

# Internationale Harmonisierung von Unfalldaten

## Fortschrittsbericht des FIA / ACEA Projekts iGLAD (Initiative for the Global Harmonization of Accident Data)

Dipl.-Ing. **D. Ockel**, Daimler AG, Sindelfingen;  
Dipl.-Math. **J. Bakker**, Daimler AG, Sindelfingen;  
Prof. Dr.-Ing. **R. Schöneburg**, Daimler AG, Sindelfingen;

### Kurzfassung

Während die Zahl der Unfallopfer in Deutschland und Europa seit Jahren fällt, stellt die globale Verkehrsicherheit ein immer größer werdendes Problem dar. Die WHO schätzt die jährliche Zahl der weltweiten Unfallopfer auf bis zu 1,3 Mio. Getötete bei ca. 20-50 Mio. Verletzten, während für Verkehrsunfälle in der Rangliste der häufigsten Todesursachen ein Anstieg von Platz 9 in 2004 auf Platz 5 in 2030 prognostiziert wird. Maßnahmen in einem oder mehreren der Felder Fahrzeug, Infrastruktur, Fahrerverhalten und Reglementierung sind also dringend erforderlich, wobei völlig unterschiedliche Ansätze regional zielführend sein werden: Während eine Trennung von Verkehrswegen hier eine überwältigend große Wirksamkeit haben kann, kann dort nur noch ein aufwändiges Assistenz- oder Rückhaltesystem einen weiteren Beitrag leisten. Hilfe bei der Priorisierung von Maßnahmen nach Aufwand-Nutzen Gesichtspunkten bieten belastbare, also detaillierte und national repräsentative Unfalldatenbanken, wie sie in Projekten zur „In-Depth“ Unfalldatenerhebung, wie z.B. GIDAS (German In-Depth Accident Study), befüllt werden.

Ähnliche Projekte zur Unfalldatenerhebung finden erfreulicherweise aktuell weltweit Verbreitung. Handlungsbedarf besteht nun beim Umfang und der Vergleichbarkeit regional erhobener Daten, um einerseits einen übergreifenden minimalen Qualitätsstandard sicher zu stellen, andererseits eine sinnvolle Gegenüberstellung der Wirksamkeiten von Maßnahmen in verschiedenen Märkten zu ermöglichen, oder aber anhand von Pre-Crash Daten die Effizienz neuartiger Sicherheitssysteme zu prognostizieren.

Das UN Programm „Decade of Action for Road Safety 2011-2020“ unterstützend, haben FIA und ACEA ein gemeinsames Projekt „iGLAD“ zur weltweiten Harmonisierung von Unfalldaten auf den Weg gebracht, um mit internationalen Experten der Unfalldatenerhebung eine Einigung auf einen gemeinsamen Datensatz zu erzielen und lokal spezifische, ergänzende Parameter zu systematisieren. Weiteres Ziel ist es, Erfahrungen zu

Finanzierungsmodellen zur Sicherstellung langfristig tragfähiger Projekte auszutauschen und daraus lokal adaptierte Handlungsempfehlungen abzuleiten. Dieser Beitrag erläutert die Fortschritte dieser Bemühungen unter Berücksichtigung bereits bestehender Ansätze im europäischen Umfeld.

## Motivation

Die Anzahl und Entwicklung der Verkehrstoten weltweit kann nur grob geschätzt werden und beträgt für 2010 zwischen 900.000 und 1.300.000. Der weitere Verlauf dieser Zahlen innerhalb der nächsten 10 Jahre hängt von vielen Faktoren ab. Die bisherige Entwicklung in „High-income countries“ lässt darauf schließen, dass insbesondere durch Arbeit an Infrastruktur, Fahrzeugtechnik und Gesetzgebung ein wesentlicher Beitrag zur Verbesserung der Situation geleistet werden kann. Auch erste „In-Depth“ Unfalldaten aus Ländern wie Indien und China bestätigen dies. Abbildung 1 zeigt die Entwicklung der Verkehrstoten von 1990 ab bis 2010 als prognostische Weiterführung nach dem TFEC Modell [1]. Das Traffic Fatalities and Economic Growth (TFEC) Projekt der World Bank entwickelte dieses Modell auf Basis von Verkehrs-, Bevölkerungs- und ökonomischen Daten [2].

