

# INFORMATION- VISUALISIERUNG

Grafische Aufbereitung und  
Analyse von statistischen Daten

## 18. Wissenschaftliches Kolloquium

gemeinsam mit der  
Deutschen Statistischen Gesellschaft

19. - 20. November 2009  
Wiesbaden

## Programmheft

Statistisches Bundesamt

## Donnerstag, 19. November 2009

- 11<sup>00</sup> **Begrüßung**  
Roderich Egeler  
*Präsident des Statistischen Bundesamtes,  
Wiesbaden*
- 11<sup>15</sup> **Einführung in das Thema**  
Prof. Dr. Claus Weihs (Moderator)  
*Technische Universität Dortmund*
- 11<sup>45</sup> **Informationsvisualisierung: Methoden und Perspektiven**  
Prof. Dr. Heidrun Schumann  
*Universität Rostock*
- 12<sup>15</sup> **Diskussion**
- 12<sup>30</sup> **Mittagspause**
- 13<sup>30</sup> **Interaktive Analyse von räumlich-zeitlichen Daten**  
Prof. Dr. Stefan Wrobel  
*Fraunhofer-Institut für Intelligente Analyse- und  
Informationssysteme IAIS, Sankt Augustin*
- 14<sup>00</sup> **Diskussion**
- 14<sup>15</sup> **Visuelle Web-Kommunikation**  
Sibylle von Oppeln-Bronikowski und Tanja Raschke  
*Statistisches Bundesamt, Wiesbaden*
- 14<sup>45</sup> **Diskussion**
- 15<sup>00</sup> **Kaffeepause**
- 15<sup>30</sup> **Datenanalyse mit interaktiven Grafiken**  
Dr. Martin Theus  
*Telefónica o2 Germany GmbH & Co. OHG, München*
- 16<sup>00</sup> **Diskussion**
- 16<sup>15</sup> **Text-Fakten-Integration in Informationssystemen**  
Dr. Maximilian Stempfhuber  
*Rheinisch-Westfälische Technische Hochschule (RWTH)  
Aachen*
- 16<sup>45</sup> **Diskussion**
- 17<sup>00</sup> **Gerhard-Fürst-Preis**  
Laudatio: Prof. Dr. Hans Wolfgang Brachinger  
*Universität Fribourg, Schweiz*  
Preisverleihung durch Präsident Roderich Egeler
- 18<sup>00</sup> **Ende des ersten Veranstaltungstages /  
Sektempfang im Foyer**
- 20<sup>00</sup> **Geselliges Beisammensein im Restaurant Lumen**

## Freitag, 20. November 2009

- 09<sup>00</sup> **Informationsgrafiken und Karten in deutschen Zeitungen**  
Thomas Heumann  
*Frankfurter Allgemeine Zeitung (FAZ)*
- 09<sup>30</sup> **Diskussion**
- 09<sup>45</sup> **Visualisierung von fehlenden Werten**  
Dr. Matthias Templ  
*Technische Universität Wien & Statistik Austria*
- 10<sup>15</sup> **Diskussion**
- 10<sup>30</sup> **Kaffeepause**
- 11<sup>00</sup> **Das Data Mining System RapidMiner  
(Open Source Data Mining)**  
Dr. Ingo Mierswa  
*Rapid-I GmbH, Dortmund*
- 11<sup>30</sup> **Diskussion**
- 11<sup>45</sup> **Visualisierung und statistische Modellierung**  
Prof. Dr. Adalbert Wilhelm  
*Jacobs University Bremen*
- 12<sup>15</sup> **Diskussion**
- 12<sup>30</sup> **Ende der Veranstaltung**

# Abstracts / Kurzfassungen der Beiträge

## Einführung in das Thema

*Prof. Dr. Claus Weihs*

*geb. 1953, hat den Lehrstuhl für Computergestützte Statistik an der Fakultät Statistik der Technischen Universität Dortmund seit 1995 inne. Davor arbeitete er mehr als 9 Jahre als Statistiker in der Chemischen Industrie (CIBA-Geigy, Schweiz). Er promovierte 1986 in Mathematik an der Universität Trier. 1978 - 1985 arbeitete er als Diplom-Mathematiker an der Fakultät für Gesellschafts- und Wirtschaftswissenschaften der Universität Bonn an der Entwicklung eines Softwaresystems für Ökonometrische Modellierung. Seine jetzigen Forschungsinteressen umfassen Klassifikationsmethoden, Statistische Prozesskontrolle, Statistische Versuchsplanung, Zeitreihenanalyse, Musikdatenanalyse und Visualisierung. Er ist seit 2004 Vorsitzender der deutschen Gesellschaft für Klassifikation (GfKl).*

### 1 Einführung

Dieser Vortrag gibt einerseits einen Einblick in die Ergebnisse des so genannten Milestones-Projekts (<http://datavis.ca/milestones/>, Friendly, 2005), in dem die historische Entwicklung der visuellen Darstellung statistischer Daten bis etwa zum Jahre 2000 dargestellt wird, andererseits werden neuere Entwicklungen vorgestellt. Daneben werden noch andere Möglichkeiten der Anordnung von grafischen Techniken diskutiert wie z. B. in Form eines Periodensystems ([visual literacy, http://www.visual-literacy.org/periodic\\_table/periodic\\_table.htm](http://www.visual-literacy.org/periodic_table/periodic_table.htm)).

### 2 Historischer Überblick

Jahr	Name	Nationalität	Erfindung
<b>Anfänge</b>			
1644	Langren	Niederlande	Entfernungsplot
1741	Süssmilch	Deutschland	Bevölkerungsstatistik
1748	Achenwall	Deutschland	Statistik
1786	Playfair	England	Zeitdarstellung und Säulendiagramm
<b>1800-1849: Beginn der modernen Datengrafik</b>			
1801	Playfair,	England	Tortendiagramm
1832	Herschel	England	Streudiagramm

<b>1850-1899: Das goldene Zeitalter der Datengrafik</b>			
1879	Perozzo	Italien	Stereogramm
<b>1900-1949: Moderne dunkle Zeiten</b>			
<b>1950-1974: Wiedergeburt der Datenvisualisierung</b>			
1969	Tukey	USA	EDA: Stamm&Blatt, Box-Plot PRIM-9 (Projektion, Rotation, Isolierung, Maskierung)
1974	Friedman, Fishkeller, Tukey	USA	
<b>1975-2000: Visualisierung hochdimensionaler Daten</b>			
1981	Kleiner, Hartigan	USA	Mosaik Darstellung
1983	Tufte	USA	Grafik Ästhetik

Der hier gezeigte Ausschnitt der Entwicklung der grafischen Darstellungen beschränkt sich im Wesentlichen auf diejenigen Erfindungen, die eine wirklich neue Sicht auf statistische Daten ermöglichten, nämlich Säulendiagramm, Tortendiagramm und Zeitreihenplot, die alle Playfair zugerechnet werden, das Streudiagramm, das Herschel zugerechnet wird, und 3D-plots (Perozzo), Box-Plot (Tukey), Mosaic Plot (Hartigan & Kleiner) sowie die interaktive multivariate Grafik. In dem Vortrag werden noch viele andere Darstellungen diskutiert, insbesondere auch solche Darstellungen aus der Kartografie, die mehr zeigen als geografische Positionen und Eigenschaften.

