

Mittelstand-Digital  
**Zentrum  
Franken**



## Digitalisierungsprojekt

Automatisierung von europaweiten Ausschreibungsprozessen  
für Büroeinrichtung mithilfe von KI

## Inhalt

Abbildungsverzeichnis .....	2
1. Einleitung .....	3
2. Ziele und Aufgabenstellung .....	5
3. Organisation und Ablauf des Projekts.....	6
3.1. Beteiligte .....	6
3.2. Roadmap.....	8
3.3. Readiness Check .....	9
4. Rahmenbedingungen.....	10
4.1. Was ist eine europaweite Ausschreibung? .....	10
4.2. Prozesses zur Erstellung der Ausschreibungsunterlagen .....	11
4.3. Datenstruktur und Datenquellen verstehen .....	13
5. Umsetzung der ausgewählten Prozessschritte .....	16
5.1. Stückzahlen auszählen .....	16
5.2. Mengengerüst.....	20
5.3. Technische Vorbeschriebe .....	22
5.4. User interface .....	27
6. Ergebnisse und Ausblick.....	29
Literaturverzeichnis .....	30

## Abbildungsverzeichnis

Abbildung 1: Planung von Büroflächen (AI generiert mit Adobe Firefly) .....	3
Abbildung 2: Roadmap für den Projektablauf (eigene Darstellung).....	8
Abbildung 3: Fünf Dimensionen des Reifegrads (eigene Darstellung in Anlehnung an (Hellge, 2024)) .....	9
Abbildung 4: Ergebnisse des Readinesschecks für die AURA GmbH .....	9
Abbildung 5: Prozess der Erstellung von Ausschreibungen (Eigene Darstellung) .....	11
Abbildung 6: Grundriss der Aura in AutoCAD (Aura GmbH) .....	14
Abbildung 7: Auszug aus Warenwirtschaftssystem.....	15
Abbildung 8: Graphische Darstellung der Möbel .....	15
Abbildung 9: Grundriss aus einem Planungsprojekt der Aura (Aura GmbH).....	16
Abbildung 10: Lösungsansätze zum Erfassen der Stückzahlen aus einem Grundriss (Eigene Darstellung) .....	17
Abbildung 11: Analyse der Metadaten Layer (rechts) und Elemente (links) aus dem Grundriss ...	19
Abbildung 12: Pipeline zur Automatisierung der Erstellung eines Mengengerüsts in einem Ausschreibungsprozess (Eigene Darstellung).....	21
Abbildung 13: Pipeline zur Automatisierung der Erstellung des technischen Vorbeschriebs in einem Ausschreibungsprozess (Eigene Darstellung).....	23
Abbildung 14: RAG Mechanismus (T. Honroth, 2024) .....	24
Abbildung 15: Geplantes Userinterface (eigene Darstellung).....	28

## 1. Einleitung



Abbildung 1: Planung von Büroflächen (AI generiert mit Adobe Firefly)

Der Projektbericht befasst sich mit der Entwicklung und Einführung eines Systems zur Automatisierung europaweiter Ausschreibungen für ein Architekturbüro, das sich auf die Planung von Büroflächen spezialisiert hat. Ausschreibungen spielen sowohl im öffentlichen als auch im privaten Beschaffungswesen eine zentrale Rolle, da sie Projekte transparent und wettbewerbsorientiert vergeben und Unternehmen die Möglichkeit bieten, sich um Aufträge zu bewerben. Insbesondere bei anspruchsvollen Bauvorhaben, wie der Gestaltung von Bürogebäuden, erfordert der Ausschreibungsprozess eine sorgfältige Vorbereitung der Unterlagen sowie eine präzise Abstimmung der Anforderungen und Spezifikationen. Für Architekturbüros stellt dies eine besondere Herausforderung dar: Neben einer Fülle an Informationen müssen länderspezifische Vorgaben berücksichtigt und strikte Fristen eingehalten werden. Dieser aufwendige Prozess birgt zudem ein hohes Fehlerrisiko. Um den administrativen Aufwand zu senken und die Qualität der Unterlagen zu gewährleisten, kommt zunehmend Künstliche Intelligenz (KI) zum Einsatz. KI-gestützte Lösungen ermöglichen eine automatisierte Erstellung und Verwaltung von Ausschreibungen, was nicht nur Zeit spart, sondern auch die Einhaltung relevanter Vorschriften sicherstellt und die Fehleranfälligkeit reduziert.

Die [Aura GmbH – Innovative Arbeitswelten](#) aus Nürnberg ist ein erfahrener Partner für innovative Büro- und Objekteinrichtungen. Mit über 30 Jahren Erfahrung bietet das Unternehmen maßgeschneiderte Raumkonzepte, die Funktionalität, Ästhetik und Effizienz verbinden. Durch strategische Planung und individuelle Beratung schafft Aura moderne Arbeitswelten, die sowohl

Produktivität als auch Mitarbeitermotivation fördern. Dabei wird auf ergonomische Lösungen und eine nachhaltige Optimierung der Büroflächen Wert gelegt. Es ist jetzt der richtige Zeitpunkt, den nächsten Schritt zu gehen und durch den Einsatz von KI die Effizienz und Innovationskraft weiter zu steigern. Die Automatisierung dieser Aufgaben hat für das Architekturbüro eine große Bedeutung, da sie es ermöglicht, den Fokus auf die kreative und funktionale Bürogestaltung zu legen – die eigentliche Kernkompetenz. Gleichzeitig sorgt eine effizientere Bearbeitung von Ausschreibungen dafür, dass Projekte schneller realisiert werden können. Dies steigert nicht nur die Zufriedenheit der Auftraggeber, sondern erhöht auch die Wettbewerbsfähigkeit auf dem europäischen Markt.

## 2. Ziele und Aufgabenstellung

Die Erstellung von Ausschreibungen ist mit erheblichem Arbeitsaufwand und einem beträchtlichen Fehlerrisiko verbunden. Besonders zeitintensiv gestaltet sich der Prozess, da fehlerhafte Dokumente im schlimmsten Fall dazu führen können, dass die Ausschreibung komplett neu aufgesetzt werden muss. Vor diesem Hintergrund hat sich Aura dazu entschieden, gemeinsam mit dem Mittelstand-Digitalzentrum Franken nach KI-gestützten Lösungen zu suchen, um diesen Herausforderungen effizient zu begegnen. Im Erstgespräch mit Aura ergaben sich folgende Fragestellungen:

1. Welche Möglichkeiten gibt es die Ausschreibungsprozesse mithilfe von KI zu unterstützen?
2. Wo kann KI grundsätzlich eingesetzt werden?
3. Wie können die Ausschreibungen mithilfe der KI neutralisiert werden?
4. Wie können Medienbrüche überwunden werden?

### Zieldefinition:

„Automatisierung von europaweiten Ausschreibungsprozessen für Büroeinrichtung mithilfe von KI“

Im Einzelnen verfolgt das Projekt folgende spezifische Ziele:

1. **Prozessanalyse und Dokumentation:** Ausschreibungen sind komplexe Vorgänge, die aus vielen Einzelschritten bestehen. Diese Schritte müssen präzise erfasst und in einem Flow-Diagramm übersichtlich dargestellt werden, um den Gesamtprozess klar zu strukturieren.
2. **Bedarfsanalyse für Künstliche Intelligenz:** Es wird ermittelt, in welchen Prozessabschnitten der Einsatz von KI sinnvoll ist und welche Aufgaben alternativ durch andere Digitalisierungs- oder Automatisierungsmaßnahmen effizient unterstützt werden können.
3. **Auswahl und Umsetzung von Lösungen:** Die vielversprechendsten Ansätze werden ausgewählt und mit Python implementiert, um eine robuste und anpassungsfähige Lösung zu gewährleisten.
4. **Entwicklung eines Benutzerinterfaces:** Zur Verbesserung der Benutzerfreundlichkeit wird ein Entwurf für ein intuitives User Interface entwickelt, das eine einfache und effiziente Bedienung des Systems ermöglicht und in einem nachgelagerten Projekt weiterentwickelt werden kann.

Alle Schritte und Ergebnisse des Projekts wurden systematisch in einem [GitHub-Repository](#) dokumentiert, um die Nachvollziehbarkeit und Reproduzierbarkeit der Forschung sicherzustellen.



### 3. Organisation und Ablauf des Projekts

Um die erfolgreiche Umsetzung dieses Projekts zu gewährleisten, ist eine klare und strukturierte Vorgehensweise notwendig. In diesem Abschnitt werden die organisatorischen Aspekte des Projekts vorgestellt, um ein Verständnis für den Ablauf und die beteiligten Akteure und die Organisation zu schaffen.

#### 3.1. Beteiligte

Ein wesentlicher Faktor für den Projekterfolg ist die Zusammenarbeit der beteiligten Akteure. In diesem Abschnitt werden die wichtigsten Beteiligten des Projekts vorgestellt:



**Eileen Badaou**

[AURA Innovative Arbeitswelten GmbH](#)

Leitung Kommunikation, Social-Media-Marketing, People & Culture

Projektorganisation und -koordination. Bindeglied zwischen den technischen und fachlichen Bereichen. Vorstellung des Istzustandes sowie der Anforderungen und Ziele des Projekts.



**Gunnar Bauer**

[AURA Innovative Arbeitswelten GmbH](#)

Geschäftsführender Gesellschafter

Strategie und Visionen. Festlegung der inhaltlichen Ziele und Ansprechpartner für den Bereich Ausschreibungen.



**Lukas Körber**

[AURA Innovative Arbeitswelten GmbH](#)

Sachbearbeiter im Innendienst

Projektinitiator und Ansprechpartner für den Fachbereich Innendienst mit Spezialisierung auf das Thema Ausschreibungen.



## **Fabio Günthert**

[AURA Innovative Arbeitswelten GmbH](#)

Fachplaner

Ansprechpartner für den Fachbereich Planung und IT. Einordnung und Beratung im Bereich der planungsrelevanten technischen Umsetzbarkeit.



## **Sandra Nuiß**

[Mittelstand Digitalzentrum Franken](#)

AI Strategist und Research Asisstant

Wissenschaftliche Mitarbeiterin mit dem Schwerpunkt Large Language Modells. Zuständig für technische Umsetzung des Projekts.