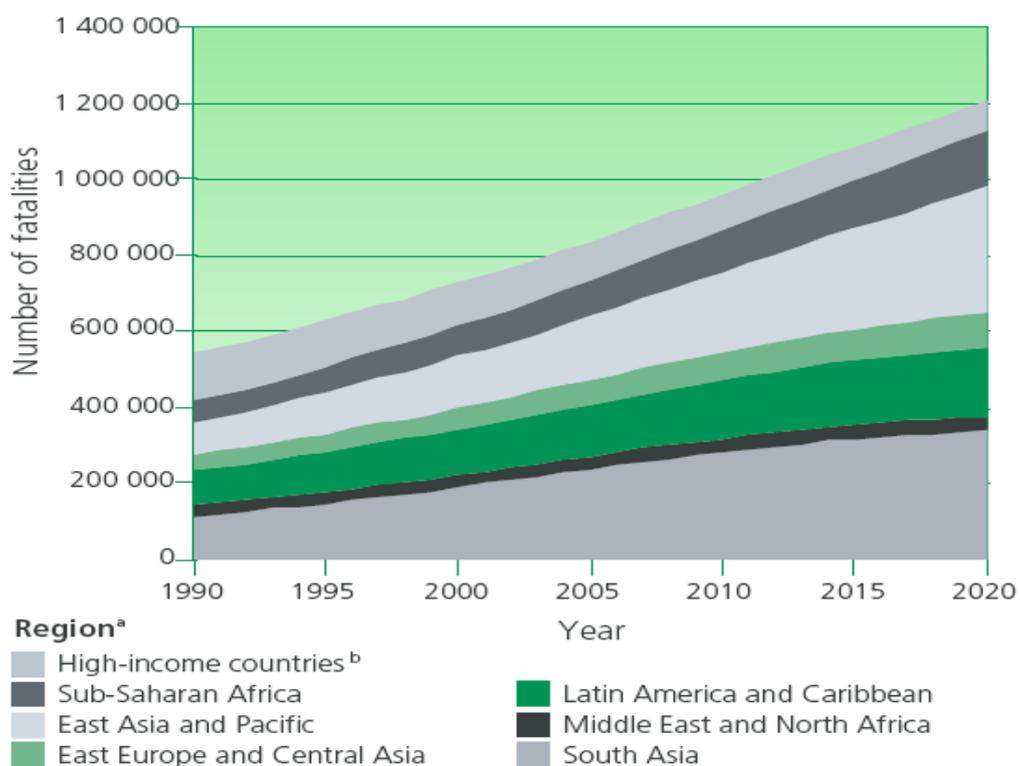


Abbildung 1: Entwicklung der Verkehrstoten weltweit (Quelle: WHO Bericht 2004).

Schon bei diesen grundlegenden Betrachtungen wird ersichtlich, dass es ein erhebliches Defizit an Daten gibt, um eine differenzierte und umfassende Einschätzung der Situation zu ermöglichen. Auf der anderen Seite hat sich herausgestellt, dass eine gute und detailreiche Datenbasis notwendig ist, um gezielt und effektiv die Verkehrsunfallsituation zu verbessern. Insbesondere bei der Verbesserung der Fahrzeugsicherheit sind Realunfalldaten eine wesentliche Grundlage. Aus diesem Grund haben sich in „High-Income countries“ schon seit einigen Jahren Projekte etabliert, die eine umfassende Erhebung von Unfalldaten durchführen.

Die Europäische Union hat sich mit dem Road Safety Action Programme [3] die Halbierung der Verkehrstoten jeweils in der ersten und zweiten Dekade dieses Jahrtausends zum Ziel gesetzt. Eine erfolgreiche Umsetzung wird zunehmend schwieriger, da die Wirkung einzelner Maßnahmen geringer werden, der Aufwand hingegen steigt. Deshalb ist es umso wichtiger, das Verkehrsunfallgeschehen genau zu analysieren und dafür eine geeignete europaweite Datengrundlage zu schaffen.

