



Technische
Hochschule
Nürnberg

Fakultät Informatik

**Vergleichsanalyse von multimodalen Large
Language Models und einer
OCR/YOLO-Pipeline zur
Dokumentenklassifikation für die
Bundesagentur für Arbeit**

Bachelorarbeit im Studiengang Informatik

vorgelegt von

Lukas Müller

Matrikelnummer 3698673

Erstgutachter: Prof. Dr. Natalie Kiesler

Zweitgutachter: Prof. Dr. Korbinian Riedhammer

© 2026

Dieses Werk einschließlich seiner Teile ist **urheberrechtlich geschützt**. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Autors unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen sowie die Einspeicherung und Verarbeitung in elektronischen Systemen.

Kurzdarstellung

Kurze Zusammenfassung der Arbeit, höchstens halbe Seite. Deutsche Fassung auch nötig, wenn die Arbeit auf Englisch angefertigt wird.

Inhaltsverzeichnis

1 Einleitung	1
1.1 Motivation	1
1.2 Problemstellung	2
1.3 Zielsetzung	2
2 Theoretische Grundlagen	3
2.1 Vision-Language Models (VLMs)	3
2.1.1 Vom Sprachmodell zum multimodalen Modell	3
2.1.2 Architekturkomponenten	3
2.2 Vorstellung der Modelle	3
2.2.1 Qwen-2.5-VL	3
2.2.2 Pixtral-12B	3
2.3 Fine-Tuning	3
2.3.1 Parameter-Efficient Fine-Tuning (PEFT)	3
2.4 Information Extraction und Metriken	3
2.4.1 Information Extraction	3
2.4.2 Metriken	3
2.5 Verwandte Arbeiten	3
2.6 Rahmenbedingungen und Infrastruktur	3
3 Datenbasis	5
3.1 Dokumentarten	5
3.1.1 KG5b	6
3.1.2 Ausbildungsvertrag	8
3.1.3 Sonstige Dokumente	9
3.2 Datenschutz	10
4 Methode und Implementierung	11
5 Evaluation	13
6 Diskussion	15
7 Zusammenfassung	17

Abbildungsverzeichnis	19
Tabellenverzeichnis	21
List of Listings	23
Glossar	25

Kapitel 1

Einleitung

1.1 Motivation

Die öffentliche Verwaltung in Deutschland steht vor einer der größten Entwicklungen ihrer Geschichte. Durch den demografischen Wandel verliert der öffentliche Sektor in den kommenden Jahren eine signifikante Anzahl an erfahrenen Sachbearbeitern, während die Erwartungshaltung der Bürger an digitale und schnell verfügbare Dienstleistungen steigt. Wegen des Personalmangels und gleichzeitig wachsendem Leistungsanspruch rückt die automatisierte oder teilautomatisierte Bearbeitung von Verwaltungsprozessen in den Vordergrund. Diese Unterstützung für die Sachbearbeiter ist keine Option, sondern eine Notwendigkeit um in den kommenden Jahre Handlungsfähig zu bleiben.

Ein Beispiel hierfür ist die Bearbeitung von Kindergeldanträgen bei der Familienkasse der Bundesagentur für Arbeit. Mit einem jährlichen Aufkommen von mehreren Millionen Anträgen und starken saisonale Spitzen, etwa mit Beginn der Ausbildungen im Herbst, stößt die Familienkasse an ihre Kapazitätsgrenze.

Um sich dieser Herausforderung zu stellen, wurde bereits im Oktober 2025 ein erstes teilautomatisiertes System produktiv gesetzt. Dieses System basiert zu einem großen Teil auf einer Kombination aus OCR-Erkennung und Objektklassifizierung mittels YOLO-Modellen.

Obwohl dieses System einen wichtigen Schritt macht, zeigen sich während dem operativen Einsatz und der Entwicklung schon erste Grenzen. Besonders die Erkennung der handschriftlich ausgefüllten Dokumente stellen für diese Pipeline eine enorme Herausforderung dar.

