



IDSt-White Paper

Fachausschuss VII „AG Automatisierte Steuererklärung“

Warum eine automatisierte Steuererklärung keine Illusion ist!

Fachausschuss VII

„AG Automatisierte Steuererklärung“

White Paper¹

Warum eine automatisierte Steuererklärung keine Illusion ist!

¹ Das White Paper spiegelt die Auffassungen des gesamten Fachausschusses VII wider, aus dem heraus es innerhalb einer Arbeitsgruppe unter dem Vorsitz von Heiko Höbbel entstanden ist. Weitere Mitglieder der Arbeitsgruppe sind: Kristiina Coenen, Christian Engel, Michael Heil, Bianca Hempel, Till Hoffmann, Florian Rath, Bernd Sacher und Jörg Schindler.

Inhalt

1. Zielsetzung und Motivation.....	3
2. Ausgangssituation.....	4
3. Projekt Setup und Rahmenbedingungen.....	5
4. Standard-Vorgehensweise zur Evaluation im Unternehmen.....	8
4.1 Abgrenzung und Definition der Daten	8
4.2 Identifizierung der Datenquellen und Funktionen.....	10
4.3 Technische Anbindung der Datenquellen (Schnittstellen)	10
4.4 Akquise, Aufbereitung und Bereitstellung der Daten.....	11
4.5 Generierung standardisierter Daten	12
4.6 Technische Anbindung der Compliance-Systeme	13
4.7 Verarbeitung und Versendung durch die Compliance-Lösung.....	14
5. Automatisierte Steuererklärung unter Kontrollaspekten.....	15
6. Abschließende Bewertung und Petitum	19

Abbildungsverzeichnis

Abbildung 1: System- und Datenschichtenmodel mit Schnittstellenbetrachtung.....	4
Abbildung 2: Möglicher End-to-End-Prozess einer automatisierten Steuererklärung	7
Abbildung 3: Ansätze zur Quelldaten-Identifizierung.....	9

1. Zielsetzung und Motivation

Die interdisziplinär besetzte Arbeitsgruppe „Automatisierte Steuererklärung“ – angesiedelt im Fachausschuss VII (Innovative Technologien) des Instituts für Digitalisierung im Steuerrecht (IDSt) – hat es sich zum Ziel gesetzt, Automatisierungsoptionen und eine Referenz im Bereich der automatisierten Steuerdeklaration zu erarbeiten. Die aus der bisherigen Arbeit gewonnenen Erkenntnisse sind in dem vorliegenden White Paper niedergelegt.

Das White Paper ist wie folgt aufgebaut: Zunächst werden Zielsetzung und Motivation dargelegt (**Kapitel 1**). Ausgehend vom Status quo (**Kapitel 2**) wird unter bestimmten Rahmenbedingungen (**Kapitel 3**) eine beispielhafte Vorgehensweise entwickelt, mit welcher sich die Umsetzung und Möglichkeiten einer „Automatisierten Steuererklärung“ unternehmensindividuell evaluieren lassen (**Kapitel 4**). Am Ende sollen Kontrollaspekte strukturiert und erläutert werden (**Kapitel 5**).

Mit Fokus auf die Ertragsbesteuerung müssen dabei gleichermaßen technische, organisatorische und steuerrechtliche Aspekte beleuchtet werden. Die alleinige Betrachtung von Buchhaltungs- und Steuerprozessen bzw. eine prozesslosgelöste Betrachtung von Technologie wird der Komplexität einer automatisierten Steuererklärung dabei nicht gerecht. Vielmehr gilt es einen ganzheitlichen Blick auf Daten, Prozesse, Technologien und Kontrollen sowie auf bestehende Ansätze und Plattform-Lösungen² zu werfen.

Die Arbeitsgruppe ist davon überzeugt, dass die Automatisierung im Bereich der Steuererklärungserstellung wesentliche Effizienzen heben und die Qualität der ertragsteuerlichen Deklaration verbessern kann. Mit dem White Paper sollen Bedenken bzgl. Fehleranfälligkeit sowie Durchführbarkeit einer voll- bzw. teilautomatisierten Erstellung der Ertragsteuererklärung abgebaut bzw. beseitigt sowie das konzeptionelle Verständnis verbessert werden.

Übergreifend lassen sich mittels einer Automatisierung des Steuerklärungsprozesses vornehmlich folgende **Ziele** erreichen:

- Sicherstellung der **Compliance und Datenqualität**
- Gesteigerte **Prozesseffizienz**
- Effizienter **Personal- und Ressourceneinsatz**
- Verbessertes **Kontrollumfeld**
- Vervollständigung des **digitalen Datenpools**

Die im Rahmen der Erstellung dieses White Papers gewonnenen Erkenntnisse der Arbeitsgruppe sollen ergänzend in einem **POC (Proof-of-Concept)** konkretisiert werden, welcher im Nachgang zur Veröffentlichung des White Papers als Referenz zur Verfügung gestellt werden soll.

² Lösungen wie beispielsweise Alteryx (<https://www.alteryx.com/de> (25.11.2023)) oder KNIME (<https://www.knime.com/> (25.11.2023)).

2. Ausgangssituation

Die **Ausgangssituation** hin zu einer automatisierten Steuererklärung stellt sich im Wesentlichen wie folgt dar:

- Die Festsetzung der Ertragsteuern eines Veranlagungszeitraums unterliegt einem gewissen Lebenszyklus, in dem es im Zeitablauf beispielsweise zu Änderungen aufgrund von angepassten Steuererklärungen oder durch Folgerungen aus Betriebsprüfungen kommen kann.
- Ertragsteuererklärungen sind derzeit überwiegend das Ergebnis von manuellen Prozessen, in denen Informationen verschiedenster Unternehmensbereiche verarbeitet werden.
- Die Daten sowie ergänzende Detailinformationen, welche für die Erstellung einer Ertragsteuererklärung erforderlich sind, werden daher häufig nicht durch die Steuerfunktion oder den steuerlichen Berater des Unternehmens verantwortet, sondern entstehen in steuerfremden Abteilungen.
- Die zugrundeliegenden (Daten-)Formate sowie die Quellsysteme divergieren und sind in der Regel nicht an die nachgelagerten Lösungen (wie beispielsweise Deklarations-Lösungen) angebunden.
- Kontrollen finden vielfach (noch) manuell, nicht selten unter Nutzung umfangreicher Excel-Tabellen, statt.

Im Rahmen dieser Ausarbeitung haben die Autoren ein Grundszenario entwickelt, welches in Abbildung 1 beschrieben ist. Den Kern bildet eine generelle Datenschicht (System-Ebene 2) mit entsprechenden Datenschnittstellen (vor allem Interface-Ebenen 1 und 2), worüber sich unterschiedliche ertragsteuerliche Prozesse unterstützen lassen.

