



Leitfaden Datenanalyse im Prüfungsprozess

Herausgeber:

ISACA Germany Chapter e.V.
Storkower Straße 158
10407 Berlin

www.isaca.de
info@isaca.de

Autorenteam:

- Phillip Bär, RLT IT- und Systemprüfung GmbH WPG
- Sebastian Bechinger
- Dr. Anne Hofmann, KfW
- Klaus Jakobi, Genossenschaftsverband – Verband der Regionen e.V.
- Andreas Kietzmann, Wilh. Werhahn KG
- Jutta Köhn, ROSER GmbH WPG
- Henning Lieder, KPMG AG WPG
- Mirko Ludwig, Trinavis GmbH & Co. KG WPG
- Carsten Meyer, Landessparkasse zu Oldenburg
- Daniel Missal, Ebner Stolz GmbH & Co. KG WPG
- Thomas Möllers, INSO Projects GmbH
- Thorsten Rumpf, Rödl & Partner GmbH, Nürnberg
- Alexander Rühle, zapliance GmbH
- Dr. Ingo Stock, BMW Group
- Johannes Martin Wagner
- Robert Wischnat, Fresenius SE & Co. KGaA

Die Inhalte dieses Leitfadens wurden von Mitgliedern des ISACA Germany Chapter e.V. in Kooperation mit dem BSI erarbeitet und sind sorgfältig recherchiert. Trotz größtmöglicher Sorgfalt erhebt die vorliegende Publikation keinen Anspruch auf Vollständigkeit. Sie spiegelt die Auffassung des ISACA Germany Chapter wider. ISACA Germany Chapter e.V. übernimmt keine Haftung für den Inhalt.

Der jeweils aktuelle Leitfaden kann unter <https://www.isaca.de/de/fg-veroeffentlichungen> kostenlos bezogen werden. Alle Rechte, auch das der auszugsweisen Vervielfältigung, liegen beim ISACA Germany Chapter e.V.

Stand:

September 2019 (Final nach Review und Überarbeitung durch ISACA-Fachgruppe Datenanalyse)

Leitfaden

Datenanalyse im Prüfungsprozess

Vorwort

»Wer einen Fehler begangen hat und ihn nicht korrigiert,
begeht einen weiteren Fehler.«

Konfuzius (chinesischer Philosoph –
vermutlich 551 bis 479 v. Chr.)

Das Aufspüren von Fehlern oder allgemeiner von Abweichungen bzw. Auffälligkeiten ist eine Kernaufgabe des Prüfers. Durch die stetig voranschreitende Automatisierung und Digitalisierung der Wirtschaft spielen in diesem Zusammenhang Daten und die Datenanalyse eine immer wichtigere Rolle. Als Fachgruppe für die Datenanalyse sind wir jeden Tag mit diesem Problem konfrontiert. Nachdem gängige Methoden durch die hohen Datenmengen immer schwerer zu bedienen sind, wurden auf der Suche nach neuen und einfacheren Methoden neue Programme entdeckt bzw. entwickelt.

Durch die steigende Komplexität und Menge an Daten erschwert sich deren Analyse. Einfache Methoden lassen sich nicht mehr anwenden und immer mehr Fragestellungen lassen sich nur noch durch den intelligenten Einsatz der Datenanalyse effektiv und effizient bewältigen.

In der Vergangenheit wurde in der Prüfung vielfach das Verfahren der manuellen Belegprüfung angewendet. Inzwischen haben sich Tabellenkalkulationsprogramme etabliert, mit der ad hoc exportierte Listen aus den Mandantenprogrammen schnell und unkompliziert analysiert werden können. Das Datenvolumen in den IT-Systemen aber steigt weiter. Selbst bei kleineren oder kleinen mittelständischen Unternehmen kommen einige Hunderttausend Buchungssätze pro Jahr schnell zusammen. Ohne IT-Unterstützung wäre eine effiziente, d.h. mit vertretbarem Aufwand durchgeführte Datenauswertung nicht mehr möglich. Noch umfangreicher stellt sich das Mengengerüst eines Unternehmens dar, wenn Materialwirtschaft und/oder Produktion mit einem ERP-System abgewickelt werden, dann ergeben sich weitaus größere Datensatzanzahlen.

Gleichzeitig steigen die Anforderungen bzgl. Qualität und Effizienz von Prüfungen. Auch in dieser Hinsicht ist durch die Anwendung der Datenanalyse eine Verbesserung der Aussagequalität möglich, da die Analysen im Gegensatz zur Stichprobenprüfung über den vollständigen Datenbestand erfolgen können. Dieser Aspekt tritt besonders im Rahmen der Internen Revision immer mehr in den Vordergrund. Denn

durch die Betrachtung des gesamten Datenbestandes können tiefere Einblicke erlangt werden, die ohne Datenanalyse gar nicht möglich sind.

Nur nachvollziehbare Analyseschritte führen zu einer belastbaren Aussagequalität. Die Optimierung des Prüfungsvorgangs durch Datenanalysen leistet einen wesentlichen Beitrag zur Prüfungssicherheit von Prüfungshandlungen. Qualitätssicherungsmaßnahmen, vergleichbar mit denen der Softwareentwicklung, unterstützen hierbei die Aussagekraft der Analysen. Die Wiederverwendbarkeit von Analysen steigert zusätzlich die Effizienz der Prüfung. Einmal erstellte Analysefunktionen können mehrfach für verschiedene Daten angewendet werden. Damit reduziert sich der Prüfungsaufwand, und der Automatisierungsgrad der Prüfung steigt.

Wir als Fachgruppe Datenanalyse haben in diesem Leitfaden unsere gemeinsame Erfahrung aus Interner Revision, Wirtschaftsprüfung, Beratung und anderen Funktionen zusammengetragen und würden uns sehr freuen, wenn dieser Leitfaden wertvolle Anregungen für Sie bereithält.

Wir wünschen Ihnen eine angenehme und aufschlussreiche Lektüre. Über einen Austausch bzw. Rückmeldung bezüglich des Leitfadens würden wir uns sehr freuen.

Ihre Fachgruppe Datenanalyse
FG-Datenanalyse@isaca.de

Inhaltsverzeichnis

1. Einleitung	3
2. Hauptprozessschritte des Datenanalyseprozesses	4
I. Entscheiden über Auftragserteilung zur Datenanalyse	6
II. Planen der Datenanalyse	8
III. Daten anfordern und Daten bereitstellen	11
IV. Daten übernehmen und Daten validieren	15
V. Aufbereiten der Daten	16
VI. Analysieren der Daten	18
VII. Erstellen der Berichterstattung.....	20
VIII. Dokumentieren der Datenanalyse und	22
Durchführen der Qualitätssicherung	
IX. Sicherstellen der Informationssicherheit und Cyber Security.....	24
3. Fazit und Ausblick	25
4. Abkürzungsverzeichnis	26
5. Glossar	27
6. Literaturverzeichnis	29

1. Einleitung

Dieser Leitfaden stellt eine methodische Vorgehensweise zur Analyse von komplexen Daten vor. Er greift die im Vorwort benannten Herausforderungen auf und liefert Antworten auf aktuelle Fragestellungen beim Umgang mit Daten. Ferner bildet er ein fachliches Fundament für den Umgang mit verwandten Themen, wie beispielsweise Data Mining oder Continuous Auditing / Monitoring.

Dabei versteht sich dieser Leitfaden keineswegs als zwingend anzuwendendes und starres Vorgehensmodell, sondern als Grundlage für ein Konzept, das situativ und flexibel auf die jeweilige Auftragssituation anzupassen ist. Ferner soll er zu einem besseren Verständnis der handelnden Personen untereinander beitragen und ihre Schnittstellen zueinander beschreiben. Der Leitfaden ist dabei bewusst unabhängig von Unternehmensorganisationen, IT-Systemen und Datenanalysewerkzeugen gehalten, um in einem möglichst breiten Umfeld anwendbar zu sein. So kann in vielen Fällen die Datenanalyse auch direkt in der datenhaltenden Anwendung (z.B. ERP- oder BW-System) erfolgen. In diesem Fall müssen z.B. die Prozessschritte zur Datenextraktion ggf. nur eingeschränkt angewendet werden.

Des Weiteren schließt der Leitfaden als praxisorientiertes Vorgehensmodell dabei aus Sicht der Fachgruppe Datenanalyse eine Lücke. Denn bisher gibt es nur vereinzelte Veröffentlichungen, die sich grob in folgende Klassen abbilden lassen:

- ▶ Veröffentlichungen, die sich mit den Chancen (Risiken) der Datenanalyse als Prüfungswerkzeug allgemein befassen
- ▶ Artikel in Form von Erfahrungsberichten zur Anwendung der Datenanalyse in spezifischen Einsatzgebieten
- ▶ Umsetzung der Datenanalyse im Rahmen von spezifischen Prüfungsframeworks

