

Verbundvorhaben FedXtract
Föderierte KI-Lösung zur Verarbeitung komplexer Dokumente

Sachbericht 2024

Condat AG

FKZ: 01IS21037A

Schlussbericht - Teil I: Kurzbericht

Gefördert durch das Bundesministerium für Bildung und Forschung (BMBF)

Projektträger

Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR)

Zuwendungsempfänger: Condat AG, Alt-Moabit 91d, 10559 Berlin	Förderkennzeichen: 01IS21037A
Projektleiter: Rolf Fricke	Tel.: +49 172 3991959 E-Mail: rf@condat.de
Laufzeit des Vorhabens: 01.11.2021 bis: 30.04.2024	
Datum: 18.08.2024	

Projektpartner:

1. Condat AG
2. Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS
3. ControlExpert GmbH

Ursprüngliche Aufgabenstellung sowie wissenschaftlicher und technischen Stand, an den angeknüpft wurde

Jedes Jahr werden immer noch Milliarden von Dokumenten aufwändig manuell verarbeitet, wie z.B. in der Schadenabwicklung bei Versicherungen. Nachdem diese bisher aufgrund der großen Bandbreite von Dokumentenlayouts und (Scan-) Qualität nur bedingt mit Standardmethoden zur Texterkennung und -interpretation analysiert werden konnten, ermöglichen die neu aufkommenden KI-Verfahren zur Dokumenterkennung nun eine automatische Verarbeitung. Bei kleineren Organisationen wie KMUs scheitert deren Einsatz jedoch noch daran, dass zum Training der KI-Modelle umfangreiche Datenmengen benötigt werden. Eine gemeinsame Entwicklung der Modelle durch verschiedene Unternehmen ist dabei kaum möglich, weil sie die Datenhoheit der Beteiligten verletzen würde. Weiterhin fehlt vielen Organisationen noch das Know-How, um KI-Modelle selbst zu trainieren und einzusetzen. Aus diesem Grund hat FedXtract eine Grundlage bereitgestellt, mit denen KMUs in die Lage versetzt werden, ein föderiertes Training von KI-Modellen ohne komplexes KI Know-How, unter Wahrung der Datenhoheit durchführen können. In dieser Lösung wird mittels sogenannten Föderierten Lernens (FL) eine unternehmensübergreifende, kontinuierliche Verbesserung eines globalen KI-Modells durch mehrere, verteilte Anwender ermöglicht. Dabei werden lediglich die inkrementellen Verbesserungen der Modelle ausgetauscht, so dass die sensiblen Daten lokal bleiben. Jeder Anwender kann dann in einem anschließenden Schritt ein User-spezifisches KI-Modell entwickeln, indem das globale, gemeinsame Modell über sogenanntes Transfer Learning (TL) für den eigenen Anwendungsfall optimiert wird.

Ablauf des Vorhabens

Da bei der Antragstellung die Maßnahmen zur Eindämmung der Covid-19-Pandemie mit ihren Reisebeschränkungen und Abstandsregeln bereits bekannt oder zumindest absehbar waren, konnte das Projekt diese schon in der Planung berücksichtigen. Deshalb wurde das FedXtract Projekt auch weitgehend dem Plan entsprechend durchgeführt. Die im Vorhaben entwickelte Lösung setzt auf Open Source (OS) - Frameworks zum FL und DL/ML auf. Dabei wurde die Basisplattform federführend von Condat bereitgestellt, während sich die FhG auf die Entwicklung der FL/DL Verfahren und ControlExpert auf die Bereitstellung der gelabelten Daten und Validierung konzentrierte. Die Schritte zur Entwicklung der FedXtract-Lösung wurden vom Condat-Projektteam unter Kooperation mit den Partnern wie folgt umgesetzt:

- Entwicklung einer Architektur,
- Auswahl eines Open Source (OS) FL-Frameworks,
- Integration und Koordination der Verfahren und Konzepte,
- separate Entwicklung der einzelnen Komponenten,
- dynamisches Zusammenfügen der FL/DL Bausteine durch klar definierte Schnittstellen,
- Durchführung des verteilten Testbetriebs
- Entwicklung einer Best Practice und
- Validierung durch 2 exemplarische Anwendungs-Beispiele, nachfolgend auch als Showcases bezeichnet.

Im Rahmen des Projektes wurden die Komponenten von Condat und den Partnern integriert, zusammengeführt und ihr Zusammenwirken getestet. Die dabei entstandenen Prototypen und Anwendungsbeispiele wurden auf verschiedenen öffentlichen Online- und Präsenz-Präsentationen vorgestellt, insbesondere zum Abschlusstreffen mit dem Projektträger DLR im Mai 2024.

