



Mittelstand-Digital  
**Zentrum  
Hannover**

# Zukunft.Digital

Digitalisierung  
von der Idee zur Umsetzung  
**Ausgabe 01/2026**



**Aus der Praxis**

**Mit Plan zur Präsenz –  
Social Media strategisch  
gedacht**

Seite 06

**Aus der Praxis**

**Digitalisierung der Abfall-  
logistik in temporären  
Ballungszentren**

Seite 18

**Aus der Praxis**

**Zuverlässige  
Finanzprognosen mit  
Künstlicher Intelligenz**

Seite 30



Das Mittelstand-Digital Netzwerk bietet mit den **Mittelstand-Digital Zentren** und der **Initiative IT-Sicherheit in der Wirtschaft** umfassende Unterstützung bei der Digitalisierung. Kleine und mittlere Unternehmen profitieren von konkreten Praxisbeispielen und passgenauen, anbieterneutralen Angeboten zur Qualifikation und IT-Sicherheit. Das Bundesministerium für Wirtschaft und Energie ermöglicht die kostenfreie Nutzung der Angebote von Mittelstand-Digital.

Weitere Informationen finden Sie unter:  
[www.mittelstand-digital.de](http://www.mittelstand-digital.de)

Gefördert durch:



Bundesministerium  
für Wirtschaft  
und Energie

Mittelstand-  
Digital



aufgrund eines Beschlusses  
des Deutschen Bundestages

# Zukunft.Digital

Digitalisierung  
von der Idee zur Umsetzung  
**Ausgabe 01/2026**

# Inhalt

## Aus der Praxis

**Mit Plan zur Präsenz –  
Social Media strategisch gedacht**  
Seite 06

**Interview:  
Die größte Herausforderung ist  
der Spagat zwischen Effizienz  
und Authentizität**  
Seite 10

**Flexibel, nachhaltig, smart:  
Lademanagement neu gedacht**  
Seite 12

**Nachhaltigkeit messbar machen:  
KI-gestützte Evaluierung von Patenten  
anhand der 17 UN-Nachhaltigkeitsziele**  
Seite 16

**Digitalisierung der Abfalllogistik  
in temporären Ballungszentren**  
Seite 18

**Wissensmanagement:  
KI-basierter Chatbot liefert unter-  
nehmensinterne Informationen**  
Seite 22

**Automatisierte Vertragsanalyse  
mittels KI bei der Prowind GmbH**  
Seite 26

**Zuverlässige Finanzprognosen  
mit Künstlicher Intelligenz**  
Seite 30

**Kalk, Sand und KI: Multisensorische  
Defekterkennung für die Produk-  
tion mineralischer Baumaterialien**  
Seite 34

**KI für schnellere Angebote:  
Zeichnungen auslesen, Herstell-  
zeiten prognostizieren, Risiken  
quantifizieren**  
Seite 38



# Editorial



**Prof. Dr.-Ing. Berend Denkena**  
Konsortialleiter des  
Mittelstand-Digital Zentrums  
Hannover

Diese Ausgabe von Zukunft.Digital ist eine besondere: Sie ist unsere letzte Magazin-Ausgabe in der aktuellen Förderperiode, die im Mai 2026 endet. Anlass genug, auf mehr als ein Jahrzehnt intensiver Zusammenarbeit mit zahlreichen Unternehmen zurückzublicken.

Seit Dezember 2015 begleitet das Zentrum Unternehmen dabei, strategisch sinnvoll in innovative Technologien und KI-Anwendungen zu investieren und sich zukunftssicher aufzustellen. Passgenau, anbieterneutral und nutzenorientiert unterstützen wir u. a. mit Informationsveranstaltungen, in Workshops und Firmengesprächen.

Insgesamt 80 gemeinsame Projekte mit Unternehmen sind seit dem Start unseres Zentrums im Jahr 2015 entstanden. Neun Projekte, die in der jüngeren Vergangenheit realisiert wurden, stellen wir Ihnen in dieser Ausgabe vor. Sie stehen exemplarisch für unsere Arbeit: technologieoffen, anwendungsnah und stets am konkreten Bedarf der Unternehmen orientiert.

Die Bandbreite der Themen zeigt eindrucksvoll, wie vielfältig digitale Lösungen im Mittelstand eingesetzt werden können –

vom KI-gestützten Lademanagement über intelligente Wissenssysteme sowie automatisierte Qualitäts- und Defekterkennung bis hin zu datenbasierten Prognosen in Finanzen, Produktion oder Logistik. So unterschiedlich die Anwendungsfelder sind, eines verbindet alle Projekte: Digitalisierung und KI entfalten ihren Nutzen dort, wo sie aktuelle Probleme lösen und den Arbeitsalltag der Mitarbeitenden in den Unternehmen spürbar verbessern.

Diese Ausgabe markiert jedoch keineswegs einen Abschluss, sondern vielmehr einen Übergang. Auch in Zukunft möchten wir für Sie da sein und Sie bei der digitalen Transformation unterstützen. Hierfür richten wir uns aktuell neu aus. Schauen Sie gern weiterhin auf unserer Website vorbei, um immer auf dem aktuellen Stand zu sein.

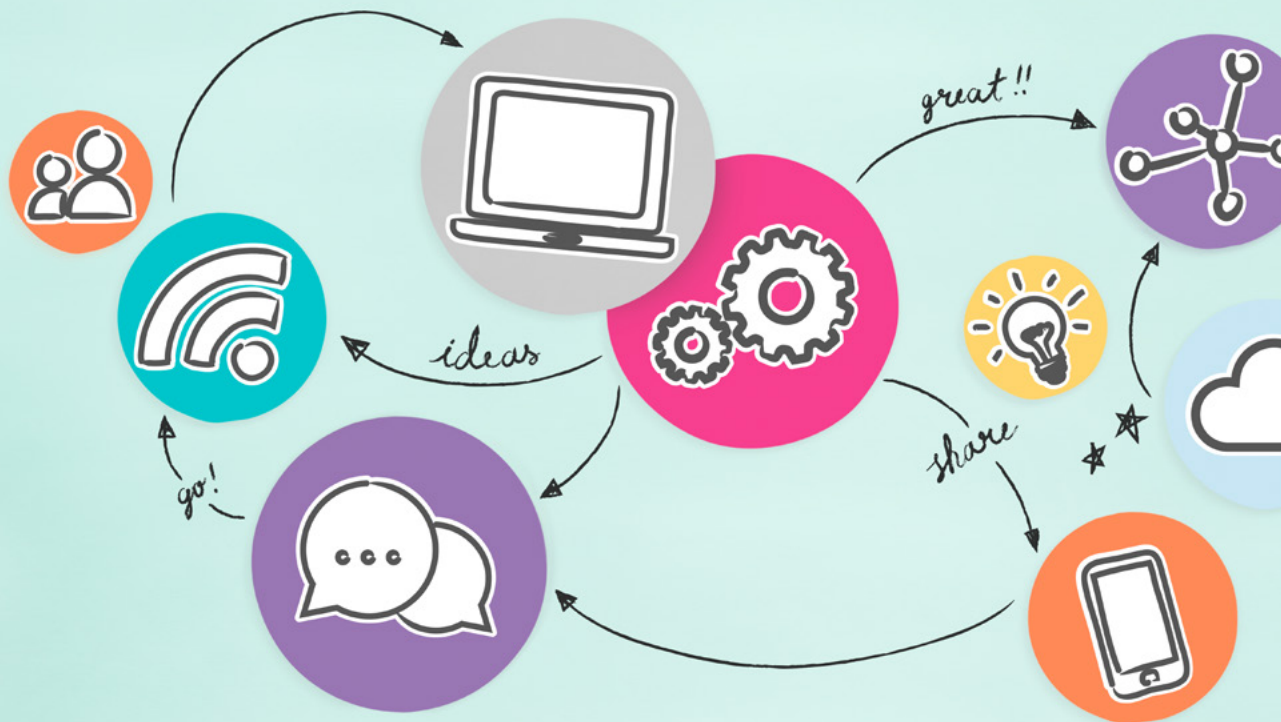
Denn unser Anspruch bleibt auch im Jahr 2026 unverändert: den Mittelstand verlässlich, praxisnah und unabhängig auf dem Weg in die digitale Zukunft zu begleiten.

Ich wünsche Ihnen eine anregende Lektüre – und lade Sie ein, mit uns auf das Erreichte zurückzublicken und zugleich den Blick bereits wieder nach vorn zu richten.

*Berend Denkena*

# Mit Plan zur Präsenz – Social Media strategisch gedacht

Projekt-Abschlussbericht von Dr. Levke Walten



## Projektüberblick

Die FORCHE commercium UG ist ein wachsendes E-Commerce-Unternehmen, das seine internen Prozesse optimieren und neue Vertriebskanäle erschließen möchte. Aufgrund der Zielgruppe und bestehender Erfahrungen rückte Instagram als potenzieller Vertriebskanal in den Fokus, insbesondere aufgrund der integrierten Shop-Funktionen. Am Beispiel der FORCHE UG wurde eine praxisnahe Content-Strategie erarbeitet, die Themen wie visuell ansprechende Inhalte, den Einsatz von KI zur automatisierten Beitragserstellung, Social-Selling-Funktionen und messbare Erfolgskennzahlen umfasst. Das Ziel bestand darin, einen übertragbaren Strategieentwicklungsprozess zu entwickeln, welcher der FORCHE UG Orientierung bietet und anderen Unternehmen als Best Practice dient.



Die zunehmende Digitalisierung eröffnet Unternehmen neue Möglichkeiten, mit Kund\*innen zu interagieren, Produkte zu vertreiben und sich am Markt zu positionieren. Im E-Commerce reicht es heute nicht mehr aus, einen funktionierenden Onlineshop zu haben – entscheidend ist eine nahtlose Verbindung aus effizienter Abwicklung, digitaler Sichtbarkeit und kanalübergreifender Kommunikation. Gerade im stark von Reichweite, Vertrauen und Kundenerlebnis geprägten Online-Handel sind digitale Strategien unerlässlich, um im Wettbewerb bestehen zu können.

Dabei kommt Social Media eine wachsende Bedeutung zu: Plattformen wie Instagram und Facebook haben sich zu Informations- und Unterhaltungsmedien sowie zu zentralen Kontaktpunkten zwischen Unternehmen und Konsument\*innen entwickelt. Sie bieten Reichweite, Interaktionsmöglichkeiten und die Chance, die eigene Marke erlebbar zu machen. All das wird durch integrierte Verkaufsfunktionen ergänzt. Diese Funktionen ermöglichen es, Produkte direkt dort zu verkaufen, wo sich die Zielgruppe aufhält.

Gerade kleine und mittlere Unternehmen (KMU) stehen dabei vor besonderen Herausforderungen. Begrenzte Ressourcen, fehlendes Fachwissen und die Vielfalt an Plattformen erschweren einen strategischen Einstieg. Gleichzeitig bieten soziale Medien niedrigschwellige Einstiegsmöglichkeiten in datenbasiertes Marketing, Kundenbindung und automatisierte Vertriebsprozesse – insbesondere, wenn moderne Technologien wie KI zur Unterstützung eingesetzt werden. Die Möglichkeit, Inhalte automatisiert zu erstellen, zielgerichtet auszuspielen und deren Wirksamkeit messbar zu machen, eröffnet auch kleinen Teams professionelle Optionen im digitalen Raum.

## Herausforderung und Zielsetzung

Hier setzt das Projekt mit der FORCHE commercium UG an. Als wachsendes Unternehmen im Bereich Baby- und Kleinkindprodukte stand FORCHE vor der Herausforderung, Social Media nicht nur als Kommunikationsmittel, sondern als strategisches Instrument für das Unternehmenswachstum zu nutzen. Am Beispiel des Unternehmens wurde ein praxisnaher Entwicklungsprozess für eine Social-Media-Strategie erprobt, der als modellhafter Ansatz für andere Unternehmen dienen kann – mit dem besonderen Fokus auf den gezielten Einsatz von Instagram.

Das Ziel bestand darin, eine übertragbare Vorgehensweise zu entwickeln, die zeigt, wie eine Strategieentwicklung auch mit begrenzten Mitteln gelingen kann. Dabei wurden Methoden wie die SMART-Ziele, SWOT-Analyse, Buyer Personas und KI-gestützte Tools eingesetzt. Damit liefert das Projekt praxisnahe Impulse für Unternehmen, die Social Media systematisch und wirkungsvoll in ihre Unternehmensentwicklung integrieren möchten.

## Lösungsweg

Der Strategieentwicklungsprozess im Projekt mit der FORCHE commercium UG folgte einem strukturierten und methodisch fundierten Vorgehen, das auf die spezifischen Anforderungen kleiner Unternehmen im digitalen Handel zugeschnitten war (Abbildung 1). Das Ziel bestand darin, eine Social-Media-Strategie zu entwickeln, die nicht nur den aktuellen Stand des Unternehmens berücksichtigt, sondern auch zukünftiges Wachstum sowie neue Anforderungen im kanalübergreifenden Vertrieb berücksichtigt. Dabei wurde bewusst auf praxisnahe Methoden gesetzt, die sich auch mit begrenzten personellen und finanziellen Ressourcen umsetzen lassen.



Abbildung 1: Strategieentwicklungsprozess im Projekt mit der FORCHE commercium UG

Zu Beginn stand die gemeinsame Erarbeitung eines Verständnisses für die Rolle von Social Media im unternehmerischen Kontext. In einem ersten Workshop wurden zentrale Unternehmensziele gesammelt und hinsichtlich ihrer Relevanz für Social Media geprüft. Auf dieser Grundlage erfolgte eine präzise Zieldefinition auf Basis der SMART-Methode. Die Formulierung konkreter, messbarer und realistischer Ziele, etwa zur Steigerung der Markenbekanntheit, zur Reichweitenentwicklung oder zur Nutzung von Shop-Funktionen, bildete das Fundament für die spätere Erfolgsmessung.

Ein weiterer methodischer Schwerpunkt lag auf der Zielgruppenanalyse. Mithilfe von Buyer Personas wurden typische Kundenprofile entwickelt, die zentrale Bedürfnisse, Interessen und digitale Nutzungsgewohnheiten abbilden. Für FORCHE bedeutete das vor allem, die Perspektive junger Eltern besser zu verstehen –

beispielsweise ihre Erwartungen an Produktsicherheit, Alltagsentlastung und ihr Informationsverhalten in sozialen Medien. Die Personas dienten als Grundlage für die Auswahl geeigneter Inhalte, Formate und Kommunikationsstile und halfen dabei, den späteren Content konsequent an den Erwartungen der Zielgruppe auszurichten.

Zur Bewertung interner und externer Einflussfaktoren wurde ergänzend eine SWOT-Analyse durchgeführt (Abbildung 2). Diese ermöglichte eine strukturierte Betrachtung der Stärken und Schwächen des Unternehmens sowie der Chancen und Risiken im Social-Media-Kontext. Zusätzlich wurde eine Wettbewerbsanalyse vorgenommen, um Einblicke in bestehende Marktmechanismen, relevante Content-Strategien und Plattformnutzung vergleichbarer Anbieter zu gewinnen. Dabei wurden ausgewählte Wettbewerber im Hinblick auf Inhalte, Posting-Frequenz, visuelle Gestaltung und Reichweitenkennzahlen auf Instagram untersucht. Ziel war es, Best Practices zu identifizieren und daraus Rückschlüsse für die eigene Positionierung und Content-Ausrichtung abzuleiten.

Die Analyseergebnisse flossen direkt in die strategische Positionierung der Marke auf Instagram ein. Für FORCHE zeigte sich dabei unter anderem das Potenzial, durch authentische, zielgruppennahe Kommunikation die Markenbindung zu stärken und sich als vertrauenswürdiger Anbieter im sensiblen Produktbereich rund um Baby- und Kleinkindbedarf zu positionieren.

Ein zentrales Ergebnis des Projekts war die Entwicklung einer Content-Strategie, die auf klar definierten Content-Säulen basiert. Diese Säulen spiegeln verschiedene Inhalte und Kommunikationsziele wider – von informativen Beiträgen über inspirierende oder unterhaltende Formate bis hin zu verkaufsorientierten Inhalten. Besonderes Augenmerk galt dabei der visuellen Qualität und Authentizität der Beiträge, da diese für die Ansprache der Zielgruppe auf visuellen Plattformen wie Instagram von entscheidender Bedeutung sind.

Für FORCHE bedeutete das vor allem, die Produktkommunikation stärker auf den Alltag junger Eltern auszurichten und über ansprechende Bilder sowie klare Botschaften Vertrauen aufzubauen. Aufbauend auf dieser Struktur wurde ein Redaktionsplan erstellt, der Themen, Anlässe, Posting-Frequenzen und saisonale Schwerpunkte systematisch abbildet. Dies ist eine wichtige Grundlage, um regelmäßig Inhalte zu veröffentlichen und gleichzeitig eine konsistente Markenwirkung zu erzielen.

Ein innovativer Aspekt des Vorgehens war der gezielte Einsatz von Künstlicher Intelligenz (KI) zur Unterstützung der Content-Erstellung. Es wurden verschiedene KI-gestützte Tools getestet, beispielsweise für das automatische Generieren von Bildunterschriften und Hashtags oder für die Gestaltung von Grafiken. Diese Technologien helfen insbesondere kleinen Unternehmen dabei, trotz begrenzter Ressourcen regelmäßig qualitativ hochwertige Inhalte zu veröffentlichen.



Abbildung 2: SWOT-Analyse

Begleitend zur Strategieentwicklung wurden Kennzahlen zur Erfolgsmessung definiert. Neben klassischen Metriken wie Reichweite, Interaktionsrate oder Follower-Wachstum wurden auch konversionsorientierte KPIs wie Klickzahlen auf Shop-Links oder Umsätze über Instagram in den Blick genommen. Mithilfe dieser messbaren Ziele kann die FORCHE UG die Wirksamkeit ihrer Maßnahmen kontinuierlich überprüfen und bei Bedarf gezielt anpassen.

## Nutzen für den Mittelstand

Das Projekt mit der FORCHE commercium UG zeigt exemplarisch, wie kleine und mittlere Unternehmen Social Media strategisch und zugleich ressourcenschonend in ihre Geschäftsprozesse integrieren können. Es verdeutlicht, dass auch Betriebe ohne eigene Marketingabteilung fundierte Strategien entwickeln können, wenn sie strukturiert vorgehen und auf praxistaugliche Methoden zurückgreifen.

Hervorzuheben ist insbesondere die Übertragbarkeit des methodischen Ansatzes: Die eingesetzten Instrumente wie SMART-Zieldefinition, SWOT-Analyse und Buyer Personas sind branchenunabhängig einsetzbar und lassen sich flexibel an unterschiedliche Unternehmensgrößen und Zielgruppen anpassen. Dadurch können KMU ihre Social-Media-Aktivitäten strategisch ausrichten, unnötige Streuverluste vermeiden und vorhandene Ressourcen zielgerichtet einsetzen.

Ein weiterer Nutzen ergibt sich aus dem Zusammenspiel von Strategieentwicklung und Digitalisierung: Der Einsatz KI-gestützter Tools senkt die Einstiegshürden und ermöglicht es auch kleinen Teams ein professionelles Auftreten. Gleichzeitig steigert die Integration von Social-Selling-Funktionen über Plattfor-

men wie Instagram die Relevanz sozialer Medien als echte Vertriebskanäle – nicht nur zur Markenbildung, sondern auch direkt umsatzwirksam.

Das Projekt liefert somit nicht nur eine individuelle Lösung für die FORCHE UG, sondern auch eine Blaupause für andere Unternehmen, die Social Media systematisch, wirkungsvoll und zukunftsorientiert in ihre Geschäftsstrategie integrieren möchten.

### Autorin

**Dr. Levke Walten**

Wissenschaftliche Mitarbeiterin am Institut für Marketing und Management (M2) der Leibniz Universität Hannover und KI-Trainerin und Expertin für Marketing und Organisationsentwicklung am Mittelstand-Digital Zentrum Hannover

# Die größte Herausforderung ist der Spagat zwischen **Effizienz** und **Authentizität**

Daniel Forche gründete 2021 die FORCHE commercium UG mit dem Ziel, alltagsnahe Lösungen für Familien zu entwickeln. Das Unternehmen mit dem Motto „Von Eltern für Eltern“ konzentriert sich auf die Gestaltung von Produkten für Babys und Kleinkinder, bei denen Funktionalität, Sicherheit und schlichtes Design im Vordergrund stehen.