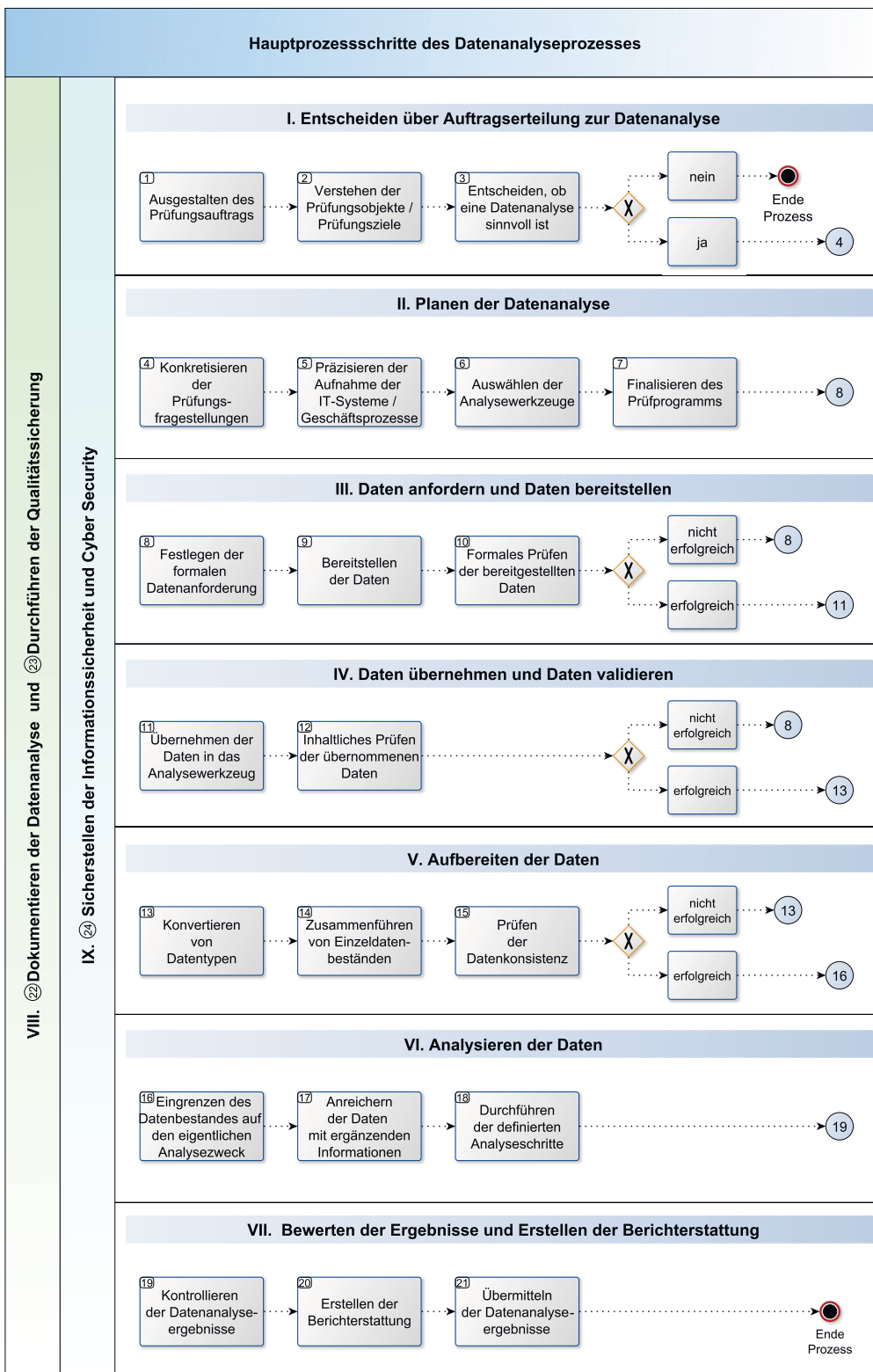
Der Leitfaden versteht sich als übergeordnetes Vorgehensmodell zur Datenanalyse im Prüfungsprozess und ist neben der Anwendung im Rahmen der klassischen Datenanalyse somit auch auf die Bereiche Big Data, Process Mining, Continuous Auditing / Monitoring anwendbar. Obwohl diese Themen schon seit mehreren Jahren in der Wissenschaft diskutiert werden und in der Praxis Anwendung finden, werden diese jedoch häufig nur sporadisch bzw. ohne methodische Grundlage verwendet. Gleichzeitig stehen Revisoren beim Umgang mit diesen Themen vor komplexen Herausforderungen. So existiert in der heutigen Unternehmenswelt kaum eine Information, die nicht in digitaler Form abgebildet ist. Dabei können diese Daten eine Vielzahl unterschiedlicher Formen und Strukturen aufweisen und verschiedenen Quellen entstammen. Dieser Umstand erfordert vom Prüfer ein konsequentes Umdenken in seinem Vorgehen bei der Sammlung und Analyse relevanter Informationen.

Datenanalysen in der Prüfung bzw. in der Revision werden nach Auffassung der Verfasser dieses Leitfadens in den nächsten Jahren weiter an Bedeutung gewinnen, da sie eine Möglichkeit darstellen, den Prüfungs- bzw. Revisionsprozess zu optimieren und diesen effektiv und effizient zu gestalten.

Zur Erläuterung der Begrifflichkeiten befindet sich am Ende des Dokuments ein Glossar.

In den folgenden Kapiteln wird der Prozess zur Datenanalyse im Grundprinzip beschrieben.

2 Hauptprozessschritte des Datenanalyseprozesses



Der Leitfaden gliedert sich in neun Hauptprozessschritte, die sich wiederum in einzelne Unterprozessschritte unterteilen.

Die Hauptprozessschritte lauten:

- I. Entscheiden über Auftragserteilung zur Datenanalyse
- II. Planen der Datenanalyse
- III. Daten anfordern und Daten bereitstellen
- IV. Daten übernehmen und Daten validieren
- V. Aufbereiten der Daten
- VI. Analysieren der Daten
- VII. Bewerten der Ergebnisse und Erstellen der Berichterstattung
- VIII. Dokumentieren der Datenanalyse und Durchführen der Qualitätssicherung
- IX. Sicherstellen der Informationssicherheit und Cyber Security

Im Hauptprozessschritt I wird das Vorgehen zur Initiierung der Datenanalyse und zur Festlegung der Rahmenbedingungen beschrieben. Der daran anschließende Hauptprozessschritt II umfasst die Planung der Datenanalyse. Die Datenanforderung und der Datenerhalt werden im Hauptprozessschritt III erläutert. Die Datenvalidierung ist Gegenstand von Hauptprozessschritt IV.

Die Hauptprozessschritte V und VI gehen mit Beschreibungen zur Aufbereitung und Analyse von Daten auf die Kernelemente der Datenanalyse ein. Die Bewertung der Ergebnisse sowie die Erstellung von Berichten und die Aufbereitung von Datenanalyseergebnissen werden im Hauptprozessschritt VII dargestellt.

Daneben sind im Vorgehensmodell zwei Hauptprozessschritte als übergreifende und durchgängige Tätigkeiten vorgesehen. Dies ist zum einen die Dokumentation der Datenanalyse sowie die Durchführung von Maßnahmen zur Qualitätssicherung, die unter Hauptprozessschritt VIII dargestellt sind. Zum anderen ist dies die Berücksichtigung von Informationssicherheitsanforderungen, die im Hauptprozessschritt IX zu finden ist.

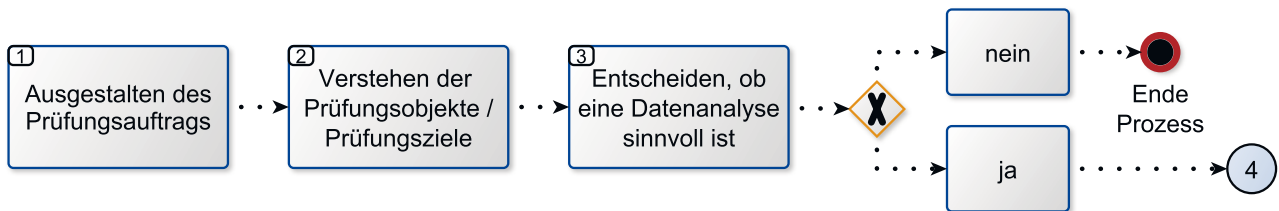
Die hier aufgeführten Hauptprozesse sind grundsätzlich in zeitlich logischer Abfolge dargestellt und jeder Hauptprozessschritt endet mit einem definierten Ergebnis.

Das beschriebene Vorgehensmodell resultiert zum einen aus den Anforderungen der Praxis, zum anderen trägt es aber auch verschiedenen Auftraggeber-Auftragnehmer-Konstellationen Rechnung. Grundsätzlich ist es möglich, das Vorgehensmodell abzuwandeln und den jeweiligen Rahmenbedingungen anzupassen. Je nach Aufgabenstellung bzw. Auftrag kann es unterschiedliche Schwerpunkte geben, die die Betonung oder Auswahl bestimmter Schritte erfordern oder das Weglassen bestimmter Schritte erlauben. Beispielsweise kann im Rahmen von Jahresabschlussprüfungen eine ausführliche Auftragsannahme-Stufe obsolet sein.

In den folgenden Kapiteln werden die einzelnen Prozessschritte detailliert beschrieben. Als Hilfestellung haben wir die potenziell beteiligten Rollen aufgeführt. Die Rollen sind bei der Anwendung dieses Leitfadens auf die jeweilige Situation anzupassen.

I. Entscheiden über Auftragserteilung zur Datenanalyse

I. Entscheiden über Auftragserteilung zur Datenanalyse



Das Ziel dieses Hauptprozessschrittes ist es, eine Entscheidung zu treffen, inwieweit die Datenanalyse als Prüfungstechnik sinnvoll im Rahmen der Prüfung eingesetzt werden kann.

Dieser Hauptprozessschritt führt ausgehend vom allgemeinen Prüfungsauftrag über die Gewinnung eines Verständnisses der Prüfungsobjekte und Prüfungsziele zur Entscheidung über den Einsatz der Prüfungstechnik pro Prüfungsobjekt und Prüfungsziel.

1. Ausgestalten des Prüfungsauftrags

Involvierte Parteien: Auftraggeber, Dateneigentümer, Prüfungsleiter

Die Ausgangslage für eine mögliche Durchführung der Datenanalyse ist der allgemeine Prüfungsauftrag. Der Prüfungsauftrag ist dabei nach ITAF 1001 »ein Prüfungsvorhaben, eine Aufgabe oder Tätigkeit, wie z. B. eine Prüfung, eine Selbstbeurteilung des Kontrollsystems, eine Betrugsermittlung oder ein Beratungsauftrag. Ein Auftrag kann mehrere Aufgaben oder Tätigkeiten zum Erreichen zusammengehöriger Prüfungsziele umfassen«.

Der allgemeine Prüfungsauftrag ist beispielsweise:

- Eine in der Jahresplanung vorgesehene Prüfung der Internen Revision
- Die Jahresabschlussprüfung
- Eine Qualitätssicherungsmaßnahme
- Eine Ad-hoc- bzw. Sonderprüfung
- Eine Prozessprüfung

Der Prüfungsauftrag legt das Prüfungsobjekt und die Prüfungsziele fest. Das Prüfungsobjekt kann ein gesamtes Unternehmen, eine Unternehmenseinheit oder auch ein Geschäftsprozess bzw. ein spezieller Bereich sein. Außerdem ist der relevante Zeitraum (z. B. ein Geschäftsjahr) im Prüfungsauftrag festgelegt.

