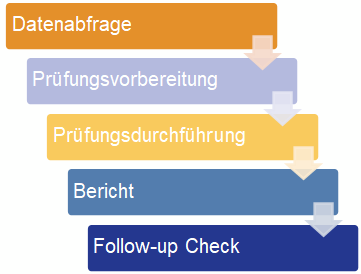
# Whitepaper Datenanalyse

1. **Einleitung**
   1. **Aktuelle Trends (1 Seite)**
      * Neue Technologien
        1. Massendaten, digitale Geschäftsmodelle, höhere Anforderungen
      * Digitale Risiken
        1. DGSVO
        2. Datensicherheit
      * Corona (Aktuelle Situation)
        1. Lockdowns
        2. Reiseverbote
        3. Prekäre Liquide Situation
2. **Remote Prüfung**
   1. **Möglichkeiten und Grenzen (1-2 Seiten)**
      * Grundsätzliche Herausforderungen und Voraussetzungen
        1. Technische Möglichkeiten (Software) und Verbindung (Bandbreite) - > Video & File Sharing
        2. Auditees mit eingeschränkter Verfügbarkeit (Zeitverschiebung, Kurzarbeit etc.)
        3. Fehlendes technisches Know-How bei Revisoren oder Auditee
      * Unterschiede zur Prüfung vor Ort
        1. Fehlende direkte Kommunikation
        2. Keinerlei Besichtigungen oder Demonstrationen vor Ort
      * Grenzen der Remote Prüfung
        1. Prozessbesichtigungen sind nicht möglich
        2. Kommunikation verlangsamt sich, der Prozess wird formeller
        3. Geprüfte Einheit ist schnell überfordert (Stichprobenlawine)
      * Agile Prüfung durch zeitnahes analysieren von Risiken
        1. Iterativer Prozess zur Abwägung zwischen Möglichkeiten und Anforderungen
        2. Keine Prüfung nach Checkliste, Fokus (Risiken) kann und wird sich während des Audits ändern
3. **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. So können Abteilungen, wie Controlling, Compliance und Risikomanagement direkt oder peripher involviert sein. Zudem muss bei einer remote Prüfung häufig die IT-Abteilung hinzugezogen werden, um Software oder VPN Zugänge bereit zu stellen.

Zuletzt sollte sich der Prüfungsleiter klarmachen, wer die finalen Adressaten der Revisionsarbeit sind. Wird der Bericht nur an die lokale Einheit kommuniziert, oder werden auch höhere Ebenen in der Organisation informiert? Das finale Ergebnis der Prüfung muss auf alle Adressaten zugeschnitten sein.

**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 sollen. So sollte auch eine Überlastung der geprüften Einheit vorgebeugt werden. Es sollte beachtet werden, dass bestimmte Abteilungen, wie die Finanzbuchhaltung in viele verschiedene Prozesse eingebunden sein können. Checklisten

**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.

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 einen Instant-Messenger-Dienst durchgeführt haben. Nach der Klärung des Scopes, kann entschieden werden, über welche Infrastruktur die gemeinsame Arbeit durchgeführt werden soll.

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

Logistische Anforderungen umfassen Equipment, Software und Policies. Equipment beschreibt Grundsätzliches wie Laptops, Drucker, Scanner, Headsets (Telefon) und Internet Verbindung 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.

Das 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 Email ist aufgrund dessen ab zu sehen. Sollte noch kein File-Sharing Tool zur Verfügung stehen, so gibt es am freien Markt professionelle Varianten von Dropbox, Slack, oder Google Cloud. Video-Chats bieten die Möglichkeit die fehlende Nähe, zumindest in Teilen zu überbrücken, da das Gegenüber zumindest sichtbar gemacht wird.

**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 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 die fremden Prozesse in eigene Datenanforderungen übersetzt werden 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.

