



Fallstudie: Digitalisierung von Prüfprozessen mit KI

David Israel

Thomas Domsch

Fallstudie: Digitalisierung von Prüfprozessen mit KI

Nicht erst seit der Finanzkrise nimmt die Bedeutung der Bereiche Audit und Compliance stetig zu. Allerdings ist eindeutig festzustellen, dass das Thema seitdem deutlich mehr Geschwindigkeiten bekommen hat.

Seitens des Regulators, der Aufsicht sowie der externen und internen Prüfer wächst der Druck, mit der zunehmenden Digitalisierung und der damit einhergehenden Beschleunigung der Geschäftstätigkeit schrittzuhalten. Der aktuelle, risikoadjustierte Prüfungsansatz setzt umfangreiches Expertenwissen voraus und basiert in der Regel auf einer intelligenten, aber stichprobenhaften Prüfung. Meist ist eine vollständige Analyse aufgrund der Datenmenge oder der begrenzten Ressourcen schlicht nicht möglich.

Im Rahmen dieses Umfelds haben wir die Commerzbank im Bereich Audit unterstützt, Lösungen für Tools zur Prozessautomatisierung und technischen Prüfungsunterstützung zu entwickeln.

Aufgrund eines engen Zeitplans von drei Monaten haben wir uns auf die folgenden zwei Themen fokussiert:

- Digitalisierung des Berechtigungsmanagements
- Automatisierter Abgleich von Vertragsdokumenten

Die entwickelten Lösungen sind allerdings nicht auf diese Themengebiete beschränkt, sondern ein Einsatz in anderen Bereichen ist genauso möglich. Im Folgenden werden die Herausforderungen, das gewählte Vorgehen und die entwickelten Lösungen zur näheren Erläuterung beschrieben und die damit erzielten Verbesserungen aufgezeigt.

Digitalisierung des Berechtigungsmanagements durch Automatisierung von Berechtigungsprüfungen

Herausforderung und Idee

Berechtigungsprüfungen der IT-Revision finden traditionell stichprobenartig statt und erfordern eine Vielzahl an manuellen Arbeitsschritten. Der zuständige Prüfer fordert dazu einen Report zu einer bestimmten Anwendung vom Berechtigungsmanagement an, bereitet die Daten auf und sucht nach Auffälligkeiten. Die genauen Kriterien und der Fokus der Prüfung sind dabei nicht immer von vornherein klar und können sich maßgeblich unterscheiden.

Es wurde die Idee entwickelt, typische Berechtigungsprüfungen und -auswertungen zu automatisieren und Datenanalysetechniken zur Effizienzsteigerung zu nutzen. Dies soll der IT-Revision ermöglichen, Prüfungen mit deutlich geringerem Zeitaufwand, regelmäßiger und umfangreicher durchzuführen.

Lösungsansatz und Vorgehen im Detail

Zur zielgenauen Entwicklung der Automatisierungslösung und zur Eingrenzung des Scopes für einen ersten Prototypen wurden zusammen mit der IT-Revision typische und relevante Berechtigungsprüfungen festgelegt. Diese umfassen z.B. Soll-Ist-Abgleiche, Prüfungen auf unzulässige Berechtigungsvergaben, wie Verstöße gegen die Funktionstrennungsvorschriften nach MaRisk¹, oder die Erzeugung einer Übersicht zur Verbreitung von bestimmten Berechtigungen. Außerdem wurde für den Prototypen ein Datenabzug aus einem der Berechtigungsmanagementsysteme als Datengrundlage für die Auswertungen

¹ https://www.bafin.de/SharedDocs/Veroeffentlichungen/DE/Rundschreiben/2017/rs_1709_marisk_ba.html (Absatz BTO), Abgerufen am 15.11.2019.

herangezogen. Die Anbindung der weiteren Systeme sowie von Anwendungen mit eigenem Berechtigungsmanagement soll anschließend bei zukünftigen Weiterentwicklungen des Tools erfolgen.

