

Hakim Harrach

Risiko-Assessments für Datenqualität

VIEWEG+TEUBNER RESEARCH

Ausgezeichnete Arbeiten zur Informationsqualität

Herausgeber:

Dr. Marcus Gebauer, Rüdiger Giebichenstein

Bewertungskommission des

Information Quality Best Master Degree Award 2009:

Prof. Dr. Holger Hinrichs, FH Lübeck (Kommissionsvorsitz)

Dr. Marcus Gebauer, Hannover Re AG und Vorsitzender der DGIQ

Prof. Dr. Knut Hildebrand, HS Darmstadt

Bernhard Kurpicz, OrgaTech GmbH

Prof. Dr. Jens Lüssem, FH Kiel

Michael Mielke, Deutsche Bahn AG und Präsident der DGIQ

Prof. Dr. Felix Naumann, HPI, Uni Potsdam

Prof. Dr. Ines Rossak, FH Erfurt

Die Deutsche Gesellschaft für Informations- und Datenqualität e.V. (DGIQ) fördert und unterstützt alle Aktivitäten zur Verbesserung der Informationsqualität in Gesellschaft, Wirtschaft, Wissenschaft und Verwaltung. Zu diesem Zweck befasst sie sich mit den Voraussetzungen und Folgen der Daten- und Informationsqualität. Sie fördert zudem durch Innovation und Ausbildung die Wettbewerbsfähigkeit der Unternehmen sowie die des unternehmerischen und akademischen Nachwuchses in Deutschland.

Die vorliegende Schriftenreihe präsentiert ausgezeichnete studentische Abschlussarbeiten in der Daten- und Informationsqualität. Veröffentlicht werden hierin die Siegerarbeiten des jährlich stattfindenden „Information Quality Best Master Degree Award“.

Hakim Harrach

Risiko-Assessments für Datenqualität

Konzept und Realisierung

Mit einem Geleitwort von Prof. Dr. Jens Lüssem

VIEWEG+TEUBNER RESEARCH

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

1. Auflage 2010

Alle Rechte vorbehalten

© Vieweg+Teubner Verlag | Springer Fachmedien Wiesbaden GmbH 2010

Lektorat: Ute Wrasmann | Anita Wilke

Vieweg+Teubner Verlag ist eine Marke von Springer Fachmedien.

Springer Fachmedien ist Teil der Fachverlagsgruppe Springer Science+Business Media.

www.viewegteubner.de



Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlags unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, 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.

Umschlaggestaltung: KünkelLopka Medienentwicklung, Heidelberg

Druck und buchbinderische Verarbeitung: STRAUSS GMBH, Mörlenbach

Gedruckt auf säurefreiem und chlorfrei gebleichtem Papier.

Printed in Germany

ISBN 978-3-8348-1344-2

Geleitwort

Hakim Harrach thematisiert in dem vorliegenden Buch die Nahtstelle zwischen Datenqualitätsmanagement und Risikomanagement – eine Datenstelle, die gerade in Zeiten knapper werdender Ressourcen stärker in den Fokus rücken wird.

Die Arbeit von Hakim Harrach liefert einen Überblick in die Gebiete des Datenqualitäts- und des Risikomanagements, um dann aufbauend Verbindungen zwischen diesen bislang häufig getrennt voneinander bearbeiteten Themen zu ziehen.

Es werden damit wesentliche Impulse für Risikomanager geliefert, die Risiken, die aus mangelhafter Datenqualität entstehen, in ihr Risikomanagementsystem integrieren wollen und müssen. Das Buch liefert aber auch Impulse für Datenmanager bzw. Datenqualitätsmanager, die nicht nur abschätzen wollen, welche Auswirkungen mit einer mangelhaften Datenqualität einhergehen, sondern auch aktiv entsprechende Risiken steuern wollen.

In der Arbeit wird ein standardisiertes Vorgehen für ein Risiko-Assessment mit Fokus auf Datenqualität vorgestellt, das in ein einfach anwendbares IQ-Assessment-Tool mündet.

Hakim Harrach stößt damit die Tür weit auf zu einer stärker risikoorientierten Betrachtung des Themas Datenqualität und leistet damit einen hervorragenden Beitrag zur Arbeit der Deutschen Gesellschaft für Informations- und Datenqualität e.V.