Prüfungsziele können sowohl aus unternehmensexternen als auch aus -internen Vorgaben resultieren, z. B. aus:

- Regulatorischen oder branchenspezifischen Vorgaben
- Unternehmensinternen Richtlinien
- Informationssicherheitsanforderungen
- Anderen Zielsetzungen, z. B. Erreichen von Wirtschaftlichkeitsvorgaben, Aufdeckung von Betrug bzw. dolosen Handlungen

2. Verstehen der Prüfungsobjekte / Prüfungsziele

Involvierte Parteien: Prüfungsleiter, Datenanalyst, Dateneigentümer, Fachabteilung

Um entscheiden zu können, ob die Prüfungstechnik Datenanalyse für den Prüfungsauftrag sinnvoll ist, benötigt der Prüfer ein Verständnis des Prüfungsobjekts, der Prüfungsziele sowie eine Einschätzung der die Prüfungsziele betreffenden Risiken (vgl. [ITAF 1201]).

Ein wesentlicher Bestandteil ist die Erlangung eines Verständnisses über die Geschäftsprozesse und die eingesetzten IT-Systeme. Das Verständnis kann u.a. durch folgende Prüfungshandlungen erlangt werden:

- ▶ Allgemeine Auseinandersetzung mit dem Prüfungsobjekt (z. B. Sichtung und Analyse zugänglicher Informationen)
- ▶ Auswertung von Prüfungsberichten aus vorherigen Prüfungen
- ▶ Gespräche mit den Fachabteilungen und den Dateneigentümern
- ▶ Gespräche mit den IT-Verantwortlichen
- ▶ Prüfung einer existierenden oder Erstellung einer neuen Prozess- oder Verfahrensdokumentation
- ▶ Identifikation interner und externer Vorgaben

Das Ergebnis dieser Phase ist eine Auflistung aller Kombinationen aus Prüfungsobjekt und Prüfungsziel einschließlich etwaiger Risiken sowie unterstützender IT-Systeme und zugehöriger Datenquellen.

3. Entscheiden, ob eine Datenanalyse sinnvoll ist

Involvierte Parteien: Prüfungsleiter, Datenanalyst

In diesem Prozessschritt muss für alle identifizierten Kombinationen aus Prüfungsobjekt und Prüfungsziel entschieden werden, ob die Prüfungstechnik Datenanalyse eingesetzt werden soll. Dabei sind insbesondere folgende Punkte zu beachten¹:

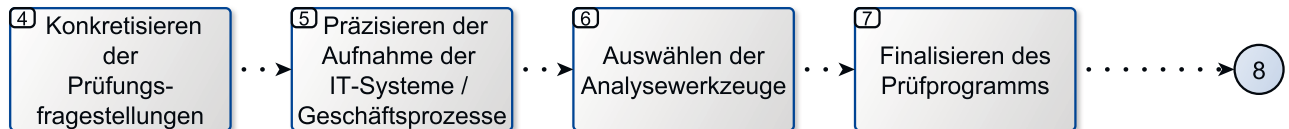
- ▶ Datenverfügbarkeit (auf Basis einer vorläufigen Einschätzung)
- ▶ Vorliegen von Massengeschäftsvorfällen bzw. Massendaten, die unter Effizienz Gesichtspunkten ohne Datenanalyse häufig nicht sinnvoll zu prüfen sind
- ▶ Verfügbarkeit geeigneter Analysewerkzeuge
- ▶ Verfügbarkeit von spezifischen Erfahrungen / spezifischem Fachwissen bzgl. Datenanalyse
- ▶ Verfügbarkeit von Datenanalyse-Ressourcen (z. B. Mitarbeiter / Dienstleister, Auswertungssysteme) im Prüfungszeitraum

Nur wenn sich ein sinnvoller (d. h. effektiver und effizienter) Einsatz der Datenanalyse abzeichnet, sollte diese in der weiteren Prüfungsplanung berücksichtigt werden.

¹ Vgl. [IDW PH 9.330.3, Tz. 16].

II. Planen der Datenanalyse

II. Planen der Datenanalyse



In diesem Hauptprozessschritt erfolgt eine detaillierte Planung der Datenanalyse hinsichtlich der Prüfungsobjekte, der Prüfungsziele sowie der Analysemethoden und Analysewerkzeuge. Das Ziel ist die vollständige Beschreibung der

durchzuführenden Prüfung in einem Prüfprogramm. Dieser Hauptprozessschritt ist geprägt durch einen intensiven Informationsaustausch zwischen den definierten, in den Prüfungsauftrag eingebundenen Parteien.

4. Konkretisieren der Prüfungsfragestellungen

Involvierte Parteien: Prüfungsleiter, Datenanalyst

Der im Prozessschritt (1) genannte Prüfungsauftrag ist zumeist allgemein gehalten. Um die Datenanalyse systematisch, effektiv und effizient durchführen zu können, ist eine Konkretisierung der Prüfungsplanung notwendig (vgl. [ITAF 2201]).

Der erste Schritt zur Konkretisierung ist die Ableitung der Prüfungsfragestellungen aus den identifizierten Kombinationen aus Prüfungsobjekt und Prüfungsziel. Die Ableitung der Prüfungsfragestellungen erfolgt risikoorientiert und ist in dieser Phase noch unabhängig von den zu untersuchenden IT-Systemen. Die Prüfungsfragestellungen werden im Laufe der Prüfung durch die Datenanalysen beantwortet.

Das Ergebnis dieser Phase ist die Auflistung der Prüfungsfragestellungen pro Prüfungsobjekt und Prüfungsziel. Zusätzlich ist eine Auflistung der adressierten Risiken pro Fragestellung hilfreich.

5. Präzisieren der Aufnahme der IT-Systeme / Geschäftsprozesse

Involvierte Parteien: Prüfungsleiter, Datenanalyst, Anwendungseigentümer, Dateneigentümer, Fachabteilung

Im Prozessschritt (2) wurde bereits ein grundlegendes Verständnis über die zu prüfenden Geschäftsprozesse und Prüfungsobjekte erlangt. Die Vertiefung dieses Verständnisses ist notwendig, um die definierten Prüfungsfragen in den Kontext der IT-Systeme und Geschäftsprozesse setzen zu können. Die Kernfragen lauten: Welche Daten zu welchem Geschäftsprozess liegen in welchem IT-System und wie sind die Daten zu interpretieren?

Dazu ist es zunächst hilfreich, die IT-Systeme den Prüfungsfragestellungen zuzuordnen und anschließend die Prüfungsfragestellungen nach den IT-Systemen zu gruppieren. Im Ergebnis sind für jedes IT-System die relevanten Prüfungsfragestellungen definiert.

Für jedes IT-System sollten danach im Hinblick auf die Prüfungsfragestellungen die Datenflüsse sowie die Datenstrukturen dokumentiert werden. Dies beinhaltet auch die dazugehörigen Schnittstellen zu vor- und nachgelagerten IT-Systemen, sofern einzelne Prüfungsfragestellungen mehrere IT-Systeme betreffen.²

Im Rahmen der Präzisierung sollte sich der Datenanalyst mit den zu untersuchenden IT-Systemen, insbesondere mit relevanten Tabellen und Spalten sowie den Beziehungen zwischen den Tabellen, vertraut machen.

Das notwendige Verständnis über die zu prüfenden Geschäftsprozesse und Prüfungsobjekte kann anhand der folgenden Prüfungshandlungen vertieft werden:

- ▶ Durchsicht vorhandener Prozessdokumentationen
- ▶ Gespräche mit den Verantwortlichen der Fachabteilungen und verantwortlichen Anwendungs- und Dateneigentümern des Quellsystems bzw. der Quellsysteme
- ▶ Gegebenenfalls Gespräche mit beteiligten Dienstleistern
- ▶ Analyse der Datenstruktur

Elementarer Bestandteil ist dabei die Erhebung des Datenumfanges sowie die Einschätzung der Daten auf deren inhärenten Risikogehalt, z. B. in Bezug auf den Datenschutz und deren Vertraulichkeit. Auch die Relevanz und Verlässlichkeit der Daten ist in diesem Schritt zu beurteilen. Dabei kann z. B. auf die Ergebnisse einer Prüfung der allgemeinen IT-Kontrollen zurückgegriffen werden.

Das Ergebnis dieser Phase ist ein tiefgreifendes Verständnis über die zu untersuchenden Geschäftsprozesse und die sie unterstützenden IT-Systeme mit ihren Datenstrukturen sowie die Beurteilung, welche Prüfungsfragen durch die Prüfungstechnik Datenanalyse beantwortet werden können.

² Vgl. [Herde et al. 2007, S. 1214].

6. Auswählen der Analysewerkzeuge

Involvierte Parteien: Prüfungsleiter, Datenanalyst, Dateneigentümer, Compliance-Funktion

Das Ziel dieser Phase ist die Auswahl geeigneter Datenabzugs- und Datenanalysewerkzeuge auf Basis der während des Prozessschrittes (5) gewonnenen Kenntnisse sowie der Ableitung der konkreten Prüfungsschritte zur Beantwortung der Prüfungsfragestellungen.

Dazu erstellt der Datenanalyst in Abhängigkeit des Prüfungsobjekts, der Prüfungsfragestellungen, des Datenumfangs sowie des Prüfungsbudgets eine Liste möglicher Datenabzugs- und Datenanalysewerkzeuge.

Anhand der Eigenschaften und Funktionen der Werkzeuge sowie ggf. weiterer Rahmenbedingungen wie Know-how, Verfügbarkeit oder regulatorische Anforderungen werden die für die Aufgabenstellung am besten geeigneten Werkzeuge ausgewählt.

Passend dazu wird das Vorgehen zur Datenbeschaffung definiert.

Regulatorische Anforderungen (u.a. im Anwendungsumfeld geltende Gesetze und Richtlinien wie z. B. DSGVO) sind durch den Prüfer bereits in der Prüfungsplanung zu berücksichtigen (z. B. Einbeziehung des Betriebsrats) und beinhalten auch Aspekte wie Pseudonymisierung oder Anonymisierung von personenbezogenen oder vertraulichen Daten.³

Idealerweise sollten personenbezogene oder vertrauliche Daten bereits während des Datenabzugs vor dem Speichern der Daten verfälscht und / oder verschlüsselt werden.⁴

Das Ergebnis dieser Phase ist die Definition und Bereitstellung der notwendigen Datenabzugs- und Datenanalysewerkzeuge.

7. Finalisieren des Prüfprogramms

Involvierte Parteien: Prüfungsleiter, Datenanalyst, Auftraggeber

Das Ziel dieser Phase ist das finalisierte Prüfprogramm, das durch die zu beteiligenden Parteien genehmigt bzw. zur Kenntnis genommen werden soll.

Die Finalisierung des Prüfprogramms ergänzt die im Prozessschritt (4) vorgenommene Konkretisierung des Auftrags um alle in den Prozessschritten (5) und (6) gewonnenen Erkenntnisse.