Die in Python entwickelte Lösung liest nun diesen Datenabzug vollautomatisch ein und bereitet die Berechtigungsdaten zur weiteren Verwendung auf. Daraufhin werden die definierten Prüfungen durchgeführt und die entsprechenden Auswertungen berechnet. Dies geschieht nach anfänglicher Spezifikation der Prüfungen und der zu prüfenden Systeme vollautomatisch. Zur Lauffähigkeit und Performanz des Prototyps auf einem herkömmlichen Arbeitsplatzrechner wurden außerdem diverse Optimierungen, u. a. zum Speicherbedarf, durchgeführt. Die Ergebnisse werden dann in verschiedenen Formaten zur Weiterverarbeitung bereitgestellt.

Ergebnisse

Die Ergebnisdaten mit gefundenen Abweichungen und Verstößen bzw. Übersichten zur Verbreitung von bestimmten Berechtigungen werden konsolidiert abgespeichert. Die Ergebnisse können somit maschinell weiterverarbeitet oder z.B. auf Dashboards visualisiert werden. Falls vom Anwender gewünscht, wird in einem weiteren Schritt eine formatierte Excelausgabe durch das Tool erzeugt. Diese enthält auch eine Übersichtsseite zur effizienten Auswertung der Ergebnisse durch den Nutzer.

Digital Auditing mit einem automatisierten Abgleich von Vertragsdokumenten

Herausforderung und Idee

Der Abgleich von Verträgen mit den zu ihnen erfassten Werten in einem Banksystem oder allgemein einer Datenbank ist traditionell ein zeitaufwendiger und ressourcenintensiver Prozess. Ein Mitarbeiter der internen Revision sucht den abzugleichenden Vertrag und in diesem die relevanten Stellen und Informationen

und gleicht die gefundenen Werte mit der Datenbank ab. Diese Prüfungshandlung soll durch den Einsatz von Datenanalyse-, Texterkennungs- und Machine-Learning-Techniken automatisiert werden. Dadurch wird sich eine erhebliche Effizienzsteigerung sowie eine deutlich breitere Abdeckung der geprüften Verträge erhofft.

Für die Fallstudie standen mehr als 10.000 Dokumente zur Verfügung und diese waren wiederum mal als Text oder nur als Bild abgelegt und lagen in verschiedensten Formaten, Dateitypen und Scanqualitäten vor.

Lösungsansatz und Vorgehen im Detail

Da der Datensatz unterschiedliche Dokumenttypen beinhaltet, von denen auch nicht alle für uns relevant sind, wurde entschieden, eine Vorklassifikation mit Hilfe von Metadaten vorzunehmen. Zu diesem Zweck und zur einfacheren Handhabung der Dokumente werden die gesammelten Metadaten außerdem in einem Excel-Sheet zur Verfügung gestellt. Zur Zusammenstellung dieses Metadaten-Sheets wurden zuerst die verfügbaren Metadaten aus der Dokumentendatenbank extrahiert. Neben diversen IDs und Zeitstempeln war hier auch bereits eine rudimentäre Unterscheidung nach Dokumenttyp vorhanden. Zusätzlich konnten wir weitere Metadaten aus den Dokumentendateien selbst auslesen. Der gesamte Prozess des Sammelns und Aufbereitens der Metadaten läuft nach Vorkonfiguration maschinell und automatisiert mittels eines selbstgeschriebenen Python-Skripts. Das resultierende Excel-Sheet erleichtert die Analyse der Dateien erheblich und ermöglicht eine Filterung oder Suche nach bestimmten Dokumenten bzw. nach bestimmten Eigenschaften.