Deutlich wird dies bereits in dem 2010 eingeführten Euro NCAP Advanced reward [4] für neue Entwicklungen von Sicherheitssystemen in Pkw. Dieses Programm möchte vielversprechende Techniken honorieren, die eine weitere Verringerung der Anzahl der Straßenverkehrstoten ermöglichen. Eine Bewertung der Systeme erfolgt aufgrund von Effektivitätsanalysen, die auf Basis von Unfalldaten erstellt werden. Hier zeigt sich schnell, dass eine genaue Hochrechnung einer solchen Effektivität auf Europa nur mit umfassenden Daten aus verschiedenen europäischen Ländern möglich ist, da sich das Verkehrsunfallgeschehen in den einzelnen Ländern erheblich unterscheiden kann.



Abbildung 2: Eine Auswahl von „In-Depth“ Unfalldatenerhebungen weltweit.

In-depth Unfallerhebungen konnten sich schon in einigen Ländern weltweit über einen längeren Zeitraum etablieren. Vor allem in neuen Märkten sind zur Zeit weitere Erhebungen im Entstehen. Abbildung 2 gibt einen Überblick über eine Auswahl von „In-Depth“ Erhebungen weltweit.

Die einzelnen Erhebungsprojekte unterscheiden sich stark in Anzahl der aufgenommenen Fälle pro Jahr, Erhebungsumfang (Anzahl der Faktoren), Stichprobencharakteristik, Organisation und Finanzierung. Dies erschwert den Vergleich von Daten aus verschiedenen Ländern erheblich und z.B. Analysen, die Aussagen für eine größere Region wie Europa treffen sollen, sind zum einen recht ungenau und erfordern zum anderen ein umfassendes Wissen mehrerer Datenbanken und deren Analyse.

Die nationalen Statistiken eines Landes erfüllen in diesem Kontext zwei wichtige Funktionen: Sie bilden ein verbindendes Element zwischen den „In-Depth“ Stichproben des jeweiligen Landes und sie ermöglichen als Randverteilung eine Hochrechnung der „In-Depth“ Daten auf die jeweilige Nation. Eine Grundvoraussetzung für die Hochrechnung auf nationale Ebene ist, dass die Variablen der nationalen Daten auch in der „In-Depth“ Stichprobe vorhanden sein müssen. Vereinheitlichte und harmonisierte nationale Daten ermöglichen dann eine Hochrechnung auf multinationale Ebenen.

Der Harmonisierungsprozess bei nationalen Daten ist glücklicherweise schon recht fortgeschritten. Projekte wie IRTAD [5] und CARE [6] haben hier bereits sehr gute Arbeit geleistet und liefern elementare Größen wie Anzahl der Getötete, Unfälle mit Getöteten, Art der Verkehrsbeteiligung für unterschiedliche Länder in standardisierter Form, so dass z.B. auch die Definition von „getötet im Straßenverkehr“ vereinheitlicht ist. Durch Ergänzung von nationalen Daten mit harmonisierten „In-Depth“ Daten ist eine Erweiterung des Analysefeldes möglich. Harmonisierte „In-Depth“ Daten und nationale Daten zusammen ermöglichen einen Vergleich verschiedener Länder und eine Identifikation von Schwerpunkten in einzelnen Ländern.

Darüber hinaus kann im erweiterten Kreis eines multinationalen Konsortiums von „In-Depth“ Projekten eine Kommunikation entstehen, um auf bewährte Prinzipien und Erfahrungen aus dem Bereich der Verkehrssicherheit in anderen Ländern zurückgreifen zu können, mit dem Ziel, Stellhebel zur Verbesserung der globalen Verkehrssicherheit schnell und effektiv zu identifizieren. Nicht zu vernachlässigen ist auch, dass ein vereinfachter Zugang zu, und verbesserte Analysemöglichkeiten mit einem großen multinationalen Datenpool, treibende Kräfte für eine übergreifende Forschung auf dem Gebiet der Fahrzeugsicherheit sein können. Bisherige Ansätze zum Aufbau eines multinationalen Pools von „In-Depth“ Unfalldaten zielen auf sehr umfangreiche Datenschemata ab [7], [8] und bauen nicht auf

bereits vorhandene, etablierte Erhebungen auf. Daraus resultiert ein erheblicher Aufwand, sowohl finanziell als auch zeitlich, um eine profunde Datenbasis zu erstellen. Dies stellt eine erhebliche Hürde dar, da typischerweise erst nach einigen Jahren eine ausreichende Datenbasis vorliegt, und deshalb die Finanzierung von „In-Depth“ Unfallprojekten langfristig auf eine solide Basis gestellt werden muss. Durch die Verwendung bereits vorhandener Daten steht von Beginn an ein Datenfundus zur Verfügung.