**Herr Forche, mit der Weiterentwicklung des Unternehmens rückte bei Ihnen auch die Frage in den Fokus, wie Social Media gezielt eingesetzt werden kann, um neue Kund\*innen zu erreichen und gleichzeitig interne Abläufe effizienter zu gestalten – ein Thema, das im gemeinsamen Projekt mit dem Mittelstand-Digital Zentrum Hannover eine zentrale Rolle spielte. Was war für Sie der ausschlaggebende Grund, sich mit Social Media und KI-Unterstützung zu beschäftigen?**

Als junges Unternehmen mit einer klaren Zielgruppe – Eltern mit Babys – war von Anfang an klar: Ohne starke Präsenz in sozialen Medien verliert man heute den direkten Draht zu potenziellen Kund\*innen. Gleichzeitig ist der Aufwand, regelmäßig hochwertigen Content zu erstellen, enorm. KI bietet hier einen entscheidenden Vorteil, denn sie hilft, Zeit zu sparen, kreative Prozesse zu beschleunigen und Inhalte datenbasiert zu optimieren. Genau das wollte ich für unsere Marke nutzen. Das Projekt bot die ideale Gelegenheit, diese Ansätze strukturiert zu entwickeln und mithilfe der Expertise des Zentrums in die Praxis umzusetzen.

**Wie sah Ihre Social-Media-Strategie vor dem Projekt aus – und was hat sich im Laufe der Zusammenarbeit verändert?**

Vor dem Projekt war die Strategie eher reaktiv: Es gab einzelne Posts, oft ohne klaren roten Faden, wenig Planung und keine echte Wiedererkennung. Durch die KI-gestützte Planung und die systematische Entwicklung eines Redaktionsplans mit emotionalem Storytelling, klarer Bildsprache und inhaltlicher Struktur wurde aus unregelmäßigem Posting ein konsistenter Markenauftritt. Der Fokus liegt jetzt auf Vertrauen, Mehrwert und dem Aufbau einer Community – nicht auf plumper Werbung. Dabei hilft uns insbesondere der strategische Blick auf die Bedürfnisse unserer Zielgruppe, der in der Projektphase geschärft wurde.

**Welche Learnings nehmen Sie aus dem Projekt mit?**

Ein zentrales Learning ist: KI kann eine Markenidentität nicht ersetzen, aber diese verstärken, wenn sie bewusst eingesetzt wird und zur Markenstimme passt. Inhalte müssen auf echter Zielgruppenkenntnis beruhen, damit KI wirklich sinnvoll unterstützen kann. Außerdem gilt: Struktur schlägt Spontaneität. Mit einer durchdachten Planung lässt sich nicht nur viel effizienter arbeiten, sondern auch eine nachhaltigere Wirkung erzielen. Emotionalität und Relevanz sind zudem der Schlüssel zu Interaktion.

**Welche Herausforderungen sind Ihnen bei der Entwicklung und Umsetzung einer Social-Media-Strategie begegnet – und wie sind Sie damit umgegangen?**

Die größte Herausforderung ist der Spagat zwischen Effizienz und Authentizität. KI kann zwar viel, aber sie darf die Markenstimme nicht verwässern. Die Lösung sind klare Briefings, eine saubere Brand Voice und menschliche Kontrolle. Das zweite Thema ist die Themenfindung. Der Schlüssel liegt darin, nicht vom Produkt aus zu denken, sondern von den Problemen und Emotionen unserer Zielgruppe. Und drittens ist Kontinuität wichtig. Diese haben wir mithilfe klarer Workflows, eines Redaktionsplans und wiederverwendbarer Content-Formate sichergestellt.

**Welche konkreten Potenziale sehen Sie im Einsatz von KI zur Erstellung und Automatisierung von Social-Media-Inhalten?**

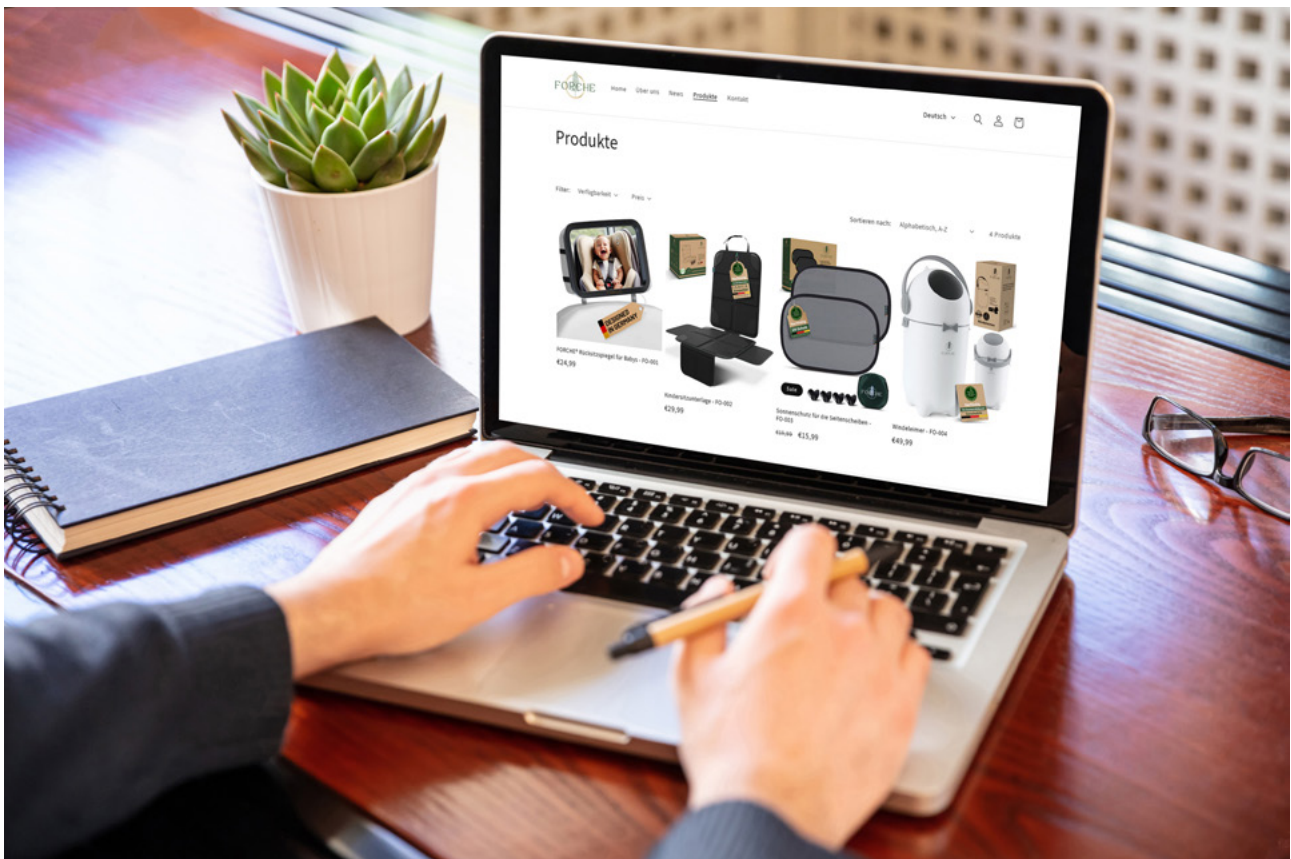
Wir sehen vor allem in drei Bereichen großes Potenzial im Einsatz von KI zur Erstellung und Automatisierung von Social-Media-Inhalten: Erstens steigert KI die Effizienz, da Inhalte schneller und mit gleichbleibender Qualität erstellt werden können. Zweitens ermöglicht KI eine hohe Skalierbarkeit: Wir können Inhalte gezielt für verschiedene Plattformen und Zielgruppen anpassen, ohne jedes Mal bei null anfangen zu müssen. Drittens liefert sie auf Knopfdruck kreative Impulse, beispielsweise in Form von Vorschlägen für Bildmaterial, alternative Textvarianten oder pas-

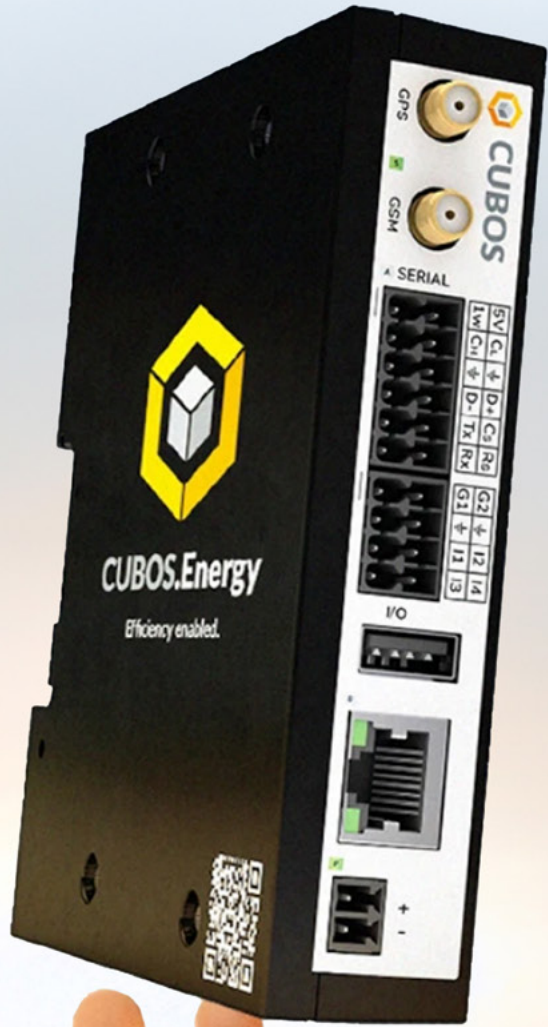
sende Hashtags, für deren Recherche man sonst viel Zeit investieren müsste. Darüber hinaus unterstützt KI uns dabei, datenbasiert zu arbeiten, indem sie aufzeigt, welche Inhalte besonders gut funktionieren und wo noch Optimierungspotenzial besteht.

**Welche nächsten Schritte planen Sie, um die im Projekt entwickelten Ansätze langfristig im Unternehmen zu verankern und weiter auszubauen?**

Wir planen, die KI-gesteuerte Content-Produktion weiter zu automatisieren. Dazu sollen feste Workflows, Templates und regelmäßige Performance-Auswertungen eingeführt werden. Parallel dazu soll das Wissen intern dokumentiert und an neue Teammitglieder weitergegeben werden. Unser Ziel ist es, nicht nur weiterhin sichtbar zu bleiben, sondern langfristig eine echte Community rund um unsere Marke aufzubauen – mit KI als zuverlässigem Sparringspartner im Hintergrund und mit einem klaren Fokus auf den echten Austausch mit Eltern.

**Vielen Dank, Herr Forche, für Ihre Zeit und die Einblicke in Ihre Arbeit.**





**CUBOS.Energy**  
Efficiency enabled.

GPS

GSM

SERIAL

I/O



# Flexibel, nachhaltig, smart: Lademanagement neu gedacht

Projekt-Abschlussbericht von Phil Köhne und Paulina Merkel

## Projektüberblick

Im Projekt mit dem Unternehmen CUBOS Service GmbH wurde untersucht, ob mittels KI-Methoden ein Lademanagementsystem aufgebaut werden kann, das datenbasierte Entscheidungen für das Laden von Elektrofahrzeugen ermöglicht. Auf Basis von rund 20.000 Ladevorgängen wurden mithilfe von Clustering-Algorithmen typische Nutzergruppen identifiziert und durch Prognosemodelle das Ladeverhalten analysiert. Ziel war es, die Standzeit und den Energiebedarf frühzeitig vorherzusagen, um Ladeprioritäten datenbasiert zu steuern. Die KI-Methode soll so die Vermeidung von Lastspitzen unterstützen, die Nutzung erneuerbarer Energien verbessern und Energiekosten senken, was einen praxisnahen Mehrwert für mittelständische Anbieter von Ladeinfrastruktur bietet.

Die Elektrifizierung des Verkehrs stellt Betreiber von Ladeinfrastruktur vor neue Herausforderungen. Gemeinsam mit der Firma CUBOS Service GmbH wurden KI-gestützte Ansätze untersucht, um Ladevorgänge effizienter zu steuern. Ziel ist es, Stromangebot, Nutzerverhalten und betriebliche Anforderungen intelligent zu verknüpfen – für ein flexibles, stabiles und kosten-effizientes Lademanagement.

## Unternehmen und Produkt

Die CUBOS Service GmbH ist ein familiengeführtes Unternehmen mit Hauptsitz in Wolfsburg, das sich seit 2018 für den Umstieg auf erneuerbare Energien und E-Mobilität einsetzt. Mit eigenentwickelter Hardware und Software bietet es seit 2021 Komplettlösungen für B2B-Kunden.

## Herausforderung und Zielsetzung

Das Ziel des KI-Anwendungsprojektes liegt darin, zu untersuchen, ob KI-basiert eine Priorisierung beim Laden von Elektrofahrzeugen möglich ist. Ausgangspunkt bildet ein umfangreicher Datensatz mit rund 20.000 anonymisierten Ladevorgängen. Auf dieser Basis soll ein Machine-Learning-Modell erstellt werden, das die Standzeit von Fahrzeugen prognostiziert. Ziel ist es, typische Ladeverhalten zu erkennen und daraus intelligente Steuerungsstrategien abzuleiten. Die Methode soll nicht nur den operativen Aufwand reduzieren, sondern auch durch präzise Vorhersagen zur Vermeidung von Lastspitzen und zur besseren Nutzung erneuerbarer Energien beitragen. Langfristig kann der Ansatz helfen, Energiekosten zu senken, den Eigenverbrauch zu erhöhen und die Netzstabilität durch gezielte Lastverschiebung zu verbessern.

## Lösungsweg

Im Folgenden wird der Lösungsweg zur Entwicklung eines KI-gestützten Lademanagementsystems beschrieben. Zuerst erfolgt eine Datenaufbereitung. Aufbauend darauf werden Methoden des überwachten und unüberwachten Lernens genutzt, um den individuellen Ladebedarf zu prognostizieren bzw. die Nutzer in Gruppen einzuteilen. Ziel ist es, aus den historischen Ladevorgängen verwertbare Vorhersagen und Strategien zur datenbasierten Priorisierung abzuleiten.

## Datenvorverarbeitung

Die Datengrundlage besteht aus einer Excel-Datei mit rund 60.000 Ladevorgängen. Erfasst sind Zeitpunkte, Energiemengen, Ladestations-IDs und Fahrzeugkennungen via RFID. Es wird vereinfachend angenommen, dass Fahrzeuge bei ausreichender Standzeit vollständig geladen werden. Zur Plausibilisierung wurde die geladene Energiemenge über Ladeleistung und Ladedauer abgeschätzt. Einschränkungen wie reduzierte Ladeleistung bei Überlast wurden nicht separat berücksichtigt. Für die Modellierung wurden irrelevante Spalten entfernt, Zeitstempel verarbeitet, Einträge vor dem 01.01.2021 ausgeschlossen und fehlende oder fehlerhafte Werte aussortiert. Es blieben nur Ladevorgänge mit plausibler Energiemenge (0–70 kWh), Standzeit (0,5–17,5 h) und Ladeleistung (0–22 kWh/h) erhalten. Das Resultat waren ca. 20.000 gültige Ladevorgänge.

## Clustering

Clustering ist ein Verfahren des unüberwachten Lernens im Bereich des maschinellen Lernens. Ziel ist es, Datenpunkte mit ähnlichen Eigenschaften in Gruppen (sogenannte Cluster) einzuteilen – ohne vorherige Labels oder Klasseninformationen. Die Idee dahinter: Ähnliche Objekte sollen im selben Cluster landen, unterschiedliche in verschiedenen. Clustering wird eingesetzt, um Strukturen, Muster oder Gruppen in großen Datensätzen zu entdecken, etwa bei Kundensegmentierung, Anomalieerkennung oder eben beim Ladeverhalten von Fahrzeugen.

Der K-Means-Algorithmus funktioniert, indem zunächst eine feste Anzahl an Clustern ( $k$ ) vorgegeben wird. Anschließend werden zufällig  $k$  Startpunkte als sogenannte Cluster-Zentren gewählt. Jeder Datenpunkt wird dem nächstgelegenen Zentrum zugewiesen. Danach werden die Zentren neu berechnet, indem der Mittelwert aller zugehörigen Punkte bestimmt wird. Dieser Vorgang der Zuweisung und Neuberechnung wiederholt sich iterativ, bis sich die Zuordnungen nicht mehr verändern oder ein Abbruchkriterium erfüllt ist. K-Means ist rechnerisch effizient und gut interpretierbar, setzt jedoch voraus, dass die Anzahl der Cluster bekannt ist und dass die Daten weitgehend kugelförmige Cluster aufweisen.

Die Bestimmung der Anzahl der Cluster erfolgte hier mit der Ellbogen-Methode. Diese Methode ist ein einfaches und visuell unterstütztes Verfahren zur Bestimmung der optimalen Anzahl von Clustern bei Clustering-Algorithmen wie K-Means. Ziel ist es, eine Balance zwischen Modellkomplexität und Erklä-

rungsqualität zu finden, also eine Anzahl an Clustern zu wählen, die möglichst viel Struktur in den Daten erfasst, ohne zu übersegmentieren. Der Ellbogen für den vorliegenden Fall ist in Abbildung 1 dargestellt. WCSS (Within-Cluster Sum of Squares) ist die Summe der quadrierten Abstände aller Punkte zu ihrem jeweiligen Cluster-Zentrum und misst die Kompaktheit der Cluster. Bei der Ellbogen-Methode wird WCSS für verschiedene Anzahlen von Clustern berechnet, um den Punkt zu finden, an dem eine weitere Erhöhung der Clusteranzahl nur noch geringe Verbesserungen bringt: den sogenannten Ellbogen. Er stellt einen sinnvollen Kompromiss dar: Die Clusterzahl ist groß genug, um die Daten sinnvoll zu strukturieren, aber klein genug, um Überanpassung zu vermeiden. Es zeigt sich in diesem Fall ein Knick bei  $k=2$ , was auf eine zweigeteilte Gruppierung hindeutet.

Die Einteilung der Nutzenden in zwei Cluster mittels K-Means führt zu den dargestellten Ergebnissen in Tabelle 1. Das erste Cluster („Kurzzeit-Nutzende“) umfasst etwa 63% der Nutzenden und zeichnet sich durch eine homogene Verteilung mit einer durchschnittlichen Standzeit von rund 4,5 Stunden aus (Maximum bei ca. 6,1 Stunden). Die geringe Standardabweichung von etwa 1,05 weist auf eine gleichmäßige Nutzung hin. Das zweite Cluster („Langzeit-Nutzende“) besteht aus etwa 37% der Nutzenden und weist eine deutlich längere durchschnittliche Standzeit von ca. 7,8 Stunden auf (Maximum bei ca. 14 Stunden). Die größere Streuung deutet auf eine heterogene Gruppe hin. Das könnte daran liegen, dass einige Nutzende ihr Fahrzeug über Nacht oder für längere Zeiträume laden. Die sich daraus ergebende Verteilung der Nutzenden ist in Abbildung 2 dargestellt.

## Prognosemodelle

Im Rahmen der Untersuchung wurde ebenfalls geprüft, inwieweit sich die Standzeit sowie die geladene Energiemenge eines Ladevorgangs bereits zu Beginn prognostizieren lässt. Hierfür wurden verschiedene Verfahren des überwachten maschinellen Lernens eingesetzt, darunter lineare Regression, Regressionsbäume und

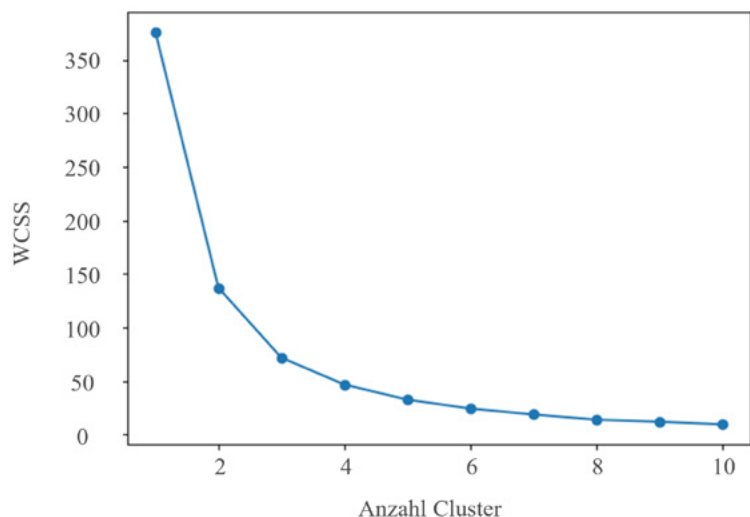


Abbildung 1: Ellbogen für die vorliegenden Daten

Random Forest. Die dafür verwendeten Merkmale wie Ankunftszeit, Wochentag, Nutzer- und Ladestations-IDs wurden numerisch kodiert, sodass die Daten in einem ML-tauglichen Format vorlagen.