Metadaten-Sheet

- ✓ Handling der Dokumente stark vereinfacht
- ✓ Überblick über Kategorien und Umfang der Dokumente
- ✓ Filterung oder Suche nach bestimmten Dokumenten bzw. nach bestimmten Eigenschaften
- ✓ Einfache Vorklassifikation der Dokumente möglich

Die eigentlichen Verträge enthalten alle ein Factsheet, welches die relevanten Größen in tabellarischer Form auflistet. Für den Abgleich der Werte mit dem System ist es möglich, sich auf dieses Factsheet zu beschränken. Dieses ist in den gescannten Dokumenten allerdings nicht immer an gleicher Stelle zu finden. Zur Suche dieser für uns relevanten Seiten in den teils umfangreichen Vertragsdokumenten haben sich textbasierte Verfahren als wenig geeignet herausgestellt. Zuverlässige Texterkennung in gescannten Dokumenten benötigt eine relativ hohe Auflösung und üblicherweise noch diverse Vorverarbeitungsschritte, was zu einer langen Laufzeit des gesamten Prozesses führt. Die Verarbeitung von mehreren tausend Dokumenten war so nicht mehr in akzeptabler Zeit möglich.

Da sich die tabellarischen Factsheets optisch deutlich von den übrigen Seiten der Verträge unterscheiden, haben wir uns dazu entschieden, eine Seitenklassifikation auf niedrig aufgelösten Bildern mit Hilfe eines Bilderkennungsverfahrens durchführen. Dazu wurde von uns ein Convolutional Neural Network (CNN) entwickelt, welches die Factsheets von den übrigen Seiten der Dokumente unterscheidet. Zum Training des CNNs wurden die vorher mittels Metadaten identifizierten Verträge im Standardformat mit bekannter Seitenabfolge verwendet und zusätzlich synthetische Trainingsdaten durch Rotation und Spiegelung der Seiten sowie zufälliger Änderung von Kontrast und Helligkeit erzeugt. Ergebnis ist ein trainiertes Modell, welches in der Lage ist, die Seiten der Vertragsdokumente optisch zu klassifizieren und damit auch bei ungewöhnlicher Dokumentenstruktur die tabellarischen Factsheets zu finden.

Seitenklassifikation mit Convolutional Neural Network (CNN)

-  Textbasierte Verfahren mit zu langer Laufzeit bei vollständiger Analyse von umfangreichen Dokumenten
-  Daher Verwendung von Bilderkennungsverfahren zur Unterscheidung der Seiten
-  Training auf Verträgen im Standardformat mit bekannter Seitenabfolge
-  Zusätzlich synthetische Trainingsdaten erzeugt durch Rotation und Spiegelung der Seiten sowie zufälliger Änderung von Kontrast und Helligkeit
-  Ergebnis: maschinelles Finden der Factsheets auch bei ungewöhnlicher Dokumentenstruktur möglich

Nun kann im letzten Schritt mittels Optical Character Recognition (OCR) zu jedem gescannten Vertrag gezielt der Text auf den zuvor gefundenen Factsheets ausgelesen und anschließend weiterverarbeitet werden. Dies bedeutet einen erheblich geringeren Zeit- und Ressourcenaufwand im Vergleich zu einer rein textbasierten Suche über den vollständigen Inhalt der Dokumente.

Optical Character Recognition (OCR)

- ✓ Texterkennung in gescannten Dokumenten möglich
- ✓ Gute Ergebnisse nach geeigneter Vorverarbeitung der Seiten
- ✓ Auslesen der relevanten Vertragsdaten unter Ausnutzung der Tabellenstruktur der Factsheets

Ergebnisse

Mit dem entwickelten Prototypen ist es möglich, aus einer großen Menge an gescannten Vertragsdokumenten systematisch die relevanten Vertragswerte auszulesen und einen maschinellen Abgleich mit den Einträgen in einer beliebigen Datenbank durchzuführen. Durch den hohen Automatisierungsgrad des gesamten Prozesses kann eine durch einen menschlichen Prüfer mit herkömmlichen Methoden nicht zu bewerkstelligende Menge an Dokumenten effizient analysiert und abgeglichen werden.

Übergreifende Anwendung der Ergebnisse

Die beiden behandelten Themen sind nur beispielhafte Ausschnitte aus einer Vielzahl von anfallenden Tätigkeiten und Analysen. Trotzdem verdeutlicht das hier aufgezeigte Vorgehen, dass zur Digitalisierung meist individuelle Lösungen pro Prozess und ggf. sogar pro Arbeitsschritt nötig sind.