### **Methodische Aspekte**

Ausgangssituation für eine Harmonisierung sind verschiedene „In-Depth“ Datenerhebungen, die sich in Anzahl, Art und Ausprägungen der Faktoren voneinander unterscheiden. Für die Harmonisierung der Daten soll ein gemeinsames Datenschema erzeugt werden, das als zusätzliche Schicht auf die unterschiedlichen Datenschemata aufgesetzt wird (Abbildung 3). Dieses gemeinsame Datenschema muss genau spezifiziert werden und bildet in Verbindung mit den in dieses Schema überführten Daten den Standarddatensatz. Die Anzahl und Ausprägungen der Faktoren des Standarddatensatzes sind im Mittelfeld zwischen nationaler Statistik und „In-Depth“ Datenerhebung anzusiedeln. Der Standarddatensatz beschränkt sich also auf die wesentlichsten Inhalte. Dies erleichtert die Erzeugung des Standarddatensatzes aus einer bereits vorhandenen „In-Depth“ Erhebung, bietet jedoch auf der anderen Seite nicht die Detailtiefe wie der ursprüngliche „In-Depth“ Datensatz. Sollte bei einer Analyse der Standarddatensatz nicht ausreichen, muss immer noch auf den jeweiligen „In-Depth“ Datensatz zurückgegriffen werden, wobei der Standarddatensatz als Zwischenschicht die weitere Analyse erleichtern kann.

Ohne die Verwendung eines Standarddatensatzes stellt eine vergleichende Auswertung verschiedener „In-Depth“ Datensätze einen erheblichen Aufwand dar, der zusätzlich sehr fehleranfällig ist, da oft die genauen Kontextinformationen der Kodierung der Daten dem Datenanalysten nicht geläufig sind. Wird hingegen jeder einzelne Datensatz zunächst auf eine einheitliche Form gebracht, fällt dieser Aufwand nur einmal an und der so resultierende umfassende Datensatz kann leicht analysiert werden. Später hinzukommende Daten können (semi-) automatisch in den Standarddatensatz transformiert werden. Hinzu kommt, dass die Transformation des einzelnen Datensatzes in den standardisierten Datensatz von den Experten erzeugt wird, die die Daten auch erheben, was zu qualitativ besseren und konsistenteren Ergebnissen führt. Insbesondere können so regionale Unterschiede bei der Interpretation der Faktoren genauer und mit dem richtigen Kontextwissen berücksichtigt werden.

