DIE WIRTSCHAFTSPRÜFUNG
VEREINIGT MIT DEN ZEITSCHRIFTEN „DER WIRTSCHAFTSPRÜFER" UND „DER WIRTSCHAFTSTREUHÄNDER"
Der Berater der Wirtschaft
Betriebswirtschaftliches Archiv und Fachorgan für das wirtschaftliche Prüfungs- und Treuhandwesen
Herausgegeben vom Institut der Wirtschaftsprüfer in Deutschland e. V.

Jahrgang 57 | Düsseldorf, 1. Oktober 2004 | Nummer 19

# Zur Beurteilung des Fraud-Risikos im Rahmen der Abschlussprüfung

*Von StB Stephan Knabe, Dr. Sebastian Mika, Prof. Dr. Klaus-Robert Müller, Dr. Gunnar Rätsch und WP Prof. Dr. Wienand Schruff, alle Berlin*

## 1. Probleme bei der Aufdeckung von Fraud und die Verantwortung des Abschlussprüfers

Eine von September bis November 2003 durchgeführte Umfrage[1] unter 1000 großen deutschen Unternehmen aus allen Wirtschaftsbereichen ergab, dass rund zwei Drittel der befragten Unternehmen nach eigenen Angaben in den letzten drei Jahren Opfer wirtschaftskrimineller Handlungen wurden. Die Einzelschäden betrugen dabei bis zu 85 Millionen Euro. Vier von fünf Unternehmen erwarten sogar, dass das Ausmaß wirtschaftskrimineller Delikte hierzulande in nächster Zeit noch steigen wird.

Für die spektakulären Fraud[2]-Fälle der letzten Jahre[3] wurden in den Augen der Öffentlichkeit unzureichende Prüfungsleistungen als eine der Ursachen gesehen. Zwar liegt die primäre Verantwortung für die Vermeidung oder ggf. die Aufdeckung von Fraud beim Management, jedoch waren in den genannten Fällen Mitglieder des Top-Managements als Täter oder Mittäter involviert. Im Ergebnis werden in der öffentlichen Diskussion immer wieder Forderungen nach einer erhöhten Prüfungsqualität in der Abschlussprüfung laut. Während der Gesetzgeber den Weg beschreitet, die Arbeit von Abschlussprüfern und Aufsichtsräten stärker zu reglementieren und zu überwachen[4], wird nur selten erörtert, den Prüfungsansatz in Bezug auf Fraud zu verändern. Ohne eine Diskussion möglicherweise notwendiger Modifikationen des Prüfungsansatzes bliebe das eigentliche Problem unberücksichtigt.[5]

In Deutschland verlangen das HGB und der für die Aufdeckung von Unregelmäßigkeiten im Rahmen der Abschlussprüfung einschlägige *IDW PS 210* vom Abschlussprüfer, die Abschlussprüfung so anzulegen, dass Unrichtigkeiten oder Verstöße gegen gesetzliche Vorschriften und sie ergänzende Bestimmungen des Gesellschaftsvertrages oder der Satzung, die sich auf die Darstellung des sich nach § 264 Abs. 2 HGB ergebenden Bildes des Vermögens-, Finanz- und Ertragslage des Unternehmens wesentlich auswirken (materieller Fraud), bei gewissenhafter Berufsausübung erkannt werden.[6] Dabei verlangen die Berufsgrundsätze vom Abschlussprüfer die Abgabe eines hinreichend sicheren Urteils, wobei seitens des Berufsstandes eine Urteilssicherheit von durchschnittlich mindestens 95 % gefordert wird.[7]

Der Prüfer hat die Abschlussprüfung mit einer kritischen Grundhaltung gegenüber dem geprüften Unternehmen, seinen gesetzlichen Vertretern und Mitarbeitern zu planen und durchzuführen. Hierfür muss er im Rahmen eines risikoorientierten Prüfungsansatzes Systemprüfungen und aussagebezogene Prüfungshandlungen in Form analytischer Prüfungshandlungen und Einzelfallprüfungen von Geschäftsvorfällen und Beständen durchführen. Ergibt eine demgemäß durchgeführte Abschlussprüfung keine wesentlichen Mängel, kann der Abschlussprüfer die Buchführung, den Abschluss sowie ggf. den Lagebericht als ordnungsgemäß akzeptieren.[8]

1 Vgl. *KPMG,* Wirtschaftskriminalität in Deutschland 2003/2004, *www.kpmg.de/forensic* (Stand: 8. 8. 2004).
2 Engl. = Betrug, List, Täuschung, Unterschlagung.
3 Beispielsweise Parmalat, Ahold, Worldcom, Enron, Comroad, Flowtex, Balsam/Procedo.
4 Verpflichtung zur externen Qualitätskontrolle gem. § 57a WPO; wohl auch Regierungsentwurf des Gesetzes zur Kontrolle von Unternehmensabschlüssen (Bilanzkontrollgesetz – BilKoG – vom 21. 4. 2004, Art. 1, § 342b HGB): durch die anlassbezogene und stichprobenartige Überprüfung der Rechtmäßigkeit der Jahres- bzw. Konzernabschlüsse durch eine private Prüfstelle („Enforcement-Verfahren") wird mittelbar auch die Arbeit der Aufsichtsräte und Abschlussprüfer überwacht.
5 Vgl. *Schruff,* Zur Aufdeckung von Top-Management-Fraud durch den Wirtschaftsprüfer im Rahmen der Jahresabschlussprüfung, WPg 2003, S. 901 ff.
6 Vgl. § 317 Abs. 1 Satz 3 HGB; *IDW Prüfungsstandard: Zur Aufdeckung von Unregelmäßigkeiten im Rahmen der Abschlussprüfung (IDW PS 210),* WPg 2003, S. 655 ff., Rn. 12.
7 Vgl. *IDW* (Hrsg.), WP-Handbuch 2002, 12. Aufl., Düsseldorf 2000, Bd. I, Rn. R 28.
8 Vgl. *IDW PS 210,* WPg 2003, S. 655 ff., Rn. 17.

Bei den aussageorientierten Prüfungshandlungen sieht sich der Prüfer mit dem Problem konfrontiert, dass bei fraudulenten Unternehmen zur Verschleierung der Taten häufig Dokumente (z. B. Rechnungen, Verträge und Kontoauszüge) gefälscht werden. Oftmals ist die Feststellung der Echtheit von Prüfungsbelegen letztendlich nur mit kriminaltechnischen Mitteln möglich. Da die Abschlussprüfung nicht als Misstrauensauftrag im Sinne kriminalistischer Ermittlungen verstanden wird, darf der Abschlussprüfer gemäß den Berufsgrundsätzen die Authentizität der ihm vorgelegten Unterlagen und Dokumente unterstellen, solange keine Anzeichen vorliegen, die Zweifel an der Echtheit der Dokumente aufkommen lassen könnten.[9]

Da eine pflichtgemäße Abschlussprüfung nicht darauf ausgelegt ist, explizit nach gefälschten Prüfungsnachweisen zu suchen, werden nach berufsständischem Standard im Rahmen eines risikoorientierten Prüfungsansatzes Systemprüfungen durchgeführt.[10] Vereinfacht gesagt, wird bei Vorliegen eines funktionierenden Internen Kontrollsystems davon ausgegangen, dass die Geschäftsvorfälle ordnungsgemäß erfasst und verarbeitet und Manipulationen auf den nachgeordneten Ebenen aufgedeckt werden. Dieser Prüfungsansatz greift dann zu kurz, wenn in dem geprüften Unternehmen ein Mitglied des Managements (oder ein anderer Mitarbeiter mit entsprechenden Einflussmöglichkeiten) die existierenden Kontrollen und Prozesse umgeht oder ihm unterstellte Mitarbeiter instrumentalisiert (sog. „Management Override").

Fraud ist ein globales Phänomen mit wachsender Bedeutung[11], was eine dynamische Entwicklung der entsprechenden Prüfungsstandards zur Folge hat. Der Prüfer muss nach allen aktuellen Standards schon in der Prüfungsplanung (und später prüfungsbegleitend) das Fraud-Risiko explizit in die Betrachtung einbeziehen. Sobald auf die stetige Funktionsfähigkeit des Internen Kontrollsystems nicht mehr vertraut werden kann oder andere Anzeichen für Fraud vorliegen, muss der Abschlussprüfer seinen Prüfungsansatz entsprechend anpassen – z. B. durch eine Ausweitung der Prüfungshandlungen – unter Umständen bis hin zur Vollprüfung riskanter Prüfungsfelder oder zur Einbeziehung forensischer Spezialisten.

Es ist daher eine gesonderte Fraud-Risiko-Analyse durchzuführen, die sich weder ausschließlich auf die Prüfung einzelner Belege noch allein auf die Systemprüfung des Internen Kontrollsystems stützt.

Die folgenden Ausführungen werden einen Ansatz vorstellen, der es ermöglicht, Fraud-Muster zu erkennen und dadurch Schlussfolgerungen für das Risiko von materiellem Fraud zu ziehen – und zwar unabhängig von Größe, Branche, Rechtsform oder z. B. Nationalität des geprüften Unternehmens.

## 2. Fraud-Risiko-Modell

Motivationen für dolose Handlungen können wirtschaftlicher Druck auf das Unternehmen oder ein finanzieller Engpass des Täters sein.[12] Dabei muss der Täter sein Verhalten vor sich selbst rechtfertigen und rationalisieren können. Typische Rationalisierungen sind „Es ist zum Wohl der Firma.", „Es ist nur vorübergehend." oder „Es steht mir zu.".

