# Whitepaper Datenanalyse

1. **Einleitung**

Das Jahr 2020 wird uns sicherlich im Gedächtnis bleiben. Nicht nur aufgrund der direkten Gefahr der Pandemie und der resultierten Toten, sondern auch durch die profunden Einschnitte in unsere Freiheit zu leben und zu arbeiten. Es war das Jahr in dem die Themen Digitalisierung, Home-Office und Remote Prüfungen auf die höchste Stufe der Agenda gesetzt wurden. Viele lang gehegte Ressentiments und Bedenken wurden durch die entstehenden Sachzwänge überrollt. Gestiegene Anforderungen, Reiseverbote und eine bei vielen Firmen verschlechtertet wirtschaftliche Situation, haben eine einzigartige Arbeitsatmosphäre geschaffen.

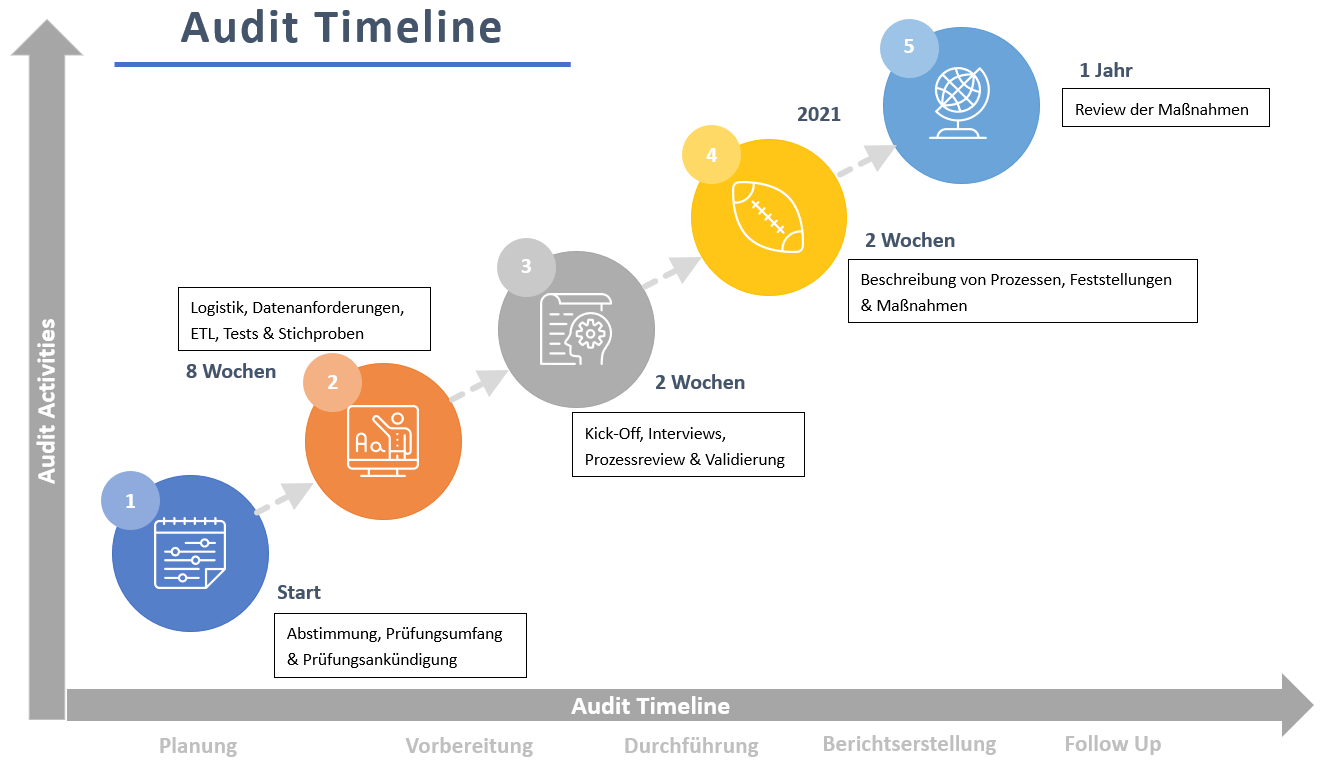
Durch das Aufkommen von Covid und die damit verbundenen Reisebeschränkungen, in Kombination mit verpflichtenden Prüfungen aus regulatorischen, rechtlichen und wirtschaftlichen Gründen werden Revisionsabteilungen gezwungen nach alternativen zu traditionellen Prüfungen zu suchen.

Firmen, welche bisher kaum oder wenig in eine digitale Infrastruktur und entsprechendes Know-how investiert haben, stehen nun vor der Herausforderung diese Rückstände in Rekordzeit aufzuholen. Es müssen nicht nur die entsprechende Software ausgewählt, entsprechende Trainings durchgeführt, sondern ganze Prüfungsprogramme und Prozesse digitalisiert werden.

Zu traditionellen Datenquellen wie ERP Systemen kommen zunehmend alternative Quellen wie dem IOT und Cloudsystemen. Dies tragen zu einer Vervielfältigung der zur Verfügung stehenden Informationen bei. Dieser Mengen von traditionellen „strukturierten Daten“ benötigen spezifizierte Datenanalysetools und selbst mit diesen stellt „Big Data“ eine gewaltige Herausforderung dar.

Zu diesen Problemen gesellen sich bekannte digitale Risikofaktoren und Regularien, wie die Datensicherheit und der DGSVO, welche Revisionsabteilungen weitere Hürden in den Weg legen. Die online Übertragung von sensiblen Informationen zwischen Revision und geprüfter Einheit birgt ein nicht zu unterschätzendes Risiko.

Dieses Whitepaper analysiert die Auswirkungen der Pandemie auf den Prüfungsprozess und beleuchtet hier im Besonderen die Datenanalyse, mit dem Ziel Risiken und Chancen aufzuzeigen. Es soll versuchen, eine effektive Blaupause für den geänderten Prüfungsprozess aufzuzeigen. Zuerst wird der Remote Prüfungsprozess exemplarisch beleuchtet und der Ablauf einer Prüfung durchgespielt. Anschließend werden die Infrastruktur und die Anforderungen der Datenanalyse anhand von Audicon Produkten aufgezeigt. Zuletzt wird gibt es einen Ausblick auf aufkommende Technologien und ein Fazit.