### 3.2. Roadmap

Die Projekt-Roadmap diente als zentraler Leitfaden für die gesamte Laufzeit des Projekts und bot einen klaren Überblick über die wichtigsten Meilensteine und Arbeitspakete. Durch ihre strukturierte Darstellung des zeitlichen Ablaufs wurden sowohl die Arbeitsphasen als auch kritische Deadlines transparent abgebildet. Dies ermöglichte nicht nur eine gezielte Koordination der einzelnen Teams, sondern stellte sicher, dass alle Projektbeteiligten stets über den aktuellen Fortschritt informiert waren. Regelmäßige Updates halfen dabei, potenzielle Engpässe frühzeitig zu identifizieren und die nötigen Ressourcen rechtzeitig bereitzustellen, um den erfolgreichen Projektabschluss sicherzustellen:

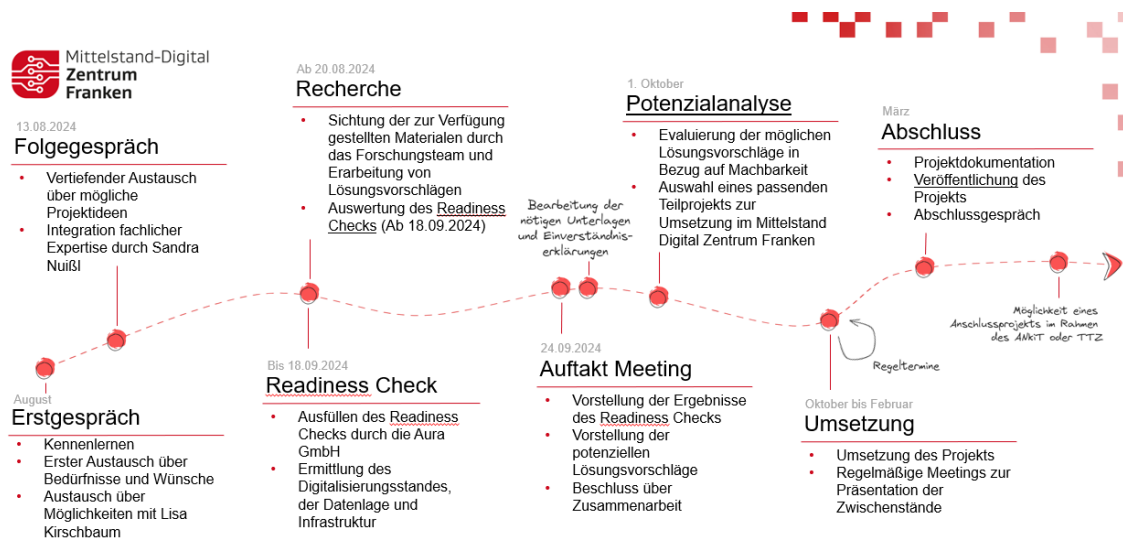


Abbildung 2: Roadmap für den Projektablauf (eigene Darstellung)

Das Projekt startete im August 2024 mit einem Auftakttreffen, bei dem alle Beteiligten einander vorgestellt wurden und ihre Erwartungen an das Vorhaben austauschten. In einer weiteren Besprechung am 13. August wurden erste Ideen konkretisiert und weiterentwickelt. Ab dem 20. August begann eine intensive Recherchephase, die am 18. September durch den „Readiness Check“ der Aura GmbH ergänzt wurde. Dieser diente dazu, den Digitalisierungsstand des Partners systematisch zu bewerten. Nach der Auswertung der Ergebnisse am 24. September wurde die Entscheidung für die weitere Zusammenarbeit getroffen. Die Umsetzung des Projekts fand zwischen Oktober 2024 und Februar 2025 statt, begleitet von regelmäßigen Meetings zur Überprüfung des Fortschritts und zur Präsentation der Zwischenergebnisse.

Das Projekt wurde am 24. März 2025 erfolgreich abgeschlossen. Die Abschlussphase umfasste die detaillierte Dokumentation und Veröffentlichung der Projektergebnisse sowie ein abschließendes Gespräch, um Erkenntnisse und weitere Perspektiven zu diskutieren. Zudem besteht die Option, das Projekt im Rahmen von ANkIT oder TTZ weiterzuführen, um die erzielten Ergebnisse auszubauen und neue Entwicklungen zu fördern.

### 3.3. Readiness Check

Um die Effektivität der implementierten Methoden frühzeitig zu bewerten, wurde ein initialer [Readiness Check](#) durchgeführt. Dieser speziell für mittelständische Unternehmen entwickelte Test hilft, potenzielle Herausforderungen frühzeitig zu erkennen und erforderliche Anpassungen vorzunehmen. Er bietet eine fundierte Grundlage, um gezielt passende Angebote des Mittelstand-Digital Zentrums Franken auszuwählen. Der Check bewertet fünf zentrale Dimensionen, die für eine erfolgreiche digitale Transformation entscheidend sind:



Abbildung 3: Fünf Dimensionen des Reifegrads (

Anhand von 25 Fragen und allgemeinen Angaben zum Unternehmen wurde

folgender Reifegrad ermittelt, welcher in der Abbildung 5 visualisiert wurde: Die Analyse zeigt, dass der Bereich „Mitarbeiter“ den höchsten Reifegrad aufweist, gefolgt von den Kategorien „Strategie“, „Technologien“ sowie „Produkt und Dienstleistung“. Im Gegensatz dazu birgt der Bereich „Organisation und Prozesse“ noch erhebliches Potenzial für Optimierungen durch Digitalisierung, Automatisierung und den Einsatz künstlicher Intelligenz. Im Rahmen des Digitalisierungsprojekts wird hier angesetzt, um innovative Ansätze zur Automatisierung von Prozessen zu entwickeln und die Effizienz weiter zu steigern.



Abbildung 4: Ergebnisse des Readinesschecks für die AURA GmbH

## 4. Rahmenbedingungen

Vor dem Start eines Projekts ist eine umfassende Analyse erforderlich, um das Thema vollständig zu verstehen. Dabei gilt es, die verfügbaren Datenquellen zu identifizieren, deren Qualität zu bewerten und Möglichkeiten zur Nutzung sowie Zugänglichkeit dieser Quellen zu klären. Zudem müssen die relevanten Prozesse detailliert skizziert und die wechselseitigen Zusammenhänge zwischen den einzelnen Komponenten systematisch erfasst werden.

Aus diesem Grund werden folgende Punkte im Rahmen dieses Kapitels beleuchtet:

- Was ist eine europaweite Ausschreibung?
- Wie sieht der Prozessablauf bei der Erstellung einer europaweiten Ausschreibung aus?
- Welche Datenquellen werden für die Erstellung der Ausschreibungen herangezogen und wie können diese an die Pipeline angebunden werden?

### 4.1. Was ist eine europaweite Ausschreibung?

Eine europaweite Ausschreibung ist ein Vergabeverfahren, das erforderlich ist, wenn der geschätzte Auftragswert eines Projekts über den von der EU festgelegten Schwellenwerten liegt. Dieser Prozess ist für alle EU-Mitgliedstaaten verbindlich und soll sicherstellen, dass öffentliche Aufträge fair, transparent und wettbewerbsorientiert vergeben werden. Insbesondere bei größeren Aufträgen wie der Ausstattung von Büroflächen mit Möbeln, wird dieser Mechanismus genutzt, um eine breite Marktteilnahme und optimale Preis-Leistungs-Verhältnisse zu gewährleisten (Maisenbacher, 2024).

1. **Bedarfsanalyse und Leistungsbeschreibung:** Die Anforderungen an die zu liefernden Büromöbel werden genau definiert.
2. **Veröffentlichung der Ausschreibung:** Diese erfolgt im Supplement zum Amtsblatt der Europäischen Union (TED), sodass sie europaweit zugänglich ist.
3. **Einreichung der Angebote:** Unternehmen aus der EU können ihre Angebote innerhalb einer vorgegebenen Frist einreichen.
4. **Eignungs- und Zuschlagskriterien:** Die Angebote werden basierend auf zuvor definierten Kriterien wie Preis, Qualität und Nachhaltigkeit bewertet.
5. **Zuschlag:** Der Vertrag wird an den Anbieter vergeben, der die Kriterien am besten erfüllt.
6. **Dokumentation:** Der gesamte Prozess wird dokumentiert, um Transparenz und Nachvollziehbarkeit sicherzustellen.

Das Digitalisierungsprojekt konzentriert sich auf einen spezifischen Teilbereich des ersten Prozessschritts, der „Bedarfsanalyse und Leistungsbeschreibung“. Ziel ist es, die Prozesse der Leistungsbeschreibung zu automatisieren und dadurch zu beschleunigen, um eine effizientere und fehlerfreie Abwicklung zu gewährleisten.

## 4.2. Prozesses zur Erstellung der Ausschreibungsunterlagen

Zu Beginn des Projekts wurde der gesamte Prozess zur Erstellung einer Ausschreibung aufgezeichnet, um diesen besser zu verstehen und die Möglichkeiten für den Einsatz von künstlicher Intelligenz zu identifizieren. Der Prozess ist in der untenstehenden Graphik veranschaulicht:

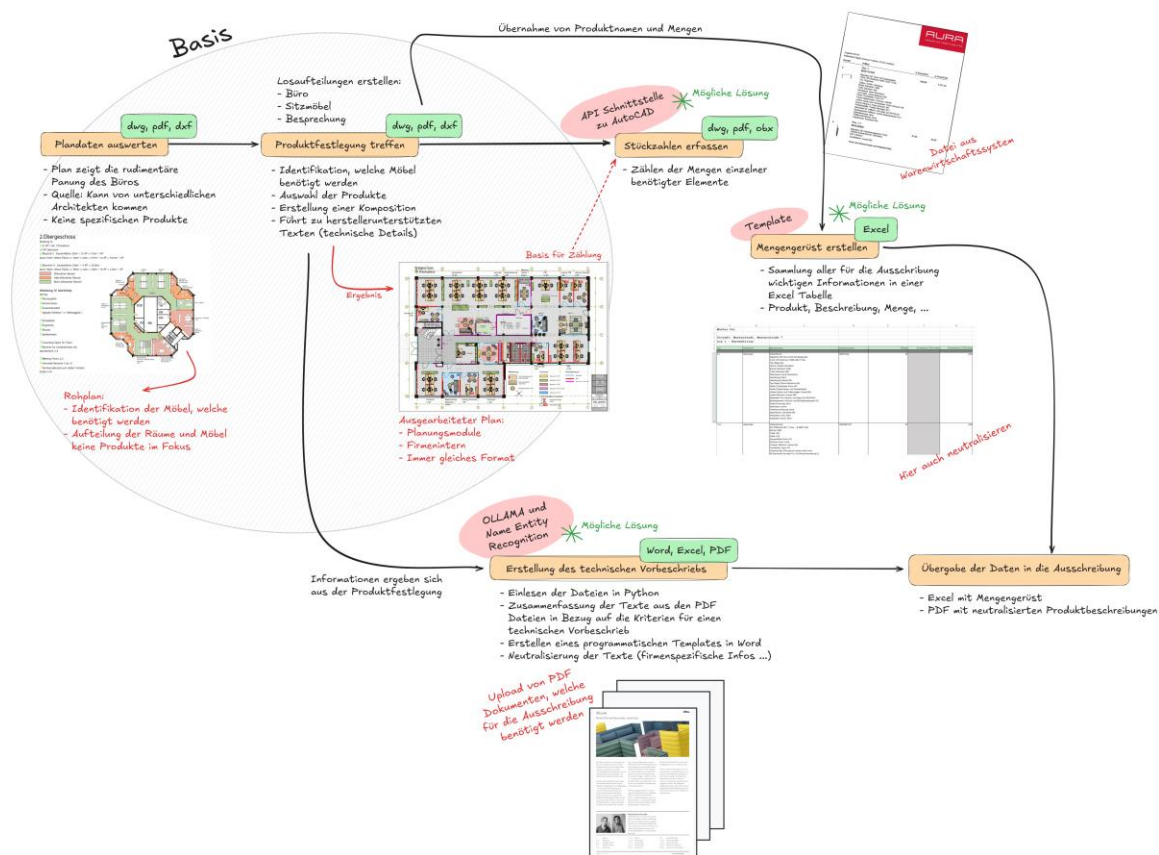


Abbildung 5: Prozess der Erstellung von Ausschreibungen (Eigene Darstellung)

Die Unterschritte, wie sie im Flussdiagramm dargestellt sind, können zum großen Teil im Rahmen des Digitalisierungsprojekts durch den Einsatz von Künstlicher Intelligenz oder einer programmatischen Implementierung einer Pipeline automatisiert werden.