Das Prüfprogramm umfasst eine Auflistung von durchzuführenden Prüfungshandlungen, die alle Anforderungen aus dem Auftrag adressiert, die umgesetzt werden können. Gegebenenfalls sollte begründet werden können, wenn eine Umsetzung nicht möglich ist (z. B. aufgrund der verfügbaren Datenstruktur).

Je nach eingesetzten Analysewerkzeugen können häufig folgende Funktionen zur Durchführung der Prüfungshandlungen genutzt werden:

- ▶ Selektion / Filterung / Extraktion
- ▶ Summierung / Verdichtung / Gruppierung
- ▶ Anreicherung um zusätzliche Informationen wie z. B. Berechnung von zusätzlichen (Daten-)Feldern
- ▶ Verbindung / Verknüpfung von Tabellen
- ▶ Splitten von Tabellen

Zur Unterstützung des Datenanalyseprozesses bieten viele Analysewerkzeuge noch weitere Funktionen (z. B. statistische Methoden, mathematische Operationen).

Sollten die vordefinierten Funktionen des Analysewerkzeugs nicht ausreichend sein, ist es dem Datenanalysten i.d.R. möglich, innerhalb des Analysewerkzeugs den Funktionsumfang durch individuelle Programmierungen zu erweitern.

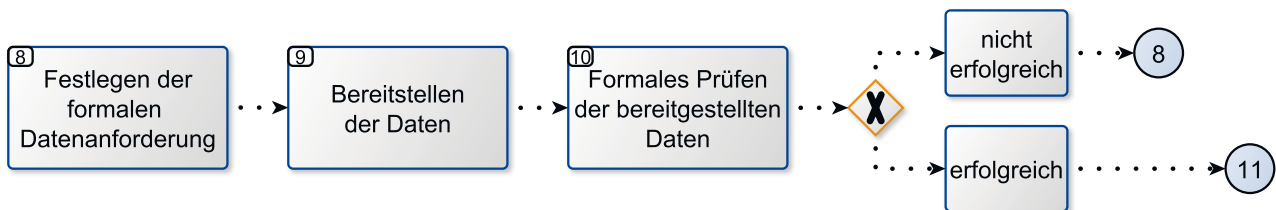
Das Prüfprogramm beinhaltet insbesondere die Erstellung einer Zeitplanung für die Analyse und die Definition von Ansprechpartnern und Verantwortlichkeiten. Des Weiteren werden Art und Umfang der Berichterstattung sowie die Ressourcenplanung festgelegt.

³ Vgl. [Keller 2009, S. 90].

⁴ Vgl. [Geschonneck et al. 2011, S. 2679-2680].

III. Daten anfordern und Daten bereitstellen

III. Daten anfordern und Daten bereitstellen



Dieser Hauptprozessschritt beschreibt die Abläufe zur Bereitstellung der Rohdaten für die Datenanalyse. Ziel ist es, die für das Prüfprogramm benötigten Daten im gewünschten Detaillierungsgrad, Umfang und Format zu definieren und anschließend zu erhalten.

8. Festlegen der formalen Datenanforderung

Involvierte Parteien: Datenanalyst, Anwendungseigentümer, Dateneigentümer

Das Ziel der Festlegung der formalen Datenanforderung ist es, eine detaillierte Beschreibung (z. B. Checkliste, Anforderungsliste) zu erstellen, die zum einen für den Datenbereitsteller eine präzise Vorgabe für die Bereitstellung der Daten darstellt und zum anderen für den Datenempfänger eine Prüfung auf formale Vollständigkeit und Richtigkeit sowie eine Nachvollziehbarkeit der Datenbereitstellung ermöglicht.

Die Anforderung der Daten sollte die Ergebnisse aus Prozessschritt (7) berücksichtigen und eine Liste der zu liefernden Dateien, die Beschreibungen der Datensatzstrukturen, der Datenrelationen und der Datenfelder bzw. Dateninhalte sowie die Selektionskriterien beinhalten.

Die formale Datenanforderung hat drei Bestandteile:

- ▶ Beschreibung der zu analysierenden Daten
- ▶ Beschreibung der Rahmenbedingungen
- ▶ Für die Validierung relevante Informationen

Die Beschreibung der zu analysierenden Daten enthält u.a.:

- ▶ Die Prüfungsobjekte in Form von Tabellen und Datenfeldern
- ▶ Gegebenenfalls zu den Prüfungsobjekten gehörende bzw. in Beziehung stehende, weiter beschreibende Tabellen und Daten
- ▶ Gegebenenfalls einschränkende Kriterien (Filterkriterien), wie beispielsweise:
 - ▶ Unternehmensorganisation
 - ▶ Auswertungszeitraum
 - ▶ Weitere Datenfelder



8. Festlegen der formalen Datenanforderung

Involvierte Parteien: Datenanalyst, Anwendungseigentümer, Dateneigentümer

Zu den Rahmenbedingungen zählen etwaige Vorgaben hinsichtlich:

- ▶ Dateiformaten
- ▶ Datenbanktypen
- ▶ Steuerzeichen, wie z. B. Feldtrennzeichen
- ▶ Übertragungswege
- ▶ Zielplattform bzw. Zielmedium
- ▶ Mechanismen für die Datensicherheit (z. B. Verschlüsselung, Zugriffsbeschränkung)
- ▶ Begrenzung von Dateigrößen (z. B. durch Splittung von Ergebnisdateien)
- ▶ Art und Weise von Datenanonymisierungen / -pseudonymisierungen
- ▶ IT-System-Restriktionen, wie z. B. frequenzstarke und -arme Zeiten bzw. Performance-Aspekte
- ▶ Verfügbarkeit und Gebrauch zeitgesteuerter Datenabrufe
- ▶ Art und Umfang von Dokumentationen:
 - ▶ Protokolle des Datenabzugs (Dokumentation von Auswahl- / Filterkriterien, wie z. B. Informationen zum Datum bzw. Zeitraum des Datenabzugs)
 - ▶ Information, ob die Daten nach dem Datenabzug manuell nachbearbeitet worden sind (z. B. im Rahmen einer Pseudonymisierung von Mitarbeiternamen)
 - ▶ Sonstige für den Datenabzug relevante Informationen
 - ▶ Erklärungen der Datenlieferanten zur Vollständigkeit
- ▶ Umgang mit Besonderheiten in der Datenlieferung (Sonder- und Steuerzeichen, nicht vermeidbare Datenfehler)

Für die Validierung relevante Informationen:

- ▶ Prüfsummen aus dem Quellsystem, wie z. B. Anzahl der Datensätze nach Filterkriterien, CRC-Code
- ▶ Weitere Prüfkriterien / Validierungsmöglichkeiten (z. B. Abstimmsummen, GuV / Bilanz)

9. Bereitstellen der Daten

Involvierte Parteien: Anwendungseigentümer, Datenanalyst, Dateneigentümer

Das Ziel dieses Prozessschrittes ist die vollständige und richtige Datenbereitstellung der Rohdaten entsprechend der Datenanforderung. Diese erfolgt entweder durch den vereinbarten Datenbereitsteller (Anwendungseigentümer) oder durch den Prüfer / Datenanalysten.

Erfolgt die Bereitstellung der Daten durch einen internen oder externen Anwendungseigentümer, kann eine Begleitung durch den Prüfer / Datenanalysten sinnvoll sein. Kann die Datenbereitstellung nicht durch den Prüfer / Datenanalysten begleitet werden, so sollte eine Vollständigkeitserklärung durch den Datenbereitsteller vorliegen.

Wird der Datenabzug durch den Prüfer / Datenanalysten selbst durchgeführt, sollte sichergestellt werden, dass eine ausreichende Expertise im Umgang mit dem Quellsystem vorhanden ist. Im Vorfeld müssen dabei auch Festlegungen über Zugangsberechtigungen zu den Quellsystemen getroffen und nachvollziehbar umgesetzt werden, damit Fehler und Fehlbedienungen (auch Beeinträchtigungen der Systemressourcen) in den Quellsystemen vermieden werden.

Mögliche Vorteile der Datenbeschaffung durch den Prüfer / Datenanalysten⁵:

- ▶ Einfache Sicherstellung der Einhaltung von Vorgaben
- ▶ Gegebenenfalls Verbesserung des Verständnisses über das IT-System und der Datenstrukturen seitens des Prüfers / Datenanalysten
- ▶ Geringe Manipulationsgefahr

Mögliche Vorteile der Bereitstellung durch Anwendungseigentümer:

- ▶ Kein Risiko für den Prüfer / Datenanalysten hinsichtlich des Verschuldens von Systembeeinträchtigungen
- ▶ Schnellere Durchführung der Datenbereitstellung aufgrund tieferer Kenntnisse über das IT-System und die Datenstrukturen seitens des Geprüften
- ▶ Zeitersparnis für den Prüfer / Analysten

Welcher Verfahrensweise der Vorzug gegeben wird, obliegt im pflichtgemäßen Ermessen des Prüfers / Datenanalysten. Dabei sollten auch risikoorientierte Aspekte bedacht werden (Wesentlichkeit der Daten, Manipulationsgefahr etc.).

⁵ Vgl. [Herde et al. 2007, S. 1215].

10. Formales Prüfen der bereitgestellten Daten

Involvierte Parteien: Datenanalyst, Anwendungseigentümer

Um frühzeitig mögliche Fehler bzw. Unvollständigkeiten der Datenlieferung zu erkennen, wird in diesem Prozessschritt eine formale Prüfung der bereitgestellten Daten durchgeführt.⁶

Zunächst werden die bereitgestellten Daten im Originalzustand gesondert gesichert. Die Weiterverarbeitung erfolgt mit Kopien der bereitgestellten Daten.

Diese werden mit den Datenanforderungen aus Prozessschritt (8) abgeglichen. Hierzu gehören die Prüfung der äußerlichen Vollständigkeit der Datenlieferung (z. B. zeitliche Eingrenzung der Daten auf ein Geschäftsjahr, Kundennummern, Buchungskreise) und die korrekte Anwendung weiterer Filterkriterien. Hilfreich hierbei können folgende Prüfungstätigkeiten sein:

- ▶ Dateieigenschaften abgleichen (CRC-Code, Name, Datum etc.)
- ▶ Review des Logs, ggf. Fehlerlog nachverfolgen
- ▶ Review von technischen Datenbankabfragen (SQL-Statements)

Eine Dokumentation der einzelnen Prüfschritte, der Abweichungen sowie der diesbezügliche Umgang ist fakultativer Bestandteil der Prüfungsdokumentation.