**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) und 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.

**Durchführung von Tests und Selektion von Stichproben**

Nach der Eingrenzung der Datenmenge, sollten die originalen Daten unbearbeitet gesichert werden, um deren Integrität sicher zu stellen. Anschließend erfolgt eine Zuordnung der Daten zu den jeweiligen Tests und eine Aufteilung in verschiedene Kategorien aufgeteilt werden. Templates, Script-Repositories oder vorgefertigte Arbeitspakete können helfen den Testing-Prozess zu automatisieren. Besonders bei Standardvorgängen, oder wiederkehrenden Fragestellungen, kann es sehr effizient sein einmal durchgeführte Prüfungen aufzuzeichnen.

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: potentielle 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 minimalistisch gehalten sein.

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.

**Prüfungsdurchführung**

**Kick Off**

Das Kick-Off Meeting dient der Vorstellung aller beteiligten Personen. 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 hervor zu heben und um sicher zu stellen, dass die Prüfungsvorbereitung zur allgemeinen Zufriedenheit abgeschlossen wurde.

**Scheduling 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.

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.

**Review der Prozesse (Beispiel: Einkauf)**

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.

Der gesamte Prozess sollte anhand von Beispielen durchgespielt werden. Durch die fehlende Möglichkeit des „über die Schulter 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.

**Durchsicht aller Tests und Stichproben (Validierung von Ergebnissen)**

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 Papier-Form erst digitalisiert werden müssen, ist hier die Arbeitsbelastung des Auditees am höchsten. Aufgrund dessen 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 interner Umbuchungen entpuppen, erzeugt durch eine Neuziehung nicht nur viel Mehrarbeit bei den Revisoren und der geprüften Einheit, sondern auch für viel Unmut bei letzterem.

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 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.

**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 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 ihrer Ursachen aufgeführt sein. Bevor der Report finalisiert werden kann, muss jede Feststellung mit dem Auditee validiert werden, vor allem da Videokonferenzen leichter zu Missverständnissen führen können. Jegliches bekommene Feedback sollte eingearbeitet, oder zumindest zur Notiz genommen werden.

**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.

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-Up 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.

Im Idealfall wird jede beschlossene Maßnahme von der geprüften Einheit nicht nur akzeptiert, sondern als Mehrwert war genommen und ihre Auswirkung ist für die Revision, im Rahmen des Follow-Up messbar nach zu prüfen.

**Follow-Up**

Virtuelle Durchsprache der bearbeiteten Maßnahmen

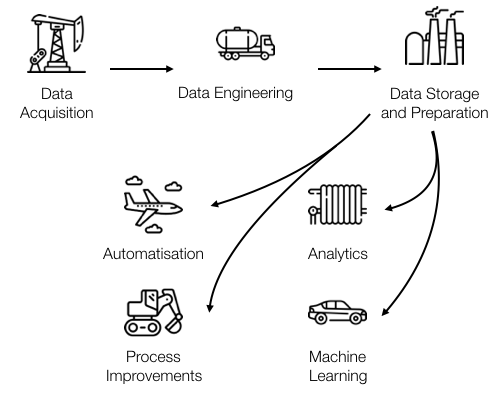
Wiederholung der Datenanalysen zur Identifikation von Verbesserungen

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 Digital Firmen 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 der 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 nutzbare gemacht werden können.

Obwohl die Revision kein Treiber dieser Entwicklung ist, so sieht sie sich als Funktion doch 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 einer Aussage zur Qualität des Prozesses und der dazugehörigen Kontrollen treffen. Dieses Instrument wird seit hunderten von Jahren eingesetzt. Die angeforderte Stichprobe enthält häufig die beteiligte Prozessdokumentation und die originalen Belege, welche die Leistung und die durchgeführten Kontrollen belegen.