Die lineare Regression diente als Referenzmodell und zeigte bereits eine solide Prognoseleistung, insbesondere im Hinblick auf die geladene Energie. Die Vorhersage der Standzeit erwies sich über alle Verfahren hinweg als deutlich schwieriger. Während der einfache Regressionsbaum ohne Optimierung nur unzureichende Ergebnisse lieferte, konnte durch Hyperparameter-Tuning (HPT) die Vorhersagequalität verbessert werden. Am besten schnitt das Random-Forest-Modell mit HPT ab. Insgesamt zeigte sich, dass der Energiebedarf besser prognostizierbar ist als die Dauer des Aufenthalts.

Da die Standzeit trotz Regression nur begrenzt vorher-sagbar war, wurde ergänzend ein Klassifikationsansatz verfolgt. Auf Basis zuvor identifizierter Cluster (kurz- und langfristige Ladevorgänge) wurde mithilfe eines RandomForestClassifier versucht, die jeweilige Kategorie vorherzusagen. Auch hier verbesserte HPT die Ergebnisse.

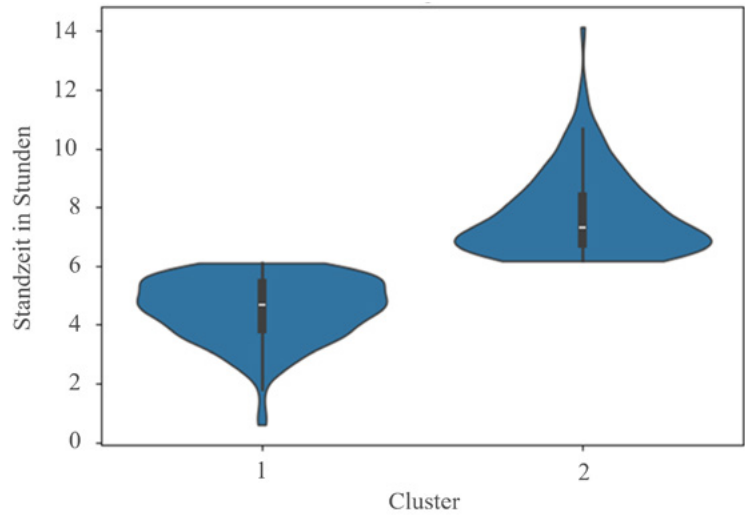


Abbildung 2: Dichteverteilung der Standzeit für Cluster 0 und Cluster 1

Cluster	Anzahl Nutzende	Durchschnittliche Standzeit (Stunden)	Standardabweichung (Stunden)	Median (Stunden)	Maximalwert (Stunden)	Cut-Off (Stunden)
1	230	~4,54	±1,05	~4,67	~6,11	6,15
2	146	~7,77	±1,38	~7,32	~14,13	

Tabelle 1:

Ergebnisse des Clusterings mit K-Means.

Insgesamt zeigen die Ergebnisse, dass sowohl Regressions- als auch Klassifikationsverfahren einen wertvollen Beitrag zur datenbasierten Prognose im Lademanagement leisten können. Die Genauigkeit ist zwar begrenzt, insbesondere bei der Vorhersage der Standzeit. Dennoch könnten diese Modelle in der Praxis als Entscheidungshilfen für die Priorisierung von Ladevorgängen dienen.

### Nutzen für den Mittelstand

Die Ergebnisse zeigen, dass KI-Methoden das Lademanagement in mittelständischen Unternehmen optimieren können und eine praxisnahe Lösung zur effizienteren Nutzung der Ladeinfrastruktur bieten. Durch die datenbasierte Prognose von Standzeiten und Ladeverhalten lassen sich Ladevorgänge besser planen, Lastspitzen vermeiden und Energiekosten senken. Gleichzeitig ermöglicht das System eine intelligentere Nutzung von Eigenstrom – etwa aus Photovoltaik – und steigert so den Eigenverbrauch. Das reduziert nicht nur Betriebskosten, sondern trägt auch zur Netzstabilität und Nachhaltigkeit bei. Für Unternehmen bedeutet das weniger manuellen Aufwand, mehr Planungssicherheit und einen aktiven Beitrag zur Energiewende.

### Autor\*innen

**Phil Köhne**  
 Projektingenieur am IPH – Institut für Integrierte Produktion Hannover gGmbH

**Paulina Merkel**  
 Projektingenieurin am IPH – Institut für Integrierte Produktion Hannover gGmbH und KI-Trainerin am Mittelstand-Digital Zentrum Hannover

# Nachhaltigkeit messbar machen: KI-gestützte Evaluierung von Patenten anhand der 17 UN-Nachhaltigkeitsziele

Projekt-Abschlussbericht von Manuel Savadogo

Nachhaltigkeit ist ein zentrales Leitmotiv moderner Innovationspolitik und ein Schlüssel zur langfristigen Wettbewerbsfähigkeit. Patente als Schutzrechte für technologische Entwicklungen nehmen dabei eine strategische Rolle ein – insbesondere dann, wenn sie nachweislich zu nachhaltigen Zielen beitragen. Der politische und gesellschaftliche Druck auf Unternehmen, nachhaltig zu wirtschaften, wächst stetig. Gleichzeitig steigt das Volumen an Patentanmeldungen kontinuierlich an. Die manuelle Bewertung dieser Anmeldungen im Hinblick auf Nachhaltigkeitsmerkmale ist jedoch ressourcenintensiv und fehleranfällig. Vor diesem Hintergrund wurde ein Projekt initiiert, das die Potenziale Künstlicher Intelligenz (KI) nutzt, um Patente systematisch und automatisiert auf ihre Übereinstimmung mit den 17 Zielen für nachhaltige Entwicklung der Vereinten Nationen (SDGs) zu prüfen. Dieses Vorhaben zielt nicht nur auf eine Effizienzsteigerung, sondern auch auf eine höhere Objektivität und bessere Nachvollziehbarkeit der Analyseergebnisse ab.

## ZIELE FÜR NACHHALTIGE ENTWICKLUNG



Abbildung 1: Die 17 Ziele für nachhaltige Entwicklung (SDGs)

## Unternehmen/Produkt

Die EZN GmbH, seit 1986 eine eigenständige GmbH, unterstützt seit Jahrzehnten Erfinder\*innen, Unternehmen und Hochschulen bei der Verwertung technischer Innovationen – insbesondere im Bereich des Schutzrechtsmanagements. Sie transferiert Ergebnisse aus Forschung und Entwicklung in die Realität und unterstützt bei allen Fragen rund um die Entwicklung neuer Produkte und Verfahren von der Idee bis zum Vermarktungserfolg. Hierbei wird explizit eine Beratung zu Patenten, Marken, Designs und anderen Schutzrechten angeboten. Darüber hinaus führt EZN Technologiebewertungen und sogenannte Freedom-to-Operate-Analysen.

## Herausforderungen und Ziel

Bislang werden Patente hinsichtlich ihrer Relevanz für Nachhaltigkeitsziele vollständig manuell geprüft. Dies ist insbesondere bei großen und komplexen Patenten sehr zeit- und ressourcenintensiv. Betrachtet man allein die 17 UN-Nachhaltigkeitsziele mit ihren 169 Unterzielen, wird der immense zeitliche Aufwand deutlich. Zudem besteht die Gefahr subjektiver Bewertungen, insbesondere wenn mehrere Mitarbeitende mit unterschiedlichem Hintergrund involviert sind. Die zunehmende Komplexität der Nachhaltigkeitsanforderungen – etwa durch europäische Regulierungen wie die Corporate Sustainability Reporting Directive (CSRD) – verschärft diese Problematik. Viele KMU verfügen

weder über das Fachpersonal noch über geeignete Tools, um Patente systematisch auf Nachhaltigkeitsmerkmale zu prüfen. Vor diesem Hintergrund wurde das Ziel formuliert, ein KI-gestütztes Tool zu entwickeln, das eine objektive, skalierbare und reproduzierbare Bewertung ermöglicht. Das prototypische Tool soll darüber hinaus nicht nur Aufschluss darüber geben, ob das Patent zu den 17 UN-Nachhaltigkeitszielen beiträgt, sondern auch inwieweit das Patent die einzelnen Nachhaltigkeitsziele unterstützt. Die Lösung sollte modular aufgebaut sein, sodass sie später problemlos auf weitere Dokumententypen oder andere regulatorische Zielsysteme angepasst werden kann.



## Lösungsweg/Vorgehensweisen

Im ersten Schritt des Projekts wurden die vorliegenden Patente hinsichtlich ihrer Relevanz und Qualität analysiert. Dies ermöglicht einen Überblick über die vorliegende Datenmenge. Zur Datenvorverarbeitung kamen u.a. Lemmatization, Stopword Removal und Named Entity Recognition zum Einsatz. Zusätzlich wurde eine umfangreiche Analyse der 17 UN-Nachhaltigkeitsziele durchgeführt. Die hierbei erzielten Ergebnisse bilden die Grundlage für das geplante KI-Tool zur Evaluierung der Patente. Anschließend wurde die Recherche nach geeigneten Ansätzen und Modellen zur technischen Umsetzung durchgeführt.

Die technische Umsetzung erfolgte final mit einem lokal ausführbaren Large Language Modell (LLM) auf Basis des Llama-Modells. Diese Modelle zeichnen sich durch ihre hohe Performance bei gleichzeitig geringem Ressourcenverbrauch aus. Das System kann anhand eines strukturierten Prompts Patente analysieren, nach SDG-relevanten Kriterien bewerten und konkrete Textstellen mit Relevanzbezug herausfiltern. Die KI wurde durch Methoden der Natural Language Processing (NLP) trainiert und mit branchenspezifischen Begriffen angereichert. Der Bewertungsalgorithmus nutzt eine Scoring-Matrix, die jedem der 17 UN-Ziele einen gewichteten Score zuordnet. Eine benutzerfreundliche Oberfläche sowie Funktionen zur Visualisierung der Ergebnisse ermöglichen einen unmittelbaren Einsatz in KMU. Die Resultate lassen sich als Excel exportieren und enthalten das Scoring der einzelnen Nachhaltigkeitsziele auf einer Skala von 1 bis 100. Neben Zeit- und Kostenersparnis, die sich aus der Anwendung des KI-Tools ergibt, wird so auch die ESG-Strategie nachhaltig gestärkt. KMU kann hierbei beispielsweise bei der Erfüllung der

EU-Taxonomie, die festlegt, welche wirtschaftlichen Aktivitäten als ökologisch gelten, geholfen werden. Darüber hinaus zeigen nachhaltige Patente, dass Unternehmen Verantwortung für gesellschaftliche Herausforderungen übernehmen.

## Nutzen für den Mittelstand

Parallel zur Entwicklung des Prototyps wurde ein Leitfaden entwickelt. Dieser bietet insbesondere kleinen und mittleren Unternehmen (KMU) eine schrittweise und leicht nachvollziehbare Anleitung zur Entwicklung eines KI-basierten Tools zur Evaluierung unternehmensinterner Dokumente. Der Leitfaden befasst sich akribisch mit der Entwicklungsumgebung und dem Programmiercode, da dies insbesondere für interessierte Unternehmen von Bedeutung ist. Darüber hinaus werden die Vorteile nachhaltiger Patente, die Integration individueller Evaluationskriterien und die Optimierung des Prompt-Designs detailliert beschrieben.

Die in Zusammenarbeit mit der EZN GmbH entwickelte prototypische Lösung ist eine optimale Referenz für viele KMU, die tagtäglich denselben Herausforderungen gegenüberstehen. Die Reduzierung des manuellen Aufwands und die Optimierung der Dokumentenanalyse durch den Einsatz von KI sind Themen, mit denen sich aktuell viele Unternehmen auseinandersetzen. In Zeiten steigender Digitalisierung ist dieses Thema aktueller denn je.

Mit Hilfe der entwickelten Lösung können KMU Ansätze der KI nutzen, um ihre unternehmensinternen Produkte zu optimieren. Der Leitfaden bietet ihnen die Möglichkeit, sich gründlich mit den vorliegenden Dokumenten und deren Analyse auseinanderzusetzen und mit dem entwickelten Modell eine aufwandsarme und reproduzierbare Evaluierung hinsichtlich initial definierter Kriterien – wie bspw. den 17 UN-Nachhaltigkeitszielen – durchzuführen. Dies führt zu einer Steigerung der Produktivität und Effizienz von KMU und stärkt das Vertrauen in die Nutzung von Ansätzen und Methoden der Künstlichen Intelligenz. Darüber hinaus werden KMU durch den Einsatz (oder Anwendung/Nutzung) des Leitfadens für das Thema Programmierung und Datenverarbeitung sensibilisiert. Der Leitfaden und der sich daraus ergebene Mehrwert zeigen Unternehmen, dass es vorteilhaft sein kann, sich intensiver mit den Potenzialen von KI hinsichtlich der Optimierung unternehmensinterner Prozesse auseinanderzusetzen.

### Autor

#### Manuel Savadogo

KI-Experte am Mittelstand-Digital-Zentrum  
Hannover und Projektingenieur am IPH –  
Institut für integrierte Produktion  
Hannover gGmbH



# Digitalisierung der Abfallogistik in temporären Ballungszentren

Projekt-Abschlussbericht von Laura Obermann und Dr.-Ing. Mark Mennenga

## Projektüberblick

Im Projekt mit der Nordseebad Carolinensiel-Harlesiel GmbH wurden digitale Ansätze zur Optimierung der Abfallogistik in einem stark touristisch geprägten Ort an der Nordsee entwickelt. Die Herausforderung bestand in der Handhabung saisonal stark schwankender Abfallaufkommen, während die bestehende Entsorgungsstruktur weder digital unterstützt ist, noch belastbare Daten über das tatsächliche Müllaufkommen liefert. Vor diesem Hintergrund erfolgte auf Basis einer Situationsanalyse die Entwicklung eines Füllstandssensors für Müllbehälter zur lokalen Speicherung von Messdaten. Darüber hinaus wurde eine App entwickelt, welche Mitarbeitende in der Nebensaison die Kommunikation und bedarfsgerechte Leerung erleichtert. Ziel war es, Leerfahrten zu vermeiden und so Ressourcen zu sparen. So konnten erste Digitalisierungsimpulse für eine datengestützte Abfallogistik gesetzt werden.

*Hinweis: Bei den hier veröffentlichten Daten handelt es sich aus Gründen der Vertraulichkeit um abgeänderte, fiktive Daten.*

Temporäre Ballungszentren wie Küstenorte mit starkem Ferien- und Sommertourismus stehen in der Abfallwirtschaft vor besonderen Herausforderungen. Während in der Hauptsaison täglich hunderte Gäste für ein stark erhöhtes Müllaufkommen sorgen, leben in der Nebensaison nur wenige Menschen dauerhaft am Ort. Entsprechend schwanken die Anforderungen an die Entsorgungslogistik erheblich. Gleichzeitig fehlen in vielen kleineren Kommunen verlässliche Daten über das tatsächliche Abfallaufkommen, insbesondere für frequentierte Orte wie Strände, Promenaden oder andere öffentliche Plätze.

Parallel dazu gewinnen digitale Lösungen in der kommunalen Infrastruktur an Bedeutung, da sie neue Möglichkeiten eröffnen, Abläufe transparenter, flexibler und besser an reale Nutzungsmuster anzupassen. Insbesondere in saisonal stark schwankenden Nutzungskontexten können datenbasierte Ansätze helfen, bestehende Routinen zu hinterfragen und operative Entscheidungen stärker am tatsächlichen Bedarf auszurichten. Die Abfallwirtschaft entwickelt sich damit zunehmend zu einem Anwendungsfeld digitaler Transformation im kommunalen und touristischen Umfeld, bei dem Effizienz, Praktikabilität und Anpassungsfähigkeit im Vordergrund stehen.

## Unternehmen

Die Nordseebad Carolinensiel-Harlesiel GmbH ist die zentrale Organisation für den Tourismus im gleichnamigen Urlaubsort an der niedersächsischen Nordseeküste und trägt maßgeblich zur Entwicklung und Vermarktung des

staatlich anerkannten Nordseeheilbades bei. Als kommunales Tourismusunternehmen sorgt sie für einen attraktiven Aufenthalt von Gästen aus ganz Deutschland und dem Ausland, indem sie das vielfältige Freizeit-, Kultur- und Erholungsangebot koordiniert und weiterentwickelt. Der Betrieb umfasst neben klassischen Aufgaben wie Gästeservice, Information und Marketing auch die Betreuung eigener Einrichtungen wie ein Schwimmbad, einen Campingplatz und einen Wohnmobilstellplatz direkt am Meer. Zudem ist das Unternehmen verantwortlich für die Pflege und Entleerung der Abfallbehälter an öffentlichen Plätzen wie Promenade, Strand und Hafen und stellt damit sicher, dass zentrale touristische Orte sauber und einladend bleiben.

Mit rund 40 Mitarbeitenden arbeitet die GmbH eng mit Einwohnern, Ferienhaus-Inhabern, Gastronomie- und Hotellerie-Betrieben sowie Vereinen zusammen und fungiert als Schnittstelle zwischen unterschiedlichen lokalen Akteuren. Die Organisation pflegt einen intensiven Dialog mit Bewohnern und touristischen Leistungsträgern, um gemeinsam die Zukunft des Ortes nachhaltig zu gestalten und die hohe Lebens- und Aufenthaltsqualität zu sichern. Carolinensiel-Harlesiel zeichnet sich durch eine einzigartige Mischung aus historischer Hafenumgebung, maritimer Kultur, Familien- und Gesundheitstourismus sowie Naturerlebnis im UNESCO-Weltnaturerbe Wattenmeer aus, die vom Unternehmen aktiv unterstützt und weiterentwickelt wird.

## Problemstellung/Zielsetzung

Im Zuständigkeitsbereich der Nordseebad Carolinensiel-Harlesiel GmbH fehlte vor Projektbeginn eine bedarfsorientierte Planung. Leerungstouren wurden überwiegend erfahrungsbasiert geplant, belastbare Daten zum tatsächlichen Abfallaufkommen an zentralen Sammelpunkten lagen nicht vor. In der touristisch stark geprägten Kommune mit rund 1.600 Einwohnern außerhalb der Saison, jedoch bis zu 187.000 Gästen und über eine Million Übernachtungen jährlich, führt diese Situation zu erheblichen Herausforderungen. Saisonale Überlastungen in der Hauptsaison sowie ineffiziente Routinen in der Nebensaison verursachen unnötige Leerfahrten, zusätzliche Fahrzeugbelastung und einen erhöhten Einsatz personeller Ressourcen.

Ziel des gemeinsam zwischen dem Mittelstand-Digital Zentrum Bremen-Oldenburg und dem Mittelstand-Digital Zentrum Hannover durchgeführten Projekts war es, unnötige Leerfahrten in der Abfalllogistik zu vermeiden und dadurch personelle, logistische und ökologische Ressourcen zu schonen. Hierfür sollten digitale Technologien eingesetzt werden, um Transparenz über Nutzung und Füllstände der Abfallbehälter zu schaffen und Leerungen bedarfsgerecht steuern zu können. Durch die Reduktion überflüssiger Fahrten und eine verbesserte Abstimmung der Entsorgung auf das tatsächliche Abfallaufkommen sollten insbesondere Kraftstoffverbrauch, Emissionen und Arbeitsaufwand gesenkt werden. Auf diese Weise leistet das Projekt einen Beitrag zu einer effizienten und umweltverträglicheren Abfallentsorgung in touristisch geprägten, temporären Ballungszentren.