Genau an diesem Punkt setzt die Thesis an. Durch den rasanten Fortschritt im Bereich der Generativen künstlichen Intelligenz öffnen sich neue Lösungen für Bearbeitung der Anträge. Insbesondere VLMs, die visuelle Informationen und Texte simultan verarbeiten, stellen eine vielversprechende Möglichkeit dar, die YOLO-OCR-Pipeline zu ersetzen. Sie können flexibel auf unbekannte Layouts reagieren, handschriftlichen Text extrahieren und den Kontext interpretieren.

Mit der Entwicklung der VLMs ergibt sich die Chance auf eine weitere Verbesserung für die Sachbearbeiter, jedoch muss getestet werden, ob diese Technologie bereits reif genug ist, um im öffentlichen Sektor eingesetzt zu werden.

1.2 Problemstellung

Trotz der Einführung der teilautomatisierten Pipeline im Oktober 2025 steht die Verarbeitung der Dokumente noch vor großen technischen Herausforderungen. Schon während der Entwicklung der YOLO/OCR-Pipeline kamen Probleme auf. Vornehmlich durch die hohe Varianz der einzelnen Dokumentarten stößt die aktuelle Pipeline an ihre Grenzen.

Ein Kernproblem stellt die Layout-Heterogenität der Ausbildungsverträge dar. Die verschiedenen Firmen und Kammern nutzen meist alle unterschiedliche Vertragsmuster, wodurch das Training der YOLO-Modelle komplex ist. Um mit den YOLO-Modellen ein zufriedenstellendes Ergebnis zu erhalten, benötigt es viele Trainingsdaten, was sehr zeit- und ressourcenintensiv ist. Herkömmliche OCR-Modelle liefern oft nur unstrukturierten Text, wodurch semantischer Kontext verloren geht. Bei schwierigen Layouts wird zum Beispiel die Unterscheidung von mehreren Daten wie Beginn und Ende der Ausbildung zunehmend problematisch, da diese eventuell falsch zugeordnet werden (semantic gap). Erschwerend kommt die Qualität der Eingabedaten hinzu. Die Dokumente liegen häufig als niedrig aufgelöste Scans oder Fotos vor und werden handschriftlich ausgefüllt.

Neben den Problemen der Erkennung stellt die Infrastruktur ein weiteres Problem dar. Da es sich um sensible Sozialdaten handelt, ist die Nutzung leistungsfähiger Cloud-APIs ausgeschlossen. Die Lösung muss demnach On-Premise sein, was einen Konflikt zwischen Modellgröße und Ressourcenverbrauch darstellt

1.3 Zielsetzung

- Forschungsfragen

Kapitel 2

Theoretische Grundlagen

2.1 Vision-Language Models (VLMs)

2.1.1 Vom Sprachmodell zum multimodalen Modell

2.1.2 Architekturkomponenten

2.2 Vorstellung der Modelle

2.2.1 Qwen-2.5-VL

2.2.2 Pixtral-12B

2.3 Fine-Tuning

2.3.1 Parameter-Efficient Fine-Tuning (PEFT)

2.4 Information Extraction und Metriken

2.4.1 Information Extraction

2.4.2 Metriken

2.5 Verwandte Arbeiten

2.6 Rahmenbedingungen und Infrastruktur

Kapitel 3

Datenbasis

3.1 Dokumentenarten

Im Rahmen der Kindergeldbeantragung sind verschiedene Nachweise gültig. Zu den anerkannten Dokumententypen zählen der offizielle Vordruck der Bundesagentur für Arbeit (KG5b)[?] sowie Ausbildungsverträge. Zusätzlich laden Kunden häufig weitere Unterlagen, wie beispielsweise Schulbescheinigungen, im Portal hoch. Da diese für den Kindergeldantrag nicht im Fokus stehen, werden sie im Folgenden unter der Kategorie „Sonstiges“ zusammengefasst.