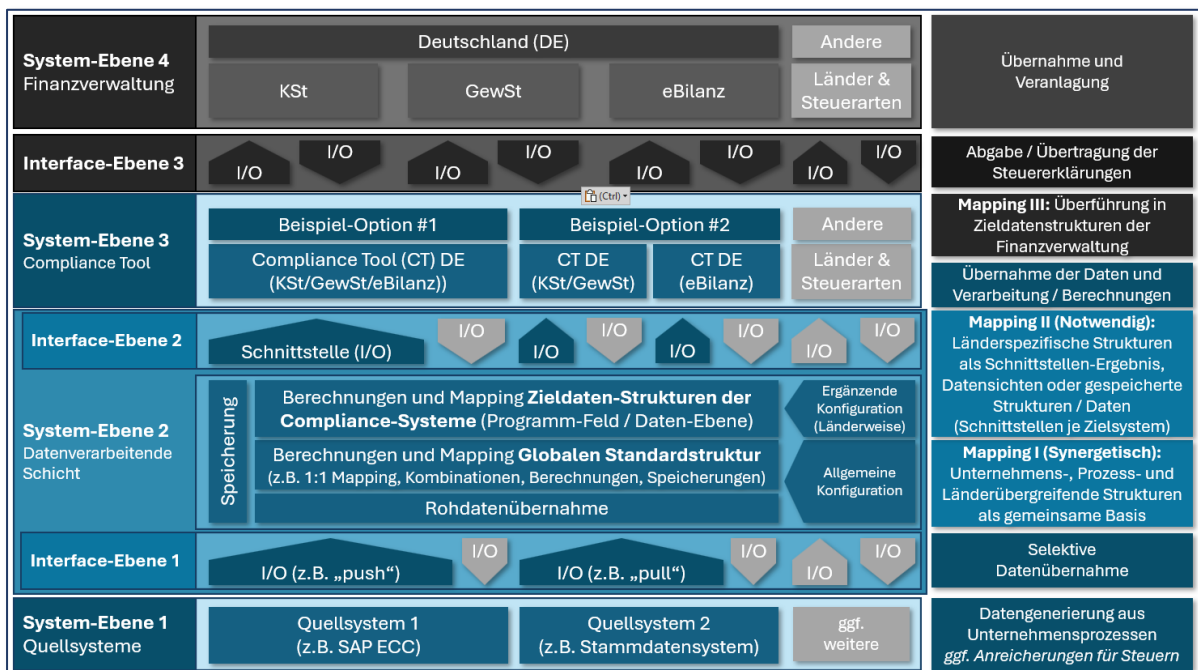


Abbildung 1: System- und Datenschichtenmodell mit Schnittstellenbetrachtung

Die insbesondere in der Systemebene 2 verwendete Begrifflichkeit des „**Mappings**“ bedarf einer Definition: „*Mapping bezeichnet den Prozess, durch den Beziehungen oder Übereinstimmungen zwischen zwei Systemen oder Datensätzen hergestellt werden. In der Finanzwelt kann dies das Zuordnen von Konten, Transaktionen oder Datenpunkten zwischen verschiedenen Berichtsstandards oder -systemen bedeuten. Es dient dazu, Konsistenz, Effizienz und Genauigkeit in der Datenverarbeitung und -analyse sicherzustellen.*“³

Diese von ChatGPT generierte Grundaussage ist zutreffend: In dem hier betrachteten Szenario geht es um die Gewinnung von Daten, deren Transformation und ihrer Nutzung zur Unterstützung von verschiedenen ertragsteuerlichen Prozessen.

Da ein einfaches Konten-Mapping (also das Zuordnen des Saldos eines Hauptbuchkontos zu einem steuerlichen Sachverhalt) aufgrund mangelnder Differenzierung für eine automatisierte steuerliche Qualifikation nicht ausreicht, sind Mapping-Konzepte erforderlich, die alle verfügbaren und relevanten Daten bzw. Datenattribute nutzen. Entsprechend sind auch Attribute wie z.B. sog. Bewegungsarten, Dokumentarten, Partner-Nummern, Tags, etc. für das Mapping heranzuziehen.

Ein sachgerechtes Mapping ist dabei nicht nur ausschlaggebend für das Gelingen der Automationsbemühungen, sondern auch eine äußerst herausfordernde Aufgabe: So existieren im Bereich der Ertragsteuer komplexe Zielstrukturen (Formularfelder der Körperschaftsteuer- und Gewerbesteuererklärung und eBilanz), die sich dazu im Zeitablauf ändern. Veränderungen unterliegen auch die Quellsysteme, die zudem häufig mehr als einen Rechtsraum abdecken müssen und daher nicht für steuerliche Anforderungen einer Jurisdiktion optimiert sind bzw. werden können.

3. Projekt Setup und Rahmenbedingungen

Für die Erarbeitung einer **Standard-Vorgehensweise** zur Evaluierung der Machbarkeit einer automatisierten Steuererklärung im Unternehmen wurde im Rahmen eines „Proof-of-Concept (PoC)“ eine **Arbeitsumgebung** aufgebaut und anschließend durchlaufen. Um die bestehende Komplexität zu reduzieren, mussten dabei diverse Annahmen getroffen werden.

Übergeordnete Determinanten:

- Rechtsraum: Deutschland
- Bereich: Ertragsteuern
- Steuerart: Körperschaftsteuer
- Steuerpflichtiger:
Stand-Alone Kapitalgesellschaft (keine ertragsteuerliche Organschaft)

³ ChatGPT Abfrage vom 22.09.2023, <https://chat.openai.com/>.

Technologisch wurde das folgende Szenario unterstellt (Arbeitsumgebung):

- Ein ERP-System (hier beispielhaft SAP ECC)
- Ein führendes Stammdatensystem
- Eine serverbasierte ETL⁴-Landschaft
- Eine Deklarationslösung, in welche Daten übernommen, dort aggregiert, ergänzt, transformiert und an die Finanzverwaltung übertragen werden können (hier beispielhaft AMANA und DATEV).

Das angenommene Vorhandensein mindestens zweier Quellsysteme (ERP- und Stammdatensystem) macht eine Datenverarbeitung erforderlich. Unter Datenverarbeitung soll hierbei sowohl die Standardisierung, das Mapping und die Interpretation von Eingangsdaten, als auch die Berechnung von neuen Daten verstanden werden.

Zusammenfassend wurden folgende Rahmenbedingungen (**Projekt Setup**) zugrunde gelegt:

- **Steuerart:** Körperschaftsteuer einer Stand-Alone Kapitalgesellschaft.
- **Prozesse:** Datengewinnung, Datenvorbereitung/-aufbereitung, Datenübergabe, Steuererklärungsabgabe und Datenwiederverwendung.
- **Datenklassifizierung und Betrachtung:** Transaktionale Daten/Stammdaten.
- **Kontroll-Opportunitäten:** Ausführungen zu möglichen regelbasierten und automatisierbaren Kontrollen.

