

Gefördert durch:



Bundesministerium
der Justiz und
für Verbraucherschutz



Bundesanstalt für
Landwirtschaft und Ernährung

aufgrund eines Beschlusses
des Deutschen Bundestages

Laufzeit des Vorhabens: 01.01.2020 - 31.12.2021

Wissenschaftlicher Abschlussbericht Smart_HEC

Vorhabentitel: „Der HeizCheck für alle mit KI –
Empowerment durch eine smarte Bewertung der
Heizkostenabrechnung“

Förderkennzeichen: 28V2304A19, 28V2304B19, 28V2304C19,
28V2304D19)

von

Yvonne Scheurer, Oliver Albrecht
co2online gGmbH, Berlin

Dr. Jana Diels, Marlene Münsch, Prof. Dr. Christian Thorun
ConPolicy GmbH, Berlin

Markus Otto
SEnerCon GmbH, Berlin

Julian Balling, Frank Binder, J. Nathanael Philipp, Robert Sachunsky,
Dr. Andreas Niekler, Prof. Dr. Gerhard Heyer
Universität Leipzig

Berlin, Leipzig

01.02.2022

Inhaltsverzeichnis

Abbildungsverzeichnis	4
Tabellenverzeichnis	5
Abkürzungen	6
1. Einführung.....	8
1.1. Ausgangslagen.....	8
1.2. Zielsetzung	8
1.3. Verbundpartner.....	10
2. Forschungsentwicklung	12
2.1. KI-Entwicklung: Maschinelles Lernen für die Layout-Analyse, Klassifikation und Datenextraktion aus HKAs.....	12
2.1.1. Anforderungsanalyse	12
2.1.2. Werkzeug- und Technologieauswahl.....	15
2.1.3. Systemarchitektur und Datenfluss	20
2.1.4. Eingeflossene Forschungs- und Entwicklungsleistung	25
2.1.5. Exemplarische Herausforderungen und Lösungsstrategien	27
2.1.6. Eingesetzte Ressourcen und Ressourcenbedarf bei Anpassungen.....	30
2.1.7. Erzielte Ergebnisse	31
2.1.8. Zwischenfazit und weiterer Bedarf an Forschung und Entwicklung	31
2.2. Verbraucherforschung zur Nutzerzentrierung, Verständlich- und Anwendungsfreundlichkeit der Anwendung	34
2.2.1. Online-Fokusgruppen zur Erhebung der allgemeinen Nutzererwartungen an die Anwendung	36
2.2.2. Quantitative Kurzumfrage zur quantitativen Erhärtung ausgewählter Fokusgruppenergebnisse.....	37
2.2.3. Quantitative Online-Umfrage zum Thema Datenschutz.....	38
2.2.4. UX-Test zur Navigation	39
2.2.5. Fokusgruppen zum Thema Handlungsmotivation und Nudging auf der Ergebnisseite	40
2.2.6. Kreativ-Workshop zur Anreicherung der Start- und Ergebnisseite mit verhaltenswissenschaftlichen Erkenntnissen	42
2.2.7. UX-Test zum Thema Conversion	44
2.2.8. Quantitative Online-Umfrage zur Umsetzungswahrscheinlichkeit von Handlungsempfehlungen	45
2.2.9. Übergeordnete Bewertung des nutzerzentrierten Forschungsansatzes.....	46

3. Anwendungsentwicklung	48
3.1. Berechnung und Bewertung	48
3.1.1. Inhaltliche Betrachtung einer HKA.....	48
3.1.2. Datenbasis der Heizkostenanalyse	54
3.1.3. Technische Umsetzung.....	56
3.1.4. Bewertungsgrundlagen	58
3.1.5. Zusammenfassung: Bewertung der HKA im Rahmen des Prototyps	63
3.2. Entwicklung des Prototyps des smarten HeizChecks.....	63
3.2.1. Anforderungsanalyse und Konzeption.....	63
3.2.2. Prototyping.....	66
3.2.3. Umsetzung.....	74
3.2.4. Beta-Phase	78
3.2.5. Zwischenfazit: Ergebnisse der nutzerzentrierten Prototypenentwicklung.....	80
4. Verwertungskonzept.....	82
4.1. Einleitung.....	82
4.2. Zielgruppen	82
4.3. Gemeinwohlorientiertes Geschäftsmodell.....	83
4.3.1. Betriebskosten	83
4.3.2. Refinanzierung/Vergütungsvarianten	84
4.4. Platzierung der Anwendung und Reichweitenaufbau	88
4.4.1. Integration in bestehende Beratungsplattformen	88
4.4.2. Steigerung der Sichtbarkeit und Multiplikator*innen	90
4.5. Konzept zur Wirkungsmessung	92
4.6. Bestehende Verwertungsrisiken	95
5. Fazit und Empfehlungen an die Politik.....	96
5.1. Entwicklung des Prototyps zur smarten Heizkostenanalyse mit Hilfe von KI.....	96
5.2. Nutzerzentrierung bei KI-Entwicklungen	98
5.3. Strategische Steuerung des Einsatzes von KI-Verfahren.....	99
5.4. Weiterentwicklung und Digitalisierung der Verbrauchsinformationen der HKA.....	100
6. Literaturverzeichnis	104

Abbildungsverzeichnis

Abbildung 1: Ablaufschema smarte Bewertung der HKA.....	9
Abbildung 2: Trainingsfehler verschiedener Trainings derselben Modelle wie beim Validierungsfehler.	21
Abbildung 3: Validierungsfehler verschiedener Trainings derselben Modelle wie beim Trainingsfehler.	22
Abbildung 4: Genauigkeit (Precision) und Vollständigkeit (Recall) für die Kennwertextraktion gemittelt über alle MDL – zu drei exemplarischen Entwicklungsständen der KI-Komponenten.....	23
Abbildung 5: Wöchentliche Entwicklung der durchschnittlichen Anzahl an automatisch erkannten Kennwerten pro durchgeführter HKA-Analyse.	24
Abbildung 6: Ausschnitt einer HKA mit Beispielen für die Anordnung der Kontextklassen „Energiekosten“, „Betrag in EUR“, „Brennstoffkosten“ (jeweils weiß hinterlegt). Sonstige Inhalte sind transparent hinterlegt.	27
Abbildung 7: Fehler bei der Binarisierung einer HKA.....	28
Abbildung 8: Visualisierung der Annotationen für den Kennwert „Kosten des Energieträgers“: Kontextklassen (türkis), Kennwertklasse (rot), für diesen Kennwert nicht relevanter Text (lila)	29
Abbildung 9: Exemplarische Darstellung des Nutzereinbindungskonzept anhand des Projekt-Zeitstrahls	35
Abbildung 10: Bandtacho zur Einordnung des Gebäudeverbrauchs (EVKW).....	61
Abbildung 11: Quantile Heizspiegel für Verbrauchsklassenermittlung	62
Abbildung 12: User Story Map zur Veranschaulichung der Anwendungsschritte des smarten HeizChecks	64
Abbildung 13: Umsetzung verschiedener Interaktionsstile in Form von Mockups für die Fokusgruppengespräche	65
Abbildung 14: Entwürfe des Anwendungsschritts „Erfassen“ vor und nach den UX-Tests...	67
Abbildung 15: Entwürfe des Anwendungsschritts „Validieren“ vor und nach den UX-Tests.	67
Abbildung 16: Entwürfe des Anwendungsschritts „Ergänzen“ vor und nach den UX-Tests .	68
Abbildung 17: Erster Entwurf der Ergebnisseite mit mehrstufiger Informationsarchitektur ...	69
Abbildung 18: Neu-Entwurf der Ergebnisseite mit Erklärtexten und allen Inhalten auf einer Seite.....	70
Abbildung 19: Teaser der Handlungsempfehlungen in Listenansicht mit integrierten Nudges (Einsparungen und Klimaeffekt)	71

Abbildung 20: Detailansicht der Handlungsempfehlungen mit integrierten Nudges (Kosten-Nutzen-Analyse, konkrete Einsparpotenziale in Euro und CO ₂ -Emissionen, Übersetzung der CO ₂ -Emissionen in verständliche Einheiten)	72
Abbildung 21: Handlungsempfehlungen mit Weiterleitung zu Verbraucherschutzorganisationen	72
Abbildung 22: Zweistufiges Intro-Overlay	74
Abbildung 23: Verknüpfung zwischen Desktop und Smartphone/Tablet, um Bilder der HKA hochzuladen	76
Abbildung 24: Hilfestellungen für die Werteeingabe im Schritt „Ergänzen“	77
Abbildung 25: Letzte Überarbeitung des Schritts „Erfassen“ durch Umsetzung der Dialog-Abfrage	79
Abbildung 26 Feedback zur Beta-Version, 43 Antworten	79
Abbildung 27: Schwankungen der Conversion-Rate (lila) und Anfragen pro Woche (gelb) .	81
Abbildung 28: Beispielhafte Integration des Spendenaufrufs in die Online-Ratgeber von co2online	86
Abbildung 29: Integration des Förderlogos des BMJV zur Erhöhung der Glaubwürdigkeit der Anwendung	89

Tabellenverzeichnis

Tabelle 1: Funktionale Anforderungen an die KI-Komponente	14
Tabelle 2: Nicht-Funktionale Anforderungen an die KI-Komponente	15
Tabelle 3: Programmbibliotheken und deren genutzte Lizenzen für die KI-Komponente (Prototyp)	19
Tabelle 4: Auflistung der nötigen Kennwerte für die Heizkostenanalyse.....	56
Tabelle 5: Vorläufige Betriebskostenaufstellung des smarten HeizChecks.....	84

Abkürzungen

API	„Application Programming Interface“ – Programmierschnittstelle
BMJV	Bundesministerium der Justiz und für Verbraucherschutz
COCO	„Common Objects in Context“ – Bezeichnung für ein Datenformat zur Verarbeitung von Objektkoordinaten und Objektklassen auf Bilddaten
CO ₂	Kohlendioxid
CPU	„Central Processing Unit“ – Hauptprozessor (eines Computers oder Serverknotens), der üblicherweise nicht für die Berechnungen zum maschinellen Lernen ausgelegt ist (siehe auch: GPU)
CSV	„Comma-separated Value“ – ein strukturiertes tabellarisches Datenformat, in dem Datenwerte durch Kommata getrennt aufgelistet werden
EVKW	Energieverbrauchskennwert
GPU	„Graphical Processing Unit“ – dezidiert Grafikprozessor, welcher üblicherweise für die Berechnungen im maschinellen Lernen eingesetzt wird, da er dafür besser geeignet ist als eine CPU
GT	„Ground Truth“ – bezeichnet die Grundwahrheit für das maschinelle Lernen von Klassifikationsproblemen, die zumeist vorab manuell erstellt, d. h., annotiert werden muss und als Grundlage für Training und Evaluation der Neuronalen Netze dient
HeizkostenV	Heizkostenverordnung
HKA	Heizkostenabrechnung
HPC	„High Performance Computing“ – Hochleistungsrechnen
JSON	„JavaScript object notation“ – ein universell nutzbares strukturiertes Datenformat nach dem Schlüssel-Wert-Prinzip
KI	Künstliche Intelligenz – gemeint sind hier Verfahren der „schwachen künstlichen Intelligenz“, also beispielsweise maschinelles Lernen
kWh	Kilowattstunde
mAP	„Mean average precision“ – mittlere durchschnittliche Genauigkeit, ein Maß für die Wiedererkennung von Klassen
mAR	„Mean average recall“ – mittlere durchschnittliche Vollständigkeit, ein Maß für die Wiedererkennung von Klassen
Mask R-CNN	Spezielle Architektur eines Neuronalen Netzes für die Objekterkennung auf Bilddaten (He et al. 2017), siehe auch: R-CNN.
MDL	Messdienstleister
OCR	„Optical Character Recognition“ – Automatisches Erkennen von Text und Schriftzeichen auf Bilddaten (Scans, Fotos)

OCR-D	Eine von der Deutschen Forschungsgemeinschaft langfristig unterstützte „Förderinitiative zur Weiterentwicklung von Verfahren der Optical Character Recognition (OCR)“; zugleich ein Technologierahmenwerk für die automatische Layoutanalyse und Inhaltserschließung von Dokumentensammlungen.
PDF	„Portable Document Format“ – ein standardisiertes und weit verbreitetes Datenformat für elektronische Dokumente, sowohl für Preprints (direkt digital generierte Dokumente) als auch für Scans (gedruckte und wieder eingescannte Dokumente)
R-CNN	„Region-based convolutional neural networks“ (regions-bezogene Faltungsnetze) – eine Architekturfamilie neuronaler Netzwerke, die zur Objekterkennung auf Bilddaten genutzt wird, s. a.: Mask R-CNN
Smart_HEC	Kurzform des Vorhabentitels „Der HeizCheck für alle mit KI – Empowerment durch eine smarte Bewertung der Heizkostenabrechnung“
UX	User Experience (Nutzererfahrung) – beschreibt die Erlebnisqualität bei der Anwendung einer Software oder einer App aus Sicht der Anwender*innen
XML	„Extensible Markup Language“ eine Auszeichnungssprache für Dokumente, ermöglicht strukturierten Zugriff auf Dokumentenelemente

1. Einführung

In dem Forschungsprojekt „Der HeizCheck für alle mit KI – Empowerment durch eine smarte Bewertung der Heizkostenabrechnung“ (Smart_HEC) entwickelten die Universität Leipzig, co2online, SEnerCon und ConPolicy über den Projektzeitraum vom 01.01.2020 - 31.12.2021 gemeinsam **eine prototypische Anwendung zur automatisierten Erkennung und Bewertung von Heizkostenabrechnungen (HKA) auf Basis von Künstlicher Intelligenz (KI)**. Das Projekt wurde vom Bundesministerium der Justiz und für Verbraucherschutz (BMJV) im Rahmen der Förderung von Vorhaben zur verbraucherbezogenen Forschung und Entwicklung zu „Anwendungen künstlicher Intelligenz zur Unterstützung des Verbraucheralltags (consumer enabling technologies)“ im Rahmen des Programms zur Innovationsförderung im Verbraucherschutz in Recht und Wirtschaft gefördert.

Im Folgenden sollen die wissenschaftlichen Erkenntnisse skizziert werden, die im Projektzeitraum im Bereich des maschinellen Lernens (KI) (s. 2.1.) sowie der Verbraucherforschung (s. 2.2.) gewonnen wurden. Es soll ferner gezeigt werden, wie die wissenschaftlichen Erkenntnisse in der Anwendungsentwicklung Berücksichtigung fanden (s. 3.2.) und auf welchen Grundlagen der Rechenkern zur Bewertung der Heizkostenabrechnung fußt (s. 3.1.). Zum Abschluss werden Verwertungsperspektiven der Projektergebnisse im Rahmen eines Verwertungskonzepts präsentiert, das die Verbundpartner gemeinsam zum Ende der Projektlaufzeit erarbeitet haben (s. 4.).

1.1. Ausgangslagen

Jährlich erhalten rund **16 Millionen Haushalte in Deutschland eine HKA**. Theoretisch könnten Verbraucher*innen in der HKA alle notwendigen Informationen finden, um Rückschlüsse auf die Preise der eingekauften Energie und Dienstleistungen, den energetischen Zustand des Gebäudes und das eigene Heizverhalten zu ziehen. In ihrer jetzigen Form sind die HKA jedoch in der Realität nicht verständlich aufbereitet, da jeder Messdienstleister (MDL) unterschiedliche Formate und Begrifflichkeiten verwendet. Das Resultat: **Intransparenz und Orientierungslosigkeit für Verbraucher*innen**. Fehler werden nicht erkannt, Energie und Dienstleistungen werden überteuert eingekauft, Anreize zum Energiesparen nicht gesetzt. Dies kostet Verbraucher*innen Milliarden Euro jährlich und schadet dem Klima durch erhebliche CO₂-Emissionen, verursacht durch hohe Heizenergieverbräuche der Gebäude.

1.2. Zielsetzung

Im Rahmen des Vorhabens Smart_HEC sollte eine prototypische Anwendung entstehen, die klassische HKA durch den Einsatz von KI in verständliche und für Verbraucher*innen

transparente, nachvollziehbare Aussagen und Bewertungen – in **smarte Heizkosten-bewertungen** – verwandelt. Die Verbundpartner haben zur Projektbeginn folgendes Ablaufschema definiert: Die Anwendung soll

- mit Hilfe **optischer Zeichenerkennung (OCR)** die HKA auslesen,
- mittels **maschinellen Lernens (KI)** die relevanten Informationen extrahieren und durch den Abgleich mit Datenbanken die für Verbraucher*innen relevanten Informationen aufbereiten,
- diese Informationen in verständlicher und transparenter Weise durch ein **virtuelles Assistenzsystem** ausgeben und
- personalisierte **Handlungsempfehlungen** aussprechen.

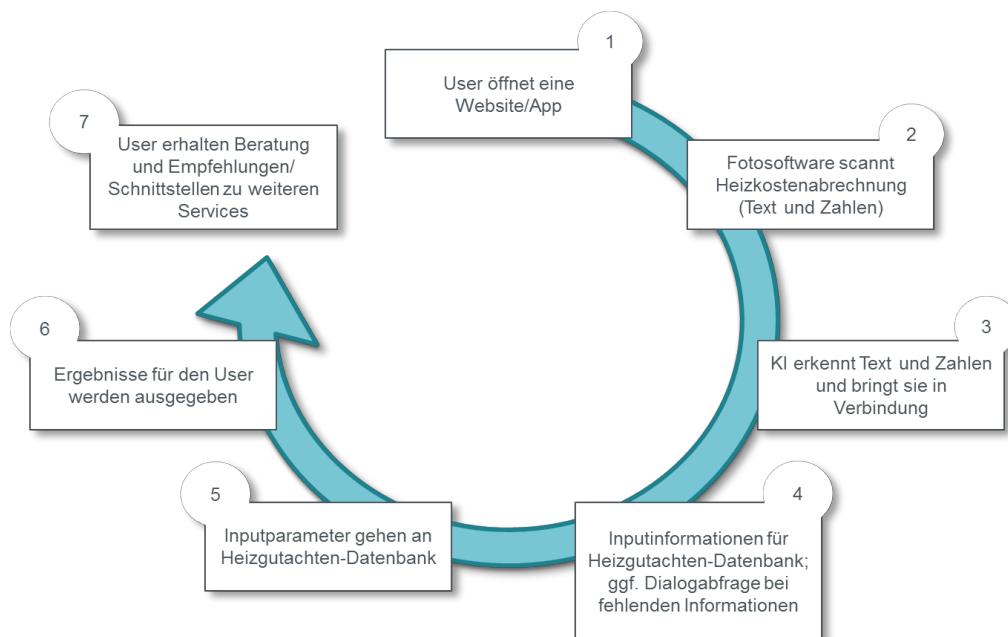


Abbildung 1: Ablaufschema smarte Bewertung der HKA

Um dieses Ziel zu erreichen, wurden die folgenden **Arbeitsziele** und dazugehörigen **Forschungsfragen** adressiert:

Technische Arbeitsziele:

- Wie müssen Standardverfahren der OCR weiterentwickelt werden, um Texte aus HKA, die durch viele Tabellen, Grafiken und Sonderzeichen charakterisiert sind, zu extrahieren?
- Wie ist maschinelles Lernen für die Erkennung von HKA so einzusetzen, dass die Verfahren nicht nur HKA eines Anbieters verstehen, sondern die große Vielzahl von HKA abdecken?
- Wie müssen die Lernverfahren ausgestaltet sein, um auch zukünftig, wenn Anbieter HKA verändern, anwendbar zu bleiben?

Nutzerzentrierung, Verständlichkeit und Anwendungsfreundlichkeit:

- Wie muss das virtuelle Assistenzsystem textlich, grafisch und in der Nutzerinteraktion ausgestaltet sein, damit Verbraucher*innen durch den Einsatz von Smart_HECs einfacher, schneller und nutzerfreundlicher an relevante Informationen kommen, um ihre Kosten und ihren Verbrauch zu reduzieren sowie Fehler zu erkennen und basierend auf diesen Informationen zu handeln?

Niedrigschwelliger Zugang:

- Wie kann gewährleistet werden, dass das Angebot von möglichst vielen Verbraucher*innen genutzt wird?
- Wie kann das Angebot gerade auch für verletzbare Verbrauchergruppen attraktiv gemacht werden?

Vorbildlicher Datenschutz:

- Wie kann sichergestellt werden, dass den Nutzenden der Smart_HECs zwar personalisierte Informationen zur Verfügung gestellt werden, gleichzeitig aber ihr Recht auf informationelle Selbstbestimmung gewahrt bleibt?

Glaubwürdiges und erfolgsversprechendes Verwertungskonzept:

- Wie könnte der Prototyp in einem gemeinwohlorientierten Geschäftsmodell verwertet werden?
- Welche Akteure sollten eingebunden werden, um die Smart_HECs nach der Förderphase erfolgreich in den Markt und zu den Verbraucher*innen zu bringen? Welche Vergütungsvarianten weisen die größte Akzeptanz auf?

1.3. Verbundpartner

Das Verbundvorhaben wurde von vier Verbundpartnern umgesetzt:



Die gemeinnützige co2online GmbH setzt sich dafür ein, den klimaschädlichen CO₂-Ausstoß zu senken. Mit onlinebasierten Informationskampagnen, interaktiven EnergiesparChecks und Praxistests motiviert co2online Verbraucher*innen, mit aktivem Klimaschutz Geld zu sparen.

Das Forschungsvorhaben wurde von co2online geleitet. Die Partnerin verantwortet die konzeptionelle Entwicklung und Programmierung des neuen Heizkosten-Rechners.



Die ConPolicy GmbH ist ein 2010 gegründetes Forschungs- und Beratungsunternehmen. Seine Dienstleistungen richten sich an öffentliche Einrichtungen, Verbände und Unternehmen, die sich mit verbraucherpolitischen Fragen befassen.

ConPolicy stellte sicher, dass zu jedem Zeitpunkt der Produktentwicklung das Feedback der Nutzer*innen und Anstöße aus der Verbraucherforschung einbezogen wurden.



Die SEnerCon (SigmaEnergyConsulting) GmbH berät seit 1992 erfolgreich Hauseigentümer*innen und Mieter*innen bei der Senkung des Heizenergie- und Stromverbrauchs mit Hilfe von webbasierten Beratungsinstrumenten.

Die Fachexpert*innen waren für die inhaltliche Analyse der HKA verantwortlich. Sie sorgten für transparente Verbrauchsinformationen und individuelle Beratungsangebote.



UNIVERSITÄT
LEIPZIG

Die Universität Leipzig wurde 1409 gegründet und gehört zu den forschungsstarken medizinführenden Universitäten in Deutschland.

Das Forschungsteam entwickelte ein innovatives Extraktionsverfahren basierend auf Machine Learning (KI), um HKAs unterschiedlicher MDLs automatisch erkennen zu können.

Das Vorhaben wurde zudem durch **drei Praxispartner** begleitet. Der Deutscher Mieterbund, FINANZTIP und die Verbraucherzentrale Nordrhein-Westfalen begleiteten die Anwendungsentwicklung durch fachliches Feedback, Hinweise aus Sicht ihrer Zielgruppe sowie die Integration ihrer Dienstleistungsangebote als weiterführende Services in den Prototypen.

2. Forschungsentwicklung

Für die Entwicklung eines smarten Prototyps für die Heizkostenanalyse mit Hilfe von KI wurde ein Neuronales Netz für die Extraktion der Verbrauchsinformationen aus der HKA entwickelt und trainiert. Die Anwendungsentwicklung wurde außerdem durch Methoden der Verbraucherforschung begleitet. Die jeweiligen Forschungsergebnisse werden nachstehend präsentiert.

2.1. KI-Entwicklung: Maschinelles Lernen für die Layout-Analyse, Klassifikation und Datenextraktion aus HKAs

Im Rahmen dieses Projekts wurde eine **Klassifikations-Pipeline** entwickelt, um verschiedene Werte einer HKA automatisiert auszulesen und **34 unterschiedliche Kennwerte zu klassifizieren**. Dabei kommen gängige Verfahren der Bildverarbeitung (Binarisierung, Zuschchnitt, Begradigung und Rauschentfernung), der automatischen Textklassifizierung (OCR) sowie der Künstlichen Intelligenz (Layout-Erkennung) zum Einsatz. Ziel der Verarbeitungspipeline ist die Erkennung, Klassifikation und Transformation der gegebenen Kennwerte vom unstrukturierten Eingabeformat (Bilder) in ein strukturiertes Datenausgabeformat (PageXML). Durch das gewählte Konzept von Kontextklassen und Kennwertklassen können tabellarische bzw. zueinander in Beziehung stehende Strukturen ausgelesen werden. Die Kontextklassen bilden dabei die relevanten Regionen wie Spalten und Reihen.

Das Prinzip ähnelt dem eines Fadenkreuzes: Ein Kennwert wird als solcher klassifiziert und ausgelesen, wenn dieser senkrecht unter bzw. in der Nähe einer mit einem Kontext markierten Spalte und horizontal bzw. in der Nähe einer mit einem Kontext markierten Reihe liegt.

2.1.1. Anforderungsanalyse

Funktionale Anforderungen

Die Informationsextraktion soll robust über unterschiedliche MDL mit unterschiedlichen HKA-Layouts realisiert werden. Insgesamt wurden sechs große unterschiedliche MDL berücksichtigt: Techem, Brunata, Ista, Minol, BFW und Kalo. Außerdem wurde ein weiterer MDL „Sons-tige“ eingeführt, in dem alle HKAs konsolidiert werden, die von anderen MDL erstellt werden.

ID	Anforderung	Bedeutung/Motivation	Anforderung erfüllt?
FA1	Die Verarbeitungspipeline soll PDFs und Bilder verarbeiten können.	Als Eingabe kann der/die Nutzer*in die HKA als PDF hochladen oder als Bild abfotografieren.	Ja
FA2	Die zu bearbeitenden Bilder sollen binarisiert werden (schwarz/weiß).	Um Geschwindigkeits- und Speichervorteile zu erhalten, werden die Bilder in schwarz-weiß verarbeitet.	Ja
FA3	Die zu bearbeitenden Bilder sollen entzerrt werden.	Von Nutzer*innen schief fotografierte HKAs müssen für die OCR entzerrt werden.	Ja
FA4	Die zu bearbeitenden Bilder sollen auf den relevanten Bereich zugeschnitten werden.	Bilder müssen auf den relevanten Teil zugeschnitten werden, um störende Bereiche (Hintergrund) zu vermeiden.	Ja
FA5	Die zu bearbeitenden Bilder sollen auf Zeilenebene segmentiert werden.	Die Segmentierung auf Zeilenebene führt zu besseren OCR-Ergebnissen.	Ja
FA6	Der Text der segmentierten Bilder soll mit OCR extrahiert werden.	Extraktion des Textes	Ja
FA7	Die für die Kennwert-Klassifikation genutzten Kontexte sollen extrahiert werden (textuelle Klassifikation).	Als Eingabe zur Layoutklassifikation werden Kontexte benötigt. Diese sollen automatisiert extrahiert und gekennzeichnet werden.	Eingeschränkt ¹
FA8	Neuronale visuelle Layout-Klassifikation der Textsegmente in 34 Kennwerte.	Tatsächliche positionelle Erkennung (Koordinaten) der gegebenen Kennwerte auf dem Bild	Eingeschränkt ²

¹ Die Kontexte für Kennwertklassen werden textuell klassifiziert. Taucht der Kontextbegriff häufiger im Bild auf, werden für den gesuchten Kennwert mehr Kontexte klassifiziert als nötig.

² Im Durchschnitt werden 14 der 34 Kennwerte auf einem dem Netz bekannten Layout gefunden. Die 34 Kennwertklassen verteilen sich über mehrere Seiten einer HKA.

ID	Anforderung	Bedeutung/Motivation	Anforderung erfüllt?
FA9	Gefundene Kennwertregionen müssen auf das komplette Schriftsegment ausgedehnt werden.	Vorhergesagte Boxen/Masken müssen auf das gefundene Segment erweitert werden, sodass Kennwerte nicht abgeschnitten werden.	Ja
FA10	Extraktion der erkannten Kennwerte.	Erneute OCR auf gefundenen Regionen zum Auslesen der Kennwerte unter Berücksichtigung der erwarteten Zeichen	Ja
FA11	Gefundene Kennwerte sollen mit semantischen Regeln nachverarbeitet und validiert werden.	Am Ende der Verarbeitung soll eine Plausibilitätsprüfung stattfinden: Wird als Kennwert ein Datum erwartet, sollte der gefundene Kennwert auch das entsprechende Format haben.	Ja
FA12	Der Datenaustausch und die Kommunikation zwischen Nutzerinterface und Backend erfolgt mit Hilfe eines Webservices.	Die Bilder bzw. PageXML-Dateien werden über Webschnittstellen transferiert.	Ja
FA13	Es sollen 6+1 MDL verarbeitet werden.	Es sind Daten für sechs große MDL vorhanden, außerdem sollen alle anderen MDL als „Sonstige“ zusammengefasst werden.	Ja
FA14	Neuronale visuelle Layout-Klassifikation soll den MDL inkl. 34 Kennwerte erkennen.	Es sind Daten für die 6 größten MDL vorhanden, außerdem sollen alle anderen MDL als Sonstige zusammengefasst werden.	Eingeschränkt ³

Tabelle 1: Funktionale Anforderungen an die KI-Komponente

³ Die Konsolidierung der unterschiedlichen MDL-Layouts und der Erkennung der 34 Kennwerte in einem Neuronalen Netz, statt in je einem Netz pro MDL, wurde implementiert, trainiert und getestet. Zwar ist es möglich, den MDL sicher zu erkennen (innerhalb der definierten sieben Klassen für die sechs großen MDL und „Sonstige“), allerdings kann das integrierte Netz die Kennwertregionen nicht mehr ausreichend gut identifizieren, wodurch die Kennwerte nicht mit ausreichender Güte extrahiert werden können. Dieser Ansatz wurde daher wieder verworfen.

Nicht-funktionale Anforderungen

ID	Anforderung	Bedeutung/ Motivation	Anforderung erfüllt?
NFA1	Die Verarbeitung soll in Echtzeit stattfinden.	Der/die Nutzer*in soll innerhalb von zwei bis drei Minuten das Ergebnis der Bildanalyse erhalten.	Eingeschränkt ⁴
NFA2	Zur Verarbeitung der Dokumente soll Open-Source Software genutzt werden.	Open-Source Software bietet besseren Zugang zu Innovation und hat weniger Fehler.	Ja
NFA3	Die Umsetzung soll so erfolgen, dass möglichst wenige Rechenressourcen benötigt werden.	Zusätzliche Kosten für leistungsstarke Hardware sollen vermieden werden.	Ja
NFA4	Das System muss auch mit geschwärzten/überklebten Bereichen auf der HKA umgehen können.	Auf Grund des Datenschutzes anonymisieren Nutzer*innen möglicherweise ihre personenbezogenen Daten. Dies soll die Klassifikation des Systems nicht beeinflussen.	Ja

Tabelle 2: Nicht-Funktionale Anforderungen an die KI-Komponente

2.1.2. Werkzeug- und Technologieauswahl

Zur Texterkennung (OCR) wird dabei auf das **Open-Source Projekt OCR-D**⁵ (Neudecker et al., 2019; Engl, 2020) zurückgegriffen. OCR-D basiert auf einem **modularen Ansatz**, d. h., OCR-D stellt Schnittstellen für bereits existierende Werkzeuge zur Verfügung, bzw. entwickelt eigene, so dass jedes Modul, genannt Prozessor, in einem Workflow⁶ nacheinander gereiht ausgeführt wird und auf die Ergebnisse des vorherigen Prozessschrittes zurückgreifen kann. Die einzelnen Prozessoren können dabei entweder nacheinander, über die Kommandozeile gestartet oder in einem Makefile zusammengefasst und automatisch ausgeführt werden. Zu Beginn eines Workflows wird neben dem Bild eine Metsdatei benötigt. Die einzelnen Prozessoren erstellen entweder neue Bildderivate oder eine PageXML-Datei.

Zur leichten Verwendung stellt OCR-D neben den Tools selbst auch einen **Docker-Container** mit allen notwendigen Abhängigkeiten zur Verfügung.

⁴ Die Verarbeitungsdauer ist abhängig von der Seitenzahl der HKA sowie von der Bildgröße bzw. Auflösung der Scans oder Fotos. Eine Einhaltung der genannten Zeitgrenzen kann im Prototypen nicht zugesichert werden.

