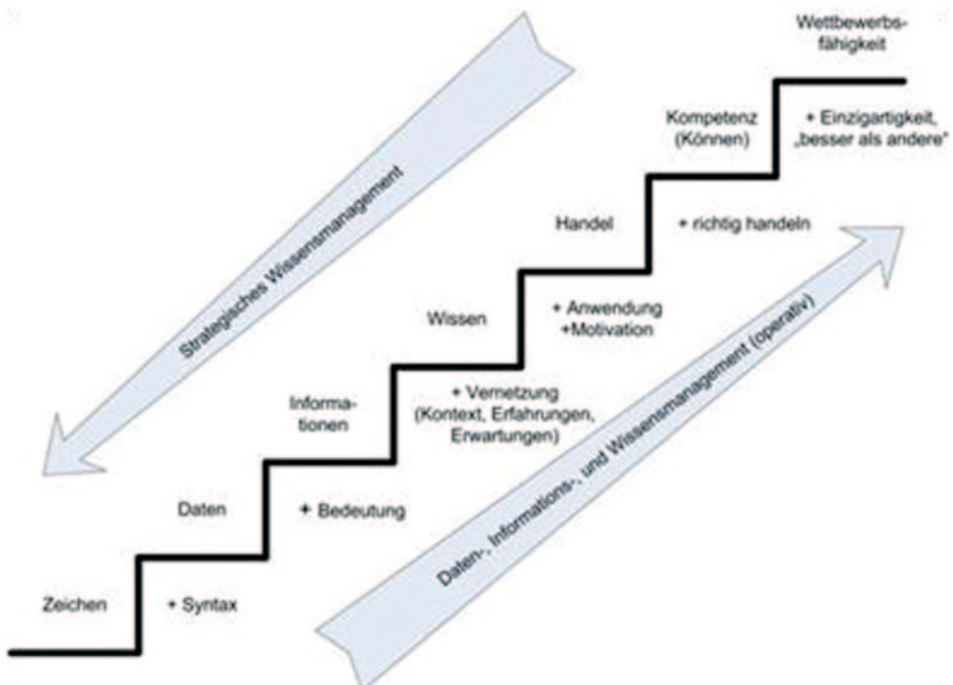


Abb. 1.5 Wissenstreppe. (North 2005, S. 32)

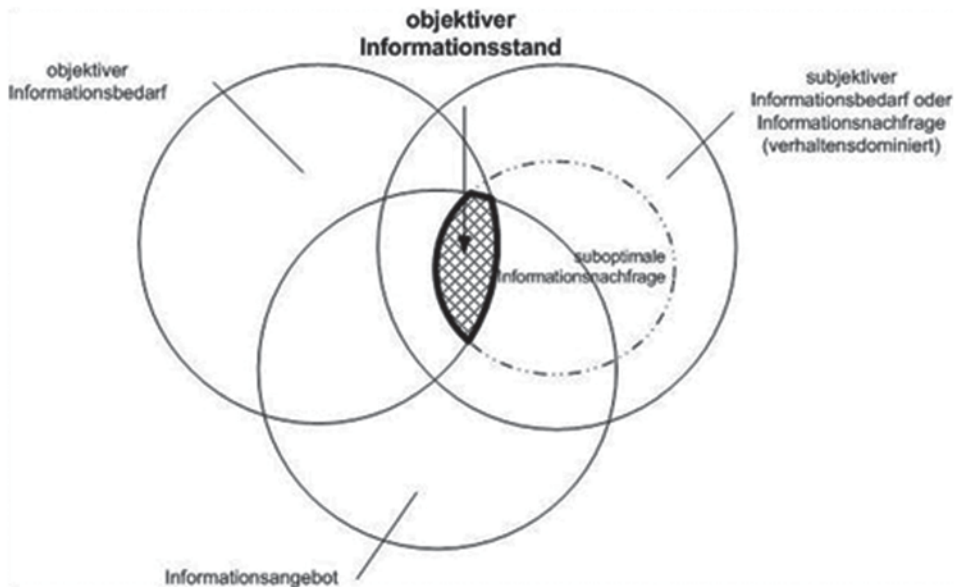


Abb. 1.6 Analyse des Informationsstandes. (Picot et al. 2003, S. 82)

Die Ziele der betrieblichen Informationswirtschaft lassen sich wie folgt zusammenfassen (Krcmar 2005, S. 51):

- Ausgleich von Informationsnachfrage und Informationsangebot,
- Versorgung der Entscheidungsträger mit relevanten Informationen,
- Gewährleistung einer hohen Informationsqualität,
- Dokumentation von Willensbildungs- und Willensdurchsetzungsprozess,
- Gestaltung der Informationswirtschaft als Querschnittsfunktion des Unternehmens,
- Einsatz von Informationstechnologie zur Unterstützung der informationswirtschaftlichen Aufgabenerfüllung,
- zeitliche Optimierung der Informationsflüsse,
- Beachtung des Wirtschaftlichkeitsprinzips.