1. **Der Remote Prüfungsprozess**

**Planung und Vorbereitung**

**Abstimmung mit allen Interessensgruppen**

Zum Beginn jeder Prüfung ist zu klären, welche Parteien in den Revisionsprozess involviert sind. Wie bei einem regulären Audit handelt es sich hier primär um das Revisionsteam und die geprüfte Einheit. Die Einbindung weiterer Stakeholder ist bei jeder Firma verschieden:

* Audit Committee: Review
* Geschäftsführung: Berichtsempfänger
* Management des Auditees: Prüfungsziel
* IT Abteilung: Software & Infrastruktur
* Controlling, Risikomanagement & Compliance (2nd Life of Defence)
* …

Der Prüfungsleiter sollte sich klar machen, wer die finalen Adressaten der Revisionsarbeit sind. Das Top Management ist mehr an zusammengefasten Ergebnissen und Bewertungen interessiert, während die 2nd Line of Defence und die geprüfte Einheit möglichst detaillierten Beschreibungen der Sachverhalte benötigen.

Definition of Done: Ansprechpartner und Interessengruppen sind klar definiert.

**Klärung des Scopes und der primären Risiken**

Während kleinere Organisationen ihre gesamte Risikomatrix auf eine betriebliche Einheit anwenden können, so würde dies im Regelfall den Rahmen einer Prüfung sprengen. Prüfungen werden deshalb häufig in organisatorische Elemente (Beispiel: Finanzbuchhaltung) oder Prozesse aufgeteilt (Beispiel: Einkauf). Vor Beginn der Prüfung muss geklärt werden, welche Teilgebiete in die Prüfung aufgenommen werden. Eine Überlastung der geprüften Einheit sollte vermieden werden. Bestimmte Abteilungen, wie die Finanzbuchhaltung sind häufig in einer Vielzahl von Prozessen involviert. Nach der Klärung des Prüfungsumfanges, können vorhandene und anwendbare Sollvorgabe und Richtlinien gesichtet werden.

Die Universität Duisburg-Essen führte zusammen mit dem DIIR eine Umfrage zum Thema Remote-Audit durch. Die Umfrage kam zu dem Ergebnis, dass sich für Remote-Audits vor allem Finanz- und IT-Audits eignen, während Beratungsaktivitäten und Fraud-Untersuchungen durch fehlende Anwesenheit besonders erschwert werden. Konsequenterweise sollten Revisionsabteilungen sich auf Prüfgebiete fokussieren, welche ihre Effektivität bei einem Remote-Ansatz maximieren.

Definition of Done: Eine Liste mit Netto-Risiken wurde erstellt.

**Prüfungsankündigung**

Sobald der Scope der Prüfung intern abgestimmt wurde, ist es notwendig diesen an das Management des Auditees zu kommunizieren. Je schneller dies passiert, desto eher kann mit der eigentlichen Prüfungsplanung begonnen werden. Dies umfasst auch eine klare Definition für die geplante Dauer der Prüfung und der unterschiedlichen Teilbereiche.

Ein Control-Self-Assesment in Form von Fragebögen kann helfen, um Risiken und Kontrollen der geprüften Einheit zu identifizieren und zu dokumentieren. Damit können die Schwerpunkte der Prüfung weiter spezifiziert werden.

Für Remoteprüfungen ist eine klare Kommunikation des neuen (digitalen) Prozesses unerlässlich. Selbst wenn bereits Erfahrungen mit der internen Revision gemacht wurden, ist es höchst unwahrscheinlich, dass alle Mitarbeiter Interviews über ein Videokonferenz-Tool wie Skype durchgeführt haben. Nach der Klärung des Scopes kann entschieden werden, über welche Infrastruktur die gemeinsame Arbeit durchgeführt werden soll.

Definition of Done: Die Planung der Prüfung wurde klar kommuniziert.

**Logistik, gemeinsame Tools, Prozesse und Sicherheitsanforderungen**

Logistische Anforderungen umfassen Equipment, Software und Policies. Equipment beschreibt Grundsätzliches wie Laptops, Drucker, Scanner, Headsets (Telefon) und Internetverbindung mit ausreichender Bandbreite. Obwohl diese Punkte trivial erscheinen mögen, kann die Zusammenarbeit schnell an schlechter Soundqualität oder abreisenden Verbindungen scheitern. Die genutzte Software umfassen Data Analytics (Beispiel: IDEA oder Excel), File-Sharing und direkte Kommunikation, während Policies das Audit Manual und interne Richtlinien miteinschließen.

File-Sharing nimmt bei einer Remoteprüfung einen besonderen Stellenwert ein. Sensible Daten werden über interne und externe Netze versendet und sind aufgrund dessen gewissen Risiken ausgesetzt. Von einem Austausch per E-Mail ist daher abzusehen. Sollte noch kein File-Sharing Tool zur Verfügung stehen, so gibt es am freien Markt professionelle Varianten von Dropbox, Slack oder Google Cloud, welche selbstverständlich unter den Gesichtspunkten des Datenschutzes und der Datensicherheit zu betrachten sind. Zuletzt bieten Video-Chats bieten die Möglichkeit, fehlende Nähe zumindest in Teilen zu überbrücken, da der Gesprächspartner und seine Reaktionen zumindest sichtbar gemacht werden.

Definition of Done: Es existiert eine funktionierende Infrastruktur für die gemeinsame Zusammenarbeit.

**Definition notwendiger Daten (Prozesse, Prüfgebiete)**

Nach Klärung der eingebundenen Prüfgebiete ergibt sich die Frage, welche Daten die benötigten Prozesse abbilden. Abhängig vom Typ des eingesetzten ERP- oder IT-Quellsystems kann der Zeitaufwand erheblich abweichen. Bei populären und standardisierten Lösungen, wie SAP können vorgefertigte Datenmodelle verwendet werden. Ältere Systeme oder Eigenentwicklungen erfordern ein hohes Maß an Einarbeitung.