⁵ Vgl.: <https://ocr-d.de/>

⁶ Vgl.: <https://ocr-d.de/en/workflows>

Auflistung der verwendeten OCR-D-Prozessoren:

- **Binarisierung:** Dieser Prozessor wandelt ein Farbbild in ein Schwarz-Weiß-Bild um. Die Konvertierung ist nötig, da spätere Prozessoren (Begradigung, Segmentierung oder Erkennung) sonst schlechtere Ergebnisse liefern. Im Workflow wurde Olena (wolf) verwendet.
- **Rauschentfernung:** Dieser Prozessor entfernt Flecken (Vorder- und Hintergrund). Im Verlauf des Projekts wurde erkannt, dass die Rauschentfernung (Ocropy) die Ergebnisse der Begradigung verschlechtert. Daher wurde es wieder aus dem Workflow entfernt.
- **Begradigung:** Dieser Prozessor begradigt ein schief gescanntes bzw. fotografiertes Bild. Im Projektverlauf wurden unterschiedliche Implementierungen der Begradigung getestet. Ocropy arbeitet zwar genauer, braucht dafür jedoch wesentlich mehr Zeit. Deshalb wurde letztendlich Tesseract⁷ für diesen Schritt verwendet.
- **Zuschnitt:** Dieser Prozessschritt schneidet das Bild auf den für die spätere Analyse relevanten Teil zu und entfernt unnötigen Hintergrund. Im Workflow wurde die OCR-D Implementierung ocrd-anybaseocr-crop verwendet, da Tesseract anfällig für stark kontrastierenden Noise ist und schlechtere Ergebnisse liefert.
- **Segmentation & OCR:** In diesem Prozessschritt werden mittels Tesseract die Textregionen segmentiert und die eigentliche Texterkennung durchgeführt. Im Projektverlauf wurden auch hier unterschiedliche Implementierungen getestet. Im Test wurde Tesseract (CPU) mit der GPU-Alternative Calamari⁸-OCR^[10] (Wick et al., 2021) verglichen. Letztendlich wurde Tesseract in den Workflow integriert, da die Qualität der Ergebnisse deutlich besser ist. Dies liegt vor allem daran, dass kein hinreichend gutes deutsches Modell zur Textklassifikation mit Calamari verfügbar ist.
- **Textuelle Kontextklassifikation:** Dieser Prozessor klassifiziert Textregionen als Kontexte, wenn diese mit einem Wort des Wörterbuchs eines Kennwerts unter Beachtung einer bestimmten Editierdistanz (abhängig von der Wortlänge) übereinstimmt. Außerdem wird ein Kennwert (Energieträger) textuell klassifiziert.
- **Kennwerterkennung & morphologische Nachverarbeitung:** Nach der Text- und Kontextklassifikation können die 34 Kennwerte mit der entwickelten KI-Komponente klassifiziert werden. Da die vorhergesagten Boxen teilweise Textregionen abschneiden, wird die erkannte Region durch Dilatation auf den relevanten Bereich ausgedehnt.

⁷ Vgl.: <https://github.com/tesseract-ocr/tesseract>

⁸ Vgl.: https://github.com/OCR-D/ocrd_calamari

- **Semantische Nachverarbeitung und Validierung:** Dieser Prozessor validiert die gefundenen Kennwerte mit im Voraus definierten Regeln.

Neben den genannten OCR-D Komponenten sind im Anwendungsprototypen sowie in der Datenvorbereitung und -auswertung folgende weitere Technologien und Werkzeuge im Einsatz:

- **LAREX⁹:** Larex ist ein Werkzeug für das seitenweise Annotieren von Dokumenten. Es entstehen Daten im PageXML-Format. Das Werkzeug wird zur Trainingsvorbereitung eingesetzt, um die Trainingsdaten zu annotieren und somit die Grundwahrheit (Ground Truth, GT) zu erschaffen, anhand derer die KI-Komponente das Extrahieren von Kennwerten aus den HKA erlernt.
- **Mask R-CNN** (He et al., 2017; Abdulla, 2017): Mask R-CNN ist ein neuronales Netz, das sowohl Regionen als auch Pixelmasken vorhersagt. Mask R-CNN besteht aus einem ResNet (50/101) für die Extraktion von Bildmerkmalen und einem Feature Pyramid Network (FPN). Die Implementierung basiert auf TensorFlow 1.15.
- **Docker¹⁰:** Docker ist eine Technologie sowie ein Werkzeug zur Bereitstellung von web-basierten Anwendungen mittels Containervirtualisierung. Im Projekt wurden die Anwendungen COCO-Explorer, LAREX und das SmartheC_Backend als Docker-Image bereitgestellt.
- **COCO-Explorer¹¹:** Der Coco-Explorer ist ein Werkzeug zur Visualisierung von COCO-Dateien (Lin et al., 2014). Es können Ground Truth (GT), False Positives (FP), True Positives (TP) und True Negatives (TN) Regionen angezeigt und miteinander verglichen werden. Das Werkzeug dient der Inspektion der (internen) Zwischenergebnisse der KI-Komponente, sowie zum manuellen visuellen Abgleich einzelner Datenbeispiele mit der Grundwahrheit.
- **PyCocoTools¹²:** Die PyCocoTools sind eine Schnittstelle und Werkzeugsammlung zum effizienten Laden, Verarbeiten und Visualisieren von Bilddaten mit annotierten oder erkannten Objekten im COCO-Format.

Programmbibliotheken und Lizenzen der KI-Komponente

Je nach gewähltem „Workflow“ kommen beim Betrieb der KI-Komponente zur Kennwertextraktion im vorliegenden Smart_HEC-Prototypen unterschiedliche Werkzeuge zum Einsatz. Die benötigten Komponenten werden beim Bauen des Docker-Images herangezogen. Die

⁹ Vgl.: <https://github.com/OCR4all/LAREX>

¹⁰ Vgl.: <https://www.docker.com/>

¹¹ Vgl.: <https://github.com/i008/COCO-dataset-explorer>

¹² Vgl.: <https://github.com/cocodataset/cocoapi>

folgende Auflistung orientiert sich am Workflow „smarthecc-sparse-tesseract-deu-postdec“¹³, welcher zum Redaktionsschluss standardmäßig aktiviert ist.¹⁴ Zu beachten ist, dass die einzelnen Komponenten jeweils weitere Abhängigkeiten haben können, demnach also weitere Drittanbieterbibliotheken nutzen, die hier nicht allesamt vollständig dokumentiert wurden.¹⁵

Nr.	Programmbibliothek und Verwendungszweck(e)	Version	Genutzte Lizenz
1.	OCR-D / core ¹⁶ : Verarbeitungsframework für Dokumentenanalyse und Texterkennung, mit Anpassungen für Smart_HEC zur Steuerung und Ausführung der KI-Komponente	PR #652	Apache License 2.0
2.	OCR-D / ocrd_olena ¹⁷ : Binarisierung		Apache License 2.0
3.	Olena / scribo ¹⁸ : Binarisierung via ocrd_olena		GPL-2.0 License
4.	qurator-spk / sbb_binarization ¹⁹ : Binarisierung	PR #31	Apache License 2.0
5.	OCR-D / ocrd_anybaseocr ²⁰ : Zuschnitt		Apache License 2.0
6.	OCR-D / ocrd_tesseract ²¹ : Begradigung; Segmentierung und Texterkennung zu allen Segmenten der HKA		MIT License
7.	tesseract-ocr / tessdata ²² : Vortrainierte Modelle für Tesseract, nötig für OLR und OCR		Apache License 2.0
8.	OCR-D / ocrd_segment ²³ : Textuelle Kontextklassifikation; Kennwerterkennung & morphologische Nachverarbeitung; Semantische Nachverarbeitung und Validierung	maskrcnn-cli (branch)	MIT License

¹³ Vgl.: https://git.informatik.uni-leipzig.de/smarthecc/smarthecc_backend/-/blob/master/smarthecc-sparse-tesseract-deu-postdec.mk

¹⁴ Vgl.: https://git.informatik.uni-leipzig.de/smarthecc/smarthecc_backend/-/blob/master/Dockerfile#L83

¹⁵ Beispiel: Die Abhängigkeiten der Komponente **ocrd_segment** in der für Smart_HEC angepassten Fassung sind gelistet in: https://github.com/OCR-D/ocrd_segment/blob/master/maskrcnn-cli/requirements.txt.

¹⁶ Vgl.: <https://github.com/OCR-D/core>

¹⁷ Vgl.: https://github.com/OCR-D/ocrd_olena

¹⁸ Vgl.: <https://gitlab.lrde.epita.fr/olena/olena/-/tree/master/scribo>

¹⁹ Vgl.: https://github.com/OCR-D/sbb_binarization

²⁰ Vgl.: https://github.com/OCR-D/ocrd_anybaseocr

²¹ Vgl.: https://github.com/OCR-D/ocrd_tesseract

²² Vgl.: <https://github.com/tesseract-ocr/tessdata>

²³ Vgl.: https://github.com/OCR-D/ocrd_segment/tree/master/maskrcnn-cli

Nr.	Programmbibliothek und Verwendungszweck(e)	Version	Genutzte Lizenz
9.	shapely / shapely ²⁴ : Manipulation geometrischer Objekte	>= 1.7.1	BSD-3-Clause License
10.	scikit-image ²⁵ : Bildverarbeitung in Python		Modified BSD License
11.	numpy ²⁶ : Array-Berechnungen in Python		BSD License
12.	rapidfuzz ²⁷ : Zeichenkettenvergleich in Python		MIT License
13.	opencv-python-headless ²⁸ : Wrapper-Paket für OpenCV in Python		MIT License
14.	pycocotools ²⁹ : Python Paket der COCO API	ocrd-fork ³⁰	FreeBSD
15.	Imgaug ³¹ : Bildaugmentierung für Neuronale Netze		MIT License
16.	tensorflow-gpu ³² : Machine-Learning Framework	1.15.4	Apache License 2.0
17.	h5py ³³ : Lesen und Schreiben von HDF5-Dateien mit Python	2.10	BSD License
18.	Mask R-CNN ³⁴ : Implementierung von Mask R-CNN für Python 3, Keras und TensorFlow	bertsky/Mask_RCNN (fork)	MIT License

Tabelle 3: Programmbibliotheken und deren genutzte Lizenzen für die KI-Komponente (Prototyp)

Hinweis zur Lizenzierung

Mindestens eine der Programmbibliotheken, die für den Betrieb des Prototyps des smarten HeizChecks herangezogen werden, wird unter einer Copyleft-behafteten Lizenz genutzt: Die Komponente „olena / scribo“ kommt unter der „GNU General Public License version 2“ zum Einsatz. Bezogen auf den lauffähigen Prototypen ist daher bei seiner Weitergabe auch der

²⁴ Vgl.: <https://github.com/shapely/shapely>

²⁵ Vgl.: <https://pypi.org/project/scikit-image/>

²⁶ Vgl.: <https://pypi.org/project/numpy/>

²⁷ Vgl.: <https://pypi.org/project/rapidfuzz/>

²⁸ Vgl.: <https://pypi.org/project/opencv-python-headless/>

²⁹ Vgl.: <https://pypi.org/project/pycocotools/>

³⁰ Vgl.: <https://pypi.org/project/ocrd-fork-pycocotools/>

³¹ Vgl.: <https://pypi.org/project/imgaug/>

³² Vgl.: <https://pypi.org/project/tensorflow-gpu/>

³³ Vgl.: <https://pypi.org/project/h5py/>

³⁴ Vgl.: https://github.com/matterport/Mask_RCNN

zugrunde liegende Quellcode (zumindest der betroffenen Komponenten, möglicherweise auch des Gesamtsystems) offenzulegen. Alternativ kann entweder die betroffene Bibliothek durch ein vergleichbares Werkzeug ersetzt werden, oder es kann mit den Herausgebern der Bibliothek möglicherweise eine alternative Lizenzierung z. B. im Rahmen dualer Lizenzierung angestrebt bzw. vereinbart werden. Bezogen auf die o. g. weiteren Bibliotheken können darüber hinaus weitere Verpflichtungen bestehen, die sich aufgrund der vielfältigen genutzten Lizenzen unterscheiden können und gegebenenfalls im Einzelfall zu prüfen sind (Nennung von Urhebern, Einschränkungen hinsichtlich der Lizenzwahl etc.). Auch hier wäre bei Bedarf die Nutzung alternativer Lizenzen in Absprache mit den jeweiligen Urhebern bzw. Herausgebern zu prüfen.

2.1.3. Systemarchitektur und Datenfluss

Aufbereitung der Daten und Annotation

Die Datengrundlage der HKAs wurde im ersten Halbjahr des Projekts in zwei Phasen zusammengetragen. Zunächst wurden durch die Verbundpartner SENERCon und co2online HKAs von Brunata (74), Ista (72), Techem (72) und einem anderem MDL (1) anonymisiert und zur Verfügung gestellt. Diese lagen bereits aus einer anderen Aktion zur Erstellung von Heizgutachten vor. Es handelt sich um Scans von hoher Qualität, die teilweise allerdings mehrere Jahre alt sind und im weiteren Verlauf als „**Altdaten**“ bezeichnet werden.

In der zweiten Phase wurden durch co2online zusätzlich HKAs von anderen MDL durch verschiedene Mailings und Mitmachaktionen gesammelt. Diese sind im Allgemeinen von geringerer Qualität und kommen den zu erwartenden Echtdaten näher. Insgesamt wurden dabei HKAs von BFW (15), Brunata (31), Ista (57), Kalo (22), Minol (13), Techem (64) sowie Sonstige (74) gesammelt. Diese werden im weiteren Verlauf als „**Neudaten**“ bezeichnet. Im Schnitt besteht jede HKA aus zwei Seiten.

Die Annotation der Daten erfolgt mittels **Larex**. Hierfür wurden die HKAs zunächst durch den OCR-D Workflow geleitet, allerdings nur bis einschließlich des OCR-Schrittes: Binarisierung, Begradigung, Zuschnitt, Segmentation & OCR.

Für die Binarisierung wird `ocrd-olena-binarize`, für die Begradigung `ocrd-tesseract-deskew`, für den Zuschnitt `ocrd-anybaseocr-crop` und für die Segmentation und OCR `ocrd-tesseract-recognize` verwendet.

Anschließend mussten die Bilder/PageXMLs für jeden Kennwert vervielfältigt werden, da Larex sonst nicht in der Lage ist, für Kontext/Ziel jedes Kennwerts Annotationen zu erzeugen. Für jeden Kennwert kann nun getrennt die Annotation der Kontexte und des/der Ziele erfolgen.

Nach Abschluss der Annotation wurden die Annotationen in COCO-Daten umgewandelt, die für das Training des Mask R-CNNs benötigt werden.

Training und Evaluation der KI-Modelle

Zu Beginn des Projektes wurde **für jeden MDL ein separates Modell des Mask R-CNN trainiert**. Im späteren Verlauf wurde das Mask R-CNN auf allen Datenmessdienstleisterunabhängig trainiert mit dem Ziel, die 34 Kennwerte und den jeweiligen MDL vorherzusagen. Dazu musste das Training signifikant verlängert werden.

Das Training wurde auf dem **HPC-Cluster** (High-Performance Computing Cluster) des ZIH der TU Dresden³⁵ durchgeführt. Neben den dort verfügbaren Softwaremodulen wurde lediglich das erweiterte ocrd_segment, Mask R-CNN sowie die nicht als Module verfügbaren Abhängigkeiten benötigt.

Folgendes Trainingsregime hat sich dabei als Bestes erwiesen. Zunächst den ResNet-Backbone (He et al., 2016) des Mask R-CNN mit den Gewichten aus einem auf den COCO-Daten trainierten Netz zu initialisieren. Anschließend wurde zunächst auf den Altdaten trainiert. Das Training selbst erfolgt dabei in drei Teilen (sog. Stages). Zunächst wurden nur die Ausgabeschichten (Heads) des Netzes für 40 (100) Epochen trainiert – d. h., dass alle anderen Gewichte im Netz nicht verändert werden. Im zweiten Teil wird dann das ResNet für 80 (200) Epochen fein-trainiert. Im dritten Schritt wird das gesamte Netz für 40 (100) Epochen trainiert. Jedes Training besteht aus jeweils 1.000 Schritten pro Epoche.

Nach Abschluss des Trainings auf den Altdaten wurde das gleiche **dreiteilige Training** auf dem bereits trainierten Netz mit den Neudaten durchgeführt.

Für das Mask R-CNN, das messdienstleisterunabhängig trainiert wurde, wurde zusätzlich noch ein drittes Mal auf allen Daten, d. h., Altdaten und Neudaten, trainiert.

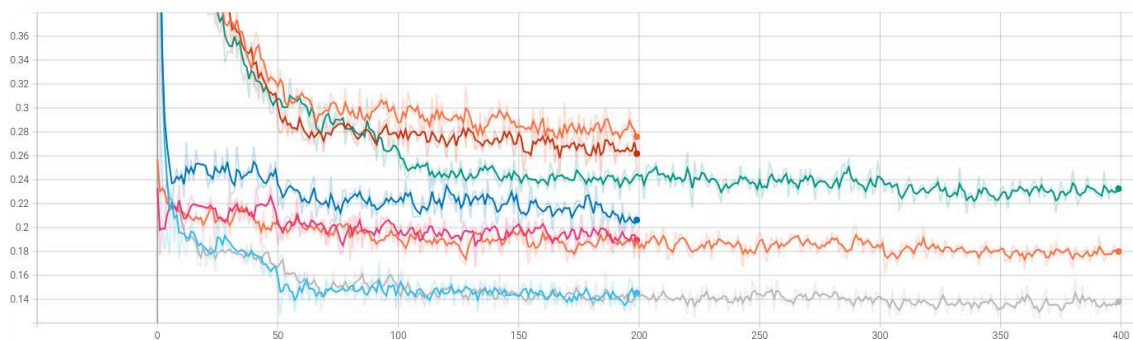


Abbildung 2: Trainingsfehler verschiedener Trainings derselben Modelle wie beim Validierungsfehler.

³⁵ Vgl.: <https://tu-dresden.de/zih/>

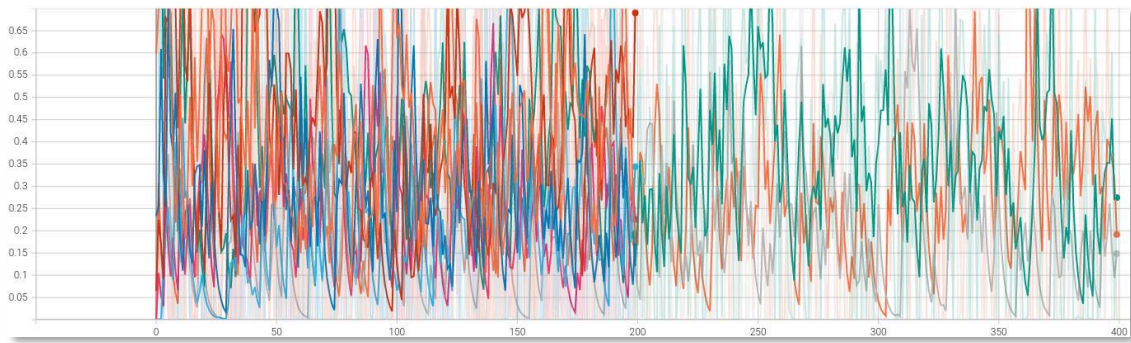


Abbildung 3: Validierungsfehler verschiedener Trainings derselben Modelle wie beim Trainingsfehler.

In Abbildung 2 und Abbildung 3 sind für verschiedene Trainingsdurchläufe der jeweilige **Trainingsfehler und der Validierungsfehler** zu sehen. Dabei wird deutlich, dass sich der Trainingsfehler zu Beginn stark reduziert und später anfängt zu oszillieren. Der Validierungsfehler hingegen oszilliert immer auf niedrigem Niveau. Dies deutet darauf hin, dass das Validierungsset aus zu wenigen Datensätzen besteht und generell zu wenig Daten vorhanden sind.

Nach Abschluss des Trainings wird jeweils eine Vorhersage mit dem trainierten Modell durchgeführt. Die Ergebnisse werden ebenfalls als COCO-JSON gespeichert.

Die **Auswertung der Prädiktionen** erfolgt mittels PyCocoTools³⁶. Dabei werden die vorhergesagten Kennwerte jedes MDL mit den Werten der Grundwahrheit verglichen. Als Metriken zur Beurteilung der Klassifikationsgüte des Netzes werden **mAP (mean Average Precision)** sowie **mAR (mean Average Recall)** genutzt. MAP gibt dabei die über alle Klassen hinweg gemittelte Klassifikationsgenauigkeit an. Der mAR gibt das Verhältnis von erkannten Objekten zur Grundwahrheit, gemittelt über alle Klassen, an. Dadurch entstehen für jeden MDL getrennt nach Altdaten, Neudaten und für jeden Kennwert ein **Precision- und ein Recallwert**. In Abbildung 4 sind diese aggregiert zu sehen:

- „**Training 2**“ entspricht dem Stand der ersten Beta-Phase und integriert zum ersten Mal Trainingsdaten für alle 34 Kennwerte über alle MDL;
- „**Training 3**“ basiert auf einer nachkorrigierten Datengrundlage und ist zum Abschluss des Berichtszeitraumes im Prototypen im Einsatz;
- „**Training Multi**“ zeigt die erzielten Ergebnisse mit einer Modellvariante, die alle sieben Messdienstleisterklassen in einem neuronalen Netz vereint. Aufgrund der deutlich schlechteren Erkennungsergebnisse kann diese Modellvariante zum Abschluss des Berichtszeitraumes ohne weitere Trainingsiterationen nicht aktiv eingesetzt werden.

³⁶ Vgl.: <https://github.com/cocodataset/cocoapi/tree/master/PythonAPI/pycocotools>

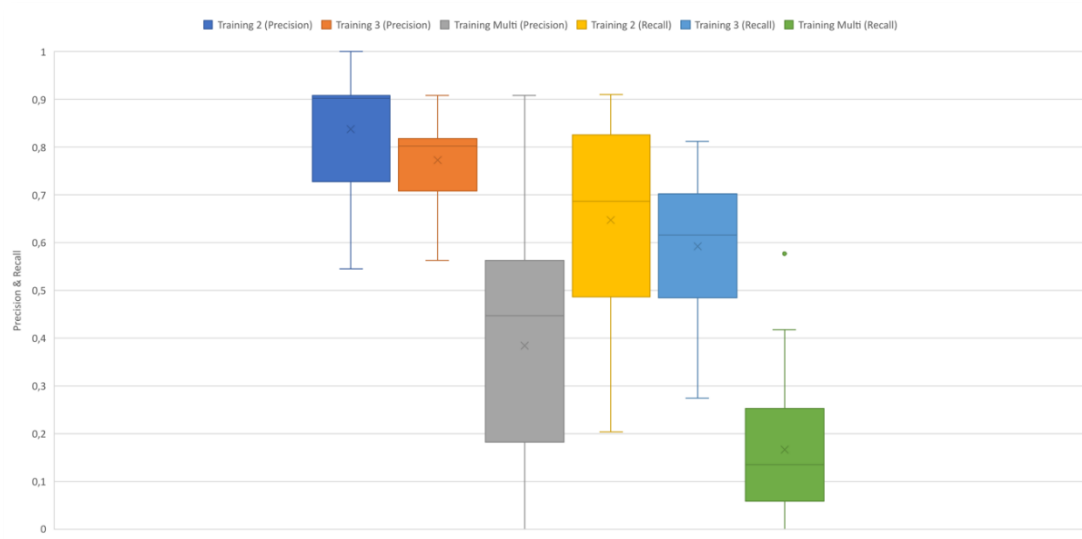


Abbildung 4: Genauigkeit (Precision) und Vollständigkeit (Recall) für die Kennwertextraktion gemittelt über alle MDL – zu drei exemplarischen Entwicklungsständen der KI-Komponenten.

In Abbildung 5 ist die Entwicklung der **Anzahl erkannter Kennwerte** erkennbar – gemittelt über alle innerhalb der dargestellten jeweiligen Kalenderwochen analysierten HKAs. Der deutliche Anstieg in Kalenderwoche 2021-44 ist in verbesserten Annotationsdaten und deren Berücksichtigung in erneuten Trainingsdurchläufen begründet (entsprechend „Training 3“ in Abbildung 4).

Nach der quantitativen Evaluation wurde eine **qualitative Evaluation** durchgeführt. Dazu wurden die vorhergesagten Daten im COCO-Explorer visualisiert und manuell mit der Ground Truth verglichen.

Als letzter Schritt werden alle annotierten Kontexte in den PageXMLs durchsucht. Dabei werden alle als Kontext markierten Wörter kopiert und separat für jeden Kennwert gespeichert. Diese Liste dient zur Laufzeit als Grundlage für die Kontextklassifizierung.

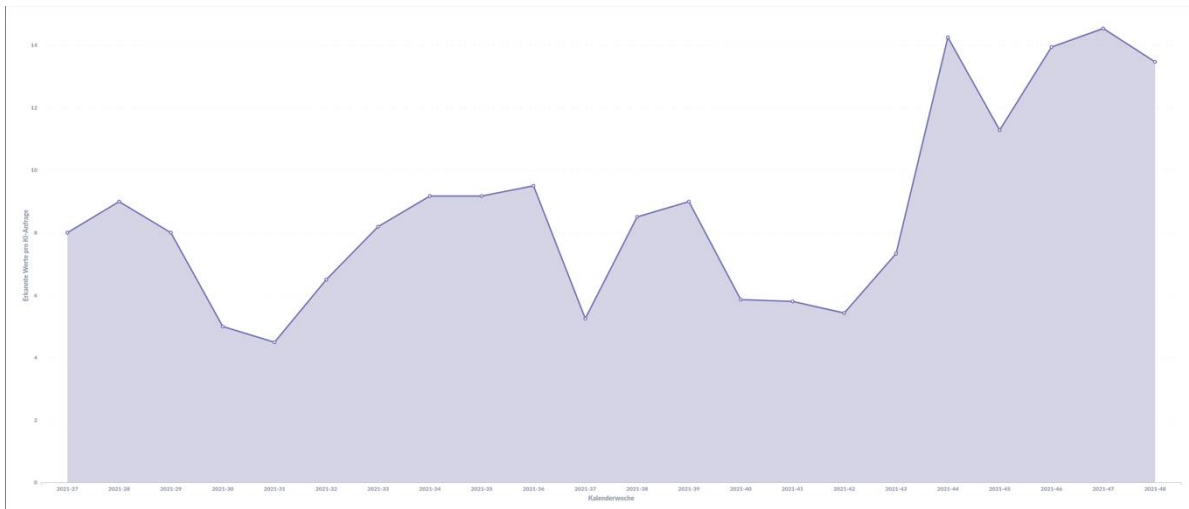


Abbildung 5: Wöchentliche Entwicklung der durchschnittlichen Anzahl an automatisch erkannten Kennwerten pro durchgeführter HKA-Analyse.

Prototyp eines Produktsystems

Das **Laufzeitsystem (Backend)** verknüpft die bisher beschriebenen Komponenten. OCR-D verfügt über eine Webschnittstelle, an die ein oder mehrere Bilder als base64 encodiertes JSON gesendet werden können. Diese gibt das jeweilig zuletzt erzeugte Bildderivat (base64 encodiert) und PageXML als JSON zurück.

Dabei ist zu beachten, dass im Fall der messdienstleisterabhängigen Variante **für jeden MDL ein separates Backend** gestartet wird. In diesem Fall gibt es sieben verschiedene Backend-Prozesse.

Die Bilder durchlaufen den oben beschriebenen OCR-D Prozess: Binarisierung, Begradigung, Zuschchnitt, Segmentierung & OCR, Kontextklassifikation, Mask R-CNN, re-OCR, Nachkorrektur.

Der zur Kontext-Klassifikation genutzte **OCR-D Prozessor** (ocrd-segment-classify-formdata-text) geht alle gefunden Textregionen durch und schaut, ob diese Kontext(e) für einen der 34 Kennwerte enthält. Dazu wurde eine Wortliste, welche aus den annotierten Kontexten aus Larex aufgebaut wird, erstellt. Zur Laufzeit werden Textregionen als Kontext markiert, wenn diese durch einen Stringvergleich unter Berücksichtigung einer maximalen Editierdistanz mit einem Wort der aufgebauten Wortliste übereinstimmen. Dies könnte in Zukunft auch durch ein trainierbares Modell ersetzt werden, um Kontexte, welche fälschlicherweise als solche markiert wurden, zu minimieren. Neben der Kontextklassifikation werden textuelle Kennwerte (momentan nur Energieträger) von diesem Prozessor direkt klassifiziert.

Der Mask R-CNN OCR-D Prozessor (ocrd-segment-classify-formdata-layout) erzeugt auf Basis eines trainierten Modells Vorhersagen für die 34 Kennwerte und schreibt diese ins

PageXML. Die Klassifikation der Kennwerte erfolgt mit einem aktiven Kennwert pro Bild, d. h., um 34 Kennwerte auf einem Bild zu erkennen, muss dieses insgesamt 34-mal klassifiziert werden.

Die re-OCR wird lediglich auf den Regionen durchgeführt, bei denen das Mask R-CNN etwas erkannt hat und stellt dabei sicher, dass die OCR auf die Kennwerte passt, d. h., dass bei Kennwerten, bei denen eine Zahl erwartet wird, auch nur solche Zeichen zugelassen werden.

Der zur **Nachkorrektur** genutzte OCR-D Prozessor (ocrd-segment-postcorrect-formdata) prüft die OCR-Hypothesen aus dem vorherigen Schritt und plausibilisiert diese. So wird z. B. sichergestellt, dass tatsächlich ein Datum von der OCR erkannt wird, wenn ein Kennwert ein Datum enthalten sollte. Das gesamte Backend ist als **Dockerimage** gepackt und verfügt über eine docker-compose-Konfiguration. Dadurch ist es möglich, alle sieben verschiedene Backends durch einen einzelnen Befehl zu starten.

Von den oben aufgelisteten OCR-D Prozessoren läuft lediglich orcd-segment-classify-formdata-layout auf GPU. Der Unterschied zur Bearbeitung einer einzelnen Seite beträgt in etwa 25 - 30 Sekunden (auf der GPU ist die Laufzeit in etwa 45 - 50 Sekunden und auf demselben Rechner ohne GPU-Nutzung 70 - 75 Sekunden). Je mehr Kennwerte erkannt werden, desto höher ist der Geschwindigkeitsvorteil mit GPU.

2.1.4. Eingeflossene Forschungs- und Entwicklungsleistung

Einige der genutzten OCR-D Prozessoren (Binarisierung, Begradigung, Zuschchnitt, Segmentation und OCR), v. a. für die Bildvorverarbeitung, konnten direkt in den Workflow integriert werden. Dennoch mussten **Anpassungen und Erweiterungen am OCR-D-Werkzeugkasten** vorgenommen werden, um die Projektziele zu erreichen.

Während der Projektlaufzeit wurden folgende Werkzeuge von OCR-D angepasst bzw. neu entwickelt:

- **OCR-D:**
 - **OCR-D core:** Implementation eines Workflow-Servers:
 - Basiert auf Flask und uWSGI
 - Multiprocessing mit entweder GPU- oder CPU-Worker (Performanzoptimierung)
 - Endpunkt zur Bearbeitung gesendeter Bilder/Start des Workflows
 - **OCR-D anybaseocr-crop:** Verbesserungen und Fehlerbehebung im Zuschchnitt-Prozessor

- **Classify-Formdata-Layout:** neu erstellter OCR-D Prozessor, Wrapper für Vorhersagen der 34 Kennwerte mit trainiertem Mask R-CNN Modell:
 - Parallele Nachverarbeitung der vorhergesagten Masken
 - Ausdehnung auf den relevanten Bereich, sodass kein Text abgeschnitten wird
 - Non-Maximum Suppression für bestimmte Klassen
- **Classify-Formdata-Text:** neu erstellter OCR-D Prozessor (s. Kontextwortextraktion)
- **Classify-Formdata-Dummy:** neu erstellter OCR-D Prozessor, vereinigt die 34 annotierten Kennwerte aus LAREX in einer XML-Datei
- **Segment-Postcorrect-Formdata:** neu erstellter OCR-D Prozessor, plausibilisiert und validiert OCR-Hypothesen auf Basis vorher definierter Regeln
- **LAREX:** Fehlerbehebung und Vorbereitung für GT-Annotation
- **Mask R-CNN:** für den erfolgreichen Einsatz von Mask-RCNN zur Kennwerterkennung mussten einige Änderungen am ursprünglichen Programmcode vorgenommen werden:
 - Eingabe der Textregionen und Kontexte im 4. bzw. 5. Kanal des Eingabebildes (vgl. Abbildung 6)
 - Aktualisierung der Python-Abhängigkeiten
 - Vorhersagen für unterschiedliche Klassen innerhalb eines Batches erlauben
 - Parametrisierung der Batch-Size mittels Konfigurationsdatei
 - Meta-Informationen zu Bildern bzw. Klassen zur Laufzeit bereitstellen
 - Nachverarbeitung der Vorhersagen mittels Informationen aus Image-Meta
 - Inferenz-DataGenerator: laufzeitoptimierter Datengenerator für Vorverarbeitung und Bereitstellung der zu klassifizierenden Bilder
 - Detect-Generator: laufzeitoptimierter Datengenerator für die schnelle/parallelisierte Klasseninferenz mittels `Model.predict_generator` (Objektinstanz von Inferenz-DataGenerator)
 - Separate Bildklassifikation, genutzt für die Vorhersage des MDL
- **COCO-Explorer:** Visualisierung der Prädiktionen vs. Ground Truth
 - Anpassungen zur besseren Benutzbarkeit (optimierte Labelpositionierung (keine Überschneidungen), Dokumentensuche implementiert)
- **Shell-Skripte zur Evaluation des Modells bzw. Workflows:** Automatisierung der einzelnen Evaluationsschritte und Aggregation in Ergebnistabelle

Abrechnung erstellt am 27.08.2013

Energiekosten	Menge	Gas	Betrag EUR	Weitere Kosten der Heizungsanlage	Datum	Betrag EUR
Bezüge:				Übertrag Brennstoffkosten		6.474,48
Brennstoffkosten	8.539 m³		6.474,48	Betriebsstrom	11.06.2013	476,45
				Kaminkehrer+Messung	15.05.2013	94,97
				Wartungsservice	04.06.2013	128,28
				Eichservice	04.06.2013	120,25
				Verbrauchsabrechnung		750,20
				Summe der Kosten		8.044,63
				Nutzerwechselgebühr		38,94
				Gesamtbetrag Liegenschaft		8.083,57

* auf einzelne Nutzer zu verteilende Kosten

Abbildung 6: Ausschnitt einer HKA mit Beispielen für die Anordnung der Kontextklassen „Energiekosten“, „Betrag in EUR“, „Brennstoffkosten“ (jeweils weiß hinterlegt). Sonstige Inhalte sind transparent hinterlegt.