Die wesentlichen Ergebnisse des Vorhabens und Zusammenarbeit mit anderen Stellen

Condat hat in seinem Teilprojekt folgende innovative Funktionen entwickelt, die über den State of the Art hinausgehen:

- Validierung und Bewertung verschiedener Open Source Frameworks für FL mit Tabelle nach verschiedenen Auswahlkriterien:
<https://www.fedxtract.net/best-practice/##fl5>
- Best Practice für KMU zur Planung und Aufsetzen eines Federated Learning-Projektes:
<https://www.fedxtract.net/best-practice/>
- Zwei exemplarische Anwendungs-Beispiele:
 - Exemplarisches Federated Learning-Beispiel für eine Document Recognition Lösung mit Nvidia-Flare und LayoutXLM (federführend durch FhG und ControlExpert):
<https://www.fedxtract.net/beispiele/beispiel-layout-xlm/>
 - Exemplarisches Federated Learning-Beispiel für eine Document Recognition Lösung auf dem Simulator mit Skalierung von Clients mit Nvidia-Flare und dem Donut-Modell (federführend durch Condat): <https://www.fedxtract.net/beispiele/beispiel-donut-modell/>
- Weiter entwickelte und verbesserten KI- und Trainings-Methoden sowie fine-getunte Modelle (z.B. Donut und LayoutXLM).

Mit dieser Best Practice unterstützt FedXtract die Nutzung der neuen KI/FL-Technologien für kleinere Unternehmen wie KMUs um zusammen mit anderen Organisationen KI-Modelle für ihren Bedarf zu trainieren. Die Best Practice und exemplarischen Beispiele bilden die Basis für die Verwertung der beiden KMU Partner, die dazu Beratung und Customizing für die Industrie, Medien, Healthcare und den öffentlichen Dienst anbieten werden. Auch wenn die Weiterentwicklung von KI- und Trainings-Methoden nicht primäres Ziel des Projektes war, hat es einige Verbesserungen insbesondere bei den Training-Methoden gegeben, die zusammen mit den fine-getunten Modellen (Donut, LayoutXLM) kostenpflichtig vermarktet werden können.

Über die Vorstellung des Projektes auf der FedXtract Website <https://www.fedxtract.net/> und Veranstaltungen (u.a. IBC, IPTC, FKTG, Xinnovations, Semantic Media Web, Meetups) online und vor-Ort gab es viele Begegnungen mit anderen Firmen und Organisationen, die innovative Konzepte im Bereich KI, FL und ML/DL nachfragen. Dabei hatte Condat u.a. Erfahrungsaustausch hinsichtlich der Konzepte und Ausnutzung weiterer Potentiale mit Firmen, die ihre Themen- und domänenspezifische Expertise einbringen konnten. Dazu gab es Gespräche mit der FKTG, DAI-Labor/Berlin, dem W3C-Büro, der EBU und dem IRT.

Verbundvorhaben FedXtract
Föderierte KI-Lösung zur Verarbeitung komplexer Dokumente

Schlussbericht 2024

Condat AG

FKZ: 01IS21037A

Schlussbericht - Teil II: Eingehende Darstellung

Das diesem Bericht zugrunde liegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01IS21037A gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei der Autorin/beim Autor.

Projektträger

Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR)

Zuwendungsempfänger: Condat AG, Alt-Moabit 91d, 10559 Berlin	Förderkennzeichen: 01IS21037A
Projektleiter: Rolf Fricke	Tel.: +49 172 3991959 E-Mail: rf@condat.de
Laufzeit des Vorhabens: 01.11.2021 bis: 30.04.2024	
Datum: 18.08.2024	

Projektpartner:

1. Condat AG
2. ContFraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS
3. ControlExpert GmbH

1. Im Vorhaben durchgeführte Arbeiten, insbesondere im Vergleich zur ursprünglichen Vorhabensbeschreibung

Im Projekt wurde zunächst von allen Partnern eine übergreifende Architektur für die FL-Lösung entwickelt, auf deren Basis die FL/DL/ML Verfahren sowie die Anwendungsbeispiele konzipiert, entwickelt und erprobt werden konnten. Nachfolgende Arbeiten wurden entsprechend dem Projektplan von Condat durchgeführt.

AP1 – Experimentelle Software Entwicklung

In den Tasks 1 und 2 wurde von Condat zusammen mit den Partnern die Softwarearchitektur aufgesetzt und konfiguriert, in der alle Komponenten des FL-Frameworks mit den verschiedenen Open Source Tools, die das Projekt benötigt, ablauffähig sind. Condat hat eine Auswahl und Untersuchung der zu diesem Zeitpunkt aktuellen Technologien und Tools zum Komponenten Framework, Training und Prozesskontrolle durchgeführt. Dabei wurden auch die bei den Projektpartnern bereits genutzten Technologien und vorhandenes Know-how berücksichtigt. Zunächst wurden die verfügbaren OS FL-Frameworks gesichtet, die als Basis für das Projekt in Frage kamen: <https://www.fedxtract.net/best-practice/##fl5>. Auf Basis der Anforderungen für das Projekt und einer Liste von Auswahlkriterien kamen nach eingehender Analyse drei Frameworks in eine Vorauswahl, die FL-Frameworks von Flower, PySift und Nvidia. Nach intensiver Analyse inklusive Testinstallationen schnitt das Framework Nvidia Flare am besten ab. Dieses Framework wurde dann ab Q2 installiert und die für das Projekt erforderliche verteilte Konfiguration erstellt, um verbliebene offene Fragen zu klären. Hier waren insbesondere die Randbedingungen und Anforderungen von den parallel durch die Partner entwickelten Modellen, Trainingspipelines, DL-Technologie und Validierungskonzept zu analysieren. Das Ergebnis war der Aufbau eines FL-Frameworks auf Basis von Nvidia Flare, mit dem ab Q1/2023 eine unter den Partnern verteilte Trainings-Pipeline aufgebaut wurde, durch die ein zyklisches, verteiltes Training eines gemeinsamen DL-Modells möglich wurde.