Sobald bekannt ist, welche Datenquellen benötigt werden, ergibt sich zwangsweise die Frage nach der Beschaffung und dem Datenzugriff. Eine Möglichkeit wäre die Bereitstellung der Daten durch die geprüfte Einheit. Auf der einen Seite hat dies den Vorteil, dass die Prüfer selbst keinen Zugang zu dem geprüften System benötigen. Auf der anderen Seite hat die lokale Einheit häufig nicht die technische Expertise und Infrastruktur, um die Daten zeitnah zu beschaffen. Zudem ergibt sich das Risiko der mangelnden Datenintegrität und manueller Beeinflussung.

Ein Abzug durch die Prüfer erfordert Benutzeraccounts mit ausreichenden Berechtigungen und eine direkte Verbindung zum ERP-System. Dies hat den Nachteil, dass die Prüfer, die fremden Prozesse in eigene Datenanforderungen übersetzen müssen. Zudem benötigt dies einen gewissen Vorlauf durch die geprüfte Einheit. Die Vorteile sind, dass fehlende Daten schnell ergänzt werden können und dass die Prüfer die Integrität und Vollständigkeit der Daten sicherstellen können.

Definition of Done: Alle benötigten Daten sind bekannt und verfügbar.

**Validierung, Einschränkung und Auswertung der Daten (Data Mining)**

Der letzte Schritt der Vorbereitung ist der ETL-Prozess. Er umfasst den eigentlichen Abzug der Daten (Extract), Anpassung von Datenwerten, Umrechnungen und Eliminierung von Duplikaten (Transform) sowie die Bereitstellung in der gewählten Analysesoftware (Load).

Hilfreich ist zudem eine systematische Überprüfung der Daten auf die gewünschten Organisationseinheiten (Beispiel: Buchungskreis) und den Prüfungszeitraum. Eine gemeinsame Inspektion von Besonderheiten bei Prozessen (Beispiel: Belegarten) kann einer späteren Mehrarbeit vorbeugen und unterstützt das Verständnis der Daten.

Definition of Done: Die vollständige Datenmenge steht zur Analyse bereit.

**Durchführung von Tests und Selektion von Stichproben**

Nach der Eingrenzung der Datenmenge sollten die originalen Daten unbearbeitet gesichert werden, um deren Integrität sicherzustellen. Anschließend erfolgt eine Zuordnung der Daten zu den jeweiligen Tests und eine Aufteilung in verschiedene Kategorien. Templates, Script-Repositories oder vorgefertigte Arbeitspakete können helfen, um den Prozess des Testens zu automatisieren. Besonders bei Standardvorgängen oder wiederkehrenden Fragestellungen kann es sehr effizient sein, einmal durchgeführte Prüfungen aufzuzeichnen. Dies führt zu einem Repertoire an Standardtests, welche vor Beginn jeder Prüfung angewendet werden können.

Je nach Größe und Risiko des jeweiligen Tests kann es notwendig sein, die Ergebnisse vollständig zu validieren. Da dies bei tausenden Datensätzen jedoch unrealistisch ist, müssen zwangsweise Stichproben zur Hilfe genommen werden. Eine mögliche Struktur wäre wie folgt: A) Vollständige Klärung (Beispiel: Zahlungen ohne Verbindlichkeiten); B) Stichprobe zur Überprüfung (Beispiel: potenzielle Doppelzahlung); C) Grobe Durchsicht zur Validierung (Beispiel: Maverick Buying).

Die Auswertungen und Stichproben müssen den Fachabteilungen als Arbeitspakete zugeschickt werden. Einerseits muss sichergestellt werden, dass einzelne Abteilungen nicht überlastet werden (Beispiel: Finanzbuchhaltung). Andererseits ist eine qualitativ hochwertige Dokumentation unerlässlich. Dies schließt eine genaue Beschreibung des Sachverhalts und der Fragestellung der Revisoren ein. Analysierten Daten aus den ERP-Systemen sind für Sachbearbeiter häufig unverständlich. Das sich hinter dem Feld „KUNNR“ die Referenznummer eines Kunden verbirgt, ist für einen Sachbearbeiter, welcher nur die direkten Benutzeroberflächen kennt, nicht unbedingt klar. Aufgrund dessen sollten die Unterlagen klar dokumentiert und auf das Nötigste reduziert sein.

Die Aufteilung der Arbeit kann nach unterschiedlichen Methoden erfolgen. In reiferen Revisionsabteilungen mit einem höheren Level an Know-how können die Aufgaben analog zu den Schwerpunkten im Team verteilt werden. Sollte es jedoch nur einen oder wenigen Experten für die Datenanalyse geben, so sollten Arbeitspakete als eine Art interner Service erarbeitet und an die verschiedenen Teilbereiche übergeben werden.

Eine Audit-Management Lösung unterstützt die Revision bei der Dokumentation, da individuelle Prüfgebiete zugeordnet, bearbeitet und mit erhaltenen Ergebnissen versehen werden können. So kann der Prüfungsleiter den Überblick über den Fortschritt der Prüfung behalten und es entsteht zudem ein Audit-Trail aus den dokumentierten Ergebnissen.

Definition of Done: Zur Vorbereitung benötigte Tests und Stichproben wurden erstellt.

**Prüfungsdurchführung**

**Kick Off**

Das Kick-Off Meeting dient der Vorstellung aller beteiligten Personen. Hier trifft das Revisionsteam auf das Management der geprüften Einheit und kann sich mit den Zuständigen der geprüften Teilbereiche vertraut machen. Es hebt die Kommunikation von einer abstrakten Ebene auf ein persönliches Niveau. Zudem bietet es den Prüfern noch einmal die Chance, ihre Methodologie darzulegen und alle Fragen der geprüften Einheit zu beantworten. Es sollte zudem genutzt werden, um den speziellen Charakter der Remote-Prüfung hervorzuheben und um sicherzustellen, dass die Prüfungsvorbereitung zur allgemeinen Zufriedenheit abgeschlossen wurde.