2.1.5. Exemplarische Herausforderungen und Lösungsstrategien

Umgang mit geringer Anzahl an HKAs

Für den komplexen Anwendungsfall der automatischen Erkennung und Extraktion von Kennwerten einer HKA über unterschiedliche Layouts der einzelnen MDL hinweg, ist der Datenstamm von 500 Dokumenten (ca. 84 Dokumente/MDL im Fall einer Gleichverteilung) sehr gering. Gerade der Datenstamm kleinerer MDL wie beispielsweise BFW (15) enthält eigentlich zu wenig Beispiele, um eine für maschinelle Lernverfahren notwendige **Einteilung des Datensatzes in Training- sowie Testdaten** vorzunehmen. Die Aufteilung des Datensatzes vermeidet Overfitting (Überanpassung auf den Trainingsdaten) und maximiert die Generalisierungseigenschaften des Neuronalen Netzes.

Die trotzdem durchgeführte Einteilung führt in diesem Fall dazu, dass das Mask R-CNN nur schwer trainierbar ist und der Validierungsfehler oszilliert, siehe Abbildung 3.

Umgang mit Scans vs. Fotos

Die Anforderung, dass Nutzer*innen ihre HKA einfach mit dem Handy fotografieren und hochladen können und die zugleich erforderliche Möglichkeit der Analyse von PDFs (Scans) führt dazu, dass der gesamte OCR-Workflow mit **sehr unterschiedlichen Bilddaten** umgehen muss. Dabei ist zu beachten, dass viele der OCR-D-Prozessoren nicht für den Umgang mit Fotografien gedacht sind, sondern lediglich mit gescannten Dokumenten. So hat sich z. B. gezeigt, dass die Binarisierung großflächige Schatten komplett geschwärzt hat, siehe Abbildung 7. Solche Fehler können im Betriebszustand nicht gänzlich ausgeschlossen werden.

The image shows a document from 'techem' with a table of energy costs. A pink box highlights a section titled 'Ausschnitt: Wert'. The table contains various energy-related data points, including 'Energiekosten', 'Brennstoffkosten', and 'Energiekosten des Energieträgers'. The table is organized into columns with headers like 'Energiekosten', 'Brennstoffkosten', and 'Energiekosten des Energieträgers'. The data is presented in a structured format with rows for different energy sources and their associated costs.

Abbildung 7: Fehler bei der Binarisierung einer HKA

Annotation der Kennwerte in den verschiedenartigen HKAs

Der Erfolg von KI-Projekten hängt entscheidend von der verfügbaren und eingesetzten Datengrundlage ab. Dabei spielt vor allem die sogenannte **Grundwahrheit** („Ground Truth“, GT) eine zentrale Rolle. Für die Erstellung der GT wurde eine **Annotationsrichtlinie**³⁷ erarbeitet, die den Prozess der Annotation beispielhaft beschreibt. In Abbildung 8 ist die Annotation für den Kennwert „Kosten des Energieträgers“ (energietraeger_kosten) visualisiert. Für diesen Kennwert (6.474,48) bilden *Energiekosten*, *Betrag in EUR* und *Brennstoffkosten* die Kontextklassen. An diesem Beispiel ist das Prinzip des Fadenkreuzes gut zu erkennen.

³⁷ Vgl.: <https://hackmd.io/@bertsky/Sy4oxZajL>

BRUNATA metrona Heizkosten- und Warmwasserkosten-Abrechnung

Liegenschafts-Nr. A-0066
 Nutzeinheit-Nr.
 Verwaltungs-Nr. 661361

505205

Abrechnungszeitraum 01.07.2012 - 30.06.2013
 Ihr Nutzungszeitraum 01.07.2012 - 30.06.2013

Abrechnung erstellt am 27.08.2013

Kostenaufstellung

Energiekosten	Menge	Gas	Betrag EUR	Weitere Kosten der Heizungsanlage	Datum	Betrag EUR
Bezüge:				Übertrag Brennstoffkosten		6.474,48
Brennstoffkosten	8.539 m³		6.474,48	Betriebsstrom	11.06.2013	476,45
				Kaminkehrer+Messung	15.05.2013	94,97
				Wartungsservice	04.06.2013	128,28
				Eichservice	04.06.2013	120,25
				Verbrauchsabrechnung		750,20
				Summe der Kosten		8.044,63
				Nutzerwechselgebühr		38,94
				Gesamtbetrag Liegenschaft		8.083,57

* auf einzelne Nutzer zu verteilende Kosten

Abbildung 8: Visualisierung der Annotationen für den Kennwert „Kosten des Energieträgers“: Kontextklassen (türkis), Kennwertklasse (rot), für diesen Kennwert nicht relevanter Text (lila)

Die dargestellte Annotation muss für jede HKA im Trainingsdatenbestand und jeden Kennwert einzeln vorgenommen werden. Um 34 Kennwerte einer HKA zu annotieren, müssen insgesamt 34 Bilder mit Kontext- und Kennwertklassen versehen werden. Bei 495 HKAs im Trainingsdatenbestand ergeben sich somit **16.830 zu annotierende Bilder**. Für diese Aufgabe wurden zwei studentische Hilfskräfte hinzugezogen, die insgesamt 500 - 600 Stunden mit der Annotation von HKAs und den anschließenden Verbesserungen verbrachten.

Das Annotationsschema wurde im Verlauf des Projekts mit mehreren Iterationen der Qualitätssicherung, bestehend aus Kennwertvorhersage und anschließender Evaluation der Ergebnisse, deutlich verbessert. Eine letzte **qualitative Evaluation des Annotationsschemas** steht zum Projektende aus. Dennoch zeigt die Steigerung der durchschnittlich vorhergesagten Anzahl an Kennwerten pro HKA in Abbildung 5 die Wirksamkeit der durchgeführten Verbesserungen.

Trainingsregime

Durch den unterschiedlichen Datenbestand mit sehr gut digitalisierten HKAs (Altdaten), weniger gut digitalisierten HKAs (Neudaten) und der geringen Anzahl an Trainingsdaten wurden unterschiedliche Trainingsregimes getestet, bevor das vorab skizzierte Regime als das Beste identifiziert werden konnte.

So wurde zunächst ein Modell auf Altdaten sonstiger MDL (angereichert mit HKAs aller MDL) trainiert und diese als Basis genommen für ein Training für die einzelnen MDL. Dabei wurde jeweils getestet, wie sich die unterschiedliche Anzahl an Epochen und Stages auswirkt. So hat sich auch generell gezeigt, dass es für das Trainieren der Gruppe der Sonstigen MDL wichtig ist, die zugehörigen Daten mit Daten anderer HKAs anzureichern.

2.1.6. Eingesetzte Ressourcen und Ressourcenbedarf bei Anpassungen

Ressourcenverbrauch und Rechenzeit

Das Trainieren der KI-Modelle erstreckte sich zeitlich über 1,25 Projektjahre. Die Abschätzung des Ressourcenverbrauchs und der eingesetzten Rechenzeit beschränkt sich somit auf diese 15 Monate. Für das Training der KI-Modelle wurden darin **durchschnittlich 25 CPU-Tage pro Monat** ermittelt, basierend auf einem zweimonatigen Referenzzeitraum im zweiten Halbjahr des zweiten Projektjahres. Für das Training wurde das HPC-Cluster des ZIH der TU Dresden genutzt (IBM Power9, NVIDIA Volta V100),³⁸ dabei waren für das vorliegende durchgehend durchschnittlich 1 GPU und 16 CPU-Kerne für das Projekt aktiv. Das entspricht in etwa 25% der angegebenen Leitung eines HPC-Knotens.

Dementsprechend ergibt sich ein geschätzter Stromverbrauch von **5.400 Kilowattstunden** (kWh) für das Trainieren der KI-Modelle im oben genannten Zeitraum von 1,25 Jahren.³⁹

Im Falle einer erforderlichen **Anpassung der KI-Komponente**, zum Beispiel an neuartige Daten, können Änderungen an verschiedenen Modulen bzw. Verarbeitungsschritten erforderlich sein. Der abzuschätzende Aufwand hängt dabei unter anderem von folgenden Überlegungen ab, die z. B. in Leistungsbeschreibungen bei Änderungsanforderungen einfließen können.

Anpassungen an Prozessoren der Bildvorverarbeitung

Werden Anpassungen an Prozessoren vorgenommen, die vor der Layouterkennung (classify-formdata-layout) für die Bildvorverarbeitung verantwortlich sind und das Koordinatensystem des Bildes beeinflussen, müssen alle Bilder und deren GT-Annotationen gemäß den Änderungen transformiert bzw. neu annotiert werden. Sobald Änderungen an den GT-Daten vorgenommen werden, muss außerdem ein neues Modell trainiert werden. Ändern sich nicht nur die Koordinaten der untersuchten Kennwertklassen, sondern auch die der Kontextklassen, müssen zusätzlich das zur Kontextklassifikation genutzte Wörterbuch angepasst werden. Dazu können die als Kontext klassifizierten Wörter mit Hilfe eines Shell-Skripts automatisch extrahiert und anschließend in einer Tabelle zusammengefasst werden.

³⁸ Vgl.: https://doc.zih.tu-dresden.de/jobs_and_resources/power9/

³⁹ $(15 * 25 * 24)h * (0,25 * 2.400)W / 1.000 = 5.400 kWh$

Erweiterung des Datenstamms um einen Kennwert

Falls zusätzliche Klassen im Layout der MDL erkannt werden sollen, müssen die Kontexte und Kennwerte zu diesen Klassen in allen Daten annotiert werden. Zusätzlich muss die zu erkennende Klasse in der Konfiguration des Mask-RCNN (OCR-D-Segment) und in den Shell-Skripten zur Evaluation und Vervielfältigung hinzugefügt werden. Weiterhin müssen die Regeln zur Nachkorrektur der neuen Klasse ergänzt werden. Anschließend müssen ein neues Training (ggf. Erhöhung der Trainingsiterationen), die Kontextextraktion sowie die Erweiterung des zur Kontextklassifikation genutzten Wörterbuch durchgeführt werden.

Erweiterung des Datenstamms um einen MDL

Kommt ein weiterer MDL zur Datenbasis hinzu, so muss dieser in die Verzeichnisstruktur aufgenommen werden. Zusätzlich müssen die notwendigen Schritte bis zum LAREX-Import (Vorverarbeitung mittels OCR-D, Vervielfältigung in 34 Kennwerte) erneut gerechnet und anschließend annotiert werden. Nach der Annotation müssen die Kontext- und Kennwertklassen in das Zielformat COCO umgewandelt werden, sodass ein neues Modell trainiert werden kann (ggf. Erhöhung der Trainingsiterationen). Wurden durch das Hinzufügen des neuen MDL neue Kontexte annotiert, so muss das zur Kontextklassifikation genutzte Wörterbuch aktualisiert werden. Das neu trainierte Modell müsste zusätzlich als Backend in der Docker-Compose Datei mit eigenem Port aufgenommen werden.

2.1.7. Erzielte Ergebnisse

Bezugnehmend auf die oben genannten **Technischen Ziele** wurde ein OCR-D Workflow aufgebaut, der es ermöglicht, HKAs zu verarbeiten und dabei geforderte Kennwerte zu erkennen, zu extrahieren und dabei auch sicherzustellen, dass der erkannte Wert dem entspricht, was für den Kennwert erwartet wird.

Für die KI-Komponente konnte gezeigt werden, dass es möglich ist, ein Neuronales Netz darauf zu trainieren, HKAs unterschiedlicher MDL zu analysieren und dabei auch den MDL zu erkennen. Dadurch ist es auch möglich, eine zukünftige Layoutänderung eines MDL relativ einfach mit dem bereits bestehen Modell abzudecken. Entweder das Neuronale Netz ist bereits jetzt in der Lage, die Kennwerte zu erkennen und zu extrahieren, oder es kann durch ein einfaches, zusätzliches Training erreicht werden. Dazu müssen dann entsprechend die HKAs annotiert werden.

2.1.8. Zwischenfazit und weiterer Bedarf an Forschung und Entwicklung

Die im Forschungsprojekt entwickelte KI-Komponente des Prototyps zeigt, in welchem Rahmen die automatische Extraktion von Kennwerten aus HKAs derzeit **allein mit Open Source-**

Komponenten möglich ist. Die Anforderungsbereiche „**Erkennungsqualität**“, „**Laufzeit/Performanz**“ der HKA-Analysen, sowie das Thema „IT-Sicherheit“ machen weitere Anpassungen und spätere Wartungsarbeiten erforderlich, wenn der Anwendungsprototyp dauerhaft genutzt werden soll.

Die in diesem Projekt prototypisch umgesetzte KI-Komponente kann insbesondere hinsichtlich der Erkennungsqualität und Laufzeit/Performanz **erweitert und gegebenenfalls verbessert werden**. Dazu müssten entsprechende Anpassungen mit anschließender Evaluation getestet werden. Diese Anpassungsmöglichkeiten werden nachfolgend nach ihrem geschätzten Aufwand aufsteigend priorisiert aufgeführt:

Erweiterung der Trainingsdaten und Anpassungen an der Trainingskonfiguration

Die Vorhersagegenauigkeit der messdienstleisterabhängigen Modelle variiert stark in Abhängigkeit des jeweiligen MDL und dem vorherzusagenden Kennwert. Wesentliche Gründe hierfür sind:

- Geringe Anzahl an Trainingsbeispielen bzw. Annotationen für diesen Kennwert des jeweiligen MDL
- Die Zulassung mehrerer Prädiktionen für bestimmte Kennwerte (Konfigurationsproblem): Für bestimmte Kennwerte (`nebenkosten_messgeraet_miete`) sind mehrere Annotationen in der GT vorhanden. Zur Laufzeit wird für jeden Kennwert maximal eine Prädiktion zugelassen.

Mehr Trainingsdaten dürften also die Qualität der Modelle substanziell verbessern. Weiterhin zeigt vor allem das messdienstleisterunabhängige Modell große Schwierigkeiten, die Kennwerte über unterschiedliche Layouts der HKAs zu erkennen. Lösungsansätze für eine Verbesserung dieses Umstands könnten die nachfolgenden Komponenten enthalten.

ResNet101 als Backbone (He et. al., 2016)

Das in diesem Projekt genutzte Mask R-CNN verfügt über zwei unterschiedliche Feature-Extraktions-Netze, die zur Merkmalsextraktion spezifischer Klassen genutzt werden. Im aktuellen Stand des Projekts wird das **ResNet50** genutzt. Dieses besteht aus 56 Schichten mit insgesamt 850.000 trainierbaren Parametern, die Merkmale zur Klassifikation aus Bildern extrahieren. Im Gegensatz dazu besteht das **ResNet101** aus 110 Schichten mit insgesamt 1.700.000 trainierbaren Parametern. Je höher die **Anzahl an trainierbaren Parametern**, desto höher die Kapazität, vorhandene Zusammenhänge zu lernen. Vor allem bei der Konsolidierung der sieben unterschiedlichen Modelle in ein finales Netz könnte sich **eine tiefere Netzarchitektur** positiv auf die Vorhersagequalität auswirken. Nachteilig würde sich dies

jedoch auf die Verarbeitungsgeschwindigkeit im Layoutklassifikator während der Inferenz des Trainings sowie auf den Ressourcenbedarf auswirken.

Erweiterte Datenaugmentation

Um die Vorhersagegenauigkeit und Kennwertwiedererkennung zu optimieren kann auf Methoden der Bildaugmentation zurückgegriffen werden. Dabei werden die Bilder während des Trainings mit Methoden der Bildbearbeitung, wie beispielsweise durch Anwendung des Blur-Effekts oder durch Anpassung des Kontrasts bzw. der Helligkeit der Bilder, zufällig und künstlich erweitert. Dieses Vorgehen kann helfen, die Genauigkeit des Klassifikators auch bei Bildern oder Scans mangelnder Qualität zu gewährleisten.

Efficient-Det (Tan et al., 2020)

Der Einsatz neuronaler Netze ist Gegenstand zahlreicher aktueller Forschungsarbeiten. In den letzten Jahren entstanden dabei neue Konzepte und Netzarchitekturen mit dem Ziel, den Ressourcenverbrauch während des Trainings und der Inferenz zu minimieren bzw. die Vorhersagegenauigkeit verglichen mit bisher verfügbaren Modellarchitekturen zu verbessern. Einen vielversprechenden Ansatz liefern Tan und Le (2020) mit der Entwicklung des Efficient-Net bzw. dem Efficient-Det. Die Genauigkeit von Efficient-Det konnte im Vergleich zu Mask R-CNN bei Reduzierung der FLOPS (floating point operations per seconds) um den Faktor 16 gesteigert werden. Der Einsatz von Efficient-Det könnte die Performanz der Layouterkennung weiter steigern.

Multimodale Modelle (Miller et al., 2020)

Ein weiterer vielversprechender Ansatz zur **Verbesserung der Vorhersagegenauigkeit** über verschiedene Layouts der HKAs hinweg leistet die Arbeit von Miller et al. (2020) mit der Klassifikation von Bildern basierend auf der Kombination visueller sowie textueller Eingabedaten. Dieser Ansatz könnte auf das Projekt Smart_HEC übertragen werden, indem Kontextklassen für Texte und Bilder der HKA die visuelle Eingabe repräsentieren. Der Vorteil des **kombinierten Ansatzes** liegt darin, dass textuelle Kontextinformationen, also die Interpretation verschiedener Wörter, gelernt werden kann. Im aktuellen Ansatz von Smart_HEC werden lediglich positionelle Merkmale der Kontext- und Kennwertklassen gelernt. Da diese in den verschiedenen Layouts der HKAs variieren, verschlechtern sich die Ergebnisse des konsolidierten und messdienstleisterunabhängigen Modells.

TensorFlow 2-Portierung

Da die ursprünglichen Entwickler*innen das Mask R-CNN-Projekt verlassen haben, wird das gesamte Projekt derzeit nicht weiter gepflegt. Hinzu kommt, dass es auf TensorFlow 1.15 basiert, welches im Januar 2021 mit Version 1.15.5 letztmalig ein Update erhalten hat.⁴⁰

Aus diesem Grund wäre eine Portierung auf TensorFlow 2.x (aktuell 2.7) dringend erforderlich. Es gibt einzelne Versuche aus der Community dies zu erreichen, welche aber über rudimentäre Ansätze nicht hinausgekommen sind. Des Weiteren gibt es im Mask R-CNN noch grundsätzlichen Entwicklungsbedarf. So gibt es z. B. die Funktion `batch_slice`, die eine einzelne Batch aufteilt und darauf einzeln die Berechnungen durchführt. Hier wäre es dringend geraten, dass diese Berechnungen ohne eine Aufteilung der Batch funktionieren, da diese nicht Teil des TensorFlow-Graphs ist. Diese Portierung sollte von einem TensorFlow-Experten durchgeführt werden.

2.2. Verbraucherforschung zur Nutzerzentrierung, Verständlich- und Anwendungsfreundlichkeit der Anwendung

Ein weiteres erklärtes Ziel des Forschungsvorhabens bestand darin, die Entwicklung der technischen Anwendung mit einer **kontinuierlichen Nutzerforschung** zu verknüpfen. Konkret sollte die Nutzerzentrierung sowie die empirische Analyse der **Verständlichkeit und Anwendungsfreundlichkeit der prototypischen Anwendung des smarten HeizChecks** durchweg sichergestellt werden, um im Ergebnis einen Service zu entwickeln, der Verbraucher*innen dazu befähigt, in einfacher und intuitiv verständlicher Weise ihren Heizkostenverbrauch zu analysieren und basierend auf dem Ergebnis dieser Analyse ihr Heizverhalten anzupassen.

Ausgehend von dem jeweiligen technischen Entwicklungsstand der prototypischen Anwendung wurden **verschiedene qualitative und quantitative empirische Forschungsmethoden** eingesetzt. Die jeweils erzielten Ergebnisse wurden dann in die Weiterentwicklung der Anwendung gegeben. Durch dieses Wechselspiel wurde die Nutzerzentrierung und -orientierung während des gesamten Forschungsprozess sichergestellt.

Zu Beginn des Forschungsvorhabens wurde ein dezidiertes **Nutzereinbindungskonzept** erarbeitet. Dieses diente dazu, die zur Verfügung stehenden **Methoden** mit den erwarteten **Forschungsbedarfen** zu verknüpfen, erste **Forschungsfragen** zu generieren und eine überblicksartige zeitliche Terminierung der Erhebungen vorzunehmen.

⁴⁰ Vgl.: <https://github.com/tensorflow/tensorflow/releases/tag/v1.15.5>

Für die Konzepterstellung wurden Erkenntnisse und Erfahrungswissen aus dem existierenden Online-Ratgeber HeizCheck von co2online⁴¹ sowie aus relevanter wissenschaftlicher und grauer Literatur zu verschiedenen Befragungsinstrumenten und Nutzereinbindung ausgewertet (Baur & Blasius, 2014; Moser, 2013). Im weiteren Forschungsverlauf diente das Konzept als Benchmark und Orientierungspunkt, um den Projektfortschritt, aber auch etwaige Abweichungen bemessen zu können. Abbildung 9 zeigt exemplarisch den Zeitstrahl der Nutzereinbindung. Dieser wurde kontinuierlich aktualisiert und an den aktuellen Projekt- und Forschungsverlauf angepasst.

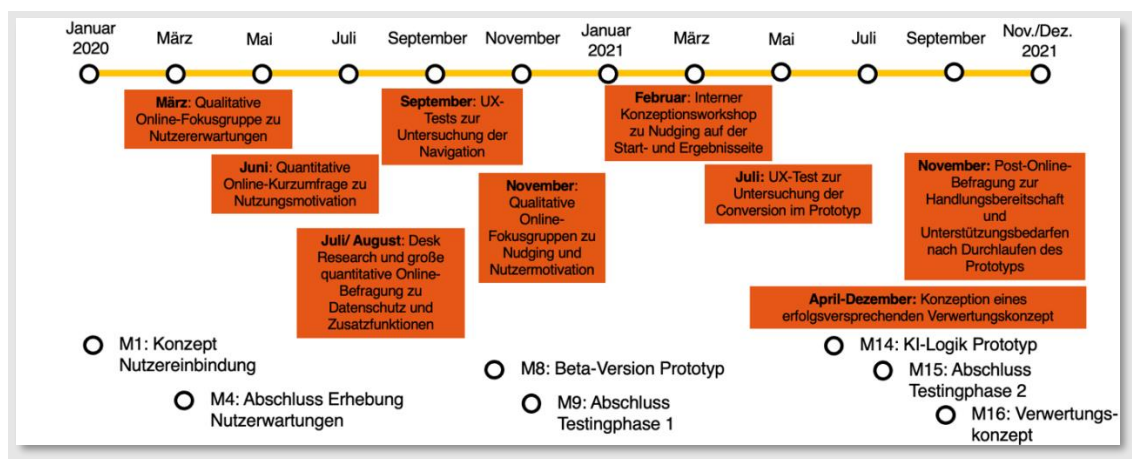


Abbildung 9: Exemplarische Darstellung des Nutzereinbindungskonzept anhand des Projekt-Zeitstrahls

Die schlussendlich durchgeführten empirischen Nutzererhebungen orientierten sich jeweils am Entwicklungsstand der Anwendung und den Forschungsbedarfen, die sich hieraus ergaben.

Zusätzlich war eines der Ziele der Nutzerzentrierung, einen **niedrigschwelligen Zugang** zur Anwendung und damit zur Heizanalyse für alle Verbrauchergruppen, insbesondere verletzte Verbraucher*innen, zu ermöglichen. Aus diesem Grund wurde parallel zur Erarbeitung des Nutzereinbindungskonzepts eine **Literaturrecherche** durchgeführt. Hierbei wurden wissenschaftliche Erkenntnisse zu **Barrierefreiheit und -reduktion**, Nutzungsgewohnheiten für und Anforderungen an **digitale Anwendungen** von **verletzlichen Verbrauchergruppen** zusammengetragen (z. B. Dietrich et al., 2017; Dinter et al., 2014; European Commission, 2016; Kettner & Thorun, 2019). Mit Hilfe dieser Erkenntnisse sollte im Laufe der weiteren Anwendungsentwicklung sichergestellt werden, dass der niedrigschwellige Zugang zum smarten HeizCheck für unterschiedliche Nutzergruppen gewährleistet wird.

⁴¹ Der HeizCheck ist einer von vielen Online-Ratgebern (EnergiesparChecks), die Verbraucher*innen beim Energiesparen unterstützen. co2online betreibt den HeizCheck bereits seit 2001. Vgl.: <https://www.co2online.de/service/energiesparchecks/heizcheck/>

Nachfolgend werden die einzelnen **Forschungsschritte** inklusive der jeweiligen Forschungsfragen und Zieldefinition, des angewandten Forschungsdesigns, der erzielten Ergebnisse sowie der Ergebnisverwendung innerhalb der Anwendungsentwicklung dargestellt. Abschließend erfolgt eine übergeordnete Bewertung des nutzerzentrierten Forschungsansatzes.

2.2.1. Online-Fokusgruppen zur Erhebung der allgemeinen Nutzererwartungen an die Anwendung

Zielstellung der Erhebung

Das Ziel der Online-Fokusgruppen bestand darin, möglichst konkrete Einsichten in **Motivationen und Hürden** beim Thema HeizCheck sowie die **Erwartungen** von Verbraucher*innen an die Nutzung und die notwendigen Funktionen eines smarten HeizChecks zu erhalten.

Zudem sollten erste Erkenntnisse dazu generiert werden, wie Verbraucher*innen dem Thema **Künstliche Intelligenz** begegnen und ob hier bspw. Vorbehalte bestehen. Übergeordnet sollte besprochen werden, was sich Verbraucher*innen allgemein von **Online-Beratungen** erwarten, wie sie angesprochen oder in welchem Stil sie durch die Anwendung geführt werden wollen.

Forschungsmethode

Es wurde die **explorative Methode der Fokusgruppe** gewählt. Aufgrund der pandemischen Situation wurden Online-Fokusgruppen anstelle von physischen Gruppengesprächen durchgeführt.

Zur Vorbereitung der Fokusgruppen fand ein **Kreativ-Workshop** statt, in dem die **Grundstruktur des Prototyps** anhand einer User Journey erarbeitet wurde (s. a. 3.2.1.). Hierfür wurden durch co2online **Personas** von fünf unterschiedlichen Nutzertypen entwickelt, um ein besseres Verständnis der Ausgangslage und Wohnsituation der Nutzer*innen, ihren möglichen Motivationen, eine Analyse der HKA durchzuführen, sowie weiteren Wünschen zu erlangen.

Auf Basis dieser Vorüberlegungen wurde das Fokusgruppenkonzept entwickelt und Teilnehmende an den Fokusgruppen rekrutiert. An den jeweils zweistündigen Gesprächen nahmen neun Personen aus der Allgemeinbevölkerung sowie sieben Nutzende der co2online-Beratungsangebote teil.

Ergebnisse

Kernergebnisse der Fokusgruppen waren:

- Die Motivation, die eigenen Heizkosten zu prüfen, liegt für den Großteil im Wunsch, **Kosten zu sparen** bzw. **Fehler in der Abrechnung zu identifizieren**. Nachhaltigkeit ist für viele ein interessanter Aspekt, wenn auch nachrangig.
- Diejenigen, die noch keine Heizanalyse mittels des bestehenden HeizChecks auf co2online durchgeführt haben, haben die **Sorge**, dass der Vorgang sehr **kompliziert und zeitintensiv** ist. Diejenigen, die bereits einen HeizCheck gemacht haben, finden ihn intuitiv verständlich.
- Die Möglichkeit, einen **Scan der Abrechnung hochladen** zu können, findet **Zuspruch**. Ein Wunsch liegt zudem in der verständlichen und individualisierbaren Darstellung der Ergebnisse.
- **Datenschutzbedenken sind nachrangig**. Jedoch besteht der Wunsch, dass die Adresse auf dem Scan geschwärzt wird.
- Hinsichtlich der Ansprache und Interaktion wird ein **sachliches Design im Formularstil** einem Abfragestil mit „smarten Fragen“ oder einem Chatbot-Stil vorgezogen.
- Insgesamt ist erkennbar, dass die Testpersonen sich wünschen zu erfahren, wo Fehler in der Abrechnung liegen, wie hoch ihr Verbrauch im Vergleich zu anderen ist bzw. wo Einsparpotenziale liegen.

Die Ergebnisse der Fokusgruppen zu Wünschen und Erwartungen der Nutzer*innen flossen anschließend in die Design-Entwicklung (Prototyping) ein (s. a. 3.2.2.).

2.2.2. Quantitative Kurzumfrage zur quantitativen Erhärtung ausgewählter Fokusgruppenergebnisse

Zielstellung der Erhebung

Die Ergebnisse der Fokusgruppen legten nahe, dass Verbraucher*innen den HeizCheck vornehmlich durchführen würden, um Kosten zu sparen. Die Motivation, etwas für den Klimaschutz zu tun und CO₂ einzusparen, war hingegen nachrangig. Um die allgemeine Ausrichtung des smarten HeizChecks festzulegen, sollten diese qualitativen **Erkenntnisse** quantitativ mit einer größeren Stichprobe **erhärtet und validiert** werden. Die quantitative Kurzumfrage diente auch dazu, **weitere Motivationsgründe** für die Durchführung des smarten HeizChecks quantitativ zu erheben.

Forschungsmethode

Zu diesem Zweck wurden zwei **quantitative Online-Kurzumfragen** mit dem Erhebungstool Pinpoll durchgeführt. Diese wurde auf der Webseite von co2online zur Beantwortung

freigeschaltet. Insgesamt nahmen 267 Personen an der ersten Umfrage und 34 Personen an der zweiten Umfrage teil.

Ergebnisse

Die Ergebnisse zeigen, dass die Testpersonen ein **besseres Verständnis** ihrer HKA und ihres Verbrauchs erlangen wollen. Sie wünschen sich, **Einsparpotenziale zu erkennen und Kosten zu senken** sowie ihren **Verbrauch mit anderen Haushalten zu vergleichen**. Klimaschutzaspekte wurden auch als Motivationsgründe genannt, allerdings von einer geringeren Anzahl der Teilnehmenden. Implikation dieser Ergebnisse war, dass im weiteren Projektverlauf ein großes Augenmerk auf das Herausstellen relevanter Vergleiche und möglicher Kosteneinsparungen gelegt werden sollte.

2.2.3. Quantitative Online-Umfrage zum Thema Datenschutz

Zielstellung der Erhebung

Erklärtes Ziel bei der Entwicklung des smarten HeizChecks war es, die Einhaltung der Grundsätze der Datensparsamkeit und der Transparenz von Datenschutz und Datensicherheit kontinuierlich mitzudenken und sicherzustellen.

Vor diesem Hintergrund sollte eine groß angelegte quantitative Umfrage Einblicke in die **Nutzererwartungen und Bedenken bezüglich des Datenschutzes** beim smarten HeizCheck liefern. Zudem sollte die Befragung Aufschluss über **Nutzerpräferenzen bezüglich entwickelter Lösungsansätze** zur Demonstration eines glaubwürdigen und guten Umgangs mit Daten erheben. Zuletzt sollten Nutzungswahrscheinlichkeiten für verschiedene datensparsame Zusatzfunktionen approximiert werden.

Forschungsmethode

Zur Erhebung der Nutzerbedarfe in Bezug auf das Thema Datenschutz wurde eine **Online-Befragung** durchgeführt. Zur Vorbereitung der Erhebung wurde ein **Datenschutz-Prüfkatalog** durch co2online erstellt. Hierin wurden **Datenschutzanforderungen und Anonymisierungs-standards** definiert. Zudem wurden potenziell relevante **Datenschutzherausforderungen** identifiziert und mögliche Lösungsansätze entwickelt.

Die Online-Befragung lief in zwei Wellen. An der Befragung nahmen insgesamt 217 Personen teil. Diese setzten sich zusammen aus Abonnent*innen der Projektinformationen zum Forschungsvorhaben Smart_HEC, bestehenden HeizCheck-Nutzer*innen und Nutzer*innen von anderen co2online-Beratungsangeboten.