Ziel der Arbeit ist die Extraktion spezifischer Informationen aus den genannten Dokumentenklassen. Um die gewonnenen Informationen aus dem Dokument bereitzustellen, wird ein flaches JSON-Schema verwendet. Die Wahl dieses Formats wird dadurch begünstigt, dass moderne LLMs durch ihr Training bereits eine hohe Zuverlässigkeit in der Generierung valider JSON-Strukturen aufweisen.

Die Qualität der Dokumente, die im Onlineportal hochgeladen werden, ist sehr unterschiedlich. Neben Scans mit guter Belichtung und hoher Auflösung enthält der Datensatz auch Fotos, die aus verschiedenen Winkeln und Entfernung aufgenommen wurden. Eine zusätzliche Herausforderung ist, dass die Dokumente häufig handschriftlich ausgefüllt sind.

In den Test- und Trainingsdatensätzen kommen diese Probleme in unterschiedlichen Konstellationen vor. Der Testdatensatz umfasst insgesamt 60 Dokumente, die sich gleichmäßig auf 20 Ausbildungsverträge, 20 KG5b-Formulare und 20 sonstige Dokumente verteilen. Der Trainingsdatensatz setzt sich aus 165 KG5b-Formularen, 227 Ausbildungsverträgen und 218 sonstigen Dokumenten zusammen.

Die Erstellung Ground Truth erfolgte in einem zweistufigen Verfahren. Zunächst wurden die JSON-Strukturen für den Testdatensatz vollständig manuell erstellt. Dieser Testdatensatz diente anschließend dazu, den System-Prompt zu optimieren, bis die Ergebnisse des Basis-Modells eine zufriedenstellende Qualität erreichten. Um den Annotationsaufwand für den Trainingsdatensatz zu reduzieren, wurde dieser optimierte Prompt für ein sogenanntes Pre-Labeling genutzt. Das Modell generierte dabei erste Vorschläge für die JSON-Schemas,

welche im Anschluss manuell validiert und korrigiert wurden. Dieser modellgestützte Annotationsprozess (Model-Assisted Labeling) ermöglichte eine effiziente Erstellung der Trainingsdaten bei gleichbleibend hoher Datenqualität.

3.1.1 KG5b

Der Vordruck KG5b ist, wie bereits erwähnt, ein offizielles Dokument der Bundesagentur für Arbeit, welches als Bescheinigung der Ausbildungsstätte dient. Volljährige Kinder weisen damit gegenüber der Familienkasse den Fortbestand ihrer Ausbildung nach, was die Voraussetzung für den weiteren Kindergeldbezug ist.

Da es sich um ein standardisiertes Formular handelt, ist das Layout gleichbleibend. Theoretisch vereinfacht dies die Extraktion der Informationen, da relevante Felder an denselben Positionen zu erwarten sind. In der Praxis ergibt sich jedoch eine hohe Varianz durch den Ausfüllprozess. Obwohl der Vordruck digital ausfüllbar angeboten wird, erfolgt die Bearbeitung in der Regel handschriftlich. Zudem erzwingen Unterschriften und Firmenstempeln einen doppelten Medienbruch. Das Dokument wird ausgedruckt, bearbeitet und anschließend wieder digitalisiert.

Abbildung 3.1 zeigt ein exemplarisch ausgefülltes KG5b-Formular mit Markierung der zu extrahierenden Felder.