Generell erfolgt das Kick-Off nach dem Abschluss der Vorbereitung und der Analyse der Daten. Standardisierte Analysen wurden durchgeführt und Stichproben sind vorbereitet. Dies bedeutet nicht, dass es zu keinen weiteren Analysen kommt. Durch das Fortschreiten der Prüfung und mit erkannten Themen und Risiken wird es notwendig neue „Ad-Hoc“ Analysen durchzuführen.

Definition of Done: Allen Beteiligten ist der Zeitplan und die Vorgehensweise klar.

**Planung von Interviews**

Durch den unflexiblen Charakter der Remote-Prüfung ist ein klarer Zeitplan unerlässlich. So früh wie möglich sollten die Ansprechpartner für die einzelnen Prüfgebiete geklärt und deren Verfügbarkeit sichergestellt werden.

Für jeden Termin sollten eine klare Agenda und Erwartungen kommuniziert werden. Um den Verlust von persönlicher Nähe und nonverbaler Kommunikation auszugleichen, sollte versucht werden die Meetings mit aktiver Kamera zu führen. So ergibt sich zumindest ein Mindestmaß an sichtbarer Körpersprache. Da es bei einer Remote-Prüfung schwieriger ist offene Fragen in spontanen Anschlussgesprächen zu klären, sollten Online-Termine präziser vorbereitet werden als klassische Interviews.

Termine sollten pünktlich begonnen werden und zeitnah beendet werden. Dies umfasst auch einen vertrauten Umgang mit der genutzten Software. Da die Konzentration bei Online-Meetings schnell nachlassen kann, ist es nicht ratsam die Termine mehr als eine Stunde ohne Pause anzusetzen. Zudem sollte die Teilnehmerzahl reduziert bleiben, um die Meetings effizient zu halten und um den Abstand zwischen den Teilnehmern zu gewährleisten.

Definition of Done: Interviews zu allen wichtigen Themengebieten und Personen wurden vereinbart.

**Review der Prozesse**

Um alle Risiken eines Prozesses zu überprüfen, sollte der Prozess in einzelne Schritte aufgeteilt werden. So können alle individuellen Risiken und Kontrollen für die Prozessschritte überprüft werden. Diese lassen sich im Einkauf grob in die Kreditorenauswahl, die Anlage der Stammdaten im ERP-System, die Anlage und Freigabe der Bestellung, Waren- und Rechnungseingang und die Zahlung aufteilen. Für jeden dieser Schritte sollten bestehende Risiken und Kontrollen dokumentiert werden.

Der gesamte Prozessfluss sollte anhand von Beispielen durchgespielt werden. Durch die fehlende Möglichkeit „über die Schulter zu schauen“, können viele Kontrollen nicht direkt nachvollzogen werden. Sämtliche unterstützenden Dokumente müssen digital zur Verfügung gestellt werden. Revisoren sollten ein zusätzliches Maß an Detailtiefe und Dokumentation anwenden, da offene Fragen nicht in spontanen Terminen geklärt werden können.

Wie schon bei der Vorbereitung, sollten einzelne Mitglieder des Revisionsteams Teile des Gesamtprozesses übernehmen, um sich mit den jeweiligen Risiken, Analysen und Ansprechpartnern vertraut zu machen. Im Team sollte ein regelmäßiger Austausch geführt werden, um die Verzahnung der Prozessschritte zu verstehen und um einen Tunnelblick auf die eigenen Vorgänge zu vermeiden.

Definition of Done: Alle Teilschritte der geprüften Prozesse wurden aufgenommen.

**Validierung von Prozessschwächen durch Tests und Stichproben**

Um finale Feststellungen und Empfehlungen zu finden, müssen Verdachte erhärtet und ermittelte Ergebnisse validiert werden. Dies geschieht in Form von Arbeitspaketen und Stichproben, welche der geprüften Einheit zur Verfügung gestellt werden, mit der Aufgabe diese zu validieren oder zu erklären.

Da angeforderte Dokumente zusammengetragen und Unterlagen in Papierform erst digitalisiert werden müssen, ist hier die Arbeitsbelastung des Auditees am höchsten. Deshalb sind Fehler soweit möglich auszuschließen und es sollten flexible Lösungen für die Dokumentation angestrebt werden.

Eine Stichprobe von 10 potenziellen Doppelzahlungen, welche sich im Nachhinein größtenteils als Stornos oder interne Umbuchungen entpuppen, erzeugt durch eine Neuziehung nicht nur viel Mehrarbeit bei den Revisoren und der geprüften Einheit, sondern sorgt auch für viel Unmut bei letzterem. Diesem Risiko kann durch eine detailgetreue und umfangreiche Vorbereitung begegnet werden.

In dieser Phase der Prüfung ergeben sich automatisch neue Fragestellungen, welche aus dem Prozessverständnis und den bereits durchgeführten Stichproben entstehen. In einem agilen und risikoorientierten Prüfungsansatz sollten diese Fragestellungen nun durch Ad-Hoc Datenanalysen vertieft und beantwortet werden. Sollte sich Beispielsweise während einer Prüfung zeigen, dass hohe Risiken im Exportgeschäft entstehen, so können zusätzliche Analysen in diesem Bereich die bestehenden Kontrollschwächen aufzeigen.

Dieser Prüfungsschritt sollte damit enden, belastbare Zahlen für den finalen Bericht zu liefern. Dies sollte nicht nur durch eine Validierung der Ergebnisse durch den Revisor erreicht werden, sondern auch durch die Einsicht des Auditees und ein gemeinsames Verständnis der Daten. Ergebnisse und Fortschritt der Prüfung sollten in wöchentlichen Update-Meetings mitgeteilt werden. Der Abschluss der Prüfungshandlungen sollte ein Closing-Meeting darstellen, in welchem alle vorläufigen Ergebnisse der Revisionsarbeit grob skizziert werden und welches das Ende der Prüfungshandlungen festlegt.

Definition of Done: Feststellungen und Kontrollschwächen wurden erfasst.