Prof. Dr. Jens Lüssem

Vorwort

Die Erstellung der Arbeit haben zahlreiche Personen ermöglicht und unterstützt, bei denen ich mich herzlich bedanke. Dabei gilt mein besonderer Dank an dieser Stelle Herrn Prof. Dr. Jens Lüsse, Institut für Informatik III der Universität Bonn, der die Arbeit von der Themenfindung bis zur Veröffentlichung betreute und mir die entsprechenden Möglichkeiten in einer Bank bot.

Bedanken möchte ich mich ebenfalls bei meinem Kommilitonen Pierre Schmitz, der mir mit seiner Zeit, den konstruktiven Diskussionen und Vorschlägen mit vollster Hilfsbereitschaft bei der Erstellung der vorliegenden Arbeit zur Seite stand.

Im Rahmen dieser Arbeit wurden zahlreiche Interviews mit Mitarbeitern einer Bank durchgeführt. Bei meinen zahlreichen Interviewpartnern möchte ich mich für das Interesse und die aufrichtige Beantwortung meiner nicht immer unkritischen Fragen bedanken.

Nicht zuletzt möchte ich mich bei meinen Eltern und Freunden für die entgegengebrachte Geduld und Unterstützung bedanken.

Hakim Harrach

Zusammenfassung

Die vorliegende Untersuchung diskutiert die Kohärenz zwischen Datenqualitätsmanagement und Risikomanagement unter dem Aspekt, ein Konzept zur Realisierung eines standardisierten Risiko-Assessments für Datenqualität zu entwickeln.

Ziel der Arbeit war es, ein Konzept für die Realisierung von automatisierten Risiko-Assessments zu erarbeiten und dessen Umsetzung anhand der Implementierung eines entwickelten Prototyps namens IQ-Assessment Tool in einer Bank zu dokumentieren.

Die Arbeit wird in vier Teile gegliedert. Im ersten Teil werden definitorische Grundlagen über Daten, Wissen, Information, Qualität, Datenqualität, Qualitätsmanagement, Datenqualitätsmanagement, Metadaten und Metadaten-Management vorgestellt. Weiterhin werden grundsätzliche Konzepte über das Risikomanagement in der IT sowie das Datenqualitätsmanagement beschrieben. Der zweite Teil der Arbeit präzisiert diese Konzepte dahingehend, dass Datenqualitätsprozesse in die Risikomanagementprozesse integriert werden. Hierbei werden die Schnittstellen für die Anwendung von Risiko-Assessments aufgezeigt und gleichwohl ein Konzept zur Realisierung der standardisierten Risiko-Assessments präsentiert. Im dritten Teil wird das erarbeitete Konzept durch die Implementation des Prototyps hinsichtlich des Anwendungsbeispiels Banken umgesetzt. Dabei wird eine modifizierte Variante des Prototyps eingeführt und detailliert beschrieben. Im letzten Teil der Arbeit werden die Ergebnisse abschließend diskutiert sowie mögliche Ansätze zur Weiterentwicklung des IQ-Assessment Tools aufgezeigt.

Abstract

In this thesis we will discuss the coherence between data quality management and risk management to develop a concept for standardized data risk assessments.

The goal of this thesis is to work out a concept for the realization of automated data risk assessments and to implement a prototype following this concept. The prototype in the bank is called IQ-Assessment Tool.

The thesis is divided into four parts. The first part introduces the subjects of data, knowledge, information, quality, data quality, quality management, data quality management, metadata and metadata management. Furthermore, we describe basic concepts for risk management in the IT and data quality management. The second part of this thesis is to formulate these concepts to integrate data quality processes into risk management processes. At this point we will illuminate the interfaces for an application data risk assessment and at the same time we describe a concept for implementing a standardized risk assessment for data quality. Part three describes the prototype that has been implemented in the bank based on the proposed data risk assessment concept. For the implementation the prototype has been modified. In the last part of this thesis we will finally discuss the resulting works and complete the thesis with an outlook on further technical and theoretical extensions for the IQ-Assessment Tool.