### Planung und Produktfestlegung:

Der erste Schritt im Planungsprozess umfasst die grundlegende Büroplanung, bei der festgelegt wird, welche Möbel und Büroeinrichtungen benötigt werden. Architekten und Planungsteams konzentrieren sich darauf, die Fläche optimal zu nutzen und sicherzustellen, dass alle funktionalen und ergonomischen Anforderungen berücksichtigt werden. In dieser Phase werden noch keine spezifischen Hersteller oder Produkte berücksichtigt. Stattdessen werden

geometrische Formen verwendet, um die geplanten Möbelstücke, wie Tische und Stühle, zu visualisieren. Nachdem der Grundriss erstellt ist, erfolgt die Auswahl konkreter Produkte, die für das Büro geeignet sind. Diese Auswahl wird in einem detaillierteren Plan und dem Warenwirtschaftssystem dokumentiert, der als Grundlage für die Erstellung weiterer Dokumente dient, die im Ausschreibungsprozess benötigt werden.

*Dieser Prozessschritt kann im Rahmen des Digitalisierungsprojekts nicht berücksichtigt werden, weshalb die beiden vorliegenden Grundrisse als festgelegte Ausgangsbasis dienen. Sie bilden den Input für die weiteren Schritte im Projekt, bei denen die nachfolgenden Prozesse automatisiert und optimiert werden.*

#### Erfassung der Stückzahlen:

Nachdem der Grundriss zusammen mit den zugehörigen Produktfestlegungen erstellt wurde, bildet dieser die Grundlage für die Erfassung der benötigten Stückzahlen. Bisher erfolgte diese Erfassung manuell, indem die Stückzahlen direkt aus den Grundrissen abgelesen wurden. Dieser Prozess war jedoch mit erheblichem Arbeitsaufwand verbunden und birgt das Risiko von Fehlern. Das Ergebnis dieses Schrittes ist eine detaillierte Liste, die die Bezeichnungen der Möbelstücke und deren Anzahl in den verschiedenen Bereichen des Projekts aufführt. Diese Liste stellt sicher, dass für andere Arbeitsschritte die genauen Mengen und spezifischen Möbelprodukte vorliegen.

*Im Rahmen des Digitalisierungsprojekts wurde zunächst eine Lösung mit Computer Vision und OCR-Algorithmen in Erwägung gezogen, um die Möbelstückerkennung aus den Grundrissen zu automatisieren, was jedoch mit Fehlerpotenzial und hohem Implementierungsaufwand verbunden war. Stattdessen wurde die Python-API von AutoCAD genutzt, um direkt die Metadaten aus den Grundrissen auszulesen, was eine präzisere und ressourcenschonendere Lösung darstellt.*

#### Erstellung des technischen Vorbeschriebs:

Die Erstellung des technischen Vorbeschriebs ist einer der aufwendigsten Schritte im gesamten Prozess. Dieser folgt einer standardisierten Struktur und beschreibt detailliert die technischen Kriterien jedes einzelnen Möbelstücks, wie Abmessungen, Materialien, Farben, Serviceaspekte sowie Nachhaltigkeitsmerkmale und schematische Skizzen. Für jedes Möbelstück muss ein separates Dokument angelegt werden. Die benötigten Informationen stammen aus verschiedenen Quellen, wie Herstellerspezifikationen in PDF-Dateien, Webseiten oder internen Portalen, und müssen stichpunktartig zusammengefasst und in eine Vorlage übertragen werden. Wichtig ist, dass keine herstellereinspezifischen Informationen in den Dokumenten enthalten sein dürfen, weshalb eine Neutralisierung der Daten erforderlich ist.

*Dieser Prozessschritt konnte im Rahmen des Digitalisierungsprojektes erfolgreich umgesetzt werden: Ein standardisiertes Template wurde mit Python erstellt, während die Quelldateien durch ein Sprachmodell zusammengefasst und neutralisiert wurden. Die neutralisierten Daten wurden anschließend in das Template integriert, um eine effiziente und fehlerfreie Erstellung der technischen Vorab-Vorträge sicherzustellen.*

#### Erstellung des Mengengerüsts:

Der abschließende Schritt bei der Vorbereitung der Unterlagen für die Ausschreibung ist die Erstellung eines Mengengerüsts. Dabei handelt es sich um eine Excel-Tabelle, die den Produktnamen, eine detaillierte Produktbeschreibung sowie die benötigte Menge enthält. Zusätzlich sind weitere Spalten für die vorgeschlagenen Preise der Bewerber erforderlich, ebenso wie eine Formel zur Berechnung der Gesamtpreise. Zudem müssen die Herstellerinformationen in dieser Tabelle so aufbereitet werden, dass sie für einen fairen Wettbewerb neutralisiert sind.

*Dieser Prozessschritt konnte im Rahmen des Digitalisierungsprojektes ebenfalls erfolgreich umgesetzt werden: Mithilfe eines standardisierten Templates konnte die Formatierung in Python abgebildet werden. Die Produktinformationen ließen sich aus dem Warenwirtschaftssystem auslesen, transformieren, mithilfe eines Sprachmodells neutralisieren und in das Template übertragen.*

### **4.3. Datenstruktur und Datenquellen verstehen**

Im Rahmen des Digitalisierungsprojekts wurden verschiedene Datenquellen in die Pipeline integriert, die erforderlich waren, um den gewünschten Output zu generieren. Um einen klaren Überblick zu gewährleisten und gleichzeitig weitere Potenziale der genutzten Plattformen aufzuzeigen, werden diese nachfolgend aufgeführt.

#### Produktinformationen für technische Vorbeschreibe

Die Quellen der Daten sind äußerst vielfältig. Der Großteil der Informationen stammt aus PDF-Dokumenten der Hersteller. Darüber hinaus können jedoch auch Inhalte von Websites oder Informationen aus Portalen, auf die Planungsbüros Zugriff haben, eine Rolle spielen.

**Herangehensweise:** Für das Digitalisierungsprojekt wurden die Dateien aus einem zentralen Ordner geladen, in dem sie abgelegt wurden. Durch diesen strukturierten Ansatz konnten die enthaltenen Informationen effizient ausgelesen und für die weitere Verarbeitung genutzt werden.

**Potenziale:** Zusätzlich könnte die Möglichkeit des Web-Scrapings integriert werden, um Informationen direkt von Websites und Portalen zu extrahieren. Darüber hinaus wäre die Einbindung einer Schnittstelle zu einer bestehenden Cloud-Lösung eine sinnvolle Erweiterung. Dadurch müssten die Dokumente nicht lokal in den Code geladen werden, sondern es könnte direkt auf die in der Cloud gespeicherten Projektdaten zugegriffen werden. Dies würde die Effizienz und Flexibilität der Datenverarbeitung weiter steigern.

### Auto CAD



Abbildung 6: Grundriss der Aura in AutoCAD (Aura GmbH)

AutoCAD ist eine Software für computergestütztes Design (CAD), die hauptsächlich für technische Zeichnungen, Konstruktionen und 3D-Modelle verwendet wird. Entwickelt von Autodesk, wird sie weltweit in Branchen wie Architektur, Maschinenbau und Bauwesen eingesetzt, um präzise Pläne und Designs zu erstellen. Ein Beispielhafter Grundriss ist links in Abbildung sechs zu sehen. Dieser wird benötigt, um die Stückzahlen auszulesen.

**Herangehensweise:** Im Rahmen des Digitalisierungsprojekts wurden Dateien aus AutoCAD im DXF-Format exportiert und anschließend in Python mithilfe einer passenden Bibliothek eingelesen. Auf diese Weise konnten die enthaltenen Metadaten effizient extrahiert und weiterverarbeitet werden.

**Potenziale:** Für eine weitergehende Automatisierung könnte AutoCAD künftig direkt über eine bestehende API-Schnittstelle mit dem Python-Code verbunden werden. Dadurch würden die manuellen Download- und Upload-Schritte entfallen, und Medienbrüche könnten effektiv überwunden werden. Zusätzlich ist es erforderlich, interne Standards für die einheitliche Benennung der Metadaten der Möbelstücke zu etablieren, um eine konsistente und zuverlässige Verarbeitung sicherzustellen.

### Warenwirtschaftssystem (Concept Office)

Das Warenwirtschaftssystem unterstützt bei der Verwaltung der Produktinformationen sowie Beständen. Durch die Integration aller relevanten Daten bietet es Unternehmen eine



transparente Übersicht und unterstützt eine bedarfsgerechte Planung. Im Rahmen der Ausschreibungsprozesse wird das Warenwirtschaftssystem genutzt, um die für ein Projekt benötigten Möbel samt ihrer detaillierten Produktbeschreibungen und Positionsnummern auszulesen. Diese Informationen dienen als Grundlage für die Erstellung des Mengengerüsts. Je nach Projekt können diese Daten mehrere hundert Seiten umfassen, was eine strukturierte und effiziente Verarbeitung unerlässlich macht.

**Herangehensweise:** Die Daten aus dem Warenwirtschaftssystem wurden zunächst als Excel-Tabelle exportiert und anschließend mit Python eingelesen und weiterverarbeitet. Dabei wurden die ausgelesenen Informationen neutralisiert und in ein programmatisch erstelltes Excel-Template übertragen. Das fertige Dokument wurde schließlich erneut exportiert, um es für den weiteren Ausschreibungsprozess bereitzustellen.