### 3 Neuere Geografische Darstellungen

Zusätzlich werden Darstellungen an Beispielen vorgestellt, die auch nicht-geografische statistische Daten als Landschaften repräsentieren. Beispiele sind Emergente Self-Organizing Maps (ESOMs, Ultsch, 1993) und nichtlineare Darstellungen, wie sie etwa schon von Sammon (1969) vorgeschlagen wurden. Eine weitere aktuelle grafische Darstellung verändert die Geografie so, dass die Größe einer weiteren Variablen beachtet wird (Diffusionsmethode, Gastner und Newman, 2004). Dabei werden z. B. die Staaten der USA so dargestellt, dass ihre Fläche proportional zu der Anzahl ihrer Wahlmänner ist (<http://www-personal.umich.edu/~mejn/election/2004/>).

### Literatur

Friendly, M. (2005): Milestones in the History of Data Visualization: A Case Study in Statistical Historiography; in: Weihs, C., Gaul, W. (eds.): *Classification – The Ubiquitous*

*Challenge, Handbook of Computational Statistics: Data Visualization*, Springer, 2005.

Gastner, M.T., Newman, M. E. J. (2004): Diffusion-based method for producing density equalizing maps, *Proc. Natl. Acad. Sci. USA* 101, 7499-7504

Sammon, J. W. (1969): A non-linear mapping for data structure analysis. *IEEE Transactions on computers* C18, 401-409

Ultsch, A. (1993): Self-organizing neural networks for visualization and classification. In: Opitz, O., Lausen, B., Klar, R. (eds): *Information and classification – concepts, methods, and applications*, Springer, Berlin, 307-313

## **Informationsvisualisierung: Methoden und Perspektiven**

*Prof. Dr. Heidrun Schumann*

*geb. 1954, studierte, promovierte und habilitierte an der Universität Rostock. Seit 1992 ist sie Professorin für Computergraphik am Fachbereich Informatik in Rostock. Sie ist Expertin auf dem Gebiet der Visualisierung von Daten und Informationen und hat hierzu in Kooperation mit Wolfgang Müller das erste geschlossene Lehrbuch verfasst (H. Schumann und W. Müller: *Visualisierung – Grundlagen und allgemeine Methoden*, Springer-Verlag 2000). Außerdem beschäftigt sie sich mit adaptiven visuellen Schnittstellen für mobile handhelds und smart environments. Sie war von 1994-2000 Sprecherin der DFG-Forscherguppe „Mobile Visualisierung“. Weitere Informationen sind abrufbar unter <http://www.informatik.uni-rostock.de/~schumann/>.*

Ziel der Informationsvisualisierung ist es, abstrakte Daten graphisch so zu repräsentieren, dass strukturelle Zusammenhänge und relevante Eigenschaften intuitiv erfasst werden können. Damit soll die interaktive Exploration komplexer Datenmengen unterstützt werden.

Aktuelle Themen adressieren vor allem Komplexität und Umfang heutiger Datensätze, eine stärkere Berücksichtigung des Anwenders sowie die Verknüpfung von visuellen und automatischen Methoden. Das neue Forschungsfeld „Visual Analytics“ setzt diesen Trend um und wird derzeit sowohl auf nationaler als auch internationaler Ebene durch verschiedene Programme unterstützt.

Im Vortrag werden diese Entwicklungen kurz aufgezeigt und mit konkreten Methoden sowie Anwendungsbeispielen, insbesondere aus dem Rostocker Umfeld, hinterlegt. Dabei wird

auch auf das Problem der automatischen Anpassung einer visuellen Repräsentation eingegangen, um dadurch die Ziele eines Nutzers oder die gegebene Infrastruktur besser zu berücksichtigen. Der Vortrag schließt mit einem kurzen Ausblick sowohl auf künftige Forschungsfragen als auch auf neue Perspektiven in der Anwendung.

## **Interaktive Analyse von räumlich-zeitlichen Daten**

*Prof. Dr. Stefan Wrobel*

*geb. 1962, ist Leiter des Fraunhofer-Instituts für Intelligente Analyse- und Informationssysteme IAIS und Professor für Informatik an der Universität Bonn. Er studierte Informatik in Bonn und Atlanta, Georgia, USA (M.S., Georgia Institute of Technology), und promovierte an der Universität Dortmund. Nach Stationen in Berlin und Sankt Augustin wurde er Professor für Informatik an der Universität Magdeburg, bevor er im Jahr 2002 den Ruf auf seine aktuelle Position annahm. Das Fraunhofer-Institut IAIS beschäftigt sich mit intelligenten Algorithmen und Systemen zur Analyse großer Datenmengen, von Sensordaten über Geschäftsdatenbanken bis hin zu multimedialen Dokumenten. Seine Forschungsergebnisse nutzt das Institut in zahlreichen Kundenprojekten in den Geschäftsfeldern Marktforschung und Marketing, Unternehmensplanung und Controlling, Digital Media Asset Management, Prozessintelligenz, Präventive Sicherheit und High-tech-Erlebnisräume.*

Die Analyse raumbezogener Daten stellt seit langer Zeit einen wichtigen Arbeitsschwerpunkt in Statistik und Datenanalyse dar, ebenso wie die Betrachtung von Zeitreihen und anderen dynamischen Veränderungen. Hierbei stellen Anwendungsbereiche, in denen zeitliche Veränderung mit starkem Raumbezug gleichzeitig analysiert werden müssen, von jeher eine besondere Herausforderung dar. Sie rücken in den letzten Jahren durch die Verfügbarkeit neuartiger Datenerhebungstechniken auf der Basis von zum Beispiel GPS-Satelliten- oder Mobilfunkdaten vermehrt ins Zentrum des Interesses. Ob im Bereich der Verkehrsplanung oder bei der Analyse eines Vertriebsnetzes, bei der Untersuchung der Ausbreitung von Pandemien oder der Entwicklung bedrohter Tierarten, bei einer Untersuchung von Wahlergebnissen oder in der Kriminologie - überall müssen räumliche und zeitliche Dimension gemeinsam betrachtet werden.