Inhaltsverzeichnis

Geleitwort	v
Vorwort	vii
Zusammenfassung	ix
Abstract	xi
Abbildungsverzeichnis	xvii
Tabellenverzeichnis	xix
Listingverzeichnis	xxi
Abkürzungsverzeichnis	xxiii
1. Einleitung und Motivation	1
1.1 Maßnahmen der Gesetzgebung	2
1.1.1 Sarbanes-Oxley Act	3
1.1.2 Data Quality Act	3
1.1.3 KonTraG	3
1.1.4 Basel II	4
1.2 Zielsetzung und Abgrenzung	4
1.2.1 Zielsetzung und Kernfragen	5
1.2.2 Abgrenzung der Arbeit	5
1.3 Gliederung und Inhalt der Arbeit	6
I Grundlagen	9
2. Grundlagen und Begriffe	11
2.1 Gesamtmodell: Daten, Wissen und Information	11
2.2 Qualität und Datenqualität	14
2.3 Qualitätsmanagement und Datenqualitätsmanagement	18
2.4 Metadaten und Metadaten-Management	21
2.5 Zusammenfassung	22

3. Risikomanagement in der IT	25
3.1 Begriffliche Grundlagen	26
3.2 Regelkreis des Risikomanagements.....	27
3.2.1 Risikoidentifikation	29
3.2.2 Risikoanalyse.....	30
3.2.3 Risikohandhabung	32
3.2.4 Risikoüberwachung	33
3.3 Risikomanagementprozess.....	34
3.4 Reifegradmodelle.....	35
3.4.1 Capability Maturity Model (CMM)	35
3.4.2 BOOTSTRAP.....	37
3.4.3 SPICE (ISO 15504).....	38
3.5 Zusammenfassung.....	39
4. Datenqualitätsmanagement	41
4.1 Data-Ownership-Konzept	41
4.2 PDCA-Zyklus	44
4.3 Datenqualität-Assessment.....	47
4.4 Ziele des Datenqualitätsmanagements in einem Unternehmen	51
4.4.1 Anforderungen an das Datenqualitätsmanagement.....	51
4.4.2 Funktionen des Datenqualitätsmanagements	52
4.5 Schlussbemerkung	53
II Konzept.....	55
5. Anwendung von Risikomanagement auf Datenqualität.....	57
5.1 Risikomanagement in Bezug auf die Datenqualität.....	57
5.2 Design des Datenqualität-Regelkreises	58
5.3 Risikoidentifizierung	62
5.4 Risikoanalyse	64
5.5 Risikohandhabung.....	65
5.6 Risikoüberwachung.....	67
5.7 Zusammenfassung.....	68

6. Ansatz für ein Gestaltungsmodell.....	69
6.1 Rahmenbedingungen	69
6.1.1 Ziele eines Risiko-Assessments	69
6.1.2 Zentrale Problemfelder	70
6.2 Datenqualität-Netzwerk.....	71
6.3 Realisierung eines Risiko-Assessments	73
6.3.1 Fragebogen an den Datennutzer	76
6.3.2 Fragebogen an den Datenerfasser	76
6.3.3 Fragebogen an den Datenverarbeiter	77
6.3.4 Fragebogen an den Datenprüfer.....	77
6.4 Auswertung des Konzepts	78
6.5 Empfehlungen für das weitere Vorgehen (Grundanalyse).....	81
6.6 Schlussbemerkung.....	82
III Umsetzung und Evaluation	85
7. Anwendung auf eine Bank.....	87
7.1 Einführung.....	87
7.2 Risikobewertungsmatrix.....	88
7.3 Applikation des Risiko-Assessments in einer Bank.....	91
7.4 Implementierung eines proaktiven Datenqualitätsmanagements.....	96
7.5 Zusammenfassung	100
8. Evaluation	103
8.1 Einleitung	103
8.2 Grundlagen und Theorien der Software-Evaluation	105
8.3 Durchführung	107
8.4 Ergebnisbewertung.....	108
8.5 Schlußbemerkung	112
IV Zusammenfassung und Ausblick.....	113
9. Zusammenfassung.....	115
9.1 Zusammenfassung der Ergebnisse	115
9.2 Abschließende Beurteilung	116

10. Ausblick.....	119
10.1 Weiterführende Arbeiten	119
10.2 Fazit.....	120
V Anhang	121
A. Technische Realisierung.....	123
A.1 Fragebogen.....	124
A.2 Detailergebnisse.....	124
B. Anwendungen.....	127
B.1 Implementation.....	128
B.1.1 Datentransport.....	128
B.1.2 Auswertung.....	128
B.2 Detailergebnisse	129
B.2.1 Logischer Ablauf eines Fragebogens	130
B.2.2 Hilfetexte	131
B.2.3 Benutzer-Identifikation.....	132
C. Evaluationsfragebogen	133
C.1 Fragebogen im Original	133
C.2 Ergebnisse der Befragung	140
Literaturverzeichnis.....	143
Index.....	149