**Berichtserstellung**

**Beschreibung der Prozesse und Feststellungen**

Da die Berichtserstellung selten vor Ort durchgeführt wird, halten sich hier die Unterschiede zur Vor-Ort-Prüfung in Grenzen. Der Bericht sollte eine Beschreibung des neuen Ansatzes und dessen Auswirkung auf den Scope der Prüfung enthalten. Mehrarbeit und Schwierigkeiten des Auditees und der Revisoren durch die Remoteprüfung sollten erwähnt werden, um den Eindruck zu vermeiden, es handele sich um eine reguläre Prüfung per Telefon.

Die Struktur des Berichts und der Prüfung sollten einander entsprechen. Es handelt sich also um eine Ausformulierung ermittelter und validierter Risiken. Der Bericht beschreibt die geprüften Prozesse und zeigt für jeden Prozessschritt die jeweiligen Feststellungen, Risiken und Empfehlungen.

Jede Feststellung sollte im Bericht zusammen mit allen quantitativen Auswertungen und ihren Ursachen aufgeführt sein. Ursachen („Root-Cause“) sind bei einer Prüfung aus der Entfernung häufig noch schwieriger zu erkennen. Hier kann eine stärkere Einbeziehung des Auditees Abhilfe schaffen.

Bevor der Bericht finalisiert werden kann, muss jede Feststellung mit dem Auditee validiert werden, vor allem da Videokonferenzen leichter zu Missverständnissen führen können. Jegliches erhaltene Feedback sollte eingearbeitet oder zumindest zur Notiz genommen werden.

Definition of Done: Alle Feststellungen wurden ausformuliert und akzeptiert.

**Gemeinsames Festlegen von Maßnahmen**

Die gemeinsame Erstellung von Maßnahmen nimmt traditionell einen kleineren Teil der Revisionsprüfung ein, stellt aber letztendlich das signifikanteste Ergebnis dar. Eine Prüfung ohne resultierende Maßnahmen ist wie die Diagnose einer Krankheit ohne die Verschreibung eines Heilmittels: informativ, aber zwecklos. Dies gilt für Remote Prüfungen genau wie für klassisches Field-Work.

Identifizierte Risiken und Feststellungen müssen durch eine entsprechende Maßnahme entschärft werden. Dies kann durch die Einführung neuer Kontrollen, dem Verbot von Prozessvarianten oder der Verbesserung identifizierter Missstände erfolgen.

Da die Maßnahmen einerseits von der geprüften Einheit ausgeführt und anderseits von der Revision, während des Follow-Ups überprüft werden müssen, bilden sie häufig das größte Konfliktpotential. Maximalforderungen aus der Sicht der Revision treffen auf die Realitäten des Geschäfts und die Ressourcen der geprüften Einheit. Aufgrund dessen ist das Ergebnis, mit Ausnahme von hoch kritischen Risiken häufig ein Kompromiss aus beidem. Ein kollaborativer Ansatz und eine transparente Zusammenarbeit können einer brisanten Abschlussbesprechung vorbeugen.

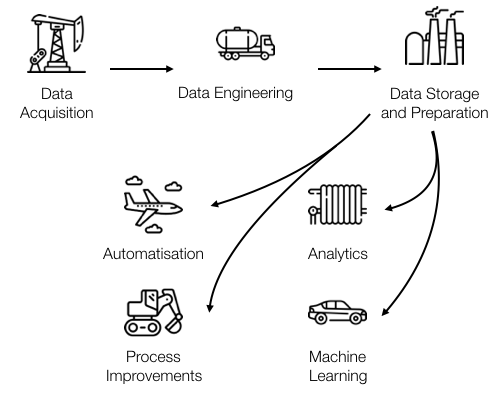
Da ein Telefon- oder Videogespräch die Chance von Missverständnissen erhöht und die Verständigung erschwert, sollte besonderes Augenmerk auf die Körsprache der beteiligten gelegt werden. Deshalb ist anzuraten, regelmäßig den gemeinsamen Stand und das Verständnis der Gesprächspartner zu überprüfen und das besprochene nach jedem Abschnitt zusammen zu fassen.

Im Idealfall wird jede beschlossene Maßnahme von der geprüften Einheit nicht nur akzeptiert, sondern als Mehrwert wahrgenommen und ihre Auswirkung ist für die Revision im Rahmen des Follow-Ups messbar nachzuprüfen. Je stärker eine Maßnahme quantitativ messbar ist, Beispielsweise in der Anzahl von Rechnungen ohne Bestellbezug, desto leichter ist eine Umsetzung zu dokumentieren und zu prüfen. Zusätzlich sollte für jede Maßnahme ein Verantwortlicher der geprüften Einheit und eine Deadline für die Umsetzung der Maßnahme festgelegt werden.

Definition of Done: Zu jeder Feststellung existiert eine überprüfbare Maßnahme.

1. **Infrastruktur der Datenanalyse**

„The worlds most valuable resource is no longer oil, but data” titelte der Economist im Jahre 2017. Während die Revolution mit dem Umgang von Daten lange von dominierenden digitalen Unternehmen wie Google, Apple oder Facebook vorgetrieben wurde, so hat sich die Nutzung verfügbarer Daten in den letzten Jahren demokratisiert. Infrastruktur wie Azure oder AWS stellt bereits eingerichtete Plattformen bereit und einfacherer Zugang zu Technologien wie Tensorflow oder IBM Watson Studio erleichtern die Analyse.

Getrieben durch mobile Endgeräte und die Verbreitung des „Internet-of-Things“, hat sich die globale Datenmenge von geschätzten 2 Zettabyte im Jahr 2010 auf circa 60 im Corona-Jahr 2020 vervielfacht. Es wird prognostiziert, dass sich diese Menge in den nächsten 3 Jahren verdoppeln wird. Die Fragestellung, welche Firmen auf der ganzen Welt umtreibt, ist wie unternehmenseigene und gesammelte Daten für die eigenen Ziele nutzbar gemacht werden können.

Obwohl die Revision kein Treiber dieser Entwicklung ist, so sieht sie sich in ihrer Funktion exponentiell wachsenden Datenmengen und gestiegenen Anforderungen gegenüber. Die Audicon GmbH bietet Software-Lösungen für weite Teile des Prüfungsprozesses an, um Teilen dieser Herausforderung Herr zu werden.