Ergebnisse

Übergeordnete Ergebnisse der Befragung waren:

- In Bezug auf unterschiedliche Datenschutzthemen wünschen sich die Befragten **Informationen zur etwaigen Datenweitergabe an Dritte** und zur Art der für die Anwendung benötigten Daten.
- Was die Datenschutzbedenken angeht, so zeigen die Ergebnisse, dass die **Sichtbarkeit der persönlichen Daten auf dem Scan** der HKA **große Bedenken** auslöst, die anonyme Speicherung von Daten hingegen für die Mehrheit unproblematisch ist.
- Als besonders **glaubwürdige Lösungsansätze**, um Vertrauen in den guten Datenschutz der Anwendung zu generieren, werden ein **Hinweis auf das Schwärzen** der persönlichen Daten, **kurze Informationen zum Datenschutz** und eine Erinnerung an den guten Umgang mit den persönlichen Daten zum Zeitpunkt der Dateneingabe gesehen.
- Auch das Abbilden von bestimmten **Siegeln und Kooperationspartnern** kommt bei Nutzenden gut an. So bewerten Befragte besonders die Kooperation mit der Verbraucherzentrale NRW, das SSL-Siegel sowie den Verweis auf einen sicheren Server als vertrauenswürdig.
- Die Befragung ergab zudem, dass das **Zwischenspeichern von Eingaben** und der **jährliche Erinnerungsservice** die **beliebtesten Zusatzfeatures** der Anwendung sind. Auch wünschen sich viele der Befragten datensparsame Alternativen zum Speichern der Ergebnisse und zur Bewertung der Anwendung.

Die Ergebnisse der Datenschutzbefragung flossen in diverse Aspekte der Datenschutzkommunikation in der Anwendung ein. Zudem wurden auch in der Entwicklung durchgehend Grundsätze der Datensparsamkeit beachtet (s. a. 3.2.3.).

2.2.4. UX-Test zur Navigation

Zielstellung der Erhebung

Basierend auf der Gestaltung einiger Seiten des zukünftigen smarten HeizChecks konnte die Interaktion der Nutzenden mit dem Design-Prototypen untersucht werden.

Ziel der UX-Tests war es, möglichst konkrete Einsichten in das **Verständnis der Nutzer*innen der einzelnen Seitenelemente** sowie der **Navigation** zu gewinnen. Auch die **Nutzerführung, Bedienbarkeit und Funktionalität** der Anwendung sollten beurteilt werden. Insbesondere **Schwachstellen der Navigation und Missverständnisse** sollten so aufgedeckt und besprochen werden.

Forschungsmethode

Vor dem Hintergrund der formulierten Zielstellung wurden fünf je einstündige UX-Tests vor Ort bei co2online durchgeführt. Die Testpersonen wurden dabei nach der Methode des Thinking Aloud gebeten, frei ihre Assoziationen und Kritikpunkte während der Interaktion mit den Seiten der Anwendung zu äußern.

Ergebnisse

Die zentralen Ergebnisse der UX-Tests waren:

- Die **Bedienung** der Seiten der Anwendung machte den Testpersonen **Spaß**, das Design wird als ansprechend empfunden.
- Es ergeben sich **diverse Verständnisprobleme**, die zum Teil auf die komplexe Materie, zum Teil jedoch auch auf den Aufbau der Anwendung zurückzuführen sind.
- Aufbau und Funktionalität von Elementen, wie etwa den **verschiedenen Optionen zum Erfassen der HKA** oder die Funktion des **Hochladens und Speicherns**, sind teilweise **unklar**.
- Das mitunter notwendige **Ergänzen von Werten** erweist sich als **kritische Stelle** und die Gefahr des Abbruchs zeigt sich hierbei als hoch.
- Empfehlungen zur weiteren Entwicklung sind zudem die **Einbindung von Hilfe-Features**, um den Abgleich mit dem Scan und das Ergänzen von Werten zu erleichtern.

Die Ergebnisse verdeutlichten, dass Nutzer*innen an vielen Stellen der Anwendung detaillierte Hilfestellungen benötigen. In Folge dieser Ergebnisse wurden die begleitenden Hilfestellungen ausgebaut (s. a. 3.2.2.).

2.2.5. Fokusgruppen zum Thema Handlungsmotivation und Nudging auf der Ergebnisseite

Zielstellung der Erhebung

Das übergeordnete Ziel des smarten HeizChecks ist es, Bürger*innen zum Einsparen von Heizenergie zu motivieren und somit den Klimaschutz voranzutreiben. Um dies zu erreichen, soll die Ergebnisseite mit der Analyse des persönlichen Heizenergieverbrauchs sowie den Tipps so gestaltet sein, dass Personen ins Handeln kommen, d. h., ihr Verhalten anpassen und Heizenergie einsparen. Die Hypothese ist, dass Personen einerseits direkt über **konkrete Tipps zum Handeln motiviert** werden können. Andererseits sollen Verhaltensveränderungen auch indirekt über ein **gesteigertes Wissen** zu den individuellen Verbräuchen und einem gesteigerten Interesse an der Thematik erreicht werden. Prinzipiell sind beide Wege legitim und sollten nebeneinander optimiert werden.

Vor diesem Hintergrund war es Ziel der Fokusgruppen, folgende Fragen zu thematisieren und explorative Einsichten zu generieren:

- **Motiviert die Ergebnisseite** in ihrem aktuellen Stand die Leute dazu, zu handeln? Wenn ja: Warum? D. h., welche Bereiche, Informationen und Aspekte sind dafür verantwortlich? Wenn nein: Was hält die Leute davon ab? Welche Hürden und gefühlten Risiken stehen dem entgegen?
- Ausgehend von Fragestellung 1: Welche **Vorstellungen und Ideen** bestehen bei den Nutzer*innen um (a) die Motivation weiter zu steigern bzw. (b) die Hürden abzubauen? Wie kann die Ergebnisseite ihrer Meinung nach dahingehend optimiert werden?
- Wie wirken sich **verschiedene Darstellungsversionen** der Ergebnisseite bzw. erste Nudging-Ideen auf die Motivation der Nutzer*innen aus, ihr Verhalten zu überdenken und zu verändern?

Forschungsmethode

Zur Beantwortung der oben genannten Forschungsfragen wurde erneut die **explorative Methode der Fokusgruppe** gewählt. Aufgrund der pandemischen Situation wurden Online-Fokusgruppen anstelle von physischen Gruppengesprächen durchgeführt.

Insgesamt nahmen 16 Testpersonen an den vier zweistündigen Gesprächen teil. Die Rekrutierung erfolgte zum einen über einen E-Mail-Verteiler von co2online sowie, zum anderen, über eine ebay-Kleinanzeige. Grundannahme für diese Art der Rekrutierung war, dass es sich bei den co2online-Nutzer*innen um stärker an Heizthemen interessierte Personen, bei den ebay-Teilnehmer*innen eher um einen Querschnitt der Bevölkerung mit geringerem Vorwissen und Interesse an Heizthemen handelt. Die Forschungsfragen konnten somit für beide Zielgruppen beantwortet werden.

Ergebnisse

Die Ergebnisse der Fokusgruppen lassen sich folgendermaßen zusammenfassen:

- Es besteht **Interesse an den inhaltlichen Ergebnissen zur eigenen Wohnung** – besonders, wenn ein klarer Bezug zum eigenen Haushalt hergestellt wird, Kennwerte eindeutig beschriftet sind und relevante Vergleichsgruppen gewählt werden.
- **Klare und frühzeitige Informationen zu Tipps** und Bezug auf (Mieter*innen-)Rechte und Pflichten sind nötig, um die Verständnis- und Bezugshürden zu Ergebnissen zu reduzieren.
- Der **Aufbau der Seite vom Allgemeinen zum Detail** folgt dem natürlichen Informationsverhalten. Auch besteht Interesse daran, Details zu verstehen. Allerdings müssen Tipps früh platziert werden, um die Motivation aufrechtzuerhalten.

- Die **Tipps** (insbesondere noch wenig bekannte Informationen) machen neugierig. Sie sollten **prägnant kommuniziert**, mit einer konkreten Aufforderung und einer **Möglichkeit, Details zu erfahren**, verbunden werden.

Als besonders relevante Motivationshebel erwiesen sich folgende Aspekte:

- **Einsparpotenziale:** Geldeinsparungen sollten hervorgehoben werden, eine Relevanz zeigte sich bereits ab einem Betrag von 80 Euro. Doppelte Einsparpotenziale (CO₂ und Geld) sollten betont werden, um die Bedeutung der Einsparung weiter zu erhöhen.
- **Relativieren:** Einsparpotenziale sollten den Anschaffungskosten, dem erwarteten Aufwand bzw. der Amortisationsdauer gegenübergestellt werden, um das Potenzial für Nutzer*innen erfassbarer zu machen.
- **Einsparrechner:** Die Teilnehmenden wünschen sich eine Darstellung der aufaddierten Einsparpotenziale von unterschiedlichen Maßnahmen, bspw. mit Hilfe eines Rechners, um das Gesamtpotenzial erfassen zu können.
- **Priorisierung:** Die Ergebnisse zeigen, dass individualisierte und zu priorisierende Tipps sowie Möglichkeit zum Filtern Nutzer*innen dabei unterstützen können, ins Handeln zu kommen.

Aufbauend auf den Ergebnissen der Fokusgruppen entstand Ende 2020 ein zweiter Entwurf der Ergebnisseite des smarten HeizChecks (s. a. 3.2.2.).

2.2.6. Kreativ-Workshop zur Anreicherung der Start- und Ergebnisseite mit verhaltenswissenschaftlichen Erkenntnissen

Zielstellung

Abweichend von den bislang angewandten empirischen Erhebungsmethoden diente der interne Workshop der **theoretischen Weiterentwicklung der Ergebnisseite** anhand von verhaltenswissenschaftlichen Erkenntnissen.

Aufbauend auf der **wissenschaftlichen Literatur zum Nudging** und den Erkenntnissen der bislang durchgeführten **Nutzertests** wurden Umsetzungsideen für die Gestaltung der Start- und Ergebnisseite entwickelt. Ziel für die Startseite war es, nach kreativen Wegen zu suchen, um Nutzer*innen dazu zu bewegen, den smarten HeizCheck durchzuführen. Für die Ergebnisseite sollten Vorschläge dazu entwickelt werden, wie Nutzer*innen dazu angeregt werden können, basierend auf den Ergebnissen der Heizkostenanalyse ihr Heizverhalten zu ändern und die gegebenen Tipps umzusetzen.

Methode

Der **Kreativ-Workshop** wurde von ConPolicy organisiert und mit Vertreter*innen von co2online und SEnerCon durchgeführt. Zu Beginn des dreistündigen Workshops stand eine theoretische Einführung in die Potenziale des Nudging-Ansatzes. Basierend hierauf wurden zunächst kritische Bereiche der Start- und Ergebnisseite herausgearbeitet und für diese Lösungs- und Verbesserungsansätze entwickelt.

Ergebnisse

Im Ergebnis wurden die kritischen Bereiche identifiziert und zugehörige Weiterentwicklungswege skizziert.

- Folgende **Barrieren und Lösungsansätze** für die Nutzung der Anwendung wurden auf der **Startseite** identifiziert:
 - Eine mögliche Barriere ist die Wahrnehmung **begrenzter Handlungsmöglichkeiten** der einzelnen Nutzer*innen. Für diese kritische Stelle wurde herausgearbeitet, dass es wichtig sei zu betonen, dass jeder Beitrag wichtig sei und die Nutzer*innen hierfür **keine Expertise benötigen**, da die KI es ihnen erleichtert.
 - Die Gefahr besteht, dass Nutzer*innen Sorge haben, die **Durchführung könne lange dauern** oder sie könnten diese nicht bewältigen. Hierfür wurde die Lösung entwickelt, eine **Vorschau des Prozesses mit Zeitangaben** darzustellen und zu erklären, welche Dinge hierfür jeweils notwendig sind.
 - Eine weitere Barriere besteht darin, dass Nutzer*innen den **Vorteil der Anwendung nicht hinreichend verstehen**. Hierfür wurde die Lösung entwickelt, die **Einsparung und Individualisierung der Beratungsergebnisse** zu betonen.
- Folgende Barrieren und Lösungsansätze für die Nutzung der Anwendung wurden auf der **Ergebnisseite** identifiziert:
 - Die wichtigsten Barrieren bei der vergleichenden Darstellung der Tipps stellt das **fehlende Orientierungsvermögen** der Nutzer*innen dar und die Gefahr, nicht die geeigneten Tipps zu identifizieren und daher gar nicht erst weiter auf die Detailseiten der Tipps zu klicken. Daher wurde die Lösung einer **vergleichenden Darstellung der Tipps** auf der Ergebnisseite erarbeitet. Die Darstellung sollte einen intuitiven Vergleich der Kosten zu den Einsparpotenzialen der einzelnen Tipps ermöglichen. Das Wording in diesem Bereich sollte überarbeitet werden, sodass Nutzer*innen stärker dazu animiert werden, die

Detailansicht des Tipps anzuklicken. Die Einsparpotenziale sollten möglichst konkret in Euro angegeben werden. Zudem sollten die Tipps individualisiert sein.

- Die Gefahren auf den Detailseiten der Tipps bestehen darin, dass die Nutzer*innen **Sorge vor Kosten oder Aufwand** haben könnten, während sie das Potenzial nicht erkennen. Daher wurde erarbeitet, auch hier eine **vergleichende Darstellung von Aufwand zu Kosten** zu verwenden. Die Betonung sollte auf möglichen Einsparpotenzialen liegen.

Im Nachgang wurde ein zwei-stufiges Intro-Overlay für die Anwendung entwickelt, in der ein Augenmerk auf die Vorschau des weiteren Prozesses gelegt wurde. Darüber hinaus wurde eine vergleichende Darstellung der Tipps als Liste entwickelt (s. a. 3.2.2.).

2.2.7. UX-Test zum Thema Conversion

Zielstellung der Erhebung

Aufbauend auf der programmierten Beta-Version der Anwendung sollten UX-Tests Hinweise dahingehend liefern, ob Nutzer*innen die Beta-Version des smarten HeizChecks bis zum Ende ohne Abbrüche durchführen können (i. e. **Conversion**), ob die Nutzerführung und Bedienbarkeit intuitiv verständlich sind und wo gegebenenfalls kritische Stellen für einen potenziellen Abbruch bestehen.

Forschungsmethode

Vor dem Hintergrund der formulierten Zielstellung wurden **sieben je einstündige UX-Tests** durchgeführt. Die Teilnehmer*innen wurden über einen E-Mail-Verteiler von co2online rekrutiert. Es nahmen Mieter*innen und Wohnungseigentümer*innen (50 % Männer und 50 % Frauen) im Alter von 30 bis 70 Jahren an den Tests teil. Es wurden Digital Natives und Not-Natives sowie Expert*innen und fachfremde Personen zu Heizthemen (laut Selbsteinschätzung) zu gleichen Teilen rekrutiert, um Unterschiede in der Conversion dieser Gruppen zu überprüfen.

Ergebnisse

In den Tests gelang es allen Nutzer*innen, den smarten HeizCheck bis zum Ende durchzuführen. Der Kenntnisstand bezüglich Heizthemen hatte hierbei keinen Effekt, jedoch **beeinflusste die Digitalaffinität** sowohl die **Conversion** als auch die **Motivation der Testpersonen** maßgeblich. Es zeigte sich, dass der **Prozess des Fotografierens bzw. Hochladens** der Abrechnung sowie die **Eingabe fehlender Werte herausfordernde Stellen** in der Nutzung darstellen. Zudem wurden Optimierungspotenziale bei der Einführung in den smarten

HeizCheck sowie des Bearbeitungsflusses ersichtlich. Übergeordnet lobten die Testpersonen das Design und die Ergebnisdarstellung und -tiefe der Anwendung.

Die Ergebnisse hatten einige Veränderungen in der Anwendungsentwicklung zur Folge. So wurde u. a. der QR-Code zur Erfassung der HKA ersetzt (s. a. 3.2.3. und 3.2.4.).

2.2.8. Quantitative Online-Umfrage zur Umsetzungswahrscheinlichkeit von Handlungsempfehlungen

Zielstellung der Erhebung

Im Fokus der abschließenden empirischen Nutzererhebung stand die übergeordnete Frage, ob und, wenn ja, in welchem Maße Nutzer*innen ausgehend von den Ergebnissen des smarten HeizChecks **Veränderungen in ihrem Heizverhalten** anstoßen bzw. welche **Hürden** sie bei der Umsetzung von konkreten Empfehlungen sehen.

Konkret sollten die folgenden Fragen beantwortet werden:

- **Wie intensiv beschäftigen** sich Nutzer*innen mit den **Ergebnissen und Empfehlungen**? Sind Nutzer*innen im Anschluss an das Durchlaufen des smarten HeizChecks **motiviert, ins Handeln zu kommen** und wie wahrscheinlich ist es, dass sie dies tatsächlich in nächster Zeit tun?
- Welchen **Empfehlungen** würden die Nutzer*innen am ehesten nachkommen und welche **Hürden** halten sie von der Umsetzung (einzelner Maßnahmen) ab?
- Welche **Unterstützungsangebote** wünschen sich Nutzer*innen und welche Angebote und Funktionen finden sie attraktiv?

Forschungsmethode

Zur Beantwortung der Forschungsfragen wurde eine **quantitative Online-Umfrage** programmiert. Der Link zur Umfrage wurde Nutzer*innen nach der Durchführung des smarten HeizChecks zur Verfügung gestellt. Hierdurch konnte sichergestellt werden, dass die Testpersonen die Anwendung und insbesondere die Ergebnisseite gut in Erinnerung hatten und beides gut und realistisch einschätzen und bewerten konnten. Insgesamt nahmen 20 Testpersonen an der Umfrage teil. Darunter waren sowohl Mieter*innen als auch Wohnungseigentümer*innen sowie Personen mit bestehendem Bezug zu co2online als auch ohne Bezug. Aufgrund der sehr geringen Stichprobengröße sind die Ergebnisse wenig belastbar und müssen unter Vorbehalt interpretiert werden.

Ergebnisse

Die Ergebnisse der quantitativen Online-Befragung lassen sich folgendermaßen zusammenfassen:

- Die Ergebnisse zeigen, dass eine **intensive Beschäftigung mit den Ergebnissen nicht die Regel** ist (etwa die Hälfte der Befragten bemerkt das PDF, davon gibt die Hälfte der Personen an, das PDF tatsächlich heruntergeladen zu haben).
- Ein Großteil der Befragten zeigt sich inhaltlich wenig überrascht und **zufrieden mit den Ergebnissen**. Allerdings sieht über ein Drittel der Befragten auch keinen Handlungsbedarf.
- Die **Empfehlungen/Tipps** auf der Ergebnisseite werden **von allen Befragten wahrgenommen**, allerdings werden sie von vielen nur **überflogen**. Obwohl die Tipps von fast allen als **hilfreich bewertet** werden, sind nur wenige bereit in die zeitnahe Umsetzung zu gehen. Hinsichtlich der **Umsetzungswahrscheinlichkeit konkreter Tipps** zeigt sich, dass Empfehlungen zu „Smarten Thermostaten“ und „Duschen statt Baden“ von vielen umgesetzt würden.
- Eine Checkliste oder Anleitung mit Schritten zur Umsetzung der Tipps sowie mehr Beispiele zur Umsetzung von Tipps sind die **meistgewünschten Angebote zur Unterstützung** auf der Ergebnisseite.
- Hinsichtlich der Einbindung von **Informationen zum CO₂-Minderungspotenzial** zeigt sich, dass die Hälfte der Befragten CO₂-Informationen wahrnimmt, für einige diese Information jedoch wenig relevant ist.
- Fazit der Befragung ist, dass die eher passive Beschäftigung mit der Ergebnisseite noch **keinen unmittelbaren Schluss auf eine Verhaltensveränderung** zulässt. Allerdings lassen die Befragungsdaten aufgrund der sehr kleinen Stichprobengröße keine validen Aussagen zu und sie sollte daher idealerweise wiederholt werden.

Die offenen Aussagen der Nutzer*innen zeigen auch, dass die **Probleme noch eher zu Beginn der Anwendung liegen**. Daher sollte die Priorität nach Abschluss des Projekts darauf liegen, die Erfassung und Eingabe zu optimieren. Zwei Optimierungspotenziale auf der Ergebnisseite wurden darin identifiziert: (1) sicherzustellen, dass sich Nutzer*innen weiterhin mit den Ergebnissen auseinandersetzen und (2) die Tipps noch besser auf deren individuelle Wohn- und Heizsituation zuzuschneiden.

2.2.9. Übergeordnete Bewertung des nutzerzentrierten Forschungsansatzes

Ein Ziel des Forschungsprojekts war es, Nutzer*innen ein virtuelles Assistenzsystem an die Hand zu geben, das an ihren Bedürfnissen an eine Online-Beratung ausgerichtet ist, die

Bewertung der eigenen HKA maßgeblich vereinfacht und diese Informationen in verständlicher und transparenter Weise ausgibt sowie personalisierte Handlungsempfehlungen ausgibt. Vor diesem Hintergrund wurde eine kontinuierliche und bedarfsorientierte Nutzerzentrierung im Forschungsprojekt Smart_HEC umgesetzt.

Hierzu kann übergeordnet festgehalten werden, dass die **konstante Einbindung und Befragung** der Verbraucher*innen die Entwicklung und Ausgestaltung der Anwendung an mehreren Stellen maßgeblich beeinflusste und lenkte. Somit konnte gewährleistet werden, dass der Entwicklungsprozess nicht an den Bedürfnissen, Wünschen, Perspektiven, aber auch Herausforderungen der Verbraucher*innen vorbei verlief. Dies wiederum ist entscheidend, wenn es um die Bewertung der **Akzeptanz der Anwendung** am Markt geht. Ein weiterer entscheidender Bestandteil der Nutzererhebung war zudem die Untersuchung der **Verständlichkeit** von Informationen. Hierbei zeigte sich, dass Nutzer*innen ein sehr unterschiedlich ausgeprägtes Vorwissen und Versiertheit bezüglich des verwendeten Fachvokabulars haben. Die Erstellung der Texte für die Anwendung orientierte sich daher stets an den Ansprüchen an Verständlichkeit und bemühte sich um die Verwendung einfacher Sprache. Zuletzt wurden Handlungsempfehlungen so entwickelt, dass sie auf den individuellen Bewertungsfall abzielen und in diesem Sinne **personalisiert** sind. Dieses Ziel kann somit als erfüllt bewertet werden.

Weiterhin sollte ein **niedrigschwelliges Angebot** entwickelt werden, das von möglichst vielen Verbrauchergruppen, besonders auch von **verletzlichen Verbraucher*innen**, genutzt wird bzw. werden kann. Diesem Anspruch wurde durch den gesamten Forschungsprozess hindurch versucht, zu genügen. Allerdings zeigten die Erhebungen wiederholend, dass die Ansprüche von Verbraucher*innen mit Vorbildung zum Thema Heizen bzw. solchen, die sich sehr für das Thema interessieren, und solchen, die eher zufällig auf den smarten HeizCheck stoßen, mitunter sehr weit auseinander lagen. Während sich die einen Profi-Tipps und Expertenmeinungen wünschten, lag bei den anderen eine Präferenz darauf, nur die unbedingt notwendigen Informationen in möglichst einfacher Art und Weise zu erhalten. Die Schwierigkeit bestand demnach darin, eine prototypische Anwendung zu entwickeln, die entweder alle Verbrauchergruppen gleichermaßen anspricht oder sich auf eine bestimmte Zielgruppe konzentriert und den smarten HeizCheck entsprechend ihrer Bedürfnisse zu optimieren. Schlussendlich wurde mit dem zweiten Entwurf der Ergebnisseite eine Variante entwickelt, in der zum einen weniger versierte und interessierte Nutzer*innen **schnell einen Überblick bekommen** (Zusammenfassung der Analyseergebnisse zu Beginn der Ergebnisseite, wiederholte Platzierung von Fazits). Zum anderen erhalten Nutzer*innen mit Detailinteresse **zusätzliche Informationen** zur Aufschlüsselung der Bewertung (in etwa zum Warmwasser und zum Heizen) sowie den Berechnungsgrundlagen bei Bedarf ebenfalls auf

der Ergebnisseite. Dafür wurden Zusatzinformationen in Form von Tooltips integriert. Durch dieses Vorgehen wurde dem Anspruch genügt, **möglichst allen Verbrauchergruppen ein adäquates und verständliches Analyseergebnis zu präsentieren.**

Ferner zielte das Forschungsprojekt darauf ab, dass mit dem smarten HeizCheck eine Anwendung entsteht, die nicht nur den gesetzlichen **Vorgaben zum Datenschutz** genügt, sondern die – darüber hinaus – durch eine **hohe Transparenz in Bezug auf den Umgang mit personenbezogenen Daten** Vertrauen bei den Nutzer*innen ausruft und sie in ihrer informationellen Selbstbestimmung stärkt. Diesem Ziel wurde gemäß des nutzerzentrierten Forschungsansatzes dadurch begegnet, dass Nutzer*innen kontinuierlich nach ihren **Datenschutzpräferenzen** befragt wurden und auf diese Präferenzen in der Entwicklung entsprechend eingegangen wurde (s. a. 3.2.3.).

3. Anwendungsentwicklung

Die Entwicklung der prototypischen Anwendung inklusive des Rechenkerns zur Bewertung der HKA erfolgte unter Einsatz **agiler und nutzerzentrierter Methoden**. So wurde die Anwendung iterativ und unter Einbeziehung regelmäßigen Nutzerfeedbacks entwickelt und getestet. Nachfolgend wird dargestellt, wie die daraus gewonnenen Erkenntnisse in die Entwicklung der Entwürfe und die Umsetzung des Prototyps eingeflossen sind und so die Anwendung „der smarte HeizCheck“ entstand.

3.1. Berechnung und Bewertung

Im Rahmen dieses Projekts wurde unter anderem eine **Programmierschnittstelle (API, Web-Schnittstelle) und ein Rechenkern zur Berechnung von Kennwerten** entwickelt, die eine differenzierte Bewertung von HKAs zu den Aspekten Wohnungsverbrauch, Energiekosten und Dienstleistungen sowie dem Zustand des Gebäudes ermöglichen. Im folgenden Abschnitt wird dazu zunächst auf die allgemeinen strukturellen und inhaltlichen Voraussetzungen von HKAs und deren Informationsgehalt eingegangen. Im Anschluss folgt die Beschreibung der technischen Implementierung der API sowie die angewendeten Methoden zur Bewertung der einzelnen Verbräuche und Kosten.

3.1.1. Inhaltliche Betrachtung einer HKA

Allgemein

In zentral beheizten Gebäuden ist die Erstellung einer HKA in Deutschland verpflichtend. Die rechtliche Grundlage dafür bildet seit 1981 die **Heizkostenverordnung (HeizkostenV)**. Sie

ist für alle Haus- und Wohnungseigentümer*innen verbindlich und regelt die verbrauchsabhängige Abrechnung der Heiz- und Warmwasserkosten.

In der HKA sind alle im Abrechnungszeitraum angefallenen **Heiz- und Heiznebenkosten sowie der angefallene Energieträgerverbrauch bzw. die Menge an gelieferter Wärme** aufgeführt. Zu den abrechenbaren Heiznebenkosten gehören Wartungskosten der Heizung, Gebühren der Heizkostenverteilung und für den/die Schornsteinfeger*in sowie die Betriebsstromkosten der Heizungsanlage. Die Kosten für die Verbrauchsmessung und für das Erstellen der Abrechnung gehören ebenfalls zu den Heiznebenkosten. Alle anderen im Abrechnungszeitraum angefallenen Kosten gehören zu den Betriebskosten und werden meist separat abgerechnet. In der HeizkostenV ist geregelt, dass mindestens 50 % der anfallenden Heizkosten verbrauchsabhängig abgerechnet werden müssen. Daraus ergibt sich die Notwendigkeit, dass die Verbräuche entsprechend für das Gebäude und die einzelnen Wohnungen gemessen werden müssen.

Der **Gesamtheizenergieverbrauch des Gebäudes** ergibt sich aus der im Abrechnungszeitraum verbrauchten Energieträgermenge (z. B. Erdgas oder Heizöl) beziehungsweise der Menge an gelieferter Wärme (Fern- und Nahwärme). Der Anteil, der dann auf die Erzeugung des Warmwassers entfällt, muss gemäß HeizkostenV seit 01.01.2014 über einen separaten Wärmemengenzähler erfasst werden. In der Praxis sind auch heute noch Abrechnungen zu finden, in denen die angefallene Warmwassermenge des Gebäudes klassisch in Kubikmetern erfasst wird. In diesen Fällen muss die zur Warmwassererzeugung eingesetzte Energiemenge gemäß den Vorgaben aus der HeizkostenV berechnet werden. Der Anteil an Heizenergie, der auf die Raumwärme entfällt, ergibt sich aus der Differenz zwischen dem Gesamtverbrauch und dem gemessenen beziehungsweise berechneten Warmwasseranteils, er wird also nicht separat gemessen.

In den Wohnungen wird die Raumwärme über **Heizkostenverteiler** an jedem Heizkörper erfasst. Hier haben sich mittlerweile hochauflösende digitale Zähler gegenüber den alten Verdunster-Röhrchen durchgesetzt. Sie messen die vom Heizkörper abgegebene Wärme nicht nur deutlich genauer, sondern sind oft auch schon fernauslesefähig, was die Anwesenheit der Bewohner*innen bei der Auslesung überflüssig macht. Der Warmwasserverbrauch wird in den Wohnungen überwiegend über klassische Kubikmeter-Zähler je Zapfstelle gemessen.

Mit all diesen Messwerten, die sich auf der HKA wiederfinden lassen, lässt sich somit der Verbrauch sowie die Heizkosten des Gebäudes als auch der Verbrauch der Wohnung im Verhältnis zum Rest des Gebäudes berechnen und bewerten. Dies schafft die Grundlage für eine **differenzierte Beurteilung der verschiedenen Handlungsfelder**. Ist der Gesamtenergieverbrauch des Gebäudes zu hoch? Wird der Energieträger und die Dienstleistungen gemäß

des Wirtschaftlichkeitsgebots eingekauft? Liegt Einsparpotenzial im Nutzer*innenverhalten in der Wohnung? Im folgenden Abschnitt wird jeweils auf die genannten Handlungsfelder und die dafür nötigen Werte der HKA eingegangen.

Daten zum Gebäude

Aus einer klimaschutztechnischen Betrachtung heraus sind die Gebäudeverbrauchsdaten am interessantesten. **Wärmetechnische Modernisierungen** im deutschen Gebäudebestand sind dringend nötig. Über den Umweg der Wohnungseigentümer*innen und Mieter*innen soll hier eine Brücke zu den für das Gebäude Verantwortlichen und damit handlungsfähigen Institutionen beziehungsweise Personen geschlagen werden. Aussagen zum allgemeinen sanierungstechnischen Zustand lassen sich auf Basis weniger Kennzahlen zum Gebäude grob ableiten und mit ähnlichen Gebäuden vergleichen.

Anhand der Postleitzahl und dem Abrechnungszeitraum werden dem Gebäude die jeweils lokal vorherrschenden Witterungsdaten zugeordnet, die über den Deutsche Wetterdienst per Schnittstelle zur Verfügung gestellt werden. Beim Abrechnungszeitraum spielt es keine Rolle, ob es sich dabei um exakt ein Kalenderjahr handelt. Wichtig ist lediglich, dass eine vollständige Heizperiode und die Sommerphase enthalten sind. Die so ermittelten Daten dienen als Basis für eine Witterungsbereinigung, die in der **Berechnung des gebäudespezifischen Energieverbrauchskennwertes (EVKW)** verwendet wird. Mit Hilfe des EVKWs lässt sich das Gebäude mit ähnlichen Gebäuden an beliebigen anderen Standorten in Deutschland vergleichen.

Das bei den Nutzer*innen der Anwendung abgefragte Gebäudealter dient der Einordnung des Gebäudes in eine spezifische **Baualtersklasse**. Die Baualtersklassen orientieren sich an den Wärmeschutz- und Energieeinsparverordnung ab 1977 beziehungsweise der zeitlichen Phase vor deren Einführung. Die Beurteilung des Gebäudeverbrauchs findet dann später auf Basis der durchschnittlichen Verbräuche im Gebäudebestand der jeweiligen Baualtersklasse statt.

Die **beheizte Wohnfläche** des Gebäudes ist notwendig, um den Heizenergieverbrauch pro Quadratmeter ermitteln zu können. Nur so lässt sich der Verbrauch mit Gebäuden anderer Größe vergleichen. Die Angabe zur Gebäudefläche ist Teil der HKA und befindet sich im Abschnitt der Heizkostenverteilung. Die Grundkosten sowohl für Raumwärme als auch Warmwasser (sofern vorhanden) erfolgt über die jeweiligen Flächenanteile der Wohnung im Verhältnis zum Gesamtfläche des Gebäudes. Daher sind beide Angaben auszuweisen.

Der im Gebäude zum Einsatz kommende **Energieträger** und dessen Einheit, in der der Verbrauch gemessen wird, sind weitere essenzielle Informationen für die Bewertung des Gebäudes. Für die Berechnung des EVKW ist es nötig, die im Energieträger enthaltene

Energiemenge zu berechnen. Dazu verfügen alle im Bereich der Wärmeerzeugung zum Einsatz kommende Energieträger über jeweils spezifische Umrechnungsfaktoren, über die sich die erzeugte Wärmeenergie in kWh berechnen lässt. Die durch den Einsatz des verbrauchten Energieträgers **angefallenen CO₂-Emissionen** ergeben sich aus dem für den Energieträger festgelegten CO₂-Faktor, der in Kilogramm je kWh angegeben wird. Zur Ermittlung der Gesamtmenge an Emissionen im Abrechnungszeitraum wird der Energieträgerverbrauch in kWh umgerechnet und mit dem spezifischen CO₂-Faktor multipliziert.