Fortschritte im aus der Informatik entstandenen Gebiet der sogenannten Visuellen Analytik haben in den letzten Jahren gezeigt, dass insbesondere bei der geeigneten Kombination von interaktiv-visuellen Elementen mit computergestützten

Analyseverfahren eine deutliche Erleichterung und verbesserte Ergebnisse bei der Analyse solcher Problemstellungen möglich sind. Der Vortrag gibt eine kurze Einführung in das Gebiet und seine Grundlagen im Hinblick auf raum-zeitliche Daten und illustriert dann anhand mehrerer Fallstudien, wie mit entsprechenden Methoden und Systemen raum-zeitliche Daten analysiert werden können. Unsere Beispiele reichen dabei von der Außenwerbung bis zur Biologie, und demonstrieren Probleme mit unterschiedlichen räumlichen Datentypen und zeitlichen Strukturen. Eine der vorgestellten Studien wurde im Jahr 2008 mit dem Preis der Deutschen Marktforschung gewürdigt.

## **Visuelle Web-Kommunikation**

*Sibylle von Oppeln-Bronikowski*

*geb. 1953, eine erwachsene Tochter. Studium der Volkswirtschaftslehre in Heidelberg, Diplom mit dem Schwerpunkt Wirtschafts- und Sozialstatistik bei Prof. Wagenführ. Seit 1980 im Statistischen Bundesamt. 2004 bis 2008 Leitung der Abteilung Preise, Außenhandel, Verkehr und Verdienste. Seit 2008 Leitung der Abteilung Grundsatzfragen der Bundes- und internationalen Statistik, Informationsverbreitung. Schwerpunkt der beruflichen Laufbahn: Strategische Ausrichtung der Statistik und der Kommunikation sowie internationale Beziehungen.*

*Tanja Raschke*

*geb. 1979, Studium der Informatik (Studiengang Computer-visualistik) in Koblenz, Diplom. Seit 2006 im Statistischen Bundesamt als wissenschaftliche Mitarbeiterin im Bereich Online-Dienste.*

Kommunikation dient der Übermittlung von Information. Dies kann auf unterschiedliche Arten geschehen. Visuelle Formen bieten gegenüber einigen anderen Darstellungsarten (z. B. textuellen Formen) häufig Vorteile. Sie ermöglichen, komplexe Sachverhalte auf einfache Weise und leicht verständlich darzustellen („From eyes to brain“). Werden diese Visualisierungen mit Interaktivität kombiniert, erhält der Nutzer die Möglichkeit, sich Daten individuell zu erschließen und auf seine Bedürfnisse anzupassen.

### **1 Interaktive Anwendungen bei Destatis**

#### **1.1 Animierte Bevölkerungspyramiden**

Die Ergebnisse der 11. koordinierten Bevölkerungsvorausbe-  
rechnung werden in Form von animierten Alterspyramiden  
angeboten. Diese existieren in zwei Varianten mit unter-  
schiedlichem jährlichen Wanderungssaldo (+100.000 oder  
+200.000 Personen) und jeweils unter der Annahme, dass  
die Geburtenhäufigkeit annähernd konstant bei 1,4 Kindern  
je Frau liegt. Die Entwicklung des Altersaufbaus kann mit  
den animierten Bevölkerungspyramiden über 100 Jahre ver-  
folgt und zu jeder Zeit angehalten werden. Mit einem Schie-  
beregler können einzelne Jahre gezielt ausgewählt werden.  
Die Stärke der Geburtsjahrgänge kann durch Überfahren mit  
der Maus angezeigt werden; ein einzelner kann durch Ankli-  
cken fixiert und somit über die Zeit beobachtet werden. Er-  
gänzend dazu werden Länderpyramiden mit dem Altersauf-  
bau der Bevölkerung nach Ländern angeboten.

### 1.2 Persönlicher Inflationsrechner

Der Warenkorb des Verbraucherpreisindex besteht aus etwa  
700 Gütern. Für jede Güterart – also zum Beispiel für Bücher,  
Kinokarten oder für Benzin – wird die Preisentwicklung be-  
rechnet. Der gesamte Verbraucherpreisindex ist dann ein  
gewichteter Mittelwert aus der Preisentwicklung bei allen  
etwa 700 Güterarten. Die Teuerungsrate hängt nicht nur da-  
von ab, wie sich die Preise verändern. Entscheidend ist  
auch, mit welchem Gewicht die Preisentwicklung der einzel-  
nen Güterarten in den Gesamtindex eingeht. Die Gewichte  
gelten für den durchschnittlichen Konsumenten. Diesen gibt  
es aber nur in der Theorie. Die amtlich berechnete Teuerung  
lässt sich daher nicht direkt auf die eigene Situation über-  
tragen. Wie stark ein einzelner Haushalt von der Inflation be-  
troffen ist, hängt von seinem individuellen Konsumverhalten  
ab. Um den Einfluss der Gewichtung auf einfache Weise zu  
veranschaulichen, kann man seine persönlichen Ver-  
brauchsgewohnheiten mit einem experimentellen Ver-  
braucherpreisindex abbilden. In diesem kann man beispiel-  
haft einige Gewichte des Warenkorbes verändern. Mit Hilfe  
von Schiebereglern können die Ausgaben für bestimmte Wa-  
ren und Dienstleistungen angepasst werden. In einer Grafik  
sieht man dann, wie sich die Änderung des Warenkorbge-  
wichtes auf die gemessene Preisentwicklung auswirkt. Mit  
dem persönlichen Inflationsrechner kann man somit ein in-  
dividuelles Wägungsschema zusammenstellen und auf die-  
ser Grundlage eine persönliche Inflationsrate berechnen.

### 1.3 Konjunkturmonitor

Der Konjunkturmonitor des Statistischen Bundesamtes bil-  
det dynamisch die Bewegung ausgewählter Indikatoren in  
einem Quadrantensystem der konjunkturellen Grundphasen  
ab. Damit bietet er eine Alternative zur gewohnten grafi-

schen Darstellung konjunkturrelevanten Daten auf der Zeitachse. Mit dem Konjunkturmonitor kann man sich den Wirtschaftskreislauf und seine Bestimmungsfaktoren interaktiv erschließen. Im Zeitablauf bewegen sich die ausgewählten Wertepaare im Uhrzeigersinn durch die Quadranten des Monitors. Über zusätzliche Schaltflächen können für jeden einzelnen Indikator auch die zugrundeliegenden Konjunkturdaten tabellarisch angezeigt werden. Ebenso ist es möglich, eine Grafik in konventioneller Darstellung auf einer Zeitachse zu generieren. Über entsprechende Selektion können jeweils die gewünschten Indikatoren zugeschaltet oder ausgeblendet werden.