Das Zielformat für die Extraktion wird durch das folgende JSON-Schema definiert:

```

1 {
2   "file_name": {
3     "type": "kg5b",
4     "name_child": "name, vorname",
5     "birthday_child": "DD.MM.YYYY",
6     "start_date_apprenticeship": "DD.MM.YYYY",
7     "end_date_apprenticeship": "DD.MM.YYYY",
8     "date_document": "DD.MM.YYYY",
9     "stamp_company": true,
10    "signature_company": true,
11    "signature_child": true,
12    "signature_legal_guardian": true,
13    "apprenticeship_ended": true,
14    "exam_month": "MM"
15  }
16 }
```

Listing 3.1: JSON-Schema für KG5b

3.1 Dokumentenarten

7

Familienname und Vorname der kindergeldberechtigten Person
Mustermann, Mathias
Kindergrd-Nr. 0283475234

Erklärung zum Ausbildungsverhältnis

Angaben zum Kind
Familienname: Mustermann, Vorname: Max

Angaben zum Ausbildungsverhältnis
Berufsausbildung als Apotheker
Das Ausbildungsdauer voraussichtlich bis 01.10.2025
Das Ausbildungserhältnis ist noch nicht beendet. (Bitte Nachweise beifügen)

start_date_apprenticeship 01.03.2002
birthday_child
exam_month September
end_date_apprenticeship 01.10.2025
apprenticeship_ended ja nein

a) Hat der/die Auszubildende die Abschlussprüfung bestanden? ja nein
Falls ja:
Letzter Prüfungstag war der _____
Die offizielle schriftliche Unterichtung über das Gesamtergebnis der Abschlussprüfung (Aushändigung oder Zugang der Bescheinigung bzw. vorläufige Bescheinigung des Prüfungsausschusses über das Bestehen der Prüfung oder Ähnliches) erfolgte am _____

b) Ist die Abschlussprüfung nach Ablauf der vertraglichen Ausbildungszeit abgelegt worden? ja nein
Falls ja:
Wurde das Ausbildungserhältnis bis zur offiziellen schriftlichen Unterichtung über das Gesamtergebnis der Abschlussprüfung verlängert? ja nein

c) Wurde das Ausbildungserhältnis aus anderen Gründen beendet? ja nein
Falls ja:
Das Ausbildungserhältnis endete am _____

d) Wurde der/die Auszubildende vom Ausbildungsbetrieb als Arbeitnehmer(in) übernommen? ja nein
Falls ja:
Beginn des Arbeitsverhältnisses: _____

Bestätigung des Ausbildungsbetriebes
Die obigen Angaben zum Ausbildungserhältnis sind richtig und vollständig.
Unterschrift: _____
Datum: 28.07.2025
stamp_company: Apotheke am Markt
Marktstr.
66666 Markt
Tel.: 66 666 6666 66
date_document
signature_company: Firmenstempel der Ausbildungsstätte

Hinweis an den Kindergeldberechtigten: Bitte füllen Sie Punkt e) erst aus, nachdem der Ausbildungsbetrieb die Angaben zum Ausbildungserhältnis bestätigt hat!

e) Hat das Kind bei einem anderen Arbeitgeber ein Arbeitsverhältnis begonnen oder wird es demnächst ein Arbeitsverhältnis bei einem anderen Arbeitgeber beginnen?
 ja nein
(voraussichtlicher) Beginn des Arbeitsverhältnisses: _____

We versichern, dass unsere Angaben vollständig sind und der Wahrheit entsprechen. Uns ist bekannt, dass wir alle Änderungen, die für das Anspruch auf Kinderfördergeld von Bedeutung sind, unverzüglich der Familienkasse mitzuteilen haben. Den Inhalt des Merkblattes Kinderfördergeld (zu finden unter www.bzst.de oder www.familienkasse.de) haben wir zur Kenntnis genommen.

Hinweis zum Datenschutz: Ihre Daten werden gemäß der §§ 31, 62 bis 79 Einkommensteuergesetz und der Regelungen der Abgabenordnung bzw. aufgrund des Bundeskindergeldgesetzes und des Sozialgesetzbuchs verarbeitet. Zweck der Verarbeitung der Daten ist die Prüfung Ihres Anspruchs auf Kinderfördergeld. Nähere Informationen über die Verarbeitung Ihrer Daten auf Basis der gesetzlichen Grundlagen und die Rechte, die Ihnen entstehen, finden Sie in der Datenschutzerklärung, die Sie durch die Kontaktdaten des Datenschutzbeauftragten bereitgestellt sind. Kinderdaten werden in der Regel nach dem Ende der Kinderfördergeldzahlung noch für 6 Jahre aufbewahrt.

