



Hans-J. Lenz

# Manipulationen und Moneten– Datentrickserei im digitalen Zeitalter

SACHBUCH



Springer

# Manipulationen und Moneten – Datentrickserei im digitalen Zeitalter

Hans-J. Lenz

# Manipulationen und Moneten – Datentrickserei im digitalen Zeitalter



Springer

Hans-J. Lenz  
Freie Universität Berlin  
Berlin, Deutschland

ISBN 978-3-658-43847-0      ISBN 978-3-658-43848-7 (eBook)  
<https://doi.org/10.1007/978-3-658-43848-7>

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

© Der/die Herausgeber bzw. der/die Autor(en), exklusiv lizenziert an Springer Fachmedien Wiesbaden GmbH, ein Teil von Springer Nature 2024

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, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von allgemein beschreibenden Bezeichnungen, Marken, Unternehmensnamen etc. in diesem Werk bedeutet nicht, dass diese frei durch jedermann benutzt werden dürfen. Die Berechtigung zur Benutzung unterliegt, auch ohne gesonderten Hinweis hierzu, den Regeln des Markenrechts. Die Rechte des jeweiligen Zeicheninhabers sind zu beachten.

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. Der Verlag bleibt im Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Covermotiv: © [stock.adobe.com/MuseColor/ID/692534607](https://stock.adobe.com/MuseColor/ID/692534607)

Covergestaltung: deblik, Berlin

Planung/Lektorat: Petra Steinmueller

Springer ist ein Imprint der eingetragenen Gesellschaft Springer Fachmedien Wiesbaden GmbH und ist ein Teil von Springer Nature.

Die Anschrift der Gesellschaft ist: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

Das Papier dieses Produkts ist recycelbar.

# Vorwort

Seien wir ehrlich, haben wir nicht alle schon einmal in der Schule bei Klassenarbeiten in der einen oder anderen Form geschummelt?

Nimmt es da Wunder, dass solche Schummeleien im späteren Berufsleben als Fälscherei und Falschinterpretation von Wirtschafts-, Technologie- und Gesellschaftsdaten ein ebenso altes wie aktuelles Problem darstellen? Erinnerung sei beispielsweise nur an den Diesel-Skandal, der nicht nur die Volkswagen AG betraf, oder an die krass gefälschten makroökonomischen Daten Griechenlands, als das Land in die Eurozone eintreten wollte und zu Unrecht vorgab, die Maastricht-Kriterien voll eingehalten zu haben, die besagen, dass die Staatsverschuldung im Betrachtungszeitraum 3 % vom Bruttosozialprodukt nicht übertreffen darf. Auch Fake News und „alternative Fakten“ sind in diesem Zusammenhang zu nennen.

Vergleichbare Unredlichkeiten zeigen sich sogar im honorigen Wissenschaftsbereich, stärker noch in dessen



**Abb. 1** Spickzettel eines Studenten in einer Klausur, Fachbereich Wirtschaftswissenschaft, Freie Universität Berlin, März 2013

nahem Umfeld. Hier sei nur an die unzähligen gefakten Experimente des seinerzeit führenden niederländischen Sozialpsychologen D. A. Stapel, ehemals Universität Tilburg, erinnert. Dazu müssen auch die Werbekampagnen der Tabakindustrie rund um das Thema *Rauchen und Krebsrisiko* genannt werden, vorgeblich auf neuesten Forschungsstudien beruhend. Sie zielten, wie wir noch im Kap. 13 sehen werden, auf Verwirrung der Bevölkerung hinsichtlich der Schädlichkeit des Rauchens ab.

Ursächlich für Unkorrektheiten, Manipulationen und Schummeleien sind neben fachlicher Unkenntnis insbesondere Absicht, also Vorsatz, motiviert durch Geldgier, Machtstreben oder Prestigesucht der Beteiligten. Hinzukommt in letzter Zeit zunehmend Hass in Form von Hassmails und -tiraden. Während die Unkenntnis noch gerade verzeihbar erscheint und durch bessere Ausbil-

dung und Schulung *on the job* abstellbar erscheint, dennoch aber mehr als ärgerlich ist, haben die anderen Motive schlicht einen unredlichen, oft sogar kriminellen Hintergrund.