Angebot 64915  
Einmalkunde, 90461 Nürnberg



Anzahl	Artikel	E-Preis/EUR	G-Preis/EUR
30	<b>Pos. 2.1.1</b> <b>W851B1710T</b>  DIS STS16,10H,BU, T-Fuss B:1600 T:434 Breite: 1600 Tiefe: 434 Höhe: 415 Korpusfarbe: Snow WY Position Fuss: Rechts Frontart: Melamin 19mm S01 Frontfarbe: Snow WY Türanschlag: Öffnung von links nach rechts Griffvariante Schiebe-Tür: Griffleiste beidseitig 11 Griffleiste: beidseitig Griffleistenfarbe: Pearl Snow ZW Sockel: Pearl Snow ZW Integrierte Rückwand: Gleiche Farbe wie Korpus Adaptive Rückwand: Nein Schloß: Mit Schloß S00 Innenzylinder: Mit Innenzylinder WZ1 Schlüssel-Typ: Kippschlüssel Schließnummernbereich: Nein Schließkreise: S-Nr. 18001-18100 Schublade L: Nein Stromanschlüsse: Nein power	250,76	7.522,80
30	<b>Pos. 2.1.2</b> <b>W8H2T070AT</b>  MTU Hutch,20H,TU, T-Fuss B:1600 T:209 Breite: 1600 Tiefe: 209 Höhe: 750 Typ Kabinett Stoff: Acoustic Kabinett Stoff: Atlantic AT Kabinett Stofffarbe: 66057 AT06 Stoff Kissen: Atlantic AT Stofffarbe Kissen: 66057 AT06 Plattenfarbe: Snow WY Plattenkante: Gleiche Farbe wie Platte Korpusfarbe: Snow WY Integrierte Rückwand: Gleiche Farbe wie Korpus Adaptive Rückwand: Nein Ambiente Licht: Nein Stromanschlüsse: Nein power	514,69	15.440,70

Abbildung 7: Auszug aus Warenwirtschaftssystem

**Potenziale:** Um den Automatisierungsgrad weiter zu steigern, könnte die Integration einer API-Schnittstelle genutzt werden, um die Daten direkt aus dem Warenwirtschaftssystem abzurufen, anstatt sie manuell als Excel-Datei zu exportieren. Dadurch ließe sich nicht nur zusätzliche Zeit einsparen, sondern auch Medienbrüche effektiv vermeiden, was den gesamten Prozess effizienter und nahtloser gestalten würde.

### Graphische Darstellungen (pCon.Planner )



Abbildung 8: Graphische Darstellung der Möbel

Dieses Planungstool wird für die Erstellung und den Abruf der 3D Graphiken genutzt, welche in den technischen Vorbeschrieb einfließen.

**Herangehensweise:** Aus diesem Tool werden lediglich die grafischen Darstellungen der Möbelstücke übernommen und in technischen Spezifikationen eingefügt, um eine visuelle und technische Darstellung der Möbelstücke zu kombinieren.

**Potenziale:** Auch dieser Schritt könnte über eine API-Schnittstelle abgewickelt werden, um den Automatisierungsgrad weiter zu erhöhen.

## 5. Umsetzung der ausgewählten Prozessschritte

Nach der Analyse der Rahmenbedingungen und der Datengrundlage des Projekts wird im Folgenden die Umsetzung der ausgewählten Prozessschritte detailliert beschrieben und in einem [Github Repository](#) festgehalten. Dabei wurden verschiedene Ansätze evaluiert, um die effizienteste und zielführendste Lösung zu ermitteln. Erkenntnisse aus den explorativen Ansätzen, die potenziell von Interesse sein könnten, werden ebenfalls berücksichtigt. Das zentrale Ziel bestand darin, eine maximale Automatisierung der Prozesse unter Berücksichtigung der vorhandenen Ressourcen zu realisieren. Darüber hinaus wird ergänzend auf identifiziertes Optimierungspotenzial eingegangen.

### 5.1. Stückzahlen auszählen

Die erste Teilaufgabe des Digitalisierungsprojekts besteht darin, die Möbelstücke sowie deren Stückzahlen aus dem Grundriss, wie in Abbildung neun dargestellt, zu extrahieren. Bislang wurde dieser Prozessschritt in der Regel manuell durchgeführt: Die Pläne wurden ausgedruckt, die einzelnen Produkte händisch gezählt und in eine Liste übertragen. Dieses Vorgehen ist jedoch äußerst zeitaufwändig und fehleranfällig.

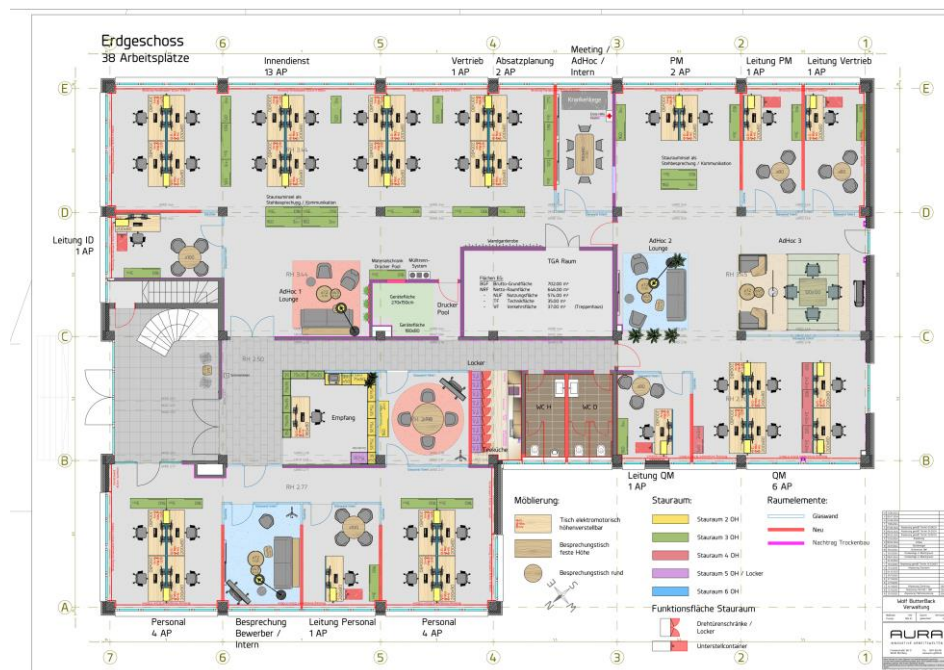


Abbildung 9: Grundriss aus einem Planungsprojekt der Aura (Aura GmbH)

Im Rahmen einer explorativen Auseinandersetzung mit dieser Problemstellung wurden insgesamt vier verschiedene Ansätze identifiziert, wie die Stückzahlen effizienter aus dem vorliegenden Grundriss extrahiert werden könnten. Diese Ansätze sind in Abbildung zehn zusammengefasst

und werden im Folgenden ausführlich erläutert. Die blau hinterlegten Kommentare enthalten Hinweise zu den Algorithmen, die für die Umsetzung der einzelnen Schritte herangezogen und kombiniert werden könnten, um die Pipeline aufzubauen.

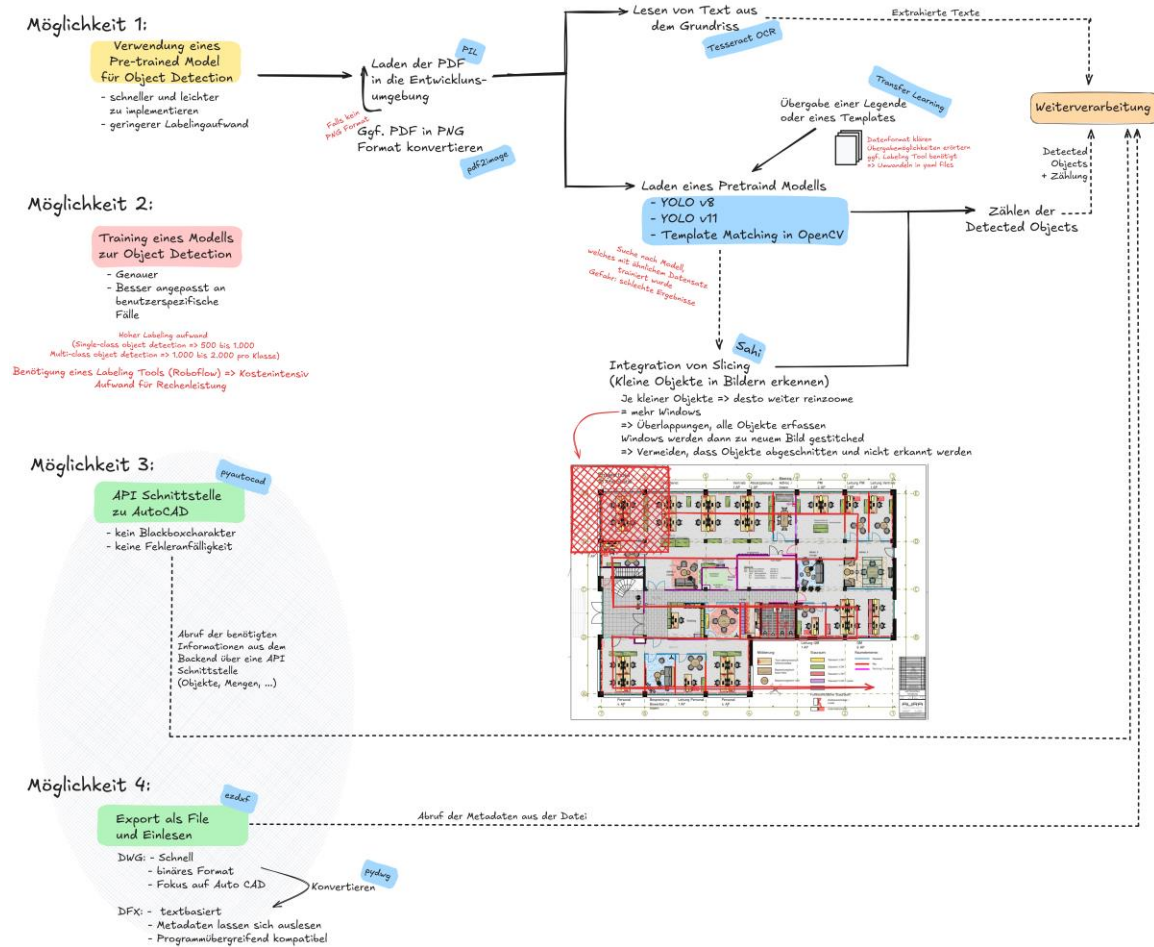


Abbildung 10: Lösungsansätze zum Erfassen der Stückzahlen aus einem Grundriss (Eigene Darstellung)

### Möglichkeit 1: Verwendung eines Pretrained Modells

Im ersten Ansatz wurde eine Lösung entwickelt, die Object-Detection- und OCR-Algorithmen kombiniert, um die Erfassung der Stückzahlen aus den Grundrissen zu automatisieren. Diese Herangehensweise basiert auf der Anwendung von sogenannten Pretrained Models aus den Computer-Vision-Technologien zur Verarbeitung von Bilddateien. Unter einem pretrained Modell versteht man ein Machine Learning Model, das bereits auf umfangreichen Datensätzen vortrainiert wurde und als Grundlage dient, um spezifische Aufgaben durch Feintuning oder Anpassungen effizient zu lösen. Der Grundriss wird hierfür mithilfe von Python eingelesen, wobei es essenziell ist, dass die Datei im PNG-Format vorliegt. Dies liegt daran, dass Object-Detection-Algorithmen speziell für die Verarbeitung von Bilddaten entwickelt wurden. Sollte der Grundriss nur in einem PDF-Format vorliegen, ist eine programmgesteuerte Konvertierung in ein

unterstütztes Bildformat problemlos möglich. Der Prozess zur automatisierten Erfassung der Stückzahlen aus den Grundrissen lässt sich des Weiteren in zwei zentrale Teilbereiche unterteilen.