Falls die gelieferten Daten nicht der Datenanforderung entsprechen, ist eine Nachbesserung bzw. erneute Lieferung der Daten durch den Datenbereiter erforderlich. Dabei sollten die festgestellten Datenfehler (z. B. falsche Selektionskriterien, falsches Datenformat) konkret benannt werden, um erneute Fehler in den bereitgestellten Daten zu vermeiden.

Im Bedarfsfall müssen Anpassungen der Datenanforderung oder des Prüfprogramms (z. B. bei nicht verfügbaren Daten) vorgenommen werden.

Grundsätzlich sollten die Daten so zur Verfügung gestellt werden, dass eine manuelle Nachbearbeitung der gelieferten Daten durch den Datenanalysten nicht erforderlich ist.

⁶ Vgl. [IDW PH 9.330.3, Tz. 29 u. 32].

IV. Daten übernehmen und Daten validieren

IV. Daten übernehmen und Daten validieren



Das Ziel dieses Hauptprozessschrittes ist es, die Daten in das Analysewerkzeug zu übernehmen und diese zu validieren.

11. Übernehmen der Daten in das Analysewerkzeug

Involvierte Parteien: Datenanalyst

Ergibt die formale Validierung der Daten keine Fehler bzw. werden die Fehler als akzeptiert dokumentiert, können die Daten weiterverarbeitet werden.

Die Daten sollten in einem ersten Schritt vollständig in das Analysewerkzeug übernommen werden. Eine Bereinigung um einzelne Felder sollte erst in einem späteren Schritt innerhalb des Analysewerkzeugs erfolgen.

12. Inhaltliches Prüfen der übernommenen Daten

Involvierte Parteien: Datenanalyst, ggf. Anwendungseigentümer

Nach der Übernahme in das Analysewerkzeug sind die Daten zu validieren.⁷ Das bedeutet, dass zu prüfen ist, ob die Datenanforderung aus Prozessschritt (8) vollständig und richtig umgesetzt wurde. Gegebenenfalls können auch durch den Import noch Datenveränderungen erfolgen, die es zu beachten gilt (z. B. fehlende Führungsnullen). Die Validierung der Daten kann durch eine oder mehrere Prüfungshandlungen erreicht werden. Dazu zählen beispielsweise:

- ▶ Abgleich der Anzahl von Datensätzen
- ▶ Ermittlung von Abstimmsummen
- ▶ Abgleich eines Hashwertes (z. B. CRC-Code)
- ▶ Stichprobenweiser Abgleich einzelner Datensätze zwischen mehreren Datenbeständen
- ▶ Kombination der genannten Maßnahmen

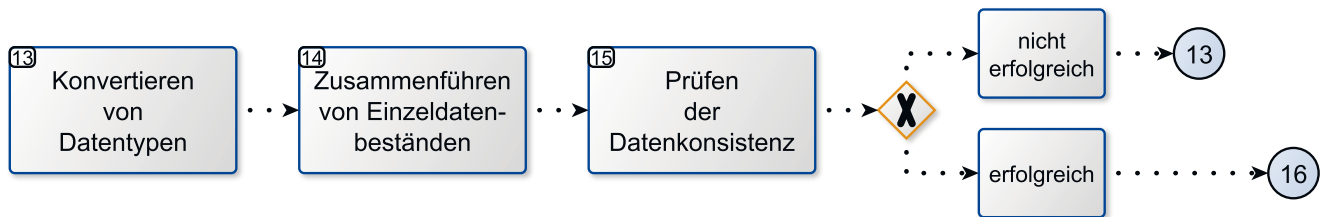
Der Prüfer muss verifizieren, ob die Datenfelder wie gewünscht / erwartet befüllt sind, d. h. plausibilisieren, ob beispielsweise unter der Feldbezeichnung »brutto« auch der Bruttowert enthalten ist. Außerdem ist zu beurteilen, ob die richtige Einschränkung vorgenommen wurde, ob z. B. die erwarteten Vertriebsgebiete selektiert wurden.

Im Falle von fehlenden Daten ist erneut zu prüfen, ob die fehlenden Informationen für die Erreichung des Prüfungsziels notwendig sind und nachgefordert werden müssen.

⁷ Vgl. [Sayana 2003, S. 2].

V. Aufbereiten der Daten

V. Aufbereiten der Daten



Das Ziel dieses Hauptprozessschrittes ist es, die zu analysierenden Daten so aufzubereiten, dass sie für die Durchführung der geplanten Analysen in ihrer Struktur, ihrem Format und in ihren Inhalten geeignet sind. Handelt es sich dabei um mehrfach verwendbare, standardisierte Analysen, sollten

definierte Input-Vorgaben eingehalten werden. Erfolgt keine angemessene Datenaufbereitung, kann es zu fehlerhaften Ergebnissen bei der Datenanalyse oder bei der Interpretation der Ergebnisse kommen.⁸

13. Konvertieren von Datentypen

Involvierte Parteien: Datenanalyst

Nach der Übernahme in das Analysewerkzeug liegen häufig nicht alle Informationen in dem für die Folgeauswertungen benötigten Datentypen (Zeichenkette, Zahlen, Zeit- oder Datumsformate) oder Datenformate vor. Exemplarisch sind hier z. B. Führungsnullen, in Zeichenketten enthaltene Leerzeichen, Dezimaltrennzeichen, unterschiedliche Datumsformate mit ein- oder zweistelligen Tagesziffern, Zahlen- oder Zeichenformate genannt.

In diesem Prozessschritt werden diese Datentypen daher in ein auswertungsfähiges Format konvertiert. Bei einer Konvertierung des Datentyps ist u. U. die Angabe des Datenformats erforderlich (z. B. die Nennung der Anzahl an Dezimalen). Die Konvertierung sollte dabei nach Möglichkeit so erfolgen, dass der ursprüngliche Dateninhalt erhalten bzw. wiederherstellbar bleibt.

Bei der Datenkonvertierung sollte der Datenanalyst beurteilen und dokumentieren, ob mögliche inhaltliche Verfälschungen der Daten eine Auswirkung auf den Prüfungsgegenstand haben.

Die Ziele der Konvertierung sind insbesondere die Vereinheitlichungen der Formate der (verknüpfenden) (Primär-)Schlüselfelder für alle Tabellen sowie die Vereinheitlichungen von Tabellenstrukturen für Tabellen, die gleichartige Inhalte haben (z. B. Lagerbestandslisten unterschiedlicher Lagerstandorte).

⁸ Vgl. [Lieder & Goldshteyn 2013, S. 49].

14. Zusammenführen von Einzeldatenbeständen**Involvierte Parteien: Datenanalyst**

Falls z. B. systembedingt oder aufgrund des Datenvolumens mehrere Datenbestände bereitgestellt wurden, kann eine Zusammenführung dieser Datenbestände erforderlich und sinnvoll sein. Es können hier die Verlängerung von Datenbeständen (z. B. von ursprünglich zwölf Monatsdateien zu einer kompletten Jahresdatei) oder die Zusammenführung von verschiedenen Tabellen (z. B. Kundenadressen mit Kundenumsätzen oder Belegköpfe mit Belegsegmenten) anhand eines Primärschlüssels als Beispiele aufgeführt werden.

Voraussetzung für eine Verlängerung einer Ausgangsdatei (auch als Anhängen von Dateien bezeichnet) ist eine identische Datensatzstruktur, d. h., Datenfeldbezeichnungen (z. B. Kontonummer), Datentypen (z. B. Zeichenkette) und Datenformate (z. B. achtstellige Zeichenkette ohne führende Nullen) sind gleich.

Voraussetzung für eine Zusammenführung (auch als Verknüpfung bezeichnet) ist ein eindeutiger und typgleicher Primärschlüssel.

15. Prüfen der Datenkonsistenz**Involvierte Parteien: Datenanalyst, Dateneigentümer**

Im Rahmen der Konsistenzprüfung werden die Ergebnisse der Konvertierung der Datentypen und der Zusammenführung von Einzeldatenbeständen geprüft.⁹ Im Rahmen der vorstehenden Prozessschritte (13) und (14) muss auch ein etwaiges Fehlerprotokoll der eingesetzten Datenanalyse- bzw. Konvertierungssoftware beachtet und ggf. bearbeitet werden.

Um die korrekte Zusammenführung und Konvertierung zu prüfen, kann sich der Prüfer / Datenanalyst z. B. folgender Fragestellungen bedienen:

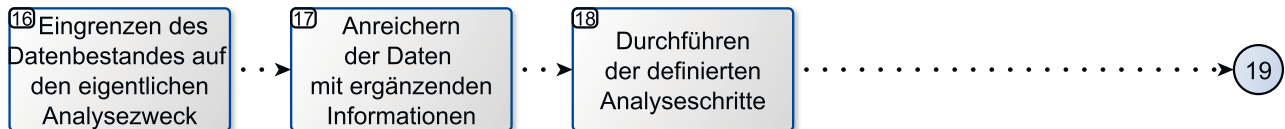
- ▶ Stimmt bei der Verlängerung einer Ausgangsdatei die Gesamtzahl an Datensätzen mit der Summe der einzelnen Einzeldatensätze überein?
- ▶ Gibt es zu jedem Transaktionsdatensatz einen Stammdatenbeleg oder existieren nach einer Verknüpfung leere Datensätze?
- ▶ Existieren fehlerhafte Daten, die vom Analysetool als solche markiert wurden?

Ergeben sich Inkonsistenzen, müssen zunächst die Ursachen ermittelt und abgestellt werden und dann die entsprechenden Prozessschritte wiederholt werden. Hier können sich ggf. bereits Feststellungen zur Ordnungsmäßigkeit und / oder Nachvollziehbarkeit ergeben, wenn die Ursachen für die Inkonsistenzen nicht abgestellt oder erklärt werden können.

9 Vgl. [Töller & Herde 2012, S. 598].

VI. Analysieren der Daten

VI. Analysieren der Daten



Die Ziele dieses Hauptprozessschrittes sind die Durchführung der originären Datenanalyse, die Bewertung der Ergebnisse und, falls erforderlich, eine Abschätzung des über den bisher geplanten Analyseumfang hinausgehenden, weiteren Datenanalysebedarfs.

Fehler bei der Durchführung der Datenanalyse führen in aller Regel zu fehlerhaften (Datenanalyse-)Ergebnissen, die u.U. erst im nachgelagerten Prozessschritt der Bewertung der Ergebnisse identifiziert werden.