Manipulation von Daten ist dabei nur eine Variante des viel mehr umfassenden Datenbetrugs. Denn Daten sind, wenn man so will, der Schatten aller menschlichen Handlungen und spiegeln so auch unredliche Handlungen wider. Im Folgenden wollen wir die wesentlichen Phänomene, teilweise sind es unglaubliche Tricksereien und Vertuschungen, aus dem Bereich des Datenbetrugs an Hand von Fällen vorstellen. An vielen Stellen wird gezeigt, wie jeder Datenbetrügereien auf die Spur kommen kann oder wie Mann oder Frau es vermeiden kann, selbst betrogen zu werden. Damit wird eine Brücke zur forensischen Kriminalistik geschlagen, soweit sie sich auf numerische Daten bezieht. Dazu rechnen auch die Tätigkeitsbereiche von Compliance-Beauftragten, Justiziarern in großen Unternehmen sowie Kriminalbeamten in den regionalen LKAs. Im Kern sind die einzelnen Tatbestände weiß Gott nicht neu, in ihrer Gesamtschau überraschen sie und sind gleichermaßen abschreckend.

Dieses Buch wäre so nicht zustande gekommen, hätte der Verfasser es während seiner gesamten beruflichen Laufbahn und auch danach nicht immer wieder mit Datenqualität und der falschen Nutzung von Daten zu tun bekommen. Als Hochschullehrer für Statistik und Wirtschaftsinformatik an der Freien Universität Berlin betraf das Lehre, Forschung und Beratung Dritter. Hinzu trat die Erfahrung aus Gastvorträgen in aller Welt und den von 2017–2022 jährlich durchgeführten Ganztageskursen *Betrugserkennung aus Datenbeständen* im Rahmen des Masterstudiengangs *Forensische Kriminalistik* an der Steinbeis Hochschule, Berlin. Für den dazu nötigen Gedankenaus-

tausch, die Thematik zu beschreiben und einzugrenzen, bin ich deren ehemaliger Leiterin, Frau B. Galley, Berlin, zu sehr großem Dank verpflichtet. Den eigentlichen Anstoß, sich noch intensiver mit der Datentrickserei zu befassen, gab schließlich Hermann Engesser im Jahr 2016, damals Programmleiter des Fachjournals Informatik-Spektrum beim Springer Verlag, Heidelberg. Er hatte bei einem gemeinsamen Arbeitsfrühstück in Berlin, bei dem wir über letzte Details zu meinem Buchprojekt *Business Intelligence* sprachen, die Idee, Glossen über Datenbetrug für die neu geschaffene Kolumne *Achtung. Datentrickserei!* im Informatik Spektrum vom Verfasser schreiben zu lassen. Dreizehn Glossen habe ich seitdem publiziert. Ihm gilt mein aufrichtiger Dank. Dieser Dank gilt ebenso seinem Nachfolger, Chefredakteur Peter Pagel, der zusammen mit Professor Dr. Thomas Ludwig, Deutsches Klimarechenzentrum, Hamburg, mich seitdem tatkräftig unterstützte, weitere Glossen thematisch festzulegen. Beiden verdanke ich etliche fachliche Verbesserungsvorschläge meiner Manuskripte.

Angeregt durch Diskussionen und das Interesse an meiner Kolumne im Informatik Spektrum, das meine Kollegen aus der Informatik als Leser mir bekundeten, entschloss ich mich, die bislang erschienen dreizehn Glossen zu einem Buch über Datenbetrug zusammenzufassen. Die Chefredaktion des Informatik Spektrums gab spontan ihre Zustimmung. Eine reine Aneinanderreihung der publizierten Artikel schloss ich aus. Stattdessen habe ich jeden einzelnen Beitrag überarbeitet, teilweise aktualisiert oder, falls notwendig, inhaltlich erweitert sowie durch Diagramme und Fotos ergänzt. Die Überarbeitung trifft insbesondere auf die Literaturhinweise zu, die in dieser Form anfangs bewusst nicht vollständig enthalten waren, da der Glossencharakter der dreizehn Kolumnen im Vordergrund stand.



Die Autorenrechte an Fotos zu wahren, erwies sich als ein unerwartetes mühseliges Puzzlespiel. Denn es galt, Ketten von Rechteinhabern bei Verlagen und Agenturen hartnäckig bis zur „Quelle“ zu verfolgen.

Bei der Durchsicht des Manuskripts hat mich mein Kollege Prof. Dr. U. Kockelkorn auf ein paar Unklarheiten im Text aufmerksam gemacht. Meiner früheren Klassenkameradin, der ehemaligen Redakteurin beim Südwestdeutschen Rundfunk, Frau B. Wenke, danke ich sehr dafür, dass sie das Skript umfassend kritisch, „quasi redaktionell“ durchsah.