signature_legal_guardian: Mustermann
Unterschrift der kindergeldberechtigten Person bzw. der gesetzlichen Vertretung
signature_child: Max Mustermann
Unterschrift des wählenden Kindes

All Eingaben löschen **Drucken** **Speichern**

(a) Erste Seite

(b) Zweite Seite

Abbildung 3.1: KG5b-Formular

Eine besondere Herausforderung bei der Verarbeitung stellen die binären Informationen dar, wie etwa die Checkbox für das Feld `apprenticeship_ended`. Hier muss das Modell visuell unterscheiden, ob ein Kasten leer, angekreuzt oder durchgestrichen ist. Ebenso kritisch ist die Detektion von `signature_company` und `stamp_company`. Im Gegensatz zu Textfeldern ist hier nicht der textliche Inhalt des Stempels relevant, sondern lediglich dessen Vorhandensein. Die Stempel sind oft blass oder eingefärbt, was besonders bei Schwarz-Weiß Scans zu einer schlechten Qualität führt. Zudem liegen die Felder `date_document` und `signature_company` nah beieinander. Unterschriften sind regelmäßig größer als ihr vorgesehener Platz, wodurch das Feld `date_document` überdeckt wird. Zusätzlich werden die Stempel häufig zusammen mit der Unterschrift auf der linken Seite platziert.

7

3.1.2 Ausbildungsvertrag

Im Gegensatz zu den standardisierten KG5b-Formularen weisen die Ausbildungsverträge eine signifikant höhere Varianz auf. Dies ist auf die Vielzahl unterschiedlicher zuständiger Stellen (z.,B. Industrie- und Handelskammern, Handwerkskammern, Ärztekammern) und Firmen zurückzuführen, die jeweils ein eigenes Layout definieren. Die Vielfalt der Dokumentenstruktur reicht dabei von formularbasierten Layouts bis hin zu unstrukturierten Fließtexten.

Eine weitere Herausforderung stellt der Umfang der Dokumente dar. Ausbildungsverträge umfassen häufig mehrere Seiten, wobei relevante Informationen und irrelevante Informationen, wie zum Beispiel gesetzliche Rahmenbedingungen, AGBs oder Rechtsbelehrungen, gemischt auftreten. Dies erschwert die Fokussierung des Modells auf die relevanten Daten.

Analog zu den KG5b-Formularen liegt auch hier durch die notwendigen Unterschriften und Stempel ein Medienbruch vor. Ein wesentlicher Vorteil gegenüber den handschriftlich ausgefüllten Anträgen besteht jedoch darin, dass die inhaltlichen Daten der Verträge fast ausschließlich maschinenschriftlich vorliegen, was die Zeichenerkennung erleichtert. Im Folgenden ist ein synthetischer Ausbildungsvertrag der Industrie- und Handelskammer abgebildet.

Das Zielformat für die Extraktion wird durch das folgende JSON-Schema definiert:

```

1 {
2   "file_name": {
3     "type": "vertrag",
4     "name_child": "name, „vorname“",
5     "birthday_child": "DD.MM.YYYY",
6     "start_date_apprenticeship": "DD.MM.YYYY",
7     "end_date_apprenticeship": "DD.MM.YYYY",
8     "date_document": "DD.MM.YYYY",
9     "stamp_company": true,
10    "signature_company": true,
11    "signature_child": true,
12    "signature_legal_guardian": true
13  }
14}

```

Listing 3.2: JSON-Schema Vertrag

Die Felder stimmen weitestgehend mit denen der KG5b-Formulare überein, jedoch entfallen die Felder `exam_month` und `apprenticeship-ended`, da diese selten auf Verträgen zu finden sind.