#### 1.4 Neu: Preiskaleidoskop/Voronoi-Diagramm

Zur Verdeutlichung des komplexen Zusammenhangs zwischen Wägungsschema und Preisentwicklung wurde eine neuartige Darstellungsform – ein so genanntes Voronoi-Diagramm – gewählt, da für Außenstehende die starke Zusammenfassung der Preisentwicklungen von etwa 700 Güterarten manchmal schwer nachzuvollziehen ist. In der neu entwickelten Anwendung werden die Gewichte und die Preisentwicklungen der Produktgruppen im Warenkorb zusammenfassend in einer Art „Kaleidoskop“ visualisiert. Die Höhe der Warenkorbgewichte wird durch die Größe von organisch geformten Teilsegmenten innerhalb eines Kreises dargestellt. Die Preisveränderungen der einzelnen Gütergruppen werden durch Farbkodierung gekennzeichnet. So bietet es einerseits einen schnellen Überblick über die Bedeutung der Gütergruppen und über deren Preisentwicklungen. Andererseits kann man sich bei Interesse auch ein sehr detailliertes Bild über die Preisentwicklungen in bestimmten Bereichen verschaffen, indem man hineinzoomt und so von einer gröberen zu einer feineren Betrachtungsebene wechselt. Die Darstellungsform ist eine gelungene Art, die komplexe Indexkonstruktion der Preisstatistik intuitiv verständlich zu machen.

## 2 Weitere Inhalte

Daten werden zunehmend als Indikatoren veröffentlicht. Um diese Daten der Öffentlichkeit umfassender und einfacher zugänglich zu machen, wurde im Destatis-Webangebot ein eigener Bereich „Indikatoren“ neu angelegt bzw. ausgebaut. Dort werden Tabellen und Diagramme für Konjunkturindikatoren in neuer Form und mit neuer Navigation angeboten, ebenso gibt es Links zu anderen Indikatorensystemen.

In diesem Bereich sowie für die Darstellung des Verbraucherpreisindex nach Abteilungen wurde eine Reiternavigation eingeführt. Sie dient der Übersichtlichkeit und verbessert

das Verständnis der Zusammenhänge der verschiedenen Datenangebote (Monatswerte/Jahresdurchschnitte, Indizes/Veränderung zum Vorjahr, unterschiedliche Abteilungen).

Um Nutzer darin zu unterstützen, bestehende Inhalte unseres Webangebotes weiterzuverwenden (Stichwort „Mashup“), bieten wir zur Einbindung eines unserer Diagramme in eigene Webseiten ein Code-Schnipsel an, das einfach kopiert und eingefügt werden kann. Hierüber erreichen wir einen größeren Nutzerkreis, der unsere Daten dennoch in unserem Design erfasst. Auch von Eurostat werden Inhalte unserer Homepage eingebunden. Dazu werden Inhalte aus einem unserer RSS-Feeds weiterverarbeitet und dessen Inhalt im Kontext der Eurostat-Webseite dargestellt.

Zunehmend wichtig ist auch die Darstellbarkeit und Nutzbarkeit unseres Webangebotes auf mobilen Endgeräten durch den stetig steigenden Anteil mobiler Internetnutzer (z. B. via iPhone).

All diese Entwicklungen zeigen deutlich, dass der Bereich Internet ein sich dynamisch entwickelnder Markt ist, der permanenter Beobachtung bedarf. Auch außerhalb der amtlichen Statistik gibt es stets eine Reihe interessanter Neuerungen, die dem statistischen Umfeld neue Möglichkeiten eröffnen und zeitnah Einfluss auf aktuelle Entwicklungen nehmen sollten.

## **Datenanalyse mit interaktiven Grafiken**

*Dr. Martin Theus*

*geb. 1967 in Geldern (Niederrhein), verheiratet, 2 Kinder. 1986 – 1992 Studium der Statistik mit Nebenfach Informatik an der Universität Dortmund, 1992 Diplom, 1992 – 1993 Software Entwickler, Continentale Lebensversicherung, München, 1994 – 1997 Mitarbeiter am Lehrstuhl von Prof. A. Unwin, Universität Augsburg, 1996 Promotion, 1997 – 1998 Statistics Research, AT&T Labs, Florham Park, NJ, 1999 – 2001 Data Warehouse Analyst, VIAG-Interkom (heute O2), München, 2002 – 2006 Wiss. Assistent am Lehrstuhl von Prof. A. Unwin, Universität Augsburg, 2003 Visiting Lecturer an der Iowa State University, Ames, IA, 2004 – 2005 Lehrauftrag an der Universität München, seit 2007 Senior Project Manager, Telefónica O2-Germany, München*

Grafiken haben in der Statistik und Datenanalyse schon lange ihren festen Platz. Generell kann man zwischen Grafiken zur Exploration, zur Diagnostik und zur Präsentation unter-

scheiden. Gerade im noch jungen Gebiet der explorativen Datenanalyse haben sich interaktive Grafiken sehr bewährt.

Interaktive Grafiken – im Gegensatz zu statischen Grafiken – bieten die Möglichkeit mittels Selektionen und linked Highlighting gezielt die Eigenschaften von ausgewählten Untergruppen in den Daten zu untersuchen. Die Untersuchung bedingter Verteilungen ist ja auch in vielen Verfahren der klassischen Statistik von fundamentaler Bedeutung.

Der große Vorteil der interaktiven statistischen Grafiken liegt darin, dass viele bedingte Ansichten der Daten in sehr kurzer Zeit erstellt werden können, die eine möglichst vollständige Betrachtung eines Datensatz ermöglichen. Im Gegensatz dazu versuchen klassische Methoden oft nur eine einzige Hypothese zu bestätigen/widerlegen.

Zusätzlich bieten interaktive statistische Grafiken eine schnelle Variation von Parametern der Grafiken; so wie zum Beispiel die Intervallbreite bei Histogrammen oder die Sortierreihenfolge in Balkendiagrammen. Interaktive Abfragen bilden das interaktive Äquivalent zu aufwendigen Skalen und Legenden in statischen Grafiken.

Dieser Vortrag bietet eine kurze Übersicht über die grundlegenden Funktionen interaktiver statistischer Grafik sowie deren effizienten Einsatz in der Datenanalyse. Für alle praktischen Beispiele wird die Software Mondrian eingesetzt, welche eine Vielzahl interaktiver Funktionen zur grafischen Datenanalyse implementiert.

