



# **Verantwortung in der digitalen Datenanalyse**

**Transparenz als Grundlage  
für eine glaubwürdige Prüfung**

**Herausgegeben vom**

Deggendorfer Forum zur digitalen Datenanalyse e. V.

**Mit Beiträgen von**

Christoph Haas, Prof. Dr. Siegfried Handschuh, Uwe Nadler,  
Remo Rechkemmer, Dr. Ralf Steinberger, Dr. Frank Wittig

**ERICH SCHMIDT VERLAG**

**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.

**Weitere Informationen zu diesem Titel finden Sie im Internet unter**

[ESV.info/978 3 503 17096 8](http://ESV.info/9783503170968)

Gedrucktes Werk: ISBN 978 3 503 17096 8

eBook: ISBN 978 3 503 17097 5

Alle Rechte vorbehalten

© Erich Schmidt Verlag GmbH & Co. KG, Berlin 2017

[www.ESV.info](http://www.ESV.info)

Dieses Papier erfüllt die Frankfurter Forderungen der Deutschen Nationalbibliothek und der Gesellschaft für das Buch bezüglich der Alterungsbeständigkeit und entspricht sowohl den strengen Bestimmungen der US Norm Ansi/Niso Z 39.48-1992 als auch der ISO Norm 9706.

Druck und Bindung: Strauss, Mörlenbach

## Vorwort

Das Wort „*postfaktisch*“ ist zu Recht zum „*Un*“-Wort des Jahres 2016 gewählt worden. Gleichzeitig werden Fakten, die aus Daten gewonnen werden können, immer mehr zur treibenden Kraft der aktuellen wissenschaftlichen und technologischen Entwicklungen. Beides zeigt eindrucklich, welchen Stellenwert die sachbezogene Analyse vorhandener Daten inzwischen hat.

Das Deggendorfer Forum widmet sich schon seit Jahren der Analyse von betrieblichen Daten. In 2016 wurde mit dem Leitmotiv *Transparenz aus Verantwortung* eine der großen Herausforderungen der digitalen Datenanalyse in den Mittelpunkt gestellt. Auf Einladung von Herr Prof. Dr. Jürgen Ernstberger fand das Forum an der Technischen Universität in München statt. Die Beiträge in diesem 12. Tagungsband des Deggendorfer Forums basieren auf den Vorträgen der Referenten, die wie in jedem Jahr aus Wirtschaft, Forschung und Praxis berichteten.

Transparenz ist in der Tat ein bedeutender Aspekt des Themas Datenanalyse geworden. Wenn Daten analysiert und Informationen daraus gewonnen werden, ist es wichtig, dass der ganze Prozess nachvollziehbar und transparent bleibt. Nur so kann die nötige Sicherheit für die Richtigkeit der abgeleiteten Informationen und Schlussfolgerungen gewährleistet werden. Die Analyse der Daten ist oftmals sehr komplex, sie spiegelt zwangsläufig die Heterogenität des Teils der Wirklichkeit wider, der gerade analysiert werden soll.

Strukturelle Probleme und Kontrolldefizite im deutschen Gesundheitssystem präsentierte der Bestsellerautor Dr. Frank Wittig in seinem Vortrag *Der Bock und der Gärtner: Kontrolle im deutschen Gesundheitssystem*.

Herr Christoph Haas berichtete unter dem Titel *IT-Sicherheit und Datenanalyse: Wie trinkt man aus einem Feuerweherschlauch?* über die Anatomie gezielter Angriffe auf IT-Systeme, deren Auswirkungen auf die Sicherheit dieser Systeme aber auch die Möglichkeiten mit der Datenanalyse solchen Angriffen zu begegnen.

Über künftige Veränderungen im Bereich der Wirtschaftsprüfung und den Einfluss neuer Technologien auf den Prüfungsalltag referierte WP Remo Rechkemmer, Vorstandsassistent PWC München, in seinem Vortrag: *Transformation der Wirtschaftsprüfung durch die digitale Datenanalyse*.

Prof. Dr. Siegfried Handschuh hielt seinen Vortrag über *Who Controls the Vocabulary, Controls the Knowledge – Wissensextraktion aus Dokumenten*. Er stellte grundsätzliche Fragen: Wie hängen Wörter und Vokabular zusammen bei der Generierung von Wissen, was ist die Bedeutung eines Wortes und wie gewinnt man Wissen aus Texten?