16. Eingrenzen des Datenbestandes auf den eigentlichen Analysezzweck

Involvierte Parteien: Datenanalyst

Das Ziel dieses Prozessschrittes ist es, den vorbereiteten Datenbestand auf den eigentlichen Analysezzweck einzugrenzen, indem zum einen nicht benötigte Informationen »ausgeblendet« werden und zum anderen, falls erforderlich, eine Verdichtung erfolgt.¹⁰ An dieser Stelle kann ggf. auch eine Stichprobenziehung erforderlich / sinnvoll sein.

Eine Eingrenzung des Datenbestands kann durch eine Selektion nach bestimmten Kriterien (z. B. die eines Buchungskreises) erfolgen oder die Eingrenzung des Datenbestandes geschieht im Bereich der bei der Datenanalyse nicht verwendeten Datenfelder (z. B. kann aus Datenschutzgründen der Mitarbeitername ausgeschlossen werden).

Einzelne Datensätze (d. h. Tabellenzeilen) können durch Summierung verdichtet werden oder es können Tabellen wiederum nach bestimmten Kriterien gesplittet werden.

Die Ermittlung von Stichproben kann zum einen selbst das Ziel der Analyse sein, zum anderen kann es auch die Vorgabe sein, die weitere Datenanalyse auf Basis von Stichproben durchzuführen.

17. Anreichern der Daten mit ergänzenden Informationen

Involvierte Parteien: Datenanalyst

Das Ziel dieses Prozessschrittes ist die Anreicherung der erhaltenen und vorbereiteten Daten mit weiteren Daten aus anderen, i. d. R. öffentlich zugänglichen Quellen. Beispielsweise können interne Daten zu getätigten Transaktionen mit extern verfügbaren Informationen zu Wochentagen angereichert werden, um zu prüfen, ob Transaktionen an Wochenenden oder Feiertagen ausgeführt wurden.

Liegen Buchungsdaten mit möglichen wetterabhängigen Schwankungen vor, kann eine Anreicherung mit Wetterdaten sinnvoll sein (z. B. Eisdienumsatz an Sonnentagen vs. Museumsumsatz an Regentagen, Überstunden bei Winterdiensten).

Haben z. B. Entfernungen für die zu analysierenden Daten eine Relevanz (wie z. B. berechnete Transportkosten), können diese mit entsprechenden Informationen zu Entfernungen oder Fahrzeiten ergänzt werden.

¹⁰ Vgl. [Fazzone 2014, S. 1033].

18. Durchführen der definierten Analyseschritte**Involvierte Parteien: Datenanalyst**

Voraussetzung für die Durchführung der Datenanalyse ist, dass sich in den vorherigen Prozessschritten keine Indizien ergeben haben, die darauf schließen lassen, dass mit dem vorhandenen Datenmaterial und mit den ausgewählten Werkzeugen die geplanten Analysen nicht bzw. nicht vollständig oder nicht korrekt durchführbar sind.¹¹

Die im finalisierten Prüfprogramm definierten Prüfungshandlungen in Prozessschritt (7) sind mit den festgelegten Analysewerkzeugen in der festgelegten Reihenfolge abzuarbeiten.

Werden im Verlauf der Durchführung Fehler festgestellt, muss ermittelt werden, inwieweit die Fortsetzung der Analyse beeinträchtigt wird. Mögliche Szenarien können sein:

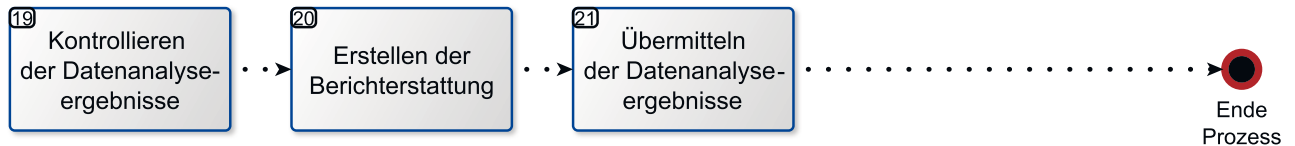
- ▶ Fortsetzung ohne Einschränkungen (hinsichtlich der Prüfungsobjekte bzw. -ziele)
- ▶ Fortsetzung mit Einschränkungen
- ▶ Wiederholungen von Prüfungsschritten mit korrigierten Daten oder Kriterien
- ▶ Änderung von Prüfungsschritten
- ▶ Unterbrechung mit Datennachforderungen
- ▶ Abbruch

Kann die Datenanalyse nicht ohne Einschränkungen hinsichtlich der Prüfungsobjekte bzw. ziele fortgesetzt werden, müssen zunächst die Auswirkungen der Einschränkungen in Bezug auf die Aussagekraft der zu erwartenden Ergebnisse untersucht werden. Darauf basierend muss eine Abstimmung mit den beteiligten Interessengruppen erfolgen.

¹¹ Vgl. [Töller & Herde 2012, S. 599].

VII. Erstellen der Berichterstattung

VII. Bewerten der Ergebnisse und Erstellen der Berichterstattung



Ziel der Berichterstattung ist die Validierung und zielgerichtete Aufbereitung der Ergebnisse der Datenanalyse zur Kommunikation an den Auftraggeber.

Art und Umfang der zu erfüllenden, internen oder externen Anforderungen in Bezug auf die Berichterstattung hängen sehr stark vom Prüfungsobjekt und Prüfungsziel ab.

19. Kontrollieren der Datenanalyseergebnisse

Involvierte Parteien: Prüfungsleiter, Datenanalytst

Die Ergebnisse von Prozessschritt (18) sind in diesem Prozessschritt zuerst einzeln sowie abschließend als Ganzes zu würdigen. Folgende Fragestellungen sollten hierbei beachtet werden:

- ▶ Sind Fehler im Verarbeitungsprozess aufgetreten?
- ▶ Sind die erzielten Ergebnisse grundsätzlich plausibel?
- ▶ Entsprechen die Ergebnisse der Datenanalyse dem erwarteten Analyseergebnis hinsichtlich Vollständigkeit und inhaltlicher Aussagekraft gemäß Prozessschritt (7)?
- ▶ Können die gewünschten Prüfungsaussagen mit der geplanten Sicherheit getroffen und mit den Datenanalysen belegt werden?
- ▶ Welche Abweichungen wurden festgestellt und wie sind diese zu bewerten?
- ▶ Sind die getroffenen Interpretationen als Ganzes haltbar?
- ▶ Sind ggf. weitere Datenanalysen notwendig?

Am Ende dieses Prozessschrittes muss Klarheit über den jeweiligen Zielerreichungsgrad der definierten Prüfungsfragestellungen pro Prüfungsobjekt herrschen.

Eine fehlerhafte Bewertung bzw. Interpretation der (Datenanalyse-)Ergebnisse, z. B. aufgrund mangelnder Erfahrung, führt in der Regel zu falschen Schlussfolgerungen im Hinblick auf die Prüfungsziele. Um dies zu vermeiden, können teilweise Einzelfallstichproben das (Datenanalyse-)Ergebnis verifizieren bzw. falsifizieren.

Bei auftretenden Fehlern oder eingeschränkten Datenanalyseergebnissen muss der Prüfungsleiter die Entscheidung treffen, die Datenanalyse abzubrechen oder unter verschärften Annahmen fortzusetzen. Im Fall der Weiterführung der Prüfung muss der Prüfungsleiter analysieren, welche Auswirkungen die Einschränkung auf das Prüfungsergebnis hat und dies mit dem Auftraggeber abstimmen.

Des Weiteren muss eine Aussage hinsichtlich der abgeschätzten Risiken und der Umsetzung der Prüfungsstrategie getroffen werden.

20. Erstellen der Berichterstattung

Involvierte Parteien: Datenanalyst, Prüfungsleiter

Nach Durchführung der Datenanalyse bedarf es der Aufbereitung der Ergebnisse zur Erstellung der Berichterstattung entsprechend den definierten Vorgaben aus Prozessschritt (7).

Die Ergebnisse der Datenanalyse sollten möglichst (wenn keine Gründe dagegen sprechen) mit dem Geprüften inhaltlich abgestimmt werden, um Fehler oder Fehlinterpretationen zu vermeiden.

Die Berichterstattung sollte sich auf den Prüfungsauftrag beziehen und das Vorgehen der Datenanalyse zusammengefasst beschreiben. Inhaltliche Ergebnisse und die Bewertung durch den Datenanalysten sollten klar und verständlich formuliert werden.

Als Teil bzw. Anlage der Berichterstattung sind häufig Detailanalysen sinnvoll. Dabei sollten Spalten- bzw. Zeilenüberschriften von Tabellen eindeutig aufbereitet werden. Die Konsistenz zwischen Detaildarstellungen und Zusammenfassungen muss sichergestellt werden.

Darüber hinaus können beispielsweise Tabellen grafisch aufbereitet oder Kreuztabellen (Pivot-Tabellen) erzeugt werden. Sofern eine Weiterbearbeitung der Daten (z. B. Formatierung, farbliche Aufbereitung) außerhalb des Analysewerkzeugs (bspw. in MS Excel) notwendig sein sollte, ist darauf zu achten, dass die Daten inhaltlich unverändert bleiben.

Im Ergebnis liegt der Berichtsentwurf vor, der mit dem Prüfungsleiter abgestimmt werden sollte.¹²

21. Übermitteln der Datenanalyseergebnisse

Involvierte Parteien: Prüfungsleiter, Datenanalyst, Auftraggeber

Die Übergabe der Datenanalyseergebnisse kann in diversen Formen erfolgen, beispielsweise als:

- ▶ Bericht
- ▶ Präsentation
- ▶ Handout
- ▶ Management Summary
- ▶ Individuelle Dokumente
- ▶ Gegenstand einer Besprechung
- ▶ Interaktives Dashboard

Im Rahmen der Übergabe / Präsentation sollten insbesondere auch weitere notwendige Prüfungshandlungen (über die Datenanalyse hinaus) mit dem Auftraggeber diskutiert werden (z. B. Einzelfallprüfungen).

Die Weitergabe der Datenanalyseergebnisse an Dritte (z. B. an die Fachabteilung oder die Compliance-Funktion) erfolgt in der Regel durch den Auftraggeber.

Mit der Übergabe / Präsentation ist der Auftrag gegenüber dem Auftraggeber abgeschlossen.

¹² Vgl. [Herde et al. 2007, S. 1218].