In diesem AP wurden dann von Condat die Komponenten der Softwarearchitektur entwickelt und schrittweise integriert. Dazu gehören das FL-Framework mit den verschiedenen Open Source Tools, die das Projekt benötigt, um die vorgesehenen Use Cases zu realisieren. Das Framework Nvidia Flare wurde stufenweise in Betrieb genommen und die für das Projekt erforderliche verteilte Konfiguration erstellt. Dabei wurde sowohl der Modus mit physisch über die Partner verteilten Clients eingerichtet, wie auch der Simulator in Betrieb genommen, der einen Test mit 2 bis 100 Clients auf einem einzigen Server erlaubt. Hier waren insbesondere die Randbedingungen und Anforderungen von den parallel durch die Partner entwickelten Modellen und Trainingspipelines zu berücksichtigen.

AP2 – Federated und Transfer Learning Komponenten

Condat hatte in diesem AP die Aufgabe, die kontinuierlich neue Forschung in den Bereichen FL (Federated Learning), DL (Deep Learning), ML (Maschine Learning) und TL (Transfer Learning) zu verfolgen und daraus für FedXtract geeignete Konzepte aufzugreifen. Condat hat in der Task 1 das ausgewählte FL-Framework Nvidia Flare weiter konfiguriert, so dass es für die verschiedenen Evaluationsszenarien eingesetzt werden kann. Dabei waren die Netzwerkkonfigurationen bei den 3 Partnern zu konfigurieren, so dass eine sichere Kommunikation und Datenaustausch möglich wurde. Zudem wurden die verschiedenen Mechanismen integriert, die von Nvidia Flare angeboten werden, um die unterschiedlichen Konfigurationen mit Clients aus spezifischen Anwendungskontexten zu unterstützen. Z.B.

werden die Möglichkeiten des FL-Frameworks genutzt, um bei den verschiedenen FL-Clients heterogene Trainingsdaten einzubinden, was insbesondere beim ControlExpert Anwendungsfall auftreten kann. In diesem Fall kann durch das Nvidia Flare ein Balancing der Daten veranlasst werden. Weiterhin wurde untersucht, welche Sicherheitsmechanismen zur Gewährleistung der Integrität und Vertraulichkeit vom FL-Framework genutzt werden sollten, so dass die bei den einzelnen Clients verwendeten Trainingsdaten sicher geschützt werden können.

AP3 – Entwicklung eines Deep Learning-basierten Modells für den Use Case Document Recognition

Condat hat zusammen mit den Partnern verschiedene Modelle und Trainings-Szenarien ausgearbeitet um die Leistungsfähigkeit der verfügbaren ML Verfahren und Methoden zum Dokument Recognition analysieren. Hierbei wurde zum einen ein Ansatz auf Basis des LayoutXLM Modells (https://huggingface.co/docs/transformers/model_doc/layoutxlm) und zum zweiten Donut, ein OCR freies Visual Understanding Modell (<https://github.com/clovaai/donut>) ausgewählt. Es wurden verschiedene Testkonfigurationen mit zentraler- und verteilter Datenhaltung verwendet sowie Tests auf dem Nvidia Simulator durchgeführt, der auch Tests mit einer großen Anzahl von Clients ermöglicht. Dabei wurden zahlreiche Testreihen mit verschiedenen Hyperparametern sowie variablen Mengen und Aufteilungen von Trainingsdaten durchgeführt. Zur Evaluierung wurde eine Metrik auf Basis eines F1-Score verwendet, der mittels Precision und Recall ermittelt wird. Damit können die Verbesserungen eines Modells durch die Trainings-Zyklen in einem bestimmten Anwendungskontext gemessen werden. Auf diese Weise wurde untersucht, wie gut die verschiedenen Trainings-Pipelines für unterschiedliche Arten von Daten (Adressen, Bankverbindung, Kfz-Teile, Rechnungen, etc.) ist.

AP4 – Content für Training und Evaluation

In diesem AP haben die Partner Datensätze zusammengestellt, um die entstandene FL-Umgebung zu validieren. Dazu wurde ein Trainingsplan erarbeitet, der verschiedene Sätze von frei verfügbaren und von CE und FHG eingebrachten Trainings- und Validierungsdaten umfasst. Dazu gehörten zunächst die von ControlExpert aus dem KFZ-Bereich eingebrachten umfangreichen Daten. Condat hat in der Task 2 größere frei verfügbare Datenmengen wie [CIFAR-10](#), [Stanford Cars](#) und [CINIC-10](#) analysiert und geeignete Datensatz-Mengen davon ausgewählt. Hierbei wurde auch entschieden, welche Aufteilungen von Daten auf die verschiedenen Clients für die verteilten FL-Trainings genutzt werden und wie heterogen diese strukturiert sein sollten.