## **Literatur**

Theus, M. (2002) Interactive Data Visualization using Mondrian, *Journal of Statistical Software*, Vol. 7, No. 11, pp 1-9.

Theus, M. und Urbanek, S. (2008) *Interactive Graphics for Data Analysis – Principles and Examples*, Chapman and Hall/CRC-Press.

Unwin, A.R. (1999) Requirements for Interactive Graphics Software for Exploratory Data Analysis, *Computational Statistics*, Vol. 14, No. 1, pp 7-22.

Wilkinson, L. (2005). *The Grammar of Graphics*, 2nd ed., Springer, New York.

# Text-Fakten-Integration in Informationssystemen

*Dr. Maximilian Stempfhuber*

*geb. 1966, Studium Informationswissenschaft und Wirtschaftsinformatik an der Universität Regensburg (Magister Artium), Promotion in Informatik an der Universität Koblenz-Landau. Von 1995 bis 2000 wissenschaftlicher Mitarbeiter am Informationszentrum Sozialwissenschaften (Bonn) im Projekt ELVIRA zur Entwicklung eines Statistikinformationssystems für Industrieverbände; ab 1998 Projektleiter. 2000 bis Juli 2009 Abteilungsleiter informationswissenschaftliche Forschung und Entwicklung am Informationszentrum Sozialwissenschaften (jetzt GESIS – Leibniz-Institut für Sozialwissenschaften); Leitung von Forschungsprojekten zu Digitalen Bibliotheken, zum elektronischen Publizieren und zur Qualitätssicherung von Forschungsdaten. Seit August 2009 Dezernent für Kommunikations- und Informationstechnologie (KIT) an der Hochschulbibliothek der Rheinisch-Westfälischen Technischen Hochschule (RWTH) Aachen. Lehraufträge an der Universität Koblenz-Landau für Softwareergonomie und Information Retrieval. Mitglied in den Vorständen von euroCRIS, der Deutschen Initiative für Netzwerkinformation e. V. (DINI) und der IuK-Initiative Wissenschaft. Mitglied in Arbeitskreisen der Allianz der deutschen Wissenschaftsorganisationen und des Knowledge Exchange Network.*

Aktuelle Förderprogramme der Deutschen Forschungsgemeinschaft und der Europäischen Kommission haben die Zielsetzung, „Virtuelle Forschungsumgebungen“ zu konzipieren und zu entwickeln, die Forscherinnen und Forschern nicht nur zeit- und ortsunabhängige Kooperation und die gemeinsame Nutzung von oft teurer Forschungsinfrastruktur ermöglichen sollen, sondern die auch Disziplinengrenzen überwinden helfen – und damit erst die Voraussetzung für die Lösung komplexer, gesellschaftlich relevanter Probleme schaffen.

Neben der Kommunikation der Forscherinnen und Forscher untereinander spielen in diesen virtuellen Forschungsumgebungen die Dokumentation von Prozessen, Daten und Ergebnissen und der Austausch von Information innerhalb der Gruppe und mit der Community eine zentrale Rolle:

- Elektronische Laborbücher dokumentieren Planung, Durchführung und Auswertung von Experimenten;
- Aus Experimenten gewonnene Forschungsdaten werden bereinigt, dokumentiert und in Formaten abgelegt, die die Langzeitverfügbarkeit sicherstellen;

- Externe Informationen, z. B. in Form von Forschungsdaten und Publikationen, werden elektronisch beschafft, analysiert, bearbeitet, mit den eigenen Beobachtungen korreliert und in den weiteren Forschungsprozess integriert;
- Computergestützte Verfahren der Datenauswertung werden konzipiert, entwickelt, dokumentiert und in Form von wiederverwendbaren Simulationsmodellen und Programmen abgelegt;
- Arbeitsberichte und wissenschaftliche Publikationen werden (gemeinschaftlich) angefertigt, im Kollegenkreis diskutiert, zur Veröffentlichung eingereicht;
- Publikationen und Forschungsdaten werden in elektronischer Form innerhalb der Gruppe, der beteiligten Einrichtungen oder weltweit (Open Access) zur Verfügung gestellt;
- Durch die nahtlose Integration des elektronischen Publikationsprozesses mit dem eigentlichen Forschungsprozess, der über weite Strecken ebenfalls digital erfolgt, und unter sofortiger und interaktiver Einbindung der Community entstehen neue, „fließende“ Publikationsformen, die in ihrer Qualität sowohl von der verwendeten Technologie als auch der zusätzlichen eingebrachten Expertise profitieren.

Soweit zumindest die Theorie.

In der Praxis muss sich die Nützlichkeit elektronisch verfügbarer Forschungsressourcen, z. B. Forschungsdaten, Publikationen und Simulationsmodellen, aber auch Kontakte, Kollaborationsbeziehungen und Expertisen, die in sozialen Netzwerken kodiert sind, an ihrer Zugänglichkeit und Integrierbarkeit in den eigenen Forschungsprozess messen.

Ein Blick auf den State-of-the-Art in den Bereichen Wissensorganisation/Semantic Web, Informationsarchitektur, Information Retrieval und Visualisierung zeigt, dass für viele Detailprobleme, die beim integrierten Zugriff auf semantisch und strukturell heterogen erschlossene und auch kodierte Informationen entstehen, Lösungsansätze existieren. Jedoch nehmen diese vorwiegend eine lokale, datentyp-orientierte Sicht ein, und machen nur peripher Gebrauch von Informationen und Verfahren, die in unterschiedlichen Phasen informationeller Prozesse oder auch für andere Datenarten angewendet werden.

Der Beitrag präsentiert das Gesamtmodell einer Informationsarchitektur für die Text-Fakten-Integration in Informati-

onssystemen, die sowohl an einem Wissenschaftlerarbeitsplatz, als Teil der überregionalen Informationsversorgung, als auch integriert in virtuelle Forschungsumgebungen zum Einsatz kommen kann. Das Modell vereinigt – auf unterschiedlichen Abstraktionsebenen – die Informationen, die im Kontext der Forschungsdokumentation erzeugt werden (z. B. in Forschungsinformationssystemen, engl. Current Research Information Systems, CRIS), mit den im Forschungsprozess selbst gewonnenen und produzierten, teilweise auch aus externen Quellen bezogenen Informationen (z. B. Forschungsdaten und Publikationen). Die hierdurch entstehenden Verbindungen zwischen kodiertem Orientierungswissen (was sind die Forschungsfelder; wer forscht zu einem Thema; wer arbeitet/publiziert zusammen?), Zugangswissen (welche Informationen existieren zu einem Thema?) und Faktenwissen (was wurde beobachtet; welche Schlussfolgerungen wurden gezogen?) können für die Integration heterogener Information genutzt werden. Damit stehen semantisches und strukturbezogenes Integrationswissen und -verfahren aus allen Bereich des Forschungs- und des Dokumentationsprozesses zur Verfügung, um einerseits die Zugänglichkeit der Forschungsinformation selbst zu verbessern (Heterogenitätsbehandlung), und andererseits Erkenntnisse aus der Datenanalyse zu verwenden, um den Zugriff auf zusätzliche, damit semantisch zusammenhängende Information anderen Typs zu unterstützen.