Auf Basis der genannten Rahmenbedingungen wurde idealtypisch eine ganzheitliche Datenverarbeitungsschicht definiert. Datenschicht meint hierbei eine Instanz zur Verarbeitung und Speicherung der Daten der Quellsysteme (hier: SAP ERP, Stammdatensystem) und Senden der transformierten Daten an die Zielsysteme (hier: AMANA oder DATEV). Der mögliche „End-to-End-Prozess“ einer automatisierten Steuererklärung stellt sich entsprechend Abbildung 2 graphisch wie folgt dar:

⁴ Extract-Transform-Load – ETL-Tools sind Software-Lösungen die Daten aus verschiedenen Quellen extrahieren, in einheitliche Formate transformieren und für die weitere Verwendung in Datenbanken, ein Data Warehouse oder andere Anwendungen laden.

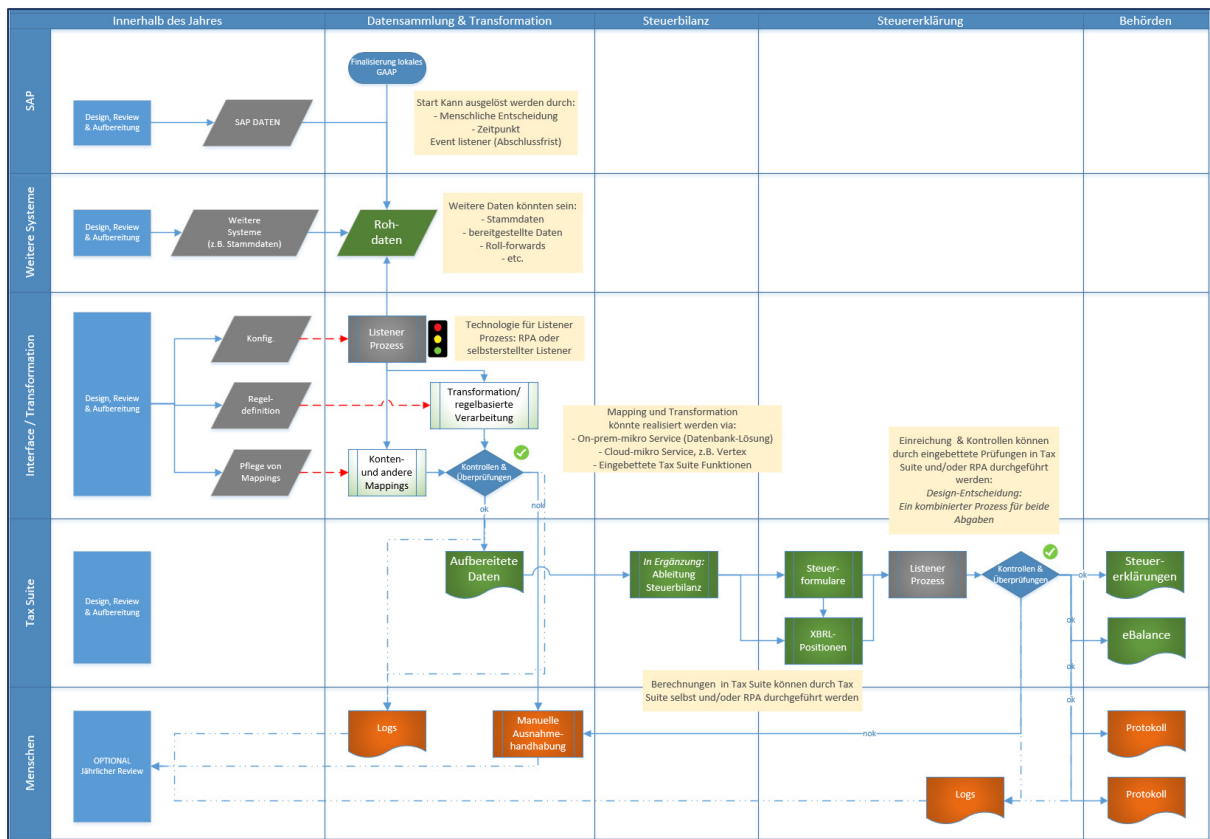


Abbildung 2: Möglicher End-to-End-Prozess einer automatisierten Steuererklärung

Wie in Kapitel 2 beschrieben, muss das Mapping hierbei nicht auf die Zuordnung von Hauptbuchkonten zu Bilanz-, Gewinn- und Verlustpositionen und/oder steuerliche Datenfelder beschränkt sein. Vielmehr können auch andere Datenattribute eines Buchungssatzes, wie z.B. Bewegungsarten, Partner-Nummern, Tax-Tags, etc. genutzt werden. Um Berechnungen und Ableitungen zu ermöglichen, lassen sich neben den Daten der Hauptbuchführung auch mehrere Quellsysteme nutzen. Dies kann durch Programmierung oder durch flexible Umgebungen, wie ETL-Lösungen umgesetzt werden.

Für die Datenvalidierung kann auf vorhandene Compliance-Lösungen zurückgegriffen werden.

4. Standard-Vorgehensweise zur Evaluation im Unternehmen

Um die Möglichkeiten der Umsetzung der „Automatisierten Steuererklärung“ unternehmensindividuell zu evaluieren, wurde innerhalb der skizzierten Arbeitsumgebung eine **Standard-Vorgehensweise** entwickelt. Hierfür wurde der Gesamtprozess zunächst in verschiedene Phasen aufgeteilt. Die Aufteilung orientiert sich dabei an den für den Gesamtprozess erforderlichen Daten. Konkret wurden acht Phasen definiert:

1. Abgrenzung und Definition der Daten (Kapitel 4.1)
2. Identifizierung der Datenquellen und Funktionen (Kapitel 4.2)
3. Technische Anbindung der Datenquellen (Schnittstellen) (Kapitel 4.3)
4. Akquise, Aufbereitung und Bereitstellung der Daten (Kapitel 4.4)
5. Generierung standardisierter Daten (Kapitel 4.5)
6. Technische Anbindung an das Compliance-System (Kapitel 4.6)
7. Verarbeitung und Versendung durch die Compliance-Lösung (Kapitel 4.7)
8. Berücksichtigung im Kontroll-Framework (Kapitel 5 – separate Darstellung)

Innerhalb der einzelnen beschriebenen Phasen wurden (1) die jeweiligen **Zielsetzungen** festgelegt, (2) die wesentlichen **Fragestellungen** aufgeworfen, (3) die zur Beantwortung gewählte **Methodik** und **Vorgehensweise** erarbeitet sowie (4) das **zu erwartende Ergebnis** formuliert. Als Resultat sollen sich damit über alle Phasen unternehmensindividuelle Aussagen zu den Anforderungen und zur Machbarkeit einer automatisierten Steuererklärung treffen lassen.