Kommt die Möglichkeit zum Begehen der Tat hinzu, sind die drei Faktoren gegeben, die ein so genanntes „Fraud Triangle" begründen.[13] Dieses Modell liegt dem hier beschriebenen Ansatz zur Fraud-Risiko-Prognose zugrunde. Bei Fraud sind demnach im Allgemeinen folgende drei Bedingungen gleichzeitig erfüllt:

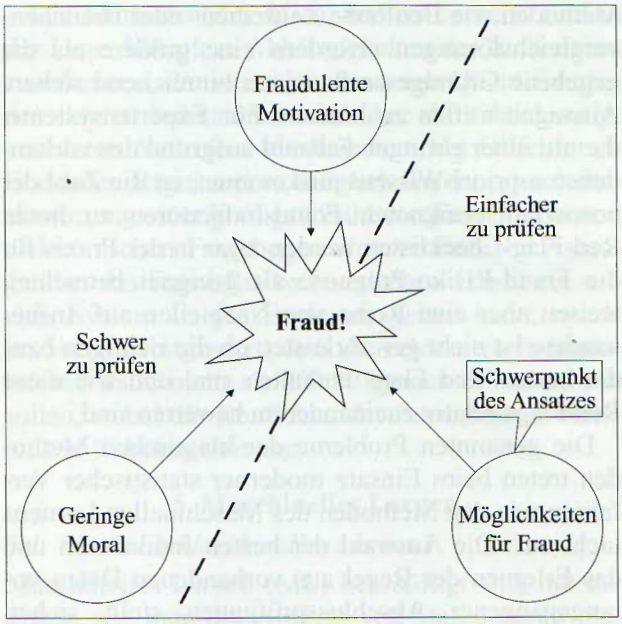
- die Möglichkeit, Fraud zu begehen. Das betrifft sowohl die Abwesenheit von Kontrollen als auch ineffektive Kontrollen oder die Möglichkeit des Managements, die Kontrollen zu überwinden (Management Override);
- das Motiv, also einen Nutzen aus der Tat oder einen Druck, die Tat zu begehen;
- eine persönliche, quasi charakterliche oder kulturelle Komponente, die es dem Täter gestattet, die Tat zu begehen bzw. zu rechtfertigen. Es gibt Individuen, die charakterlich eher bereit sind, wissentlich und gewollt betrügerische Handlungen zu begehen. Aber auch ehrliche Menschen werden unter starkem Druck in die Lage versetzt, betrügerisches Handeln zu akzeptieren.

Gemäß dem Fraud-Triangle-Ansatz ist das Fraud-Risiko hoch, wenn eine oder mehrere Personen, die durch einen privaten oder betrieblichen Umstand motiviert sind, eine Gelegenheit zur Tat finden und dabei nicht durch eine entsprechende Wirtschaftsmoral davon abgehalten werden.

Mit anderen Worten: „Ist der Boden fruchtbar und wird nicht gejätet, wächst früher oder später Unkraut darauf."

Dem Abschlussprüfer ist es in der Regel nicht möglich, die Motivationslage der potenziellen Täter für materiellen Fraud in einem Unternehmen vollständig zu erfassen.[14] Hierfür wären auch Prüfungs-

9 Vgl. *IDW PS 210*, WPg 2003, S. 655 ff., Rn. 15.
10 Vgl. *IDW PS 210*, WPg 2003, S. 655 ff., Rn. 14.
11 Vgl. *PWC*, Global economic crime survey 2003, S. 6, *www.pwc.com/crimesurvey* (Stand: 8. 8. 2004).
12 Weitere, weniger wahrscheinliche Motivationen sind Rache oder die Herausforderung, „das System zu überwinden".
13 Vgl. *IFAC*, International Standard on Auditing (ISA) 240 (rev.): The Auditor's Responsibility to consider Fraud in an Audit of Financial Statements (ISA 240 (rev.)), Rn. 51; *Krambia-Karpadis*, Enhancing the Auditor's Fraud Detection Ability – An interdisciplinary Approach, Frankfurt u.a. 2001; *Leobbecke/Eining/Willingham*, Auditors' experience with material irregularities – Frequency, nature, and detectibility, Auditing 8/1989, S. 1–28.; *Albrecht/Wernz/Williams*, Fraud – Bringing Light to the Dark Side of Business, Burr Ridge, IL, 1995.
14 Vgl. *IFAC*, ISA 240 (rev.), Rn. 18.




*Übersicht 1: Fraud-Triangle-Ansatz*

handlungen in der Privatsphäre der handelnden Personen notwendig, was aus rechtlichen und aus ethischen Gründen nicht in Betracht kommt. Die Erfassung der Wirtschaftsmoral bedarf darüber hinaus Qualifikationen und Ansätzen, die in einer regulären Abschlussprüfung nicht verfügbar sind. Dennoch kann es möglich sein, dass der Prüfer direkt oder indirekt Kenntnis über diese beiden Faktoren erhält.[15]

Es stellte sich die grundlegende Frage, ob es überhaupt möglich ist, mit den im Rahmen einer normalen Jahresabschlussprüfung verfügbaren Informationen, insbesondere über die Möglichkeiten für das Begehen der Tat, das Risiko von materiellem Fraud mit einer hinreichenden Sicherheit zu bestimmen.

## 3. Anforderungen an ein Verfahren zur Beurteilung des Fraud-Risikos

Gemäß dem Modell des Fraud Triangle ist das Fraud-Risiko gering, wenn einer der oben genannten drei Faktoren nicht gegeben ist. Wenn z. B. die Wahrscheinlichkeit der Existenz von Möglichkeiten für Fraud gering ist, dann ist auch das Fraud-Risiko insgesamt gering.

Dabei ist ein Verfahren zu identifizieren, das auf der Grundlage der im Rahmen einer Jahresabschlussprüfung verfügbaren Informationen in der Lage ist, das Fraud Risiko zu bestimmen. Da Fraud ein weltweites und branchenübergreifendes Phänomen ist und die Problemstellung allgemeingültig gelöst werden soll, stellt sich in einem globalisierten Umfeld die Frage, ob die Bestimmung des Fraud-Risikos auch landesunabhängig behandelt werden kann, d. h. ob Länderspezifika wie z. B. abweichende Gesetzeslage, wirtschaftliche Entwicklung und Wirtschaftskultur Auswirkungen auf die Beurteilung des Fraud-Risikos haben und ggf. spezielle Ansätze erfordern.

Die gleiche Frage stellt sich hinsichtlich unterschiedlicher Branchen, Größen oder Rechtsformen der geprüften Unternehmen, da auch Prüfungsstandards[16] davon ausgehen, dass sich hieraus ganz unterschiedliche Möglichkeiten und Anreize für fraudulente Handlungen ergeben.

Nicht zuletzt stellt sich die Frage der Universalität eines solchen Ansatzes auch hinsichtlich der Vielfalt möglicher Erscheinungsformen von Fraud. Es sind vielfältige Fraud-Muster bekannt, die in einer nicht bestimmbaren Menge von situationsbedingten Sonderfällen, kreativen Varianten und möglichen Verschleierungstaktiken abgewandelt werden können. Es muss also geklärt werden, ob ein Ansatz für alle Varianten geeignet sein kann oder ob Speziallösungen notwendig sind.

## 4. Die Ableitung des Verfahrens

Bei der Auswahl von geeigneten Verfahren ist grundsätzlich zu berücksichtigen, dass keine einfachen, universellen Regeln bekannt sind, Fraud-Fälle von Non-Fraud-Fällen zu unterscheiden. Es ist daher notwendig, derartige Regeln zu ermitteln.

Das Fraud-Risiko kann aufgrund der Natur der Sache, dass Fraud durch den Täter verschleiert wird, nicht direkt ermittelt werden. Es bietet sich an, aus den Daten bekannter Fraud- und Non-Fraud-Fälle statistisch ggf. existierende Muster von Eigenschaften fraudulenter Unternehmen zu ermitteln. Entsprechend dem Modell des Fraud Triangle kann angenommen werden, dass bei Vorliegen eines ähnlichen Musters bei einem unbekannten Fall eine Fraud-Risiko-Konstellation gegeben ist.

Demnach sind Indikatoren zu identifizieren, die Rückschlüsse auf das Fraud-Risiko zulassen. Bei diesen auch als Red Flags bezeichneten Indikatoren handelt es sich um Ereignisse oder Umstände, die auf Anreize oder einen Druck hinweisen, Fraud zu begehen oder Möglichkeiten für Fraud zu schaffen. Obwohl diese Fraud-Risiko-Indikatoren nicht zwangsläufig die Existenz von Fraud nachweisen, treten sie doch häufig im Zusammenhang mit Fraud auf.[17] Dabei gibt es nicht nur einige wenige Indikatoren, sondern entsprechend der Vielfalt möglicher Ausprägungen von Fraud eine Vielzahl von möglicherweise geeigneten Red Flags. Jedoch ermöglicht keiner dieser Fraud-Risiko-Indikatoren für sich alleine eine Entscheidung für die Wahl des Prüfungsansatzes.[18] Zur Systematisierung und besseren Handhabbarkeit werden die (potenziellen) Indikatoren zu Red-Flag-Checklisten zusammengefasst. Diese Checklisten – wie auch sonstige Veröffentlichungen zu möglichen Red Flags – repräsentieren i. d. R.

15 Vgl. *IFAC*, ISA 240 (rev.), Rn. 51.
16 Vgl. *IFAC*, ISA 240 (rev.), Rn. 36 f., 52 f.
17 Vgl. *IFAC*, ISA 240 (rev.), Rn. 49.
18 Vgl. *Schruff*, WPg 2003, S. 908.

die persönlichen Erfahrungen von Prüfern bzw. Autoren. Untersuchungen über die Prognosesicherheit dieser Indikatoren liegen jedoch nur vereinzelt vor. Es ist nicht bekannt, welche Eigenschaften eines Unternehmens in welcher Kombination das Risiko der verschiedenen Erscheinungsformen von Fraud am besten abschätzen lassen.