# Informationsgrafiken und Karten in deutschen Zeitungen

Thomas Heumann

*geb. 1957, Dipl.-Geograph mit Schwerpunkt Politische und Wirtschaftsgeographie. Nach Studium an der Universität Bonn 1992 Eintritt in die Redaktion der Frankfurter Allgemeinen Zeitung. Seit 1996 Leiter der Informationsgrafik-Redaktion, die für Recherche und graphische Umsetzung in Informationsgrafiken und Karten bei F.A.Z. und Frankfurter Allgemeiner Sonntagszeitung verantwortlich ist.*

## 1 Informationsgrafiken und Karten in deutschen Zeitungen

### 1.1 Quantifizierung

In deutschen Tageszeitungen wurden 2005 einer Stichprobe zufolge (Wieseth, Stefanie: Der Einsatz von Infografiken in deutschen Tageszeitungen. Eine Bestandsanalyse, Saarbrücken 2008, S. 32f.) täglich im Durchschnitt 116 Informationsgrafiken veröffentlicht. Hochgerechnet ergibt dies im Jahr rund 35 000 Grafiken von großer Themen- und Formenvielfalt. Schon allein durch Menge und Verbreitung hat Zeitungsgrafik großen Einfluss darauf, mit welchen visuellen Informationsformen die Bevölkerung vertraut ist und in welcher Weise sie wahrgenommen und genutzt werden.

### 1.2 Unterschiede zu visueller Information in Fachpublikationen

Zeitungs-Informationsgrafiken unterscheiden sich von vergleichbaren Darstellungen in wissenschaftlichen und Fachpublikationen vor allem durch

- Bestimmung für ein breites Publikum
- vielfältige, sehr unterschiedliche Themen von allgemeinem beziehungsweise verbreitetem Interesse
- häufigen Bezug auf einen aktuellen nachrichtlichen Anlass.

Für die Grafiken bedeutet dies

- Zugänglichkeit der Darstellung auch für fachfremde Leser (vorzugsweise Nutzung von gängigen Grafikformen, bei Bedarf auch Reduktion der Komplexität)
- Grafikerstellung nicht durch Spezialisten für das jeweilige Thema; die Informationsgrafiker müssen meist das ganze Themenspektrum der Zeitung abdecken
- Recherche und graphische Umsetzung unter Zeitdruck.

Hinzu kommen die Anpassung an die drucktechnischen Bedingungen und die Einbettung in das Gesamterscheinungsbild der Zeitung (Spaltenbreiten, Farben, typographische Konventionen, generell Anpassung an das Layout der Seite).

Insgesamt bedeutet dies eine Gratwanderung zwischen einerseits inhaltlicher Exaktheit und umfassender, quellentreuer Darstellung sowie andererseits der erforderlichen Zugänglichkeit und journalistischen Zuspitzung. Beispiele zeigen sowohl gelungene als auch weniger gelungene Lösungen aus verschiedenen Zeitungen.

## **2 Informationsgrafiken und Karten in der F.A.Z.**

Die F.A.Z. hat in diesem Spannungsfeld für ihre Informationsgrafik eigene Lösungen gesucht:

- Schaffung mehrerer Verständnisebenen, um auch unterschiedliche Informations- und Genauigkeitsbedürfnisse zu befriedigen
- Antizipation verschiedener Einstiegs- und Verständniswege
- Anbieten von Vergleichsdaten und -zeiträumen
- Zuschnitt ursprünglich allgemein gehaltener Darstellungen auf aktuelle Fragestellungen
- Arbeitsteilung mit dem F.A.Z.-Internetauftritt.

Anhand von Grafikbeispielen lassen sich Entwicklungsprozesse, Gestaltungskriterien und die organisatorische Einbettung in die redaktionelle Arbeit erläutern.

## **3 Zeitungs-Informationsgrafik als Vermittler wissenschaftlicher bzw. statistischer Fachinformationen**

Der Übertragungs- und Vermittlungsprozess bei der Weitergabe und Bearbeitung wissenschaftlicher bzw. statistischer Fachinformationen an ein allgemeines Publikum ist nicht unproblematisch:

- Einerseits befürchten die Quellen dieser Informationen oftmals Vergröberung oder gar Verfälschung durch die Medien
- Andererseits wird das Material häufig in einer Weise bereitgestellt, die sich geeigneter Weiterverarbeitung verschließt oder sogar Missverständnisse befördert.

Hier haben sich Verfahrensweisen wie die Vorklärung möglicher Fehlinterpretationen, die Frage nach etwaigen besser

geeigneten Datensätzen und Darstellungsverfahren etc. bewährt.

#### **4 Trends**

Neuere Trends in der Zeitungs-Informationsgrafik mit primär inhaltlichen Auswirkungen können sich zum einen aus der direkten Informationsgewinnung, -auswertung und -aufbereitung aus Datenbanken sowie der Nutzung von Geographischen Informationssystemen ergeben, sind jedoch wegen des Aufwandes auf nur wenige Zeitungen beschränkt. Generell ins Auge fällt die aufwendigere Gestaltung mit dreidimensionalen Darstellungsweisen und der Einbezug hochgradig illustrativer Elemente – bis hin zum teilweisen Abgleiten in „Infotainment“ – ins Auge.

#### **5 Fazit**

Informationsgrafiken in Zeitungen nutzen nur selten ausgefallene Darstellungsweisen oder gar innovative Visualisierungsformen. Dennoch könnte Zeitungs-Informationsgrafik – mit ihrer Aufgabenstellung, komplexe Sachverhalte für ein allgemeines Publikum verständlich zu erschließen – Impulse dafür geben, wie bei der Informationsvisualisierung Verständnishürden vermieden und unterschiedliche Nutzerinteressen bedient werden können. Zu suchen und zu fördern sind Verbesserungsmöglichkeiten an den Schnittstellen zwischen fachspezifischer visueller Information und allgemeinen Medien.