4.1 Abgrenzung und Definition der Daten

(1) Zielsetzung

Zunächst gilt es, die benötigten Daten bzw. Datenbereiche zu definieren, wozu die nachfolgende Abgrenzung dienen soll. Für die Ertragsteuererklärung benötigte Daten sind hiernach solche, die

- in den verschiedenen Vorsystemen und in den verschiedenen Modulen der Hauptbuchführung relevant sind (Quell-Daten [*SRC-Daten*]),
- zu der Erstellung der sog. CoTo-Daten (s. nächster Punkt) notwendig sind [*ETL-Daten*],
- als Grundlage zur Ermittlung der zu übertragenen Daten benötigt werden (Compliance-Tool-Daten [*CoTo-Daten*]),
- final als Datensatz an die Finanzbehörden übertragen werden müssen (Steuererklärungs- und E-Bilanz-Daten [*StE-Daten*]).

(2) Zu betrachtende Fragestellungen

- Welche Daten werden für eine automatisierte Steuererklärung bzw. einen „Proof-of-Concept“ (PoC) benötigt?
- Welchen Prozessen sind diese Daten zugeordnet?

(3) Methodik und Vorgehen

Mit Blick auf die Zielsetzung stehen folgende Methoden zur Wahl:

- Von der Quelle zum Ziel – Identifizierung aller möglichen steuerrelevanten Daten in den Quellsystemen.
- Vom Ziel zur Quelle – Identifizierung der zu ermittelnden Daten (z.B. Formulardaten) und Vergleich mit den Quellsystemen und den Verarbeitungsmöglichkeiten.
- „Two-Way-Approach“ – beide vorgenannten Methoden werden gleichzeitig angewendet.

Hinweis: Die Arbeitsgruppe hat sich entschieden, primär der ersten Methodik zu folgen, ergänzend jedoch bereits vorhandene Informationen, wie Konten-/System-Mappings und Verarbeitungsmethoden, in die Betrachtung einzubeziehen.

(4) Erwartetes Ergebnis

- Definition der Daten unterteilt nach steuerrelevanten Prozessen.
- Aussage zur Verlässlichkeit der gewählten Methode.

Abbildung 3 soll den Ansatz zur Quelldaten-Identifizierung visualisieren und skizziert dabei im Wesentlichen zwei für die Abgrenzung und Definition von Daten relevante **Entscheidungsräume**:

1. Datenbestimmungen ausgehend von der ERP-Prozesseite sowie anderer Quellsysteme und/oder ausgehend von den Formularfeldern.
2. Alle Unternehmen einer Unternehmensgruppe betrachtend oder (wie innerhalb der hier angenommenen Arbeitsumgebung) einzelne Gesellschaften fokussierend.

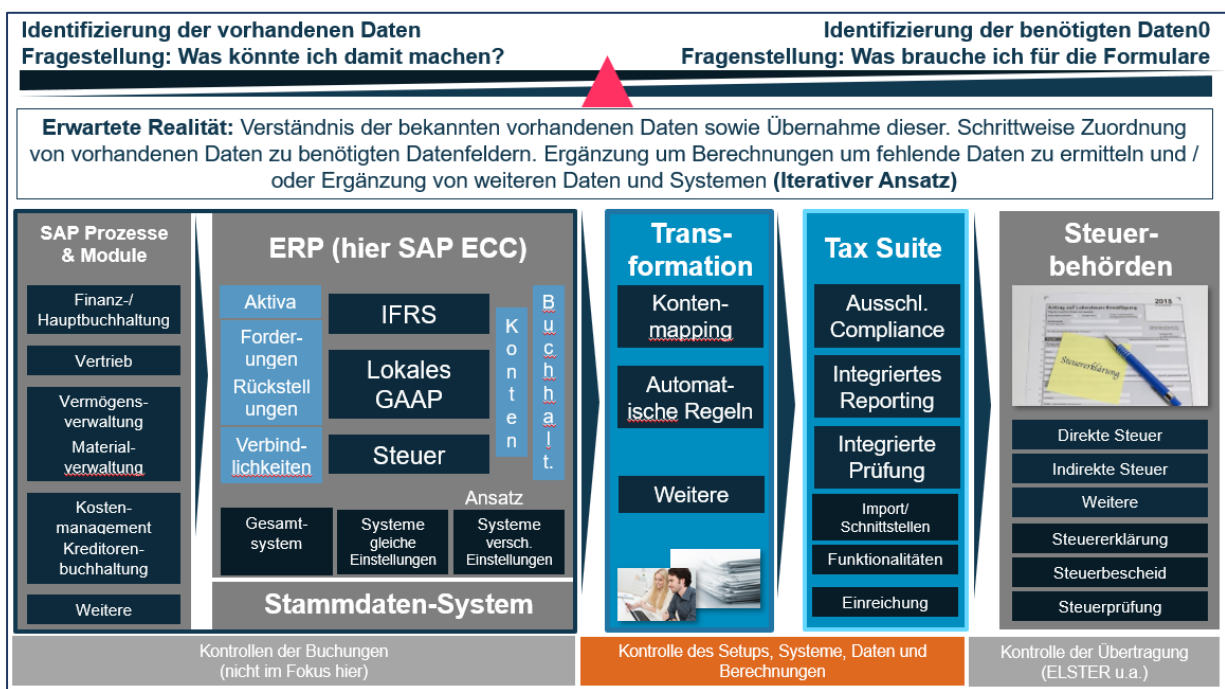


Abbildung 3: Ansätze zur Quelldaten-Identifizierung

4.2 Identifizierung der Datenquellen und Funktionen

(1) Zielsetzung

Mit Blick auf die Abgrenzung und Definition der Daten sind unterschiedliche Quellsysteme in Betracht zu ziehen. Dabei handelt es sich vornehmlich um die Finanzbuchhaltung bzw. ERP-Systeme, vorgelagerte Systeme (z.B. Einkauf, Reisekosten), führende Stammdatensysteme (z.B. zentrales Stammdatensystem für Unternehmensdaten), aber auch Systeme der Steuerfunktion, die für Steuererklärungen genutzt werden können (z.B. Steuerplanungssysteme, eigene Stammdatensysteme).

Hinweis: Dem Projekt-Setup liegt ein ERP-System des Herstellers SAP in einer umfassenden Standard-Konfiguration zugrunde, sodass neben der Finanzbuchhaltung (Modul FI) auch auf Daten der SAP-Module CO, MM, SD und AM zugegriffen werden kann. Wenn gleich als Release-Version ein ECC6-System unterstellt wurde, lässt sich die erarbeitete Standardvorgehensweise auch auf die neue SAP S/4HANA-Umgebung übertragen. Konkret wird auf SAP-Standardreports (z.B. RFBILA, Anlagen-Gitter) und SAP-Standard-Tabellen abgestellt. Für Zwecke der Stammdaten wurde ein vorliegendes zentrales Stammdaten-System unterstellt bzw. simuliert.