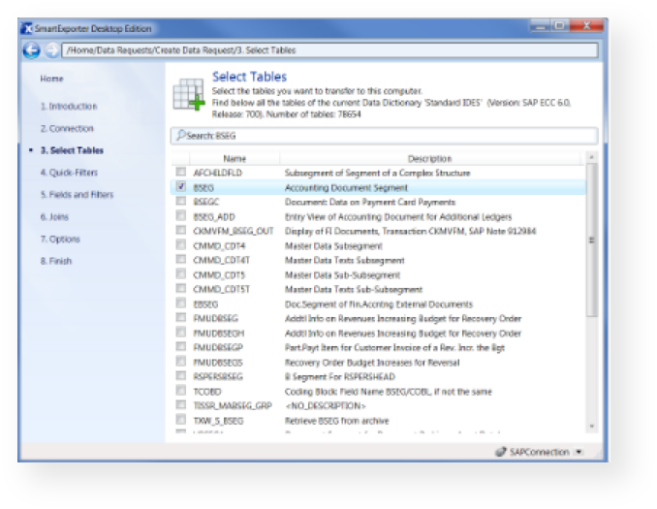
Um eine repräsentative Stichprobe zu bilden, muss ein gewisser Prozentsatz der Vorgänge untersucht werden. Hierbei wächst die Stichprobe proportional zur Anzahl der Vorgänge. Eine repräsentative Stichprobe aus einer Anzahl von 10.000 Vorgängen, mit einer Konfidenz von 95% sollte mindestens **370 Vorgänge**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

Es ist schnell ersichtlich das 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 Herr zu werden.

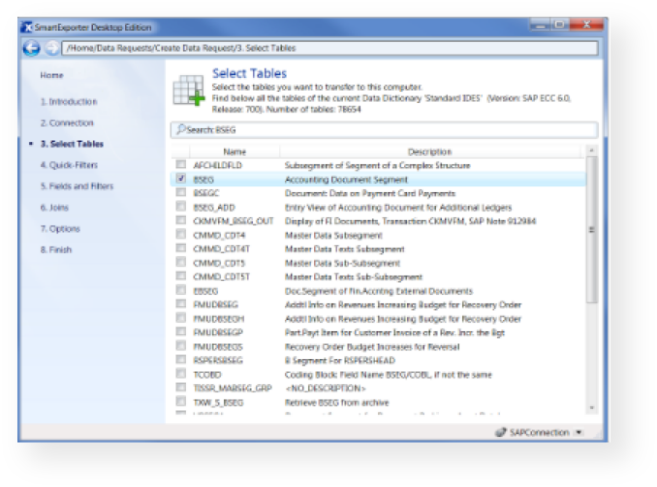
**Datenextraktion mit dem Smart Exporter**

Wie bereits beschrieben, ist der erste Schritt der Datenanalyse die Extraktion der Daten. 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 einen Nutzer mit entsprechenden Autorisierungen zu generieren, welcher über einen installierten RFC-Baustein 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 Datumfelder, 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 kann 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 Einladen in ein offenes IDEA Projekt, die Speicherung als CSV- oder TXT- Datei.

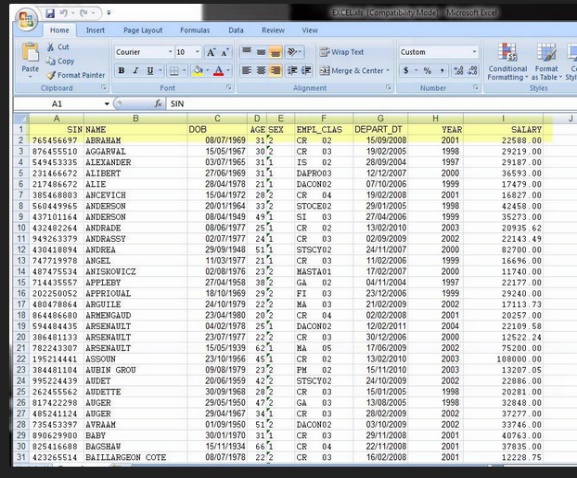
**Vorteile von IDEA (1)**

IDEA ist ein CAAT mit einer langen Geschichte und wurde als Werkzeug seit über 20 Jahren von internen und externen Prüfern, Steuerberatern und staatlichen Behörden eingesetzt. Obwohl es Excel auf den ersten Blick stark ähnelt, hat es eine Vielzahl von Vorteilen, welche es attraktiv für Prüfer und Datenanalysten machen.