VIII. Dokumentieren der Datenanalyse und Durchführen der Qualitätssicherung

Das Ziel dieses Prozessschrittes ist die Sicherstellung der anforderungskonformen Abwicklung der festgelegten Prüfungsschritte und ihrer Dokumentation im Rahmen des Prüfungsauftrags. Als Ergebnis liegt ein nachvollziehbarer Nachweis vor, dass mit den durchgeführten Arbeitsschritten der Datenanalyse die im Prüfungsauftrag beschriebenen Prüfungsziele erreicht wurden (vgl. [ITAF 1203] bzw. [ITAF 2203]).

Bei der Dokumentation und der Qualitätssicherung handelt es sich jeweils um eine Querschnittsfunktion des Datenanalyseprozesses, da entsprechende Aktivitäten in jeder einzelnen Phase eines Datenanalyseprozesses integriert sind bzw. durchgeführt werden müssen. Die nachstehenden Ausführungen beschreiben daher, allgemein und ohne besonderen Bezug zu einem bestimmten Prozessschritt, die an eine Dokumentation bzw. Qualitätssicherung zu stellenden Anforderungen.

22. Erstellen der Dokumentation der Datenanalyse

Involvierte Parteien: Datenanalyst, Prüfungsleiter

Neben der Berichterstattung sind sowohl die Planung, das prozessuale Vorgehen der Datenanalyse sowie die Ergebnisse der Analyse zu dokumentieren (vgl. [ITAF 1203.5]).

Die Dokumentation dient dabei der Gewährleistung der Nachvollziehbarkeit und Nachprüfbarkeit von Prüfungsergebnissen.¹³ Daher ist bei allen Punkten zu dokumentieren, wann und von wem die einzelnen Prüfungsschritte durchgeführt wurden.

Die Dokumentation sollte mindestens die folgenden Punkte umfassen:

- ▶ Hintergrund (Auftraggeber, Prüfung / Projekt)
- ▶ Ziel und Gegenstand der Analyse (ggf. auch eine Abgrenzung)
- ▶ Prüfungsmethode und -vorgehen
- ▶ Datenanforderung
- ▶ Fehlerprotokollierung
- ▶ Prüfungsnachweise (z. B. Protokolle (auch Systemprotokolle))
- ▶ Ergebnisse
- ▶ Gegebenenfalls relevante Schritte zur Ergebnisaufbereitung
- ▶ Maßnahmen zur Qualitätskontrolle / -sicherung

Ergänzend dazu sind an die Dokumentation folgende Anforderungen zu stellen:

- ▶ Die Dokumentation muss so gestaltet sein, dass ein sachverständiger Dritter in angemessener Zeit in der Lage ist, anhand der Dokumentation den Datenanalyseprozess in seiner Abfolge vom Prüfungsobjekt bis hin zum Prüfungsergebnis nachzuvollziehen. Sie stellt die Verknüpfung zwischen dem Prüfungsobjekt und dem Prüfungsbericht her.
- ▶ Anhand der Dokumentation muss es möglich sein, jeden einzelnen Prozessschritt in gleicher Weise zu wiederholen (Wiederholbarkeit).
- ▶ Die Dokumentation muss so gestaltet sein, dass einmal in den Dokumentationsprozess eingebrachte Informationen nicht mehr unterdrückt oder ohne Kenntlichmachung überschrieben, gelöscht, geändert oder verfälscht werden können. Bereits in den Dokumentationsprozess eingebrachte Informationen dürfen nicht ohne Kenntlichmachung durch neue Daten ersetzt werden.

Sofern eine Archivierung bzw. Aufbewahrung des Analyseergebnisses aus gesetzlichen bzw. anderen Vorgaben erfolgt, ist mit der dazugehörigen Dokumentation ebenso zu verfahren.

Die Entscheidung über die Aufbewahrung der Rohdaten ist durch den Prüfer unter Berücksichtigung der jeweiligen regulatorischen Anforderungen zu treffen. Wir empfehlen die Aufbewahrung der Rohdaten bis zum Abschluss der Prüfung bzw. der Umsetzung der Maßnahmen einer möglichen Follow-up-Prüfung.

Ein Entscheidungskriterium bzgl. der Aufbewahrung der Rohdaten kann sein, ob das Quellsystem die dauerhafte Verfügbarkeit der Daten gewährleistet.

¹³ Vgl. [IDW PH 9.330.3, Tz. 78], [IDW PS 460, Tz. 9-12].

23. Durchführen der Qualitätssicherung

Involvierte Parteien: Qualitätssicherer

Die Qualitätssicherung des Analyseprozesses beinhaltet sowohl inhaltliche als auch formale Aspekte. Je wirksamer die Qualitätskontrollen sind, desto höher ist das Vertrauen in die Richtigkeit der erzielten Prüfungsergebnisse und die Angemessenheit der Prüfungsdurchführung. Die Durchführung der Qualitätssicherung ist ebenfalls zu dokumentieren.

Qualitätskontrollen bzw. -sicherungen können auf verschiedene Art und Weise durchgeführt werden:

- ▶ Gegenseitigkeitskontrolle – Kontrolle durch eine andere Person (Vieraugenprinzip)
- ▶ Unabhängige Qualitätskontrolle

Ferner beziehen sich Qualitätskontrollen sowohl auf inhaltliche als auch auf formale Aspekte und beinhalten beispielsweise:

- ▶ Die Umsetzung des Prüfprogramms
- ▶ Die Dokumentation von Abweichungen
- ▶ Die Validierung der Analysedaten und ggf. auch der Analyseergebnisse anhand von Abstimmsummen, Stückzahlen oder Einzeldatenabgleich mit Daten aus dem Quellsystem oder anhand von Belegen, Auswertungen etc.
- ▶ Plausibilitätskontrollen – überschlägige Einschätzungen, ob ein Analyseschritt richtig durchgeführt wurde bzw. ob das Ergebnis erwartungsgemäß ausgefallen und annehmbar ist¹⁴
- ▶ Die Wiederholung des Prozessschrittes in Erwartung eines gleichen Ergebnisses, ggf. auch mit anderen Methoden oder Tools

14 Vgl. [Töller & Herde 2012, S. 599-600].

IX. Sicherstellen der Informationssicherheit und Cyber Security

Das Ziel dieses Hauptprozessschrittes ist es, geeignete Maßnahmen zur Sicherstellung einer angemessenen Informationssicherheit und Cyber Security zu ergreifen. Dabei bezieht sich die Informationssicherheit insbesondere auf den Schutz der verarbeiteten Daten bzw. Informationen des Geprüften sowie der Analyseergebnisse.

Die Cyber Security ist vor dem Hintergrund der Datenanalyse in vernetzten Umgebungen (z. B. Prüfer-Rechner mit Internetanschluss, Nutzung von Cloud-Plattformen) bzw. des

Datenaustausches zu beachten. Aus Sicht eines Angreifers stellt die Umgebung der Datenanalyse ein hoch relevantes Ziel dar, da hier häufig sehr vertrauliche Daten aus verschiedenen Quellen an einem Ort in auswertbarer Form vorzufinden sind.

Als Ergebnis liegt ein nachvollziehbarer Nachweis vor, dass im Rahmen von Datenablage, Datenverkehr, Datensicherung und Datenarchivierung die Schutzziele Vertraulichkeit, Verfügbarkeit und Integrität sichergestellt werden.¹⁶

24. Berücksichtigen von Informationssicherheits- und Cyber-Security-Anforderungen

Involvierte Parteien: Datenanalyst

Zur Erreichung der Schutzziele bedarf es einer Risikoanalyse während der Planung der Datenanalyse, bei der wenigstens die folgenden Fragen adressiert werden sollten:

- ▶ Welche Daten sind schützenswert?
- ▶ Welche Daten sind personenbezogen?
- ▶ Welche Risiken bestehen bei der Datenverarbeitung, -speicherung und -ablage?
- ▶ Welche Risiken bestehen hinsichtlich der eingesetzten Anwendungs- und Systemtechniken?

Je nach Prüfungsauftrags-Rahmenbedingungen, d. h. je nach technischen Bedingungen beim Prüfer / Datenanalysten sowie beim Geprüften, sollten ferner folgende Aspekte berücksichtigt werden:

- ▶ Beim Transfer von Daten und Analyseergebnissen zwischen dem Geprüften und dem Datenanalysten sollte darauf geachtet werden, dass die Verbindungstechniken und die Datenträger bzw. die transferierten Informationen geeignet geschützt sind, z. B. durch den Einsatz von Verschlüsselungstechniken bzw. Passwörtern.
- ▶ Bei der Datenanalyse sollte darauf geachtet werden, dass die Datenhaltung und die Analysetools durch geeignete Mechanismen gegen unbefugte Zugriffe geschützt sind.
- ▶ Bei der anschließenden Archivierung sollte ebenfalls auf den Schutz der Verbindungen und der Datenhaltungsorte geachtet werden.
- ▶ Zudem sollte eine korrekte Regelung hinsichtlich der Aufbewahrungsdauer getroffen werden.

Zu den festgestellten Risiken sind geeignete Maßnahmen zu treffen¹⁵.

¹⁵ Vgl. [BSI & ISACA 2015].

¹⁶ Vgl. [ISACA 2016, S. 14].

3. Fazit und Ausblick

Der Leitfaden wurde von Personen aus der Praxis für Personen aus der Praxis entwickelt. Er adressiert interne und externe Prüfer gleichermaßen und unterscheidet nicht zwischen Unternehmensgröße, Unternehmensbranche, Internationalisierungsgrad oder anderen Unternehmensfaktoren. Bei seiner Erstellung lag der Fokus stets auf Praxisbezug und Pragmatismus. Beschreibungen wurden z.T. generisch gehalten, um eine breite Anwendbarkeit auf die oben genannte Zielgruppe zu ermöglichen. Auch wurde eine tiefgründige, wissenschaftliche Diskussion des Themas Datenanalyse bewusst vermieden.

Dieser Leitfaden bietet Ihnen eine methodische Anleitung für strukturierte Datenanalysen in verschiedenen Umfeldern, d. h. in unterschiedlichen Unternehmensformen, in unterschiedlichen Beziehungsstrukturen (Stakeholder, Durchführer) sowie für unterschiedliche Aufgabenstellungen.