Der erste Teilbereich umfasst die **Erkennung und Klassifizierung der Möbelstücke** um diese identifizieren und zählen zu können. Dabei kann ein Object-Detection-Algorithmus wie YOLO (You Only Look Once) zum Einsatz kommen. Dieser gehört zu den bekanntesten Algorithmen im Bereich der Computer Vision, da er eine schnelle und präzise Objekterkennung ermöglicht. Ein wesentliches Merkmal von YOLO ist seine Fähigkeit, die Erkennung in einem einzigen Durchlauf durch ein neuronales Netz durchzuführen, was zu hoher Geschwindigkeit und Echtzeitfähigkeit führt. Zudem bietet der Algorithmus ein gutes Gleichgewicht zwischen Genauigkeit und Effizienz (Rahima Khanam, 2024). Um die Ergebnisse zu verbessern, kann YOLO durch SAHI (Slicing Aided Hyper Inference) ergänzt werden. SAHI optimiert die Objekterkennung bei vor allem großformatigen und hochauflösenden Bildern, indem es Bilder in kleinere Abschnitte (sogenannte Slices) unterteilt, bevor der Erkennungsalgorithmus angewendet wird. Die erkannten Objekte aus jedem Slice werden im Anschluss wieder zusammengefügt, wodurch keine Objekte abgeschnitten oder übersehen werden (Fatih Cagatay Akyon, 2022). Der zweite Teilbereich umfasst die **Texterkennung und -zuordnung**. Dieser Schritt ist erforderlich, da die Textelemente im Grundriss wichtige Zusatzinformationen wie Maße oder Objektbeschreibungen enthalten, die nicht durch die reine Objekterkennung abgedeckt werden. Hierfür kommen OCR-Algorithmen (Optical Character Recognition) zum Einsatz, die speziell für die Extraktion von Texten aus Bildern entwickelt wurden. Leistungsstarke und etablierte OCR-Engines wie Tesseract eignen sich besonders gut für diese Aufgabe, da sie robust und vielseitig einsetzbar sind (Poatek, 2023). Durch die Kombination dieser beiden Teilbereiche wird eine umfassende und effiziente Analyse der Grundrisse ermöglicht, die sowohl die visuelle als auch die textuelle Information berücksichtigt. Dennoch gibt es einige Einschränkungen: Während OCR schwer mit komplexen oder unklaren Zeichnungen umgehen kann, hat YOLO Schwierigkeiten bei der Erkennung kleiner oder dicht gruppierter Objekte. Dies lässt sich mithilfe von Sahi ausgleichen, jedoch bleibt immer ein gewisses Restrisiko für Fehler. Darüber hinaus wird ein qualitativ hochwertiger, annotierter Trainingsdatensatz für das Finetuning des Modells benötigt, was wiederum mit einem hohen Labelingaufwand verbunden ist.

Aus den genannten Gründen wurde dieser Ansatz in Rahmen des Digitalisierungsprojektes nicht weiterverfolgt.

#### Möglichkeit 2: Training eines eigenen Modells

Für die Bilderkennung und das Auslesen der Möbelstücke wäre es theoretisch möglich, ein eigenes Modell von Grund auf zu erstellen oder zu trainieren. Dies würde jedoch enorme Ressourcen in Anspruch nehmen, da dafür eine große Menge an annotierten Trainingsdaten benötigt wird, um eine ausreichende Genauigkeit zu erreichen. Zudem wäre der

Entwicklungsaufwand für die Modellarchitektur und das Training zeit- und kostenintensiv. Da bereits zahlreiche vortrainierte Modelle wie YOLO existieren, die speziell für solche Aufgaben optimiert sind, wäre dies Herangehensweise für diesen Usecase übertrieben.

#### Möglichkeit 3 & 4: Abruf der Daten über eine API / Datenupload

Da die Erstellung der Grundrisse mit AutoCAD erfolgt, bot sich eine einfachere und weniger fehleranfällige Alternative zu den zuvor genannten Optionen an. Statt auf KI-basierte Methoden zurückzugreifen, kann die **Python-API von AutoCAD**<sup>1</sup> genutzt werden, um Metadaten direkt aus den AutoCAD-Dateien auszulesen. Diese Vorgehensweise ermöglicht eine präzise und effiziente Erfassung der Möbelstückbezeichnungen und Stückzahlen mit geringerem Aufwand und reduziertem Fehlerrisiko, da die Daten direkt aus dem Backend über einen Datenbankabruf ausgelesen werden und nicht erst von einer KI ermittelt werden müssen. Um in der Explorationsphase einen ersten Einblick in die Metadaten zu erhalten, wurde der Grundriss zunächst im DXF-Format exportiert, anschließend in Python eingelesen und analysiert. Das Besondere an diesem Format ist seine plattformunabhängige Struktur, die den Austausch von 2D-

	Layer	Block Count		Blockname	Block Count
0	0	12	7	AP-200-2	37
1	A_BODEN	1	14	Lightnet_CUBIC F2_Side-Cut-Out-Base	22
2	A_Bemassung	1	43	sts160-3oh	22
3	A_Boden_BEMABUNG	1	34	locker40-5R	9
4	A_DECKE	1	1	*U72	8
5	A_Decke_AKUSTIKSEGEL	42	40	sts120-3oh	7
6	A_GRUNDRISS-2D	1	17	Pflanze	6
7	A_Klima-Lüftung	1	16	Perch	6
8	_AU_MÖBEL	188			

Abbildung 11: Analyse der Metadaten Layer (rechts) und Elemente (links) aus dem Grundriss

und 3D-Konstruktionsdaten zwischen verschiedenen CAD-Programmen erleichtert, sowie seine textbasierte Lesbarkeit, die eine einfache Bearbeitung und Automatisierung unterstützt. Neben geometrischen Daten wie Linien, Kreisen und Polygonen enthält das DXF-Format auch Metadaten, die Informationen zu Layern, Farben, Linientypen und weiteren Attributen umfassen. Diese zusätzlichen Informationen ermöglichen eine strukturierte Organisation der Zeichnungselemente und vereinfachen die Weiterverarbeitung in unterschiedlichen CAD-Systemen (Lasers, 2023). Die Analyse hat ergeben, dass die gewünschten Metadaten zwar vorhanden und auslesbar sind, jedoch kein standardisierter Prozess für die Benennung der Möbelstücke etabliert ist. Während die Layer (Abbildung 11, links) noch einheitliche und aussagekräftige Bezeichnungen aufweisen, fehlen bei den Elementen selbst sprechende Variablennamen (Abbildung 11, rechts). Die explorative Untersuchung dieses Problems

<sup>1</sup> [Autodesk - Python API Overview](#) und [Installation von pyautocad 0.2.0](#)

verdeutlicht die Notwendigkeit, einen standardisierten Prozess zur Vereinheitlichung der Benennungen einzuführen. Erst nach der erfolgreichen Etablierung eines solchen Prozesses kann mit der Integration einer API in die Pipeline begonnen werden.

#### **Key-Takeaways:**

Die explorative Analyse hat gezeigt, dass der **Einsatz von KI nicht immer erforderlich** ist. Für diesen speziellen Use Case stellt das Auslesen der Metadaten über eine **API-Schnittstelle die optimale Lösung** dar. Da in Ausschreibungsprozessen höchste Genauigkeit entscheidend ist und idealerweise keinerlei Fehler auftreten sollten, wäre der Einsatz eines Computer-Vision-Algorithmus zu unsicher. Zudem wäre der Aufwand für die Implementierung einer solchen KI nicht gerechtfertigt, da alle notwendigen Informationen und Zugänge für einen effizienten Metadatenabruf bereits vorhanden sind.

Es ist wichtig, eine **einheitliche und aussagekräftige Datenstruktur** in AutoCAD zu **implementieren**, damit die Elemente eindeutig einem bestimmten Möbelstück zugeordnet werden können. Im Anschluss sollte ermittelt werden, für welche Zwecke diese Informationen genau benötigt werden. Aus diesem Ergebnis ergibt sich, ob die Informationen möglicherweise direkt aus dem Benutzerinterface von AutoCAD abgerufen werden können oder ob die **API in die Python-Pipeline integriert** werden muss, um die Daten in diesem Prozess weiterzuverarbeiten.

## **5.2. Mengengerüst**

Der zweite Aufgabenbereich bei der Erstellung einer Ausschreibung ist die Entwicklung eines Mengengerüsts. Hierbei handelt es sich um eine von Aura erstellte Excel-Tabelle, die neben den Positionsnummern der Produkte (Möbelstücke) auch deren Eckdaten wie Maßangaben und Farben sowie die Mengenangaben enthält. Diese Datei wird in zwei unterschiedlichen Szenarien verwendet. Zum einen beinhaltet die Tabelle eine Ausgabe, die geschätzte Preisangaben umfasst. Diese werden ebenfalls aus dem Warenwirtschaftssystem abgerufen und ermöglichen eine erste Preisschätzung für das Projekt. In diesem Fall dient die Datei der internen Nutzung, und es ist nicht erforderlich, die Produktangaben zu neutralisieren. Artikelnummern müssen in diesem Kontext nicht entfernt werden. Der zweite Anwendungsfall betrifft die Weitergabe der Datei an die Unternehmen, die sich auf die Ausschreibung bewerben möchten. Diese verwenden die Tabelle, um ihre Preisschätzungen für die einzelnen Produkte einzutragen, woraufhin der Gesamtpreis automatisch berechnet wird. In diesem Fall muss die Datei zwingend neutralisiert werden, um einen fairen Wettbewerb zu gewährleisten. Das bedeutet, dass keine Artikelnummern enthalten sein dürfen und die Produktbeschreibungen nicht auf einen bestimmten Hersteller zurückzuführen sein dürfen. Zudem müssen die meisten Felder der Excel-Tabelle gesperrt werden, sodass die Bewerber ausschließlich die Spalte mit den Preisangaben

bearbeiten können. In Abbildung 12 ist der Prozess dargestellt, wie diese Aufgabenstellung programmatisch umgesetzt wurde.