(2) Zu betrachtende Fragestellungen

- Welche Daten sind über das (zentrale) ERP-System verfügbar und welche Datenfelder der unterstellten Referenz-Compliance-Systeme bzw. der auszufüllenden Formulare können damit bestückt werden?
- Welche führenden Stammdatensysteme sowie weitere Grundaufzeichnungen sind verfügbar?

(3) Methodik und Vorgehen

- Identifizierung der relevanten Module/Tabellen aus den Quellsystemen. Mappen der entsprechenden Datenfelder auf Referenztabellen sowie Reports der relevanten Quellsysteme.

(4) Erwartetes Ergebnis

- Übersicht zu den einschlägigen Quellsystemen bzw. Datenquellen.
- Mapping-Tabelle.

4.3 Technische Anbindung der Datenquellen (Schnittstellen)

(1) Zielsetzung

Mit Blick auf die Nutzung der Datenquellen besteht die Zielsetzung dieser Phase darin, technische Möglichkeiten zu beschreiben, wie die zuvor identifizierten Quelldaten [SRC-Daten] in einer Datenschicht (Extract, Transform, Load = ETL-Umgebung) verfügbar gemacht werden können.

(2) Zu betrachtende Fragestellungen

- Wie können Prozesse angestoßen werden, damit die Daten zur gewünschten Zeit, in der gewünschten Form und gewünschten Qualität zur Verfügung stehen?
- Was sind die möglichen Strategien und Technologien, die einem Unternehmen zur Verfügung stehen (SQL-Zugriff, Rest-Services, RFC, mittelbarer Datenzugang durch eine Data-Lake-Lösung, ETL oder ähnliche Ansätze)?
- Gibt es Standard-Methoden oder Best-Practice-Ansätze für die Systemanbindung und den Datenaustausch? Hinweis: Diese Fragestellung ist wichtig, da der Datenaustausch zwischen Systemen seitens der Unternehmens-IT in der Regel restriktiv gehandhabt wird und standardisiert ist.

(3) Methodik und Vorgehen

- Verständnis der bisherigen IT- und Daten-Strategie des Unternehmens bezogen auf die im Fokus stehenden Systeme und Daten sowie im Hinblick auf bestehende Schnittstellen.
- Evaluierung vorhandener Datenfluss- und Schnittstellendokumentationen.
- Auswahl von möglichen Dienstleistern und Kooperationspartnern zur Mitwirkung im Rahmen eines „Proof-of-Concept“ (soweit am Markt verfügbare Lösungen zur Anbindung von im Fokus stehenden Software-Lösungen vorhanden sind oder neue Erfahrungen aufgebaut werden müssen).

(4) Erwartetes Ergebnis

- Möglichkeiten und Methoden einer technischen Anbindung der einschlägigen Quellsysteme und Datenquellen einschl. Bewertung (Kosten-Nutzen).
- Dokumentierte Beispielprozesse.

4.4 Akquise, Aufbereitung und Bereitstellung der Daten**(1) Zielsetzung**

Diese Phase dient dazu, mittels einer (oder mehrerer) Beispieltechnologie(n) zu eruieren, wie Lade- und Verarbeitungsprozesse in Bezug auf die erforderlichen Datenquellen grundsätzlich aufgebaut werden können, sodass diese automatisierbar sind.

(2) Zu beantwortende Fragen

- Wie können Daten automatisiert akquiriert, aufbereitet und geladen werden?
- Wie könnte eine Verarbeitung in einer ETL-Landschaft aussehen?
- Wie lassen sich Regeln zur automatisierten Akquise, Aufbereitung und Bereitstellung einrichten?

Hinweis: Ein Beispiel für eine solche regelbasierte Verarbeitung ist die Qualifikation von Dividendenerträgen nach § 8b Abs. 1 i.V.m. Abs. 4 KStG. Eine Steuerfreiheit ist nur für den Fall einer Beteiligung von größer/gleich 10% zum Beginn des

Kalenderjahres gegeben. Die Verbuchung von Dividendenerträgen im Hauptbuch kann als weiteren Buchungssatzbestandteil die sog. Partner-Nummer beinhalten. Die Partner-Nummer ist in der Regel die Gesellschaftsnummer der Gesellschaft im Konzern. Werden aus dem Gesellschaftsstammdaten-System die Beteiligungsdaten importiert, so können die Daten aus der Buchhaltung und aus dem Stammdatensystem kombiniert werden. Als Ergebnis dieser Kombination kann eine automatisierte Qualifikation nach § 8b KStG erreicht werden durch Aufteilung in Fälle mit einer Beteiligung kleiner 10% und größer/gleich 10%. Diese ermittelten Daten können dann zur weiteren Verarbeitung weitergegeben werden.

(3) Methodik und Vorgehen

- Aufbau eines beispielhaften ETL-Workflows (Extract, Transform, Load).

Hinweis: Dem Setup des PoC liegt die Lösung KNIME⁵ zugrunde.

(4) Erwartetes Ergebnis

- Möglichst konfigurierte „Step-by-Step“ Beispiel-Konfiguration.
- Beschreibung der Verarbeitungs-Optionen (Strukturen, Formate, errechnete Daten, Frequenz, Abbruch-Verhalten, Umgang mit Änderungen der Daten).

4.5 Generierung standardisierter Daten

(1) Zielsetzung

Nach unserem Verständnis existiert aktuell kein einheitliches Datenmodell im ertragsteuerlichen Bereich, welches als Strukturhilfe bei der Datenextraktion oder bei der Übertragung von Daten aus einer ETL-Umgebung in eine Compliance-Lösung genutzt werden kann. Entsprechend besteht die Zielsetzung dieser Phase darin, ein Prozedere zur Überführung der Daten in eine oder mehrere Strukturen zu erarbeiten, welche von der Compliance-Lösung verarbeitet werden können.