Ab Q1/2023 wurde von den Partnern eine verteilte Trainings-Pipeline aufgebaut, und ein zyklisches, verteiltes Training von gemeinsamen DL-Modellen durchgeführt. Im letzten Projektjahr wurden dann die 2 exemplarischen Beispiele für ein verteiltes Training von KMUs entwickelt, die als Paradigma für das Aufsetzen eines Frameworks und verteiltes Training dienen

Vergleich zur ursprünglichen Vorhabensbeschreibung

Da sich das Technologie-Segment FL schnell weiter entwickelt, hat sich die Situation ergeben, dass die vorhandenen FL Frameworks Ende 2023 bereits eine gute Sicherheitsinfrastruktur bereitstellen. Damit können die nach der Best Practice aufgesetzten FL-Projekte bereits über mehr Sicherheits-Mechanismen verfügen, als wir zu Projektbeginn vorausgesehen hatten.

Ein zweite Entwicklung, von der das Projekt profitieren konnte, sind die schnelle Entwicklung von Modellen und deren Bereitstellung als OS. Z.B konnten wir auf die Modelle LayoutXLM und Donut

aufsetzen, die den nutzenden KMUs bereits ein leistungsfähigeres Document Recognition erlaubt, als wir zu Projektbeginn erwartet hatten.

2. Die wichtigsten Positionen des zahlenmäßigen Nachweises

Es gab auf Grund der Corona-Richtlinien geringfügige Veränderungen im Projektablauf, z.B. führten die fehlenden vor-Ort Termine zu einigen Kommunikations-Hindernissen. Insbesondere aufgrund der Corona-Richtlinien gab es 2 relevante Änderungen bzgl. der geplante Aufwände, zu denen bereits in den Zwischenberichten 2022 und 2023 eine Mittel-Umwidmung beantragt wurde:

- Reisekosten wurden zu erhöhten Personalaufwänden umgewidmet und
- ein Teil der Cloud-Rechenzeit Kosten wurde in mehr Personalaufwände umgewidmet.

Für die reduzierten Kosten für Reisen und Nutzung von Cloud-Services (F8050) wurde eine Umwidmung in Personalkosten beantragt, da für die Einsparungen auch jeweils erhöhte Personalaufwände erforderlich waren.

Letztlich gab es jedoch keine Probleme bzgl. der kooperativen Konzeption und Entwicklung mit den Partnern. Die Arbeits-, Zeit- und Ausgabenplanung wurde über die 30 Monate hinweg weitgehend eingehalten, alle Meilensteine konnten erreicht und die geplanten Ergebnisse fertiggestellt werden.

Darüber hinaus gab es lediglich einige Änderungen im Arbeitsplan, die sich aus dem Aufkommen neuer KI-Technologien ergeben hatten. Diese Anpassungen konnten jedoch ohne Probleme erfolgen, da die Konsortialpartner das neue Know-How über die neuen KI Methoden kontinuierlich in das Projekt transferiert hatten.

3. Die Notwendigkeit und Angemessenheit der geleisteten Arbeit

Aufgrund der hoch innovativen KI-, FL- und ML- Technologien bestand ein technisches und wirtschaftliches Risiko, das Condat alleine aus seinen Kundenprojekterlösen nicht tragen konnte und somit die für das Vorhaben beantragten Fördermittel benötigte. Die Entwicklung der Komponenten, Best Practice und Beispiele zum verteilten Training waren notwendig, da solche Funktionen und Komponenten derzeit noch nicht von anderen System-Anbietern verfügbar sind. Dies liegt vor allem daran, dass es sich bei den verwendeten KI/FL/ML- Methoden um sehr innovative Technologien handelt, die unter den für diese Lösung vorliegenden Bedingungen bisher noch nicht erprobt waren.

3. Der voraussichtliche Nutzen, Verwertbarkeit, fortgeschriebener Verwertungsplan

Allgemeine Trends

Es waren einige generelle Trends zu beobachten, die zum Teil auch Auswirkungen der Covid-19 Pandemie sind, die Folgen für die Verwertung der Projektergebnisse haben könnten:

- wachsende Anforderungen an Datenschutz und -Sicherheit
- Aufkommen neuer Technologien im FL/ML/DL Bereich, wie die Transformer Technologien
- steigende Patent/Copyright Risiken und
- erhöhte Handelsbarrieren.

Diese Trends sind für die Verwertung zu berücksichtigen, haben aber keine Auswirkung auf unsere Vermarktungsstrategie.

Wirtschaftliche Verwertbarkeit

Condat strebt an, die entstandenen Ergebnisse sowohl individuell sowie auch mit einzelnen Partnern zu verwerten. Dabei ist nachfolgend vor allem die individuelle Verwertung dargestellt, bei der folgende Projektergebnisse von FedXtract verwertet werden sollen:

Basis der Verwertung sind:

- Best Practice für KMU zur Planung und Durchführung eines Federated Learning-Projektes: <https://www.fedxtract.net/best-practice/> mit der Auswahl, Aufsetzen und Realisierung
 - ML-Methoden
 - FL-Framwork
 - Trainings-Netzwerk und
 - verteiltem Training.
- zwei exemplarische Anwendungs-Beispiele.