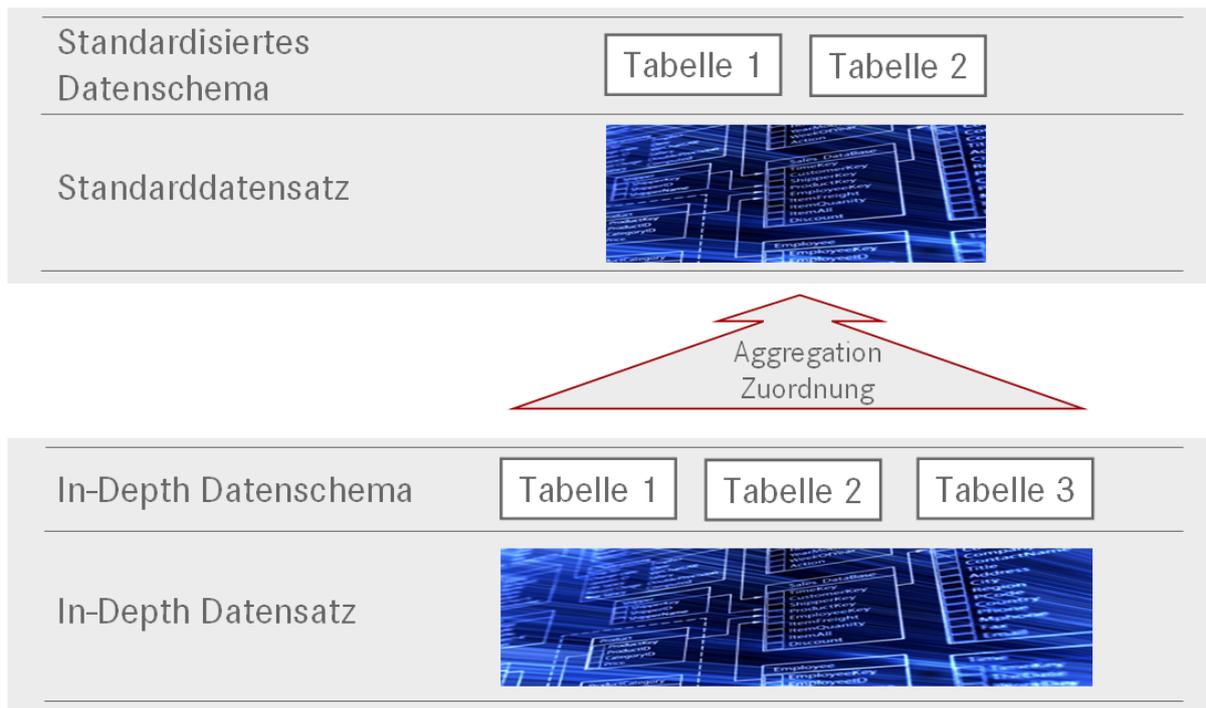


Abbildung 3: Standarddatensatz als zusätzliche Schicht auf dem „In-Depth“ Datensatz.

Die Transformation eines Datensatzes in den Standarddatensatz kann in zwei Grundschritten erfolgen: die Abbildung der einzelnen Faktoren des Datensatzes und die Homogenisierung der Stichprobe.

Schritt 1, Abbildung von Faktoren und deren Ausprägungen auf den Standarddatensatz:

- 1.1 Identifizieren der inhaltlich vergleichbaren Faktoren im „In-Depth“ Datensatz und im Standarddatensatz.
- 1.2 Finden von Gemeinsamkeiten und Unterschieden in Kontext und Beschreibung der im letzten Schritt identifizierten Faktoren.
- 1.3 Definieren eines Abbildungsverfahrens vom „In-Depth“ Datensatz zum Standarddatensatz unter Verwendung von Aggregationen und Merkmalszuordnungen.
- 1.4 Implementieren des Abbildungsverfahrens aus dem letzten Schritt in einem Tool zur (semi-) automatischen Generierung des Standarddatensatzes.
- 1.5 Dokumentieren des Abbildungsverfahrens.
- 1.6 Anwenden des Tools auf den gesamten „In-Depth“ Datensatz.
- 1.7 Validierung des erzeugten Datensatzes anhand von einigen ausgewählten Faktoren.

Schritt 2, Homogenisierung der einzelnen Stichproben innerhalb des Standarddatensatzes:

Falls eine nationale Statistik zur Verfügung steht, und ausreichend gemeinsame Faktoren mit dem „In-Depth“ Datensatz vorhanden sind (z.B. Ortslage, Unfallschwere, Art der Beteiligung, Tageszeit, Unfalltyp), können die Randverteilungen der gemeinsamen Faktoren mit geeigneten statistischen Methoden angeglichen werden, wie z.B. dem „Iterative Proportional Fitting“ [9]. Daraus können Wichtungsfaktoren berechnet werden, mit denen die gesamte Stichprobe an den Standarddatensatz angepasst werden kann. Stellt der „In-Depth“ Datensatz bereits eine eher repräsentativ gezogene Stichprobe dar (bezogen auf die nationale Statistik als Gesamtheit), kann dieser Schritt nahezu überflüssig werden. Eigenschaften einer Stichprobe, die die Repräsentativität bzgl. der nationalen Statistik begünstigen, sind z.B. identische Selektionskriterien (z.B. alle Unfälle mit Personenschaden), ein guter Querschnitt der geographischen und infrastrukturellen Gegebenheiten der Gesamtregion (des Landes) oder eine Zufallsstichprobe (z.B. durch zeitliche Stichprobenziehung mit alternierenden Schichten). Ferner sollte die Größe der Stichprobe ausreichend sein. Als grober Anhaltspunkt hierfür kann man ca. 500 Fälle pro Jahr im „In-Depth“ Datensatz ansetzen.