# Visualisierung von fehlenden Werten

*Dr. Matthias Templ*

*geb. 1975, absolvierte das Diplomstudium der Technischen Mathematik an der TU Wien (1997-2003) und 2009 das Doktoratsstudium mit Auszeichnung. Nach dem Studium arbeitete er in der Forschungs- und Entwicklungsabteilung der voestalpine Stahl. Seit 2004 ist er im Methodenreferat von Statistik Austria beschäftigt. Weiters ist er seit 2006 am Institut für Statistik und Wahrscheinlichkeitstheorie der TU Wien angestellt und in der Forschung und Lehre tätig. Er ist Mitglied in mehreren europäischen Forschungsprojekten (z. B. ESSnet SDC 2008-2009) und Projektleiter an der TU Wien für mehrere Projekte (z. B. FP7 Projekt AMELI - Advanced Methodology for European Laeken Indicators). Seine Forschungsschwerpunkte liegen im Bereich Statistische Geheimhaltung von Mikrodaten, Imputation, Visualisierung, Robuste Multivariate Statistik, Methoden für Kompositionsdaten und Computerorientierte Statistik. Seine Publikationsliste beinhaltet über 50 wissenschaftliche Beiträge seit 2006. Er ist Autor von mehreren frei erhältlichen R-Paketen.*

Ein zentrales Problem in der Aufbereitung von Daten oder während einer Datenanalyse sind fehlende Werte. Fehlende Werte werden stets imputiert, um vollständige Datensätze zu erhalten. Die Imputation kann vor der statistischen Analyse (typisch für die offizielle Statistik) oder durch eine Prozedur innerhalb einer statistischen Methode erfolgen. Unabhängig davon ist es von großem Vorteil vor dem Schätzen eines Modells die Daten zu veranschaulichen und dadurch über die zugrundeliegenden Daten zu lernen. Auch vor der Anwendung von Imputationsmethoden an fehlenden Werten sollte die (multivariate) Struktur von fehlenden Werten untersucht werden.

Visualisierungstechniken bieten hierbei entscheidende Vorteile gegenüber statistischen Tests. Zu diesem Zwecke wurden eine Vielzahl von grafischen Methoden entwickelt. Grundidee ist, die fehlenden Werte von bestimmten Variablen in nicht notwendigerweise unterschiedlichen Variablen hervorzuheben. In einem bivariaten Scatterplot können zum Beispiel beobachtete und fehlende Werte bestimmter anderer Variablen mit unterschiedlichen Farben und deren univariaten Verteilungen zusätzlich dargestellt werden.

Das R-Paket VIM (Visualisation and Imputation of Missing Values, Templ and Filzmoser, 2008, Templ and Alfons, 2008) hat eine Vielzahl an Visualisierungsmethoden für fehlende Werte implementiert. Verschiedenste Plot-Methoden wurden modifiziert, z. B. univariate Grafiken wie Histogramm und

Spineplot (siehe auch Theus, et al., 1997) oder multivariate Grafiken wie der Parallele Koordinatenplot. Neue Methoden wie z. B. der Matrixplot oder die Darstellung von Karten sind ebenso implementiert.

Mit VIM ist es möglich publikationstaugliche Graphiken zu erzeugen, aber auch die interaktive Darstellung von Graphiken wird gleichzeitig unterstützt. Diese Arbeit wurde teilweise von der Europäischen Union im Zuge des 7. Rahmenprogrammes für Forschung finanziert (Projekt AMELI, Nr. 217322).

## **Literatur**

Templ, M. and Filzmoser, P. (2008). Visualisation of Missing Values Using the R-package VIM. Research Report CS-2008-1, Department of Statistics and Probability Theory, Vienna University of Technology. <http://www.statistik.tuwien.ac.at/forschung/CS/CS-2008-1complete.pdf>.

Templ, M. and Alfons, A. (2008). VIM: Visualization and Imputation of Missing Values. R package version 1.2.4. <http://cran.r-project.org>.

Theus, M., Hofmann, H., Siegl, B. and Unwin, A. (1997). MANET: Extensions to Interactive Statistical Graphics for Missing Values. In *New Techniques and Technologies for Statistics II*, IOS-Press Amsterdam, 247- 259. <http://home.vrweb.de/~martin.theus/NTTS.pdf>

# Das Data Mining System RapidMiner (Open Source Data Mining)

*Dr. Ingo Mierswa*

*Gründer und geschäftsführender Gesellschafter von Rapid-I, ist verantwortlich für die Bereiche Technologie & Entwicklung, Personalmanagement und Marketing und arbeitet darüber hinaus als Softwareentwickler und Berater für Rapid-I. Seit acht Jahren koordiniert und leitet er die Entwicklung der Open Source Software RapidMiner mit weltweit 30 Entwicklern. Das Unternehmen Rapid-I hat einen breiten Kundenkreis in über 40 Ländern weltweit zu dem u.a. Unternehmen wie E.ON, Sanofi Aventis, GfK, Libri, Lufthansa Systems, mobilkom austria, Schott, Pervasive, Allianz, comdirekt bank, ThyssenKrupp und Siemens gehören.*

Die Visualisierung und Analyse komplexer Datensätze und Zusammenhänge wird gemeinhin als ein Feld für Spezialisten betrachtet. Diese definieren mit häufig sündhaft teuren Softwarelösungen mehr oder weniger komplexe Analyseprozesse, um beispielsweise drohende Kündigungen oder die Verkaufszahlen eines Produkts zu prognostizieren. Dass Softwarelösungen für Datenanalyse jedoch nicht zwingend teuer oder schwer zu bedienen sein müssen, daran dürfte spätestens seit der Entwicklung der Open Source Software RapidMiner wohl niemand mehr ernsthaft zweifeln.

Insgesamt beinhaltet das mittlerweile mehrfach ausgezeichnete RapidMiner mehr als 500 Operatoren für alle Aufgaben der Wissensentdeckung in Datenbanken, d.h. Operatoren für Ein- und Ausgabe sowie der Datenverarbeitung (ETL), maschinelles Lernen und Data Mining. Aber auch Methoden des Text Mining, Web Mining, der automatischen Stimmungsanalyse aus Internet-Diskussionsforen (Sentiment Analysis, Opinion Mining) sowie der Zeitreihenanalyse und Prognose stehen dem Analysten zur Verfügung. Zusätzlich beinhaltet RapidMiner mehr als 20 Verfahren, auch hochdimensionale Daten und Modelle zu visualisieren. Dieser Vortrag zeigt, dass hierbei nicht nur bei der Definition der Analyseprozesse selbst sondern auch bei der Visualisierung der Modellierungsergebnisse der Schwerpunkt stets auf ansprechende und zugleich leicht verständliche Darstellungsformen gelegt wird.