*Undurchschaubarer Datensumpf oder kristallklarer Datensee? Die Bedeutung von Information Governance für Transparenz und verantwortliches Handeln bei der Analyse von Daten.* Uwe Nadler von der IBM Deutschland erläuterte den Begriff *Information Governance* und dessen Bedeutung für die Datenanalyse.

Dr. Ralf Steinberger vom European Commission's Joint Research Centre in Ispra, Italien, berichtete über *Supporting the Analyst by Turning Multilingual Text into Structured Data*. Er stellte den von seinem Team entwickelten *Europe Media Monitor* vor, der täglich mehr als 200.000 Online-Artikel aus über 70 verschiedenen Sprachwelten maschinell auswertet.

Allen Referenten und Mitwirkenden möchte ich an dieser Stelle persönlich, sowie im Namen des Vereins, meinen Dank aussprechen, für ihr großes Engagement und die Mühe, mit der sie Ihr Wissen und ihre Erfahrungen in den Tagungsband eingebracht haben. Ohne ihre Unterstützung wäre die Herausgabe dieses Tagungsbandes nicht möglich gewesen.

Mein besonderer Dank richtet sich an die Kooperationspartner der dab: GmbH, BDO AG, DATEV eG und der Technischen Hochschule Deggendorf.

Für die Anpassung der schriftlichen Beiträge an ein einheitliches Layout bedanke ich mich bei Carmen Andraschko, Studentin der Technischen Hochschule Deggendorf, dem Erich Schmidt Verlag sowie meinem Mitstreiter Herrn Ernst-Rudolf Töller für die vielfältigen Anregungen, Diskussionen und wertvollen Kommentare.

Georg Herde

Deggendorf, im Februar 2017

# Inhaltsverzeichnis

Vorwort.....	5
Dr. Frank Wittig	
<b>Dem Missbrauch Tür und Tor geöffnet: Kontrolldefizite im (deutschen) Gesundheitssystem</b> .....	9
1    Einleitung.....	11
2    Kontrolldefizite im (deutschen) Gesundheitssystem.....	11
3    Die medizinischen Fachgesellschaften .....	12
4    Die Krankenhausfinanzierung .....	18
5    Ein Lösungsvorschlag.....	23
Christoph Haas	
<b>IT-Sicherheit durch Datenanalyse:</b>	
<b>Wie man aus einem Feuerwehrschauch trinkt</b> .....	25
1    Abstract .....	27
2    Anatomie gezielter Angriffe.....	27
3    IT-Sicherheit und Datenanalyse .....	29
4    Herausforderungen .....	35
Remo Rechkemmer	
<b>Transformation der Wirtschaftsprüfung durch die digitale Datenanalyse</b> .....	39
1    Einleitung.....	41
2    Bisherige Entwicklung der Wirtschaftsprüfung .....	41
3    Aktueller Einsatz von digitalen Vollanalysen .....	43
4    Zukünftige Auswirkung der Automatisierung .....	45
5    Zukünftige Entwicklung der Wirtschaftsprüfung .....	49
6    Thesenförmige Zusammenfassung .....	51
Prof. Dr. Siegfried Handschuh	
<b>Who Controls the Vocabulary, Controls the Knowledge</b>	
<b>Wissensextraktion aus Dokumenten</b> .....	53
1    Einleitung .....	55
2    Was ist Semantik? .....	55
3    Anwendungen .....	60
4    Zusammenfassung .....	66

Uwe Nadler

<b>Undurchschaubarer Datensumpf oder kristallklarer Datensee?</b> .....	69
1 Die Bedeutung von Information Governance im Data Lake .....	71
2 Das Information Governance Modell .....	72
3 Ausrichtung der Information Governance auf die Unternehmensziele .....	73
4 Die Kerndisziplinen einer Information Governance .....	74
5 Treiber von Information Governance-Vorhaben .....	76
6 Die Hilfsdisziplinen .....	77
7 Protokollierung & Berichte .....	79
8 Die Umsetzung von Information Governance im Data Lake Konzept .....	80
9 Am Data Lake beteiligte Rollen .....	81
10 Die Bedeutung des Katalogs für den Data Lake .....	82
11 Fazit .....	83

Dr. Ralf Steinberger

<b>EMM: Supporting the Analyst by Turning Multilingual Text into Structured Data</b> .....	85
1 Introduction .....	87
2 The Europe Media Monitor (EMM) .....	88
3 Other media monitoring services and tools .....	99
4 Text analysis quality and usage warnings .....	102
5 Transparency and responsibility .....	104
6 Conclusion and future developments .....	105