Die Best Practice und exemplarischen Anwendungsbeispiele ermöglichen Condat ihre bereits erfolgreich im produktiven Einsatz befindliche SME (Smart Media Engine) Produktlinie zu erweitern um:

- Consulting zur Beratung zur Planung und Durchführung eines konkreten FL-Trainings
- Customizing zur Realisierung von Anpassungen und Erweiterungen für eine ML/DL Lösung
- Verwertung der weiter entwickelten und verbesserten KI- und Trainings-Methoden sowie fine-getunten Modelle und (Donut, LayoutXLM).

Die SME wird bereits als SW-Lizenz zusammen mit einem Consulting und Customizing vermarktet. Da die Einführung einer FL-Lösung eine umfangreiche Analyse, Konfiguration und Einrichtung von der Planung bis hin zum Trainingsbetrieb erfordert, erwarten wir in diesem Bereich die größten Umsätze. Eine Verwertung mit einem anderen der beiden Projektpartner ist durch die sich ergänzenden IT-Lösungen und Kompetenzen sinnvoll. Auch wenn die Weiterentwicklung von KI- und Trainings-Methoden nicht primäres Ziel des Projektes war, hat es einige Verbesserungen insbesondere bei den Training-Methoden gegeben, die zusammen mit den fine-getunten Modellen (Donut, LayoutXLM) kostenpflichtig vermarktet werden können.

Condat hat bereits damit begonnen, die entstandene Best Practice Lösung bis hin zur Erreichung der Produktreife weiterzuentwickeln. Damit kommt zu den Kosten des geförderten Projektes noch einmal ein beträchtlicher Aufwand hinzu und zudem sind weitere Ausgaben für Marketing- und Vertriebsaufwendungen kalkuliert.

Verwertungsmöglichkeiten und perspektivisch aussichtsreiche Märkte

Durch das Aufkommen von ChatGPT und der generativen KI befindet sich der gesamte Markt einschließlich aller Teilbereiche wie KI, ML, DL und FL im Aufwind. Eine Reihe von ganz aktuellen Studien untersucht die Auswirkungen der KI, ML/DL auf die aktuellen Märkte im Bereich der Content Erstellung, Content Management- und Informationssysteme, die vom der neuen FL-Lösung adressiert

wird. Alle Marktforschungsinstitute prognostizieren ein starkes Wachstum, wie z.B. eine Studie von Statista, die für die KI, ML, DL ein jährliches Wachstum von über 46% von 19 Mrd € in 2023 auf über 330 Mrd € bis 2030 weltweit¹ voraussagt. Dies entspricht etwa der Einschätzung von Cap Gemini, die bis 2032 ein jährliches Wachstum des KI-Marktes von über 30% erwarten². Die Bitkom veröffentlichte für Deutschland die Ausgaben für KI-Lösungen, die in 2023 um 32 Prozent auf 6,3 Milliarden Euro angewachsen sind. Dabei entfielen 4,1 Milliarden Euro auf KI-Software, gefolgt von Dienstleistungen 1,3 Milliarden Euro und Hardware 0,9 Milliarden Euro. Für 2024 wird ein Wachstum von 30 Prozent auf 8,2 Milliarden Euro erwartet und 2025 die 10-Milliarden-Euro-Marke übersteigen³. Nach diese Untersuchungen umfassen die Dienstleistungen vor allem Content-Optimierung, Personalisierung und Automatisierung von Routineaufgaben⁴. Diese Bestrebungen befinden sich jedoch noch in den Anfängen und sind bisher kaum in den konkreten Arbeitsprozess integriert.

Wenn man die Zahlen der Bitkom von ca. 10 Mrd € in 2025 für Deutschland herunterbricht, erhält man ca. 1,8 Mrd € für KI-Dienstleistungen Deutschland, von denen nach unserer Abschätzung mindestens 10%, d.h. 180 Millionen € für die Entwicklung und Aufbau von verteilten Trainings ausgegeben werden, so dass ausreichend Budget für die von FedXtract entwickelten Leistungen vorhanden ist.

Zielmärkte

Die entwickelte Lösung bietet mit seinen zahlreichen Komponenten eine gute Grundlage für die Einbindung und Nutzung im Medien / TV Bereich. Das im Projekt entstandene System ist zunächst für die Vermarktung im TV / Medienbereich konzipiert, kann jedoch auch für andere Bereiche wie Publishing, Journalismus, Mediatheken oder andere kommerzielle Websites verwendet werden. Damit ergänzt die entstandene Lösung die von Condat bereits erfolgreich vermarktete Smart Media Engine⁵ und kann wie vorgesehen verwertet werden. Letztlich können natürlich auch verwandte Bereiche in der Industrie, Healthcare oder Pharma anvisiert werden, jedoch hat Condat dort weniger Partner und kennt die Datenbestände und Formate nicht so gut. Deshalb planen wir aufgrund unserer Marktbeobachtung und Kenntnisse für das vorliegende Projekt zunächst nachfolgende Märkte für Mitarbeiter- und Kunden-Portale zu adressieren, in denen wir für die nächsten 10 Jahre eine hohe Nachfrage nach innovativen Funktionen zum verteilten Training erwarten:

- **Nachrichtenagenturen oder Redaktionen im Medien/TV- Bereich** für die Erstellung von Nachrichten und Berichte, wobei diese ihre Inhalte in der Regel auch zusätzlich über das Internet, mobil und in der Tagespresse verbreiten. Bei der Bereitstellung der Funktionen für

¹ <https://de.statista.com/outlook/tmo/kuenstliche-intelligenz/generative-ki/weltweit> , März 2024

² <https://www.gminsights.com/de/industry-analysis/generative-ai-market> , 2024

³ <https://www.bitkom.org/Presse/Presseinformation/KI-Markt-waechst-um-ein-Drittel> , 2023

⁴ <https://cms-cdn.lmu.de/media/04-som/dss/downloads/empfehlungen-zum-umgang-mit-generativen-und-anderen-ki-tools.pdf>

⁵ siehe Condat Portfolio: <https://www.condat.de/loesungen/smart-media-engine/>

Newsrooms ist zu beachten, dass dort rund um die Uhr sehr schnelle Reaktionszeiten erforderlich sind und oft eine Ad-hoc Nutzung zur Gewinnung eines Überblicks über die bereitstehenden Quellen angeboten werden sollten.

- **Redaktionen von TV-/Video-Produktionen oder weitere ePublishing Anbieter**, die TV-Sendungen oder tiefergehende Publikationen planen, erstellen oder aufbereiten. Hier werden in Zeiträumen von 3 Tagen bis 5 Monaten Analysen und Reportagen für verschiedenste Bereiche erstellt, in denen eine hohe Qualität und Korrektheit hinsichtlich der dargestellten Informationen Priorität haben. Dazu werden von der entstandenen FedXtract-Lösung umfangreichere Funktionen mit mehr Parametern und besseren Erläuterungen der Ergebnisse angeboten.
- AV Inhalte-Anbieter wie **WebTV, Mediatheken** und **mobile Portale**, die über Internet oder mobile Distributionskanäle AV Materialien verbreiten. Dort werden sowohl Nachrichten wie auch kurze und tiefergehende Berichte erstellt. In der Regel werden dort einfachere Workflows verwendet und mehr „user generated content“ (z.B. aus Twitter, Facebook, TikTok) genutzt, so dass vor allem eine Validierung der Korrektheit von Informationen bzgl. Ort, Zeit und Quellen sowie eine Rechte- und Kostenanalyse zu unterstützen ist.

Wettbewerb

Seit Projektbeginn sind inzwischen 3 Jahre vergangen und im Bereich KI, FL, ML/DL haben sich Technologie und Märkte rasant weiterentwickelt. Es gibt inzwischen die verschiedenen FL-Frameworks wie PySift, Tensorflow, Nvidia und Flower, die einen Teil der Anforderungen von KMU abdecken können. Allerdings bleibt das Aufsetzen eines solchen FL-Projektes eine technische, organisatorische und finanzielle Herausforderung. Angesichts dieser Entwicklung hat sich das Konsortium darauf konzentriert, den KMU den Einsatz ein solches Framework zu empfehlen, und darauf aufbauend eine Best Practice für das Planen, Aufsetzen und Durchführen eines FL-Trainings mit exemplarischen Beispielen zu entwickeln. Die verschiedenen Anbieter der FL-Frameworks aus der Tabelle sind jedoch auch als unser Mitbewerb anzusehen. Hierbei ist zunächst zu berücksichtigen, dass einige Anbieter aus China kommen (z.B. WeBank <https://www.fedai.org>) und europäische KMUs vor einer solchen Lösung zurückschrecken werden. Bei anderen (z.B. Flower) handelt es sich um Projekte, so dass die Verfügbarkeit und Support für deren Produkte noch nicht absehbar ist. Dazu gibt es kostenpflichtige KI-Plattformen wie modzy (<https://www.modzy.com/>) die sogenannte ML-Marketplaces bieten, von denen bestehende ML-Modelle mittels Container-Technologie „gemietet“ oder dort deployed werden können. Letztlich richten sich jedoch alle FL-Framework Anbieter wie auch PySift, TensorFlow und Nvidia an professionelle System-Integratoren und Entwickler. Deshalb ist die Strategie von FedXtract, dass die KMUs ein solches Framework nutzen, und Condat oder die anderen FedXtract Partner ihnen ein Consulting und Customizing zur Planung und Aufsetzen des FL-Trainings einschließlich der Auswahl von Modellen und Trainingsstrategie anbieten. Hierbei fallen für die KMU keine Kosten für die OS FL-Frameworks an, sondern nur die Kosten für das Consulting und Customizing. Zudem können einzelne Modelle kostenpflichtig sein, andere wiederum auch frei zur Verfügung gestellt werden.