In nicht jedem Gebäude erfolgt die Warmwasserbereitung auch über die zentrale Heizanlage, sondern dezentral in den Wohnungen über **Durchlauferhitzer oder Boiler**. Für solche Gebäude regelt das Gebäudeenergiegesetz, kurz GEG (ehemals Energieeinsparverordnung, kurz EnEV), im Rahmen der Berechnung des EVKW, dass ein Schätzanteil für die Warmwasserverzeugung berücksichtigt werden muss. Dadurch wird ermöglicht, dass die Heizenergieverbräuche von Gebäuden mit oder ohne zentraler Warmwasserbereitung miteinander vergleichbar sind.

Ein Schätzanteil für die zur Warmwasserbereitung nötigen Energie kommt auch dann zum Einsatz, wenn im Gebäude die Warmwasserbereitung zwar zentral erfolgt, aber eine Messung im Einzelfall nicht möglich beziehungsweise unzumutbar ist. Erfolgt die Messung des Warmwassers, ist die zu ermittelnde Energiemenge abhängig von der jeweiligen Einheit in der gemessen wird. Im Idealfall geschieht dies bereits mittels direkter Erfassung der Energiemenge über einen separaten Wärmemengenzähler. Diese Messung ist seit 01.01.2014 vorgeschrieben, wird aber immer noch nicht konsequent umgesetzt. Wird stattdessen weiterhin nur die erzeugte Warmwassermenge nur in Kubikmetern gemessen, muss die darauf entfallende Energiemenge zur Erwärmung berechnet werden. Die dafür zu verwendende Formel ist in der HeizkostenV beschrieben. Um mit dieser Formel rechnen zu können, ist neben dem Warmwasserverbrauch zusätzlich die Information über die im Heizsystem eingestellte Warmwassertemperatur nötig.

Daten zur Wohnung

Neben der Bewertung des Gebäudes ist für die Nutzer*innen eine **Einordnung ihres individuellen Haushalts** essenziell. Änderungen im eigenen Nutzungsverhalten lassen sich meist sofort und ohne größere Hürden umsetzen, das meiste davon ohne oder zumindest nur mit geringen Investitionskosten. Um den jeweiligen Verbrauch der Wohnung in der Anwendung des smarten HeizChecks berechnen und bewerten zu können, sind mehrere Kennwerte nötig, die entweder direkt aus der HKA extrahiert oder zusätzlich direkt bei den Nutzer*innen in der Anwendung separat abgefragt werden müssen (s. a. 3.1.2.).

Zum Beispiel spielt die **Lage der Wohnung** innerhalb des Gebäudes für die Bewertung des Raumwärmebedarfs eine wichtige Rolle. Wohnung in exponierter Lage benötigen in der Regel mehr Heizenergie für Raumwärme als Wohnungen, die rundherum von anderen beheizten Wohnungen umgeben sind. Exponiert meint hier beispielsweise Wohnungen im Erdgeschoss oder im Dachgeschoss beziehungsweise im obersten Geschoss, wenn das Dach nicht ausgebaut ist oder ein Flachdach vorliegt. Ebenso exponiert sind Wohnungen mit überdurchschnittlich viel Außenwandfläche, zum Beispiel bei einer Außenlage in einem Reihenendhaus. Um die Lage der Wohnung später bei der Bewertung berücksichtigen zu können, wird bei den Nutzer*innen abgefragt, in welcher Etage die eigene Wohnung liegt und über wie viele Außenwänden sie verfügt.

Der **Heizenergieanteil, der in Wohnungen auf die Raumwärme entfällt**, wird nur in den seltensten Fällen direkt über einen wohnungszentralen Wärmemengenzähler gemessen. Am weitesten verbreitet ist die Erfassung von Verbrauchseinheiten je Heizkörper. Die Summe aller für die Wohnung angefallenen Verbrauchseinheiten sowie die Gesamtmenge aller Verbrauchseinheiten im gesamten Gebäude sind in der HKA im Abschnitt der Verbrauchskostenverteilung für die Heizung (Raumwärme) ausgewiesen. Zur Berechnung des Wohnungsverbrauch (Raumwärmeanteil) werden beide Werte im Verhältnis zu den jeweiligen Flächenanteilen gegenübergestellt:

```
$verhaeltnis_faktor =  
($wohnung_verbrauchseinheiten / $gebaeude_verbrauchseinheiten) /  
($wohnung_flaeche / $gebaeude_flaeche)
```

Der so gewonnene Verhältnis-Faktor aus Verbrauchseinheiten und Flächen wird mit dem auf den Quadratmeter bezogenen Raumwärmeanteil der Heizenergie des Gebäudes multipliziert. Das Ergebnis entspricht dem berechneten Heizenergieverbrauch für Raumwärme der Wohnung. Um schlussendlich den **Gesamtheizenergieverbrauch der Wohnung** zu erhalten, wird noch der auf die Warmwasserbereitung entfallende Anteil an Heizenergie für die Wohnung addiert.

Zur Ermittlung des **Heizenergieanteils, der im Abrechnungszeitraum für die Erzeugung des Warmwassers in der Wohnung angefallen ist**, wird die über den üblichen Kubikmeter-Zähler gemessene Warmwassermenge mit Hilfe der in der HeizkostenV vorgeschriebenen Berechnungsformel ermittelt.

Statistische Auswertungen zum Thema Warmwasserverbrauch in Haushalten haben gezeigt, dass der auf den Quadratmeter bezogene Heizenergiebedarf für die Warmwasserbereitung nicht hinreichend zum Vergleich genutzt werden kann, da der Warmwasserverbrauch deutlich

mehr von der Anzahl an Personen im Haushalt abhängt als von der jeweiligen Wohnungsfläche.

Daher wird im Rahmen der Anwendung die Information zur jeweiligen **Haushaltsgröße** beziehungsweise Bewohneranzahl zusätzlich in der Anwendung abgefragt. Darüber lässt sich zusätzlich der **Pro-Kopf-Verbrauch an Warmwasser** je Bewohner*in im Haushalt berechnen und im Empfehlungsteil belastbarer bewerten.

Strukturelle Herausforderungen von HKAs

Jenseits von inhaltlichen Vorgaben, welche Angaben mindestens auf einer HKA ausgewiesen sein müssen, gibt es **keine gesetzlichen Vorgaben, wie das Layout der einzelnen Seiten aussehen muss**. Auch bei den **Begrifflichkeiten** zu einzelnen Kennwerten besteht seitens der MDL gestalterischer Spielraum. Diese Faktoren machen es schwer, messdienstleisterübergreifend HKAs korrekt zu lesen und zu verstehen. Weiterhin ist es den MDL beziehungsweise Hausverwaltungen freigestellt, die Heizkosten kombiniert mit den Betriebskosten in einer Rechnung abzurechnen oder auf zwei separate Abrechnungen aufzuteilen. Bei **kombinierten Abrechnungen** kann es zu Überschneidungen vor allem bei den Heiznebenkosten kommen. So tritt es in der Praxis immer wieder der Fall auf, dass z. B. die Kosten für die Messtechnik zur Erfassung der Raumwärme- und Warmwasserverbräuche innerhalb der Betriebskosten und nicht über die Heizkosten abgerechnet werden. In solchen Abrechnungen fehlen diese Kosten somit im Kostenblock der Heiznebenkosten und verfälschen damit die spätere Bewertung. Auch sind die Begrifflichkeiten im Bereich der Heiznebenkosten oft nicht eindeutig zuzuordnen. Sind die aufgeführten Kosten laufende Wartungskosten oder versteckte Reparaturkosten, die nicht auf diesem Weg abgerechnet werden dürften?

Einige MDL integrieren auch die Beschaffungskosten für Trinkwarmwasser in die Heizkosten zur Warmwassererwärmung. Dies erfolgt dann unter der Bezeichnung **„Kaltwasser für Warmwasser“** und ist nicht immer leicht zu erkennen. Diese Kosten müssen für die spätere Bewertung der Heizkosten des Gebäudes abgezogen werden, da sie sonst zu einer stark verfälschten Bewertung führen würden.

Auch die **Verteilung der Grundkostenanteile** jeweils für Raumwärme und (sofern vorhanden) Warmwarmwasser erfolgt nicht auf allen HKAs über eine identische Gebäudeflächenangabe. Hier wird seitens der MDL für Raumwärme über die kleiner bemessenen „beheizten Wohnfläche“ des Gebäudes umgelegt wohingegen für die Umlage der Warmwassergrundkosten dann die gesamte Wohnfläche des Gebäudes erfolgt. Für die Anwendung des smarten HeizChecks wurde festgelegt, in solchen Fällen immer die Flächenangabe für die Raumwärmeumlage zu verwenden.

3.1.2. Datenbasis der Heizkostenanalyse

Die folgende Tabelle gibt einen Überblick über die Kennwerte, die für die Heizkostenanalyse im Rahmen des smarten HeizChecks abgefragt bzw. mit Hilfe der KI-Komponenten extrahiert werden:

Index	Parameter	in HKA?	KI-Extraktion?	annotiert?
1	MDL	ja	ja	nein
2	Postleitzahl	nein	-	-
3	Gebäudetyp	nein	-	-
4	Gebäudefläche	ja	ja	ja
5	Gebäudebaujahr	nein	-	-
6	Heizungstyp	nein	-	-
7	Energieträger	ja	ja	nein
8	Warmwasserbereitung Art	nein	-	-
9	Warmwasserbereitung Messung	nein	-	-
10	Warmwasser Temperatur	ja	ja	ja
11	Wohnungsfläche	ja	ja	ja
12	Wohnungslage Etage	nein	-	-
13	Wohnung Anzahl Außenwände (vereinfacht)	nein	-	-
14	Bewohneranzahl der Wohnung	nein	-	-
15	Abrechnungszeitraum (Start & Ende)	ja	ja	ja
16	Nutzungszeitraum (Start & Ende)	ja	ja	ja
17	Verbrauch Energieträger	ja	ja	ja
18	Verbrauch Energieträger Einheit	ja	ja	ja

Index	Parameter	in HKA?	KI-Extraktion?	annotiert?
19	Verbrauch Energieträger Kosten	ja	ja	ja
20	Warmwasser Verbrauch	ja	ja	ja
21	Warmwasser Verbrauch Einheit	ja	ja	ja
22	Warmwasser Verbrauch Wohnung	ja	ja	ja
23	Warmwasser Verbrauch Wohnung Einheit	ja	ja	ja
24	Heizkosten des Gebäudes Gesamt	ja	ja	ja
25	Heizkosten des Gebäudes Raumwärme	ja	ja	ja
26	Heizkosten des Gebäudes Warmwasser	ja	ja	ja
27	Prozentanteil Grundkosten für Raumwärme	ja	ja	ja
28	Prozentanteil Grundkosten für Warmwasser	ja	ja	ja
29	Verbrauchseinheiten des Gebäudes	ja	ja	ja
30	Verbrauchseinheiten der Wohnung	ja	ja	ja
31	Grundkosten Raumwärme Gebäude	ja	ja	ja
32	Grundkosten Warmwasser Gebäude	ja	ja	ja
33	Prozentanteil Verbrauchskosten Raumwärme	ja	ja	ja
34	Prozentanteil Verbrauchskosten Warmwasser	ja	ja	ja
35	Verbrauchskosten Raumwärme Gebäude	ja	ja	ja
36	Verbrauchskosten Warmwasser Gebäude	ja	ja	ja

Index	Parameter	in HKA?	KI-Extraktion?	annotiert?
37	Heizkosten der Wohnung Gesamt	ja	ja	ja
38	Grundkosten Raumwärme der Wohnung	ja	ja	ja
39	Verbrauchskosten Raumwärme der Wohnung	ja	ja	ja
40	Grundkosten Warmwasser der Wohnung	ja	ja	ja
41	Verbrauchskosten Warmwasser der Wohnung	ja	ja	ja
42	Kosten des Kaltwassers für Warmwassererzeugung	ja	nein	nein
43	Betriebsstromkosten	ja	ja	ja
44	Wartungskosten der Heizung	ja	ja	ja
45	Messgerät Art	nein	-	-
46	Messgerät Auslesen per Funk	nein	-	-
47	Messgerät gemietet	nein	-	-
48	Mietkosten der Messgeräte	ja	ja	ja
49	Messkosten Dienstleister	ja	ja	ja

Tabelle 4: Auflistung der nötigen Kennwerte für die Heizkostenanalyse

Alle nicht direkt in der HKA enthaltenen Kennwerte und diejenigen, die sich nicht geeignet extrahieren ließen, werden in der Anwendung direkt bei den Nutzer*innen abgefragt. **Auf die Extraktion entfallen insgesamt 36 Einzelwerte, von denen 34 über die Annotation der Trainingsdaten gelernt wurden.** Der MDL sowie der Energieträger werden über eine **Volltextsuche**, also ohne Annotation, erkannt.

3.1.3. Technische Umsetzung

Datenübertragung (Webservice)

Alle für die Berechnung und Bewertung nötigen Eingabewerte werden in der Anwendung des smarten HeizChecks in der Nutzeroberfläche gesammelt und zur Weiterverarbeitung

bereitgestellt. Für die Übertragung dieser Daten und den späteren Ergebnissen wurde durch SEnerCon **ein spezifischer Webservice programmiert**. Implementiert wurde der Webservice im PHP-Framework Phalcon, Version 3.4. **Phalcon** ist ein minimalistisches Framework und zeichnet sich vor allem durch eine hohe Performance bei der Verarbeitung von vielen Anfragen und hohen Datenmengen aus. Es stellt für die Datenverarbeitung vorgefertigte Module für Verbindungsaufbau, Routing und Datenbankzugriffe zur Verfügung. Für die Datenbankzugriffe verfügt Phalcon über ein System zur objektrelationalen Abbildung (englisch object-relational mapping, ORM) und nutzt eine spezielle Query Language (PHQL). Dies ermöglicht neben einer Vereinfachung bei der Implementierung vor allem eine **erhöhte Sicherheit**, da so das Einschleusen von Datenbankabfragen über Formulare weitestgehend ausgeschlossen werden kann.

Der mit Phalcon implementierte Webservice wurde so aufgebaut, dass alle sekundären Verbindungsdaten (API-Key, Response-Typ, Landeskenntung etc.) via HTTP-Header gesendet werden, während die primären Daten (Eingabedaten aus den HKA) in die Request-URL integriert sind. Die für die Verbindung genutzten sogenannten Endpunkte nehmen Daten ausschließlich im standardisierten **JSON-Format** entgegen und geben die Rückgabewerte ebenfalls ausschließlich in diesem Format zurück. Via Header-Konfiguration kann beim Aufruf des Webservices entschieden werden, ob ein strukturiertes Objekt oder eine flache Datenhierarchie in Form eines einfachen Arrays (key => value) geliefert werden soll. Dies erhöht die **Flexibilität bei der Weiterverarbeitung** der Daten auf Seiten der aufrufenden Anwendung.

Validierung

Da die einzelnen Eingabewerte im smarten HeizCheck im Idealfall zu einem Großteil automatisch durch die KI erkannt und extrahiert werden und die Prüfung auf Richtigkeit durch die Nutzer*innen nur ein optionaler Handlungsschritt ist, kommt der **nachträglichen Prüfung auf inhaltliche Richtigkeit** eine wichtige Rolle zu. Nur so kann weitestgehend sichergestellt werden, dass die auf der Ergebnisseite der Anwendung dargestellten, berechneten Kennwerte schlüssig sind. Um dies gewährleisten zu können, werden die über den Webservice übertragenen Werte so weit wie möglich über **Kreuzvalidierungen** gegeneinander abgeglichen. Zu diesem Zweck wurden von Beginn des Projektes an mehr Einzelwerte auf den HKAs zur Erkennung durch die KI definiert, als für die spätere Berechnung tatsächlich benötigt werden.

Exemplarisch sei hier die Erkennung und Prüfung der Grundkosten- und Verbrauchskostenanteile jeweils für Heizung und Warmwasser angeführt. Die Summe dieser beiden Werte muss immer jeweils genau 100 % ergeben. Bei einer Abfrage direkt in der Anwendung wäre es ausreichend, entweder nur den Grundkosten- oder nur den Verbrauchskostenanteil abzu-

fragen und den jeweils anderen Prozentsatz durch die Differenz zu 100 zu berechnen. Das KI-Modell steht beim Auffinden dieser Werte vor zwei Herausforderungen. Zum einen muss die Lokalisation des jeweiligen Wertes auf der Abrechnungsseite korrekt vorhergesagt werden und zum Zweiten muss die OCR-Extraktion den dort abgebildeten Wert korrekt erkennen und in eine valide Zahl umwandeln. Um Fehler in diesen Bereichen im Nachgang auszuschließen, wurden die KI-Modelle so trainiert, dass beide Felder extrahiert werden und die Anwendung durch eine Summierung beider Werte prüfen kann, ob das Ergebnis der erwarteten Prüfsumme von 100 entspricht. Nach diesem grundsätzlichen Ansatz werden noch weitere, voneinander abhängige Einzelwerte und Summenwerte gegeneinander geprüft.

Da nicht alle Eingabewerte über solche Kreuzvalidierungen geprüft werden können, führt der Webservice in einem weiteren Schritt noch **auf den Berechnungsergebnissen basierende Validierungen** durch. So wird unter anderem geprüft, ob der berechnete EVKW innerhalb eines plausiblen Wertebereichs liegt. Abweichungen könnten auf Eingabefehler bei der Gebäudefläche und/oder der Energieträger-Verbrauchsmenge beziehungsweise der angegebenen Einheit hindeuten.

3.1.4. Bewertungsgrundlagen

Wohnungsverbrauch

Der Heizenergieverbrauch einer einzelnen Wohnung wird durch den energetischen Zustand des Gebäudes und dem individuellen Nutzungsverhalten der einzelnen Bewohner*innen des Haushalts bestimmt. Da die Bewohner*innen keinen direkten Einfluss auf den Zustand des Gebäudes haben, konzentriert sich die Bewertung der Wohnung auf den **Anteil, der durch das Nutzungsverhalten bestimmt wird**. Dazu wird der individuellen Wohnungsverbrauch pro Quadratmeter Wohnungsfläche dem mittleren, auf den Quadratmeter bezogenen Verbrauch des gesamten Gebäudes gegenübergestellt. Der mittlere Verbrauch des gesamten Gebäudes bildet den **Durchschnittsverbrauch aller Bewohner*innen des Gebäudes** ab. Je größer das Gebäude, desto mehr gleichen sich die unterschiedlichen Nutzungsprofile der einzelnen Haushalte gegenseitig aus und die Bewertung wird genauer. In kleinen Mehrfamilienhäusern können dagegen über- beziehungsweise unterdurchschnittliche Nutzung einzelner Wohnung und/oder ein signifikanter Anteil an Leerstand zu Verzerrungen bei der Bewertung führen.

Die Bedeutung der **Wohnungslage** bei der Bewertung des Heizenergieverbrauchs wurde bereits in Abschnitt 3.1.1. dargelegt. Für die verschiedenen Kombinationsmöglichkeiten aus den Angaben zur Etage und der Anzahl an Außenwänden der Wohnung hat SEnerCon auf Basis statistischer Erhebungen entsprechende Anpassungsfaktoren entwickelt, mit deren

Hilfe der ansonsten für das gesamte Gebäude global durchschnittliche Verbrauch auf die jeweilige Wohnungslage skaliert wird. Dadurch steigt beispielsweise bei Wohnung mit exponierter Lage der erwartbare Verbrauch im Verhältnis zum einfachen Durchschnitt. Bei Wohnungen mit günstiger Innenlage reduziert sich der Wert. Die Einordnung des individuellen Wohnungsverbrauchs erfolgt also nicht mit direktem Bezug auf den durchschnittlichen Gebäudeverbrauch, sondern mit dem **individuell lagebereinigten Verbrauchswert**.

Für aussagekräftige Angaben zum **Warmwasserverbrauch** im Haushalt wurde die Berechnung des **Pro-Kopf-Verbrauchs** herangezogen. Dabei bleiben die Angaben in der für Nutzer*innen im Vergleich zur kWh leichter verständlichen Einheit Liter. Die je Wohnung gemessenen Kubikmeter Warmwasser werden also in Liter umgerechnet und zu gleichen Teilen auf alle angegebenen Bewohner*innen des Haushalts verteilt. Diesem Kennwert wird im Ergebnis der Anwendung der statistisch erhobene, durchschnittlichen Pro-Kopf-Verbrauch an Warmwasser in Deutschland gegenübergestellt. Es wird also explizit nicht mit den Nachbar*innen im eigenen Gebäude verglichen, sondern bundesweit mit allen anderen beliebigen Haushalten. Dieser Vergleichswert stammt aus einer internen Auswertung von über 25.000 Beratungen des co2online Online-Ratgebers WasserCheck⁴².

Energiepreis und Heiznebenkosten

Vermieter*innen müssen bei der Betriebskostenabrechnung den Grundsatz der Wirtschaftlichkeit (**Wirtschaftlichkeitsgebot**) beachten. Sie sind somit verpflichtet, im Interesse der Bewohner*innen beim Einkauf von Energie und Dienstleistungen die regional üblichen Marktpreise beachten und in Abwägung zwischen Leistungsumfang und Preis so günstig wie möglich einzukaufen.

Der jeweilige Energiepreis im Abrechnungszeitraum lässt sich über die auf der HKA ausgewiesene eingekaufte Energieträgermenge und dem jeweiligen Preis ermitteln. Für eine eigene **Energiepreisdatenbank** erhebt co2online regelmäßig die Energieträgerpreise in Deutschland. Je nach Energieträger gibt es dafür unterschiedliche Quellen und Eckdaten. Für Erdgas und Fernwärme werden die Preise auf Basis zweier Abnahmeszenarien für insgesamt 45 kommunale Versorgungsgebiete recherchiert. In den anderen Versorgungsgebieten wird mit einem deutschen Durchschnittspreis verglichen. Für Einfamilienhäuser werden die Preise für einen angenommenen Jahresverbrauch von 15.000 kWh ermittelt. Für Mehrfamilienhäuser wird von einem Verbrauch von 150.000 kWh pro Jahr ausgegangen.

Bei Strompreisen für Wärmepumpen haben die meisten Versorger deutlich günstigere Sondertarife als für normale Stromkund*innen. Hierfür erhebt co2online einen deutschen

⁴² Vgl.: <https://www.co2online.de/service/energiesparchecks/wassercheck/>

Durchschnittspreis aus sieben Abnahmegebieten für ebenfalls zwei durchschnittliche Abnahmeszenarien. Quelle sind die beim Vergleichsportal Verivox gelisteten Versorgungstarife für Wärmepumpenstrom.

Um die Heizölpreise zu bestimmen, werden die Erhebungen des Statistischen Bundesamts („Erzeugerpreise für leichtes Heizöl“; Liefermenge 40 - 50 Hektoliter) für aktuell 13 deutsche Bezugsorte genutzt. Der durchschnittliche Preis für Holzpellets in Deutschland stammt aus einer jährlichen Untersuchung von C.A.R.M.E.N. e. V. (Centrales Agrar-Rohstoff Marketing- und Energie-Netzwerk)⁴³. Durchschnittliche Preise für Flüssiggas stehen für insgesamt zwölf verschiedene Versorgungsgebiete zur Verfügung, die durch den Bund der Energieverbraucher erhoben und veröffentlicht werden.

Die jeweiligen Vergleichspreise werden quartalsweise ermittelt. Für die Bewertung des individuellen Energiepreises des Gebäudes wird aus den zum Abrechnungszeitraum dazugehörigen Quartalen ein mittlerer Mischpreis berechnet. Die so ermittelten Vergleichspreise eignen sich tendenziell nicht für einen begründeten Einspruch der Mieter*innen auf Basis des Wirtschaftlichkeitsgebots. Sie dienen lediglich als grobe Orientierung und zeigen maximal eine **Auffälligkeit im Bereich der Energiebeschaffung** an. Ob für das Gebäude im regionalen Einzugsgebiet günstiger eingekauft werden kann, sollte durch den/die Vermieter*in geprüft werden. Die Bewertung in der Anwendung des smarten HeizChecks dient hier primär als Einstieg in eine Kommunikation zwischen Mieter*in und Gebäudeeigentümer*in.

Bei den **Heiznebenkosten** ist die Herausforderung um ein Vielfaches größer. Dazu gibt es keine öffentlich zugänglichen Datenquellen mit detaillierten Kostenvergleichen über die verschiedenen MDL hinweg. co2online selbst hatte bis Anfang 2016 mit dem Heizgutachten eine Dienstleistung über die Heiznebenkosten statistisch ausgewertet und als Grundlage für Benchmarks genutzt werden konnte. Seitdem die Heizgutachten eingestellt wurden, ist diese Quelle versiegt. Der entwickelte Prototyp verfolgt somit auch das Ziel, **aktuelle Heiznebenkosten-Daten zu erheben, damit zukünftig wieder aktuelle Vergleichswerte zur Verfügung stehen**. Bis das in einer ausreichenden Menge der Fall ist, wird auf die Bewertung der einzelnen Heiznebenkostenpositionen verzichtet. Stattdessen wird die Gesamthöhe der Heiznebenkosten analog zum Energiepreis verglichen und dient auch dort als grobe Orientierung für die Mieter*innen.

Gebäudeverbrauch

Dem individuell für das Gebäude berechneten **EVKW** wird für die Einordnung des potenziellen energetischen Zustands des Gebäudes auf der Ergebnisseite des smarten HeizChecks ein

⁴³ Vgl.: <https://www.carmen-ev.de/>

entsprechender Vergleichswert gegenübergestellt. Die Darstellung erfolgt mit Hilfe eines Bandtachos, auf dem **fünf Verbrauchsklassen** abgebildet sind. Das der Bewertung zugrundeliegende Gebäude wird einer der Verbrauchsklassen zugeordnet:

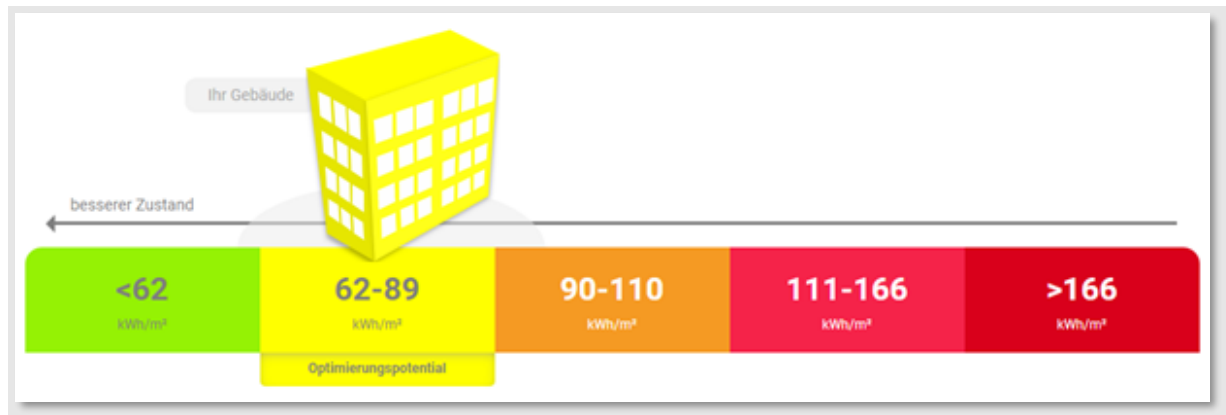


Abbildung 10: Bandtacho zur Einordnung des Gebäudeverbrauchs (EVKW)

Die für die Definition der verschiedenen Verbrauchsklassen nötigen, **statistischen Vergleichswerte** basiert auf den von SENERCon und co2online seit mittlerweile 15 Jahren erhobenen Daten für die Erstellung des bundesweiten Heizspiegels⁴⁴. Die Daten dafür stammen direkt von Verbraucher*innen, die das Online-Beratungsangebot von co2online nutzen. Diese sind über die Webseiten von co2online und vielen weiteren hundert Partnerseiten verfügbar und speisen eine intern als „**Gebäudedatenbank**“ bezeichnete zentrale Datensammlung mit über drei Millionen Energieabrechnungsdaten.

Die **Methodik der Heizspiegel-Erstellung** wurde über die vergangenen Jahre stetig weiterentwickelt. Der entsprechende Heizspiegel-Datensatz aus der Gebäudedatenbank wird mittels statistischer Methoden regelmäßig analysiert. Dazu werden im ersten Schritt die spezifischen Verbrauchswerte **Gebäudegrößenklassen** zugeordnet:

- 100 - 250 m²
- 250 - 500 m²
- 500 - 1.000 m²
- > 1.000 m²

Methodische Grundlage für die Einteilung der Datensätze in Verbrauchsgrößenklassen ist die Häufigkeitsverteilung der in der Stichprobe enthaltenen Verbrauchsdaten. Hierbei wird eine Normalverteilung angenommen. Für die Ermittlung der **Verbrauchsklassengrenzen** wird die Häufigkeitsverteilung in Quantile (0,1-, 0,5-, 0,9-Quantil) aufgeteilt (siehe Abbildung 11):

⁴⁴ Der Heizspiegel für Deutschland wird von der gemeinnützigen Beratungsgesellschaft co2online umgesetzt. Er informiert Verbraucher*innen seit 2004 über Sparpotenziale rund ums Heizen und wird vom Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit gefördert. Vgl.: <https://www.heizspiegel.de/>

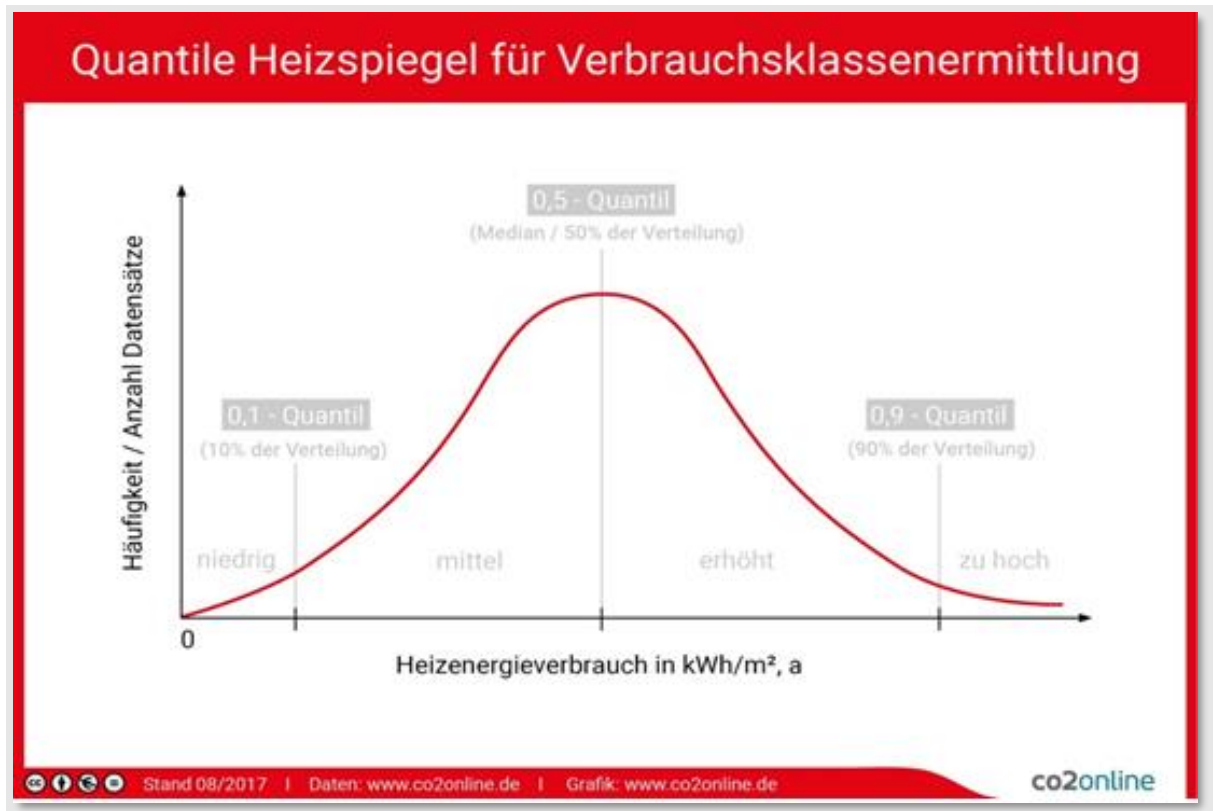


Abbildung 11: Quantile Heizspiegel für Verbrauchsklassenermittlung

Dieses Verfahren wird separat auf jeden Energieträger und die zuvor definierten Gebäudegrößenklasse angewendet. Die Verbrauchswerte werden dann auf Basis der Trendlinien und anhand der mittleren Gebäudeflächen der Gebäudegrößenklassen berechnet. Dabei werden angesetzt:

- 175 m² für die Gebäudegrößenklasse 100 - 250 m²
- 375 m² für die Gebäudegrößenklasse 250 - 500 m²
- 750 m² für die Gebäudegrößenklasse 500 - 1.000 m²
- 1.200 m² für die Gebäudegrößenklasse > 1.000 m²

Median und Perzentile bilden für die jeweilige Gebäudegrößenklasse dann den Bezug zu den **Bewertungsklassen, die im smarten HeizCheck verwendet werden:**

- **besten 10 %** (10 % der verteilten Verbrauchsdatensätze)
- **unter Durchschnitt** (10 - 40 % der verteilten Verbrauchsdatensätze)
- **Durchschnitt** (40 - 60 % der verteilten Verbrauchsdatensätze)
- **über Durchschnitt** (60 - 90 % der verteilten Verbrauchsdatensätze)
- **schlechtesten 10 %** (90 - 100 % der verteilten Verbrauchsdatensätze)

3.1.5. Zusammenfassung: Bewertung der HKA im Rahmen des Prototyps

Der als primäres Projektziel angestrebte **zentrale Rechenkern** und dessen Verwendbarkeit über eine dokumentierte API (Web-Schnittstelle) konnte erfolgreich implementiert werden. Die Grundlage für eine weitestgehend fehlerfreie und inhaltlich belastbare Berechnung und Bewertung von HKAs ist die **Vollständigkeit und Korrektheit der dazu zur Verfügung stehenden Parameter** wie Angaben zu Verbräuchen und den jeweiligen Kosten. Eine Herausforderung stellt in diesem Zusammenhang die Kombination automatisiert erfasster Einzelwerte und händischer Eingaben der Nutzer*innen dar, die eine **Fehlerquelle für die Plausibilitätsprüfung** (Validierung) sind. Durch die unterschiedlichen Wege der Kennwert-Erfassung kann sich im schlechtesten Fall eine Konstellation ergeben, bei der Nutzer*innen der Anwendung eine Vielzahl an sich teilweise aufeinander beziehenden Fehlermeldungen in geeigneter Form zur Verfügung gestellt werden muss. Dies darf im Idealfall aber nicht demotivieren und gegebenenfalls zum Abbruch der Online-Beratung führen. Speziell in diesem Prozess der **Korrektur von Fehleingaben in der Anwendung durch die Nutzer*innen** steckt zum Projektende noch weiteres Optimierungspotenzial.