Aufgrund dieses Mangels empfahl es sich, für die statistische Identifikation von Fraud-Mustern möglichst viele der als geeignet betrachteten Indikatoren in die Untersuchung einzubeziehen. Um zudem die Eigenschaften der untersuchten Fraud- und Non-Fraud-Fälle standardisiert erheben zu können, war ein möglichst umfassender Fragebogen zu entwickeln.

Als Quellen für potenziell geeignete Fraud-Risiko-Indikatoren dienten in erster Linie die Prüfungschecklisten der großen Wirtschaftsprüfungsgesellschaften[19], die im Entwicklungszeitraum aktuellen Prüfungsstandards[20] und eine Vielzahl von Veröffentlichungen[21]. Bei der Entwicklung des Fragebogens galt es zu beachten, dass die Indikatoren im Rahmen einer Jahresabschlussprüfung erhebbar sein müssen. Um die internationale Einsetzbarkeit des Ansatzes zu gewährleisten, musste der Fragebogen internationalisiert werden, d.h. länderspezifische Aspekte und Fragestellungen waren zu vermeiden bzw. ggf. zu neutralisieren.

Als Ergebnis entstand ein Fragebogen, der über 300 qualitative, prozessorientierte Indikatoren (Umfeldinformationen, (interne) prozessübergreifende Kontrollen und prozessinterne Kontrollen) erfasste, sowie ein Auswertungsbogen mit über 300 quantitativen Werten aus der Jahresabschlussanalyse (Jahresabschlussangaben, daraus abgeleitet finanz- und erfolgswirtschaftliche Kennzahlen, Zeit- sowie Branchenvergleiche).

Für die Analyse wurden in einem Zeitraum von drei Jahren insgesamt 115 Fraud- und Non-Fraud-Fälle aus Deutschland, Großbritannien, Frankreich, der Schweiz und den Niederlanden identifiziert und anonymisiert verfügbar gemacht. Zur Reduktion kontraproduktiver Effekte des Einsatzes einer Checkliste – wie Missinterpretationen oder die unbewusste Wiedergabe eines wünschenswerten Soll-Zustandes anstatt des als Ausnahme empfundenen Ist-Zustandes – wurden zwei Mittel kombiniert:

- die möglichen Antwortalternativen wurden (sinnvoll) standardisiert (das war schon aus technischen Gründen notwendig), und
- der Fragebogen wurde in einer Dialog-Situation zwischen einem Experten und einem leitenden Mitglied des Prüfungsteams beantwortet.

Die verhältnismäßig geringe Anzahl von Fällen (115) bei gleichzeitig hoher Zahl der Indikatoren (über 600) hat zur Folge, dass klassische statistische Verfahren nicht zur Anwendung kommen können. Methoden wie Benford-, Zeitreihen- oder Branchenvergleichslösungen erfordern eine größere als die gegebene Grundgesamtheit, um hinreichend sichere Aussagen treffen zu können. Für Expertensysteme, die mit einer geringen Fallzahl aufgrund des vorhandenen a-priori-Wissens auskommen, ist die Zahl der potenziell geeigneten Fraud-Indikatoren zu hoch. Red-Flag-Checklisten werden zwar in der Praxis für die Fraud-Risiko-Prognose als geeignet betrachtet, weisen aber eine Reihe von Nachteilen auf. Insbesondere ist nicht gewährleistet, ob die richtigen bzw. die besten Red Flags enthalten sind und wie diese Red Flags relativ zueinander zu bewerten sind.

Die genannten Probleme der klassischen Methoden treten beim Einsatz moderner statistischer Verfahren wie den Methoden des Maschinellen Lernens nicht auf. Die Auswahl der besten Indikatoren und das Erlernen der Regel aus vorhandenen Daten vorangegangener Abschlussprüfungen stellt sicher, dass die optimalen Indikatoren für die Prognose gewählt werden und keinerlei statistisch nicht validen, vorgefassten „persönlichen Urteile" die Objektivität der Beurteilung trüben. Eine ähnliche Studie[22] zeigt bei einem Vergleich auf der Basis der 500 umsatzstärksten US-Unternehmen (US-Fortune-500), dass bereits Künstliche Neuronale Netze (KNN, eine Teilklasse des Maschinellen Lernens und die Vorgänger der hier eingesetzten Methoden) als Mittel der Fraud-Risiko-Beurteilung im Vergleich mit Expertensystemen und klassischen statistischen Verfahren am erfolgreichsten waren.

19 Stand: 2002.

20 *AICPA,* SAS 82: Consideration of Fraud in a Financial Statement Audit (AICPA SAS 82); *IFAC,* ISA 240 a.F.: The Consideration of Fraud and Error in an Audit of Financial Statements (ISA 240 a.F.); *IFAC,* ISA 250: Consideration of Laws and Regulations in an Audit of Financial Statements (IFAC ISA 250); *Stellungnahme HFA 7/1997: Zur Aufdeckung von Unregelmäßigkeiten im Rahmen der Abschlussprüfung,* WPg 1998, S. 29 ff.

21 Vgl. z.B. *Heiman-Hoffman u.a.,* The Warning Signs of Fraudulent Financial Reporting – What do auditors believe are the best ways to spot fraud?, JoA 10/1996, S. 75–77; *Huntington/Davies/Lohse,* Wirtschaftskriminalität im Unternehmen – Betrug erkennen und bekämpfen, Frankfurt/New York 1999; *Langenbucher/Blaum,* Die Aufdeckung von Fehlern, dolosen Handlungen und sonstigen Gesetzesverstößen im Rahmen der Abschlussprüfung, DB 1997, S. 437–443; *Strand/Strawser,* Help for Internal Auditors – SAS No. 82 identifies risk factors for misappropriation of assets, Internal Auditing, January/February 2000, S. 9–14; *Odenthal,* Verfahren und Instrumente zur Aufdeckung wirtschaftskrimineller Handlungen, Österreichische Zeitschrift für Recht und Rechnungswesen 1999, S. 1019–1023; *Müller,* Wirtschaftskriminalität – Analyse eines interdisziplinären Phänomens, Der Schweizer Treuhänder 1995, S. 839–846; *Mertens,* Früherkennung von Bilanzmanipulationen bonitätsschwacher mittelständischer Unternehmen, Düsseldorf 1998; *Lohse,* Wirtschaftskriminalität – Prävention und Prüfung, Die Bank 1996, S. 196–200; *Krambia-Kapardis,* a.a.O. (Fn. 3); *Gill,* The effects of a red flags questionaire and selected elements of cognitive style on the audit planning task of assessing the probability of material management fraud, Diss., The University of Mississippi, 1992; *Baucus/Near,* Can Illegal Corporate Behavior Be Predicted? – An Event History Analysis, Academy of Management Journal 1/1991, S. 9–36; *Albrecht/Romney,* Red-flagging management fraud – a validation, Advances in Accounting 1986, S. 323–333.

22 Vgl. *Green/Choi,* Assessing the Risk of Management Fraud Through Neural Network Technology, Auditing 1/1997, S. 14–28.

Die Methoden des Maschinellen Lernens sind in der Lage, komplexe Abhängigkeiten der Indikatoren zu erkennen, die weit über das reine Quantifizieren der aktiven Indikatoren oder über die Gleichzeitigkeit deren Auftretens hinausgehen. Mit Hilfe der zu findenden Regel soll es möglich werden, durch Beantwortung einer geringen Zahl von Fragen in einem elektronischen Fragebogen eine verlässliche Aussage über das Fraud-Risiko zu erhalten. Damit ein solches System als praxisrelevant gelten kann, wurde im Vorfeld festgesetzt, dass die relevanten Indikatoren mit nicht mehr als 40–50 Fragen erfassbar sein sollen, wobei die Konfidenz der Bewertung mindestens 95 % betragen muss.

## 5. Maschinelles Lernen

### *5.1 Was kann Maschinelles Lernen?*

Maschinelles Lernen (ML) beschäftigt sich mit der Frage, wie man die Analyse von Daten durch Maschinen, d.h. meist Computer, automatisieren kann. Das Besondere dabei ist, dass nicht ein menschlicher Experte diesen Automatismus von Hand modellieren muss, sondern dass die Maschine diesen selbstständig aus vorliegenden Daten *erlernt.*[23]

Maschinelles Lernen ist deshalb so mächtig und erfolgreich bei der Lösung komplexer Probleme, weil es auf ein äußerst genau untersuchtes mathematisches Fundament bauen kann. Die Grundlagen hierzu bilden das Lernen aus Beispielen und die statistische Lerntheorie. Die zentrale Erkenntnis der statistischen Lerntheorie ist dabei verblüffend intuitiv: *Die einfachste Lösung eines Problems, die eine akzeptable Erklärung liefert, ist die beste.* Allerdings ist die hinter dieser Einsicht liegende Mathematik und deren Umsetzung in praktische Anwendungen äußerst komplex.