Als Kernthema enthält der Leitfaden die Vorgehensmethodik und den Umgang mit verschiedenen Rahmenbedingungen, um sicherzustellen, dass die Analyseergebnisse am Ende formal und inhaltlich abgesichert, nachvollziehbar und belastbar sind. Die Autoren haben hierbei die beschriebenen Lösungsschritte anhand von praktischen Anwendungsfällen verifiziert. Es wurde aus Gründen der Klarheit und Einfachheit bewusst darauf verzichtet, in dem vorliegenden Leitfaden ausführlich auf Praxisbeispiele einzugehen. Die Autoren verweisen hierfür auf geplante Veröffentlichungen, die sich im Wesentlichen Anwendungsfällen (Use Cases) und praktischen Beispielen widmen werden, die auf der hier beschriebenen Methodik basieren sollen.

Eine Veränderung der Reihenfolge oder Weglassen einzelner Schritte ist möglich und auch beabsichtigt. Je nach Aufgabenstellung sollen die sinnvollen und erforderlichen Schritte durchlaufen werden. Dabei ist eine möglicherweise damit verbundene Veränderung der Ergebnisqualität, wie beispielsweise eine Veränderung der Belastbarkeit des Ergebnisses, in ihrer Wirkung für die jeweilige konkrete Anwendung entsprechend abzuschätzen.

Obwohl bei der Erstellung dieses Leitfadens Sorge hinsichtlich der Richtigkeit gemachter Angaben getragen wurde, mag der Leitfaden nicht für alle Praxissituationen passend erscheinen. So richtet sich der Leitfaden primär an Revisoren und umfasst daher viele Hinweise zur Erfüllung formaler Anforderungen. Gleichwohl mag Datenanalyse per se auch

von »Nicht-Revisionsfunktionen« (z. B. durch eine Fachabteilung) ausgeführt werden. Die dafür notwendigen Erfordernisse mögen im Einzelnen von der hier vorgegebenen Struktur abweichen.

Dieser Leitfaden liefert die Basis für weitere praxisorientierte Diskussionen. Weitere Themen hierfür umfassen z. B. die Entwicklung von Anwendungsbeispielen für die Datenanalyse, die Identifikation und Bewertung erhältlicher Softwarelösungen, die sich für die Durchführung von Datenanalysen eignen, die Erörterung möglicher Maßnahmen zur Erhöhung der Awareness und Akzeptanz von Datenanalysen innerhalb von Unternehmen sowie die Darstellung komplexerer Datenanalyseformen (z. B. Continuous Auditing). Einige dieser Themen befinden sich in aktueller Bearbeitung durch die Fachgruppe Datenanalyse und werden in naher Zukunft in weiteren Publikationen münden.

4. Abkürzungsverzeichnis

Abkürzung	Erläuterung
AO	Abgabenordnung
BSI	Bundesamt für Sicherheit in der Informationstechnik
BW	Business Warehouse
CRC	Cyclic Redundancy Check – Zyklische Redundanzprüfung
DSGVO	Datenschutz-Grundverordnung, gemeint: VERORDNUNG (EU) 2016/679 DES EUROPÄISCHEN PARLAMENTS UND DES RATES: Verordnung des Europäischen Parlaments und des Rates zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung)
ERP	Enterprise Resource Planning – Unternehmens-Ressourcen-Planung
GoBD	Grundsätze zur ordnungsmäßigen Führung und Aufbewahrung von Büchern, Aufzeichnungen und Unterlagen in elektronischer Form sowie zum Datenzugriff
GoBS	Grundsätze ordnungsmäßiger DV-gestützter Buchführungssysteme
HGB	Handelsgesetzbuch
ITAF	Technology Assurance Framework
SQL	Structured Query Language

5. Glossar

Begriff	Erläuterung
Analysetool	Software zur Durchführung von Datenanalysen. Dies kann die datenhaltende Anwendung selbst sein (z. B. ein ERP-System oder ein BW-System), typische Office-Anwendungen, Datenbankanwendungen, spezielle Datenanalyseanwendungen oder sonstige Anwendungen/ Umgebungen, die die Datenanalyse unterstützen.
Anonymisierung	Anonymisierungen verändern personenbezogene Daten derart, dass die Einzelangaben über persönliche oder sachliche Verhältnisse nicht mehr oder nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft einer bestimmten oder bestimmaren natürlichen Person zugeordnet werden können.
Compliance	Compliance steht als Sammelbegriff für die Einhaltung von Regeln (= Regeltreue/-konformität). Die Regeln können dabei unterschiedlicher Natur sein: Gesetze und rechtliche Normen, unternehmenseigene Richtlinien oder auch ethische Kodizes.
Cyber Security	Cyber-Sicherheit befasst sich mit allen Aspekten der Sicherheit in der Informations- und Kommunikationstechnik. Das Aktionsfeld der klassischen IT-Sicherheit wird dabei auf den gesamten Cyber-Raum ausgeweitet. Dieser umfasst sämtliche mit dem Internet und vergleichbaren Netzen verbundene Informationstechnik und schließt darauf basierende Kommunikation, Anwendungen, Geschäftsprozesse und verarbeitete Informationen mit ein.
Integrität	Unter Integrität wird die Vollständigkeit und Richtigkeit von Datenbeständen, aber auch deren Redundanzfreiheit verstanden.
Konsistenz	Unter Konsistenz wird die Korrektheit von Datenbeständen bzw. Datenbanken verstanden. Damit ist die Widerspruchsfreiheit der gespeicherten Informationen zueinander gemeint (Beispiel: Das System-Anlagendatum eines Kunden kann nicht vor seinem Geburtsdatum liegen).
Primärschlüssel	Als Primärschlüssel (lat. primus, der Erste) bezeichnet man in relationalen Datenbanken ein Datenfeld (oder eine Kombination von Datenfeldern) in einer Tabelle, die jeden in dieser Tabelle gespeicherten Datensatz eindeutig beschreiben.
Prozessdokumentation	Eine Prozessdokumentation ist die grafische oder textliche Beschreibung von Geschäftsprozessen im Rahmen der Geschäftsprozessmodellierung.
Pseudonymisierung	Pseudonymisierung beschreibt das systematische Ersetzen des Namens und anderer Identifikationsmerkmale durch ein Kennzeichen zu dem Zweck, die Bestimmung des Betroffenen auszuschließen oder wesentlich zu erschweren. Die Pseudonymisierung ist durch Verwendung der genutzten Systematik (Schlüssel) umkehrbar.
Quellsystem	Ein Quellsystem bezieht sich auf eine Datenquelle, aus der Daten abgefragt werden sollen.
Rohdaten	Rohdaten entstehen bei der Datenextraktion vom Quellsystem und bilden die Grundlage der Datenanalyse. Sie umfassen die für das Prüfprogramm benötigten Daten im gewünschten Detaillierungsgrad, Umfang und Format.
Transaktion	Eine Transaktion beschreibt eine Folge von Programmschritten, die nur komplett oder gar nicht durchgeführt wird.



Begriff	Erläuterung
Validierung	<p>Als Validierung wird die Tätigkeit umschrieben, Daten oder Datenbestände auf Vollständigkeit zu prüfen. Dies kann durch</p> <ul style="list-style-type: none">▶ Abgleich der Anzahl von Datensätzen,▶ durch Ermittlung von Abstimmsummen,▶ durch Abgleich eines Hashwertes (z. B. CRC-Code),▶ durch stichprobenweisen Abgleich einzelner Datensätze zwischen mehreren Datenbeständen▶ oder einer Kombination dieser Maßnahmen <p>erfolgen.</p>
Verfahrensdokumentation	<p>Die Verfahrensdokumentation nach GoBD (bis 2015 »GoBS«) dient dazu, nachweisen zu können, dass die Anforderungen des Handelsgesetzbuches (HGB), der Abgabenordnung (AO), der GoBS und der GoBD für die Aufbewahrung von Daten und Belegen erfüllt sind.</p>

6. Literaturverzeichnis

[BSI & ISACA 2014] BSI; ISACA (Hrsg.): Leitfaden Cyber-Sicherheits-Check (Stand: März 2015).

[Fazzone 2014] *Fazzone, Renato*: E-Discovery: Unternehmensinterne Analyse unstrukturierter Daten zur Aufklärung von Kartellverstößen – ein Praxisbericht. BB 2014, S. 1032-1034.

[Geschonneck et al. 2011] *Geschonneck, Alexander; Meyer, Jörg; Scheben, Barbara*: Anonymisierung im Rahmen der forensischen Datenanalyse. BB 2011, S. 2677-2680.

[Herde et al. 2017] *Herde, Georg; Diewald, Bastian; Reiter, Robert*: Digitale Datenanalyse für Revisoren, Wirtschaftsprüfer und Steuerberater. BBK 2007, S. 1211-1120.

[IDW PH 9.330.3] IDW PH 9.330.3 – Einsatz von Datenanalysen im Rahmen der Abschlussprüfung (Stand: Oktober 2010).

[IDW PS 460] IDW PS 460 n.F. – Arbeitspapiere des Abschlussprüfers (Stand: August 2015).

[ISACA 2016] ISACA (Hrsg.): ISACA-Leitfaden Grundlagen der IT-Revision für den Einstieg in die Praxis (Stand: Juli 2016).

[ITAF 1001] ISACA ITAF IT-Prüfungsstandard 1001 – Audit Charter

[ITAF 1201] ISACA ITAF IT-Prüfungsstandard 1201 – Auftragsplanung

[ITAF 1203] ISACA ITAF IT-Prüfungsstandard 1203 – Durchführung und Überwachung

[ITAF 1203.5] ISACA ITAF IT-Prüfungsstandard 1203.5 – Durchführung und Überwachung

[ITAF 2201] ISACA ITAF Ausführungsleitlinie 2201 – Auftragsplanung

[ITAF 2203] ISACA ITAF Ausführungsleitlinie 2203 – Durchführung und Überwachung

[Keller 2009] *Keller, Thomas*: Digitale Analyse von Finanzdaten. In: Digitale Datenanalyse, interne Revision und Wirtschaftsprüfung. Hrsg. Deggendorfer Forum zur digitalen Datenanalyse e.V., Berlin 2009, S. 87-102.

[Lieder & Goldshteyn 2013] *Lieder, Henning; Goldshteyn, Michael*: Effizienzsteigerung der Abschlussprüfung durch Datenanalyse. WPg, 2013, S. 46-55.

[Sayana 2003] *Sayana, Anantha*: Using CAATs to Support IS Audit. ISACA Journal, 2003.

[Töller & Herde 2012] *Töller, Ernst-Rudolf; Herde, Georg*: Einsatz von Analysesoftware in der Prüfung. WPg 2012, S. 598-604.