## Lösungsweg

Um den beschriebenen saisonalen Schwankungen und fehlenden Transparenzen in der Abfalllogistik wirksam zu begegnen, ist ein besseres Verständnis des tatsächlichen Nutzungs- und Füllverhaltens der Abfallbehälter erforderlich. Nur wenn belastbare Informationen über Ort, Zeitpunkt und Umfang des

Abfallaufkommens vorliegen, können Leerungen vom bisherigen Routineprinzip gelöst und bedarfsgerecht geplant werden. Digitale Erfassungs- und Kommunikationslösungen bieten hierfür einen geeigneten Ansatz, da sie die notwendige Datengrundlage schaffen, um operative Entscheidungen flexibel an wechselnde Rahmenbedingungen in temporären Ballungszentren anzupassen. Zur Entwicklung einer bedarfsorientierten und digital unterstützten Abfalllogistik wurde das Projekt in mehreren Schritten umgesetzt:

### 1. Analyse der Ausgangssituation:

Der Digitalisierungsgrad wurde zunächst erfasst, um mögliche Maßnahmen auf Grundlage bestehender Daten abzuleiten. Dabei zeigte sich, dass im Bereich der Abfallentsorgung keinerlei digitale Unterstützung vorhanden ist und die wenigen verfügbaren Daten lediglich in unregelmäßigen Abständen und nur für einen zentralen Sammelcontainer vorliegen. Informationen zum täglichen oder wöchentlichen Aufkommen sowie zu einzelnen Mülleimern im Ort fehlen vollständig. Weiterhin wurde im Rahmen einer Begleitung realer Entsorgungstouren ein umfassendes Verständnis der bestehenden Abläufe erarbeitet. Dabei zeigte sich, dass die Leerung von Müllbehältern routinemäßig erfolgt, unabhängig davon, ob eine Tonne tatsächlich voll oder leer ist. Dies gilt besonders in der Nebensaison, wenn das Müllaufkommen stark zurückgeht, aber der Takt der Touren weitestgehend gleich bleibt.

### 2. Entwicklung eines Sensors zur Füllstandsmessung:

Basierend auf den Erkenntnissen wurde ein Prototyp zur automatisierten Füllstandserfassung entwickelt. Der Sensor ist in der Lage, den Füllstand eines Müllbehälters zu erfassen und lokal zu speichern. Aufgrund der oft schwierigen Netzabdeckung im Außenbereich wurde auf eine stromsparende Lösung ohne permanente Onlineverbindung gesetzt. Abbildung 1 zeigt exemplarische Messdaten des erfassten Füllstands ausgewählter Mülleimer im Projektgebiet. Die kontinuierlichen Messungen ermöglichen es, typische Entsorgungsmuster über Tages-, Wochen- und Monatsverläufe hinweg zu erkennen. Der Mehrwert der datengestützten Erfassung liegt weniger in der unmittelbaren Auslösung einzelner Leerungen, sondern in der Analyse des tatsächlichen Entsorgungsverhaltens über längere Zeiträume. Die gewonnenen Erkenntnisse ermöglichen es, Leerungsintervalle und Tourenplanung für zukünftige Haupt- und Nebensaisons zum Beispiel durch Einrichtung von Datenschnittstellen zur Erfassung des Besucheraufkommens bedarfsgerechter anzupassen und wiederkehrende Leerfahrten ohne ausreichenden Füllgrad gezielt zu vermeiden.

Datum	Uhrzeit	Füllstand [%]	Abstand zum Müll [cm]
01.08.2025	09:00	21,30	80,00
01.08.2025	16:00	52,70	50,50
01.08.2025	21:00	63,80	40,00
02.08.2025	09:00	68,60	35,50
02.08.2025	16:00	66,40	34,60
02.08.2025	21:00	84,20	20,90
03.08.2025	09:00	91,80	13,70
03.08.2025	16:00	13,90	87,00
03.08.2025	21:00	42,50	60,45
04.08.2025	09:00	52,10	50,80
04.08.2025	16:00	73,80	30,45
04.08.2025	21:00	84,20	20,90

Abbildung 1: Beispielhafte Messdaten des Mülleimerfüllstands

### 3. Entwicklung einer App für die Mitarbeitenden:

Ergänzend zur sensorbasierten Füllstandserfassung wurde eine zweite, davon unabhängige Lösung entwickelt, um auch solche Standorte in die bedarfsgerechte Abfalllogistik einzubeziehen, an denen der Einsatz von Sensorik nicht praktikabel ist. Während Sensoren insbesondere an fest installierten Abfallbehältern zum Einsatz kommen, die nicht direkt mit dem Sammelfahrzeug erreichbar sind, existieren andere Entsorgungsorte, an denen sich Mitarbeitende regelmäßig aufhalten. Dort verhindern Witterungseinflüsse und offene Bauformen der Mülleimer einen zuverlässigen Sensoreinsatz. Vor diesem Hintergrund stellte die Entwicklung einer mobilen App für die Mitarbeitenden das sinnvollste Vorgehen dar, um auch an diesen Standorten Informationen zum Füllzustand systematisch zu erfassen und in die operative Planung einfließen zu lassen. An zahlreichen Standorten sind bereits vor Beginn der regulären Entsorgungstour Mitarbeitende anwesend, welche dort aufgrund anderer Aufgaben tätig sind. Diese Präsenz kann genutzt werden, um den Füllstand der Mülleimer zeitnah und ohne zusätzlichen Personalaufwand zu erfassen: Da in der Nebensaison die Tonnen nur vereinzelt gefüllt sind, wurde eine App entwickelt, mit der Mitarbeitende untereinander kommunizieren und Füllstände manuell melden können. Dies ermöglicht es, Tonnen mit geringem oder keinem Füllstand gezielt auszulassen und nur dort zu leeren, wo es notwendig ist. Die App unterstützt die flexible Tourenplanung in schwankenden Nutzungssituationen. So kann die Entsorgungstour abgekürzt werden und die Mitarbeitenden andere Aufgaben der Grünflächenpflege im Ort erledigen. Abbildung 2 zeigt einen Screenshot der eigens entwickelten App, über die die Mitarbeitenden Meldungen zum Füllstand übermitteln können. Die Standorte sind in der Reihenfolge der Mülltour aufgelistet, sodass ein effizienter Ablauf unterstützt wird. Neben der Füllstandsklassifizierung können Bilder hochgeladen werden, um Schäden oder Auffälligkeiten zu dokumentieren. Eine Kommentarspalte ermöglicht zusätzliche Hinweise, etwa zu Reparaturen oder Austauschbedarf. Ein Zeitstempel zeigt an, wie aktuell die Meldung ist. Das Ampelsystem erleichtert die Priorisierung: Rot markiert Tonnen, die am selben Tag geleert werden sollten, Grün steht für unkritische Füllstände, und Orange signalisiert eine situationsabhängige Entscheidung – beispielsweise abhängig von der geplanten nächsten Tour oder der Nähe zu bereits roten Standorten.

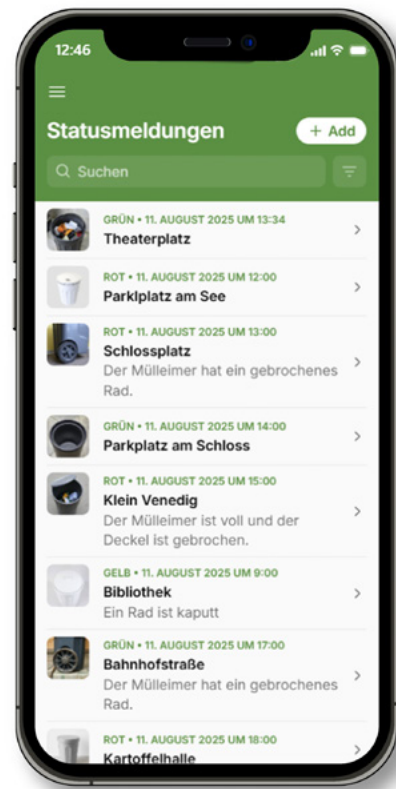


Abbildung 2: Screenshot aus der App zur Füllstandsmeldung.

### 4. Prototypische Erprobung:

Die entwickelten Lösungen wurden in einer Testphase erprobt. Die entwickelte App wurde in Abstimmung mit den Mitarbeitenden, die sie in der Praxis einsetzen, konzipiert. Nach einer ersten Entwicklungsphase erfolgte eine Vorstellung und anschließende Anpassung der Anwendung auf Grundlage des Feedbacks der Nutzenden. Die Einführung und aktive Nutzung der App ist für die Nebensaison vorgesehen. Bei der Entwicklung des Füllstandssensors lag der Fokus auf der Kombination aus technischer Funktionalität (z. B. Messgenauigkeit, Robustheit) und praktischer Anwendbarkeit. Die Erkenntnisse des Testlaufs flossen iterativ in die Weiterentwicklung ein. Abbildung 3 zeigt den entwickelten Sensoraufbau zur automatischen Füllstandsmessung in einer Mülltonne. Das Gehäuse wurde mittels 3D-Druck hergestellt und an der Unterseite des Zigarettenauffangbehälters befestigt, sodass der Sensor direkt über der Öffnung positioniert und zugleich wettergeschützt ist. Die Box muss lediglich zum Tausch der Batterien abgenommen werden. Ein Ein- und Ausschalten des Sensors ist über seitliche Schlitze möglich, ohne den Sensor zu demontieren. Im Inneren befinden sich neben der Batterie ein Ultraschallsensor, eine Speicherkarte sowie ein Microcontroller. Der Sensor ist im oberen Bereich der Mülltonne angebracht und sendet in regelmäßigen Abständen Ultraschallimpulse in Richtung des Abfalls aus. Die vom Müll reflektierten Schallwellen werden vom Sensor wieder empfangen; aus der gemessenen Laufzeit wird der Abstand zwischen Sensor und Abfalloberfläche berechnet. Auf Basis dieses Abstands lässt sich der aktuelle Füllstand der Tonne in Prozent bestimmen. Diese Messung erfolgt dreimal täglich. Die erfassten Werte – bestehend

aus dem gemessenen Abstand und dem berechneten Füllstand – werden lokal auf einer Speicherkarte abgelegt. Die Daten werden in definierten Zeiträumen, etwa nach Ferienzeiten, nach der Hauptsaison oder zwischen Saisonabschnitten, ausgelesen und ausgewertet. Die Programmierung sieht vor, dreimal täglich eine Messung durchzuführen und den Füllstand lokal zu speichern. Zwischen den Messungen versetzt sich der Controller in den Sleep-Modus, um eine Batterielaufzeit über mehrere Monate zu gewährleisten. Der entwickelte Sensor wurde gezielt an Abfallbehältern eingesetzt, die aufgrund ihrer Lage nicht direkt mit dem Müllfahrzeug erreichbar sind. Da diese Standorte von den Mitarbeitenden nicht im Rahmen der regulären Leerungstouren angefahren werden können, dient der Sensor dazu, den Füllstand dieser Mülleimer regelmäßig zu erfassen und verlässlich zu dokumentieren. Auf diese Weise wird auch für schwer zugängliche Sammelpunkte eine kontinuierliche Einschätzung des Entsorgungsbedarfs ermöglicht.



Abbildung 3: Sensor zur Füllstandsmessung in einer Mülltonne

## 5. Erkenntnisse & Ausblick:

Das Projekt hat gezeigt, dass bereits mit vergleichsweise einfachen digitalen Lösungen eine erhebliche Verbesserung der Abfalllogistik, insbesondere in Kontexten mit stark schwankender Nutzung, möglich ist. Die Kombination aus erfassten Sensordaten und strukturierter interner Kommunikation ermöglicht eine flexible, bedarfsorientierte und ressourcenschonende Organisation von Entleerungsprozessen.

Ein praktisches Beispiel ist die bedarfsgesteuerte Entleerung: Die Mülleimer übermitteln in festgelegten Intervallen ihre Füllstände, wodurch Entleerungen nur bei tatsächlichem Bedarf erfolgen. Dies führt zu einer Reduktion unnötiger Fahrten, ermöglicht eine effizientere Personalplanung und einer Vermeidung von Überfüllungen.

Der Einsatz weiterer Sensoren in Behältern verschiedener Abfallkategorien kann darüber hinaus wertvolle Erkenntnisse über das Trennverhalten der Nutzenden und den relativen Anteil einzelner Abfallarten liefern. Auf dieser Grundlage lassen sich Maßnahmen zur Verbesserung der Trennqualität ableiten, etwa durch eine gezielte Anpassung von Behälterstandorten, Volumina oder Kennzeichnungen.

Perspektivisch lassen sich die entwickelten Ansätze auf weitere Standorte mit vergleichbarer Infrastruktur übertragen, wie etwa Veranstaltungsflächen, Rastplätze oder temporäre Siedlungen. Eine vertiefte Digitalisierung fungiert dabei als Enabler für eine zunehmende Automatisierung: Durch den Wegfall manueller Dateneingaben und die softwaregestützte Analyse von Stoffströmen wird die Fehleranfälligkeit reduziert und die Reaktionsgeschwindigkeit der Logistik erhöht.

Das Fundament hierfür bilden Internet-of-Things-(IoT)-Lösungen, also vernetzte Sensoren und Geräte, die relevante Kennzahlen wie Füllstände oder Behälterpositionen autonom erfassen und in Echtzeit bereitstellen. Die Vernetzung dieser Daten in einem zentralen System ist die Voraussetzung für eine dynamische Tourenplanung: Anstatt starre Abholintervalle abzufahren, werden Routen bedarfsgerecht nur dann generiert, wenn Sensoren einen vordefinierten Füllgrad melden. Dies führt zu einer messbaren Ersparnis von Kraftstoff und Emissionen, da unnötige Leerfahrten vermieden werden. Darüber hinaus ermöglicht die Echtzeitverfügbarkeit der Daten eine vorausschauende Ressourcenplanung (Predictive Analytics), bei der durch die Analyse von Trends – etwa ein erhöhtes Aufkommen zu bestimmten Stoßzeiten – Personal und Logistikkapazitäten proaktiv statt reaktiv gesteuert werden können. Schließlich gewährleistet die automatisierte Rückmeldung der Sensoren eine lückenlose digitale Dokumentation und Transparenz über die gesamten Stoffströme, was die Grundlage für ein effizientes und revisionsssicheres Logistikmanagement bildet.

## Nutzen für den Mittelstand

Das Projekt zeigt, wie mittelständische und kommunale Entsorger durch einfache digitale Werkzeuge, wie Sensorik und App-basierter Kommunikation ihre Abläufe effizienter gestalten können. Besonders in saisonal geprägten Regionen lassen sich Leerfahrten reduzieren, Ressourcen besser planen und Prozesse flexibler gestalten. Dies ermöglicht nicht nur Kosteneinsparungen, sondern trägt auch zur Umweltentlastung und langfristigen Wettbewerbsfähigkeit bei.

### Autor\*innen

#### Laura Obermann

Wissenschaftliche Mitarbeiterin am Institut für Werkzeugmaschinen und Fertigungstechnik (IWF) der Technischen Universität Braunschweig und Expertin für Nachhaltige Produktion im Mittelstand-Digital Zentrum Hannover

#### Dr.-Ing. Mark Mennenga

Stellvertretende Institutsleitung und Abteilungsleiter am Institut für Werkzeugmaschinen und Fertigungstechnik (IWF) der Technischen Universität Braunschweig und Experte für datenbasierte Geschäftsmodelle in der Kreislaufwirtschaft im Mittelstand-Digital Zentrum Hannover



# Wissensmanagement: **KI-basierter Chatbot** liefert unternehmensinterne Informationen

Projekt-Abschlussbericht von Paolo Pappé, Philipp Harder und Robin Stöber



## Projektüberblick

Im Rahmen eines Digitalisierungsprojekts des Mittelstand-Digital Zentrums Hannover wurde gemeinsam mit der ES Edgar Schröder Unternehmensberatungsgesellschaft für Zeitarbeit mbH ein leistungsfähiger, lokal betriebener RAG-Chatbot (Retrieval-Augmented Generation) entwickelt und erfolgreich umgesetzt. Ziel war eine prototypische Lösung, die auf Basis lokaler Dokumente (PDF-Dateien mit Text- und Bildinhalten wie PowerPoint-Folien) qualitative, kontextbezogene Antworten auf individuelle Fragen liefert. Der Prototyp demonstriert, wie Unternehmen internes Wissen sicher, schnell und ohne Cloud-Abhängigkeit nutzbar machen und so Rechercheaufwände deutlich reduzieren können.

Die zunehmende Verfügbarkeit leistungsfähiger KI-Sprachmodelle eröffnet Unternehmen neue Wege im Wissensmanagement. Besonders dort, wo Informationen in heterogenen Dokumentenformaten vorliegen und sich Mitarbeitende häufig durch umfangreiche Unterlagen arbeiten müssen, bietet ein RAG-basierter Ansatz erhebliche Effizienzpotenziale. Vor diesem Hintergrund wurde gemeinsam ein prototypischer Chatbot entwickelt, der Dokumente lokal verarbeitet und Antworten transparent mit Quellenbezug bereitstellt.

## Unternehmen

Die ES Edgar Schröder Unternehmensberatungsgesellschaft für Zeitarbeit mbH mit Sitz in Varrel (Niedersachsen) berät seit ihrer Gründung im Jahr 1993 Unternehmen der Personaldienstleistungsbranche.

Das Team besteht aktuell aus 16 Fachberatern, die mehr als 300 Zeitarbeitsunternehmen betreuen. Das Leistungsspektrum umfasst unter anderem strategische Beratung, Aktenprüfung und Revision, Controlling sowie Unternehmensentwicklung. Darüber hinaus betreibt die Unternehmensberatung die Akademie der Zeitarbeit, die seit über 30 Jahren praxisnahe Online- und Präsenzseminare zu Themen wie dem Arbeitnehmerüberlassungsgesetz (AÜG), Tarif- und Arbeitsrecht sowie Vertragsmanagement anbietet.

## Problemstellung und Zielsetzung

Ein zentraler Bestandteil der Arbeit ist der Umgang mit komplexen, rechtlich geprägten Dokumenten (z. B. Gesetzestexte, Tarifverträge, Verträge, Richtlinien). Für Mitarbeitende entsteht dabei regelmäßig hoher Rechercheaufwand – insbesondere, wenn es um die schnelle Beantwortung spezifischer Fachfragen geht.

Der Ausgangspunkt dieses Projektes war daher der Bedarf, den Zugang zu vorhandenem Wissen deutlich zu beschleunigen: Relevante Passagen aus verschiedenen Dokumenten sollten ohne langwierige Recherche auffindbar sein, und auf Basis dieser Textpassagen sollte eine gestellte Frage beantwortet werden. Gleichzeitig musste der gesamte Prozess datensouverän ablaufen, also ohne Auslagerung sensibler Inhalte in externe Cloud-Dienste.

Im Projekt galt es zudem, die technische Machbarkeit eines ressourcenschonenden, lokal betriebenen Systems zu prüfen – einschließlich der Auswahl kompakter Open-Source-Sprach- und Embedding-Modelle sowie einer effizienten Hardware-Nutzung. Ein weiteres Ziel war es, die Antworten des Chatbots transparent und nachvollziehbar zu gestalten: Die für die generierte Antwort herangezogenen Textstellen (Chunks) sollten für die Nutzenden sichtbar sein und damit Vertrauen in die Resultate schaffen.

## Lösungsweg

Die Umsetzung des Projekts erfolgte in zwei sequentiellen Phasen:

### Phase 1 – Experimentelle Entwicklung und Modelloptimierung

In der ersten Phase lag der Fokus auf der Evaluation und Optimierung geeigneter Modell- und Systemparameter. Zum Einsatz kamen mittelgroße Sprachmodelle, die von Ollama bereitgestellt wurden – darunter Llama3.2 (3B), Mistral (7B) und Gemma3 (12B). Um möglichst geeignete Chunks zu erhalten, wurden unterschiedliche Embedding-Modelle getestet, darunter jina-embeddings-v2-base-de, nomic-embed-text, bge-m3 und mxbai-embed-large.

Parallel wurden verschiedene Chunk-Größen zwischen 200 und 800 Tokens sowie unterschiedliche Überlappungsstrategien erprobt. Auch das Prompt-Engineering wurde gezielt variiert – unter anderem hinsichtlich der Rollenbeschreibung, der Formatierung und von kontextuellen Hinweisen.

Alle Experimente wurden mit Hilfe von LangSmith dokumentiert und systematisch ausgewertet. Als optimale Konfiguration für den vorliegenden Anwendungsfall erwies sich die Kombination aus dem Sprachmodell Llama3.2 (3B), dem Embedding-Modell jina-embeddings-v2-base-de sowie einer Chunk-Größe von 400 Tokens ohne Überlappung. Diese Kombination überzeugte sowohl hinsichtlich Antwortqualität als auch der Verarbeitungsgeschwindigkeit.