Ziel von Lernmethoden[24] ist es, durch Beobachtung der Vergangenheit Aussagen über ungesehene Daten, d.h. die Zukunft, abzuleiten. Der Anwendungsrahmen dieser Beobachtungen und Prognosen ist dabei nahezu unbeschränkt: von der Lösung einfacher technischer Probleme (virtuelle Chemie, Prozesssteuerung) über die Analyse des menschlichen Genoms, von Hirnsignalen, von Dokumenten bis hin zur Vorhersage komplexer, dynamischer Systeme. In der Tat sind viele aktuelle Fragestellungen, wie z.B. das Finden von Genen in der Bioinformatik, ohne Maschinelles Lernen nicht mit der notwendigen Präzision und Effizienz lösbar.[25]

Der wesentliche Unterschied zu einer Expertenlösung eines Problems ist, dass auf Expertenwissen des speziellen Anwendungsbereichs verzichtet werden kann: „Die Daten sprechen zunächst für sich selbst". Andererseits gilt auch, dass es möglich ist, Wissen aus der jeweiligen Anwendungsdomäne, so genanntes a-priori-Wissen, einzubringen und bei der Modellierung zu nutzen. Einen entsprechend sorgfältigen Einsatz vorausgesetzt, lassen sich ML-Systeme auch in sicherheitskritischen Anwendungen, wie z.B. der Medizintechnik oder bei sensiblen Problemen, wie im vorliegenden Fall, nutzen. Es folgt eine kurze Übersicht typischer durch ML-Methoden lösbarer Probleme:

- *Detektion ungewöhnlicher Vorfälle,* also die Identifikation von Vorgängen, die nicht in ein festes Schema passen, z.B. zur Erkennung von Kreditkartenbetrug, Versicherungsbetrug oder Unterschlagung. Basierend auf vorhandenen Daten von erlaubtem und betrügerischem Verhalten findet die eingesetzte Methode automatisch den Zusammenhang, der es erlaubt, ungewöhnliches oder riskantes Verhalten zu entdecken bzw. vorherzusagen;
- *Definition und Berechnung von Scores,* z.B. zur Bewertung der Kreditwürdigkeit, der Erfüllung von diversen Richtlinien (z.B. Basel II), des Kredit-Ausfallrisikos oder einer Erkrankungswahrscheinlichkeit. Ziel ist z.B. bei der Kreditvergabe durch die Erhebung von wenigen relevanten Informationen, eine möglichst genaue Aussage über das mit dem Kredit verbundene Risiko zu treffen;
- *Vorhersage und Analyse* von (nicht-)linearen Zeitreihen, z.B. im Portfoliomanagement oder zur Marktanalyse; Methoden des Maschinellen Lernens gehören heute zu den am weitesten fortgeschrittenen, insbesondere im Umgang mit nichtlinearen und nichtstationären Zeitreihen und bei stark zeitvarianten Dynamiken;
- *Clustering, Klassifikation und Visualisierung:* Auch hier bietet das Maschinelle Lernen eine Vielzahl neuer leistungsfähiger Verfahren, wie z.B. die später noch näher beschriebenen Support-Vektor-Maschinen; Letztere eignen sich besonders für die robuste Analyse komplexer Daten.

### *5.2 Wie funktioniert Maschinelles Lernen?*

Maschinelles Lernen vertritt einen spezifischen Zugang zum Grundproblem der Erkennung von Regularitäten (statistischer Zusammenhänge) in einer komplexen Umwelt, den man folgendermaßen vereinfacht modellieren kann: auf Basis von empirischen Beobachtungen zweier Größen x und y versucht man den Zusammenhang dieser Größen zu schätzen, um damit bei der Beobachtung einer neuen Eingabe x' eine gute Prognose y' abgeben zu können.

Mustererkennung, Klassifikation, Regression oder Zeitreihenprognose sind einige abstrakte Beispiele,

23 Maschinelles Lernen (ML) kann sowohl als Teil der künstlichen Intelligenz (KI) als auch der angewandten Statistik verstanden werden.

24 Vgl. hier insbesondere *Müller u.a.,* An Introduction to Kernel-Based Learning Techniques, IEEE Transactions on Neural Networks 2001, Band 12:2, S. 181–201; *Schölkopf/Smola,* Learning with Kernels, MIT Press, Cambridge MA 2002.

25 Vgl. *Mjolsness/DeCoste,* Machine Learning for Science – State of the Art and Future Prospects, Science 2001, S. 2051–2055.

die sich in das Paradigma einfügen. Ganz allgemein lässt sich sagen, dass sich dieser statistische Zugang zum Schätzen von Regularitäten dadurch auszeichnet, dass er auf Probleme angewandt werden kann, bei denen die explizite Modellierung des Zusammenhangs schwierig oder unmöglich ist, dafür aber empirische Beobachtungen vorliegen.

Ein Paradebeispiel hierfür ist die Erkennung handgeschriebener Ziffern. Wie eine Ziffer im Detail geschrieben wird, hängt von einer Reihe schwer modellierbarer Faktoren ab. Daher bietet es sich an, die Klassifikation zwischen Ziffern aus Beispielen zu lernen. Es ist in diesem Fall zwar theoretisch möglich, Regeln für Form und Schreibweise einer Ziffer anzugeben, aber es wären einige zehntausend Regeln erforderlich, was praktisch natürlich völlig indiskutabel ist.

In dem vorliegenden Anwendungsfall der Beurteilung des Fraud-Risikos ist aus bereits genannten Gründen die manuelle Formulierung eines Regelwerkes in Form von Red-Flag-Listen schwierig. Betrachtet man dazu die zuvor genannten Anforderungen an das zu entwickelnde Verfahren, erscheint es unwahrscheinlich, wenn nicht unmöglich, dass sich diese auf der Basis von Red-Flag-Listen erfüllen ließen.

### *5.3 Unterschiede zur klassischen Statistik/Vorteile moderner ML-Verfahren*

Üblicherweise ist ein Ziel der Mustererkennung, auf möglichst wenigen starken Indikatoren zu arbeiten, ähnlich der Herangehensweise eines menschlichen Experten. Moderne ML-Verfahren, wie z. B. die hier eingesetzten Support-Vektor-Maschinen, setzen sich von dieser Philosophie der starken Indikatoren ab. Durch die Verwendung von so genannten Kernfunktionen zur effizienten Repräsentation der Daten in einem hochdimensionalen Raum wird implizit eine sehr große, reichhaltige Anzahl von Indikatoren erzeugt, die nachfolgenden Verarbeitungsschritten zur Verfügung stehen. Somit spiegelt sich in Kern-Algorithmen ein Paradigmenwechsel wider.[26] Anstatt mit möglichst einfachen niedrig-dimensionalen Modellen realistische Datensätze zu modellieren, besteht nun aufgrund des Einsatzes von Kernfunktionen die Möglichkeit, komplexe nichtlineare Probleme effizient mit hoher Genauigkeit zu lösen. Dabei bieten die Methoden des Maschinellen Lernens drei signifikante Vorteile gegenüber klassischen statistischen Verfahren (z. B. Entscheidungsbäumen, Dichteschätzung oder Faktoranalyse), aber auch gegenüber Verfahren, die im Allgemeinen als „neuronale Netze" bezeichnet werden:

- *Maschinelles Lernen ist präziser:* Eine höhere Genauigkeit in der Vorhersage führt zu qualitativ besseren Lösungen.
- *Maschinelles Lernen wird eingesetzt, wenn herkömmliche Verfahren versagen:* Statistisch valide Aussagen lassen sich bei modernen ML-Methoden bereits mit kleinen Stichproben erreichen. Die klassische Statistik benötigt in der Regel eine größere Anzahl von Datensätzen, um ähnlich zuverlässige Aussagen treffen zu können. Wenn also systematisch oder beispielsweise aus Kostengründen nur kleine Datenmengen zur Verfügung stehen, sind solche Probleme mit klassischen Ansätzen nicht befriedigend lösbar.
- *Anwendung auf komplexe, nicht-numerische Daten:* Die in diesem Projekt verwandten Methoden sind direkt auf komplexe, nicht-numerische Daten wie Texte, Fragebögen oder Bilder anwendbar. Für klassische Verfahren ist es in der Regel schwierig, mit diskreten Größen oder strukturierten Daten umzugehen. Solche Daten tauchen z. B. in der Chemie oder Bioinformatik oder – wie hier – bei der Analyse von Fragebögen auf.

Insbesondere gilt im vorliegenden Fall der Analyse von Daten aus der Abschlussprüfung zur Betrugserkennung, dass

(i) eine hohe Genauigkeit erreicht werden muss,
(ii) nur wenige Daten vorhanden sind (das Erheben weiterer Daten ist zeitaufwendig und käme deshalb erst zu spät zum Ergebnis) und
(iii) diese in Form von diskreten (Ja-Nein-Fragen) und kontinuierlichen Größen (z. B. Umsatz) vorliegen.

Da die dem Fraud-Problem zugrunde liegende Statistik unklar ist, mussten verschiedene moderne Verfahren evaluiert werden. Es wurden so genannte RBF-Netze, Entscheidungsbäume und kernbasierte Verfahren – z. B. Lineare-Programmier-Maschinen (LPM), Support-Vektor-Maschinen (SVM) und Kern-Fisher-Diskriminanten – getestet. Es stellte sich heraus, dass Ergebnisse der geforderten Güte nur mittels LPM und SVM zu erzielen waren.

### *5.4 Support-Vektor-Maschinen*

Support-Vektor-Maschinen zur Mustererkennung können aus einem Trainingsdatensatz lernen, zwischen zwei Klassen zu unterscheiden, z. B. Fraud und Nicht-Fraud. Dazu finden sie in einem Merkmalsraum beliebiger Dimensionalität eine Hyperebene $w$ mit minimaler Komplexität (Gleichung (1) in Übersicht 2). Nur durch die Benutzung von Kernen kann man in diesem Merkmalsraum z. B. Skalarprodukte berechnen. Letzteres wäre bei hoch-dimensionalen Merkmalsräumen sehr aufwendig bzw. sogar unmöglich. Entscheidend ist, dass sich alle Berechnungen, die für das Finden der optimalen Entscheidungsfunktion notwendig sind, durch solche Skalarprodukte ausdrücken

26 Siehe auch *Boser/Guyon/Vapnik,* A training algorithm for optimal margin classifiers, in: Haussler (Hrsg.), Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory 1992, S. 144–152; *Vapnik,* The nature of statistical learning theory, New York 1995; *Müller u. a.,* a.a.O. (Fn. 24).

lassen. Dies führt zu einer Support-Vektor-Entscheidungsfunktion f(x) (Gleichung (3) in Übersicht 2). Nun ersetzt man alle Skalarprodukte durch die Kernfunktion (der so genannte Kerntrick – Gleichung (4) in Übersicht 2). Um die optimale Funktion zu finden, muss ein quadratisches Optimierungsproblem gelöst werden (Gleichung (5) in Übersicht 2).