### **Organisatorische Aspekte**

Ziel von iGLAD ist es, einen multinationalen Datenpool zu erstellen. Dazu ist die Bildung eines Konsortiums notwendig, in dem verschiedene Projekte vertreten sind, die Unfallerhebungen durchführen. Zusätzlich sollten Vertreter von nationalen Statistiken oder Organisationen, die übergreifend nationale Statistiken harmonisieren beteiligt sein. Durch Einspeisung einer ausreichenden Menge von eigenen Daten in den Datenpool soll der Zugriff auf die gesamten Daten des Pools ermöglicht werden. Somit wird der Datenpool für die Beteiligten größer und attraktiver, je mehr Unfalldatenprojekte beteiligt sind.

Das iGLAD Konsortium sollte sowohl den Bestand des Datenpools sichern und organisatorische und technische Probleme adressieren, als auch eine Informationsplattform für beteiligte Institutionen darstellen, und so zu einem leichterem Austausch von Erfahrungen und Best Practices führen.

Die jeweiligen „In-Depth“ Erhebungen vor Ort beinhalten im Allgemeinen einen erheblich größeren Umfang an Variablen und auch innerhalb der Variablen eine größere Detailgenauigkeit. Für tiefer gehende Studien wird deshalb der jeweilige Originaldatensatz benötigt. Dadurch wird eine attraktive Plattform für die Organisationen der lokalen „In-Depth“ Erhebungen geboten, um Studien oder Daten anbieten zu können und damit die eigene Erhebung finanziell zu stärken.

Die Sicherung einer konsistenten Qualität innerhalb des Datenpools sollte hoch priorisiert werden. Dazu sollten einheitliche Plausibilitätschecks und weitere Prüfungsmechanismen aufgestellt werden und in einem Qualitätskriterienkatalog gesammelt werden. Für eine regelmäßige Überprüfung der Qualität ist die Einführung einer Zertifizierung vorstellbar. Beim Aufbau neuer Erhebungen können Erfahrungswerte aus dem Konsortium einfließen. Auch aus technischer Perspektive gewinnen neue Erhebungen einen Vorsprung durch bereits vorhandene Dokumentation (Datenschema) und Expertise aus dem Konsortium.

### **Zusammenfassung und Ausblick**

Die globale Verkehrssicherheit rückt mit dem stark wachsenden Automobilmarkt in den Schwellenländern immer mehr in den Vordergrund. Während in technologisch und infrastrukturell weiter entwickelten Ländern die Anzahl der Getöteten im Straßenverkehr kontinuierlich rückläufig sind, ist in Schwellenländern ein entgegengesetzter Trend zu beobachten. Das Projekt iGLAD adressiert diese Problematik aus Sicht der Unfallforschung, in dem die Datenbasis von Straßenverkehrsunfällen weltweit harmonisiert und vergrößert werden soll. Dies soll dazu führen, dass die Stellhebel zur Verbesserung der globalen Verkehrssicherheit schnell und effektiv identifiziert werden können. Das Projekt befindet sich noch in der organisatorisch konzeptionellen Phase. Nächste Schritte sind die Bildung eines geeigneten Konsortiums und die technische Umsetzung des Standarddatensatzes.

- [1] M. Peden et. al., WHO, World report on road traffic injury prevention, 2004.
- [2] E. Kopits, M. Cropper, Traffic fatalities and economic growth. Washington, DC, The World Bank, 2003 (Policy Research Working Paper No. 3035).
- [3] EU Commission, European Road Safety Action Programme, ISBN 92-894-5893-3, 2003.
- [4] Euro NCAP Advanced reward, <http://www.euroncap.com/rewards.aspx>, 2010.
- [5] IRTAD, International Traffic Safety Data and Analysis Group, International Transport Forum, OECD, <http://internationaltransportforum.org/irtad/index.html>.
- [6] CARE, European Road Safety Observatory, Safetynet (EU Commission DG-TREN), [http://erso.swov.nl/safetynet/content/wp\\_1\\_care\\_accident\\_data.htm](http://erso.swov.nl/safetynet/content/wp_1_care_accident_data.htm).
- [7] S. Reed, A. Morris, Glossary of data variables for fatal and accident causation databases, Deliverable D5.5 of the EU FP6 project SafetyNet contract no SI2.395465/506723, 2008.

- [8] G. Vallet et. al., STAIRS Standardisation of Accident and Injury Registration Systems, EU 4<sup>th</sup> Framework Programme Final Report, 1999.
- [9] W.E. Deming, F.F. Stephan, On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known, *Annals of Mathematical Statistics* 11 (4): 427–444. doi:10.1214/aoms/1177731829. MR3527, 1940.