3.2. Entwicklung des Prototyps des smarten HeizChecks

Die Entwicklung des Prototyps des smarten HeizChecks lässt sich in verschiedene inhaltliche **Entwicklungsphasen** gliedern, die je durch Formate der Nutzerzentrierung begleitet wurden. Die Übergänge zwischen den Phasen sind fließend und letztere Phasen überschneiden sich zum Teil zeitlich, da es sich bei der agilen Produktentwicklung nicht um einen linearen Prozess handelt. Das **Feedback der Verbraucher*innen** zur prototypischen Anwendung floss in jede Entwicklungs-Iteration mit ein und bildete die Basis für die Programmierung der neuen Anwendung (User Experience, UX). Die Online-Anwendung wurde so mit Hilfe der **Design Thinking-Methode** Schritt für Schritt entwickelt. Die Entwürfe für das Design und die Seiten der Anwendung entstanden gemeinsam mit einer erfahrenen UX-Agentur in Form von Sprints und werden im Folgenden zur Veranschaulichung beispielhaft präsentiert.

3.2.1. Anforderungsanalyse und Konzeption

Zum Start des Entwicklungsprozesses wurde ein **Konzeptions-Kick-off** mit den Verbundpartnern co2online, ConPolicy und SEnerCon und der betreuenden UX-Agentur veranstaltet. Ziel des Workshops war es, eine gemeinsame Vision und ein Grobkonzept des smarten HeizChecks zu entwickeln. Vorab erstellte das Team von co2online **Personas, die die Zielgruppe des Angebots repräsentieren**. Mit Hilfe von Personas werden beispielhafte Nutzergruppen mit je eigenen Bedürfnissen, Erwartungen und Herausforderungen skizziert. Dies sollte es dem Projektteam erleichtern, sich in die späteren Nutzer*innen und ihre Anforderungen

hineinzuversetzen. Anhand der Methode des **User Story Mapping** wurde anschließend gemeinsam die Grundstruktur des Prototyps – d. h., die wichtigen Schritte der Web-Applikation – entwickelt.

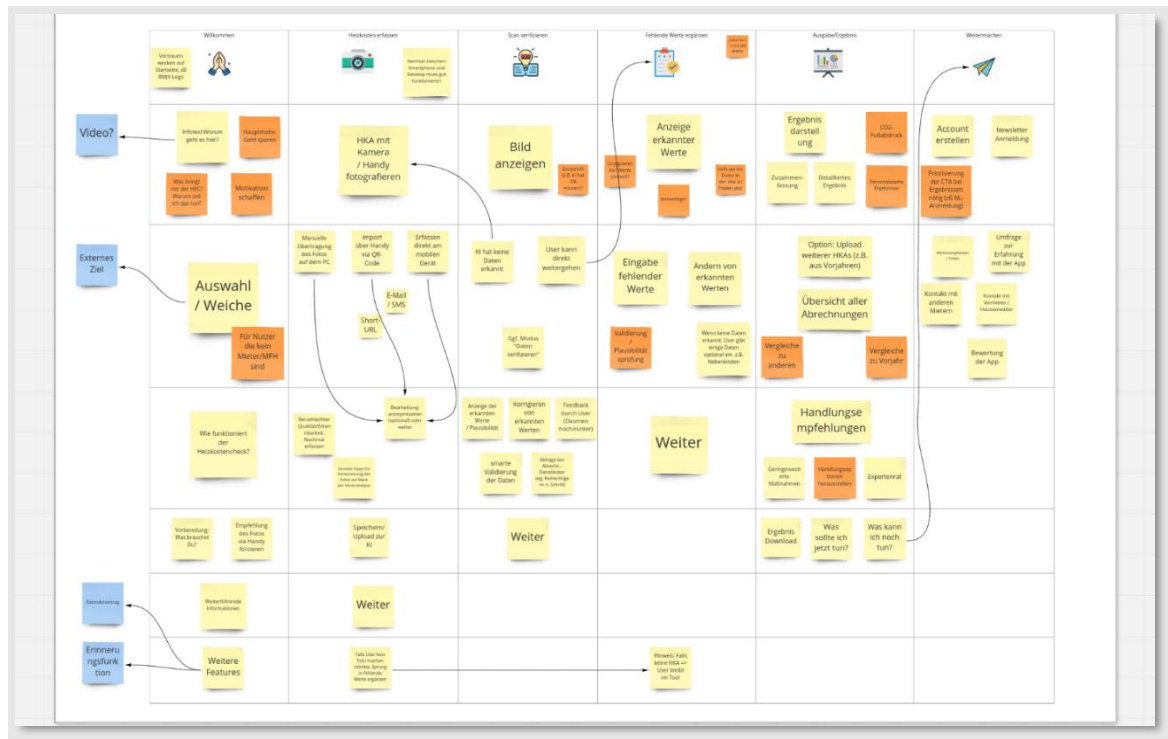


Abbildung 12: User Story Map zur Veranschaulichung der Anwendungsschritte des smarten HeizChecks

Die Konzeption der prototypischen Anwendung sollte sich von Beginn an den Erwartungen, Wünschen und Herausforderungen der Zielgruppe orientieren. Für den ersten Nutzertest, die Fokusgruppengespräche (s. a. 2.2.1.), wurden **Entwürfe in Form von Mock-ups der Anwendung** erstellt, die die Grundstruktur des Prototyps abbilden. In den Mock-ups wurden verschiedene Interaktionsstile exemplarisch umgesetzt, um zu testen, welche Form der Nutzerinteraktion und -kommunikation die Nutzer*innen bevorzugen. Die folgenden Stile wurden mit Nutzer*innen getestet: „smarte Fragen“, „Formularstil“ und „Bot-Dialog“ (s. Abbildung 13).

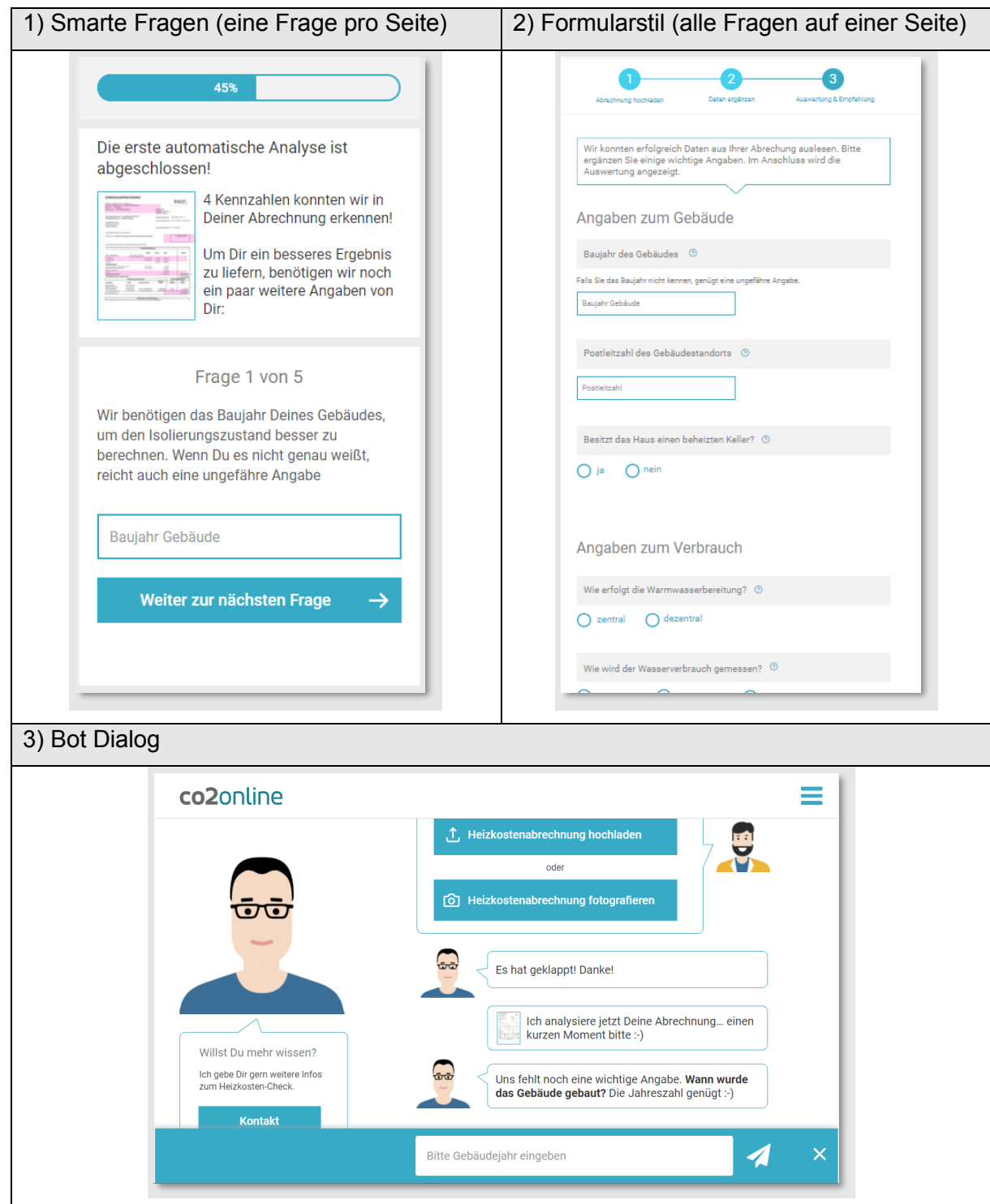


Abbildung 13: Umsetzung verschiedener Interaktionsstile in Form von Mockups für die Fokusgruppengespräche

Die Ergebnisse der Fokusgruppengespräche (s. a. 2.2.1.) dienen als Grundlage für die Weiterentwicklung der prototypischen Anwendung. Konkret wurden in der Phase des Prototypings die Entwürfe der neuen Anwendung an den Bedürfnissen der Nutzer*innen ausgerichtet und weiter konkretisiert.

Schwerpunkt der Weiterentwicklung bildeten folgende Erkenntnisse:

- Der **Formularstil** wird präferiert.
- Die Ansprache mit „Sie“ wird bevorzugt.
- Vergleiche und individuelle Maßnahmvorschläge/Tipps werden gewünscht.

3.2.2. Prototyping

Die **Design-Entwicklung in Form des Prototypings** verlief in der zweiten Jahreshälfte 2020 in kurzen Sprints in enger Zusammenarbeit mit der UX-Agentur und dem verantwortlichen Programmierer. Konkret wurden folgende Anwendungsschritte entworfen, die im Folgenden kurz vorgestellt werden:

- **„Erfassen“** bezeichnet den Schritt der Anwendung, in dem Nutzer*innen die Werte ihrer HKA erfassen.
- **„Validieren“** bezeichnet den Schritt, in dem Nutzer*innen die Möglichkeit haben, die von der KI erkannten Werte zu überprüfen und zu bestätigen.
- **„Ergänzen“** bezeichnet den Schritt in der Anwendung, in dem Nutzer*innen aufgefordert werden, zusätzliche Angaben für die Heizkostenanalyse zu machen und die von der KI nicht erkannten Werte manuell zu ergänzen.
- **„Auswerten“** steht für die Ergebnisseite der Anwendung. Hier erhalten die Nutzer*innen eine individuelle Auswertung zu ihrem Heizenergieverbrauch und Tipps dazu, wie sie ihren Verbrauch künftig senken können.

Die Entwürfe der Seiten wurden mit Nutzer*innen erprobt und auf Basis der Beobachtungen und des Feedbacks anschließend optimiert (s. a. 2.2.4.).

Anwendungsschritte „Erfassen“, „Validieren“, „Ergänzen“

Die Entwürfe für die Anwendungsschritte des Erfassens, Validierens und Ergänzens waren die ersten, die in Form des Prototypings entstanden sind und dann nach dem UX-Test wie folgt angepasst wurden:

1) Schritt „Erfassen“ vor UX-Test	2) Optimierte Version des Schritts „Erfassen“ nach dem UX-Test

Änderungen an den Entwürfen auf Basis der Ergebnisse der UX-Tests:

- Die Erfassung der HKA mit der Kamera ist als Funktion beliebt und wird in den Fokus gerückt.
- Der Prozess des Erfassens via QR-Code wird nicht verstanden und nun durch eine Animation zur Prozess-Erklärung unterstützt.

Abbildung 14: Entwürfe des Anwendungsschritts „Erfassen“ vor und nach den UX-Tests

1) Schritt „Validieren“ vor UX-Test	2) Optimierte Version des Schritts „Validieren“ nach dem UX-Test

Änderungen an den Entwürfen auf Basis der Ergebnisse der UX-Tests:

- Die Überprüfungsfunktion der Werte wurde optisch unterstützt und umstrukturiert.
- Die Markierungen durch Pfeile wurden ergänzt, um den erkannten Kennwert deutlicher zu identifizieren.
- Die Beschriftung der Felder wurde überarbeitet.

Abbildung 15: Entwürfe des Anwendungsschritts „Validieren“ vor und nach den UX-Tests

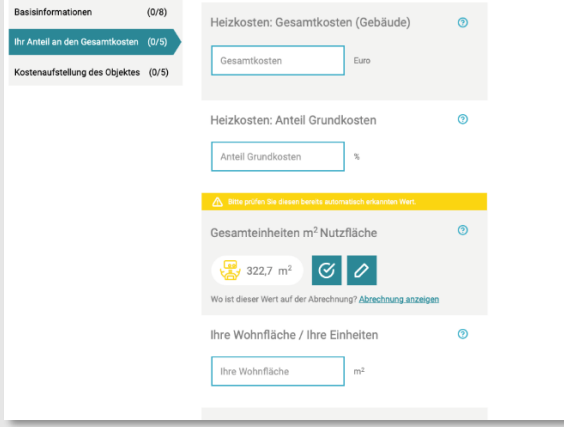
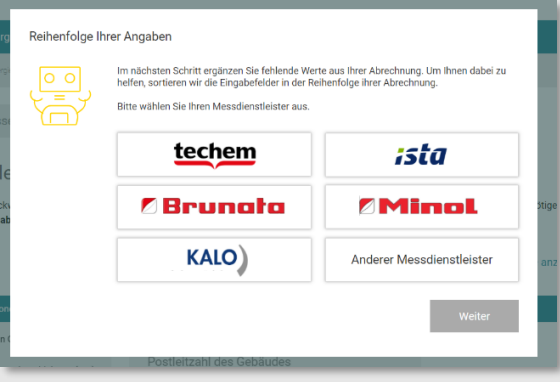
1) Schritt „Ergänzen“ vor UX-Test	2) Unterstützung des Schritts „Ergänzen“ nach dem UX-Test durch ein Intro-Overlay
	
Änderungen an den Entwürfen auf Basis der Ergebnisse der UX-Tests: - Die Suche der Kennwerte auf den Abrechnungen muss unterstützt werden: Deutliche Kommunikation der Sortierung der Werte nach Abrechnungslayout und der Benennung der Datenfelder nach den Abrechnungsbezeichnungen der jeweiligen MDL. - Die Darstellung einzelner Formularfelder sowie der dazugehörigen Einheiten (Lage der Wohnung, Außenwände) wurden überarbeitet.	

Abbildung 16: Entwürfe des Anwendungsschritts „Ergänzen“ vor und nach den UX-Tests

Anwendungsschritt „Auswerten“/Ergebnisseite

Eine besondere Herausforderung bei der Anwendungsentwicklung war die Gestaltung und Umsetzung der Ergebnisseite, da es sich bei der Heizkostenanalyse um ein hoch komplexes Thema handelt. Zeitgleich ist die Ergebnisseite des zu entwickelnden Prototyps elementar, um die Erwartungen der Nutzer*innen sowie die Ziele des Verbraucher- und Klimaschutzes (Transparenz, Empowerment, CO₂-Reduktion) zu erfüllen.

Auf Basis der Ergebnisse der ersten Fokusgruppengespräche und der quantitativen Kurzumfrage bezüglich der Verbrauchermotivation entstand ein **erster Entwurf der Ergebnisseite** (s. a. 2.2.1. und 2.2.2.). Die Aussagen der Verbraucher*innen ließen darauf schließen, dass die Ergebnis-/Bewertungsseite so gestaltet sein muss, dass sie zugleich verständlich, umfassend und individualisiert ist, um Nutzer*innen für energetische Einsparmaßnahmen zu aktivieren. So entstand die Idee, neben einer komprimierten und damit auf das Wesentliche reduzierten „**Auf einen Blick**“-Ergebnisdarstellung noch eine **erweiterte Ansicht** mit zusätzlichen Kennwerten und Darstellungsvarianten zu implementieren.

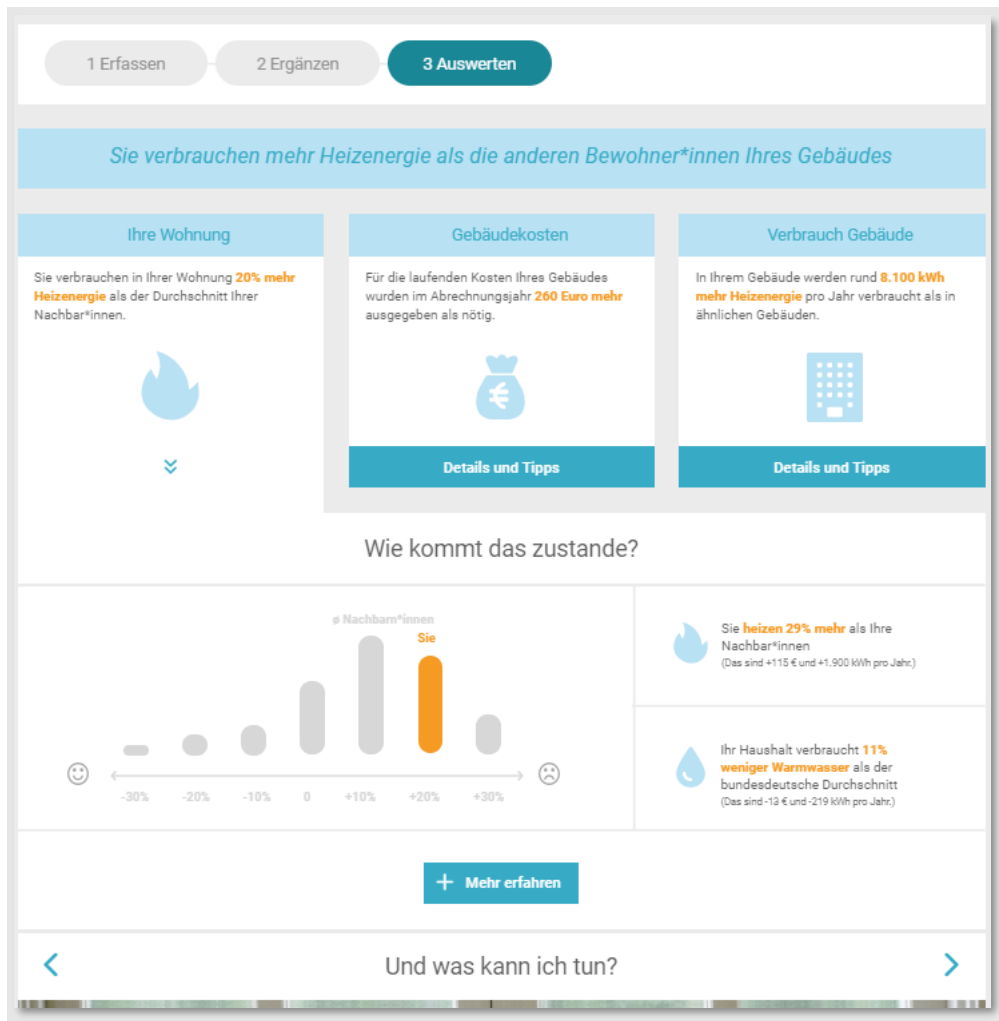


Abbildung 17: Erster Entwurf der Ergebnisseite mit mehrstufiger Informationsarchitektur

Dieser Entwurf der Ergebnisseite sowie unterschiedliche Darstellungen der integrierten Handlungsempfehlungen auf Basis der Heizkostenanalyse wurden dann in einer **zweiten Runde von Fokusgruppengesprächen** im Herbst 2020 zur Diskussion gestellt (s. a. 2.2.5.). In den Gesprächen zeigte sich, dass die verkürzte „Auf einen Blick“-Darstellung nicht geeignet ist, das notwendige Verständnis bei den Nutzer*innen zu manifestieren, diese sich aber zu einem großen Teil für mehr Informationen interessieren. Die Gespräche demonstrierten, dass die Verbraucher*innen wichtige Elemente der Ergebnisseite (in etwa der Bezug vom Heizen zum Warmwasser oder auch der Zusammenhang zwischen der eigenen Wohnung und dem bewohnten Gebäude) nicht verstanden. Außerdem wurde deutlich, dass die Befragten in dieser Darstellungsvariante schnell das Interesse verlieren und die Aussagen zum energetischen Zustand des Gebäudes ausblenden. Das Projektteam entschloss sich daher für die Konzeption eines **strukturellen Neu-Entwurfs der Ergebnisseite**, der folgende Nutzerbedürfnisse, die sich aus den Fokusgruppen und der Kurzumfrage ergaben, berücksichtigt:

- Es bedarf niedrigschwelliger Erklärungstexte, um das generelle Verständnis zur Zusammenstellung der Heizkosten zu festigen.
- Die Zusammenhänge zwischen den einzelnen Analyseebenen (Verbrauch des Haushalts, Heizkosten des Gebäudes, Verbrauch des Gebäudes) müssen textlich erläutert werden, um Nutzer*innen zu Einsparungen zu befähigen.
- Die eigene Wohnung steht an erster Stelle des Interesses bzw. der Auswertung.
- Der Bezug aller Ergebnisse zum eigenen Haushalt muss früh und immer deutlich werden.
- Vergleiche der eigenen Ergebnisse mit relevanten Vergleichsgruppen werden von den Nutzer*innen gewünscht. Zum Verständnis müssen Vergleichsgruppen und Bezugsgrößen hinreichend erklärt werden.
- Individuelle Möglichkeiten zum Geldsparen motivieren die Nutzer*innen. Hierfür können u. a. auch Nudges genutzt werden.

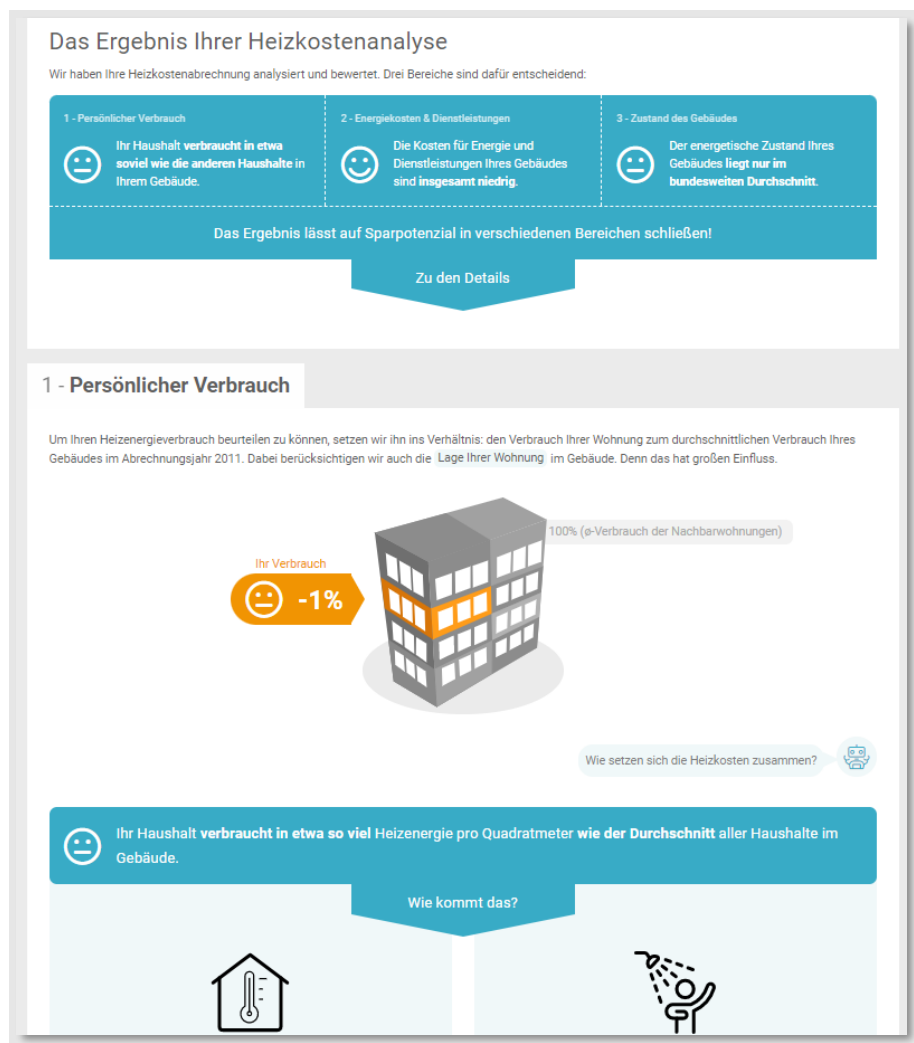


Abbildung 18: Neu-Entwurf der Ergebnisseite mit Erklärungstexten und allen Inhalten auf einer Seite

Die **Entwürfe der integrierten Handlungsempfehlungen/Tipps** wurden ebenfalls auf Basis der Fokusgruppengespräche sowie der verhaltenswissenschaftlichen Expertise des Verbundpartners ConPolicy (s. a. 2.2.5. und 2.2.6.) angepasst.

Ein Fokus lag dabei auf der Umsetzung folgender, zentraler Erkenntnisse:

- Individuelle Einsparpotenziale in Euro sind besonders wirksam.
- Es muss deutlich werden, welchen Handlungsspielraum Verbraucher*innen haben.
- Die Handlungsempfehlungen sollten priorisiert und vergleichend dargestellt werden. Dabei sind das Einsparpotenzial, die Schnelligkeit der Umsetzung und der Impact für Verbraucher*innen von besonderem Interesse.
- Kosten-Nutzen-Analysen sollten in die Darstellung integriert werden.
- Integrierte Produktempfehlungen wecken Misstrauen.
- Der mehrstufige Aufbau der Tipps überzeugt Nutzer*innen: (1) Teaser zum Überblick, (2) Mehr Infos in der Detailansicht.
- Es benötigt eine Übersetzung in Vergleichsgrößen für die Verdeutlichung der CO₂-Einsparpotenziale, damit diese für Verbraucher*innen erfassbar werden.



Abbildung 19: Teaser der Handlungsempfehlungen in Listenansicht mit integrierten Nudges (Einsparungen und Klimaeffekt)



Abbildung 20: Detailansicht der Handlungsempfehlungen mit integrierten Nudges (Kosten-Nutzen-Analyse, konkrete Einsparpotenziale in Euro und CO₂-Emissionen, Übersetzung der CO₂-Emissionen in verständliche Einheiten)

Die Ergebnisseite wurde so gestaltet, dass die **Handlungsempfehlungen/Tipps nach jedem Ergebnis bzw. Bewertungsbereich** (1- Persönlicher Verbrauch, 2- Energiekosten und Dienstleistungen, 3- Zustand des Gebäudes) **direkt integriert** wurden, um die Motivation der Nutzer*innen aufrechtzuerhalten (s. a. 2.2.5.).

Die Ausgestaltung des Beratungsergebnisses im smarten HeizCheck wurde im Projektverlauf mehrfach mit den **Praxispartnern** diskutiert und basierend auf deren Praxiserfahrung angepasst. Je nach Bewertungsergebnis bezüglich der Kosten für Energie und Dienstleistungen wird den Nutzer*innen empfohlen, eine **zusätzliche Beratung durch weitere Akteure des Verbraucherschutzes** wie die Beratungsstellen der Verbraucherzentralen oder des Deutschen Mieterbunds in Anspruch zu nehmen und ihre Rechte geltend zu machen. Dafür werden sie auf relevante Kontaktadressen im Branchenbuch von co2online (Rat und Tat) verwiesen.

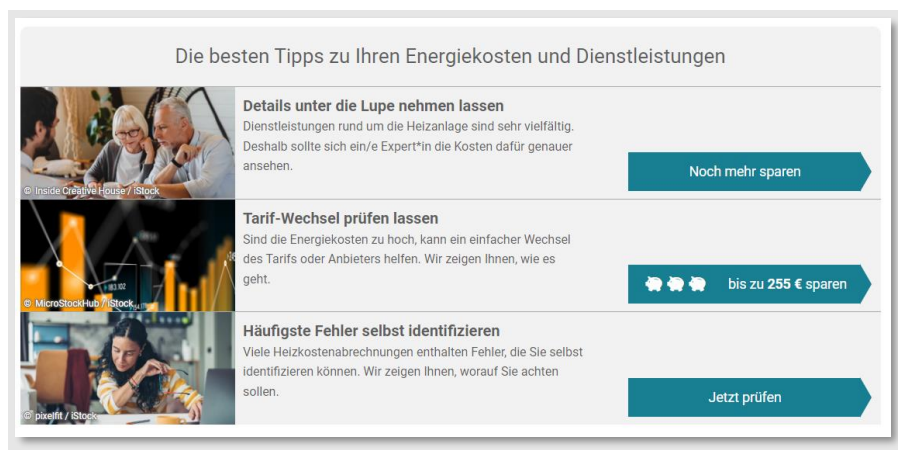


Abbildung 21: Handlungsempfehlungen mit Weiterleitung zu Verbraucherschutzorganisationen

Ein weiteres Anliegen ist es, Mieter*innen zu einem **Gespräch mit ihrem/ihrer Vermieter*in** zu motivieren, sofern der energetische Gebäudezustand überdurchschnittlich schlecht ist. Hierfür wurde ein Anschreiben inklusive der Ergebnisse der Heizkostenanalyse entwickelt, das den Nutzer*innen zum Download zur Verfügung steht.

Entwurf Intro/Landingpage

Aus den Nutzerbefragungen zeigte sich deutlich, dass eine **frühzeitige Vorteilskommunikation** notwendig ist, um Mieter*innen zur Heizkostenanalyse zu motivieren und ihre Hemmnisse zu adressieren (s. a. 2.2.1.). Zudem wurden in einem Kreativworkshop mit Hilfe der verhaltenswissenschaftlichen Expertise des Verbundpartners ConPolicy Ansätze zur Vorteilskommunikation auf der Startseite erarbeitet. Empfehlungen für die Startseite, die aus dem Workshop resultierten und anschließend umgesetzt wurden waren die Folgenden:

- Kurze und prägnante Vorstellung des Produkts bzw. des Anbieters und dessen Vorteile für die Nutzer*innen.
- Benennung des Ziels bzw. Mission des Produkts.
- Kurze Vorschau („sneak peak“) in den Prozess der Anwendung anhand eines Ablauf-Diagramms.
- Hervorhebung der Bedeutung des Themas, d. h., die Relevanz der Klimawirkung im Gebäudesektor.
- Nutzung des Effekts der sozialen Nachahmung, um neue Nutzer*innen zu gewinnen.
- Nutzung geeigneter Siegel und Logos, um die Glaubwürdigkeit der Anwendung zu stärken.