**Von Stichproben zur Datenanalyse**

Stichproben (Samples) sind ein klassisches Element der Statistik, um für einen Datensatz (Population) eine Hypothese zu validieren oder zu widerlegen. Eine Stichprobe von 10 Bestellungen soll beispielsweise belegen, dass der Prozess ordnungsgemäß durchgeführt wurde oder das Prozessschwächen entdeckt wurden. Die Revision kann anhand dieser Stichprobe eine Aussage zur Qualität des Prozesses und der dazugehörigen Kontrollen treffen.

Um eine repräsentative Stichprobe zu bilden, muss ein gewisser Prozentsatz der Vorgänge untersucht werden. Hierbei wächst die Stichprobe proportional zur Anzahl der Vorgänge. Eine repräsentative Stichprobe aus einer Anzahl von 10.000 Vorgängen, mit einer Konfidenz von 95% sollte mindestens 370 Vorgänge enthalten.

Es ist schnell ersichtlich, dass bei modernen ERP-Systemen mit Millionen von Vorgängen die Prüfung manueller Stichproben nicht zielführend ist. Es handelt sich um die bekannte Suche der Nadel im Heuhaufen. Die Anzahl der theoretisch zu überprüfenden Dokumenten sprengt den Rahmen jeder regulären Prüfung.

Allerdings ergeben sich durch Software gestützte „Computer-assisted audit tools“ (CAATs), Möglichkeiten der exponentiell gewachsenen Datenmenge bewältigen. Die Datenmenge kann durch ihre Anwendung auf eine kleinere Anzahl auffälliger Sätze reduziert werden, um Stichproben wirksamer zu gestalten. Alternativ ergeben sich die Möglichkeiten Datensätze systematisch nach problematischen Vorgängen zu untersuchen. Der Heuhaufen wird also entweder kleiner, oder wird gezielt mit einem Magneten durchsucht.

**Datenextraktion**

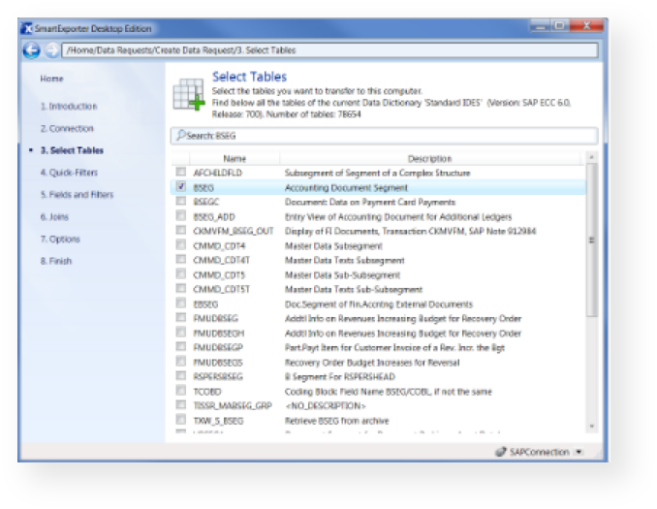
Wie bereits beschrieben, ist der erste Schritt der Datenanalyse die Extraktion der Daten. Hier werden verschiedene Arten von Daten aus unterschiedlichen Quellen und Dokumenten gesammelt und zusammengefügt. Die größte Informationsquelle für die interne Revision stellt das ERP-System der Firma da, in welchem die Geschäftsprozesse der Firma digital abgebildet werden.

Da eine große Anzahl von unterschiedlichen ERP-Systemen wie Oracle oder SAP genutzt werden, gestaltet sich der Abzug der Daten vielfältig. In erster Instanz müssen die verfügbaren Datenquellen und Datenbanken erforscht und eingegrenzt werden.

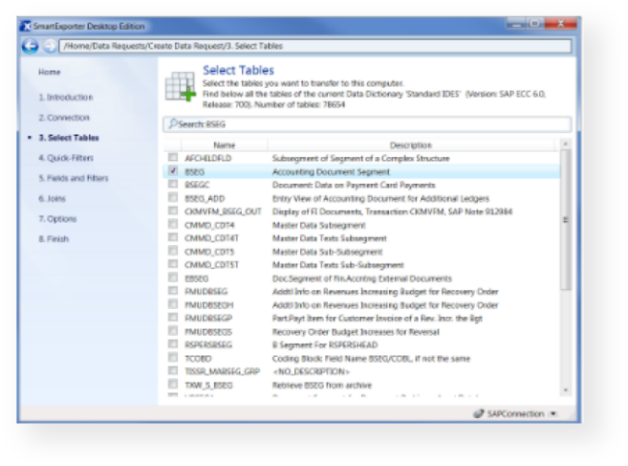
Ein Abzug kann über Datenqueries und strukturierte Sprachen wie SQL, oder über ODBC („Open Database Connector“) Schnittstellen erfolgen. Unabhängig von der Art des Abzugs muss die Datenintegrität und die Richtigkeit des Formats der jeweiligen Daten und Variablen sichergestellt werden, bevor ein Import der Daten in die gewählte Analyse Software erfolgen kann.

**Datenextraktion SAP (Smart Exporter)**

SAP speichert Stammdaten, Bewegungsdaten und Änderungsbelege automatisch in eine große Anzahl von Datentabellen. So werden beispielsweise Informationen zu den Kopfdaten von Bestellungen in der Tabelle „EKKO“ gespeichert.

Der „Smart Exporter“ bietet die Möglichkeit Tabellen direkt aus dem verbundenen SAP-System zu extrahieren. Zu Beginn ist es notwendig in dem Zielsystem einen Nutzer mit entsprechenden Autorisierungen zu generieren, welcher auf die gewünschten Tabellen zugreift. Nach einer Verbindung mit dem Zielsystem wird für die Abfrage eine Datenanforderung erstellt.