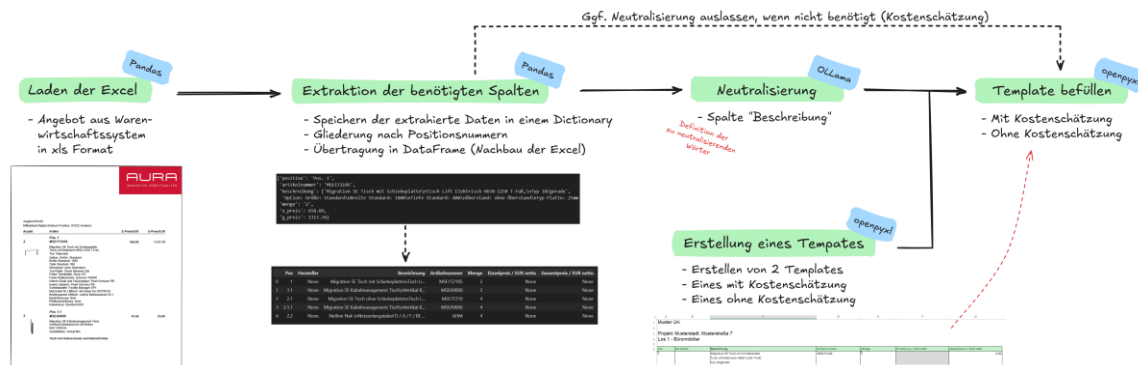


Abbildung 12: Pipeline zur Automatisierung der Erstellung eines Mengengerüsts in einem Ausschreibungsprozess (Eigene Darstellung)

Die wesentlichen Schritte lassen sich jedoch wie folgt zusammenfassen:

- Export der Produktdaten:** Die Produktdaten werden als XLS-Datei aus dem Warenwirtschaftssystem exportiert.
- Laden der Daten in Pandas:** Die XLS-Datei wird mithilfe von Python in ein Pandas DataFrame geladen.
- Datenvorverarbeitung:** Mit Pandas-Befehlen werden verschiedene Vorverarbeitungsschritte durchgeführt, um die relevanten Informationen aus der XLS-Datei zu extrahieren und in einem Dictionary zu speichern.
- Neutralisierung der Bezeichnungen:** Die Texte unter dem Schlüssel „Bezeichnung“ werden mithilfe eines OLLama-Sprachmodells neutralisiert (weitere Details hierzu sind im Kapitel „5.3. Technische Vorbeschriebe“ zu finden), sodass keine Rückschlüsse auf den Hersteller möglich sind.
- Erstellung der Zieltabelle:** Der Inhalt des Dictionary wird in ein Pandas DataFrame übertragen, das der Struktur der Zieltabelle entspricht. Es gibt zwei verschiedene Tabellenstrukturen:
  - Für die interne Kostenschätzung: (Positionsnummer, Hersteller, Bezeichnung (nicht neutralisiert), Artikelnummer, Menge, Einzelpreis, Gesamtpreis)
  - Für die Bewerber: (Positionsnummer, Bezeichnung (neutralisiert), Menge, Einzelpreis (manuell befüllbar), Gesamtpreis (Formel zur Berechnung hinterlegt))
- Export und Layout-Anpassung:** Mithilfe von Openpyxl wird das Pandas DataFrame in eine XLSX-Datei exportiert. Dabei wird das optische Layout bearbeitet, Excel-Formeln integriert und die Tabelle gesperrt, um nur die Bearbeitung der relevanten Felder durch die Bewerber zu ermöglichen.



### Key-Takeaways:

Mithilfe der erstellten **Templates** und der **Bereinigungs-Pipeline** müssen die Inhalte aus dem Warenwirtschaftssystem zu den Produkten nicht mehr manuell von einer PDF-Datei in eine Excel-Datei durch Copy-Paste übertragen werden. Dies führt zu einer **erheblichen Zeitersparnis** und minimiert gleichzeitig potenzielle Fehler, die durch manuelle Kopiervorgänge entstehen könnten. Der Einsatz des **Sprachmodells ist in diesem Prozessschritt besonders wertvoll für die Neutralisierung der Texte**. Durch den Einsatz des OLLama Modells wird eine schnelle und konsistente Verarbeitung einer großen Menge an Texten ermöglicht, was manuell nur mit erheblichem Aufwand möglich wäre.

Im Rahmen des Digitalisierungsprojekts wurde in diesem Prozessschritt ein **hoher Grad an Automatisierung durch den Einsatz von Python und künstlicher Intelligenz erreicht**. Wie in Kapitel „5.4. User Interface“ beschrieben, kann diese Pipeline über ein benutzerfreundliches User Interface für Endnutzer zugänglich gemacht werden. Derzeit müssen die XLS-Dateien aus dem Warenwirtschaftssystem manuell hochgeladen werden, um sie der Pipeline zuzuführen. Für einen noch höheren Automatisierungsgrad und zur Überwindung von Medienbrüchen könnte das Warenwirtschaftssystem jedoch über eine **API-Schnittstelle<sup>2</sup> direkt an das User Interface angebunden** werden. Dadurch könnten die benötigten Informationen direkt durch die Eingabe oder Auswahl der entsprechenden Projektnummer abgerufen werden. Diese Erweiterung würde nicht nur die Effizienz steigern, sondern auch die Benutzerfreundlichkeit verbessern, indem redundante Arbeitsschritte entfallen und potenzielle Fehlerquellen minimiert werden.

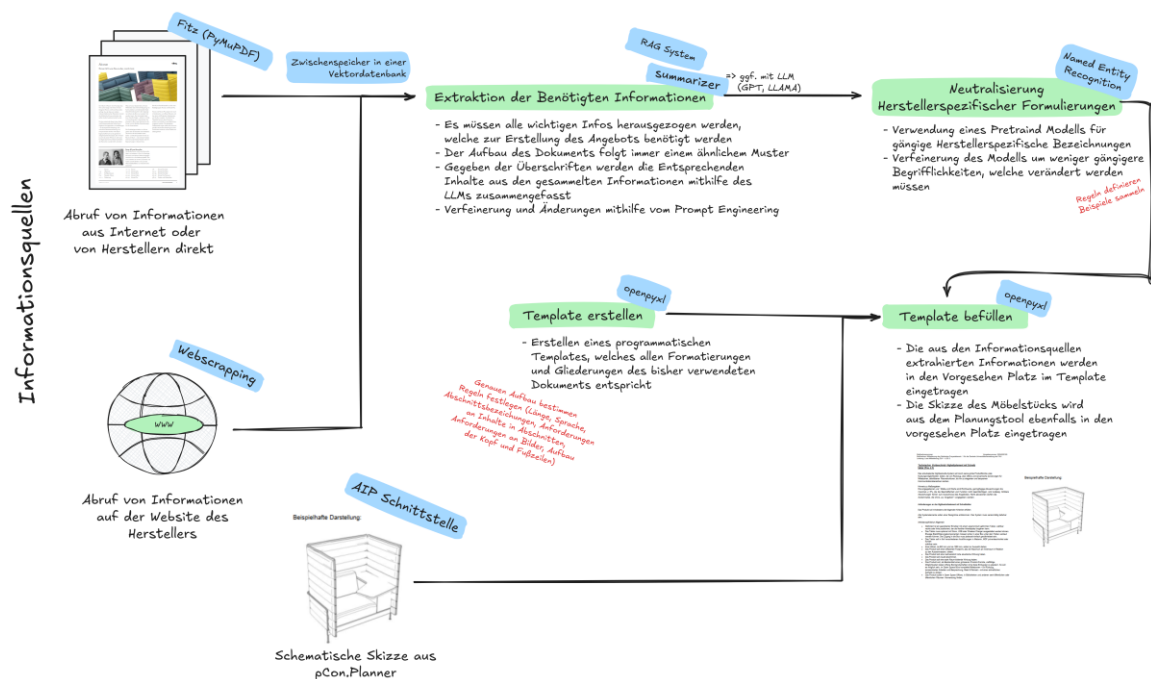
## 5.3. Technische Vorbeschriebe

Die Erstellung des technischen Vorbeschriebs stellte im Rahmen dieses Digitalisierungsprojekts die größte Herausforderung dar. Ziel war es, den technischen Vorbeschrieb für die Ausschreibungsunterlagen so weit wie möglich zu automatisieren. Der technische Vorbeschrieb ist eine detaillierte Beschreibung eines Möbelstücks, die prägnant und übersichtlich Informationen zu Maßen, Serviceleistungen, Nachhaltigkeit, Farben, Materialien, Funktionalitäten und weiteren Merkmalen zusammenfasst. Ergänzend enthält er eine schematische 3D-Skizze, um eine visuelle Orientierung zu bieten. Bisher wurde dieser Prozess manuell von den Mitarbeitenden durchgeführt. Die dafür benötigten Dokumente stammten aus verschiedenen Quellen, wie dem Internet, PDF-Dateien und dem internen Warenwirtschaftssystem. Diese wurden durchgesehen, relevante Informationen entnommen und in ein vorgegebenes Word-Template übertragen. Die 3D-Skizzen wurden aus dem Planungsprogramm pCon.Planner

---

<sup>2</sup> Weitere Informationen zu der Bereitstellung der API durch Concept Office sind [hier](#) zu finden

generiert und anschließend eingefügt. Darüber hinaus wurden sämtliche Texte, die Rückschlüsse auf den Hersteller zuließen, in neutraler Form überarbeitet. Dieser manuelle Arbeitsablauf war äußerst zeitaufwändig und fehleranfällig. Die Ineffizienz dieses Prozesses wirkte sich negativ auf die Wettbewerbsfähigkeit aus, da erhebliche Zeit für die Erstellung der Dokumente aufgewendet wurde, die für die eigentlichen Kernkompetenzen – die Planung und Umsetzung der Projekte – hätte genutzt werden können. Im Rahmen des Digitalisierungsprojekts wurde der Versuch unternommen, diesen Prozess so weit wie möglich zu automatisieren und Handlungsempfehlungen für die nächsten Schritte zu entwickeln. Die Ergebnisse der explorativen Analyse zu diesem Thema sind in Abbildung 13 dokumentiert.



**Abbildung 13: Pipeline zur Automatisierung der Erstellung des technischen Vorbeschriebs in einem Ausschreibungsprozess (Eigene Darstellung)**

Die wesentlichen Schritte lassen sich jedoch wie folgt zusammenfassen:

#### Abruf von Informationen:

Die Informationen werden entweder aus dem Internet, von Herstellern direkt oder von Hersteller-Websites gewonnen. Für den Proof of Concept wurden die Informationen manuell in die Verarbeitungspipeline eingespeist. Für ein höheres Maß an Automatisierung besteht die Möglichkeit, die Datenquellen direkt mit der Pipeline zu verknüpfen. Hierbei könnte Webscraping genutzt werden, um Informationen automatisch aus dem Internet abzurufen. Für den Zugriff auf interne Datenquellen könnte eine API-Schnittstelle eingerichtet werden, die direkt mit der Ordnerstruktur der Cloud verbunden ist.