Hinweis: Dieser Zieldefinition liegt eine bewusst gewählte Abstraktion zu Grunde: Rohdaten sollen von verschiedenen Quellen auf eine für die Unternehmenssituation geeignete Standardstruktur überführt werden. Diese wird dann für Berechnungen und Mappings genutzt, um z.B. fehlende Daten in dieser einheitlichen Struktur zu ergänzen. Diese hieraus geschaffene Datengrundlage wiederum bildet die Basis verschiedener

⁵ KNIME ist eine Open-Source-Plattform für Datenanalyse, maschinelles Lernen und die Entwicklung von Data-Science-Anwendungen. Die Abkürzung steht für "Konstanz Information Miner". Mit KNIME können Benutzer Daten aus verschiedenen Quellen importieren, diese manipulieren, analysieren und visualisieren. Die Plattform bietet eine benutzerfreundliche grafische Benutzeroberfläche, die es Datenanalysten und Wissenschaftlern ermöglicht, Datenpipelines zu erstellen, um komplexe Analyse- und Verarbeitungsaufgaben ohne Programmierkenntnisse durchzuführen. KNIME unterstützt zudem die Integration von Algorithmen für maschinelles Lernen und Künstliche Intelligenz. (basierend auf OpenAI/Chat GPT, 17.09.2023). Siehe auch <https://www.knime.com/> (17.09.2023).

systemabhängiger Import-Mappings für die unterschiedlichen Zielsysteme. Dieses Gedankenkonstrukt ist nicht limitiert auf einen Rechtsraum, ein Quellsystem oder ein Zielsystem (vgl. auch Kapitel 2).

Die Arbeitsgruppe ist sich bewusst, dass bereits Datenmodelle z.B. im Bereich der E-Bilanz oder SAF-T vorhanden bzw. im Entstehen sind (vgl. insbesondere § 147b AO). Ergänzend wird auf die Aktivitäten der IDSt-übergreifenden Arbeitsgruppe „Common Data Model“ verwiesen.

(2) Zu betrachtende Fragestellungen

- Lässt sich eine einheitliche Datenstruktur aufbauen, welche die automatisierte Verarbeitung innerhalb der jeweiligen Compliance-Lösung ermöglicht?
- Existieren Möglichkeiten einer system- bzw. herstellerübergreifenden Standardisierung?

(3) Methodik und Vorgehen

- Unter Nutzung der vorhandenen ETL-Umgebung soll die Bereitstellung der Daten in eine möglichst einheitliche Datenstruktur simuliert werden.

(4) Erwartetes Ergebnis

- Prozedere zur Überführung der bereitgestellten Daten in eine oder mehrere Strukturen, welche von der Compliance-Lösung verarbeitet werden können.

4.6 Technische Anbindung der Compliance-Systeme

(1) Zielsetzung

Die Zielsetzung dieser Phase besteht darin, technische Möglichkeiten zu beschreiben, wie die Ergebnisse, welche in der Datenschicht erzeugt werden, für eine Compliance-Lösung nutzbar gemacht werden können.

(2) Zu beantwortende Fragen

- Wie lassen sich Prozesse anstoßen, damit die Daten zur gewünschten Zeit, in der gewünschten Form und gewünschten Qualität zur Verfügung stehen?
- Was sind die möglichen Strategien und Technologien, die einem Unternehmen zur Verfügung stehen (SQL-Zugriff, Rest-Services, RFC, mittelbarer Datenzugang durch eine Data-Lake-Lösung, ETL oder ähnliche Ansätze)?
- Gibt es Standard-Methoden oder Best-Practice-Ansätze zur technischen Anbindung?

Hinweis: Häufig geben die Compliance-Lösungen Schnittstellenmethoden (Datenformate und Art des Imports, wie z.B. File-Import, RFC oder SQL-Aufruf) vor, die es ermöglichen, die Daten dann auch weiter zu verarbeiten. Die Unternehmensseite ist

dabei durch die Art der Quellen, Art der Verarbeitung sowie eigenen Sicherheitsaspekten geprägt (siehe Kapitel 4.3).

3) Methodik und Vorgehen

Bei der Betrachtung der Vorgehensweise sind mögliche Entscheidungen des Unternehmens für eine Compliance-Lösung zu berücksichtigen. Je nachdem kann sich die Methodik und das Vorgehen wie folgt unterscheiden:

- Im Fall, dass eine Tool-Entscheidung (noch) nicht vorliegt:
 - Eruierung von Lösungen/technischen Möglichkeiten mittels Recherche sowie gezielten Interviews von Softwareanbietern und Unternehmen.
 - Auswahl von möglichen Dienstleistern und Kooperationspartnern zur Mitwirkung im Rahmen eines „Proof-of-Concept“, zum Test der Methode und der Ergebnisse.
 - Anlage von einer Test-Instanz inkl. Testdaten.
 - Durchführung eines „Proof-of-Concept“ – Fokus: Zurverfügungstellung von Daten sowie deren Übermittlung und Verarbeitung.
- Im Fall, dass eine Tool-Entscheidung vorliegt:
 - Durchsicht der Systemdokumentation und der Handbücher.
 - Evaluierung vorhandener Datenfluss- und Schnittstellendokumentationen.
 - Anlage einer Test-Instanz inkl. Testdaten.
 - Durchführung des „Proof-of-Concept“ – Fokus: Zurverfügungstellung von Daten sowie deren Übermittlung und Verarbeitung.

Hinweis: Das Ziel besteht vornehmlich darin, die möglichen Funktionen zu kennen, sie zu bewerten, im Rahmen der Bewertung Stakeholder einzubinden und ein Konzept zur Umsetzung zu erarbeiten.

(4) Erwartetes Ergebnis

- Möglichkeiten und Methoden zur technischen Anbindung der Datenschicht an das Compliance-System einschl. Bewertung (Kosten-Nutzen).
- Dokumentierte Beispielprozesse.

4.7 Verarbeitung und Versendung durch die Compliance-Lösung

(1) Zielsetzung

Als Zielsetzung dieser Phase soll eine Bewertung der Umsetzbarkeit der Verarbeitung innerhalb der Compliance-Lösung erstellt und evaluiert werden bzw. die Umsetzbarkeit bei einer bereits vorhandenen Lösung untersucht werden.

Abschließend sollen die bereitgestellten und idealerweise standardisierten Daten in die Compliance-Lösung übernommen und eine automatisierte Verarbeitung simuliert

werden. Dabei sollen Erkenntnisse dazu gewonnen werden, inwieweit sich Daten automatisiert verarbeiten und entsprechend Steuererklärungen erstellen bzw. übermitteln lassen.

(2) Zu betrachtende Fragestellungen

- Was sind die bestehenden Verarbeitungs- und Automatisierungsmöglichkeiten der am Markt verfügbaren Lösungen bzw. der bereits genutzten Lösungen?
- Existieren ergänzende Automatisierungsmöglichkeiten außerhalb der Compliance-Lösung, um z.B. fehlende Funktionen zu ergänzen?

(3) Methodik und Vorgehen

- Das angestrebte Arbeitsergebnis soll durch eine Simulation mit Hilfe der bestehenden ETL-Umgebung – ggf. ergänzt um weitere Tools – erarbeitet werden.

Hinweis: In den vorherigen Schritten wurden Daten extrahiert, verarbeitet und in einen Standardformat bereitgestellt. Diese können technisch in die Compliance-Lösung übertragen oder importiert werden. Innerhalb dieser Phase soll geprüft werden, ob die Verarbeitung hin zu einer Steuererklärung und deren Versendung möglich sind. Im Fall fehlender Funktionen bietet es sich ggf. an, Alternativen wie „Robotic-Process-Automation“ (RPA)⁶ zu evaluieren.