Schließlich muss ich auch meine Alma Mater, die Freie Universität Berlin, erwähnen. Seit meiner Verabschiedung 2008 und bis zum heutigen Tag überlässt sie mir einen Arbeitsraum und einen Internetzugang.

Das Buch verdeutlicht insgesamt gesehen Breite und Vielfalt der Datentrickserei, umfangreicher als es einzelne Glossen für sich allein gesehen hätten erreichen können. Wie sagte doch unser Dichturfürst Wolfgang von Goethe im Faust Teil I so treffend: „Dann hat er die Theile in seiner Hand, fehlt leider nur das geistige Band“. Hiermit liefere ich nun das geistige Band nach!

Ich danke Herrn P. Pagel, Frau H. Jung, Herrn R. Ravindran und Frau P. Steinmüller, Vieweg/Springer Verlag, Heidelberg, aufrichtig dafür, dass sie nicht nur das Manuskript wohlwollend begrüßten, der Veröffentlichung zustimmten, sondern auch die redaktionelle und technische Drucklegung in die Wege geleitet und bis zur Herstellung fortlaufend unterstützt haben.

Berlin-Dahlem  
Dezember 2023

Hans-J. Lenz

# Inhaltsverzeichnis

<b>1</b>	<b>Die Fälscher sind unter uns – Money makes the world go ‚round‘</b>	<b>1</b>
<b>2</b>	<b>Der Libor-Skandal – Kleinvieh macht auch Mist</b>	<b>9</b>
<b>3</b>	<b>‚Datendiebstahl – Mach‘ die Augen auf oder den Beutel</b>	<b>17</b>
<b>4</b>	<b>Datentrickserei in der Wissenschaft – Ehrlich währt am Längsten</b>	<b>29</b>
<b>5</b>	<b>Krumme Dinger – Datentrickserei in Wirtschaft und Gesellschaft</b>	<b>53</b>
<b>6</b>	<b>Lug und Trug – Schattenseiten im Sport</b>	<b>63</b>
<b>7</b>	<b>Wahlbetrug – Wer die Wahl hat, hat die Qual</b>	<b>79</b>

<b>XII</b>	<b>Inhaltsverzeichnis</b>	
<b>8</b>	<b>Datentrickserei im Taxigewerbe – von Abzocke bis Steuerbetrug</b>	<b>93</b>
<b>9</b>	<b>Bilanzbetrug – Schönen, Frisieren, Betrügen</b>	<b>99</b>
<b>10</b>	<b>Betriebs- und Heizkostenabrechnungen – Schein und Sein</b>	<b>109</b>
<b>11</b>	<b>Abrechnungsbetrug in der Gastronomie – Süßer die Kassen so klingen</b>	<b>121</b>
<b>12</b>	<b>Abrechnungsbetrug im Gesundheitswesen: Die Geschäfte laufen wie geschmiert</b>	<b>131</b>
<b>13</b>	<b>Fake Science – zwielichtige Wissenschaft</b>	<b>147</b>
	<b>Literaturnachweis Kolumnen im Informatik Spektrum</b>	<b>163</b>



# 1

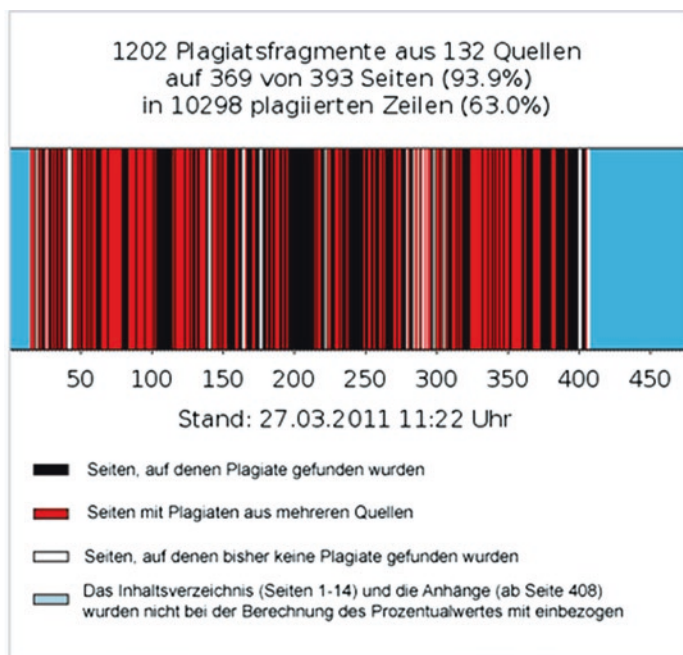
## Die Fälscher sind unter uns – Money makes the world go ,round‘

Die oft mit Datenbetrug gleichgesetzte Datenmanipulation stellt einen speziellen Fall des viel weitergehenden, umfassenden Datenbetrugs dar, der in seiner Breite in diesem Kapitel nur angerissen werden soll. Die Details dahinter beleuchten wir in den nachfolgenden Kapiteln.