Eine Datenanforderung ist eine Kombination aus Tabellen, welche über einen Quickfilter (globale Filter) eingegrenzt wurden. Diese Filter enthalten unter anderem Datumsfelder, Buchungskreise, Sprachen und Kontenpläne. Nach Erstellung dieser globalen Filter müssen für jede Tabelle die gewünschten Spalten ausgewählt werden. Hierbei werden neben den technischen Feldbezeichnungen („LIFNR“) auch die Beschreibungen der Spalten angezeigt („Lieferantennummer“), was Anwendern ohne exzessives Verständnis der SAP-Datenstruktur ermöglicht, benötigte Informationen zu selektieren. Weiterhin kann das Format des Feldes direkt für die spätere Analyse übernommen werden. Dies greift einer späteren Formatierung („Transform“) vor.

Für jede der ausgewählten Tabellen kann zudem ein lokaler Filter auf ein beliebiges Datenelement gelegt werden. Die Möglichkeit Tabellen direkt über JOINS vor der Abfrage zu verknüpfen, reduziert redundante Daten und macht eine spätere Verknüpfung von Kopf und Positionstabellen überflüssig. Die Datenanforderungen können für wiederholten Gebrauch gespeichert werden, womit der Datenabzug automatisiert wird.

Muss eine DGSVO-Konformität gewährleistet sein, so ist es möglich Daten zu verschlüsseln. Hierzu muss eine Komponente in SAP installiert werden, welche die Pseudonymisierung oder Anonymisierung von ausgewählten Feldern wie Namen oder Adressen ermöglicht. Da dies durch den lokalen SAP-Admin geschehen kann, hat der Prüfer keine direkte Sicht auf sensible Daten.

Zuletzt kann die Datenanforderung auf verschiedene Weise gespeichert werden. Möglich sind eine direkte Extraktion und Laden in ein offenes IDEA-Projekt, die Speicherung als CSV- oder TXT-Datei.

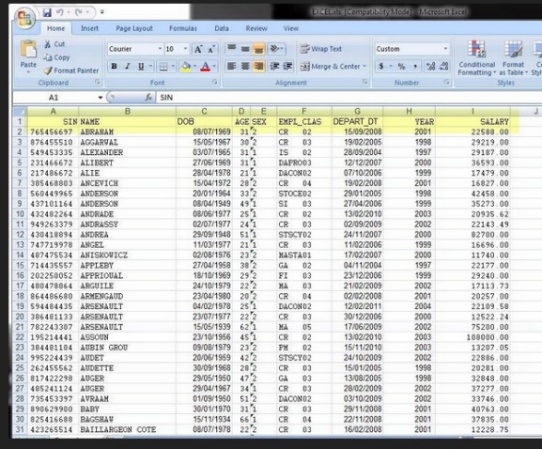
**Datenanalyse mit IDEA**

IDEA ist ein CAAT und wird seit mehr als 20 Jahren von Prüfern verwendet. Die Software bietet durch die Speicherung von Originaldateien und Speicherung aller durchgeführten Operationen in der Historie einen klaren Audit-Trail. Da die Originaldaten geschützt sind und jede Aktivität aufgezeichnet wird, bleibt die Integrität der Daten jederzeit erhalten und gleichzeitig können Prüfprojekte zwischen den Revisoren geteilt werden.

Zu Beginn einer Prüfung muss ein Projekt erstellt werden, von dem auf die Daten und das Script-Repository zugegriffen wird. Im Anschluss können die Daten in das Projekt importiert werden. Dies geschieht beispielsweise durch eine Datenanforderung aus dem Smart Exporter. Bei dieser Art des Imports werden die Daten analog zu SAP oder als vorgefertigter JOIN eingelesen. Alternativ lassen sich auch für jede Tabelle Import-Definitionen festlegen, welche den Variablentyp jedes Datenfeldes bestimmen.

Die Funktionalität wird kaum durch die Menge der Daten beschränkt. Jede Form von strukturierten Daten kann einfach geladen werden und selbst die Bearbeitung von Millionen von Datensätzen beeinträchtigt die Performance kaum.

Eine große Anzahl von vorgefertigten Ribbons mit Audit-spezifischen Prüfungsschritten ermöglicht den schnellen Einstieg in die Datenanalyse. Durch die einfache Aufzeichnung von Makros in IDEA- oder Visual-Script können Operationen wiederverwertet und automatisiert werden. Durch die Veränderung von Parametern entstehen so schnell komplexe Prüfschritte, welche sonst Stunden manueller Arbeiten in Anspruch genommen hätten.

Nach Sichtung der vollständigen Datenmenge kann durch wenige Klicks eine Datenentdeckung in Form von deskriptiven Statistiken durchgeführt werden. Hier werden beispielsweise automatisch die Durchschnittswerte für alle numerischen Variablen berechnet.

Ein nächster logischer Schritt wäre die Verknüpfung der Tabellen (Join), um beispielsweise Stammdatenfelder eines einzelnen Kreditors aus verschiedenen Quellen zu vereinen. Hier bietet IDEA den Vorteil, dass Tabellen über einen oder mehrere Primärschlüssel verknüpft werden können und der Nutzer die gewünschten Felder aus beiden Ursprungstabellen bestimmen kann (Beispiel: 10 Felder aus Tabelle 1 und 5 Felder aus Tabelle 2).

Sollten nur bestimmte Teile von Tabellen benötigt werden, so können die Daten entweder in einer Tabelle gefiltert werden oder per direkte Extraktion in eine neue Unterdatei umgewandelt werden. Hier tritt der Funktionseditor hervor. Ähnlich den Funktionen in Excel erlaubt er die Verknüpfungen von hunderten mathematischen und statistischen Operationen. Jede durchgeführte Funktion kann aufgezeichnet und zur Wiederverwertung gespeichert werden.

IDEA ist für den Remoteprozess ein geeignetes Tool, da es einen klaren Audit-Trail und eine nachvollziehbare Dokumentation erzeugt. Zudem ermöglicht es die Analyse und die Teilung von großen Datenmengen mit einem hohen Maße an wieder Verwertbarkeit für durchgeführte Analysen.

Aufzählung von Funktionen?