Abbildungsverzeichnis

Abbildung 1.1: Anlässe zur Installierung eines Risikomanagement-Systems	2
Abbildung 1.2: Grundsätzlicher Aufbau der Arbeit	7
Abbildung 2.1: Daten, Information und Wissen.....	13
Abbildung 2.2: Kategorisierung der Datenqualität.....	17
Abbildung 2.3: Funktionen des Qualitätsmanagements	19
Abbildung 2.4: Kosten-Nutzen-Faktor der Datenqualität	20
Abbildung 3.1: Klassischer Regelkreislauf eines Risikomanagement-Systems	28
Abbildung 3.2: Übersicht der Methoden zur Risikoidentifikation	29
Abbildung 3.3: Übersicht über Risikobewältigungsstrategien	33
Abbildung 4.1: Die Rollensicht im Data-Ownership-Konzept.....	42
Abbildung 4.2: Die Aufgabenübersicht der verschiedenen Rollen im Data-Ownership-Konzept.....	44
Abbildung 4.3: Der PDCA-Zyklus	46
Abbildung 4.4: Eine Prozessübersicht der Datenqualität	47
Abbildung 4.5: Integration der Datenqualitätsprozesse in die Linienorganisation ...	48
Abbildung 4.6: Die Schritte eines Datenqualität-Assessments	49
Abbildung 4.7: Funktionen des Datenqualitätsmanagements	53
Abbildung 5.1: Regelkreis der Risikomanagementprozesse	61
Abbildung 5.2: Proaktive und korrigierende Maßnahmen zur Risikoreduzierung ...	66
Abbildung 6.1: Aufbau eines stabilen Datenqualität-Netzwerkes.....	72
Abbildung 6.2: Der Ablauf des Risiko-Assessments	75
Abbildung 6.3: Schema zur Auswertung des Konzepts	78

Abbildung 7.1: Ebenen des Datenqualitätsmanagementprozesses	88
Abbildung 7.2: Register der atomaren Fragebögen	92
Abbildung 7.3: Exemplarische Auswertung des Fragebogens	93
Abbildung 7.4: Visualisierung der Ergebnisse.....	94
Abbildung 7.5: Flussdiagramm zum Verlauf des proaktiven Datenqualitätsmanagementprozesses.....	96
Abbildung 7.6: Beispielhafte Visualisierung der Befragungsergebnisse.....	98
Abbildung 7.7: Beispielhafte Visualisierung der Befragungsergebnisse.....	98
Abbildung 7.8: Kern des proaktiven Datenqualitätsmanagementprozesses	99
Abbildung 8.1: Qualitätsmodell für die Benutzbarkeit	104
Abbildung 8.2: Klassifikation von Evaluationsmethoden	105
Abbildung A.1: Architektur – Technische Sicht der Risiko-Assessments	123
Abbildung A.2: Screenshot der Fehlermeldung.....	125
Abbildung A.3: Visualisierung des Ergebnisses	126
Abbildung B.1: Allgemeiner Ablauf eines Risiko-Assessments und verwendete Programme.....	127
Abbildung B.2: Beispiel für einen Hilfetext	131

Tabellenverzeichnis

Tabelle 3.1: Die fünf Stufen des Capability Maturity Model.....	36
Tabelle 6.1: Schema zur Bewertung der Datenqualität	79
Tabelle 7.1: Risikobewertungsmatrix	89
Tabelle C.1: Graphische Aufbereitung der Evaluierungsergebnisse (Teil 1).....	140
Tabelle C.2: Graphische Aufbereitung der Evaluierungsergebnisse (Teil 2).....	141

Listingverzeichnis

Listing 6.1: Beispiel eines SQL-Statements	80
Listing 6.2: Beispiel eines SQL-Statements	80
Listing A.1: Skript zur Auswertung des Fragebogens.....	124
Listing B.1: Skript zum Datentransport.....	128
Listing B.2: Skript zur Auswertung.....	129
Listing B.3: Skript zum logischen Ablauf des Fragebogens (Teil 1)	130
Listing B.4: Skript zum logischen Ablauf des Fragebogens (Teil 2)	130
Listing B.5: Skript zu den Hilfetexten.....	131
Listing B.6: Skript zur Sicherung der Benutzer-Identifikation (BID).....	132