(1) Hyperebene im Merkmalsraum $f(x) = (w \cdot \Phi(x)) + b$
   a. x ist die Eingabe (z. B. die Antworten zum Fragebogen) und
   b. $\Phi$ die Abbildung in einem Merkmalsraum.

(2) Skalarprodukt im Merkmalsraum: $(w \cdot \Phi(x))$

(3) Support-Vektor-Entscheidungsfunktion: $f(x) = \sum_{i=1}^{l} y_i \alpha_i k(x_i, x) + b$

   a. Man nutzt dabei, dass $w = \sum_i y_i \alpha_i \Phi(x_i)$.

   b. Entscheidungen werden danach getroffen, ob $f(x)$ größer oder kleiner als Null ist.

(4) Kerntrick: $k(x_i, x_j) = (\Phi(x_i) \cdot \Phi(x_j))$.

(5) Optimierungsproblem von SVM $\min \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j k(x_i, x_j) + C \sum_{i=1}^{l} \xi_i$

   a. unter den Nebenbedingungen $y_i f(x_i) \geq 1 - \xi_i$

   b. und $\alpha_i, \xi_i \geq 0$, wobei C eine Regularisierungskonstante ist.

*Übersicht 2: Support-Vektor-Entscheidungsfunktion f(x)*

## 6. FRAT44[27]: Maschinelles Lernen zur Betrugs-Risiko-Bestimmung – erste Ergebnisse

### *6.1 Klassifizierungsleistung*

Die statistische Evaluation des endgültigen Systems zeigt, dass es mit diesem Ansatz möglich ist, mit hoher Genauigkeit das Vorliegen eines Fraud-Musters zu erkennen und damit zu helfen – gemäß dem Fraud Triangle –, das Fraud-Risiko im Prüfungsprozess zu beurteilen. Mit Hilfe der gewählten Lernmaschine konnte ein Verfahren entwickelt werden, das mit einer – gemessen an den Trainingsdaten – durchschnittlich mehr als 96,5 %igen Sicherheit die analysierten Unternehmen entsprechend ihrer Übereinstimmung mit einem bekannten Fraud-Muster klassifiziert und somit auf das Fraud-Risiko schließen lässt. Zur besseren Anwendbarkeit klassifiziert das Modell nicht nur in High-Risk-Fälle und Low-Risk-Fälle, sondern gibt zudem noch die Sicherheit an, mit der die Klassifizierung zutreffend ist. In über ¾ der Fälle liegt diese mit 99 % extrem hoch.

### *6.2 Fragebogen und wichtigste Indikatoren*

Der Fragebogen umfasste gegen Ende der Entwicklungsphase, also während der Datenerhebungen, 180 Fragen, die insgesamt 331 Fraud-Risiko-Indikatoren repräsentierten.[28] Aus der Analyse der Rechnungswesendaten ergaben sich nochmals 336 Indikatoren, so dass insgesamt 667 Indikatoren pro Fall zu bedienen waren. Da dieser Umfang für den praktischen Einsatz zu groß ist, war eine Reduktion des Fragebogens auf die wichtigsten Fragen bzw. Rechnungswesendaten notwendig – nach Möglichkeit jedoch ohne eine Reduzierung der unter Verwendung aller Indikatoren erreichten Klassifizierungsleistung.

Die Auswahl der relevanten Indikatoren ist ein hoch anspruchsvolles Problem und gegenwärtig eine der Hauptforschungsrichtungen des Maschinellen Lernens. Zur Selektion wurde ein dreifach verschachtelter Leave-One-Out-Kreuz-Validierungs-Prozess angewandt, der auf einem Rechnercluster von etwa 1000 Computern eine Rechenzeit von 13 Tagen erforderte. Das war notwendig, um für jede Indikatoren-Kombination Fehler-Balken zu erhalten und diese auf statistische Signifikanz und Stabilität zu testen. Unter Verwendung einer Linearen-Programmier-Maschine (LPM) wurde ermittelt, wie häufig jeder einzelne Indikator für die Klassifikation genutzt wurde.

Zusätzlich kommt erschwerend hinzu, dass in der Entwicklungsphase etwa 10 % aller Fälle unbeantwortete Fragen beinhalteten. Die Problematik der fehlenden Antworten lässt sich auch im praktischen Einsatz nicht ausschalten, da sie sowohl aus der Nichtbeantwortung von Fragen aufgrund Unwissens als auch aus der Nichtanwendbarkeit aller Fragen des Fragebogens auf alle Unternehmen resultieren. Es zeigte sich aber, dass mit Verfahren zur Berechnung des wahrscheinlichsten Wertes für die fehlende Antwort eine den Anforderungen mehr als genügende Leistung erzielt werden kann.

Bemerkenswert war die Erkenntnis, dass die zur Klassifizierung notwendige Information in weniger als 15 % der ursprünglichen 667 als potentiell geeignet betrachteten Fraud-Risiko-Indikatoren enthalten ist. Der Fragebogen konnte somit auf 44 Fragen reduziert werden, wobei eine Klassifizierung mit einer Sicherheit von durchschnittlich über 95 % gewährleistet blieb. Die Reduktion des Fragebogens ergab über die Liste der wichtigsten 45 Indikatoren folgende 44 Fragen, geordnet nach ihrer Bedeutung (Score[29]):

27 Fraud Risk Assessment Tool, basierend auf 44 Fragen.

28 Durch geeignete Fragestellung konnten gelegentlich mehrere Indikatoren durch eine Frage erfasst werden.

29 Der Score basiert auf einer Scorefunktion, die sich aus zwei Komponenten zusammensetzt: (a) einer Projektion der entsprechenden Frage auf den Lösungsvektor (z. B. Hyperebenenvektor), m.a.W.: wie stark wird die Lösung von der Frage beeinflusst, und (b) einem Anteil, wie wichtig die Frage für eine Vorhersage der fehlenden Antworten ist. Letzteres ist in der Praxis sehr wichtig, da in den Daten ca. 10 % fehlende Antworten enthalten waren. Daher muss ausbalanciert werden, ob es wichtiger ist, eine Frage, die es erlaubt, die fehlenden Antworten gut vorherzusagen, zu erhalten oder eine Frage zu bevorzugen, die die Lösung beeinflusst. Die Einstellung dieser Balance wurde aus den Daten gelernt.

| Nr. | Score | Titel der Frage | ep/ea[30] |
|---|---|---|---|
| 1 | 76,90 | Ungewöhnliche Geschäftsvorfälle | ep |
| 2 | 57,60 | Rechnungswesenpersonal | ep |
| 3 | 56,96 | Kündigungen wegen doloser Handlungen | ep |
| 4 | 36,68 | Persönliche Ziele der Gesellschafter/Manager | ep |
| 5 | 29,91 | Umstrukturierungen | ep/ea |
| 6 | 28,50 | Kritische Unternehmenssituation | ep/ea |
| 7 | 26,67 | Ungewöhnliche Transaktionen | ep |
| 8 | 23,13 | Wahrheitsgehalt der Angaben | ep |
| 9 | 21,78 | Persönliche Verknüpfungen | ep/ea |
| 10 | 20,87 | Entlegene Niederlassungen | ep/ea |
| 11 | 20,79 | Erlangung von Prüfungsnachweisen | ep |
| 12 | 16,99 | Wissensabfluss | ep |
| 13 | 16,38 | Führungsstil | ep |
| 14 | 16,34 | Besonderheiten Steuerberater, Rechtsanwalt, Consultant | ep |
| 15 | 16,11 | Handel von Unternehmensteilen | ep |
| 16 | 15,72 | Existenz, Verbreitung, Kontrolle, Anpassung eines Unternehmensleitbildes | ep |
| 17 | 14,63 | Rechtsform | ea |
| 18 | 14,48 | Rechnungsbegleichung (Ausgangsrechnungen) | ep |
| 19 | 13,50 | Niederlassungen / Beteiligungen | ep/ea |
| 20 | 13,30 | Rechnungsbegleichung (Eingangsrechnungen) | ep |
| 21 | 12,48 | Geschäftsvorfälle mit bestimmten Ländern | ep/ea |
| 22 | 12,31 | Bankkonten | ep |
| 23 | 11,73 | Kassenführung und -kontrolle | ep |
| 24 | 11,07 | Handhabung Buchführungsgrundsätze | ep |
| 25 | 10,59 | Anzahl der Mitglieder und Dienstjahre Top-Management | ep/ea |
| 26 | 9,83 | Einstellung zu den Prüfungsergebnissen | ep |
| 27 | 9,50 | Gesamteindruck | ep |
| 28 | 9,05 | Funktionstrennung | ep |
| 29 | 9,00 | Management Overview | ep |
| 30 | 8,77 | Informationsverhalten | ep |
| 31 | 8,37 | Leistungsabhängige Vergütung | ep |
| 32 | 8,20 | Belegerstellung | ep |
| 33 | 7,87 | Dokumentation der Netzwerk- und Anwendungsberechtigungen | ep |
| 34 | 7,87 | Prüfungsergebnisse | ep |
| 35 | 7,57 | Ausschreibungen | ep |
| 36 | 7,08 | Machtkonzentration | ep |
| 37 | 6,92 | Verhaltenskodex für Einkaufsmitarbeiter | ep |
| 38 | 6,82 | Produktänderungen | ea |
| 39 | 6,45 | Zusätzliche Aufträge | ep/ea |
| 40 | 6,40 | Nummerierung | ep |
| 41 | 6,35 | Verhalten gegenüber Aufsichtsbehörden | ep/ea |
| 42 | 6,22 | Beförderungen | ep |
| 43 | 6,15 | Installationsdatenträger | ep |
| 44 | 6,04 | Alter | ea |
| | 739,88 | (Gesamtscore der 44 wichtigsten Fragen) | |

*Übersicht 3: Die 44 wichtigsten Fragen*

Unter den 44 wichtigsten Fragen (siehe Übersicht 3) sind lediglich zwölf Fragen, die sich u. U. vor einer Erstprüfung durch die Auswertung extern verfügbarer Quellen erheben lassen. Da 32 Fragen nicht ohne Kenntnisse aus vorherigen Abschlussprüfungen des Unternehmens beantwortet werden können, eignet sich der Ansatz nicht für Auftragsannahmeentscheidungen (Erstprüfungen), wohl aber zur Unterstützung der Überlegungen der Annahme eines Folgeauftrags.