### Phase 2 – Integration in eine produktionsreife Umgebung

Auf Grundlage der Erkenntnisse aus Phase 1 wurde die finale Systemkonfiguration als Docker-Container bereitgestellt, sodass die Edgar Schröder GmbH die vorwettbewerbliche Lösung nach kleineren Anpassungen und dem Import der gewünschten Dokumente unmittelbar einsetzen konnte. Die Infrastruktur wurde mittels Docker Compose eingerichtet, um die initiale Bereitstellung möglichst einfach zu gestalten und das Gesamtsystem jederzeit mit nur einer Befehlszeile hoch- bzw. herunterfahren zu können.

Für die Extraktion und Verarbeitung von Texten aus Dokumenten kam das Tool Docling zum Einsatz, welches insbesondere durch seine einfache Implementierung und hohe Parsingqualität überzeugte. Vereinzelt traten jedoch Probleme bei der Texterkennung in eingebetteten Bildinhalten auf, was in einigen Fällen zu unvollständiger Extraktion führte.

Die Benutzeroberfläche wurde mit Streamlit realisiert und zeichnete sich durch eine nutzerfreundliche Drag-and-Drop-Funktion zum lokalen Hochladen von Dokumenten aus. Diese Dokumente wurden automatisiert in Chunks zerlegt und anschließend in der Vektordatenbank ChromaDB gespeichert. Nutzerinnen und Nutzer konnten dann direkt über die Oberfläche Fragen eingeben und erhielten sofort qualitativ hochwertige und kontextbezogene Antworten. Zusätzlich wurden die genutzten Textchunks transparent ange-

zeigt, sodass jederzeit nachvollziehbar blieb, welche Abschnitte der Ursprungsdokumente zur Antwortbildung herangezogen wurden. Der Prompt selbst hingegen war im Backend statisch definiert und konnte nicht direkt über die Oberfläche angepasst werden.

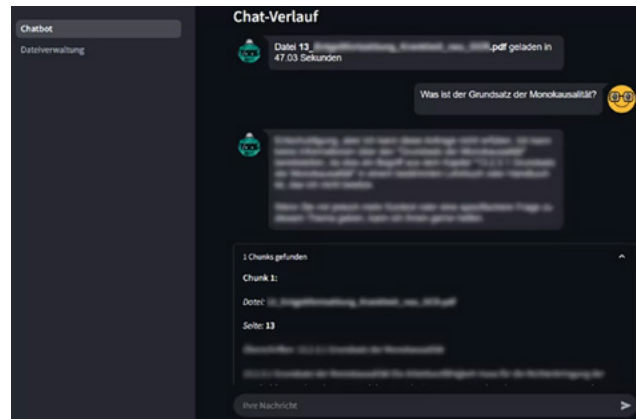


Abbildung 1: Benutzeroberfläche des Q&A-Tools mit Chat-Verlauf und Ausgabe relevanter Chunks

Abbildung 1 zeigt einen Screenshot der Benutzeroberfläche, der mit Streamlit entwickelt wurde. Die Oberfläche zeigt auf der linken Seite einen Navigationsbereich mit zwei Optionen: „Chatbot“ und „Dateiverwaltung“. Der Hauptbereich beinhaltet verschiedene Einstellungsmöglichkeiten für die Retrieval-Funktion, darunter Schieberegler für die Anzahl der abgerufenen Ergebnisse sowie den Schwellenwert für die Ähnlichkeitsbewertung (nicht in Abbildung 1 zu sehen). Unterhalb der Einstellungen befindet sich ein Bereich, der die gestellte Frage, die generierte Antwort und ermittelten Textchunks zur transparenten Nachvollziehbarkeit anzeigt. Die visuelle Darstellung verdeutlicht, wie Nutzende intuitiv die Dokumentenverarbeitung steuern und unmittelbar Ergebnisse abrufen können.

## Technische Umgebung

Das System wurde auf einem lokalen Rechner mit folgender Hardware betrieben:

- **Prozessor:** AMD Ryzen 7 5700G mit Radeon Graphics (8 Kerne)
- **Grafikkarte:** NVIDIA GeForce RTX 3060 Ti (8 GB vRAM)
- **Arbeitsspeicher:** 16 GB RAM
- **Speicher:** 1 TB NVMe SSD

Dank der Wahl kompakter, aber leistungsfähiger Modelle, ließ sich das gesamte System performant und ohne Cloud-Abhängigkeit lokal betreiben.

## Erkenntnisse & Lessons Learned

Die Entwicklung des lokal betriebenen RAG-Chatbots machte deutlich, dass die Wahl der richtigen Chunk-Größe ein zentraler



Hebel für die Antwortqualität ist: Zu kleine Textsegmente liefern zu wenig Kontext, zu große überlasten das Modell – ein Umfang von rund 400 Tokens ohne Überlappung erwies sich als tragfähiger Kompromiss. Ebenso zeigte sich, wie sensibel die Ausgabequalität auf Formulierungen im Prompt reagiert. Bereits geringfügige Anpassungen beeinflussten Relevanz und Präzision der Antworten spürbar, weshalb iteratives Prompt-Engineering unerlässlich ist. Grenzen offenbarte das Dokumenten-Parsing bei bildlastigen PDFs: Während Docling insgesamt zuverlässig arbeitete, blieb eingebetteter Text in Bildern teilweise unerkannt. Darüber hinaus wurde klar, dass die verfügbare Hardware maßgeblich über die Praxistauglichkeit entscheidet: Ohne GPU stiegen Antwortzeiten und RAM-Bedarf (auf 100 %) deutlich an. Kleinere Modelle reduzierten zwar Ressourcenverbrauch, führten jedoch zu spürbaren Qualitätseinbußen bei den Antworten. Schließlich stärkte die transparente Anzeige der genutzten Textpassagen das Vertrauen in die Ergebnisse, da die Nachvollziehbarkeit der Antworten jederzeit gegeben war.

## Fazit und Ausblick

Lokal betriebene RAG-Chatbots ermöglichen Unternehmen, internes Wissen effizient und datensouverän zu nutzen. Das Projekt hat gezeigt, dass sich mit einer sorgfältigen Auswahl von Modellen, einer passenden Infrastruktur sowie einer durchdachten Chunking-Strategie und kontinuierlichem Prompt-Engineering präzise und nachvollziehbare Antworten generieren lassen. Die Qualität der Ergebnisse hängt dabei maßgeblich von der Güte der zugrunde liegenden Dokumente ab. Mit Blick nach vorn erscheint vor allem die Integration von Wissensgraphen und Metadaten vielversprechend. Durch eine semantische Anreicherung der Dokumentbasis kann der Chatbot Beziehungen zwischen Inhalten besser verstehen und relevantere Kontexte bereitstellen.

Für Unternehmen empfiehlt sich ein schrittweises Vorgehen: Zunächst sollte ein klar umrissener Anwendungsfall – etwa im Onboarding oder im Kundenservice – definiert werden. Darauf aufbauend gilt es, geeignete Hardware bereitzustellen und insbesondere auf Struktur, Konsistenz und Aktualität der Dokumente zu achten. Mit diesem Fundament können Unternehmen die Potenziale der RAG-Technologie kontrolliert ausbauen und langfristig ein leistungsfähiges, internes Wissensmanagement etablieren.

## Autoren

### Paolo Pappe

Wissenschaftlicher Mitarbeiter am IPH – Institut für Integrierte Produktion Hannover gGmbH und KI-Trainer am Mittelstand-Digital Zentrum Hannover

### Philipp Harder

Wissenschaftlicher Mitarbeiter am IPH – Institut für Integrierte Produktion Hannover gGmbH und Experte für Digitalisierung, die Auswahl von Business Applications und KI am Mittelstand-Digital Zentrum Hannover

### Robin Stöber

Wissenschaftlicher Mitarbeiter am IPH – Institut für Integrierte Produktion Hannover gGmbH und Projekt-Ingenieur am Mittelstand-Digital Zentrum Hannover



# Automatisierte Vertragsanalyse mittels KI bei der Prowind GmbH

Projekt-Abschlussbericht von Philipp Harder und Paolo Pappe

## Projektüberblick

Wie können historisch gewachsene, unstrukturierte Datenbestände nutzbar gemacht werden? Die Prowind GmbH stand vor der Herausforderung, tausende von Bestandsverträgen effizient zu verwalten. Gemeinsam mit dem Mittelstand-Digital Zentrum Hannover wurde ein KI-gestützter Prototyp entwickelt, der Metadaten automatisch ausliest und so die Basis für eine weitreichende Digitalisierungsstrategie schuf.

Das zentrale Ergebnis des Projekts ist der erfolgreiche Machbarkeitsnachweis (Proof of Concept) für die automatisierte Vertragsanalyse mittels Künstlicher Intelligenz (KI). Es konnte gezeigt werden, dass moderne Vision Language Models (VLM) selbst bei komplexen, gescannten Dokumenten mit Handschriften eine hohe Auslesequalität erreichen. Für die Prowind GmbH lieferte das Projekt eine anbieterneutrale Bewertung der Technologie. Dies diente als entscheidende Entscheidungsgrundlage für die Geschäftsführung, um nun die operative Skalierung auf den gesamten Bestand von ca. 20.000 Verträgen mit einem externen Softwarepartner und die Integration in das System zu starten.



Die Digitalisierung im Mittelstand scheidet selten an der fehlenden Vision, sondern häufig an der Last der Vergangenheit: historisch gewachsene Datenbestände. In vielen Unternehmen schlummern in den Archiven tausende von Dokumenten, die zwar digitalisiert (gescannt) wurden, deren Inhalte aber für Computersysteme unlesbar bleiben. Es sind „stumme“ PDF-Dateien, die nicht durchsuchbar sind und deren wertvolle Informationen – wie Fristen, Klauseln oder Partnerdaten – manuell gepflegt werden müssen.

Gerade im Bereich des Vertragsmanagements birgt dies enorme Risiken und Ineffizienzen. Werden Kündigungsfristen übersehen, verlängern sich ungünstige Konditionen automatisch. Müssen für eine Due-Diligence-Prüfung hunderte Verträge gesichtet werden, bindet dies wertvolle Fachkräfte über Wochen. Das hier vorgestellte Projekt zeigt exemplarisch, wie Künstliche Intelligenz (KI) diese „stummen“ Daten zum Sprechen bringt und wie ein mittelständisches

Unternehmen durch ein Digitalisierungsprojekt mit dem Mittelstand-Digital Zentrum Hannover das Risiko einer Großinvestition minimiert.

## Unternehmen und Produkt

Die Prowind GmbH mit Sitz in Osnabrück ist ein etablierter Akteur im Bereich der erneuerbaren Energien. Das Unternehmen hat sich auf die Entwicklung, den Bau und den Betrieb von Projekten spezialisiert, die unsere Energiezukunft sichern: Windparks und Solaranlagen.

Prowind deckt dabei die gesamte Wertschöpfungskette ab – von der ersten Akquise der Flächen über die Genehmigungsplanung bis hin zur technischen Umsetzung und dem langfristigen Betrieb der Anlagen. Mit einer Mitarbeiterzahl im mittleren dreistelligen Bereich repräsentiert Prowind den typischen innovativen Mittelstand, der vor der Herausforderung steht, wachsende administrative Komplexität mit effizienten Prozessen zu bewältigen.

## Herausforderung und Zielsetzung

Die Herausforderung bei Prowind war quantitativer und qualitativer Natur. Im Laufe der Jahre war der Bestand an Verträgen auf rund 20.000 Dokumente angewachsen. Dabei handelt es sich um eine heterogene Mischung aus:

- Gestattungsverträgen: Erlaubnis zur Nutzung von Grundstücken
- Nutzungsverträgen: Langfristige Pachtvereinbarungen
- Stromlieferungsverträgen: Regelungen zur Einspeisung und Vergütung

Diese Dokumente lagen als unstrukturierte PDF-Dateien vor. Die Benennung der Dateien war historisch bedingt teilweise uneinheitlich, Ablagestrukturen in Projektordnern variierten. Besonders problematisch war die fehlende Durchsuchbarkeit nach spezifischen Kriterien. Eine einfache Frage wie „Welche Nutzungsverträge im Landkreis Osnabrück laufen 2026 aus?“ war systemseitig nicht beantwortbar und erforderte das händische Öffnen und Lesen zahlreicher Akten.

Zusätzlich zur reinen Laufzeitüberwachung sind im Bereich der erneuerbaren Energien geografische Daten essenziell. Verträge beziehen sich oft auf spezifische Flurstücke (Gemarkung, Flur, Flurstücksnummer). Diese Informationen sind in den Verträgen oft in Tabellenform oder im Fließtext versteckt und für herkömmliche Suchalgorithmen unsichtbar.

Das Ziel war es, einen Prototypen zu entwickeln, der diese Informationen mittels Large Language Models (LLM) extrahiert und strukturiert bereitstellt, um sie später in ein neues Dokumentenmanagementsystem (DMS) zu integrieren.

## Lösungsweg

In einem gemeinsamen Digitalisierungsprojekt entwickelten die Prowind GmbH und das Mittelstand-Digital Zentrum Hannover einen Prototypen zur automatisierten Vertragsanalyse. Dabei wurden zwei technische Pipelines (d. h. automatisierte Ketten von Arbeitsschritten, die nacheinander ablaufen) erprobt und miteinander verglichen:

1. Pipeline A: OCR + LLM
2. Pipeline B: Vision-Language-Model (VLM)

Der direkte Vergleich beider Ansätze war zentral, um Stärken, Grenzen und den späteren Praxiseinsatz bewerten zu können.

### Datenaufbereitung und OCR (Pipeline A)

Zunächst wurden exemplarische Verträge digitalisiert und mittels OCR (Texterkennung) in maschinenlesbaren Text überführt. Die besondere Herausforderung lag in der starken Varianz der Dokumente – von sauberen digitalen PDFs bis hin zu älteren (teilweise etwas schiefen) Scans mit handschriftlichen Passagen, etwa bei Datum oder Unterschrift.

Im Verlauf zeigte sich, dass OCR bei der Erkennung von Handschriften sowie bei Informationen aus Schaubildern, Zeichnungen oder grafischen Elementen an seine Grenzen stößt. Diese Inhalte wurden entweder unvollständig erkannt oder gar nicht in verwertbaren Text überführt.

### Einsatz von Large Language Models (LLM) (Pipeline A)

Auf Basis der OCR-Texte wurden Large Language Models (LLM) eingesetzt. Über spezifische Prompts extrahierte das Modell vordefinierte Felder aus den Vertragstexten, unter anderem:

- Vertragsart und Vertragspartner
- Laufzeitbeginn und -ende
- Verlängerungsoptionen (z. B. automatische oder optionale Verlängerung)
- Kündigungsfristen
- Geografische Daten (Gemarkung, Flur, Flurstück)

Getestet wurden sowohl lokal installierte Modelle auf geeigneter Hardware (mit leistungsfähiger Grafikkarte) als auch cloudbasierte Modelle von EU-Anbietern (u. a. aus Datenschutzgründen, z. B. mit Serverstandorten in der EU).

In der Auswertung zeigte sich, dass dieser Ansatz insbesondere bei gut lesbaren, standardisierten Textpassagen zuverlässig funktioniert. Metadaten wie Laufzeiten, Vertragspartner oder Kündigungsfristen konnten zuverlässig erkannt und strukturiert ausgegeben werden. Gleichzeitig wurde deutlich, dass die Ergebnisse stark von der vorgelagerten OCR abhängen: Bei unsaubereren Scans, handschriftlichen Ergänzungen oder Inhalten in grafischen Elementen war der OCR-Text häufig unvollständig oder fehlerhaft, was sich direkt auf die LLM-Extraktion auswirkte. Damit ist die Pipeline für viele Dokumente schnell und praxistauglich, bleibt jedoch anfällig für jene Vertragstypen, die im historischen Bestand besonders häufig vorkommen.

### Einsatz eines Vision-Language-Models (VLM) (Pipeline B)

Parallel wurde eine zweite Pipeline erprobt, die ein Vision-Language-Model (VLM) nutzt. Ein VLM verarbeitet PDFs direkt als visuelle Dokumentseiten und interpretiert Text und Layout gemeinsam. Dadurch können auch Inhalte zuverlässig ausgelesen werden, die OCR nur schlecht oder gar nicht erfasst – etwa Handschrift, Tabellenstrukturen oder Informationen innerhalb grafischer Darstellungen.

Die VLM-Pipeline erwies sich im Vergleich als deutlich robuster und lieferte spürbar bessere Extraktionsergebnisse – vor allem dort, wo OCR + LLM typischerweise an Grenzen stößt. Handschriftliche Ergänzungen (z. B. bei Datum oder Unterschrift), Informationen aus Schaubildern, Skizzen oder Zeichnungen sowie komplexe Layouts mit Tabellen oder Formularstrukturen konnten wesentlich zuverlässiger erkannt werden. Der Qualitätsgewinn ging allerdings mit einem höheren Rechenaufwand einher: Die Verarbeitung pro Dokument dauerte merklich länger als bei der OCR-LLM-Kette. Für den späteren produktiven Einsatz bedeu-



tet das, dass die überlegene Genauigkeit des VLM gegen den aktuell noch begrenzenden Zeitbedarf abgewogen und technisch optimiert werden muss.

### Validierung und Feinjustierung

Ein entscheidender Schritt war der Abgleich der KI-Ergebnisse mit einer manuellen Muster-Auswertung durch die Fachabteilung von Prowind.

- **Erfolgsmessung:** Ergebnisse wurden systematisch gegen Musterantworten geprüft. Zusätzlich wurde für jede extrahierte Information die Berechnung eines „Confidence Scores“ erprobt. In manuellen Stichproben zeigte sich jedoch, dass dieser Score nicht mit unserer Bewertung übereinstimmt und daher keine belastbare, verlässliche Metrik darstellt.
- **Erkenntnisse:** Standarddaten (z. B. Namen, einfache Datumsangaben) wurden sehr zuverlässig erkannt. Bei komplexeren juristischen Klauseln – insbesondere verschachtelten Verlängerungsoptionen – war iteratives Prompt-Engineering nötig. Die VLM-Pipeline reduzierte diese Fehler deutlich, da sie schwierige Stellen kontextreicher erfassen konnte.

### Nutzen für den Mittelstand

Das Projekt fungierte als entscheidender erster Schritt für eine langfristige Digitalisierungsstrategie.

- **Machbarkeitsnachweis (Proof of Concept):** Beide Pipelines zeigten, dass eine automatisierte Metadaten-Extraktion technisch möglich und wirtschaftlich sinnvoll ist.
- **Qualitätsgewinn durch VLM:** Der Vergleich machte deutlich, dass VLM in heterogenen, teils schlecht gescannten Vertragsbeständen qualitativ überlegen ist – besonders bei Handschrift und grafischen Dokumentanteilen.
- **Neutrale Bewertung:** Durch die Zusammenarbeit mit dem Mittelstand-Digital Zentrum Hannover erhielt Prowind eine anbieterneutrale Einschätzung der Technologie, frei von vertrieblichen Versprechen kommerzieller Anbieter.
- **Grundlage für Investitionsentscheidung:** Basierend auf den positiven Ergebnissen des Projekts – insbesondere der VLM-Qualität – entschied sich die Geschäftsführung von Prowind, einen weiteren umfangreichen Proof of Concept (POC) mit einem Start-up zu starten. Dieser umfasst die Analyse von 1.000 Verträgen und die direkte Integration in das neue Dokumentenmanagementsystem.

Das Projekt hat somit das Risiko einer Fehlinvestition minimiert und den Weg für eine Skalierung auf den gesamten Bestand von ca. 20.000 Verträgen geebnet.

*„Die Inhalte des Projektes waren für uns sehr wertvoll. Wir haben eine anbieterneutrale Einschätzung zur Machbarkeit bekommen, die nicht durch eine ‚Salesbrille‘ getrübt war. Das lieferte uns die perfekte Argumentationsgrundlage gegenüber der Geschäftsführung, um nun den nächsten Schritt in Richtung einer voll integrierten KI-Lösung zu gehen.“*

Tom Strating, Prowind GmbH

### Autoren

#### Paolo Pappè

Wissenschaftlicher Mitarbeiter am IPH – Institut für Integrierte Produktion Hannover gGmbH und KI-Trainer am Mittelstand-Digital Zentrum Hannover

#### Philipp Harder

Wissenschaftliche Mitarbeiter am IPH – Institut für Integrierte Produktion Hannover gGmbH und Experte für Digitalisierung, die Auswahl von Business Applications und KI am Mittelstand-Digital Zentrum Hannover

# Zuverlässige Finanzprognosen mit Künstlicher Intelligenz

Projekt-Abschlussbericht von Ali Soltani und Denise Wullfen

## Projektüberblick

Im Projekt mit der Termath AG wurde der Einsatz von Künstlicher Intelligenz zur Unterstützung von Finanzentscheidungen untersucht. Die KI soll Mitarbeitende durch datenbasierte Analysen und Prognosen bei der Einschätzung finanzieller Entwicklungen entlasten und unterstützen.