### Extraktion der benötigten Informationen:

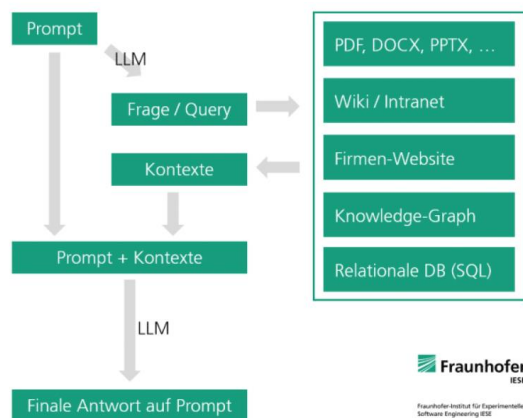
Für den Proof of Concept wurden die Daten aus den PDF-Dateien mithilfe der Fitz-Bibliothek in Python extrahiert. Dabei wurden sowohl Textinhalte als auch Bilder sowie Texte innerhalb von Bildern ausgelesen und in einer Variablen gespeichert. Zur Extraktion und Clustering der relevanten Informationen wurden zwei unterschiedliche Ansätze aus dem Bereich des Prompt Engineering<sup>3</sup> untersucht:

- Multi-Modal Prompting in Verbindung mit Directional Stimulus Prompting
- RAG (Retrieval-Augmented Generation)

### **Multi-Modal Prompting i.V.m. Directional Stimulus Prompting:**

Die Zusammenfassung der relevanten Informationen erfolgte mithilfe eines Sprachmodells (OLLAMA-Modelle). Dabei wurden die ausgelesenen Inhalte des PDF-Dokuments an das Sprachmodell direkt übergeben (Multi-Modal Prompting). Anschließend wurde gezielt nach den spezifischen Kriterien für das Möbelstück im Kontext des technischen Vorbeschriebs gesucht (Directional Stimulus Prompting). Alle wichtigen Informationen wurden systematisch zusammengefasst und in einem Dictionary gespeichert. Hierbei dienten die Kriterien als Keys, während die zugehörigen Values aus Listen von Stichpunkten bestanden, die aus den Input-Quellen zu diesem Kriterium extrahiert wurden. Die Verfeinerung der Ergebnisse und anderweitige Anpassungen erfolgen durch gezieltes Prompt Engineering im Backend, indem diese als Systemprompts hinterlegt wurden.

### **Retrieval-Augmented Generation:**



**Abbildung 14: RAG Mechanismus (T. Honroth, 2024)**

Eine weitere Möglichkeit besteht in der Implementierung eines RAG-Mechanismus (Retrieval-Augmented Generation), dessen Architektur in Abbildung 14 dargestellt ist. Dabei werden die Dokumente in einer Vektordatenbank zwischengespeichert, um den Kontext des Sprachmodells zu erweitern. Wenn eine Anfrage an das Sprachmodell gestellt wird, erfolgt eine sogenannte Similarity Search. Diese durchsucht die Vektordatenbank nach den für den Prompt relevantesten Dokumenten, filtert diese heraus und verwendet sie, um eine präzise und kontextbezogene Antwort zu generieren. Dieses

Vorgehen ist besonders praktisch für Anwendungsfälle, bei denen eine große Menge an Informationen aus unterschiedlichen Quellen effizient verarbeitet und kontextualisiert werden muss.

<sup>3</sup> Weitere Informationen zum Prompt Engineering und Prompt Engineering Techniken sind [hier](#) zu finden.

Im Rahmen dieses Digitalisierungsprojekts wurden die Möglichkeiten eines RAG-Systems (Retrieval-Augmented Generation) untersucht. Dabei wurde im Jupyter Notebook „Embedding“ mithilfe von ChromaDB<sup>4</sup> eine Vektordatenbank erstellt, in der sowohl Text- als auch Bilddaten im Vektorformat gespeichert wurden. Da jedoch für die Erstellung technischer Vorbeschriebe regelmäßig unterschiedliche Dateien benötigt werden – teilweise nur eine spezifische Datei oder sehr spezielle, aktuelle Dokumente (z. B. keine veralteten Daten) – erwies sich dieser Ansatz als ungeeignet für die weitere Implementierung. Für zukünftige Projekte könnte die Vektordatenbank jedoch eingesetzt werden, um den Prozess zu automatisieren. Dabei könnten Daten direkt aus der Cloud oder aus spezifischen Planungstools angebunden werden, wodurch manuelle Uploads und Downloads überflüssig würden.

#### Neutralisierung herstellerspezifischer Formulierungen:

Bei der Neutralisierung der herstellerspezifischen Formulierungen kommen zwei vortrainierte Sprachmodelle zum Einsatz, welche typische herstellerspezifische Begriffe erkennt und neutralisiert. spaCy v2.0's Named Entity Recognition (NER) ist ein System zur Erkennung und Klassifizierung benannter Entitäten (z. B. Personen, Orte, Organisationen) in Texten. Es nutzt eine fortschrittliche Word-Embedding-Strategie, die Subword-Features und "Bloom"-Embeddings kombiniert, um auch seltene oder unbekannte Wörter besser zu erfassen. Das System basiert auf einem tiefen, konvolutionalen neuronalen Netzwerk mit Residualverbindungen, das Sprachmuster effizient erkennt. Es verwendet einen übergangsbasierten Ansatz, um Entitäten im Text zu markieren. Das Ziel ist eine optimale Balance zwischen Genauigkeit, Geschwindigkeit und Anpassungsfähigkeit, sodass das Modell sowohl vortrainierte als auch benutzerdefinierte Entitäten verarbeiten kann (spaCy, 2024). Da nicht alle Begriffe automatisch erkannt wurden, wurden spezifische Begriffe manuell über eine Liste ergänzt. Zusätzlich wurde das Ollama-Modell mithilfe eines Direct Stimulus Prompt<sup>5</sup> erweitert, um die Erkennung weiter zu optimieren.

#### Erstellung und Befüllung des Templates:

Nachdem die relevanten Informationen gesammelt wurden, können diese in einem zusammenhängenden Dokument konsolidiert werden. Hierfür wird ein programmatisches Template entwickelt, das die Formatierungen und Gliederungen des bisherigen Dokumentlayouts abbildet. Mithilfe eines Loops werden die extrahierten Informationen an den entsprechenden Stellen im Template eingefügt. Zusätzlich wird eine schematische Skizze des Möbelstücks, die im Planungsprogramm pCon.Planner erstellt wurde, integriert.

Für ein höheres Maß an Automatisierung könnte zukünftig eine API-Schnittstelle genutzt werden, um das Bild direkt aus dem Planungstool abzurufen und nahtlos in das Dokument einzufügen.

---

<sup>4</sup> Weitere Informationen zu ChromaDB können hier nachgeschlagen werden: <https://www.trychroma.com/>

<sup>5</sup> Weitere Informationen zu Direct Stimulus Prompts unter <https://arxiv.org/abs/2302.11520>

Dies würde einen weiteren Prozessschritt eliminieren und Medienbrüche vermeiden. Darüber hinaus könnten Projektinformationen, wie Vergabenummern und Maßnahmenbeschreibungen, im Backend hinterlegt und mit den Dokumenteninhalten verknüpft werden. Dadurch ließe sich der gesamte Workflow weiter optimieren und standardisieren.

### Key-Takeaways:

Die Umsetzung des technischen Vorbeschriebs im Projekt brachte deutliche Fortschritte in der Automatisierung dieses bislang arbeitsintensiven Prozesses. Besonders hervorzuheben ist die **effiziente Datenextraktion**. Mit Python-Skripten konnten technische Informationen aus PDF-Dateien, Bildern und anderen Quellen automatisiert extrahiert, mithilfe eines **vortrainierten Sprachmodells zusammengefasst** und in strukturierte **Templates** übertragen werden. Darüber hinaus ist ein wesentlicher Meilenstein die **Neutralisierung von Herstellerinformationen**. Hierbei kamen ebenfalls Sprachmodelle zum Einsatz, die herstellerspezifischen Formulierungen automatisiert erkannten und neutralisierten.

Dennoch ergeben sich weitere **Potenziale zur Optimierung und Automatisierung**, insbesondere durch die Einbindung zusätzlicher Datenquellen und die Nutzung von Schnittstellen anstelle manueller Uploads.

Mögliche Ansätze sind:

- **Webscraping** zur automatischen Erfassung von Daten aus dem Internet.
- API-Anbindungen zur direkten **Integration von Cloud-Datenquellen** und dem **Planungsprogramm pCon.Planner** für den Abruf relevanter Informationen.

Ein entscheidender Aspekt für die erfolgreiche Umsetzung ist die Einführung einer **plattformübergreifenden gemeinsamen ID**. Diese ermöglicht es, die extrahierten Informationen eindeutig zuzuordnen und effizient zu mappen, wodurch die Daten konsistenter und der Gesamtprozess weiter automatisiert wird.

## 5.4. User interface

Für eine benutzerfreundliche und einfache Interaktion mit der erstellten Pipeline zur Erstellung der Dokumente für die Ausschreibung wurde ein Entwurf für ein User Interface erstellt, welcher in Abbildung 15 zu sehen ist. Dieses soll einen Überblick über alle wichtigen Input-Daten und die verarbeiteten Ergebnisse bieten. Über ein Dropdown-Menü lässt sich das gewünschte Projekt auswählen. Das Interface ermöglicht den Upload von relevanten Dateien, darunter Grundrisspläne im IMG- oder DXF-Format sowie Excel-Dokumente mit Mengengerüsten. Diese Daten werden über eine im Backend verarbeitet und dienen als Basis für die Erstellung der Ausschreibungsunterlagen. Die hochgeladenen Pläne werden visuell dargestellt, und die im Projekt verwendeten Möbelstücke werden automatisch aus dem hinterlegten Informationssystem abgerufen (in diesem Fall noch aus der DXF Datei. Später könnte dies über eine direkte API Verbindung umgesetzt werden). Zusätzlich bietet die Benutzeroberfläche eine strukturierte Darstellung der technischen Vorbemerkungen und Standardspezifikationen für die Ausschreibung. Diese Dokumente können entweder aus Vorlagen generiert oder individuell angepasst werden. In Zukunft könnte ein integriertes Prompt-Feld ermöglicht die direkte Bearbeitung und Ergänzung von Textbausteinen, während die fertigen Dokumente als PDF oder Word-Dateien exportiert werden können.