Im digitalen Zeitalter von Globalisierung, Netzwerken, Industry 4.0, Google, Big Data und Künstlicher Intelligenz umfasst der Datenbetrug zunächst einmal den *Datenklau*, auf den sich nicht nur Geheimdienste aus aller Welt spezialisiert haben, sondern wo neben den großen Tech-Firmen auch namhafte Industrie- und Wirtschaftsunternehmen recht aktiv waren und sind. Nutzen sie doch die anfallenden Transaktionsdaten ihrer Kunden oft stillschweigend zu umfassenden Datenanalysen für Marketingzwecke und für die Kundenprofilbildung. Die Deutsche Post, aber nicht nur die, steht stellvertretend für alle, die sich in einem juristischen Graubereich eines schwunghaften Handels mit Adressdaten befassen.

Zweitens sind *Plagiate* zu erwähnen. In Form der Produktpiraterie ist dies in der realen Welt grundsätzlich nichts Neues. Auch die Musikbranche und die Malerei sind von illegalen Nachahmern betroffen. Besonders ausgebufft erscheinen Musikstücke im Netz, die die Musik Dritter plagiiert, indem die Tonfolge schlicht rückwärts gespielt wird. Im Gegensatz zur Produktpiraterie fällt die digitale Variante jedoch viel leichter. Die Doktorarbeit des ehemaligen deutschen Verteidigungsministers von Guttenberg ist nahezu vollständig abgeschrieben. „1202 Plagiatsfragmente aus 132 Quellen auf 369 von 393 Seiten (93,9 %) in 10.298 plagiierten Zeilen (63 %)“ [1]. Die Betrügerei des Politikers hat immerhin dazu geführt, dass neuere Methoden der Text- und Zitieranalyse entwickelt wurden, um weitere plagiatsverdächtige Doktorarbeiten zu überprüfen. Die neuen Analyseverfahren führen zu geringeren Fehlerraten als die traditionellen Ansätze der Plagiatsprüfung [2]. Denn sie beruhen nicht nur auf reinen Textvergleichen, sondern analysieren zusätzlich die Zitatestruktur von Original und Plagiat. Die Abb. 1.1 zeigt eindrucksvoll, in welchem Umfang der Politiker plagiierte.

Gleichzeitig stellt von Guttenbergs Dissertation nunmehr eine Messlatte dar, mit der Methoden der Plagiatsaufdeckung hinsichtlich der Fehlklassifikationsraten zu überprüfen sind. Denn die Wahrscheinlichkeit, einen Unschuldigen an den Pranger zu stellen, sollte ebenso klein sein wie die, einen Plagiator durchkommen zu lassen. Ein positiver Nebeneffekt dieser Anstrengungen, solche Missetäter zu entlarven, ist übrigens auch die damit verbesserte Aufklärungsrate von *Selbstplagiaten*, wo Wissenschaftler Teile ihrer eigenen Texte, Diagramme, Tabellen oder Fotos – ggf. in leicht geänderter Form – erneut publizieren, ohne allerdings die ursprüngliche, eigene Quelle anzugeben. Im akademischen Umfeld machen dies unter anderem Nach-



**Abb. 1.1** Spektrum des Plagiiens in der Dissertation von Karl-Theodor zu Guttenberg [1]

wuchskräfte, um die Anzahl ihrer Publikationen zu erhöhen, und erhoffen sich so, ihre Karrierechancen zu verbessern. Das Schlagwort „Publish or perish“ – „Publiziere oder versage“ – macht das verhängnisvolle Spannungsfeld im heutigen Wissenschaftsbereich deutlich. Auch die Unart, aus eigentlich einer umfänglichen Arbeit zwei Publikationen zu machen, gehört in diesen Zusammenhang.