Viele mittelständische Unternehmen kennen die Herausforderung: Finanz- und Projektdaten liegen im ERP-System, werden aber für Auswertungen in umfangreiche Excel-Tabellen exportiert. Mit wachsender Projektzahl und Datenmenge stößt dieses Vorgehen an Grenzen: Risiken werden oft erst spät sichtbar, Liquiditätsplanungen sind aufwendig, und der Ausfall einzelner Expertinnen und Experten kann Entscheidungen verzögern.

## Unternehmen und Produkt

Die Termath AG ist ein Spezialist für sicherheitsrelevante Gebäude- und Kommunikationstechnik. Sie wurde 1955 gegründet und sitzt in Wolfsburg sowie an weiteren Standorten in Niedersachsen und Sachsen-Anhalt. 192 Mitarbeitende arbeiten an den Themen Brandschutz, Einbruchschutz, Videoüberwachung, Zutrittskontrolle, Kommunikationssysteme, IT-Services und CAD-Services. Gemeinsam mit dem Mittelstand-Digital Zentrum Hannover hat die Termath AG begonnen, Finanzdaten systematisch aus dem ERP zu extrahieren, zu bereinigen und in eine zentrale Datenbasis zu überführen.

## Problemstellung und Zielsetzung

Die Termath AG wünschte sich Unterstützung bei der Entscheidungsfindung im Finanzbereich. Der alte Prozess: Projektkosten, Bestellungen, Zeiten und Erlöse wurden manuell zusammengeführt – abhängig vom Erfahrungswissen weniger Schlüsselpersonen. Der Ausgangspunkt des KI-Anwendungsprojekts war die Herausforderung, dass relevante Informationen zur Liquiditätssteuerung in mehreren operativen Systemen verteilt vorliegen und bislang überwiegend manuell aggregiert und interpretiert werden müssen. Dies führt zu einem hohen Aufwand, einer begrenzten Prognosefähigkeit und einer Abhängigkeit von individuellem Erfahrungswissen.

Wichtig war dabei auch immer der unternehmenskulturelle Aspekt: KI soll nicht als „Black Box“, die Entscheidungen treffen. Controlling, IT und Fachbereiche arbeiten gemeinsam an Modellen und

Kennzahlen. Die ersten Erfahrungen zeigen: Wer klein beginnt – mit einem klar definierten Use Case – und konsequent auf Datenqualität und Transparenz setzt, kann KI im Finanzbereich pragmatisch nutzbar machen und zugleich einen skalierbaren Baukasten für weitere Anwendungen im Unternehmen aufbauen.

Ziel des Projekts mit dem Mittelstand-Digital Zentrum Hannover war, herauszufinden, wie ein KI-Modell diesen Prozess der Finanzprognose unterstützen und somit weniger abhängig von einzelnen Faktoren oder Erfahrungswissen einzelner Personen machen kann.

## Lösungsweg

Im Rahmen des Projekts wurde für die Termath AG ein datengetriebener Ansatz als Entscheidungsunterstützung im Finanzbereich konzeptioniert und prototypisch umgesetzt. Die Vorgehensweise im Projekt bestand darin, ein maschinelles Lernverfahren auf Basis bestehender ERP- und Finanzdaten zu trainieren, welches automatisiert beurteilen kann, ob in einem definierten Prognosezeitraum ein Liquiditätsengpass droht.

## Aufbau eines KI-Modells

Als technische Grundlage wurde ein einheitliches, Machine-Learning-fähiges Datenformat definiert und in Anforderungen aufbereitet (s. Abbildung 1), um die geeigneten Dateien aus dem ERP-System zu ermitteln. Das Format führt alle für die Engpassprognose relevanten Größen in einer einzelnen, zeitlich geordneten Tabelle zusammen. Jede Zeile dieser Tabelle repräsentiert einen Beobachtungszeitpunkt – typischerweise einen Tag oder eine Woche – und enthält sowohl die zu diesem Zeitpunkt bekannten Finanzkennzahlen als auch ein Zielattribut. Letzteres gibt an, ob innerhalb des festgelegten Vorhersagehorizonts ein Engpass eingetreten ist. Die Spalten des Templates umfassen den Zeitstempel, den aktuellen liquiden Saldo, die Summe der offenen Forderungen, die Summe der offenen Verbindlichkeiten, die monatlichen Fixkosten, die in den nächsten dreißig Tagen erwarteten Projektkosten sowie den Lagerwert als Maß für gebundene Liquidität.

### Anforderungen an Daten und Zielgrößen für ML-Modelle

C Anforderungen an Datenaufbereitung	C ML-Zielgrößen & Modelle
<ul style="list-style-type: none"> <li>○ Eindeutige IDs für Rechnungen, Projekte, Bestellungen</li> <li>○ Eindeutige Kodierung für Zahlungsstatus</li> <li>○ Fehlende Werte kennzeichnen oder imputieren</li> <li>○ Datenhistorie mind. 12 Monate</li> <li>● Zeitachsen-Synchronisation: gleiche Zeitgranularität (täglich, wöchentlich, monatlich)</li> <li>● Einheitliches Datumsformat (YYYY-MM-DD)</li> <li>● Stabiler Exportprozess (CSV/Excel/API)</li> </ul>	<ul style="list-style-type: none"> <li>□ Input: Cashflow + offene Posten</li> <li>○ Ziel: Zeitpunkt Engpass</li> <li>□ Input: Historischer Saldo</li> <li>● Ziel: Engpass (ja/nein)</li> <li>■ Modelltyp: Klassifikation (XGBoost, IsoForest)</li> <li>■ Modelltyp: Zeitreihe (Prophet, LSTM)</li> </ul>

Abbildung 1: Anforderungen für ML-Modelle

Ergänzt wird dies durch eindeutige Identifikatoren für Kunden, Rechnungen, Bestellungen und Projekte, die eine Rückverfolgbarkeit der Daten zu den Quellsystemen ermöglichen. Das Label „Engpass“ ist als binäre Zielvariable kodiert: Der Wert 1 bedeutet, dass der prognostizierte Liquiditätssaldo innerhalb des Vorhersagefensters negativ wird, kennzeichnet also einen Engpass. Dieses Label wird nicht manuell vergeben, sondern algorithmisch aus historischen Zahlungsströmen und Kontostandsverläufen berechnet. Dazu werden für jeden Stichtag die erwarteten Ein- und Auszahlungen über den Prognosehorizont mit dem aktuellen Saldo verrechnet. Diese Datenquellen wie Rechnungen, Bestellungen und Projektstatus müssen ebenfalls bestimmte Anforderungen erfüllen, wie in Abbildung 2 ersichtlich. Auf Basis dieses einheitlichen Datenformats wurde eine durchgängige ML-Pipeline in Python implementiert. Die Daten werden zunächst eingelesen, chronologisch sortiert und um Zeitmerkmale wie Jahr, Monat und Wochentag ergänzt, um saisonale oder periodische Effekte modellierbar zu machen. Identifikatoren und der Zeitstempel selbst werden vor dem Training aus der Feature-Matrix entfernt, da sie dem Modell keine numerischen Informationen liefern. Sie bleiben jedoch im Datensatz erhalten, um sie später interpretieren zu können.

**Training als Grundlage für KI-Erfolg**

Der Datensatz wird zeitbasiert in einen Trainings- und einen Testbereich aufgeteilt. Dadurch wird das Modell ausschließlich auf früheren Daten trainiert und auf späteren Daten validiert, was einer realistischen Einsatzsituation entspricht. Da Liquiditätsengpässe im Vergleich zu normalen Betriebszuständen typischerweise selten sind, wird das Klassenungleichgewicht durch eine gewichtete Verlustfunktion berücksichtigt. Als Kernalgorithmus wurde ein XGBoost-Klassifikator eingesetzt. XGBoost ist ein Gradient-Boosting-Verfahren, das eine Vielzahl von Entscheidungsbäumen sequenziell trainiert, wobei jeder neue Baum die Fehler der vorherigen korrigiert. Dadurch ist das Modell in der Lage, komplexe, nichtlineare Zusammenhänge zwischen den finanziellen Kennzahlen zu lernen. Es kann beispielsweise Situationen erkennen, in denen ein hoher Kontostand durch gleichzeitig sehr hohe kurzfristige Zahlungsverpflichtungen überkompensiert wird.

Das Modell gibt für jeden Beobachtungszeitpunkt eine Wahrscheinlichkeit für das Eintreten eines Engpasses aus. Durch einen einstellbaren Schwellenwert wird diese Wahrscheinlichkeit in eine Entscheidung überführt. Die Bewertung erfolgt anhand von verschiedenen Metriken. So können datenbasierte Prognosen durchgeführt werden. Die Auswertungen zeigen, dass das Modell Engpasssituationen mit einer Trennschärfe von stabilen Phasen unterscheiden kann. Zusätzlich wird eine Schwellenwertanalyse durchgeführt, um den Zielkonflikt zwischen Fehlalarmen und verpassten Engpässen explizit darzustellen und betrieblich geeignete Einstellungen zu ermöglichen.

Ein weiterer wesentlicher Bestandteil des Prototyps ist die Erklärbarkeit. Mithilfe der Feature-Importance-Auswertung des XGBoost-Modells sowie einer Fall-back-Analyse mittels Permutation Importance lässt sich ermitteln, welche Eingangsgrößen den größten Einfluss auf die Engpassprognose haben. In der Regel sind dabei der liquide Saldo, die offenen Verbindlichkeiten, die offenen Forderungen und die kurzfristigen Projektkosten die dominanten Treiber. Diese Informationen sind entscheidend, um die Modellentscheidungen gegenüber Fachabteilungen nachvollziehbar zu machen und um konkrete Handlungsoptionen abzuleiten.

<b>C</b> Rechnungen (Ausgang)	<b>C</b> Bestellungen (Einkauf)	<b>C</b> Projektstatus
<ul style="list-style-type: none"> <li><input type="checkbox"/> Rechnungsnummer</li> <li><input type="checkbox"/> Kunde</li> <li><input type="checkbox"/> Rechnungsdatum</li> <li><input type="checkbox"/> Betrag</li> <li><input type="checkbox"/> Zahlungsziel</li> <li>Format: CSV/Excel</li> <li>Verwendung: Cash-In Prognose</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Bestellnummer</li> <li><input type="checkbox"/> Lieferant</li> <li><input type="checkbox"/> Betrag</li> <li><input type="checkbox"/> Fälligkeit/Zahlungsziel</li> <li><input type="checkbox"/> Status</li> <li>Format: CSV/Excel</li> <li>Verwendung: Cash-Out Prognose</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Projekt-ID</li> <li><input type="checkbox"/> Start-/Enddatum</li> <li><input type="checkbox"/> Erwarteter Umsatz</li> <li><input type="checkbox"/> Erwartete Kosten</li> <li><input type="checkbox"/> Ist-Kosten</li> <li>Format: CSV/Excel</li> <li>Verwendung: Projekt-Cashflow-Analyse</li> </ul>
<ul style="list-style-type: none"> <li><input checked="" type="checkbox"/> Status („bezahlt“ / „offen“)</li> </ul>		<ul style="list-style-type: none"> <li><input checked="" type="checkbox"/> Fortschritt (%)</li> </ul>

Abbildung 2: Anforderungen an Datenquellen (Auszug)

Der Modellierungsprozess (s. Abbildung 3) ist transparent abbildbar. Die gesamte Lösung wurde in einem reproduzierbaren Jupyter-Notebook umgesetzt und ausgeführt. Dabei kamen Python-Bibliotheken wie Pandas und Numpy für die Datenverarbeitung, Scikit-Learn für Metriken und Hilfsfunktionen, XGBoost für das eigentliche Lernverfahren und Matplotlib für die Visualisierung zum Einsatz. Mithilfe des bereitgestellten Excel-Templates mit klar definierten Headern, Datentypen und dem Engpass-Label kann die Termath AG ihre realen ERP-Daten in das vorgegebene Format überführen und den Prototypen mit Echtzeiten betreiben.

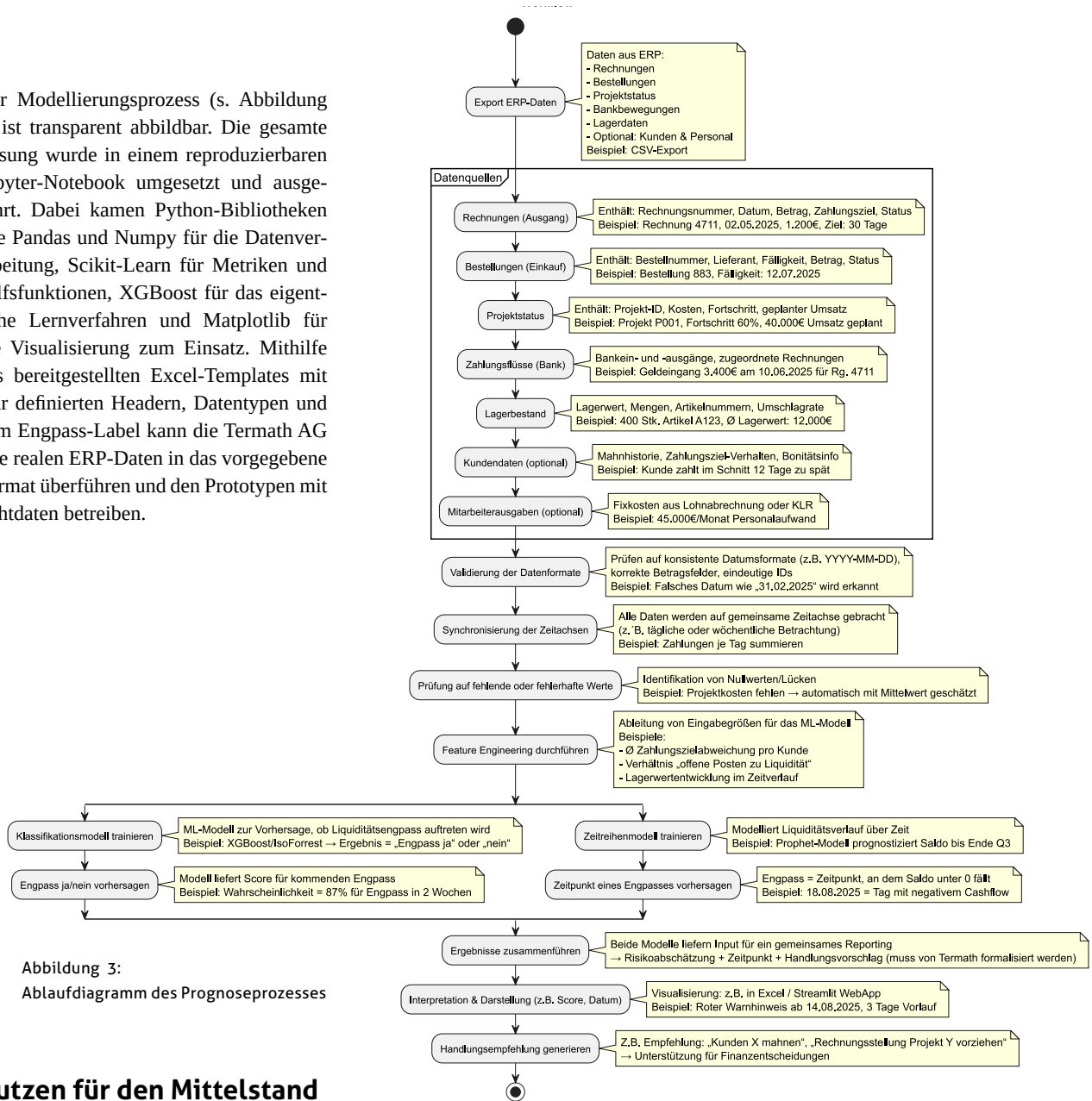


Abbildung 3: Ablaufdiagramm des Prognoseprozesses

## Nutzen für den Mittelstand

Das Projekt zeigt, dass eine KI-gestützte Liquiditätsprognose auf Basis vorhandener Unternehmensdaten realisierbar ist und sowohl quantitative Vorhersagen als auch erklärbare Entscheidungsgrundlagen liefern kann. Dies kann insbesondere KMU zukünftig dabei unterstützen, Finanzentscheidungen aufwandsarm auf Basis von Unternehmensdaten, nicht nur aufgrund von Erfahrungswissen zu treffen. Dies ermöglicht eine zuverlässige Entscheidungsunterstützung für Unternehmen.

Eine KI-basierte Prognose bietet im Vergleich zu reinem Expertenwissen mehrere Vorteile. Die Prognosegenauigkeit kann durch die Mustererkennung höher sein, da KI-Modelle wiederkehrende Muster wie saisonale Schwankungen erkennen. Zudem können heterogene Datenquellen automatisiert verarbeitet werden, was weniger zeitaufwendig ist. Darüber hinaus besteht so die Möglichkeit für KMU, durch Szenarioanalysen drohende Liquiditätsengpässe frühzeitig zu identifizieren. Ein weiterer großer Vorteil ist die Unabhängigkeit von einzelnen Mitarbeitenden, was besonders für Unternehmen mit hohem Fachkräftemangel oder einem Führungswechsel interessant sein kann, da das Expertenwissen nicht immer im selben Maße zur Verfügung steht. Transparente und datenbasierte Entscheidungsgrundlagen ermöglichen eine Unterstützung für das Management in KMU, beispielsweise wenn es um neue Investitionen geht.

### Autor\*innen

#### Ali Soltani

Projektingenieur am IPH – Institut für Integrierte Produktion Hannover gGmbH und Experte für Künstliche Intelligenz und ihre Entwicklung am Mittelstand-Digital Zentrum Hannover.

#### Denise Wullfen

Wissenschaftliche Mitarbeiterin am IPH – Institut für Integrierte Produktion Hannover gGmbH und Expertin für die Einführung von Künstlicher Intelligenz am Mittelstand-Digital Zentrum Hannover.

# Kalk, Sand und KI: Multisensorische Defekterkennung für die Produktion mineralischer Baumaterialien

Projekt-Abschlussbericht von Dr. Claudia Niederée und Sofiane Laridi

## Projektüberblick

In diesem Projekt konnte gezeigt werden, dass Methoden der künstlichen Intelligenz (KI) eingesetzt werden können, um auch kleine Defekte in mineralischen Baumaterialien, insbesondere in Kalksandsteinen und Porenbeton, erfolgreich und frühzeitig bereits während der Produktion zu erkennen. Das Projekt wurde in Zusammenarbeit mit der Forschungsvereinigung Kalk-Sand e.V. durchgeführt, die ihre Domänenexpertise in das Projekt eingebracht und ihr Technikum für praxisnahe Experimente zur Verfügung gestellt hat. Nach einer Analyse relevanter Defektarten und des Produktionsprozesses wurde eine Datenerfassung (RGB-Bilder und Thermal-Bilder) durchgeführt. Auf der Grundlage der annotierten Daten wurde ein leistungsstarkes CNN-Modell zur Defekterkennung trainiert, das einen Patch-basierten Ansatz nutzt, um auch sehr kleine Defekte erkennen zu können. Zudem wurden unterschiedliche Sensorfusionsstrategien erprobt, um sich ergänzende Bildinformationen von multiplen Sensoren z. B. Thermalbilder und RGB-Bilder optimal zu kombinieren. Die Analyse der Strategien zur Sensorfunktion hat experimentell gezeigt, dass die Fusion unterschiedlicher Sensordaten, das Analyseergebnis der eingesetzten KI verbessern kann. Es hat sich aber auch gezeigt, dass es dabei einen großen Unterschied macht, wie bzw. wann im Analyseprozess die Fusion durchgeführt wird.