### *6.3 Die Bedeutung von Rechnungswesendaten für die Fraud-Risiko-Prognose*

Interessanterweise befand sich unter den 100 wichtigsten Indikatoren kein einziger aus der Analyse der Rechnungswesendaten, der aus ihnen berechneten Kennzahlen und der Zeit- bzw. Branchenvergleichswerte. Tatsächlich verschlechterten diese Werte das Ergebnis sogar. Das beste Ergebnis, das mit der isolierten Analyse der Rechnungswesendaten erzielt werden konnte, war ein Fehler von ca. 21% bzw. eine Klassifizierungsleistung von 79% mit Support-Vektor-Maschinen und Lineare-Programmier-Maschinen. In Anbetracht des betriebenen Aufwandes (finanz- und erfolgswirtschaftliche Analyse, Zeit- und Branchenvergleiche) zeichnet sich damit eine Grenze der Klassifizierungsfähigkeit durch Jahresabschlussanalyse ab. Sie liegt auf jeden Fall weit entfernt von der seitens des Berufsstandes geforderten Urteilssicherheit von 95%.

Es stellte sich die Frage, ob die Analyse der Rechnungswesendaten zusätzlich zur Analyse der qualitativen Daten eine Verbesserung der Klassifizierungsleistung zur Folge hätte. Verwendet man die Rechnungswesendaten zusätzlich zu den Daten aus dem Fragebogen, verschlechtern sich die Ergebnisse um 3–4%, d.h. die Rechnungswesendaten-Information verwirrt die Methoden mehr, als dass sie hilft. Folglich war nur der Entwicklungsfragebogen zu reduzieren, die Analyse von Rechnungswesendaten ist im vorliegenden System irrelevant.

Ursache für die geringe Bedeutung von Abschlussdaten für die Fraud-Risiko-Beurteilung ist wahrscheinlich die Tatsache, dass es ja im Regelfall genau die Rechnungsweseninformationen sind, in denen die dolosen Handlungen verschleiert werden sollen. Es ist ein vorrangiges Bestreben des Täters, die Rechnungswesendaten so darzustellen, dass es gerade nicht zu Auffälligkeiten kommt. Natürlich gelingt dies nicht vollständig und auch nicht jeder Täter ist überhaupt fachlich und organisatorisch in der Lage, die Abbildung seiner Handlung im Rechnungswesen zu verschleiern. Dennoch hat nicht jeder erhöhte Aufwand bzw. geringere Umsatzerlös gleich seine Ursache in Fraud. Dass es überhaupt zu

30 Erhebbarkeit ex post oder ex ante.

einer Klassifizierungsleistung von 79 % auf der Basis von Rechnungswesendaten kommt, ist wahrscheinlich auf die mutmaßlich hohe Korrelation einer schwierigen wirtschaftlichen Situation, die sich verhältnismäßig gut aus Abschlussinformationen herauslesen lässt, und dem Fraud-Risiko zurückzuführen. Dieser Zusammenhang lässt sich jedoch ebenso mit qualitativen Fragen an den Wirtschaftsprüfer feststellen, zumal sich eine schwierige wirtschaftliche Situation nicht nur und vor allem nicht sofort in den Abschlüssen widerspiegelt.

Die Erhebung der Rechnungswesendaten gestaltet sich zudem in der Praxis sehr schwierig. Es bedarf in einem internationalen Umfeld eines großen Aufwandes, die Informationen zu normieren und geeignete Werte, z. B. für die Branchenanalyse, zu beschaffen.

In Anbetracht dieses deutlichen Ergebnisses relativiert sich die Empfehlung der Prüfungsstandards, die Beurteilung des Fraud-Risikos auf analytische Prüfungshandlungen zu stützen[31] – zumindest hinsichtlich Zeit- und Branchenvergleichen von Rechnungswesendaten.

### *6.4 Stationarität von Fraud*

#### *6.4.1 Internationalität von Fraud*

Ein Ziel der Arbeit war es, den Ansatz länderunabhängig zu gestalten. Um zu testen, ob die Leistungsfähigkeit des Ansatzes bei der Anwendung auf internationale Fälle bestehen bleibt, wurde die Lernmaschine mit 100 Fällen aus Deutschland und Großbritannien trainiert und mit 15 ihr unbekannten Fällen aus Frankreich, der Schweiz und den Niederlanden in einem Blindtest konfrontiert.[32] Von diesen 15 internationalen Fällen wurden elf Fälle als sicher klassifiziert gekennzeichnet. Diese elf Fälle wurden richtig klassifiziert. Von den übrigen vier nicht als sicher gekennzeichneten Fällen war lediglich eine Beurteilung tatsächlich falsch. So wurde gezeigt, dass die mit den deutschen und englischen Fällen trainierte Maschine ohne weiteres auf Fälle aus Frankreich, der Schweiz und den Niederlanden angewendet werden konnte.

Da das Verfahren auf dem Prinzip der Erkennung von Fraud-Mustern basiert, lässt sich ableiten, dass die Muster zumindest in diesen EU-Ländern sehr ähnlich sind. Das ist umso beachtlicher, als die wirtschaftlichen und rechtlichen Gegebenheiten in den getesteten fünf Ländern durchaus differieren.

#### *6.4.2 Stationarität hinsichtlich der Branche, Rechtsform, Unternehmensgröße*

Ziel des Ansatzes war es, mit hinreichender Sicherheit prüfungspflichtige Unternehmen jeder Rechtsform, Branche und Größe hinsichtlich deren Fraud-Risiko richtig zu klassifizieren. Die internationalen Prüfungsstandards gehen davon aus, dass die Aussagefähigkeit von Fraud-Risiko-Indikatoren von der Größe, Rechtsform oder Branche abhängt, so dass es letztlich beim Prüfer liegt, diese jeweils in Relation zu setzen.[33]

Da bis auf den Ausschluss von Unternehmen des Finanzsektors[34] keinerlei vorgegebenen Beschränkungen[35] hinsichtlich der Einbeziehung von Branchen, Rechtsformen und Größen bestanden, kann aufgrund der sehr hohen Klassifizierungsleistung von über 95 % davon ausgegangen werden, dass die dem Fraud-Risiko zugrunde liegenden Muster weitgehend unabhängig von der Rechtsform, Branche und Größe eines prüfungspflichtigen Unternehmens sind.

#### *6.4.3 Unabhängigkeit von der Art der Tat*

Da Fraud-Fälle in die Untersuchung dann einbezogen wurden, wenn für die Abschlussprüfung wesentlicher Fraud vorlag, war die Art der Tat für die Einbeziehung nicht entscheidend. Unter den Fällen befanden sich alle Arten bestätigungsvermerksrelevanter Unregelmäßigkeiten, wie Bilanzmanipulationen, Vermögensschädigungen und Gesetzesverstöße[36], wobei die Bilanzmanipulationen bei weitem überwogen. Dabei traten die Tatarten häufig in Kombination miteinander auf, was auch nahe liegend ist, da eine Vermögensschädigung im Jahresabschluss verschleiert wird und eine Bilanzmanipulation Vermögensschädigungen zur Folge haben kann.

Daher scheint es nicht erforderlich zu sein, für die verschiedenen und vielfältigen Erscheinungsformen von Fraud spezielle Prognosemethoden zu entwickeln, da es sich bei den angegebenen Fraud-Indikatoren offenbar um unterschiedliche Symptome derselben Ursachen handelt.

---

31 Vgl. *AICPA,* Statement on Auditing Standards (SAS 99): Consideration of Fraud in a Financial Statements Audit (Stand: 2002), Rn. 28 ff.; *IFAC,* ISA 240 (rev.), Rn. 53 f.; *IDW PS 210,* WPg 2003, S. 655 ff., Rn. 44.

32 Bis zur Durchführung dieses Tests waren diese 15 neuen Fälle unter Verschluss gehalten worden. Die Vorhersage erfolgte „blind" und es war garantiert, dass das FRAT44-System diese Daten nie zuvor gesehen hat.

33 Vgl. *IFAC,* ISA 240 a. F., Rn. 36 f.; *IFAC,* ISA 240 (rev.), Rn. 51 f.

34 Aufgrund der hohen Spezifität von Prozessen und Jahresabschlussinformationen wurden Unternehmen des Finanzsektors (Banken, Versicherungen, Leasing-Gesellschaften etc.) ausgeschlossen.