Auf Basis dieser Arbeiten entstanden die Entwürfe für ein **zweistufiges Intro-Overlay** als Einstieg in den smarten HeizCheck sowie der Entwurf einer **Landingpage** zur Produktvermarktung, die nach Projektabschluss umgesetzt werden kann. Dabei wurde neben der **Vorteilskommunikation** darauf geachtet, **Transparenz über den Umgang mit Daten** sowie eine **Schritt-für-Schritt-Anleitung** zu integrieren.



Abbildung 22: Zweistufiges Intro-Overlay

3.2.3. Umsetzung

Aus den mit Hilfe der breiten Expertise der Verbundpartner und den Ergebnissen der Nutzer-tests entstandenen, stetig optimierten Entwürfen der Anwendung entstand in der Umsetzungsphase **der Prototyp des smarten HeizChecks**. Im Prototypen wurden die KI-Module zur Erkennung sowie der Rechenkern zur Heizkostenanalyse per Web-Schnittstelle integriert.

Die Verbundpartner haben sich bereits frühzeitig darauf verständigt, den Prototypen in Form einer **Web-Applikation** umzusetzen. Die Entwicklung einer mobilen App zur Installation ist für

die Ziele des Projekts nicht sinnvoll, da es sich um eine Anwendung handelt, die keinen Nutzeraccount vorsieht und in den meisten Fällen lediglich ein Mal pro Jahr verwendet wird.

Die Umsetzung verfolgt einen starken **responsiven Ansatz**. Alle Gerätetypen, -klassen und -größen sollen die Eingabe möglichst einfach durchführen können. Der Prototyp wurde im Projektzeitraum in ca. 36 % der Fälle mit mobilen Geräten verwendet. Der Prototyp läuft nach dem ersten Aufruf **komplett eigenständig im Browser der Nutzer*innen**, ähnlich einer progressiven Webapp. Nur bei der KI- und Ergebnisberechnung wird eine Serverkommunikation benötigt. Nutzer*innen werden also während der Eingabe nicht durch unnötige Wartezeiten unterbrochen und können ihre Eingaben schnell und einfach durchführen. Dennoch werden die Eingaben der Nutzer*innen im Hintergrund mit dem Server synchronisiert. So können diese auch bei Verbindungsabbrüchen wieder hergestellt werden. Im UX-Design der Anwendung wurde auf „instant Feedback“ Wert gelegt. Fehlerhafte Eingaben sollen sofort erkannt werden und **dem/der Nutzer*in ein direktes und eindeutiges Feedback gegeben werden**. So wird Frustpotenzial bei der Nutzung minimiert und die Nutzer*innen brechen den Ablauf seltener ab. Auf die Nutzung von Cookies wird verzichtet. Es wird lediglich eine anonyme ID im Browser gespeichert, damit Nutzer*innen auch bei einer Wiederkehr in die Anwendung ihre Ergebnisse erneut aufrufen können.

Im Rahmen dieses Berichts sollen **drei Besonderheiten der Umsetzung** näher beleuchtet werden, die eng mit den Ergebnissen der begleitenden Verbraucherforschung zusammenhängen: die Erfassung der HKA, die Ausgestaltung der Hilfestellungen und die Umsetzung der Datenschutzanforderungen.

Erfassung der HKA

Um die Werte der HKA für die Analyse zu erfassen können Nutzer*innen zwischen **drei Optionen** wählen: Sie können (1) die HKA mit dem Smartphone/Tablet fotografieren, (2) Bilddateien der HKA hochladen oder (3) die Werte der HKA komplett manuell erfassen. Da die HKA in den meisten Fällen per Post verschickt wird und nur analog vorliegt, wurde die **Fotografier-Funktion im Prototypen** mit viel Aufwand umgesetzt. Die meisten Nutzer*innen starten den HeizCheck am Rechner. Es galt also, technisch eine Verbindung zwischen der Anwendung am Rechner und dem Smartphone/dem Tablet der Nutzer*innen herzustellen, so dass diese Bilder der Abrechnung machen und hochladen können. Der **Gerätewechsel zwischen Desktop und Smartphone/Tablet** wurde zunächst mit Hilfe eines **QR-Codes** gelöst. Bei den UX-Tests zeigten sich allerdings schwerwiegende Verständnis- und Bedienungsschwierigkeiten beim Scan des QR-Codes durch die Tester*innen (s. a. 2.2.7.). Um den Prozess zu vereinfachen wurde der QR-Code auf Basis der letzten Testiteration verworfen. Die Verknüpfung zwischen Desktop und Smartphone/Tablet wurde nun mit Hilfe eines **Kurzlinks und der**

Eingabe einer eindeutigen Kennziffer gelöst, um die Handhabung für die Nutzer*innen zu erleichtern.

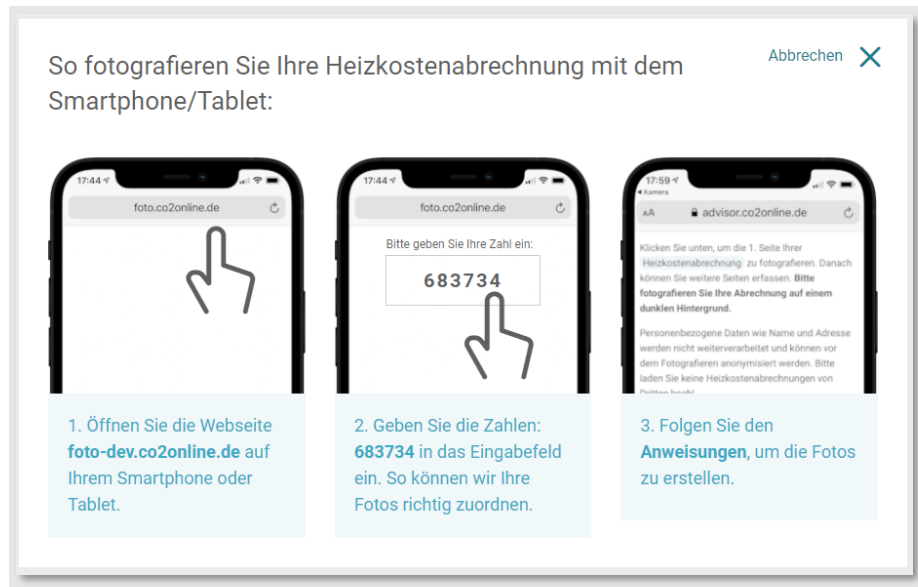


Abbildung 23: Verknüpfung zwischen Desktop und Smartphone/Tablet, um Bilder der HKA hochzuladen

Hilfestellungen im Prototypen

Bei den Nutzertests war zudem deutlich ersichtlich, dass die größte Herausforderung in der Anwendung für Nutzer*innen darin besteht, die **korrekten Werte auf ihrer HKA zu finden**, um diese in dem Anwendungsschritt „Ergänzen“ einzutragen (s. a. 2.2.4. und 2.2.7.). Da die Nutzer*innen bei der Kennwertsuche zu einem großen Teil verloren waren und sehr viel Zeit investieren mussten, handelt es sich bei dieser Aufgabe mutmaßlich um die **größte Abbruchstelle** für den smarten HeizCheck. Entsprechend dieser Erkenntnisse hat das Entwicklerteam **umfassende und dynamische Hilfestellungen für die Werteingabe in der Anwendung** entwickelt und umgesetzt. Die Hilfestellungen wurden für je 31 Kennwerte von sechs unterschiedlichen MDL/Abrechnungslayouts individuell angepasst und mit Beispielbildern inkl. Markierungen der Kennwerte in Musterabrechnungen versehen. Zudem wird den Nutzer*innen kommuniziert, wo sie den Kennwert üblicherweise in der Abrechnung finden. Um die Werteingabe weiter zu vereinfachen, wurden die Kennwerte so gruppiert und betitelt, wie sie sich auch auf der Abrechnung finden. Zuletzt wurde die Reihenfolge der Kennwertabfrage gemäß den Vorlagen der sechs MDL festgelegt.

Basisinformationen (4/7)

Aufstellung der Gesamtkosten (1/4)

Aufteilung der Gesamtkosten (3/15)

Sonstige Angaben (7/10)

Heizkosten EUR

Warmwasserkosten EUR

Wohnfläche (Gebäude)

Hier finden Sie den Wert: Seite 3, oben links.

5. Aufteilung der Gesamtkosten

Kostenart	Gesamt- betrag	Gesamteinheiten der Liegenschaft	Beitrag pro Einheit	Beitrag des Gebäudes	Beitrag des Gebäudes an den Gesamtbetrag
Heizkosten	1.000,00	1.000,00	1,00	1.000,00	100,00%
Warmwasserkosten	2.000,00	2.000,00	1,00	2.000,00	200,00%
Gesamt	3.000,00	3.000,00	1,00	3.000,00	300,00%

Bitte geben Sie hier die Kosten für die genutzte Raumwärme („Heizen“) des Gebäudes an.

Abbildung 24: Hilfestellungen für die Werteingabe im Schritt „Ergänzen“

Umsetzung Datenschutzanforderungen

Die Datenschutzanforderungen wurden in der Konzeptionsphase mit Hilfe des Datenschutzbeauftragten von co2online definiert und passende Maßnahmen bzw. Lösungsansätze durch den Verbund vorbereitet. Die Prototypenentwicklung folgt dem Anspruch der **Einhaltung vorbildlicher Danteschutzstandards**. Das bedeutete auch, die Erwartungen und Bedenken der Nutzer*innen zu erheben und in die Anwendungsentwicklung einfließen zu lassen. Auf Basis der Ergebnisse der quantitativen Umfrage zum Thema Datenschutz wurde bei der Anwendungsentwicklung darauf geachtet, die **Datenschutzkommunikation proaktiv zu gestalten (push-Ansatz)** und den Nutzer*innen relevante Datenschutzinformationen transparent und verständlich zur Verfügung zu stellen (s. a. 2.2.3.). Zudem erhalten Nutzer*innen Datenschutzhinweise immer genau dort in der Anwendungsoberfläche, wo Daten eingegeben/verarbeitet werden und mögliche Fragen entstehen, statt auf eine zentrale Datenschutzerklärung zu verweisen.

Die Empfehlungen und hohen Datenschutzstandards wurden wie folgt umgesetzt:

- **Upload und Speicherung der HKA:** Nutzer*innen erhalten wichtige Hinweise zur Speicherung ihrer HKA zum Zweck der Qualitätssicherung und Verbesserung der automatischen Heizkostenanalyse direkt in der Anwendung:
 - Umfang sowie Dauer der Speicherung und Datenverarbeitung
 - Erklärung zur Notwendigkeit der Speicherung
 - Hinweise zur Möglichkeit, personenbezogene Daten durch Schwärzen zu anonymisieren
 - Verlinkung zur Datenschutzerklärung des smarten HeizChecks⁴⁵

⁴⁵ Vgl.: <https://www.co2online.de/datenschutzerklaerung/#c166491>

- Möglichkeit, die Analyse manuell fortzusetzen und damit die Speicherung abzulehnen
- **Datenübertragung:** textliche Hinweise auf verschlüsselte und sichere Datenverarbeitung sowie Sichtbarmachung durch Integration des SSL-Siegels in der Anwendung
- **Datensparsamkeit:** Abgefragt werden nur die Daten, die für die Berechnung des Ergebnisses benötigt werden. In Hilfetexten wird erklärt, warum einzelne Kennzahlen benötigt werden.
- **Keine Weiterverarbeitung personenbezogener Daten:** Der Name, die Wohnadresse oder andere personenbezogene Daten der Nutzer*innen sind für die Analyse der HKA unerheblich und werden daher nicht weiterverarbeitet. Für die Heizkostenanalyse wird nur die Postleitzahl verwendet.
- **Eingabe der E-Mail-Adresse:** Die Beratungsergebnisse und Handlungsempfehlungen werden den Nutzer*innen ohne Eingabe weiterer Daten wie z. B. der E-Mailadresse zur Verfügung gestellt. Die E-Mailadresse wird nur für die Inanspruchnahme zusätzlicher Funktionen (z. B. Newsletter-Anmeldung) abgefragt (berechtigtes Interesse).

Das in den UX-Tests und über das Online-Feedbackformular (s. a. 3.2.4.) gesammelte Feedback lässt bisher darauf schließen, dass die nicht stark ausgeprägten Bedenken keine weiteren Anpassungen erfordern und die gut sichtbar platzierten Datenschutzinformationen hilfreich sind. Die Umsetzung einer **integrierten Anonymisierungsfunktion zur Schwärzung personenbezogener Daten ist daher aktuell nicht geplant**, da sie den Ablauf der Erfassung der HKA in dem smarten HeizCheck mitunter wesentlich verkomplizieren würde.

3.2.4. Beta-Phase

Die **Beta-Version des smarten HeizChecks wurde Anfang Juli auf www.co2online.de veröffentlicht und später in das Angebot www.heizspiegel.de integriert**. Die Beta-Phase wurde zunächst in kleinerem Rahmen durch verschiedene Zielgruppenmailings angekündigt, um dann die Nutzerzahlen nach und nach zu steigern und begleitend weiteres Nutzerfeedback einzusammeln.

Auf Basis der zweiten Runde an **UX-Tests**, die mit der Beta-Version durchgeführt wurden, konnte diese weiter optimiert werden (s. a. 2.2.7.). Die wichtigsten Änderungen umfassten:

- Schritt „Erfassen“:
 - Integration einer **Dialog-Abfrage**, um die Optionen zur Erfassung zu verdeutlichen
 - Umstellung des Fotografier-Prozess: Verknüpfung des Smartphones/Tablets mittels **Kurzlink**, Ersetzung des QR-Codes

- Integration von Ladeanimation und Fortschrittsanzeige nach Upload der HKA
- Schritt „Ergänzen“:
 - Integration einer **Guided Tour**, um die Anwendungsoberfläche vorzustellen und Nutzer*innen bei der Eingabe der Kennwerte zusätzlich zu unterstützen
 - Prominente Platzierung der Hilfestellungen beim Klick in das Eingabefeld
 - Überarbeitung der Validierungslogik und dazugehöriger Fehlermeldungen

Wie wollen Sie Ihre Heizkostenabrechnung erfassen?

Dateien hochladen
 Voraussetzung ist, dass Ihre Abrechnung bereits digital als Bilddateien (JPG/PNG) oder PDF vorliegt.

Mit Smartphone/Tablet fotografieren
 Voraussetzung ist, dass Sie ein Smartphone/Tablet mit Internetverbindung und Kamerafunktion haben.

oder

Sie wollen Ihre Heizkostenabrechnung nicht hochladen oder fotografieren?

Angaben manuell erfassen
 Die nötigen Angaben aus Ihrer Abrechnung können Sie auch selbst herausuchen und eingeben. Dieser zusätzliche Schritt dauert circa 15 Minuten.

Abbildung 25: Letzte Überarbeitung des Schritts „Erfassen“ durch Umsetzung der Dialog-Abfrage

Des Weiteren wurden die Nutzer*innen eingeladen, über ein **Online-Formular Feedback zur Anwendung** zu hinterlassen. Das Bild zur Interaktion mit der Anwendung ist zum Berichtszeitraum noch nicht eindeutig. Für eine umfassende Auswertung der Zufriedenheit und Conversion müssen erst die Nutzungszahlen gesteigert werden. Dies erfolgt nach Projektende mit der offiziellen Einführung des neuen HeizChecks und der vollständigen Implementierung des Produkts in das Online-Beratungsangebot von co2online.



Abbildung 26 Feedback zur Beta-Version, 43 Antworten

3.2.5. Zwischenfazit: Ergebnisse der nutzerzentrierten Prototypenentwicklung

Das Forschungsvorhaben verfolgte das Ziel, mit dem zu entwickelnden Prototypen Transparenz für Verbraucher*innen zu schaffen und bestehende Einsparpotenziale für sie zu heben, indem die Analyse der eigenen Heizkosten bzw. des eigenen Heizenergieverbrauchs einfach und schnell ermöglicht wird (s. **Arbeitsziele der Nutzerzentrierung, Verständlichkeit und Anwendungsfreundlichkeit**). Dieses Ziel konnte durch die konsequente Nutzerzentrierung bei der Anwendungsentwicklung sowie technische Fortschritte im smarten HeizCheck durch den Einsatz von KI zu einem wesentlichen Teil erfüllt werden. Der Ansatz des wiederholten Testens wurde bis zum Projektende weiterverfolgt, um größere und kleinere Optimierungsmaßnahmen am Prototypen kontinuierlich umzusetzen.

Um zu beurteilen, ob Nutzer*innen durch die Entwicklung des KI-Prototyps einfacher an relevante Informationen kommen, wurde der Fortschritt der Nutzer*innen in der Anwendung ausgewertet und die Erreichung der Ergebnisseite als Conversion ermittelt. Nach einem knappen halben Jahr der Veröffentlichung des Prototyps als Beta-Version ist eine **Conversionsrate von 38 %** erreicht worden. Dies betrifft den Anteil an Nutzer*innen, die durch eine erste Eingabe oder das Hochladen von Bildern eine Beratung gestartet haben und am Ende ein Beratungsergebnis, also eine komplette Heizkostenanalyse, erhalten haben. Die **Conversionrate bei der Nutzung der automatischen Bilderkennung liegt dabei mit 47 % deutlich höher als bei den Nutzer*innen, die keine automatische Erkennung nutzen (31 %)**. Zum Vergleich: Bei einem vergleichbaren Ratgeber wie dem HeizCheck von co2online liegt die Conversionrate gemittelt über mehrere Jahre bei etwa 32 %. Zu beachten ist noch, dass im Laufe der Beta-Phase des Prototyps mehrere Online-Nutzerbefragungen stattgefunden haben, die durch eine Incentivierung sowohl die Nutzungszahlen als auch die Abbruchquoten erhöht haben. Daher ist davon auszugehen, dass die **organische Conversionrate eher höher liegt**.

Positiv hervorzuheben ist, dass sich die lange Rechenzeit der KI (anderthalb bis vier Minuten) nicht negativ auf die Conversionrate auswirkt. Die Abbruchquote der Nutzer*innen, die zwar Daten hochladen aber danach keine Eingaben mehr vornehmen, liegt lediglich bei 18 %. **Verbesserungspotenzial ist vor allem bei Nutzer*innen mit manueller Eingabe aller Kennwerte** (ohne den Upload von Bildern) zu sehen. Bei den 68 % der Nutzer*innen mit manueller Eingabe, die während der Eingabe abbrechen, werden nur 54 % der nötigen Felder ausgefüllt. Der Grund ist nicht ganz klar und kann an fehlender Hilfestellung oder auch fehlender Motivation liegen.

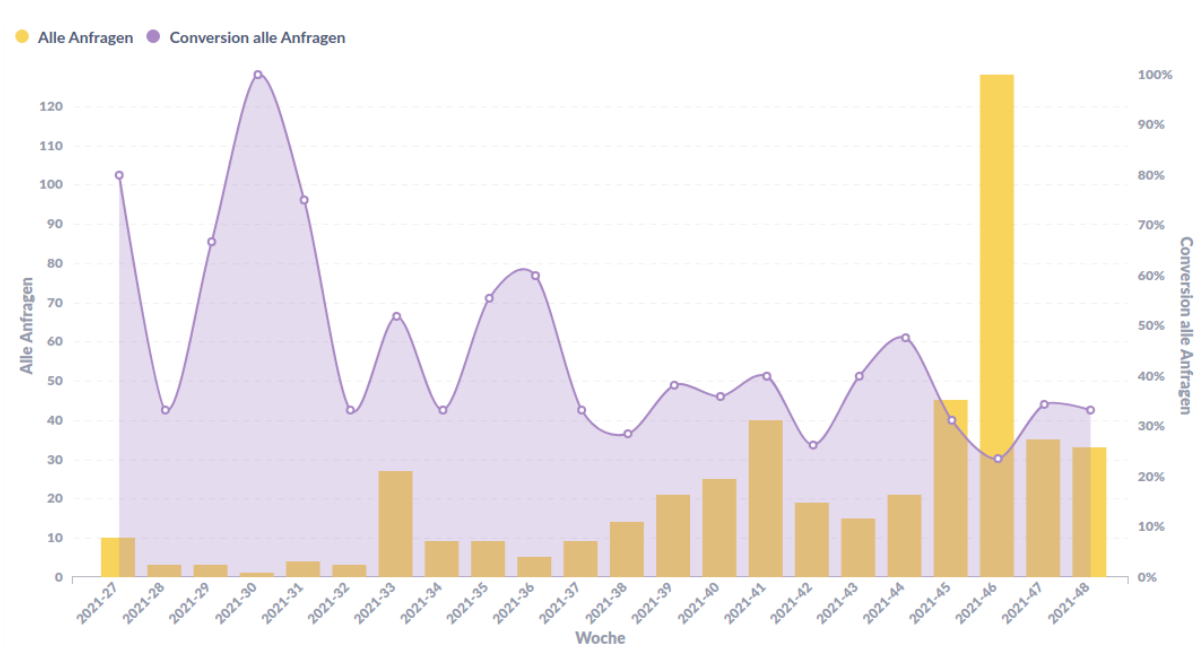


Abbildung 27: Schwankungen der Conversion-Rate (lila) und Anfragen pro Woche (gelb)

Im Vergleich zu anderen Beratungsinstrumenten von co2online konnte **eine Steigerung der Conversion durch die automatische Erkennung erreicht werden**. Es ist jedoch noch Raum für weitere Verbesserungen, die nach Projektende weiter geprüft werden sollen. Dazu gehören u. a. eine mögliche Weiterentwicklung der Schritte für die Datenerfassung (z. B. Hilfestellungen und Fehlermeldungen) und die langfristige Begleitung von Nutzer*innen, um den Heizenergieverbrauch zu senken. Dabei ist zu berücksichtigen, wie sich die organische Conversionrate entwickelt, nachdem der Prototyp als Produkt fest in das Beratungsangebot von co2online integriert wurde (s. a. 4.4.1.)

Ein weiteres Arbeitsziel bestand in der **Umsetzung vorbildlicher Datenschutzstandards**. Der entwickelte Prototyp hebt sich insofern von anderen Online-Anwendungen ab, dass relevante Datenschutzinformationen proaktiv und transparent direkt in die Anwendungsoberfläche integriert wurden. Ferner werden personenbezogenen Daten nicht weiterverarbeitet. So enthalten die hochgeladenen und gespeicherten HKAs zwar die Wohnadresse und Namen der Nutzer*innen, diese werden allerdings weder von der automatischen Bildverarbeitung noch im Zuge der Qualitätssicherung und Verbesserung der automatischen Heizkostenanalyse verwendet. Menschen mit hohen Datenschutzbedenken können den smarten HeizCheck vollumfänglich nutzen ohne personenbezogene Daten einzugeben, indem sie sich für die manuelle Eingabe der Kennwerte der HKA entscheiden. Durch zusätzliche Einhaltung der Grundsätze der Datensparsamkeit und Veröffentlichung der Analyseergebnisse ohne Abfrage unnötiger Nutzerdaten konnte das Ziel des vorbildlichen Datenschutzes erfüllt werden.

4. Verwertungskonzept

Im Rahmen des Vorhabens wurde ein Verwertungskonzept entwickelt, um die erfolgreiche Vermarktung und Inbetriebnahme des entwickelten Prototypen nach der Förderphase sicherzustellen. Das Konzept bündelt Projektergebnisse und Empfehlungen der Verbundpartner, um die Verstetigung der Anwendung des smarten HeizChecks bei zeitgleich hoher Nutzerakzeptanz sicherzustellen.

4.1. Einleitung

Ziel des Forschungsprojekts Smart_HEC war es, eine prototypische KI-Anwendung zu entwickeln, mit deren Hilfe Mieter*innen ihre Heizkostenabrechnung (HKA) prüfen und dazugehörige Sparpotenziale (Heizenergieverbrauch, Heizkosten) bequem identifizieren können. Im Projektzeitraum von 01.01.2020 - 31.12.2021 entstand der Prototyp des smarten HeizChecks, der nach Projektende als Produkt im Rahmen eines gemeinwohlorientierten Geschäftsmodells verwertet werden soll. Um zu identifizieren, in welchem Rahmen die Anwendung genutzt, verbreitet und verwertet werden kann, haben die Verbundpartner über verschiedene Quellen Wissen zusammengetragen:

- Nutzerbefragungen und -tests
- Expertengespräche mit den Praxispartnern des Projekts Deutscher Mieterbund, FINANZTIP und der Verbraucherzentrale NRW
- Literaturrecherche, u. a. zu Nudging-Ansätzen im Spendenbereich und Erkenntnissen aus dem Digitalen Marketing zum Reichweitenaufbau

Die Ergebnisse werden im Folgenden zusammengefasst und Handlungsempfehlungen abgeleitet, die als Basis für die Markteinführung des smarten HeizChecks dienen sollen.

4.2. Zielgruppen

Der Prototyp des smarten HeizChecks richtet sich in erster Instanz an die 16 Millionen Haushalte, die jährlich eine HKA erhalten. Dies **umfasst Mieter*innen und Wohnungseigentümer*innen zentral beheizter Gebäude**.

Eine weitere, indirekte Zielgruppe der Heizanalyse sind **Vermieter*innen, Hauseigentümer*innen und Hausverwaltungen**. Sie können durch Mieter*innen über die Ergebnisse der Heizkostenanalyse informiert werden. Dafür wird in der Anwendung ein Anschreiben zum Download zur Verfügung gestellt, in dem die Ergebnisse der Analyse zusammengefasst und Empfehlungen zur Verbesserung des energetischen Zustands des Gebäudes integriert werden. So werden die Empfänger*innen beispielsweise dazu motiviert, mit Hilfe einer

Energieberatung wirksame Sanierungsmaßnahmen zu identifizieren und so den Heizenergieverbrauch und die Heizkosten nachhaltig zu reduzieren.

4.3. Gemeinwohlorientiertes Geschäftsmodell

Die Anwendung wird nach Projektende im Rahmen der Klimaschutzberatung des Verbundpartners co2online verwertet und betrieben. Die **gemeinnützige Beratungsgesellschaft** besteht aus einem Team von über 40 Klimaschutzexpert*innen und berät seit fast 20 Jahren vor allem zu Klimaschutzmaßnahmen zuhause, die den CO₂-Fußabdruck nachweislich senken. Dafür initiiert das Sozialunternehmen verschiedene öffentlich geförderte cross-mediale Kampagnen und entwickelt digitale Beratungsinstrumente (Online-Ratgeber und Apps) sowie Websites und Dialog- und Themenkampagnen. Da der smarte HeizCheck in diese Klimaschutzberatung integriert werden soll, wird eine Gemeinwohlorientierung für den dauerhaften Betrieb des Produkts angestrebt. Im **Satzungszweck der gemeinnützigen Beratungsgesellschaft (gGmbH) ist die Verbraucherberatung verankert**. Eine abschließende Beurteilung und Entscheidung zur Orientierung an der Gemeinwohlökonomie hat co2online noch nicht getroffen und befindet sich zum Berichtszeitraum noch im **Prozess der nachhaltigen Geschäftsmodellentwicklung**.

4.3.1. Betriebskosten

Für den dauerhaften Betrieb der Anwendung ist eine gute Übersicht über die laufenden Betriebskosten elementar. Aufgrund einiger Unbekannten (genaue Platzierung/Implementierung des Produkts, Umsetzung der Weiterentwicklungsideen etc.) ist zum Zeitpunkt des Projektabschlusses noch nicht abschließend festzustellen, wie hoch die Betriebs- und etwaige Weiterentwicklungskosten künftig wirklich ausfallen werden. Nichtsdestotrotz wurde eine **Prognose für die Betriebskosten** erstellt, die auf Erfahrungswerten der Verbundpartner beruht.

Die laufenden Kosten lassen sich hiernach für die Bereiche **Fixkosten, Marketing und Vertrieb sowie Weiterentwicklung des Produkts** aufteilen. Um zu kalkulieren, wie hoch die Kosten pro Beratung gemäß dieser Prognose je Bereich sein werden, wurden die Gesamtkosten durch die Anzahl an Beratungen dividiert, von denen aktuell ausgegangen werden kann. Zum Berichtszeitpunkt schließen ca. 7.000 Mieter*innen und Wohnungseigentümer*innen mit Zentralheizung jährlich den HeizCheck ab. Da von einer Steigerung der Reichweite um 10 % ausgegangen wird, ist **mit 7.700 abgeschlossenen Beratungen pro Jahr zu rechnen**. Die Gesamtkosten pro Beratung liegen laut dieser Schätzung **hinsichtlich der reinen Fixkosten bei 2,25 Euro pro Beratung**. Sollen zusätzlich Marketing und

Vertriebsaktivitäten finanziert, sowie Weiterentwicklungsarbeiten durchgeführt werden, erhöhen sich die Kosten auf ca. **9,40 Euro pro Beratung**.

Betriebskostenaufstellung des smarten HeizChecks	
Prognose, Stand: 17.12.2021	
Fixkosten	
Pflege Anwendungsinhalte	2.880,00 €
Serverkosten	1.800,00 €
Wartung KI-Modelle	10.000,00 €
Wartung Online-Anwendung	2.665,60 €
Fixkosten p. a. gesamt	17.345,60 €
Fixkosten p. a./Beratung	2,25 €
Marketing & Vertrieb	
Kampagnenarbeit	6.720,00 €
Vertrieb	2.100,00 €
Mediabudget/Werbung	20.000,00 €
Marketing & Vertrieb p. a. gesamt	28.820,00 €
Marketing p. a./Beratung	3,74 €
Weiterentwicklung	
Programmierung	9.996,00 €
UX/Grafik	3.332,00 €
Produktmanagement	5.040,00 €
Content Management	960,00 €
Weiterentwicklung p. a. gesamt	19.328,00 €
Weiterentwicklung p. a./Beratung	2,51 €
Betriebskosten p. a. gesamt	65.493,60 €
10% Gemeinkosten	6.549,36 €
Kosten p. a. gesamt	72.042,96 €
Anzahl Beratungen p. a.*	7.700
Kosten p. a./Beratung gesamt	9,36 €

Tabelle 5: Vorläufige Betriebskostenaufstellung des smarten HeizChecks

4.3.2. Refinanzierung/Vergütungsvarianten

Im Laufe des Vorhabens wurden verschiedene Einnahmequellen geprüft, um den Betrieb der Online-Anwendung zu refinanzieren. Es zeigte sich deutlich, dass nicht alle Ansätze geeignet sind, da sie entweder nicht auf Akzeptanz bei Verbraucher*innen stoßen oder die mit dem Projekt verfolgten Verbraucher- und Klimaschutzziele konterkarieren.

Nutzungsentgelt / kostenpflichtiges Angebot

Grundsätzlich besteht die Möglichkeit, die Dienstleistung des smarten HeizChecks **kostenpflichtig** anzubieten. In dieser Variante könnten Verbraucher*innen die Anwendung nur nutzen, sofern sie einen festen Preis für das Angebot bezahlen. Denkbar sind auch so genannte **Freemium-Modelle**, bei denen Nutzer*innen zwar ein Teil des Angebots kostenfrei zur Verfügung gestellt wird, sie allerdings für das Vollprodukt und Erweiterungen (in etwa das Anschreiben an den/die Vermieter*in bzw. die Hausverwaltung) bezahlen müssten.

Im Zuge der Fokusgruppen zur Ergebnisseite im Forschungsprojekt wurde auch die Zahlungsbzw. Spendenbereitschaft von Nutzer*innen erhoben. Die Ergebnisse dieser Gespräche verdeutlichten eine starke **Ablehnung eines kostenpflichtigen Angebots**. Ein solches würde, laut eigenen Aussagen, von kaum jemandem genutzt. Den Ergebnissen zufolge steht die Etablierung eines kostenpflichtigen Angebots in Kontrast zu den Reichweitenzielen des smarten HeizChecks und sollte daher als Refinanzierungskanal nicht weiterverfolgt werden. Vielmehr wird angestrebt, möglichst viele Verbraucher*innen zu befähigen, Heizkosten und -energie zu sparen.

Werbeeinnahmen

Ferner lässt sich der Betrieb der Anwendung durch Werbeeinnahmen refinanzieren. Für den smarten HeizCheck spielen dabei vor allem Ansätze des **Affiliate-Marketings** eine Rolle. Hierbei stellt in der Regel ein kommerzieller Anbieter finanzielle Mittel bereit, um die eigenen Produkte über die Kanäle eines Vertriebspartners (sog. Affiliate) zu vermarkten. Ist die Vermarktung erfolgreich, erhält der Affiliate eine Provision. Im Rahmen der Gestaltung der prototypischen Anwendung ließen sich zum Beispiel **Produktlinks** zu smarten Thermostaten und Sparduschköpfen zum Sparen von Heizenergie/-kosten von verschiedenen Herstellern direkt auf der Ergebnisseite integrieren oder nach der Nutzung per E-Mail an die Nutzer*innen versenden.

Im Zuge der Fokusgruppen zur Ergebnisseite wurden die Reaktionen der Nutzer*innen auf Produktplatzierungen und Verweise auf kommerzielle Webseiten untersucht. Die Ergebnisse zeigten, dass eine gewisse **Skepsis bezüglich Produktempfehlungen** besteht und die Sorge, dass hinter dem Angebot kommerzielle Interessen stünden. Hingegen wünschen sich Nutzer*innen unabhängige Testberichte. Auch von dieser Einnahmequelle wird in Folge der gemischten Reaktionen zunächst Abstand genommen.