3.1 Dokumentarten

9

Berufsausbildungsvertrag
(§§ 10, 11 des Berufsbildungsgesetzes – BBiG)

Zwischen dem Ausbildenden (Ausbildungsbetrieb) und der/dem Auszubildenden wird nachstehender Berufsausbildungsvertrag zur Ausbildung im Ausbildungsbetrieb

Apotheker

(wenn einschlägig, bitte einschließlich Fachrichtung, Schwerpunkt, Wahlqualifikationen) und/oder Einsatzgeber nach der Ausbildungsordnung bescheinigen:

nach Maßgabe der Ausbildungsordnung¹ geschlossen.

Beispielschule
Zuständige Berufsschule
Änderungen an den vorliegenden Vertragshälften und von Auszubildenden untersagte die Erfüllung in das Verzeichnis der Betriebsausbildungsergebnisse bei der Industrie- und Handelskammer ansetzen. Die negativen Auswirkungen der tatsächlichen und zeitlichen Geltung des Ausbildungsbuches (Ausbildungsergebnis) sowie die negativen weisen Steueranträgen und Besteuerung dieses Vertrages.

Angaben zum Ausbildenden

Apotheke am Markt	Mustermann Name Musterstraße 4 Straße, Haus-Nr. 12345 Musterstadt PLZ Ort 01.03.2002 E-Mail-Adresse (Angabe freiwillig) 0800 123482394 Telefonnummer Beispiel, Pettla Name, Vorname verantwortliche Ausbilderin	Max Vorname birthday_child E-Mail-Adresse (Angabe freiwillig) Mobile/Telefonnummer (Angabe freiwillig)
-------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------

Angaben zum zu gesetzlichen Vertreter(n)²

<input checked="" type="checkbox"/> keiner <input type="checkbox"/> Eltern <input type="checkbox"/> Mutter <input type="checkbox"/> Vater <input type="checkbox"/> Vormund	Name, Vorname Straße, Haus-Nr.: mit Monaten angereichert: ³ PLZ Ort	Auf die Ausbildungsdauer wird die Berufsausbildung zurück: bzw. eine berufliche Fortbildung in Die Berufsausbildung wird in: <input checked="" type="checkbox"/> Vollzeit <input type="checkbox"/> Teilzeit ⁴ (% der Ausbildungsdauer in Vollzeit) durchgeführt.
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Soviel keine geschichtete Formulierung gewählt wird, steht dies allein der Verantwortung der Leistungsträger. Auch dort werden alle Menschen angesprochen – unabhängig von ihrem Geschlecht (siehe).

AUSFÜLLUNG FÜR AUSBILDENDE
BLATT 2 SEITE 1 VON 4

start_date_apprenticeship

Die Beschäftigung wird im Rahmen eines ausbildungsinstituierenden dualen Studiums absolviert
Das Berufsausbildungsverhältnis beginnt am **01.10.2022** und endet am **01.10.2025**

2. Probezeit
Die Probezeit beträgt **6** Monate⁵

end_date_apprenticeship
§ 2 – siehe S. 3 des Berufsausbildungsvertrages

§ 3 – Ausbildungsstätte
Die Ausbildung findet vorbehaltlich der Regelungen nach § 4 Nr. 12 dieses Vertrages in

Name/Ort/Nummern der Ausbildungsstätte und den mit dem Betrieb für die Ausbildung (bitte vereinige zusammenhängende Bau-, Montage- und sonstige Arbeitsstellen statt)

§ 4 – Pflichten des Auszubildenden
Ausbildungsmethoden aufzuhalten der Ausbildungsstätte(s) sind für den folgenden Zeitraum in den unten eingeschlossenen Ausbildungsstätte(n) vorgesehen (hierzu zählen auch Auslandsaufenthalte)

§ 5 – Pflichten der/des Auszubildenden
Führung von schriftlichen oder elektronischen Ausbildungsnachweisen folgt geführt:
Der Ausbildungsnachweis wird in: schriftlich elektronisch

§ 6 – Bestandteile der Vergütung und sonstige Leistungen
Höhe und Fälligkeit
 Das Ausbildungsvorhaben fällt in den Geltungsbereich des folgenden Tarifvertrages:
 Das Ausbildungsvorhaben fällt nicht in den Geltungsbereich eines gültigen Tarifvertrages.