**Neuerungen von IDEA 11**

Die Version IDEA 11 erlaubt es, „IDEA Script“ durch die objektorientierte Programmiersprache Python zu ersetzen. Python stellt die am weitesten verbreitete Sprache für Datenanalyse und maschinelles Lernen dar. Mit ihr wird nicht nur Nutzung komplexer Schleifen (Loops) und Listen möglich, sondern auch die Arbeit mit Modulen wie „Scikit-Learn“, „Pandas“ und „Matplotlib“. Damit werden klassische Tools aus der Data Science für die Revision nutzbar.

Weiterhin bietet IDEA die Möglichkeit Tabellen über eine ODBC-Schnittstelle an andere Programme zu übermitteln. So können beispielsweise fertige Ergebnisse an Visualisierungssoftware wie Tableau oder Power BI übergeben werden. So bietet sich der Revision eine elegante Möglichkeit ihre Ergebnisse nicht in Form von Tabellen zu präsentieren, sondern in interaktiven Dashboards.

Um gestiegenen Anforderungen, einer höheren Anzahl von Analysen und begrenzten Ressourcen Herr zu werden, kann es ab einem bestimmten Reifegrad notwendig sein, Prozesse zu automatisieren. In diesem Zusammenhang wird auch von Robotic Process Automation (RPA) gesprochen. Ein Beispiel mit den verfügbaren Tools wäre es, eine definierte Datenanforderung über den Smart Exporter zu terminieren. Diese Abfrage wird nun aus SAP automatisch durchgeführt. Nach Abschluss des Downloads und des folgenden Ladens in IDEA, können nun geplante Skripte („Task Scheduler“) durchgeführt werden. Dies bedeutet, dass eine Datenanforderung und Skripte in IDEA an einem Abend gestartet, der ETL und Analyseprozess über Nacht automatisch durchgeführt und die finalisierte Analyse am nächsten Tag in Empfang genommen werden können.

Über IDEA Lab können Prototypen und SkripteScripts veröffentlicht und importiert werden. Es bietet Caseware die Möglichkeit Prototypen von neuen Funktionen außerhalb von regulären Releases zu veröffentlichen. Zudem können IDEA User ihre Ideen und Skripte global untereinander austauschen. Diese Entwicklung ist vor allem durch ihre Nutzung von maschinellem Lernen interessant. So kann das Plugin „Outlier Detection“ Ausreißer und ungewöhnliche Elemente aus Datensätzen identifizieren, indem beispielsweise verschiedene Kategorien und Mengenkombinationen analysiert werden. Hierbei kann aus einer Reihe von Algorithmen gewählt werden (Beispiel: „Isolation Forest“). Das Plugin „Text Analysis“ (Text Mining) analysiert Textdateien. Es erstellt automatisch eine Sentiment Analyse („Sentiment Analysis“), Worthäufigkeitsbalkendiagramme („Word Frequency Bar Charts“) und eine Wortwolke („Word Cloud“). Es ist zu hoffen, dass sich das Lab in ein aktives Forum für Ideen und neue Apps entwickelt.

Zuletzt bietet IDEA 11 mit Cloudshare eine eigene Cloudlösung. Über einen gemeinsamen Teamzugriff können registrierter Nutzer ihre Projektdateien und Scripts in die Cloud laden und austauschen. IDEA 11 bietet einen mitgelieferten Speicherplatz von 10 GB und verfügt über ein eigenes Hosting in Europa mit allen benötigten Sicherheitszertifikaten.

**Prüfungsautomatisierung mit IDEA Apps (P2P)**

Um Datenanalysen zu automatisieren, oder ohne größeres Vorwissen eine gewisse Tiefe bei der Analyse zu erreichen, bietet Audicon eine Reihe von spezialisieren „Apps“ an. Diese Apps bieten über den IDEA internen Smart Analyzer einen automatisieren Datenabzug, Aufbereitung der Daten, Analyse und dynamische Selektion der Tests an.

Die „Analytics“ Apps für den Vertrieb („Order-to-Cash“) und Einkauf („Purchase-to-Pay“) bieten eine vorgefertigte Anzahl von Tests und Auswertungen für ihre jeweiligen Prozesse an. Die „TaxAudit Professional App“ ist auf die Bedürfnisse der Anlagen- und Steuerprüfung zugeschnitten. Zuletzt erlauben die „Process Mining Apps“ für den Einkauf und den Vertrieb eine günstige Lizenzierung der Celonis Software. Sie erlaubt eingeschränkte Version des Process Minings, basierend auf fixierten Datenanforderungen des Smart Exporters über Stamm-, Bewegungs- und Änderungsdaten.

Die Apps erlauben es einer Revisionsabteilung, ohne eigenes Know-how eine gewisse Flughöhe für verschiedene Themengebiete zu erreichen. Die Steuerung erfolgt hier logisch Anhand eines Workflows. Der Smart Exporter wird genutzt um mit vorgefertigten Datenanforderungen, gefiltert auf die relevanten Prüfgebiete die nötigen Daten zu beschaffen. Anschließend werden die Daten in IDEA importiert und aufbereitet. Die Tests selbst können individuell ausgewählt und zugeschnitten werden. Zuletzt können die Ergebnisse in vorgefertigte Auswertungen extrahiert und an die geprüfte Einheit übergeben werden.

1. **Fazit & Ausblick**

In den vergangenen Kapiteln wurde der Prozess der Remoteprüfung mit ihren Herausforderungen beschrieben. Im Anschluss wurde aufgezeigt wie diese über die Infrastruktur der Datenanalyse angegangen und gelöst werden können.

Remoteprüfungen können so effektiv wie Vor-Ort Prüfungen durchgeführt werden. In Kombination mit Innovationen und digitalen Analysen stellen sie durch ihre Kostenersparnisse attraktive Alternativen für Teile der Revisionsarbeit dar. Viele der in der Corona Krise eingeführten Änderungen werden die Pandemie überdauern. Die Nutzung von eingeführter Infrastruktur und aufgebautem Know-how kann es Revisionsabteilungen ermöglichen die aktuelle Krise in eine Chance für die kommenden Jahre zu verwandeln.

Neue Techniken und Technologien (Machine Learning)

Stolpersteine

**WIP:**

Text auflockern (e.g. IDEA Funktionen)

Fazit schreiben

Darstellung anpassen