Bezüglich **Texterkennung und speziell Document Recognition** gibt es eine Reihe von kommerziellen Anbietern wie Google oder ABBYY. FedXtract bietet hier zunächst die kostenfreien OS Modelle von Donut oder LayoutXLM an, mit denen wir in unserer Validierung und den Beispielen gute Ergebnisse

erzielt haben. Dabei behalten die KMUs die volle Kontrolle über die Speicherung, Nutzung und Lizenzrechte der erzeugten Ergebnisse. Dies ist ein großer Vorteil für viele Unternehmen, die strengen Datenschutz-Regelungen unterliegen, die eine Nutzung von Cloud-Diensten im Zusammenhang mit sensiblen Daten nicht erlauben.

Verwertung durch Condat

Die aus der Projektentwicklung entstandenen Ergebnisse gliedern sich mit ihrem Customizing- und Consulting Portfolio nahtlos in die Produktlinie Smart Media Engine ein, die in verwandten Bereichen bereits erfolgreich im Einsatz ist. Sie ergänzt somit unsere Kernprodukte und Kompetenzen. Aktuell hat Condat ca. 18 Mitarbeiter im Bereich TV / Medien beschäftigt, die einen jährlichen Umsatz von 2 Mio € erwirtschaften. Dieser Bereich soll u.a. gestützt durch die FedXtract Ergebnisse ausgebaut werden, wobei insbesondere folgende Effekte zum Tragen kommen:

- das angewachsene KI- und FL/ML/DL Know-How führt zu einem breiteren Leistungsspektrum und einer Ausweitung der Condat Aktivitäten
- durch einen erhöhten Bekanntheitsgrad und besseren Sichtbarkeit von Condat als Dienstleister für KI und FL/ML/DL basierte Systeme (z.B. durch die Präsentationen und Veranstaltungen) ergeben sich zusätzliche Consulting Projekte und
- durch die enge Kooperation und Synergie mit den Projektpartnern FhG und ControlExpert ergeben sich vermehrte Kontakte zu neuen Zielgruppen und in der Folge zusätzliche Projekte.

Für die Entwicklung der Anzahl der Mitarbeiter und des Umsatzes ist für die folgenden 5 Jahre nach Projektabschluss eine jährliche Steigerung von ca. 4 % geplant, die sich zu einem großen Teil aus dem FedXtract Projekt und den darin entwickelten FL/ML/DL-gestützten Verfahren ergeben.

Akquisitions- und Informationsaktivitäten

Condat wird über sein gutes Netzwerk, das auf langjährige Projekte, Partner und Organisationen aufbaut, die Verwertung in folgenden Bereichen vorantreiben:

- Medien mit TV/ Radio und Film, TV/Radio-Sender und Medien-Unternehmen: Projekte und Kontakte bei RBB, LFM, ZDF, Bayerischer Rundfunk, WDR, NDR, MDR, N-TV, DVB, Deutsche Welle
- ePublishing, Medien-Produktionen, Verlage: Projekte und Kontakte bei Cornelsen, Klett, Westermann, Springer und Holtzbrinck sowie
- öffentlicher Dienst mit Berlin- / Brandenburger Behörden sowie Bundesministerien.

4. Der während der Durchführung des Vorhabens dem Zuwendungsempfänger bekannt gewordene Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen

Während unserer kontinuierlichen Marktbeobachtung während der Projektlaufzeit konnten wir keine neuen Wettbewerber oder Lösungen erkennen, die die Verwertung des Vorhabens beeinträchtigen könnten.

Das Auftauchen neuer Technologien, wie z.B. Transformer Methoden im Bereich von ML/DL, deren Relevanz zu Projektbeginn noch nicht erkennbar war, spielt sicherlich eine wichtige Rolle. Dies führte jedoch dazu, dass die wichtigsten dieser neuen Technologien aufgegriffen und in den von Condat entwickelten Komponenten zu deren Verbesserung integriert wurden.

In unserer Analyse konnten wir feststellen, dass führende Anbieter von Content Management- und Redaktions-Systemen vorsehen, neue KI basierte Funktionen z.B. zum automatisierten Journalismus, Sprachassistenten, Storytelling und Social-Media Analyse einführen, die die Planung eines Trainings und auch verteilten Trainings erfordern. Aufgrund der dynamischen Entwicklung im TV-Medienbereich, wie z.B. dass unsere Kunden LFM (Landesmedienanstalten), ZDF, WDR, Deutsche Welle und rbb immer mehr Inhalte und Angebote auch im Web anbieten, besteht ein großes Interesse an den im vorliegenden Projekt entstandenen Know-how sowie den FL/ Trainings-Methoden.

Damit ergibt sich letztlich, das wir mit unserer Strategie auf einem guten Weg sind. Die Best Practice von FedXtract und Erfahrung von Condat ermöglichen, eine durchgängige Lösung zu konzipieren und ausreichende Datenmengen zum Training der FL/KI Funktionen bereitzustellen.