(4) Erwartetes Ergebnis

- Prozedere der Verarbeitung und Versendung durch die Compliance-Lösung.
- Aufzeigen verschiedener Optionen einschließlich Bewertung (Kosten-Nutzen).
- Dokumentierte Beispielprozesse.

5. Automatisierte Steuererklärung unter Kontrollaspekten

Die automatisierte Steuererklärung ist integrativer Bestandteil des steuerlichen Kontrollumfeldes. Damit definieren die (Kontroll-)Anforderungen, die durch ein Internes Kontroll-System (IKS) für Steuern bzw. ein „Tax-CMS“⁷ definiert sind, zugleich einen übergeordneten Maßstab zur Beurteilung aller Prozessschritte einer automatisierten Steuererklärung. Begleitend zu den vorgenannten Phasen 1 bis 7 gilt es entsprechend die jeweiligen Kontrollaspekte in die Betrachtung einzubeziehen. In Anlehnung an **Abbildung 1** sollen dabei folgende Kontrollbereiche zum Tragen kommen:

⁶ RPA steht für Robotic Process Automation und bezieht sich auf eine Technologie, bei der Software-Roboter oder Bots eingesetzt werden, um menschliche Arbeitsabläufe in digitalen Geschäftsprozessen zu automatisieren. Diese Software-Roboter sind in der Lage, wiederholbare Aufgaben und Interaktionen mit digitalen Systemen und Anwendungen auf effiziente Weise auszuführen, ohne menschliches Eingreifen. (OpenAI/ChatGPT, 17.09.2023), siehe auch https://de.wikipedia.org/wiki/Robotic_Process_Automation (17.09.2023).

⁷ Vgl. hierzu auch White Paper des Fachausschuss V des IDSt – Steuerliches Kontrollumfeld, abrufbar unter: <https://idst.tax/veroeffentlichungen/>.

1. Die **Richtigkeit des Buchwerkes** bzw. der **Quelldaten** selbst (*System-Ebene 1*).
 - Hinweis: Da diese Fragestellung für die manuelle oder teil-automatisierte Steuererklärung identisch ist und generell gelöst werden muss, wird sie hier nicht weiter betrachtet.
2. Die **Datenübertragung von den Quellsystemen an die datenverarbeitende Schicht** (*Interface-Ebene 1*).
 - Kontrollaspekt 2.1 –
Richtigkeit und Vollständigkeit der identifizierten Quellen (Konzeptphase).
 - Kontrollaspekt 2.2 –
Sicherstellung des Prozesses zur Identifizierung von Änderungen der Quellsysteme und Daten (Konzeptphase und laufende Verpflichtung).
 - Kontrollaspekt 2.3 –
Vollständigkeit der übertragenden Daten (laufende Kontrolle eingebettet in den Verarbeitungsprozess).
 - Kontrollaspekt 2.4 –
Richtigkeit der übertragenen Daten (Laufende Kontrolle eingebettet in den Verarbeitungsprozess).
3. Die **Verarbeitung der Daten in der datenverarbeitenden Schicht** (*System-Ebene 2*).
 - Kontrollaspekt 3.1 –
Die **Richtigkeit bzw. Aktualität der hinterlegten Konfigurationen** und des **Mappings** (Konzeptphase und laufende Verpflichtung).
 - Kontrollaspekt 3.2 –
Die **Richtigkeit der Verarbeitungen und Berechnungen** (laufende Kontrolle eingebettet in den Verarbeitungsprozess).
4. Die **Datenübertragung von der datenverarbeitenden Schicht an das Compliance System** (*Interface-Ebene 2*).
 - Kontrollaspekt 4.1 –
Sicherstellung des Prozesses zur Identifizierung von Änderungen der Datenstrukturen der empfangenden Compliance-Lösung (Konzeptphase und laufende Verpflichtung).
 - Kontrollaspekt 4.2 –
Vollständigkeit der übertragenden Daten (laufende Kontrolle eingebettet in den Verarbeitungsprozess).
 - Kontrollaspekt 4.3 –
Richtigkeit der übertragenen Daten (laufende Kontrolle eingebettet in den Verarbeitungsprozess).

5. Die **Verarbeitung der Daten in der Compliance-Lösung** (*System-Ebene 3*).
 - Kontrollaspekt 5.1 –
Die **Richtigkeit bzw. Aktualität der hinterlegten Konfigurationen** und des **Mappings** (Konzeptphase und laufende Verpflichtung).
 - Kontrollaspekt 5.2 –
Die **Richtigkeit der Verarbeitungen und Berechnungen** (laufende Kontrolle eingebettet in den Verarbeitungsprozess).
6. Die **Datenübertragung von der Compliance-Lösung an die Finanzverwaltung** (*Interface-Ebene 3*).
 - Kontrollaspekt 6.1 –
Es muss sichergestellt werden, dass die Steuererklärungsdaten richtig übertragen werden.
Hinweis: Da diese Fragestellung für die manuelle oder teil-automatisierte Steuererklärung identisch ist und generell gelöst werden muss, wird sie hier nicht weiter betrachtet.
7. **Übergreifende Prozesskontrolle**
 - Kontrollaspekt 7.1 –
Es muss sichergestellt werden, dass stets erkennbar ist, welchen **Prozessualen Zustand die Verarbeitung in ihren einzelnen Schritten** hat (Meilenstein-Übersicht bzw. Visualisierung und laufende Kontrolle eingebettet in den Verarbeitungsprozess).
 - Kontrollaspekt 7.2 –
Es muss sichergestellt werden, dass **Fehlerfälle bzw. ausgelöste Kontrollen** zu einer Zuführung zu einem manuellen Prozess führen („Exception-Case“ - Meilenstein-Übersicht bzw. Visualisierung und laufende Kontrolle eingebettet in den Verarbeitungsprozess).
 - Kontrollaspekt 7.3 –
Es muss dokumentiert werden, wie **Fehlerfälle identifiziert und korrigiert** wurden und wie diese in Zukunft vermieden werden („Issue Log“ mit Beschreibung der Behebung [Gegenwart und Zukunft])
 - Kontrollaspekt 7.4 –
Es muss sichergestellt werden, dass die **System- und Prozessdokumentationen** erstellt und gepflegt werden (Konzeptphase und laufende Verpflichtung).
 - Kontrollaspekt 7.5 –
Bei Änderungen müssen die **Anforderungen der GoBD**, insbesondere an die **Nachvollziehbarkeit** und den **Datenzugriff im Rahmen von Betriebsprüfungen** erfüllt werden (Projekt- und Pflegeaufgabe).