§ 13 – Vertragstafel
Vorliegende Vertragsvereinbarungen zwischen den Auszubildenden und deren gesetzlichen Vertretern und den beteiligten Unternehmen/Veranstaltern der Berufsausbildung unterliegen noch deren Erstellung auszuhören. Bei elektronischer Abfassung ist die Vertragstafel bis zu einer späteren, über die elektronische Abfassung erfolgenden Vertragserweiterung bestehend aus dem Vertrag und der Vertragserweiterung zu verstehen. Diese Auszubildende verzichtet sich, den Empfang der elektronischen Vertragstafelung selbst oder über die elektronische Abfassung der Vertragserweiterung zu fordern. Der Vertragstafelung und der Empfangserweiterung sind von den beiden Auszubildenden nach Abschluss der Berufsausbildung zu unterschreiben. Die Vertragstafelung und die Empfangserweiterung beiderseitig beurkundet wurde, drei Jahre aufzuhören.

Anlage gemäß § 4 Nr. 1 des Berufsausbildungsvertrages⁶

(a) Erste Seite
(b) Zweite Seite

Abbildung 3.2: KG5b-Formular

3.1.3 Sonstige Dokumente

Die Kategorie **Sonstiges** steht als Auffangklasse für alle restlichen Dokumente bereit. Hier gehören zum Beispiel Schulbescheinigungen, Anträge auf Eintragung bei der Handelskammer oder Studienbescheinigungen dazu. Da diese Dokumente keine Relevanz für die Weiterbehandlung des Kindergeldes bei volljährigen Auszubildenden hat, werden keine Information benötigt. Somit ist rein die Klassifizierung wichtig, weshalb das extrahierte JSON-Schema folgendermaßen aussieht:

```

1 {
2   "file_name": {
3     "type": "sonstiges"
4   }
5 }
```

Listing 3.3: JSON-Schema Sonstiges

9

3.2 Datenschutz

Da es sich bei allen verwendeten Dokumenten um Echtdaten aus den Fachverfahren handelt, enthalten diese sensible personenbezogene Daten. Der Schutz dieser Daten hat höchste Priorität.

Um die Sicherheit bei der Verarbeitung der Echtdaten zu gewährleisten, wurde ausschließlich auf die Infrastruktur der Bundesagentur für Arbeit zurückgegriffen. Die Daten liegen in einem gesicherten Netzwerkbereich, auf den nur ein autorisierter Personenkreis Zugriff hat. Durch die Verwendung von Echtdaten wird die Qualität der Ergebnisse deutlich gesteigert.

Alle in dieser Arbeit abgebildeten Dokumentenbeispiele wurden jedoch synthetisch erzeugt und beinhalten keinerlei sensible Daten.

Kapitel 4

Methode und Implementierung

Kapitel 5

Evaluation

Kapitel 6

Diskussion

Kapitel 7

Zusammenfassung

Abbildungsverzeichnis

3.1 KG5b-Formular	7
3.2 KG5b-Formular	9

Tabellenverzeichnis

List of Listings

3.1 JSON-Schema für KG5b	6
3.2 JSON-Schema Vertrag	8
3.3 JSON-Schema Sonstiges	9

Glossar

FamKa Familienkasse der Bundesagentur für Arbeit. i

JSON JavaScript Object Notation. i, 5, 6, 8, 9

LLM Large Language Model. i, 5

MLLM Multimodal Large Language Model. i

OCR Optical Character Recognition. i, 1, 2

PEFT Parameter-Efficient Fine-Tuning. i

VLM Vision Language Model. i, 1, 2

YOLO You Only Look Once. i, 1, 2