6. Angabe der verwendeten Fachliteratur sowie der benutzten Informations- und Dokumentationsdienste

[Adesso, 2020] Adesso Report “KI-eine Bestandsaufnahme”(2020), URL: <https://ki.adesso.de/ki-de/oeffentliche-downloads/ki-eine-bestandsaufnahme-2020.pdf>

[AI-Act, 2024] <https://digital-strategy.ec.europa.eu/de/policies/european-approach-artificial-intelligence>

[Amazon, 2024] Amazon Web Services, URL: <https://aws.amazon.com/de/>

[Bitkom, 2019] Studie Digital Office im Mittelstand, URL: https://www.bitkom.org/sites/default/files/2019-10/191021_studie_digital-office-im-mittelstand.pdf

[Bitkom, 2023] Andreas Streim, www.bitkom.org/Presse/Presseinformation/ChatGPT-Jedes-sechste-Unternehmen-plant-KI-Einsatz-Textgenerierung

[Datenschutzexperte, 2023] <https://www.datenschutzexperte.de/leitfaden-chatgpt/>

[DSA, 2024] <https://digital-strategy.ec.europa.eu/de/policies/digital-services-act-package>

[Devlin, 2019] Devlin et al. (2019), BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, URL: <https://www.aclweb.org/anthology/N19-1423/>

[DSGVO, 2024] Datenschutz-Grundverordnung, <https://dsgvo-gesetz.de/>

[Geyer, 2017] Geyer et al. (2017), Differentially private federated learning: A client level perspective, URL: <https://arxiv.org/abs/1712.07557>

[EY, 2019] Artificial Intelligence in Europe – How 307 Major Companies Benefit from AI, URL: <https://news.microsoft.com/de-de/ki-studie-kuenstliche-intelligenz-steht-bei-deutschen-unternehmen-hoch-im-kurs/>, Exzerpt URL: <https://www.ind-ai.net/geschaeftsprozesse/kuenstliche-intelligenz-hat-hohe-prioritaet/>

[Fraunhofer, 2020] URL: <https://www.fraunhofer.de/de/presse/presseinformationen/2020/september/dokumente-schneller-analysieren-mit-kuenstlicher-intelligenz-von-fraunhofer.html>

[G1, 2020] <https://www.heise.de/newsticker/meldung/Gesellschaft-fuer-Informatik-zeichnet-10-praegende-Koepfe-der-KI-aus-4659250.html>

[Google, 2020] Google Vision AI/Google Document AI, URL: <https://cloud.google.com/vision?hl=de> und <https://cloud.google.com/solutions/document-ai?hl=de>

[Google AI Blog, 2017] Federated Learning: Collaborative Machine Learning without Centralized Training Data, URL: <https://ai.googleblog.com/2017/04/federated-learning-collaborative.html>

[Handelsblatt, 2023] Was Unternehmen bei der Einführung der KI beachten sollten, <https://www.handelsblatt.com/technik/ki/chatgpt-was-unternehmen-bei-der-einfuehrung-der-ki->

[beachten-sollten/29353500.html](#)

[Hooker, 2020] Hooker (2020), The Hardware Lottery, URL: <https://arxiv.org/pdf/2009.06489.pdf>

[JI, 2023] “Anbieterverzeichnis für KI-Anwendungen”, <https://www.journalismuslab.de/ki-anbieter-verzeichnis/> (Stand 2024)

[Kairouz, 2019] Kairouz et al. (2019), Advances and open problems in federated learning, URL: <https://arxiv.org/abs/1912.04977>

[Manager Magazin, 2020] URL: <https://www.manager-magazin.de/unternehmen/industrie/kuenstliche-intelligenz-knapp-sechs-prozent-der-firmen-sind-nutzer-a-1305836.html>

[Mikolov, 2013] Mikolov et al. (2013), Distributed representations of words and phrases and their compositionality, URL: <http://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and>

[Namysl, 2020] Namysl, Behnke, Köhler (2020), NAT: Noise-Aware Training for Robust Neural Sequence Labeling, URL: <https://www.aclweb.org/anthology/2020.acl-main.138>

[NVIDIA Developer Blog, 2024] URL: <https://developer.nvidia.com/blog>

[Schmidhuber, 2014] Schmidhuber (2014), Deep Learning in Neural Networks: An Overview, URL: <https://arxiv.org/abs/1404.7828>

[Stanoevska-Slabev, 2016] Commercialization of Composite Software Resulting from Collaborative Research, eChallenges e-2015, Katarina Stanoevska-Slabev (MCM University of St. Gallen), Rolf Fricke (Condat AG), ISBN: 978-1-905824-53-3, 2016, Vilnius, Lithuania

[Prelle, 2019] Sacha Prelle, Condat AG, Externes Wissen: Neue, AI-basierte Methoden zur Klassifikation und Anreicherung von Inhalten. In FKTG-Magazin 04/2019, <https://www.fkt-online.de/news/news-detail/25546-externes-wissen-neue-ai-basierte-methoden-zur-klassifikation-und-anreicherung-von-inhalten/>

[Thomsen, 2020] Jan Thomsen, Rolf Fricke (Condat): Qurator – AI based Curation Technologies for News Room Systems, Präsentation EBU (European Broadcast Union), Metadata Developers, 2020