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.

Zudem ist es in seiner Funktionalität kaum durch die Menge der Daten beschränkt. Jede Form von strukturierten Daten kann einfach eingeladen 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.

**Datenanalyse mit IDEA (2)**

Zum 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.

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 allen numerischen Variablen berechnet.

Ein nächster logischer Schritt wäre die Verknüpfung der Tabellen, um beispielsweise Stammdatenfelder eines einzelnen Kreditors aus verschiedenen Quellen zu vereinen. Hier bietet IDEA den Vorteil, dass Tabellen über einen primary Key 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 direkter Extraktion in eine neue Unterdatei umgewandelt werden. Hier tritt der Funktionseditor hervor. Ähnlich der Funktionen in Excel erlaubt er die Verknüpfungen von hunderten mathematischen und statistischen Operationen. Jede durchgeführte Funktion kann aufgezeichnet und zur Wiederverwertung gespeichert werden.

Mehrfachbelegung & Duplikate

Summierung / Pivot / Schichtung / Alterstruktur

**IDEA 11 - Neuerungen (3)**

Die Version IDEA 11 erlaubt es, „IDEA Script“ durch die objektorientierte Programmiersprache Python zu ersetzen. Python stellt die am weitesten verbreiteten 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 Daten 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 her zu werden, kann es ab einem bestimmten Leven von Maturity 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. Diese bedeutet, dass eine Datenanforderung und Skripte in IDEA an einem Abend gestartet, der ETL und Analyse Prozess über Nacht automatisch durchgeführt werden und die finalisierte Analyse am nächsten Tag in Empfang genommen werden können.

Über IDEA Lab können Prototypen und Scripts 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“ Ausreiser und ungewöhnliche Elemente aus Datensätzen identifizieren, indem beispielsweise verschiedene Kategorien und Mengenkombinationen analysiert werden. Hierbei kann aus einer Reihe von Algorithmen gewählt werden (z.B. „Isolation Forest“). Das Plugin „Text Analysis“ (Text Mining) analysiert Textdateien. Es erstellt automatisch eine Sentiment Analyse, Worthäufigkeits-Balkendiagramme und eine Wortwolke. Es ist zu hoffen, dass sich das Lab in ein aktives Forum für Ideen und neue App entwickelt.

Cloudshare

* IDEA Projektdateien können in die Cloud (10GB)
* Teamzugriff für alle registrierten User
* Alle Sicherheitszertifikate / Hosting in Europa

**Prüfungsautomatisierung mit IDEA Apps (P2P)**

* + - Apps (P2P, Process Mining) – Standardisierte Standort Prüfungen
      1. Automatisierung der Prüfung
      2. Zusätzliche Prüfgebiete
    - Zusammenspiel
      1. Schematischer Ablauf einer Prüfung mit SE, IDEA und App
    - Teamwork und Kollaboration
      1. Organisation des Abteilungs Know-Hows
      2. Vermeidung von „Silos“
      3. Cloud Share?

1. **Fazit und Ausblick (1 Seite)**
   1. Alles wird digital
   2. Neue Techniken (e.g. Machine Learning)

Einsatzgebiet in der Prüfung / Fragestellung / Übersetzung einer Fragestellung in ein Datenproblem /

Datenanalyse / Interpretation / Kommunikation der Ergebnisse

**To Dos:**

* **Typische Stolpersteine, Werttreiber & Mindesterfordernisse**
  + **Für jeden Absatz**
  + **Finale Checkliste**
* Darstellunge, Zitate etc
* Checkliste als Anhang