Die Herstellung von Kalksandstein und Porenbeton sind wichtige Prozesse für die Bauwirtschaft. In Deutschland werden z.B. jährlich rund 2,3 Milliarden Kalksandsteine (entspricht etwa 8 Millionen Tonnen Kalksandsteinmaterial) produziert [1]. Kalksandstein wird aus Sand, Kalk und Wasser hergestellt. Nach dem Mischen und Formen wird Kalksandstein in einem Autoklav unter hohem Druck und Temperatur dampfgehärtet. Dadurch entsteht ein besonders druckfester und maßhaltiger Baustoff. Porenbeton besteht aus Sand, Kalk, Zement, Wasser und einem Treibmittel (meist Aluminium). Nach dem Mischen bildet sich durch das Treibmittel eine porige Struktur. Der entstandene Teig wird vorgehärtet, geschnitten und anschließend ebenfalls in einem Autoklav dampfgehärtet. Das Ergebnis ist ein leichter, wärmedämmender Baustoff. Beide Prozesse sind durch die Notwendigkeit die hergestellten Steine und andere Baumaterialien zu „backen“ mit hohen Energieverbräuchen verbunden.

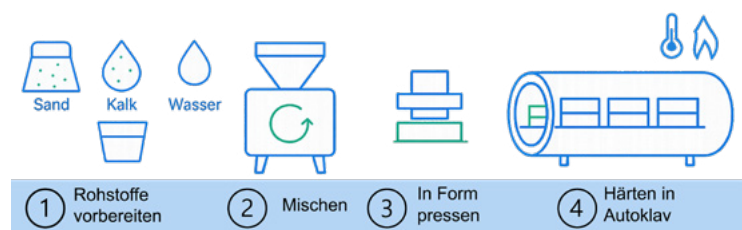


Abbildung 1: Wichtigste Schritte in der Herstellung von Kalksandstein

Neben der Verringerung des Materialeinsatzes durch verstärktes Recycling ist es eine wirksame Maßnahme zur Reduzierung des Energieverbrauchs in der Produktion und zur Erreichung der CO<sup>2</sup>-Neutralität, defekte Steine bereits vor dem Autoklavenprozess zu erkennen, auszusortieren und das Material wieder in die Produktion zurückzuführen. Auf diese Weise können unnötige Produktionsschritte für nicht nutzbare Steine vermieden werden.

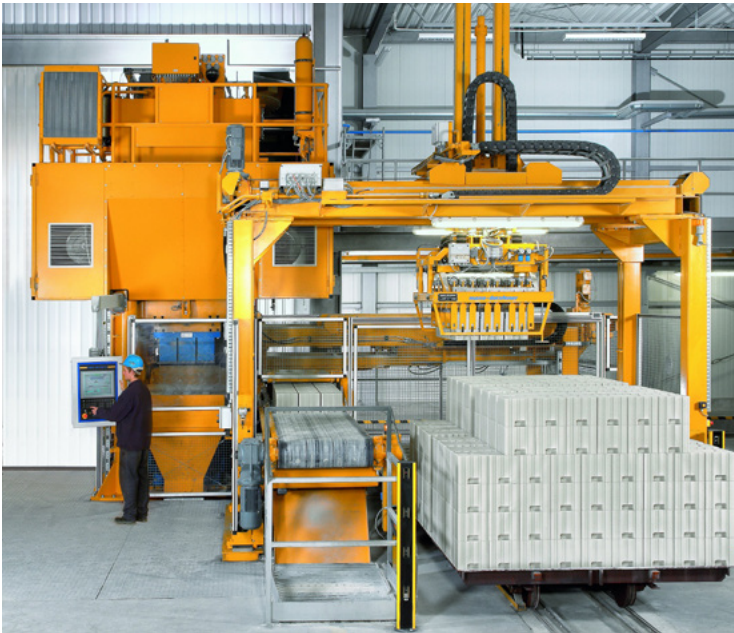


Abbildung 2: Formen und Pressen der Rohlinge und Vorbereitung zum Härten



Abbildung 3: Härten der Steine im Autoklav

## Unternehmen

Das Projekt wurde in Zusammenarbeit mit der Forschungsvereinigung Kalk-Sand e.V. durchgeführt. Die Forschungsvereinigung Kalk-Sand e.V. führt vorwettbewerbliche Forschung zu Fragestellungen der Kalksandsteinindustrie sowie auch zu verwandten Branchen durch. Um geschlossene Stoffkreisläufe, grüne Energie oder auch nachhaltiges Bauen weiter voranzutreiben, widmen sich die Forschungsarbeiten der Vereinigung insbesondere Fragestellungen zum Umweltschutz, zum Recycling von Kalksandsteinmaterial, zur Entwicklung innovativer Produkte und Verfahren, zur Energie- und CO<sub>2</sub>-Einsparung bei der Produktion und zur Steigerung der Leistungsfähigkeit von Kalksandstein. In ihrer Rolle als Forschungsvereinigung für die Kalksandsteinindustrie und durch die enge Zusammenarbeit mit zahlreichen Unternehmen in diesem Bereich ist sie in der Lage, praxisnah wichtige Impulse zu geben und Pilotlösungen für die nachhaltige Weiterentwicklung dieses Produktionsbereichs zu entwickeln.

## Problemstellung/Zielsetzung

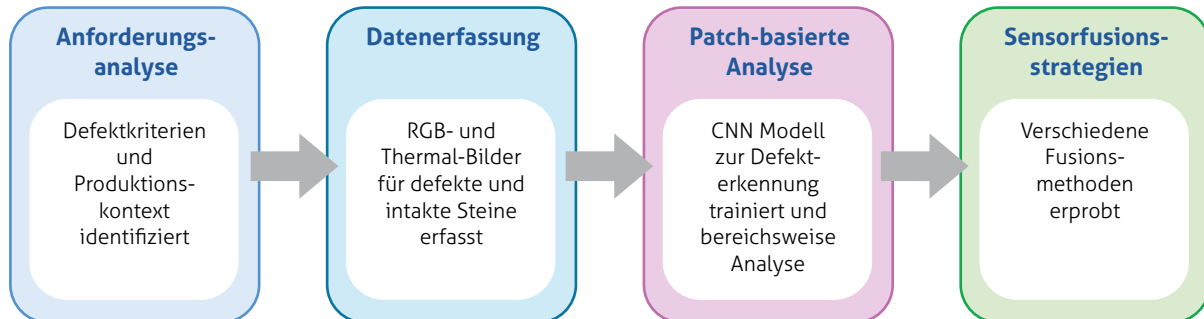
In der Produktion von Kalksandstein und Porenbeton können verschiedene Defekte, wie Risse, Aufwölbungen oder Luncker auftreten, die die mechanischen Eigenschaften beeinträchtigen und somit zu einer höheren Ausschussrate, Nacharbeitung oder Qualitätsmängeln führen. Die Unternehmen setzen zur Qualitätskontrolle vor allem auf manuelle, oft stichprobenartige Sichtkontrollen, um Defekte vor oder nach der Autoklavierung zu erkennen. Bei festgestellten Defekten erfolgt Ausschuss, Nachbearbeitung oder Zerstückelung und Recycling des Materials.

Angestrebt ist eine Verbesserung der Defekterkennung durch digitale Methoden und der Einsatz von KI. Die (automatisierte) Erkennung solcher Defekte ist jedoch nicht einfach, da a) die Grenze zwischen rauer Oberflächenstruktur und Defekt nicht einfach zu ziehen ist, b) viele unterschiedliche Arten von Defekten auftreten und c) die Wärme und Druckbehandlung im Autoklav die Defektentwicklung beeinträchtigt. Insbesondere kommt es während der Autoklavierung zu einer lokalen Spannungsumlagerung im Bereich vorhandener Risse bzw. an den Risspitzen. Dadurch wird das Risswachstum begünstigt und bestehende Risse können sich weiter ausbreiten. Zudem befinden sich viele Defekte (z.B. Risse, Luncker oder ungleichmäßige Porenverteilung) unterhalb der Oberfläche; sie sind mit bloßem Auge und herkömmlichen RGB-Kameras oft nicht zu erkennen. Um sie vor der kosten-, zeit- und energieaufwändigen Autoklavierung zu erkennen, müssen weitere bildgebende Technologien herangezogen und die verschiedenen Bilddaten fusioniert werden, sodass nicht-oberflächliche Defekte digital sichtbar werden.

Ziel des Projekts war die Erprobung und Bewertung verschiedener bildgebender Technologien inklusive datengetriebener Sensorfusion, um oberflächliche und nicht-oberflächliche Defekte frühzeitig und automatisiert vor der energie-, kosten- und CO<sub>2</sub>-intensiven Autoklavierung KI-basiert zu detektieren. Damit soll demonstriert werden, dass ein Inline-Ausschuss vor der Autoklavierung technisch möglich ist und für die Industrie die technische Grundlage für eine fundierte Entscheidung zur Einführung solcher Systeme gelegt werden.

## Lösungsweg

### KI-basierte Defekterkennung



Eine wichtige Grundlage für die Erprobung einer KI-Lösung war es die Anforderungen an eine solche Lösung in Produktionsprozess von Kalksandstein und Porenbeton zu verstehen. Hierzu wurden zum einen die unterschiedlichen Arten von Defekten gemeinsam mit Domänenexperten systematisch nach Größe, Häufigkeit, Wirkung und Auftreten analysiert. Dabei wurden die unterschiedlichen Produktionsschritte berücksichtigt. Aus dieser Analyse ergaben sich wichtige Erkenntnisse hinsichtlich der Größe, Sichtbarkeit und Tiefe von Defekten.

Eine weitere wichtige Analyse in dieser Phase des Projekts war die Betrachtung des Produktionsprozesses, wie z.B. die Stückelung und Paketierung der Steine während der Produktion, die alle Auswirkung auf die praktische Realisierung einer KI-Lösung in einem solchen Kontext haben. So sind z.B. Kalksandsteine vor dem Backen relativ empfindlich, was bei der Bewegung für eine Bilderfassung beachtet werden muss und die Produktionsumgebung ist typischerweise heiß, staubig und feucht, was zusätzliche Anforderungen an die in der Praxis eingesetzten Sensoren stellt.

Zur Erprobung von KI-Lösungen wurden Bilder von intakten und defekten Kalksandsteinen am Technikum der Forschungsvereinigung Kalk-Sand e.V. in Hannover erfasst. Dabei wurde zum einen eine RGB-Kamera mit hoher Auflösung und zum anderen ein spezialisiertes Dual-Sensor-Kamerasystem für RGB und thermische Bilder verwendet. Das verfü-

bare Dual Sensor-System hatte dabei eine niedrigere Auflösung und erfasst synchronisierte Bildpaare. Dadurch wird sichergestellt, dass jeder Pixel im thermischen Bereich der entsprechenden räumlichen Koordinate im RGB-Bereich zugeordnet werden kann. Ein Teil der Bilder wurde nach dem Erhitzen der Kalksandsteine gemacht, um eine kombinierte Nutzung der unterschiedlichen Sensoren besser erproben zu können. RGB-Sensoren sollten auf der Grundlage der Lichtreflektion oberflächliche Risse und Verfärbungen gut erfassen, während der Thermalsensor die thermische Trägheit des Materials erfasst und damit auch Löcher unterhalb der Oberfläche und Mikrorisse besser erkennen kann.

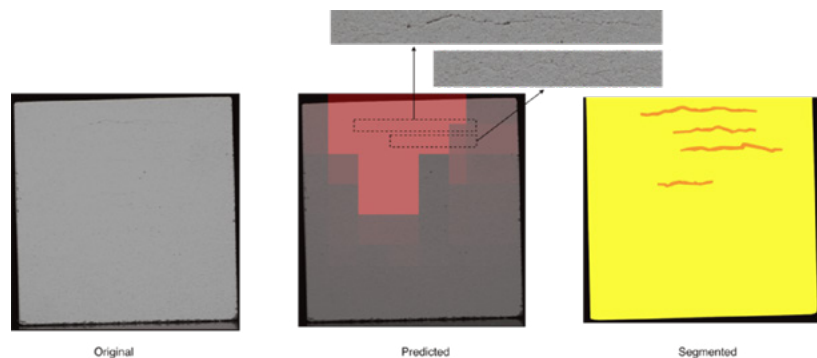


Abbildung 4: Bearbeitungsergebnisse für ein RGB-Beispielbild: Originalaufnahme (rechts), Erkennung der Defekte durch die patch-basierte Methode (Mitte, mit Details) und Visualisierung der Defekte durch pixelbasierte Segmentierung (links).

Nach einigen Vorbearbeitungsschritten wurde als Teil der KI-Methode ein Patch-basiertes Verfahren verwendet. Durch die Methode, bei der in einem Sliding-Window Verfahren kleine, überlappende Bereiche (Patches) einzeln analysiert werden, können auch kleinere Risse gut detektiert werden. Im Kern wurde als KI-Methode ein ResNet-18 Deep Convolutional Neural Network (CNN) genutzt, das mit ImageNet Daten für die allgemeine Bildererkennung vor-trainiert war. Zur Defekterkennung wurde der letzte Layer des Netzwerks mit erfassten und annotierten Bildern durch spezielles Training angepasst (Fine-Tuning). Das für die hochauflösenden RGB-Aufnahmen trainierte KI-Modell wurde mit den Testdaten evaluiert und konnte Defekte mit hoher Präzision erkennen (AUROC 0.94).

Für die Kombination der unterschiedlichen Sensoren (Thermal + RGB) wurden unterschiedliche Fusionsansätze erprobt: Early fusion (Kombination am Eingang des neuronalen Netzwerkes), mid fusion (Kombination im Inneren des neuronalen Netzwerkes) und late fusion (separate KI, Kombination der Ergebnisse). Bei der Erprobung der unterschiedlichen Fusionsstrategien hat sich gezeigt, dass die Fusion innerhalb des neuronalen Netzes (Mid Fusion) die besten Ergebnisse liefert. Durch diese Strategie ergab sich bei der Fusion eine moderate Verbesserung der Analyseleistung gegenüber der alleinigen Betrachtung von RGB-Bildern. Es hat sich aber auch gezeigt, dass die Wahl der richtigen Fusionsstrategie sehr wichtig ist: So hat z.B. die Early Fusion Strategie zu einer signifikanten Verschlechterung der Ergebnisse gegenüber der reinen Nutzung von RGB-Bildern geführt.

## Nutzen für den Mittelstand

Das Projekt schafft einen fundierten, übertragbaren Orientierungsrahmen für Unternehmen in der Produktion von mineralischen Mauersteinen und fördert den zukünftigen Einsatz KI-gestützter Defekterkennung in der Baustoffindustrie. Die gewonnenen Erkenntnisse zu geeigneten Technologien und Verfahren können von weiteren Unternehmen als Basis für eigene Entwicklungen und Investitionsentscheidungen genutzt werden.

[1] Kalksandstein Geschäftsberichte 2024+ 2024, Bundesverband Kalksandsteinindustrie e. V., [www.kalksandstein.de](http://www.kalksandstein.de)

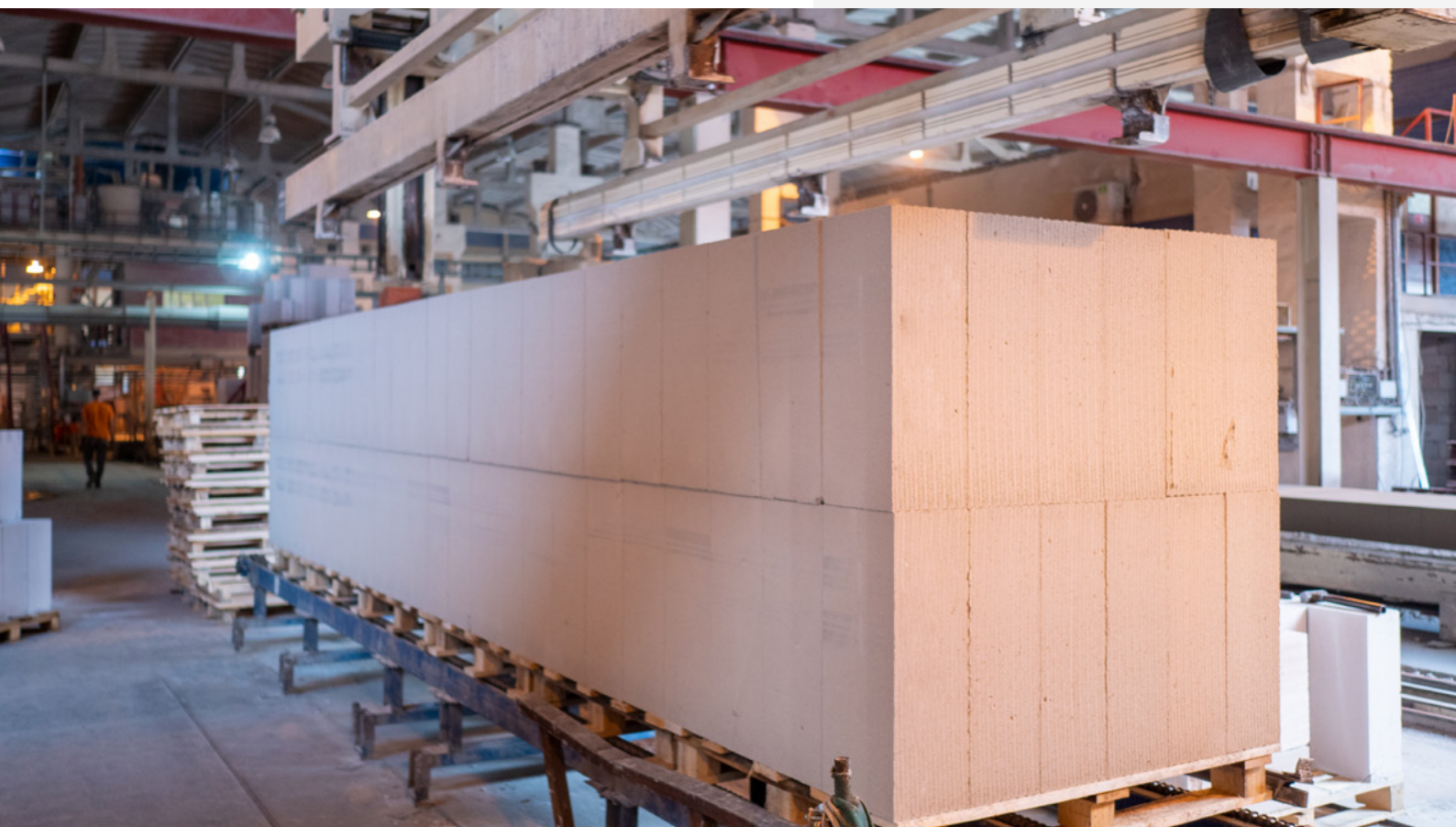
### Autor\*innen

#### Dr. Claudia Niederée

Forschungsgruppenleiterin und Geschäftsführerin am KI-Forschungszentrum L3S und Expertin für KI am Mittelstand-Digital Zentrum Hannover

#### Sofiane Laridi

Wissenschaftlicher Mitarbeiter am Forschungszentrum L3S und Experte für KI-Methoden, insbesondere Anomalie-Erkennung und föderiertes Lernen und Experte für KI am Mittelstand-Digital Zentrum Hannover



# KI für schnellere Angebote: Zeichnungen auslesen, Herstellzeiten prognostizieren, Risiken quantifizieren

Projekt-Abschlussbericht von Marcus Nein



### Projektüberblick

Im Projekt mit der INOMETA GmbH wurde die Nutzung von KI zur automatisierten Analyse technischer Zeichnungen und zur Prognose von Fertigungszeiten untersucht. In der Machbarkeitsstudie konnte gezeigt werden, dass in den unterschiedlichen Baugruppen des Variantenfertigers ab einer Datenmenge von 500 Baugruppen je Baugruppenklasse bessere Prognoseergebnisse im historischen Datensatz erzielt werden, als mit der bislang eingesetzten Methode.

### Unternehmen und Produkt

Die INOMETA GmbH zählt zu den führenden Herstellern von Leichtbauprodukten aus Aluminium sowie Faserverbundwerkstoffen. Das Unternehmen entwickelt, fertigt und veredelt rotierende Komponenten für die Folien-, Papier-, Druck- und Verpackungsindustrie. Zum Produktspektrum gehören unter anderem funktionsbeschichtete Walzen, Bauteile aus kohlenstofffaserverstärktem Kunststoff und glasfaserverstärktem Kunststoff sowie Präzisionsrohre und Druckwalzen. Neben der Produktfertigung bietet INOMETA einen umfassenden Walzenservice an und unterstützt Kunden entlang des gesamten Lebenszyklus ihrer Walzensysteme.