*Datenmanipulation* bildet den dritten großen Bereich des Datenbetrugs. Sie umfasst schon lange nicht mehr nur Telefongebühren-, Kredit- und Scheckkartenbetrug. Vielmehr bietet die Digitalisierung zusammen mit der Vernetzung und Integration von Daten in vielen Lebens-

bereichen bislang ungeahnte „Geschäftsmodelle“ für Betrüger. Der Zugriff auf offene, große Datenmengen in den sozialen Netzwerken, auf die internen Speicher der Assistenzsysteme, die in Kraftfahrzeuge und Haushalte eingebaut sind, hat quasi die Pandora-Büchse des Datenbetrugs geöffnet. Motiv ist auch hier überwiegend: Prestige, Wut und Geldgier. Betroffen vom Schönfärben und Frisieren sind nicht nur experimentell gewonnene Daten, sondern auch Langzeit-, Querschnitts- und Paneldaten, gesammelt bzw. ausgewertet von Staaten (Griechenland), Unternehmen (Deutsche Bank, Deutsche Post), Universitätskliniken (Klinikum rechts der Isar), Fernsehanstalten (ZDF) und Vereinen (ADAC). Datenmanipulationen wurden selbst bei kirchlichen Organisationen ruchbar, siehe beispielsweise die unchristliche Geldwäsche des Vatikans. Die Wahl 2011 in Russland steht stellvertretend für Wahlfälschungen in vielen anderen Ländern der Welt. Auf Wahlfälschung in seiner vollen Breite kommen wir in Kap. 7 noch zu sprechen. In der universitären und außeruniversitären Forschung ist seit jeher Datenmanipulation im Spiel gewesen. So haben beispielsweise bereits Galilei oder Newton, meine Heros in den Naturwissenschaften, ihre Messreihen aus Plausibilitätsgründen den theoretisch richtigen Funktionsverläufen angepasst. Allerdings sollte die seinerzeit nicht sehr genaue Zeitmesstechnik dabei mitberücksichtigt werden. Die Trickkiste umfasst weiterhin vorsätzlich oder aus mangelndem statistischem Fachwissen fehlerhaft entworfene Studien oder Reports. Dazu rechnen auch etliche Presseveröffentlichungen und amtliche Verlautbarungen, die sich auf das regionale Infektionsrisiko während der Corona-Pandemie bezogen. Die gemeldeten Infektionsfälle wurden schlicht auf die falsche Teilgesamtheit der Bevölkerung bezogen. Zur gängigen Datenmanipulation rechnen das Weglassen „ungeeigneter“ Mess- oder Beobachtungswerte oder das auf das Untersuchungs-

ziel ausgerichtete Erfassen weiterer Daten. Hinzu kann die Verzerrungen auslösende Auswahl von Gegenstands- sowie Orts- und Zeitbereich bei empirischen Studien kommen, die Art und Weise der Ziehung einer Stichprobe – mit der Gefahr einer völligen Verzerrung der Ergebnisse, statistisch „Sampling Bias“ genannt. Der Werkzeugkasten der Statistik bietet dem Unkundigen weiterhin ein breites Angebot an Untersuchungsmethoden. Hier seien nur Datentransformationen in Form von „Ausreißer-“, Trend-, Konjunktur- oder Saisonbereinigungen sowie die breite Palette statistischer Signifikanztests, die nicht immer auf Stichproben beruhen, genannt. Mit solchen Tests lassen sich empirische Hypothesen statistisch überprüfen, vorausgesetzt, die Verfahren werden sachlich korrekt und nachprüfbar angewendet.

Vorsicht war geboten bei Googles Parole „ $N = \text{All}$ “, wir brauchen sozusagen keine Stichproben mehr, wir erfassen alle Daten! Denn sie täuscht die Vollständigkeit der Untersuchungseinheiten nur vor, wie bei der Prognose von Epidemien in den USA geschehen. Die Voraussage basierte auf allen Google zugänglichen Daten, insbesondere Tweets. Man muss beachten, dass Twitter-Nutzer nicht die Gesamtheit aller Personen repräsentieren, sondern als Auswahlgesamtheit altersbedingt nur eine echte Teilgesamtheit, eben die, die Tweets überhaupt verschicken. Denn weder Kleinkinder noch Senioren, vielleicht bis auf eine Minderheit, wurden erfasst. Haben Sie, lieber Leser, schon einmal Kurznachrichten über Twitter, seit kurzem in X umbenannt, verschickt? Ein „beliebtes“ Mittel zum (erzwungenen) Signifikanz-Nachweis einer Vermutung oder Hypothese ist übrigens das ergebnisabhängige Heraufsetzen vom eigentlich vorgegebenen Signifikanzniveau eines Tests, d. h. statt 1 % auf 5 % oder 10 %, oder sogar die „geeignete“ Anpassung der Prüfgröße samt Test. Ein Erzübel, das kaum ausrottbar zu sein scheint, ist das wie-