35 Auch wenn keine Beschränkungen vorgegeben waren, führt die auf derzeit 115 Fälle beschränkte Zahl der Unternehmen der Grundgesamtheit zu einer tatsächlichen Limitierung, da theoretisch keine Aussage über Branchen, Rechtsformen und Unternehmensgrößen getroffen werden kann, die nicht in der Grundgesamtheit enthalten waren. Die Tatsache, dass fast ausschließlich prüfungspflichtige Unternehmen in die Untersuchung einbezogen wurden, limitiert eine Aussage zu Kleinunternehmen, Einzelunternehmen und Personengesellschaften. Da jedoch auch Nischenunternehmen, Unternehmen mit mehrheitlich öffentlicher Beteiligung, prüfungspflichtige Personenhandelsgesellschaften und freiwillig geprüfte kleine Kapitalgesellschaften (Start-Up-Unternehmen) in der Grundgesamtheit enthalten waren und dies zu keinen besonderen Schwierigkeiten bei der Erhebung und Klassifizierung führte, kann wohl die generalisierte Aussage zu Stationarität unter praktischen Gesichtspunkten aufrecht erhalten werden.

36 Vgl. *IDW PS 210,* WPg 2003, S. 655 ff., Rn. 7 ff.

### 6.5 Erklärungskomponente

Um die Aussagefähigkeit und Akzeptanz des Ansatzes weiter zu erhöhen, wurde eine Erklärungskomponente geschaffen, die den „Black Box"-Charakter eines reinen Klassifizierungstools mildert.[37] Dabei wurden verschiedene Ansätze hinsichtlich ihrer Machbarkeit diskutiert:

(1) eine Visualisierung der für die Entscheidungsfindung relevanten und im Fragebogen abgebildeten Prozessinformationen,

(2) die Benennung und das Ranking gemäß Ergebnisbeitrag der für die Entscheidung relevanten Fragen und

(3) eine auf Wahrheitswerten basierende Verknüpfung der verwendeten Indikatoren.

Variante (1) hat einen zu geringen praktischen Nutzen. Wird nur ein Prozessschritt durch die Erklärungskomponente benannt, birgt das unter Umständen für das Prüfungsteam keine neuen Informationen; werden zu viele Prozessinformationen als relevant aufgezeigt, verhindert das ggf. eine gezielte fraud-risiko-orientierte Ausrichtung der Prüfung. Variante (3) scheidet aufgrund des exponentiell mit der Zahl der Indikatoren ansteigenden Rechenaufwandes für die genaue Berechnung aus. Es gibt zwar heuristische (polynomiale) Näherungslösungen für dieses Problem, doch sind diese nicht einfach implementierbar bzw. für das vorliegende Problem adaptierbar.

Daraus ergibt sich als Lösung Variante (2), also die Ausgabe der für die Klassifizierung entscheidenden Fragen. Das FRAT44-System generiert nun nicht nur eine Risiko-Beurteilung und einen Sicherheitswert, sondern trifft durch die Rückgabe der Fragen entsprechend ihrer Klassifizierungsbeiträge auch Aussagen über die Entscheidungsfindung. Das entwickelte Erklärungsmodul liefert drei Listen:

(a) alle Fragen, die die Klassifizierung unterstützten,

(b) alle Fragen, die ihr entgegenstanden, und

(c) die Fragen, die sie sowohl unterstützten als auch ihr entgegenstanden.

So hat der Prüfer die Möglichkeit, die Risikokonstellation einzugrenzen. Wenn z. B. insbesondere die Fragen nach „Produktänderungen", „Wissensabfluss" und „Dokumentation der Netzwerk- und Anwendungsberechtigungen" die Klassifizierung unterstützt haben, deutet das auf eine Risikokonstellation im Bereich „Intellectual Property"[38] hin.

### 6.6 Vergleich mit deutschen und internationalen Prüfungsstandards

#### 6.6.1 Vergleich der Ergebnisse mit IDW PS 210

Der am 8. 5. 2003 verabschiedete *IDW PS 210* enthält eine beispielhafte Aufzählung von Red Flags zur Orientierung für den Abschlussprüfer. Diese Aufzählung entspricht weitestgehend der in ISA 240 a. F. enthaltenen Liste.[39] Da ISA 240 a. F. und ISA 250 schon zum Zeitpunkt der Entwicklung des Fragebogens zur Verfügung standen und in die Entwicklungsversion einflossen, kann nun die empirische Relevanz der Fraud-Indikatoren des *IDW PS 210* überprüft werden.

In *IDW PS 210* sind folgende Red Flags genannt:

| | FRAT44[40] |
|---|---|
| Beherrschung des Geschäftsführungsgremiums durch eine oder wenige Personen … | ✓ |
| … ohne ein wirksames Überwachungsorgan | – |
| Undurchsichtige Organisationsstrukturen | – |
| Aggressive Ausnutzung von Wahlrechten und Beurteilungsspielräumen durch die Unternehmensleitung | ✓ |
| Fehlende Bereitschaft zur Verbesserung des internen Kontrollsystems | ✓ |
| Fehlende Bereitschaft der Unternehmensleitung, unterjährig bekannt gewordene Fehler in der Buchhaltung zeitnah zu korrigieren | ✓ |
| Häufiger Personalwechsel in Führungspositionen | ✓ |
| Dauerhafte personelle Unterbesetzung der Buchhaltungsabteilung | – |
| Häufiger Wechsel des Abschlussprüfers | – |
| Schrumpfende, stagnierende oder auch stark expandierende Geschäftstätigkeit | ✓ |
| Unzureichende Kapitalausstattung sowie Abhängigkeit von einzelnen Kreditgebern | – |
| Ungünstige Ergebnisentwicklung | – |
| Risikoreiche Ertragsquellen oder steigender Anteil der nicht betrieblichen Erträge | – |
| Notwendigkeit, Gewinnminderungen im operativen Geschäft durch Sondermaßnahmen zu kompensieren | – |
| Abhängigkeit von wenigen Lieferanten oder Kunden | – |
| Geschäfte mit wesentlichen Gewinnauswirkungen (besonders gegen Jahresende) | ✓ |
| Komplizierte Geschäfte … | ✓ |
| … oder ungewöhnliche Bilanzierung von Geschäften | ✓ |
| Geschäfte mit nahe stehenden Personen und Unternehmen | ✓ |
| Im Verhältnis zur erhaltenen Leistung überhöhte Ausgaben für Vermittlungsprovisionen und für Rechts- oder Unternehmensberatung | ✓ |
| Mangelhafte Buchung oder Dokumentation von Geschäftsvorfällen | ✓ |

37 Dieser „Black Box"-Charakter hat auch Vorteile: So verlangt SAS 99, Tz. 50, „Elemente der Nichtvorhersehbarkeit" in die Prüfungshandlungen einzubeziehen, d. h. Prüfungshandlungen bei Gegenständen zu Zeitpunkten und an Orten vorzunehmen, die der Mandant nicht erwartet. Vergleichbares wird erreicht, wenn man zwar eine Klassifizierung, aber keinerlei Anhaltspunkte darüber erhält, wie diese zustande gekommen ist. Support-Vektor-Maschinen sind dabei deterministisch, d. h. in einer bestimmten Fallkonstellation wird immer dieselbe Klassifizierung erzeugt. Das Klassifizierungsurteil ist also eindeutig, jedoch aufgrund der Komplexität der Entscheidungsfindung für den Anwender – erst recht für das geprüfte Unternehmen – nicht vorhersehbar.

38 Verletzung des geistigen Eigentums (Industriespionage, Produktpiraterie).

39 *IDW PS 210* überführt ISA 240 und ISA 250 in nationales Recht; vgl. WPg 2003, S. 655 ff., Rn. 3.

40 In FRAT44 enthalten. Dieses Ergebnis ergibt sich durch den Vergleich des Wortlauts der Fragen bzw. der Antwortalternativen. Ist ein Red Flag des *IDW PS 210* in FRAT44 enthalten (✓), kann es als empirisch bestätigt betrachtet werden, andernfalls nicht (–).

| | FRAT44 |
|---|---|
| Hohe Zahl von Differenzen zwischen den Ergebnissen der Buchführung und den Bestätigungen Dritter | ✓ |
| Schwer prüfbare Buchführungssysteme | ✓ |
| Ausweichende oder schwer nachvollziehbare Auskünfte der gesetzlichen Vertreter zu Anfragen des Abschlussprüfers | ✓ |
| Fehlende Bereitschaft der gesetzlichen Vertreter, den vorhergehenden Abschlussprüfer auf Anforderung des Abschlussprüfers von der Verschwiegenheitsverpflichtung zu befreien | – |
| Hohe ergebnisabhängige Vergütungen für Mitarbeiter in leitender Funktion | ✓ |
| Unangemessen kurze Zeit zur Erstellung des Abschlusses | – |
| Die Unternehmensleitung steht unter starkem Druck, die eigenen (ggf. bereits veröffentlichten) Ergebniserwartungen oder die Erwartungen Dritter zu erfüllen (z. B. Ergebniserwartungen von Analysten oder institutionellen Investoren) | – |
| Unzureichende Wirksamkeit der Internen Revision | – |
| Fehlende oder veraltete Dokumentation des Aufbaus der Dateien oder der Programme | – |
| Zahlreiche Programmänderungen, die nicht dokumentiert, genehmigt oder getestet sind | – |
| Behördliche Untersuchungen sowie Straf- oder Bußgeldbescheide | – |
| Zahlungen für nicht spezifizierte Dienstleistungen oder Darlehen an Berater, nahe stehende Personen sowie Mitarbeiter des Unternehmens oder Behörden sowie an Personen, die dem benannten Personenkreis nahe stehen | ✓ |
| Außergewöhnliche Einkaufs- oder Verkaufspreise, hohe Provisionen | ✓ |
| Ungewöhnliche Zahlungen in bar, durch Inhaberschecks oder auf Nummernkonten oder Zahlungen ohne angemessenen Nachweis des Zahlungsgrundes | ✓ |
| Außergewöhnliche Geschäfte mit Unternehmen in Niedrigsteuerländern | ✓ |
| Zahlungen an einen Empfänger in einem anderen Land als dem Ursprungsland von bezogenen Waren oder Dienstleistungen | – |
| Nicht autorisierte oder ungenügend dokumentierte Geschäfte | ✓ |
| Nachteilige Presseberichterstattung | – |

*Übersicht 4: Vergleich von IDW PS 210 und FRAT 44*

Von den in *IDW PS 210* genannten Red Flags wurden 21 als Bestandteil von FRAT44 als statistisch signifikant bestätigt (18 statistisch nicht signifikant). Obwohl die „Messung" durch Zählung von Fragen nicht exakt ist, ist dieser Wert im Vergleich zu dem 15 %-Anteil der relevanten Fragen an den Fragen des Entwicklungsfragebogens sehr hoch. Die relevanten Fragen sind in ihrer Bedeutung – gemessen am Score (vgl. Überischt 3) – relativ gleichmäßig verteilt. Fünf der zehn wichtigsten Fragen der Endfassung des Fragebogens decken sich mit den Red Flags des *IDW PS 210,* wobei die wichtigste Frage des FRAT44 „ungewöhnliche Geschäftsvorfälle" in der gewählten Formulierung und mit den entsprechenden Antwortalternativen gleich vier Red Flags des Prüfungsstandards abbildet, was die Bedeutung dieser Frage unterstreicht.