Das System stellt somit eine effiziente Lösung dar, um Ausschreibungsunterlagen schnell und präzise zu erstellen, indem es alle relevanten Informationen an einem zentralen Ort bündelt und eine intuitive Interaktion ermöglicht.

### Key-Takeaways:

Durch eine **intuitive und benutzerfreundliche Benutzeroberfläche** ist es nicht notwendig programmatische Grundkenntnisse zu haben, da die Dokumentenerstellungspipeline im Hintergrund ausgeführt wird und der Nutzer lediglich **über das Frontend mit diesem interagiert**. Jedoch muss diese noch deployed und verprobt werden, bevor eine produktive Nutzung möglich ist. Auch wenn in der aktuellen Version Dateien durch einen **Upload** dem Programm zur Verfügung gestellt werden, ist es effizienter auch die **Datenübergabe zwischen Programmen zu automatisieren**.

Zusätzlich werden die gewünschten Schritte zur Zusammenfassung gemäß den vorgegebenen Kriterien durchgeführt, ebenso wie die Neutralisierung der Texte und die automatische Befüllung der Templates. Darüber hinaus könnte das System aber noch **erweitert** werden: Durch die Integration eines **Prompt-Feldes** in das Frontend erhalten Nutzer die Möglichkeit, Texte direkt in der Anwendung zu bearbeiten und individuell anzupassen, um ihnen den finalen Feinschliff zu verleihen.

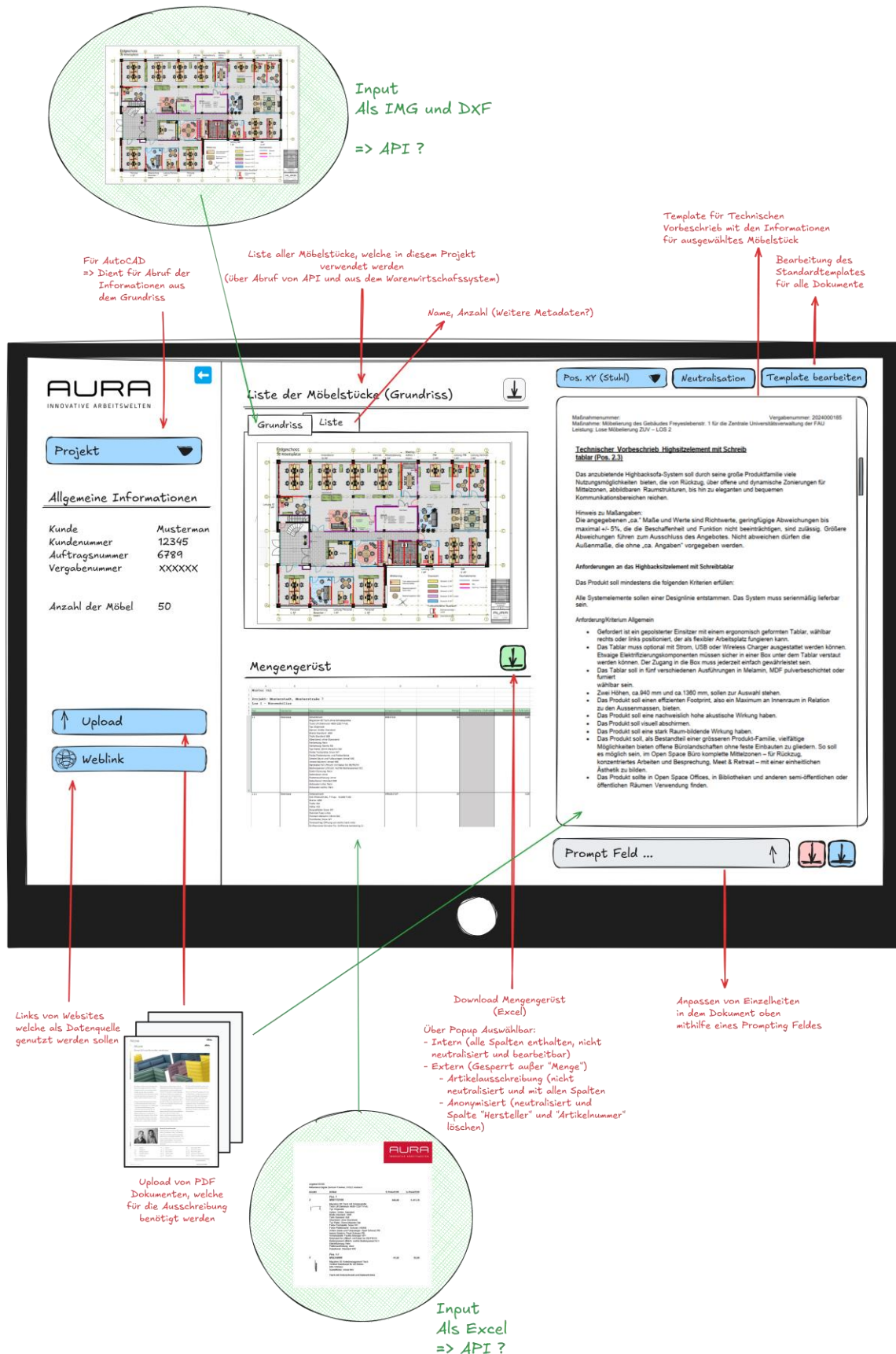


Abbildung 15: Geplantes Userinterface (eigene Darstellung)



## 6. Ergebnisse und Ausblick

Im Rahmen des Digitalisierungsprojekts „Automatisierung europaweiter Ausschreibungsprozesse für Büroeinrichtung mithilfe von KI“ wurde versucht Ergebnisse herauszuarbeiten, die den Weg für eine effizientere und fehlerreduzierte Bearbeitung solcher Prozesse ebnen. Besonders hervorzuheben ist die erfolgreiche Automatisierung der Erstellung technischer Vorbeschriebe. Durch die Kombination von Sprachmodellen und Python-Skripten konnten herstellersizifische Informationen neutralisiert und technische Details in standardisierte Templates übertragen werden. Dies führte zu einer deutlichen Zeitersparnis und einer signifikanten Reduktion von Fehlern. Ein weiteres zentrales Ergebnis ist die Optimierung der Mengengerüste. Statt zeitaufwendiger manueller Tätigkeiten wurden Daten aus dem Warenwirtschaftssystem automatisiert extrahiert, verarbeitet und in strukturierte Excel-Templates integriert. Dadurch entfiel der bisherige „Copy-and-Paste“-Ansatz, und die Effizienz des gesamten Prozesses konnte erheblich gesteigert werden. Zusätzlich wurde ein benutzerfreundliches User Interface entworfen, das eine einfache Bedienung ermöglicht und einen Überblick über alle wichtigen Projektdaten sowie die verarbeiteten Ergebnisse bietet. Dieses Interface bildet eine Grundlage für zukünftige Erweiterungen, wie etwa die Anbindung weiterer Datenquellen oder API-Schnittstellen.

Die im Projekt erzielten Fortschritte bieten eine solide Basis für zukünftige Entwicklungen. So könnte die Automatisierung durch die Integration von APIs zu Cloud-Lösungen oder Planungstools wie pCon.Planner weiter ausgebaut werden. Dies würde redundante Arbeitsschritte eliminieren und die Prozesse nahtloser gestalten. Darüber hinaus könnte eine KI-gestützte Analyse von Grundrissen zukünftig genutzt werden, um die Planung und Produktauswahl zu automatisieren. Eine solche Erweiterung könnte auch Lageparameter von Möbelstücken erfassen und für eine bedarfsorientierte Planung nutzbar machen. Ein weiterer vielversprechender Ansatz liegt in der Einführung von Web-Scraping-Methoden, um Informationen direkt von Herstellerwebsites und digitalen Plattformen zu extrahieren und so Grad der Automatisierung noch weiter zu erhöhen. Gleichzeitig würde eine standardisierte Datenstruktur, insbesondere in AutoCAD, die Zuverlässigkeit und Skalierbarkeit der automatisierten Prozesse erheblich verbessern, da hierdurch ein Abruf der Stückzahlen aus den Metadaten möglich wird.

Abschließend lässt sich sagen, dass dieses Projekt zeigt, wie der gezielte Einsatz von KI und Automatisierung nicht nur Effizienzgewinne ermöglicht, sondern auch die Innovationskraft und Wettbewerbsfähigkeit nachhaltig stärkt. Mit den aufgezeigten Potenzialen für Erweiterungen kann der Prozess der Ausschreibungen in Zukunft noch weiter optimiert und auf ein neues Niveau gehoben werden.

## Literaturverzeichnis

Fatih Cagatay Akyon, S. O. (24. 10 2022). *SLICING AIDED HYPER INFERENCE AND FINE-TUNING*. Von arxiv.org: <https://arxiv.org/pdf/2202.06934> abgerufen

Hellge, D. V. (2024). *READINESS-CHECK*. Von <https://digitalzentrum-kaiserslautern.de/unser-angebot/self-service/readiness-check> abgerufen

Lasers, M. B. (27. 06 2023). *Was ist eine DXF Datei? Wissenswertes zu Drawing Interchange File Format Dateien*. Von [https://www.mr-beam.org/blogs/news/was-ist-eine-dxf-datei?srsId=AfmBOoqbNT6jm\\_9h0TOV4TvnVFnEbf8UkseTSB-9TqvOsgEyTjqRW4s](https://www.mr-beam.org/blogs/news/was-ist-eine-dxf-datei?srsId=AfmBOoqbNT6jm_9h0TOV4TvnVFnEbf8UkseTSB-9TqvOsgEyTjqRW4s) abgerufen

Maisenbacher, B. (10. 12 2024). *Europaweite Ausschreibungen*. Von IHK Karlsruhe: <https://www.ihk.de/karlsruhe/fachthemen/recht/oeffentliches-auftragswesen/ausschreibungspraxis/europaweite-ausschreibungen-3441510> abgerufen

Poatek. (19. 06 2023). *Understanding OCR: Your (Entertaining) Guide to Optical Character Recognition*. Von <https://medium.com/poatek/understanding-ocr-your-entertaining-guide-to-optical-character-recognition-5891d31159c6> abgerufen

Rahima Khanam, M. H. (23. 10 2024). *YOLOv11: An Overview of the Key Architectural Enhancements*. Von arxiv.org: <https://arxiv.org/html/2410.17725v1> abgerufen

spaCy. (2024). *spaCy's NER model*. Von <https://spacy.io/universe/project/video-spacys-ner-model> abgerufen

T. Honroth, D. J. (13. 05 2024). *Retrieval Augmented Generation (RAG): Chatten mit den eigenen Daten*. Von [iese.fraunhofer.de](https://www.iese.fraunhofer.de): <https://www.iese.fraunhofer.de/blog/retrieval-augmented-generation-rag/> abgerufen