Zur Umsetzung solcher individuellen Lösungen brachten unter anderem unsere folgenden Kernkompetenzen einen maßgeblichen Mehrwert zur effizienten Zielerreichung:

- Schnelle, gründliche und bereichsübergreifende Prozessanalyse
- Detailliertes Prozessverständnis, insb. beim Design zur Digitalisierung und/oder Automatisierung
- Technische und fachliche Kompetenzen zur Entwicklung eines Prototypen (insb. im Zusammenhang mit der MaRisk und BAIT)
- Entwicklung von stabilen und individuellen Best Practice Lösungen

Die hier vorgestellten Lösungen bieten darüber hinaus weitreichende Anwendungsmöglichkeiten. Die folgenden aktuellen Themen stehen beispielhaft dafür:

- Vertragsreview im Zusammenhang mit der IBOR-Transition im Rahmen der Benchmark-Regulierung
- Dokumentenanalyse im Rahmen einer automatisierten Kreditvergabe
- Einsatz im Rahmen der Geldwäscheprüfung
- Digitalisierung der Systemerfassung und/oder von Kontrollprozessen
- Dokumentenkontrolle im Rahmen der Compliance- oder Personaltätigkeit
- Entscheidungsfindung und Analysetätigkeit auf der Basis externer Dokumente

Mit unserer weitreichenden Erfahrung und tiefem themenübergreifenden Know-how sind wir in der Lage, aus den vielen zur Verfügung stehenden Werkzeugen das individuell passende auszuwählen. Dadurch resultieren effiziente Lösungen,

Fallstudie: Digitalisierung von Prüfprozessen mit KI

die sowohl flexibel einsetzbar als auch beliebig skalierbar sind. So können wir einen maßgeblichen Mehrwert leisten, um von den Vorteilen der Digitalisierung und möglicher Prozessautomatisierungen im Kleinen wie im Großen zu profitieren.



Wie den vorgestellten Lösungen zu entnehmen ist, konnte Finbridge aufgrund von großer Erfahrung und Know-how in diesem Bereich ohne langwierige Einarbeitung direkt unterstützen und entscheidend dazu beitragen, die Entwicklungszeit zu verkürzen und die Qualität der Lösungen zu steigern.

Unser Angebot

Die hier vorgestellten Techniken und Lösungsansätze bieten umfangreiche Anwendungsmöglichkeiten zur Lösung einer Vielzahl von typischen Problemen in diversen Bereichen. Gern beraten wir Sie im Rahmen einer Analyse bezüglich möglicher Lösungsvarianten zur Effizienzsteigerung in den unterschiedlichsten Bereichen, z.B. Audit und Compliance. Darauf aufbauend unterstützen wir Sie ebenfalls gern bei der Entwicklung von individuellen Automatisierungslösungen sowie bei der anschließenden Installation und Einbettung in vorhandene Prozesse.

Unser Beratungsansatz fokussiert sich bei der Umsetzung auf die optimale Einbettung der automatisierten Analyse- und Prüfungstätigkeiten in Ihre vorhandene Prozesslandschaft. Die Schulung interner Mitarbeiter und eine bedarfsgerechte Anwendungsbetreuung sind für uns ebenso selbstverständlich wie ein kundenorientierter Lösungsansatz und die Umsetzung eventuell benötigter individueller Gestaltungen. Gern unterstützen wir darüber hinaus im Projekteinsatz oder/und dem regelmäßigen Betrieb.

Wir hoffen, Ihr Interesse an unserer Beratung geweckt zu haben und freuen uns auf Ihre Kontaktaufnahme!

Team



David Israel
Financial Engineer
Financial Engineering
[eMail](#) | [LinkedIn](#) | [Xing](#)



Thomas Domsch
Senior Expert
Financial Engineering
[eMail](#) | [LinkedIn](#) | [Xing](#)



Insights und Trends



Finbridge GmbH & Co. KG
Louisenstraße 100
61348 Bad Homburg v. d. H.
www.finbridge.de