# Visualisierung und statistische Modellierung

Prof. Dr. Adalbert Wilhelm

*geb. 1963, ist seit 2001 als Professor für Statistik an der Jacobs University Bremen tätig und hat dort den Stiftungslehrstuhl für Information Management der Commerzbank-Stiftung inne. Der Schwerpunkt seiner Forschungsaktivitäten der letzten fünfzehn Jahre liegt im Bereich der rechnergestützten Datenanalyse mit besonderer Betonung der Bereiche Data Mining, statistische Visualisierung und Statistiksoftware. Sein Anwendungsspektrum der Statistik und Datenanalyse ist - basierend auf dem statistischen Beratungsangebot seiner bisherigen Arbeitsgruppen und der derzeitigen Einbettung in einen geistes- und sozialwissenschaftlichen Fachbereich - sehr vielfältig und reicht von wirtschaftswissenschaftlichen Fragestellungen - insbesondere durch Kooperationen mit Unternehmen aus dem Finanz- und Versicherungsbereich - über medizinische und psychologische Auswertungen, hin zu linguistischen Datenanalysen. In den vergangenen Jahren hat er sich forschungsmäßig intensiv mit dem Thema Informations- und Wissensmanagement beschäftigt, insbesondere mit Modellen zur Erklärung der Prozesse Wissensgenerierung und Wissenstransfer in virtuellen Wissens- und Kompetenznetzwerken. In aktuellen Kooperationen mit Informatikern und Kommunikations- und Medienwissenschaftlern beschäftigt er sich mit der Analyse visuellen Materials, insbesondere von Bildern und Videosequenzen. Er ist Mitglied in den einschlägigen Fachgesellschaften und Associate Editor für Computational Statistics, Computational Statistics and Data Analysis sowie dem Journal of Computational and Graphical Statistics.*

Die Entwicklung und Bereitstellung effizienter statistischer Software hat zu einer stärkeren Verbreitung und Verwendung explorativer Grafiken geführt, aber zeitgleich auch das Modellierungsspektrum hin zu immer komplexeren statistischen Modellen erweitert. Es gibt jedoch nur wenige Arbeiten, die sich systematisch mit einem Brückenschlag zwischen den beiden Forschungsgebieten befassen, wie es zum Beispiel in Gelman (2004) angeregt und insbesondere im Hinblick auf Bayesianische Modelle praktiziert wird. Die grafische Aufbereitung von Daten führt in vielen Fällen erst durch eine entsprechende statistische Modellierung zu aussagekräftigen Ergebnissen. Andererseits ist uns allen das Anscombe-Beispiel wohlbekannt, das aus vier grafisch leicht

unterscheidbaren Datensätzen besteht, die alle in einem identischen linearen Regressionsmodell resultieren. Statistische Modellierung ohne eine entsprechende graphische Aufbereitung der Daten ist eine unsichere und riskante Angelegenheit, die graphische Aufbereitung der Daten ohne entsprechende statistische Modellierung vielfach ein unsystematisches Unterfangen. Die Kombination von graphischer Darstellung der Daten und statistischer Modellierung steht im Mittelpunkt dieses Vortrags. Dabei sollen anhand ausgewählter Beispiele die Möglichkeiten gezeigt werden, wie sich Visualisierung und Modellbildung gegenseitig ergänzen und befruchten können. Der Fokus der statistischen Modellierung liegt hierbei auf linearen Regressionsmodellen, wobei insbesondere die verschiedenen in der Ökonometrie verwandten Varianten eingehender betrachtet werden. Die Visualisierung konzentriert sich dabei auf drei wesentliche Aspekte des Analyse- und Modellierungsprozesses: vor der eigentlichen Modellierung, um Besonderheiten der Daten, Substrukturen, Ausreißer und etwaig notwendige Transformationen bestimmen zu können; während der Modellierung zur Überprüfung der Modellannahmen und der Modellqualität, und nach der Modellierung, um die einzelnen Modellparameter interpretieren und Vergleiche zwischen konkurrierenden Modellen unterstützen zu können.

## Literatur

A. Gelman, Exploratory Data Analysis for Complex Models (with Discussion by Andreas Buja and Rejoinder). *Journal of Computational and Graphical Statistics* 13, 755-787, (2004).

J.W. Tukey, Analyzing data: Sanctification or detective work? *American Psychologist*, 24, 83-91 (1969).

J. Bowers & K.W. Drake, EDA for HLM: Visualization when Probabilistic Inference Fails. *Political Analysis*, 13, 301-326, (2005).

P. Groenen, & A.J. Koning, A new model for visualizing interactions in analysis of variance. *Econometric Institute Report*, No EI 2004-06 Revision Date: 2009-07-29, Erasmus University Rotterdam, Econometric Institute, <http://econpapers.repec.org/RePEc:dgr:eureir:1765001189>, (2004).

## Tagungsort:

•> Statistisches Bundesamt  
Gerhard-Fürst-Saal  
Gustav-Stresemann-Ring 11  
65189 Wiesbaden

## Tagungsleitung:

•> *Christian König*  
Statistisches Bundesamt  
Institut für Forschung und Entwicklung  
in der Bundesstatistik  
65180 Wiesbaden

Telefon: +49 (0) 611 / 75 - 20 77  
E-Mail: christian.koenig@destatis.de

## Organisation:

•> *Marco Schwickerath*  
Statistisches Bundesamt  
Institut für Forschung und Entwicklung  
in der Bundesstatistik  
65180 Wiesbaden

Telefon: +49 (0) 611 / 75 - 38 89  
Telefax: +49 (0) 611 / 75 - 39 50  
E-Mail: kolloq@destatis.de  
Internet: [www.destatis.de](http://www.destatis.de) -> Veranstaltungen

## Moderation:

•> *Prof. Dr. Claus Weihs*  
Technische Universität Dortmund  
Lehrstuhl für Computergestützte Statistik  
Vogelpothsweg 87  
44221 Dortmund

Telefon: +49 (0) 231 / 755 - 43 63  
E-Mail: [weihs@statistik.uni-dortmund.de](mailto:weihs@statistik.uni-dortmund.de)

Erschienen im November 2009

© Statistisches Bundesamt, Wiesbaden 2009

Vervielfältigung und Verbreitung, auch auszugsweise,  
mit Quellenangabe gestattet.

Statistisches Bundesamt, Informationsvisualisierung, 2009  
Bestellnummer: 0000124-09700-1