(1) Zu betrachtende Fragestellungen

Wie sind die im Rahmen der automatisierten Steuererklärung etablierten Kontrollen im Kontext eines Tax CMS zu beurteilen bzw. die bestehenden zu erweitern?

- Wie kann deren Vollständigkeit und Richtigkeit auch im Zeitablauf gewährleistet werden?
- Gibt es Grenzen hinsichtlich der Kontrollierbarkeit bzw. der Automatisierung der Kontrollen?
- Wie können Standardkontrollen aussehen?

(2) Methodik und Vorgehen

- Die gewählten Methodiken sind für den gesamten Prozessfluss zu definieren und umzusetzen.
- Kontrollen können dabei automatisiert, teilautomatisiert oder manuell sein.
- Kontrollen können Teil der Konzeptionsphase, des laufenden Betriebes oder wiederkehrender Natur (periodisch oder in Zusammenhang mit Änderungen) sein.
- Mit Blick auf die in Kapitel 4 definierten Phasen lassen sich beispielhaft (nicht abschließend) folgende Anregungen geben:
 - **Datenübertragung** (ohne Abgabe der Steuererklärung): Eine quantitative und qualitative Richtigkeit kann etwa durch Referenzwerte (z.B. Datenvergleiche) und Plausibilitätsberechnungen (z.B. Check-Summen, Hash-Wert-Vergleiche) überprüft werden.
 - **Verarbeitung in der datenverarbeitenden Schicht**: Die einzelnen Berechnungen bzw. Berechnung von steuerlichen Werten, können hinsichtlich ihres Risikos beurteilt und in eine Matrix überführt werden. Die Kontrollen können durch Referenzwertvergleiche (z.B. Datenvergleiche, Mehrjahres-Referenzwertvergleich) und Plausibilitätsberechnungen innerhalb der datenverarbeitenden Schicht aufgebaut werden.
 - **Verarbeitung innerhalb der Compliance-Lösung**: Zwar entspricht das Kontrollumfeld grundsätzlich dem einer manuellen Erstellung, was jedoch fehlt ist die menschliche Interaktion und Beurteilung. Insofern sind sowohl die vorhandenen Prüfmethode der Compliance-Lösung, als auch die Elster-Prüfungen auszuführen und zu interpretieren. Darüber hinaus besteht bei Zugriff auf die Datenbank der Compliance-Lösung die Möglichkeit des weiteren Aufbaus von Kontrollen, z.B. als Teil der datenverarbeitenden Schicht.
 - **Übergreifende Prozesskontrollen**: Über die datenverarbeitende Schicht (und ihre Methoden, z.B. ETL) lassen sich der Zustand von Systemen und Daten abfragen und ein laufendes Berichtswesen inkl. Visualisierung darstellen. Über definierte Eventlogs lassen sich Datenbankeinträge generieren, die zur Dokumentation von Korrekturen und Fehlerbehebungen genutzt werden.

(3) Erwartetes Ergebnis

- Beschreibung notwendiger Kontrollen und Methoden in den verschiedenen automatisierten Prozessschritten.
- Beschreibung möglicher Umgebungen zur Ausführung dieser Kontrollen.
- Beschreibung der Verantwortlichkeiten.
- Dokumentation als Teil einer System- und Verfahrensdokumentation.

6. Abschließende Bewertung und Petitum

Die Automatisierung der Steuererklärung ist realisierbar und dies selbst ohne eine allumfassende Lösung eines einzelnen Anbieters. Die Kombination von vorhandenen Compliance-, ETL- und RPA-Lösungen sowie anderer Software-Produkte lassen sich gezielt dafür einsetzen, die automatisierte Steuererklärung im Unternehmen zu verwirklichen. Die Verwendung von Standard-Lösungen schafft dabei eine gewisse Herstellerunabhängigkeit und eröffnet zugleich die Option der Wiederverwendung von Methodiken für weitere Prozesse mit unterschiedlicher Skalierung.⁸ Die dabei zu erzielenden positiven Effekte hängen jedoch maßgeblich vom Agieren der beteiligten Parteien ab:

Gesetzgeber und Verwaltung sollten ein Umfeld schaffen, in welchem die automatisierte Steuererklärung nicht nur geduldet, sondern integraler Bestandteil der Steuerveranlagung ist. Dies könnte insbesondere durch eine engere Zusammenarbeit und den Austausch mit Softwareanbietern, der Zurverfügungstellung von standardisierten Berechnungsmethoden und nicht zuletzt durch die Realisierung der bereits vielfach zugesagten elektronischen Bescheid-Rückübertragung positiv befördert werden. Auch könnte das Teilen von Erfahrungswerten und Kontrollmethoden zwischen Steuerpflichtigen und Finanzverwaltung zu einer effizienten und zeitnäheren Umsetzung von Kontrollmechanismen in den Unternehmen führen.⁹ Dieses White Paper soll auch dazu dienen, diesen Dialog zu bereichern.

Ferner sollte darauf Wert gelegt werden, dass die beteiligten Behörden ihre Automatisierungs- und Digitalisierungsbemühungen synchronisieren, somit deren Umsetzungsgeschwindigkeit erhöhen und damit zugleich die Gefahr von Fehlinvestitionen auf Seiten der Unternehmen senken. Ein weiterer Baustein ist sicherlich die Etablierung eines standardisierten Datenformats, welches mit der Einführung des § 147b AO auf den Weg gebracht wurde.¹⁰

⁸ So können die Methoden, die auf eine (teil-)automatisierte Ersterstellung von Steuererklärungen hinführen, auch für spätere Korrekturprozesse, z.B. im Rahmen von Betriebsprüfungen, verwendet werden.

⁹ Voraussetzung hierfür wäre indessen wohl eine weniger restriktive Auslegung von § 88 Abs. 3 und Abs. 5 AO durch die Finanzverwaltung.

¹⁰ Vgl. hierzu auch White Paper des Fachausschuss IV des IDSt, abrufbar unter : <https://idst.tax/veroeffentlichungen/>.

Software-Hersteller sollten durch eine stärkere Unterstützung von automatisierten Prozessen und herstellerübergreifenden Standards sowie durch die Vereinheitlichung von Schnittstellen dazu beitragen, die Umsetzbarkeit zu vereinfachen.

Berater sollten durch daten- und prozessorientierte Beratungsansätze die entsprechenden Automatisierungsbemühungen flankieren und so die Qualität und Vollständigkeit steuerdeklaratorischer Prozesse vorantreiben.

Gelingt das Zusammenspiel zwischen Gesetzgeber, Verwaltung, Unternehmen, Softwareherstellern und Beratern ist der Weg zur flächendeckenden automatisierten Steuererklärung geebnet. Als IDSt werden wir uns dafür auch weiterhin verstärkt einsetzen und das Thema vorantreiben.