Die andere Hälfte der in *IDW PS 210* genannten Red Flags, die nicht in der Endfassung des Fragebogens enthalten waren, sind nicht zwangsläufig unwichtig. Sie können im Einzelfall durchaus entscheidende Anregungen für die Aufdeckung von Unregelmäßigkeiten liefern.

### *6.6.2 Ein Blick in die Zukunft: Vergleich mit ISA 240 (rev.)*

Der im Februar 2004 verabschiedete ISA 240 (rev.) orientiert sich weitgehend am US-amerikanischen Standard der AICPA (SAS 99) und enthält ebenfalls eine Aufzählung von Red Flags, die gegenüber der Indikatoren-Liste in ISA 240 a. F. dem Fraud-Triangle-Prinzip folgt[41]: die Fraud-Risiko-Faktoren sind gegliedert nach Anreizen bzw. Druck, Möglichkeiten zum Begehen der Tat und Anzeichen für die Fähigkeit, die Tat zu rechtfertigen. Dabei werden die Indikatoren jeweils getrennt für Bilanzmanipulationen und Vermögensschädigungen angegeben.

Von den 44 Faktoren, die auf ein Risiko von Bilanzmanipulationen hindeuten, waren 37 in der Entwicklungsversion des Fragebogens enthalten; von den 23 Risikofaktoren, die im Zusammenhang mit Vermögensschädigungen auftreten können, waren 19 in der Entwicklungsversion abgebildet.

Bei den nicht in der Entwicklungsversion enthaltenen Risiko-Indikatoren handelt es sich fast ausschließlich um Fragen, die zwar theoretisch bestechend und in der Literatur bekannt, jedoch im Rahmen einer Jahresabschlussprüfung schwer zu erheben sind und daher bei der Erstellung des Entwicklungs-Fragebogens nicht einbezogen wurden.

Von den auf Bilanzmanipulationen hindeutenden Red Flags erwiesen sich 19, also etwa die Hälfte, als empirisch relevant, indem sie in der Endfassung des Fragebogens (FRAT44) enthalten waren.[42] Von den Fraud-Risk-Indikatoren, die gemäß ISA 240 (rev.) im Zusammenhang mit Vermögenschädigungen auftreten, fanden sich nur drei als entscheidungsrelevant in der Endfassung des FRAT44 wieder.

| | ISA 240 (rev.) | Entwicklungsversion des Fragebogens | FRAT44 |
|---|---|---|---|
| Indikatoren Bilanzmanipulation | 44 | 37 | 19 |
| Indikatoren Vermögensschädigung | 23 | 19 | 3 |

*Übersicht 5: Vergleich der Red Flags in ISA 240 (rev.)/SAS 99[43] mit FRAT44*

41 Vgl. *IFAC,* ISA 240 (rev.), Appendix 1.

42 Natürlich konnten die in der Entwicklungsfassung des Fragebogens nicht enthaltenen Fragen auch nicht bewertet werden.

43 Die Liste in ISA 240 (rev.) ist bis auf vier zusätzliche Indikatoren identisch mit den im Anhang angegebenen Fraud-Risiko-Faktoren des SAS 99, der als Vorbild für ISA 240 (rev.) diente.

Die Gründe hierfür können vielfältig sein. Eine Hauptursache ist wohl, dass die in den Standards vorgeschlagenen Indikatoren zwar theoretisch durchaus sinnvoll sind, aber in der Praxis nicht zwangsläufig auch eindeutig beantwortet werden können. Wann liegt z. B. der Indikator, dass eine Organisationsstruktur übermäßig komplex ist, vor? Wann genau also ist die Grenze überschritten, an der differenzierte Strukturen betriebswirtschaftlich noch zweckmäßig oder rechtliche Gestaltungen noch nicht ungewöhnlich sind?

An diesem Beispiel wird das Problem von Checklisten deutlich, nämlich dass viele Fragen weit auslegbar sind. Selbst die hier gewählte (in den Standards nicht enthaltene) Vorgabe von standardisierten Antwortalternativen und deren Bewertung im Experten-Dialog kann – wie die praktische Datenerhebung zeigte – dieses Beurteilungs-Problem nicht gänzlich eliminieren. Letztlich ist hier der Abschlussprüfer auf sein pflichtgemäßes Ermessen angewiesen.[44]

Auch wenn sich der Großteil der in den Standards vorgegebenen Fraud-Risiko-Indikatoren empirisch als nicht zu einer konkreten Messung des Risikos geeignet erwies, können diese Red Flags dennoch Denkanstöße für die Prüfung geben, z. B. in der Diskussion der Fraud-Anfälligkeit des geprüften Unternehmens innerhalb des Prüfungsteams[45]. Diese Bedeutung ist nicht zu unterschätzen, handelt es sich doch bei einer Jahresabschlussprüfung um eine in erster Linie intellektuelle Leistung.

### 7. Ergebnis

ISA 240 (rev.) geht davon aus, dass „es nicht einfach ist, Fraud-Risiko-Indikatoren zu einem effektiven Vorhersagemodell zu kombinieren".[46] FRAT44 führt den Beweis, dass eine solche Kombination möglich und effektiv ist. Zwar bedarf es einer empirischen Vorgehensweise auf der Basis vieler geeigneter Indikatoren und von hochentwickelten mathematisch-statistischen Methoden, aber die notwendige Information kann aus geeigneten Red Flags abgeleitet werden.

Das FRAT-Projekt zeigt, dass es möglich ist, auf der Basis von Methoden des Maschinellen Lernens und im Rahmen der Jahresabschlussprüfung verfügbarer Informationen Fraud-Muster mit einer Sicherheit von über 95 % zu erkennen und die Risikokonstellation einzugrenzen.

Die Untersuchung zeigt, dass für Prüfungsanwendungen, also für Klassifizierungsentscheidungen mit unbestimmten Entscheidungskriterien, Methoden des Maschinellen Lernens hervorragende Ergebnisse liefern, da sie in der Lage sind, durch Mustererkennung aus den zur Verfügung stehenden Daten ein Maximum von entscheidungsrelevanten Informationen zu extrahieren.

Aufgrund der empirischen Basis führt die Anwendung von maschinellen Lernmethoden zu einer Objektivierung der Risikobeurteilung.

Da der Ansatz das Fraud-Risiko über die Erkennung von Fraud-Mustern bestimmt, kommt ihm im Prüfungsprozess eine (Früh-)Warnfunktion zu: im Falle einer eindeutigen Fraud-Klassifikation hat das Prüfungsteam zu prüfen, ob der gewählte Prüfungsansatz dem Fraud-Risiko-Niveau gerecht wird. Gegebenenfalls hat es die Prüfungshandlungen anzupassen – bis hin zur Einbeziehung von forensischen Spezialisten.

FRAT44 stellt somit ein ergänzendes, fraud-spezifisches Prüfungsinstrument im Kontext mit anderen Prüfungshandlungen dar.

Es leistet einen Beitrag zur Reduktion der so genannten „Performance-Gap" als Teil der Erwartungslücke[47]. FRAT44 kann dazu beitragen, dass Fraud-Fälle früher als bisher möglich erkannt werden.

44 Vgl. *IFAC*, ISA 240 (rev.), Rn. 50.
45 Vgl. *IFAC*, ISA 240 (rev.), Rn. 27–32.
46 Vgl. *IFAC*, ISA 240 (rev.), Rn. 50.
47 Die Erwartungslücke (Expectation Gap) setzt sich nach allgemeinem Verständnis aus einer vom Prüfer zu verantwortenden Leistungs-Lücke (Performance Gap) einerseits und einer ungerechtfertigten Übererwartung der Öffentlichkeit andererseits zusammen; vgl. *Schruff*, Aus der Facharbeit des *IDW*, WPg 2004, S. 455.

# Strukturelle Probleme bei der Implementierung des Goodwill-Impairment-Tests
## – Der Ansatz von IAS 36 im Vergleich zu US-GAAP –

*Von WP/StB Dr. Norbert Lüdenbach, Frankfurt a. M., und WP/StB Dr. Wolf-Dieter Hoffmann, Freiburg i. Br.*

### 1. Einleitung

Die Neufassung von IAS 36 (rev. 2004)[1] und die Verabschiedung von IFRS 3 vollziehen die vor drei Jahren durch SFAS 142 erfolgten Änderungen der amerikanischen Bilanzierungsregeln nach. Die planmäßige Goodwillabschreibung wird durch einen Im-

1 Nachfolgend vereinfacht zitiert als IAS 36.