Für die Zweckeignung von Information gibt es unterschiedliche Klassifikationen, die sich jedoch ähnlich sind. Exemplarisch gibt Eschenröder folgende Kriterien an (Abb. 1.7):

Für betriebswirtschaftliche Überlegungen ist also ein erweiterter Informationsbegriff notwendig, der eine Vielzahl von Anforderungen erfüllen muss, die im Bereich der Semantik und Pragmatik liegen. Eine Quantifizierung dieser Aspekte ist jedoch mit erheblichen Schwierigkeiten behaftet. Weizäcker spricht von Quantifizierungsgrenzen, weil sich die Bedeutung und Wirkung von Information nur in Bezug auf eine Situation und zeitlich erst nach erfolgter Information feststellen lässt. „Lediglich in den Fällen, wo öfter gleiche Situationen durch gleiche Informationen beeinflusst werden, läßt sich im voraus und ‚objektiver‘ die Bedeutung und Wirkung angeben“ (Weizsäcker und Maurin 1974, S. 82 f.).

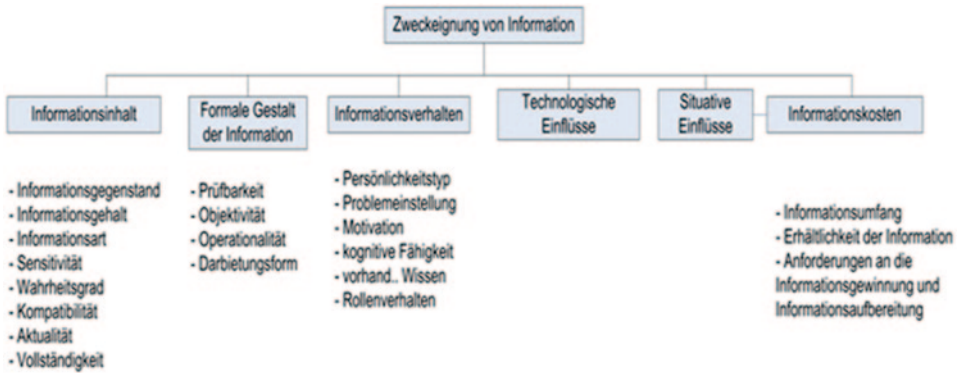


Abb. 1.7 Einflussfaktoren einer Zweckeignung von Information. (Eschenröder 1985, S. 49)

Diese Anforderungen relativieren also die Erfolgsaussichten der „objektiven Informationskonzepte“. Das Verstehen und die Zweckmäßigkeit des Inhalts einer Information hängt nämlich sowohl von der Eigenschaft ab, neu zu sein, als auch von der Eigenschaft bekannt zu sein. Weizsäcker fasst diese Eigenschaft von Information wie folgt zusammen: „Wenn wir nun annehmen, dass die Erstmaligkeit von 0 auf 100% ansteigen kann, und wenn wir postulieren, dass zur gleichen Zeit die Bestätigung von 100 auf 0% abnimmt, dann erhalten wir ein Kontinuum, das das relative Gewicht von Erstmaligkeit und Bestätigung von Situationen darstellt“ (Weizsäcker und Maurin 1974, S. 98). Der pragmatische Gehalt einer Information steigt nun zur Mitte hin an und fällt dann wieder ab, wenn sie zu häufig bestätigt wurde. Dieser Verlauf entspricht der Alltagserfahrung mit Informationen (Abb. 1.8).

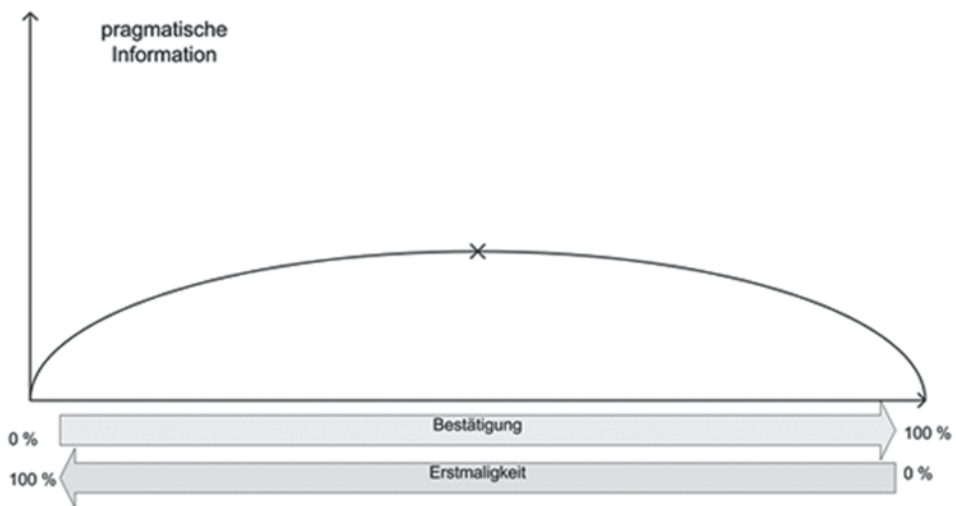


Abb. 1.8 Pragmatische Information. (Weizsäcker und Maurin 1974, S. 99; Picot et al.2003, S. 83)