### Herausforderung und Zielsetzung

Der Angebotsprozess bei variantenreichen technischen Produkten erfordert eine präzise Bewertung der Kundenanforderungen sowie eine belastbare Abschätzung von Fertigungszeiten und -kosten. Technische Zeichnungen stellen dabei die wichtigste Informationsquelle dar, sind jedoch häufig nur eingeschränkt maschinenlesbar. Relevante Merkmale wie Toleranzen, Bearbeitungskennzeichen oder funktionsrelevante Geometrien müssen bislang manuell identifiziert werden. Dies stellt einen zeitintensiven Schritt dar, der Erfahrungswissen erfordert.

Aus vergangenen Aufträgen liegen umfangreiche historische Rückmeldedaten vor, in denen reale Arbeitspläne und Durchlaufzeiten dokumentiert sind. Diese Daten bieten großes Potenzial für eine datenbasierte Prognose der Herstellzeiten, werden im Angebotsprozess bislang jedoch nicht ausreichend genutzt. Die Kombination aus manueller Zeichnungsbewertung und heuristischen Zeitschätzungen geht mit Unsicher-

heiten in der Zeitanpassbildung und erhöhten Planungsaufwänden einher, insbesondere bei komplexen oder neuen Produktvarianten.

Ziel der gemeinsamen Machbarkeitsstudie des Mittelstand-Digital Zentrums Hannover und der INOMETA GmbH war es, zu untersuchen, ob moderne KI-Methoden eine automatisierte Analyse technischer Zeichnungen ermöglichen und aus Vergangenheitsdaten präzise Bearbeitungszeitprognosen abgeleitet werden können. Zusätzlich wurde geprüft, ob neben einem reinen Prognosewert auch eine Quantifizierung der Unsicherheit möglich ist, um die Transparenz hinsichtlich der Verlässlichkeit der Prognose zu erhöhen.

Angewendet wurde hierfür das methodische Vorgehen aus dem Projekt SzenoKalk des Instituts für Fertigungstechnik und Werkzeugmaschinen (IFW) (siehe Abbildung 1). Dabei werden Baugruppeninformationen extrahiert und über Korrelationen zu den Rückmeldedaten bereits gefertigter Baugruppen identifiziert. Diese Zusammenhänge sollen für neue Baugruppen mit den verfügbaren Informationen über die durchzuführenden Fertigungsmittel in der Angebotsphase für Kosten- und Fertigungszeitprognosemodelle genutzt werden. Auf dieser Grundlage besteht die Möglichkeit, den Angebotsprozess künftig datenbasierter, schneller und objektiver zu gestalten.

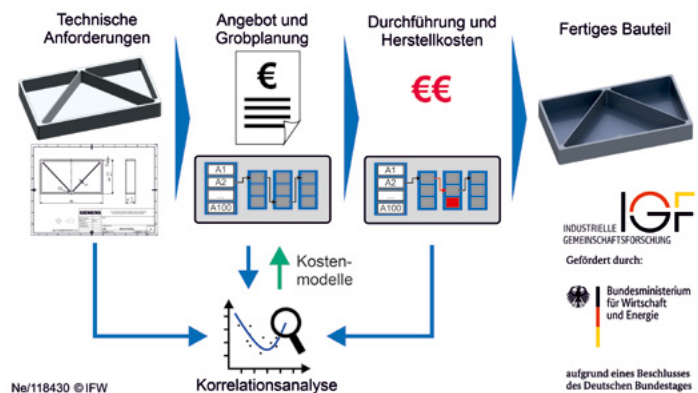


Abbildung 1: Methodisches Vorgehen zur Erstellung von Kostenprognosemodellen in der Angebotsphase

### Lösungsweg

#### Automatisierte Analyse technischer Zeichnungen

Zur Bewertung der Potenziale bei der automatisierten Auswertung von technischen Zeichnungen wurde ein KI-gestütztes Verfahren getestet, das sowohl Textinformationen als auch geometrische Symbole automatisiert erkennt. Zum Einsatz kamen zwei etablierte Computer-Vision-Modelle: PaddleOCR zur Extraktion von Schrift- und Maßangaben sowie YOLO zur Erkennung von Symbolen und Toleranzelementen. Da in der industriellen Praxis häufig keine annotierten Trainingsdaten für die Symbolerkennung vorliegen, wurde im Projekt ein synthetischer Trainingsdatensatz erzeugt. Hierfür wurden technische PDF-Zeichnungen

mithilfe des interaktiven CAD/CAM/CAE-Systems Siemens NX vom IFW automatisiert erstellt und anschließend mit der Python-Bibliothek PyMuPDF analysiert.

PyMuPDF ermöglicht den programmgesteuerten Zugriff auf die Vektordaten einer PDF-Datei und erlaubt das Auslesen grafischer Elemente einschließlich ihrer geometrischen Positionen. Auf diese Weise konnten die in den Zeichnungen enthaltenen Symbole automatisiert erkannt und deren Positionen als Trainingslabels für das Nachtrainieren des YOLO-Symbolerkennungsmodells abgeleitet werden. Eine manuelle Annotation war somit nicht erforderlich.

Um die Robustheit und Generalisierungsfähigkeit des Modells zu erhöhen, wurden die Positionen der erkannten Symbole zusätzlich systematisch variiert. Die resultierende Heatmap zeigt, an welchen Stellen der Zeichnung diese synthetisch erzeugten Symbolpositionen platziert wurden (siehe Abbildung 2). Damit lässt sich nachvollziehen, wie breit das Modell hinsichtlich möglicher Layouts und Platzierungen angelernt wurde.

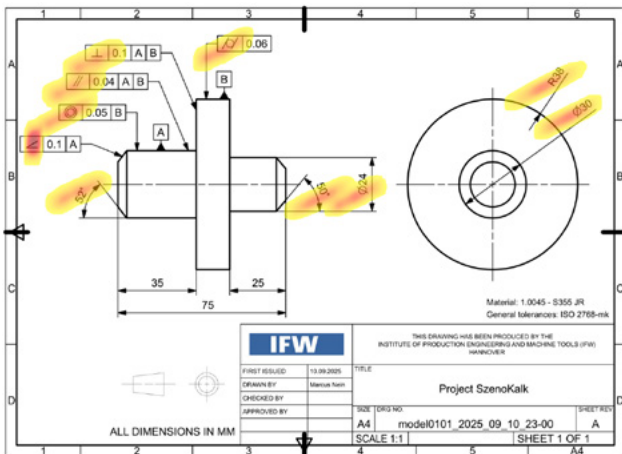


Abbildung 2: Heatmap der wechselnden Zeichenpositionen bei dem Bauteilkonfigurator

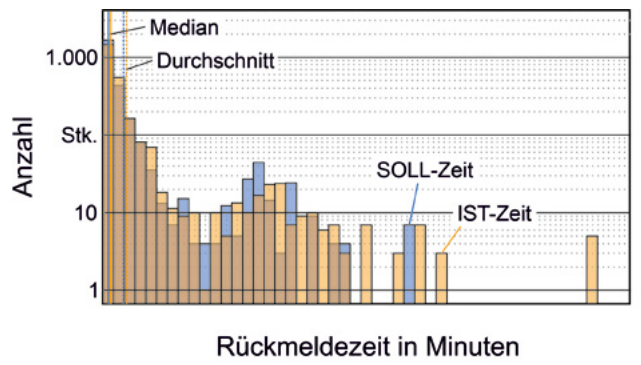
Der entwickelte Ansatz dient als prototypischer Workflow, der es ermöglicht, technische Zeichnungen maschinenlesbar aufzubereiten und daraus relevante Merkmale für weitere Auswertungen zu gewinnen. Die Vorgehensweise wurde an ausgewählten Zeichnungsbeispielen erprobt und bildet die Basis für eine spätere Integration in den Angebotsprozess. Die Übertragbarkeit auf technische Zeichnungen aus weiteren CAD-Systemen mit einer hohen Varianz an Schriftarten wird als technisch machbar eingeschätzt und stellt überwiegend eine IT-seitige Fragestellung dar.

**KI als Unterstützung in der Angebotskalkulation**

Da Herstellkosten zeitabhängigen Einflussfaktoren wie Inflation oder Lohnentwicklungen unterliegen, wird im Projekt nicht der Kostenwert selbst, sondern die zugrunde liegende Herstellzeit prognostiziert. Diese Kennzahl ist bei konstanten Fertigungstechnologien deutlich stabiler und kann im Anschluss über aktuelle Kostensätze in einen Kostenwert überführt werden.

Für die automatisierte Abschätzung von Bearbeitungszeiten werden zunächst alle verfügbaren technischen Informationen der jeweiligen Baugruppe als Eingabeparameter genutzt. Dazu zählen strukturelle Merkmale aus den technischen Zeichnungen, produkt- und fertigungsrelevante Stammdaten sowie die Rückmeldezeiten der Bauteile. Auf dieser Grundlage wird die voraussichtliche Fertigungszeit eines neuen Bauteils datenbasiert prognostiziert.

Die Basis dafür bilden umfangreiche historische Rückmeldedaten, in denen reale Rüst-, Maschinen- und Laufzeiten dokumentiert sind. Eine Analyse dieser Daten zeigt Unterschiede zwischen geplanten und tatsächlich angefallenen Zeiten. Die Verteilung dieser Abweichungen verdeutlicht, wie variantenreich die Produktgruppen sein können und welches Prognosespektrum ein Modell abdecken muss (siehe Abbildung 3).



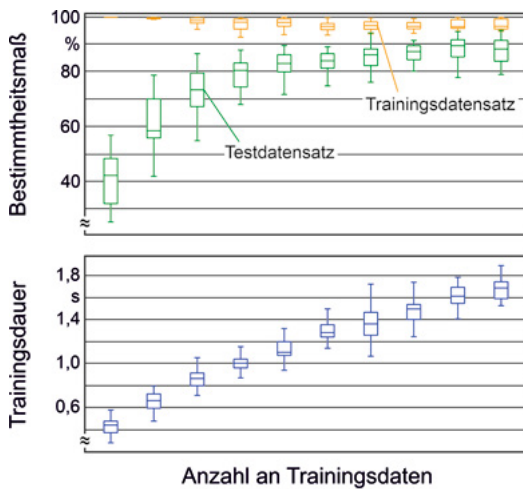
Anzahl der Baugruppen in dieser Baugruppenklasse = 2.541  
 Abweichung des Durchschnitts zwischen SOLL und IST = +14,1 %

№/118434 © IFW

Abbildung 3: Rückmeldezeitenverteilung einer Baugruppenklasse

Zur Modellierung der Herstellzeiten wurden automatisierte Machine-Learning-Verfahren (AutoML) aus der Python-Bibliothek TPOT getestet. Diese Ansätze durchsuchen eigenständig eine Vielzahl möglicher Regressionsmodelle und identifizieren jene, die die vorliegenden Datenstrukturen am besten abbilden. So entsteht ein belastbares prädiktives Modell, das Eingabeparameter der Baugruppe mit den zu erwartenden Fertigungszeiten verknüpft.

Um die Leistungsfähigkeit des Modells zu bewerten, wurden Lernkurven untersucht, die den Einfluss der Trainingsdatenmenge auf die Güte der Vorhersage darstellen (siehe Abbildung 4). Die Ergebnisse zeigten erwartungsgemäß, dass mit wachsender Datenbasis sowohl Genauigkeit als auch Stabilität der Prognosen deutlich zunehmen. Bei zunehmender Datenmenge ist eine weitere Verbesserung des Ergebnisses zu erwarten. Zusätzlich wurde die Skalierbarkeit, d. h. die Auswirkung der Anlerndauer auf die Prognosequalität, analysiert. Es konnte festgestellt werden, dass die Trainingszeit moderat ansteigt, sodass die Modelle auch bei größeren Datensätzen praxistauglich bleiben.



Ne/118435 © IFW

Abbildung 4: Lernkurven eines Regressionsmodells zur Prognose der Rückmelde einer Baugruppenklasse

Eine Besonderheit des entwickelten Ansatzes ist die Quantifizierung der Prognoseunsicherheit. Neben der Zeitvorhersage wird ein weiteres Modell eingesetzt, das den erwarteten Prognosefehler durch den Residualfehler des vorherigen Modells prognostiziert. Daraus wird ein Konfidenzintervall abgeleitet. So liegt nicht nur ein Prognosewert vor, sondern auch eine Einschätzung darüber, wie verlässlich die Prognose in einem konkreten Fall ist. Dies schafft mehr Transparenz und bietet einen entscheidenden Vorteil für die Angebotserstellung sowie die Bewertung von Risiken in der Arbeitsvorbereitung.

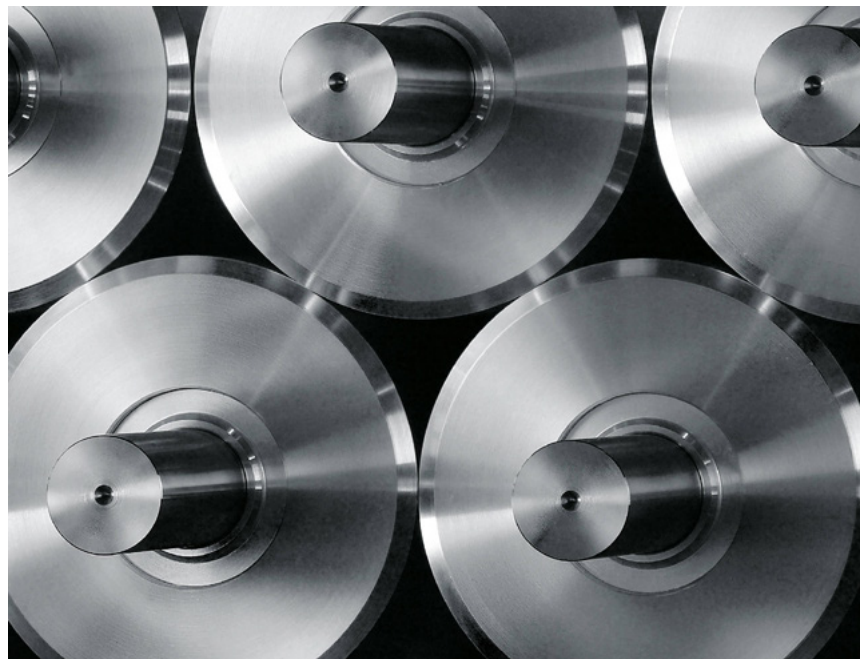
Durch die Kombination der prototypisch getesteten Ansätze zur datengetriebenen Zeitprognose, zur Unsicherheitsschätzung und zur strukturierten Verarbeitung technischer Eingangsparameter wird deutlich, wie ein zukünftiges Softwaretool zur Unterstützung der Angebotskalkulation aussehen könnte. Ein solches KI-Tool bietet zudem erhebliches Potenzial, die Bewertung neuer Kundenanfragen zu beschleunigen.

## Nutzen für den Mittelstand

Die Unterstützung der Angebotskalkulation durch KI-basierte Analyseverfahren bietet dem Mittelstand Vorteile. Durch die automatisierte Verarbeitung technischer Informationen und die datengetriebene Abschätzung von Fertigungszeiten kann der Aufwand für die Bewertung neuer Anfragen reduziert werden. Dies ergänzt bestehende Kalkulationswerkzeuge, in denen Informationen bisher vollständig manuell eingegeben und bewertet werden müssen, und erleichtert die schnelle Einschätzung der Herstellkosten komplexer Produktvarianten.

Der zusätzliche Grad an Objektivität, der durch die Nutzung historischer Fertigungsdaten und die integrierte Unsicherheitsschätzung erreicht wird, führt zu einer transparenteren Entscheidungsgrundlage im Angebotsprozess. Limitiert wird der Ansatz jedoch durch Änderungen der Produktionstechnologien, da diese bei identischen Bauteilen zu kürzeren oder längeren Bearbeitungszeiten führen können.

Darüber hinaus kann eine datenbasierte Unterstützung zu einer effizienteren Ressourcennutzung beitragen und Fachkräfte entlasten, indem sie wiederkehrende Analyseschritte reduziert und dadurch Zeit für komplexere Aufgaben schafft. Langfristig stärkt dies die Wettbewerbsfähigkeit kleiner und mittlerer Unternehmen, die in dynamischen Märkten zunehmend gefordert sind, Angebote schnell, konsistent und gut begründet zu erstellen. Damit diese Vorteile erreicht werden, sind zukünftig weitere Anstrengungen erforderlich, um die prototypischen Untersuchungen dieses Vorhabens auf industrielle Anwendungen und Datenbasen zu übertragen. Insbesondere die zunehmende Komplexität und Variantenvielfalt technischer Zeichnungen stellen weiterhin Hürden für die Anwendung der erzielten Ergebnisse in der industriellen Praxis dar.



### Autor

Marcus Nein

Wissenschaftlicher Mitarbeiter am Institut für Fertigungstechnik und Werkzeugmaschinen und Experte für KI im Mittelstand-Digital Zentrum Hannover

# Impressum

Schriftenreihe des Mittelstand-Digital Zentrums Hannover  
Zukunft.Digital – Digitalisierung von der Idee zur Umsetzung  
Ausgabe 01/2026

Institut für Fertigungstechnik und Werkzeugmaschinen  
der Leibniz Universität Hannover  
An der Universität 2, 30823 Garbsen

Herausgeber: Prof. Dr.-Ing. Berend Denkena  
Redaktion: Dr.-Ing. Michael Rehe  
Satz und Layout: Laura-Sophie Oetter

**Bibliografische Information der Deutschen Nationalbibliothek**  
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind online unter <https://www.dnb.de> abrufbar.

Dieses Werk ist urheberrechtlich geschützt. Alle Rechte, auch das des Nachdruckes, der Wiedergabe, der Speicherung in Datenverarbeitungsanlagen und der Übersetzung des vollständigen Werkes oder von Teilen davon, sind vorbehalten.

© TEWISS – Technik und Wissen GmbH, April 2026  
An der Universität 2, 30823 Garbsen  
Telefon: 0511 762 19434. Mail: [info@tewiss-verlag.de](mailto:info@tewiss-verlag.de)  
[www.tewiss-verlag.de](http://www.tewiss-verlag.de)

**Konsortialführung**  
des Mittelstand-Digital Zentrums Hannover:



## Bildnachweis

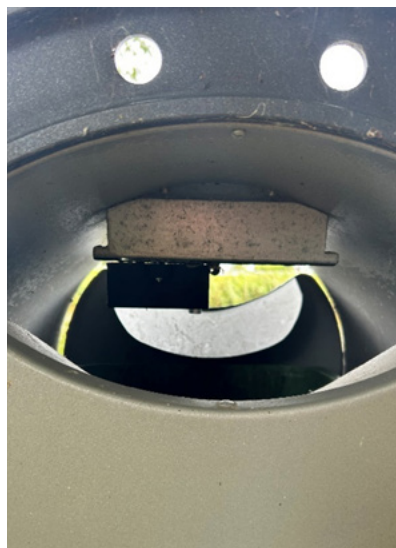
- Titel: somram/stock.adobe.com, generiert mit KI
- Seite 05, 39–40, 41 oben: Institut für Fertigungstechnik und Werkzeugmaschinen (IFW), Leibniz Universität Hannover
- Seite 06: Rawpixel/elements.envato.com
- Seite 09 Icons: Freepik.com, flaticon.com
- Seite 11: rawf8/elements.envato.com
- Seite 12: Cubos Service GmbH
- Seite 14–15, 24, 31–33: IPH - Institut für Integrierte Produktion Hannover gGmbH
- Seite 16: UNDP, Public domain, via Wikimedia Commons
- Seite 17: Generiert mit Adobe Firefly
- Seite 18: SkyNextphoto/elements.envato.com
- Seite 20: Mockup-Vorlage von Aozora/stock.adobe.com
- Seite 21, Rückseite unten: TU Braunschweig
- Seite 22: Naiem/stock.adobe.com, generiert mit KI
- Seite 25: Muhammad Adnan/stock.adobe.com, generiert mit KI
- Seite 26-27: elxeneize/stock.adobe.com
- Seite 29, Rückseite oben: SOMKID/stock.adobe.com
- Seite 30: mindandi/freepik.com
- Seite 34: generiert mit KI
- Seite 35, 36 unten, Rückseite links mittig: MASA-Group GmbH
- Seite 37: galitsin/stock.adobe.com
- Seite 38: stevepb/pixabay.com
- Seite 41 unten, Rückseite rechts mittig: Inometa GmbH

ISBN 978-3-69030-178-7 (Print)  
ISBN 978-3-69030-191-6 (Online)

ISSN 2941-3044 (Print)  
ISSN 2944-7674 (Online)



Zukunft.Digital **online**  
[digitalzentrum-hannover.de/downloads](https://digitalzentrum-hannover.de/downloads)



Mittelstand-Digital  
**Zentrum**  
**Hannover**