Lizenzmodelle

Die Online-Anwendung kann zudem als **Angebot an Dritte vermarktet** und in externe Webseiten eingebunden werden. Für die Integration und Verwendung der Anwendung bezahlen

Dritte dann **Lizenzgebühren**, die wiederum den Betrieb und die Weiterentwicklung der Anwendung refinanzieren. co2online hat seit Jahren Erfahrungen mit starken Klimaschutzpartnern, die die Online-Ratgeber (EnergiesparChecks) im Rahmen ihres eigenen Angebots nutzen.

Lizenzmodelle werden generell als geeignet bewertet, um Aufwände für den Betrieb des smarten HeizChecks anteilig zu tragen. Mögliche Partner hierfür sind Kommunen, Energieversorger, Hersteller von Effizienzprodukten im Gebäudebereich und Verbände wie z. B. der Deutsche Mieterbund, die im Rahmen der Vertriebstätigkeiten von co2online in der Vergangenheit zum Teil bereits angesprochen wurden oder in Zukunft akquiriert werden können.

Spenden / Pay-as-you-want-Service

Eine weitere Säule zur Refinanzierung des gemeinwohlorientierten Service ist die Generierung von Spendeneinnahmen. Dabei sollen Nutzer*innen gebeten werden, als **Gegenleistung für die kostenfreie Beratung eine freiwillige Spende in flexibler Höhe** (d. h., pay-as-you-want) zu entrichten. Die Möglichkeit zur Spende kann unter anderem am Ende der Heizanalyse auf der Ergebnisseite des smarten HeizChecks integriert oder als Aufruf an Nutzer*innen per E-Mail versandt werden, die den Newsletter von co2online abonniert haben. co2online testet zum Berichtszeitpunkt verschiedene Varianten des Spendenaufrufs und evaluiert, welche Botschaften die verschiedenen Zielgruppen am besten motivieren, für das Beratungsangebot zu spenden.



Abbildung 28: Beispielhafte Integration des Spendenaufrufs in die Online-Ratgeber von co2online

Die Ergebnisse der Nutzerbefragungen und die bisherigen Erfahrungen der Verbundpartner legen nahe, dass ein wesentlicher Teil der Kosten in Zukunft über Spendeneinnahmen gedeckt werden könnte. So wurde die **Spendenbereitschaft** der Nutzer*innen in Fokusgruppen zur Ergebnisseite sowie UX-Tests der Beta-Version der prototypischen Anwendung ermittelt.

Diese Gespräche zeigten, dass Nutzer*innen bei Zufriedenheit mit der Heizkostenanalyse grundsätzlich bereit sind, einen gewissen Beitrag zu spenden. Laut dem Feedback dürfte dieser Betrag fünf Euro nicht übersteigen. Zudem legen die Ergebnisse nahe, dass **anlassbezogene Spendenaufrufe** im Vergleich zu generischen Spendenaufrufen aufgrund deren höherer Nachvollziehbarkeit besonders zu Spenden motivieren (in etwa: „Spenden Sie 2 €, um die Künstliche Intelligenz auf einen neuen Messdienstleister zu trainieren.“).

Die Ergebnisse der Nutzerbefragungen und UX-Tests wurden durch Literaturrecherchen vervollständigt. Damit möglichst viele Nutzer*innen einen freiwilligen Beitrag leisten, ist ein geeignetes **Framing des Spendenaufrufs** sicherzustellen. Hierzu wurde die verhaltenswissenschaftliche Literatur zu Spendenbereitschaft und Fundraising aufbereitet. Folgende Erkenntnisse lassen sich für die Weiterentwicklung des Spendenaufrufs für den smarten HeizCheck bei co2online nutzen:

- Eine sogenannte **nicht-bindende Empfehlung für den Betrag einer Spende**, d. h., eine **spezifische Voreinstellung** (sog. Defaults) im Vergleich zur freien Wahl des Betrags, erhöht die Wahrscheinlichkeit, dass mehr Nutzer*innen eine Spende abgeben (Altmann et al., 2019). Die Höhe der Empfehlung sollte dem Kontext angemessen gewählt werden. Denn während **niedrige Vorschlagsbeträge für Spenden** eine höhere Antwortrate von Nutzer*innen bewirken (Adena, 2018), führen höhere Defaults tendenziell dazu, dass weniger Personen sich dazu entscheiden, zu spenden (Altmann et al., 2019). Aus diesem Grund empfehlen die Autor*innen, die Höhe des Defaults zu personalisieren (Altmann, 2019). Die **Freiheit, den voreingestellten Betrag für die Spende zu verändern**, sollte jedoch zwingend gewährleistet sein, da es andernfalls zu Abbrüchen kommen kann (Adena & Huck, 2020).
- Auch ein **Framing durch die Bezeichnung der Spende** kann einen gewissen Effekt auf die Spendenbereitschaft von Nutzer*innen haben und sollte daher berücksichtigt werden. So zeigen die Ergebnisse einer Literatur-Review, dass Nutzer*innen mehr spenden, wenn die Spende in einer Art und Weise präsentiert wird, die den Wert des Service und die Fairness des Angebots betonen (Gerpott, 2017). So hat etwa die Benennung „Bezahle, was es dir wert ist“ einen besonders positiven Effekt auf die Zahlungsbereitschaft der Nutzer*innen (Machado & Sinha, 2013).
- Die Abgabe und Abwicklung einer Spende sollte **einfach und ohne Hürden** möglich sein. Dazu gehört die Verwendung **einfacher, moderner Zahlungsmethoden** bei gleichzeitiger Gewährleistung von transparenten und **sicheren Transaktionsprozessen** (Pollach et al., 2005). Zudem sorgt ein in die **Webseite integriertes Spendenformular**, das idealerweise sehr reduziert gestaltet ist (Lumbres, 2021), für Vertrauen und vereinfacht den Spendenprozess (Lampe et al., 2015). Hier sollten außerdem nur

unbedingt notwendige Daten der Nutzer*innen abgefragt werden, um den Aufwand der Eingabe gering zu halten (Lampe et al., 2015).

- Das **Bewusstsein für die Notwendigkeit** einer Spende ist laut einer Meta-Analyse eines der wichtigsten Mechanismen zur Förderung der Spendenbereitschaft (Bekkers & Wiepking, 2010). Hier kann es erfolgsversprechend sein, konkrete Geschichten aus der Entwicklung zu berichten und Nutzer*innen in diesem Kontext anlassbezogen zur Spende aufzufordern (Lampe et al., 2015). Diese Erkenntnisse stimmen auch mit dem Feedback der Nutzer*innen in den Fokusgruppen überein. Nutzer*innen sind besonders interessiert daran, **exklusive Einblicke** zu erhalten und anschließend über die **Verwendung ihrer Spende** informiert zu werden (Lampe et al., 2015). In diesem Zusammenhang ist eine zeitnahe Rückmeldung mit einer Danksagung zuträglich (Lampe et al., 2015).
- Die **Kommunikation einer sozialen Norm**, wie etwa die Spendenbereitschaft anderer Personen, ist außerdem ein starker Mechanismus, der einen Konformitätsdruck auslösen und Nutzer*innen zur Spende bewegen kann (Goette & Tripodi, 2021). Dieses Prinzip wird u. a. von Wikipedia erfolgreich angewendet, indem die Anzahl an Nutzer*innen, die bereits gespendet haben, kommuniziert wird.

4.4. Platzierung der Anwendung und Reichweitenaufbau

Für den Erfolg der Anwendung und deren Nutzung von einer größtmöglichen Anzahl an Verbraucher*innen ist es entscheidend, dass die Anwendung im Internet Sichtbarkeit erlangt und deren Reichweite gesteigert wird. Ausgangspunkt sind hierbei die bestehenden Kanäle von co2online, die durch ihre langjährige Klimaschutzberatung bereits über eine hohe Reichweite verfügen – verschiedene Online-Portale erreichen durchschnittlich monatlich mehr als 330.000 Nutzer*innen. Darüber hinaus besteht ein umfangreiches Partnernetzwerk, worüber die Reichweite des Angebots weiter ausgebaut werden kann.

4.4.1. Integration in bestehende Beratungsplattformen

Der Prototyp des smarten HeizChecks soll als Produkt nach der Förderphase weiterentwickelt und in das Beratungsangebot von co2online integriert, sowie durch begleitende Kampagnen/Kommunikationsmaßnahmen verbreitet werden. Dabei sind insbesondere zwei Kanäle von Relevanz: die Webseite www.co2online.de sowie die Kampagnenseite www.heizspiegel.de.

Es wird zunächst geprüft, in welcher Form der Prototyp als fester Bestandteil des **umfangreichen Angebots der EnergiesparChecks auf co2online.de** (Online-Ratgeber)⁴⁶ etabliert werden kann. Der smarte HeizCheck ergänzt hiermit die Rechner rund um das Thema Energiesparen mit Fokus auf die Zielgruppe Mieter*innen und profitiert von den laufenden Klimaschutzkampagnen, die co2online umsetzt. Zum Berichtszeitpunkt schließen pro Jahr etwa 7.000 Mieter*innen und Wohnungseigentümer*innen eine HeizCheck-Beratung ab. So profitiert der Prototyp von der bereits bestehenden Reichweite von co2online, die nach der Projektlaufzeit erhöht werden soll (s. a. 4.4.2.).

Ferner wird angestrebt, das neue Angebot **im Rahmen der Aktion „Heizspiegel“ zu verwerthen**. Auf www.heizspiegel.de können Mieter*innen und Eigentümer*innen kostenlos prüfen, wie hoch ihre Heizkosten im Vergleich zu ähnlichen Haushalten sind. Außerdem berät der Heizspiegel Mieter*innen zu Themen wie Heizenergieverbrauch, Heizkosten und CO₂-Emissionen der Energieträger.

Für gezielte Kommunikationsmaßnahmen ist in der Projektlaufzeit zudem der Entwurf einer **Produkt-Landingpage** entstanden, die zu Werbezwecken umgesetzt werden sollte. Die Umfrageergebnisse zum Datenschutz zeigten, dass Siegel und Logos das Vertrauen in die smarte Heizkostenanalyse erhöhen und sie so für Verbraucher*innen attraktiver machen können. Aus diesem Grund wurde das **Logo des BMJV** prominent in den Prototypen integriert.

Abbildung 29: Integration des Förderlogos des BMJV zur Erhöhung der Glaubwürdigkeit der Anwendung

⁴⁶ Vgl.: <https://www.co2online.de/service/energiesparchecks/>

4.4.2. Steigerung der Sichtbarkeit und Multiplikator*innen

Sichtbarkeit

Die meisten Internetnutzer*innen gehen bei ihren Anliegen problembezogen vor und suchen eine Lösung über eine Suchmaschine. Erst später folgen Zugriffe auf eine Seite über Verweise von Websites Dritter (Kettner & Thorun, 2019; von Kossak et al., 2018). Daher ist es für die Sichtbarkeit jeder Anwendung entscheidend, dass das Angebot gut auffindbar und sichtbar auf der ersten Seite einer Suchmaschine erscheint.

Der erste Schritt für die gute Sichtbarkeit ist die **Optimierung der eigenen Webseiteneinhalte** mittels einer **Keyword-Analyse**. Diese Analyse ermöglicht es, die Suchinteressen von Nutzer*innen besser zu verstehen und Einblicke in die damit zusammenhängenden Fragen zu entwickeln (Lampe et al., 2015; von Kossak et al., 2018). So zeigen die Ergebnisse einer Studie zur Suche nach Energieberatung beispielsweise, dass Verbraucher*innen anlassbezogen nach Produkten oder ihnen bekannten Expert*innen, nicht aber nach dem Begriff Energieberatung suchen (Ellermann et al., 2020). Auch im Zuge der Befragungen im Rahmen dieses Projekts wurde deutlich, dass Nutzer*innen mit bestimmten Begrifflichkeiten besser vertraut sind als mit anderen (bspw. kennen viele den Begriff „Messdienstleister“ nicht, nutzen aber dafür andere Begriffe wie „Ablesedienst“). Zur Integration dieser Erkenntnisse in die Verwertung der Anwendung könnten systematisch Erkenntnisse zu thematisch genutzten Begrifflichkeiten der Nutzer*innen, die im Projekt gewonnen wurden, gesammelt, mit Drittquellen angereichert und aufbereitet werden. Anschließend könnten diese Erkenntnisse von der co2online-Redaktion genutzt werden, um Inhalte auf der eigenen Webseite besser an die Interessen und Begrifflichkeiten anzupassen.

Ein zweiter wichtiger Pfeiler zur Sicherstellung der Sichtbarkeit ist die **Suchmaschinenoptimierung** (SEO). Hierbei können sowohl Maßnahmen wie Crawling und Indexierung genutzt werden, um sicherzustellen, dass Suchmaschinen die Webseite gut auffinden können und die Seite entsprechend positioniert wird (von Kossak et al., 2018). Zum anderen können Maßnahmen getroffen werden, um die Darstellung von Inhalten in Suchmaschinen zu optimieren. Da der smarte HeizCheck auch zukünftig Teil der co2online-Webseite sein soll und hier schon umfassende SEO-Maßnahmen getroffen werden, werden zur Verwertung der Projektergebnisse bestehende Maßnahmen fortgeführt und bei Bedarf ausgebaut.

Multiplikator*innen

Die Anwendung sollte über die gute Sichtbarkeit hinaus mit Hilfe **relevanter Multiplikator*innen** an geeignete Zielgruppen verbreitet werden und so langfristig eine große Reichweite erzielen. Zum Aufbau der Reichweite besteht bereits ein solides Partnernetzwerk

aus Beratungsdienstleistern zu Finanz- und Umweltthemen sowie aus Mieterverbänden. Dieses Netzwerk soll auch genutzt werden, um die Anwendung zu bewerben. Während der Entwicklungsphase fanden bereits **Expertengespräche mit den Praxispartnern** Deutscher Mieterbund, FINANZTIP und der Verbraucherzentrale NRW statt, um die wirksame Verbreitung der neuen Anwendung vorzubereiten. Es ist angedacht, dass die Praxispartner je mit einem Testimonial die Qualität des smarten HeizChecks bezeugen, das u. a. auf der Produkt-Landingpage zu Werbezwecken genutzt werden kann, und gemeinsam mit co2online den Prototypen bewerben. So zeigte sich etwa FINANZTIP offen für die Idee, den neuen Service im eigenen, reichweitenstarken Newsletter anzukündigen.

Darüber hinaus ist es erstrebenswert, zur Verbreitung des Angebots weitere Kooperationen anzubahnen. Für diesen Zweck wurden mögliche **Anlässe der Beschäftigung mit Energie- und Heizthemen** von Verbraucher*innen identifiziert. Unterschieden werden können dabei u. a. persönliche, gebäudebezogene oder politische bzw. gesetzliche Anlässe (Ellermann et al., 2020). Zu persönlichen Anlässen können etwa ein Umzug oder der Wunsch einer Veränderung der Heizsituation zählen, zu gebäudebezogenen Anlässen ein Sanierungsbedarf am Gebäude oder ein Anbieterwechsel für die Wohneinheit. Zu den politischen Anlässen zählt die mediale Berichterstattung über den Energieverbrauch im Wohnsektor sowie neue gesetzliche Bau- oder Energiesparverordnungen. Aus der Psychologie und den Verhaltenswissenschaften ist außerdem bekannt, dass Verbraucher*innen nach Umbrüchen im Leben (z. B. Studium, Elternschaft, Job, Umzug, Hauskauf) besonders offen für die Ansprache mit neuen Themen sind.

Daraus lassen sich Berührungspunkte mit dem Thema und mögliche Anknüpfungspunkte für die Kommunikation rund um den smarten HeizCheck ableiten:

- **Mieterforen und (Energie-)Beratungsdienstleister** können insbesondere für Mieter*innen nach einem Umzug Anschlussfähigkeit mit dem Thema herstellen.
- **Vergleichsportale und Finanzplattformen** können bei der Suche nach neuen Produkten oder Dienstleistungsverträgen möglicher Anknüpfungspunkt für die Heizkostenanalyse sein.
- **Nachbarschaftsgruppen und Elternvereinigungen** können für junge Familien eine geeignete Austauschmöglichkeit bieten.
- **Umweltaktivistische Gruppen und Influencer** können bei jüngeren Menschen Aufmerksamkeit für die Bedeutung des Gebäudesektors beim Klimaschutz generieren.

4.5. Konzept zur Wirkungsmessung

Das kurzfristige Ziel des smarten HeizCheck ist es, Verbraucher*innen eine smarte Anwendung zur Verfügung zu stellen, mit deren Hilfe sie niedrigschwellig und in einfacher und intuitiver Weise ihre HKA überprüfen können. Konkret soll der smarte HeizCheck sie dabei unterstützen, Auffälligkeiten der Abrechnung zu erkennen, das eigene Heizverhalten einzuschätzen sowie Einsparpotenziale zu identifizieren. Langfristig soll der smarte HeizCheck dazu beitragen, nachhaltige Verhaltensveränderungen bei den Verbraucher*innen anzustoßen, energetische Sanierungen zu initiieren und in der Folge CO₂-Einsparpotenziale zu heben.

Die Frage danach, ob der smarte HeizCheck diese Ziele erreichen und somit seine Wirkung entfalten kann, sollte analog zu den gesteckten Zielen sowohl kurzfristig als auch langfristig beantwortet werden. Für diesen Zweck wurde ein Konzept zur Wirkungsmessung mit Vorschlägen geeigneter Indikatoren und Methoden zur Messung derselbigen entwickelt, das im Nachgang zum Projekt genutzt werden kann, um die kurz- und langfristigen Wirkungen der Anwendung zu evaluieren.

Kurzfristige Wirkungsmessung

Die Conversion, d. h., die Zahl der Nutzer*innen, die den smarten HeizCheck erfolgreich abschließen, gibt Aufschluss darüber, ob es sich um eine einfach zu bedienende und intuitiv verständliche Anwendung handelt. Es zeigte sich in der Beta-Phase bereits, dass die **Conversionrate bei Nutzung der Bilderkennung mit 47 %** höher liegt als die Conversionrate im bestehenden HeizCheck mit 32 %. Das deutet darauf hin, dass die **Niedrigschwelligkeit und Nutzungsfreundlichkeit des Angebots gesteigert werden konnte**.

Die Frage danach, ob die Ergebnisse des smarten HeizCheck Verbraucher*innen dabei unterstützen, ihren Verbrauch besser zu verstehen, Auffälligkeiten zu identifizieren sowie Ansatzpunkte für Einsparungen zu erkennen, kann auf verschiedene Weisen approximiert beantwortet werden:

- **Download-Zahlen der Ergebnisseite als PDF-Datei:** Die Anwendung bietet Nutzer*innen die Möglichkeit, sich die Ergebnisse der persönlichen Heizkostenanalyse als PDF-Datei herunterzuladen. Wird dieser Service genutzt, kann angenommen werden, dass die Ergebnisse eine gewisse Relevanz haben, später noch einmal angesehen oder in anderer Weise genutzt werden. Demnach können die Download-Zahlen der Analyse als Indiz für die kurzfristige Wirkung des smarten HeizChecks herangezogen werden.

- **Klickzahlen auf tiefere Ebene der Tipps:** Die Ergebnisseite gibt überblicksartig Hinweise dazu, wie die Nutzer*innen basierend auf dem Analyseergebnis ihr Heizverhalten anpassen oder welche Maßnahmen sie ansonsten ergreifen können, um ihren Heizenergieverbrauch zu senken. Jede dieser Maßnahmen kann angeklickt werden, um tiefergehende Informationen zu erhalten. Die Analyse der Frequenz der Aufrufe dieser detaillierteren Unterseiten kann genutzt werden, um die Relevanz der Tipps zu überprüfen und die kurzfristige Wirkung des smarten HeizChecks abzuschätzen.
- **Analyse der Verweildauer auf der Seite sowie des Scrollverhaltens:** Die Ergebnisseite ist in verschiedene Abschnitte unterteilt und stellt kompakt auf einer Seite die persönlichen Ergebnisse zum persönlichen Verbrauch, den Energiekosten sowie zum Zustand des Gebäudes dar. Die Analyse danach, ob die Seite bis zum Ende gescrollt wird und wie lange sich Nutzer*innen bei den einzelnen Abschnitten aufhalten, kann einen weiteren Anhaltspunkt dazu geben, ob eine kurzfristige Wirkung erzielt wird.

Die dargestellten Analyseoptionen haben den Vorteil, dass sie ohne konkretes Zutun der Nutzer*innen durchgeführt werden können und ein realitätsgetreues Bild der kurzfristigen Wirkung des smarten HeizChecks ermöglichen. Gleichzeitig erlauben die Optionen nur eine approximierte Wirkungsanalyse und es können keine verallgemeinerbaren oder kausalen Rückschlüsse gezogen werden.

Eine **Nutzerbefragung zum Abschluss des smarten HeizChecks** bietet die Möglichkeit, ab einer gewissen Stichprobengröße generalisierbare Aussagen dazu zu treffen, wie Nutzer*innen den smarten HeizCheck und die zugehörigen Ergebnisse wahrgenommen haben, wie hoch der jeweilige Erkenntnisgewinn ist und ob die Ergebnisse der Analyse eine Handlungsintention nach sich ziehen. Allerdings kann eine Befragung nicht das tatsächliche Verhalten der Testpersonen abbilden und die Ergebnisse müssen unter Einbeziehung verschiedener Biases (etwa Social-Desirability Bias oder Attitude-Behavior-Gap) interpretiert werden. Zudem besteht eine Schwierigkeit darin, Nutzer*innen für die Teilnahme an der Umfrage zu gewinnen und sie entsprechend zu incentivieren.

Langfristige Wirkungsmessung

Um abschätzen zu können, welche CO₂-Einsparungen durch den smarten HeizCheck initiiert und bewirkt werden können, muss die Langzeitwirkung des smarten HeizChecks auf das Heizverhalten von Verbraucher*innen sowie die konkrete Umsetzung von energetischen Sanierungs- und Ertüchtigungsmaßnahmen gemessen bzw. approximiert werden. Hierfür können grundsätzlich zwei Wege gegangen werden:

Auf der einen Seite kann **ex-post eine Nachbefragung der Nutzer*innen** Erkenntnisse darüber liefern, ob und, wenn ja, welche Maßnahmen zur Reduktion der Heizenergiekosten angestoßen bzw. umgesetzt wurden und welche Verhaltensweisen sich dauerhaft etabliert haben. Eine solche Nachbefragung kann verlässliche Ergebnisse über das tatsächliche Verhalten produzieren und so eine belastbare Aussage über CO₂-Einsparungen ermöglichen. Die Herausforderung besteht jedoch darin, ausreichend Testpersonen für die Nachbefragung zu gewinnen. Hierfür sollten unterschiedliche Motive zur Teilnahme an Befragungen beachtet werden:

- **Incentivierung:** Es kann davon ausgegangen werden, dass Nutzer*innen des smarten HeizChecks nur dann für eine Nachbefragung zur Verfügung stehen, wenn sich für sie hieraus ein gewisser Vorteil ergibt oder sie sich von der Teilnahme eine Wirkung versprechen. Bei Personen mit ohnehin hoher intrinsischer Motivation, etwas für den Klimaschutz zu tun oder die Arbeit von co2online zu unterstützen, genügen ggf. motivierende Botschaften. Diese sollten auf verhaltenswissenschaftlichen Erkenntnissen beruhen und bspw. den konkreten Beitrag zum Klimaschutz unterstreichen, mit sozialen Vergleichen oder Challenges das Gruppengefühl stärken. Bei Personen ohne expliziten Bezug zum Klimaschutz oder zu co2online als Anbieter des Service können Vergütungen in Form von Gutscheinen oder Gewinnspielen motivieren, an einer Nachbefragung teilzunehmen.
- **Timing:** Es kann davon ausgegangen werden, dass die Bereitschaft an einer Umfrage teilzunehmen, dann besonders hoch ist, wenn die HKA und die damit verbundenen Kosten gerade präsent sind bzw. wenn Verbraucher*innen sich mit dem Thema beschäftigen. Vor diesem Hintergrund sollte der Hinweis auf eine Nachbefragung ungefähr in den Zeitraum der Abrechnung fallen, um so die Relevanz zu erhöhen.

Die in der Nachbefragung erhobenen Ergebnisse zu angestoßenen Maßnahmen und Veränderungen können abschließend mit den jeweils zugehörigen CO₂-Einsparungen multipliziert werden, um so abzuschätzen, **welche Klimaschutzwirkung durch die Durchführung des smarten HeizCheck angestoßen werden konnte**. Bei dieser Abschätzung ist jedoch zu beachten, dass nicht von einer Kausalität ausgegangen werden kann, da andere Faktoren für Veränderungen im Heizverhalten bzw. Modernisierungsanstöße verantwortlich sein können.

Eine andere Möglichkeit, die durch den smarten HeizCheck initiierten CO₂-Einsparungen zu quantifizieren, ist die **ex-ante Approximation**. Diese greift auf die **Ergebnisse der Nutzerbefragung im Anschluss an die Durchführung des smarten HeizChecks** zurück, die bereits pilothaft während der Projektlaufzeit durchgeführt wurde. In dieser soll – wie oben

dargelegt – die Nutzerintention abgefragt werden, bestimmte Maßnahmen einzuleiten bzw. das persönliche Heizverhalten anzupassen und dauerhaft zu ändern.

Die hierbei erhobenen Angaben stellen in der Approximation die obere Grenze dessen dar, was mit dem smarten HeizCheck maximal bewirkt werden kann. Die Wirkung der CO₂-Einsparung liegt somit zwischen Null (wenn niemand eine Maßnahme umsetzt) und diesem Maximalwert, der sich aus der Wahrscheinlichkeit berechnet, dass eine Maßnahme umgesetzt wird, multipliziert mit den eingesparten CO₂-Emissionen der jeweiligen Maßnahme. Die so berechnete Klimaschutzwirkung muss zudem um das **Attitude-Behavior-Gap** bereinigt werden, das je nach Eingriffstiefe der Maßnahme unterschiedlich groß ausfällt und bemisst, wie viele der Personen, die angeben, etwas umzusetzen, dies schlussendlich auch tun. Zudem muss ein gewisser **Selection-Bias** beachtet werden, da die Umfrage eher von besonders motivierten und interessierten Nutzer*innen beantwortet wird oder von solchen, die an etwaigen Incentives interessiert sind. Die Stichprobe kann daher nicht als repräsentativ gewertet werden, wodurch die Generalisierbarkeit der Ergebnisse auf die Gesamtbevölkerung relativiert wird.

4.6. Bestehende Verwertungsrisiken

Die Verwertung des Prototyps im Rahmen eines gemeinwohlorientierten Geschäftsmodells ist mit einigen Verwertungsrisiken verbunden, die abschließend kurz dargelegt werden sollen. So kann zum Berichtszeitpunkt noch nicht beurteilt werden, ob für die Nutzung des smarten HeizChecks eine **entsprechend hohe Reichweite erlangt werden kann**. Die Nutzungszahlen des Beratungsangebots von co2online sind mitunter stark von **externen Faktoren** abhängig. Dazu zählen unter anderem die mediale Aufmerksamkeit für Themen rund um den Klimaschutz zuhause, die Entwicklung der Energiepreise und damit verbundene Kosten bei Verbraucher*innen sowie die Entwicklung und Positionierung von Konkurrenzprodukten am Markt wie beispielsweise das Portal www.heizung.de der Viessmann Climate Solutions Berlin GmbH. Ein zweites großes Risiko besteht in der **langfristigen Sicherung der Finanzierung** des Betriebs der Anwendung, der u. a. durch notwendige Weiterentwicklungen und Wartungskosten der KI-Komponente mit aller Voraussicht kostenintensiv sein wird. Die gemeinnützige Beratungsgesellschaft co2online ist von öffentlichen Geldern abhängig, die projektbezogen bewilligt werden. So wird beispielsweise der Heizspiegel für Deutschland vom Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit gefördert. Des Weiteren ist noch nicht absehbar, ob durch die Nutzer*innen ein ausreichend großer Umfang an Spendeneinnahmen generiert werden kann, um einen angemessenen Teil der Fix- und Weiterentwicklungskosten zu tragen. Auch sind private Mittel – etwa von Unternehmen – für ein Angebot, das anbieterneutral sein soll, ausgeschlossen. Die beschriebenen Finanzierungsrisiken könnten dazu

führen, dass der Betrieb des Online-Ratgebers dauerhaft untergedeckt bleibt, sofern das Angebot kostenfrei zur Verfügung gestellt werden soll. Durch die vorab skizzierten Verwertungsansätze können diese Risiken verringert, aber nicht vollständig gelöst werden.

5. Fazit und Empfehlungen an die Politik

Ziel des Forschungsvorhabens war es, eine prototypische Anwendung zur **automatisierten Erkennung und Bewertung von HKAs** zu entwickeln. Dabei sollten KI-Bilderkennungsverfahren so weiterentwickelt werden, dass die Kennwerte der nicht-standardisierten HKAs automatisch extrahiert werden können. Dieser Vorgang sollte es Verbraucher*innen ermöglichen, auf Basis ihrer HKA in einer Web-Applikation **schnell und einfach transparente Verbrauchsinformationen zu erhalten und dazugehörige Sparpotenziale zu heben**.

Im Folgenden werden zunächst in Kapitel 5.1 die **wesentlichen Ergebnisse** des Forschungsvorhabens zusammenfassend dargestellt und eingeordnet. Anschließend werden in den Kapiteln 5.2. - 5.4. Schlussfolgerungen aus den erzielten Ergebnissen gezogen und **Empfehlungen** für weitere KI-basierte Forschungsprojekte sowie an die Politik abgeleitet.

5.1. Entwicklung des Prototyps zur smarten Heizkostenanalyse mit Hilfe von KI

Im Ergebnis wurde in diesem Vorhaben der neue Prototyp des KI-basierten smarten HeizChecks durch die interdisziplinäre Zusammenstellung des Verbunds und unter kontinuierlicher Nutzereinbindung erfolgreich getestet und entwickelt. Ein besonderes Augenmerk wurde auf die **Ausgestaltung der Ergebnisse der Heizkostenanalyse** auf der Ergebnisseite des Online-Ratgebers gelegt. Nutzer*innen erhalten hier eine prägnante, aber umfassende Auswertung der Analyseebenen 1) Persönlicher Verbrauch, 2) Energiekosten und Dienstleistungen sowie 3) Zustand des Gebäudes. Direkt **integrierte und individualisierte Handlungsempfehlungen** unterstützen Nutzer*innen dabei, ihren Heizenergieverbrauch und die Heizkosten nachhaltig zu reduzieren. Die Analyse steht Nutzer*innen zum Berichtszeitpunkt gebührenfrei und ohne weitere Dateneingabe zur Verfügung. Damit wurde ein niedrigschwelliges Angebot für Verbraucher*innen geschaffen, um die **Ziele des Verbraucher- und Klimaschutzes voranzutreiben**.

Die vielfältigen **Methoden der Nutzerzentrierung** (Fokusgruppengespräche, Befragungen, UX-Tests) konnten erfolgreich eingesetzt werden, um blinde Flecken in der Anwendungsentwicklung kontinuierlich über den Projektverlauf hinweg zu identifizieren und einen Service zu entwickeln, der den **Bedürfnissen der Verbraucher*innen entspricht**. So wurde mehrfach verifiziert, welche Entwürfe bzw. Elemente des smarten HeizChecks von Nutzer*innen

verstanden und akzeptiert werden, und identifiziert, an welchen Stellen sie Herausforderungen begegnen.

Die Ergebnisse der Nutzererhebungen zeigen jedoch auch, dass die Bewertung des Heizenergieverbrauchs ein für Verbraucher*innen komplexes Thema bleibt und **ein hoher Beratungsbedarf** besteht. Auch zeigt sich, dass langanhaltenden/persistenten Verhaltensveränderungen im Bereich Heizen trotz personalisierten und motivierenden Handlungsempfehlungen eine Reihe von Hürden wie etwa der mögliche Komfortverlust oder hohe Aufwände bzw. Kosten entgegenstehen. Dieses Ergebnis demonstriert einen weiterhin hohen Forschungsbedarf hinsichtlich Verbrauchermotivation und Verhaltensveränderung beim Thema Heizen.

Die Projektergebnisse verweisen hinsichtlich der Erkennungsrate und des Abschlusses der Online-Beratung (i.e. Conversion) mit Hilfe der Bilderkennung auf **die Potenziale der erforschten KI-Methoden zur Unterstützung des Verbraucheralltags** (consumer enabling technologies). In dem Forschungsvorhaben konnte gezeigt werden, dass Verfahren der KI grundsätzlich dafür geeignet sind, die **Datenerfassung** in dem Online-Ratgeber zur Heizkostenanalyse für Mieter*innen und Wohnungseigentümer*innen mit zentraler Heizanlage erheblich zu **erleichtern**. So erhielten in der Beta-Phase des Prototyps **47 % der Nutzer*innen, die die KI-Bildanalyse nutzen, ein Analyseergebnis**. Durch manuelle Eingabe aller Kennwerte gelang es lediglich 31 % der Nutzer*innen zum Ende der Anwendung zu kommen. Bei einem vergleichbaren Service wie dem bestehenden Online-Ratgeber des HeizChecks von co2online liegt die Conversionrate gemittelt über mehrere Jahre bei etwa 32 %. Der Status quo der erfolgreich abgeschlossenen Beratungen zu den Heizkosten konnte mithin deutlich verbessert werden. Im Vergleich zum ursprünglichen Angebot des HeizChecks wurde somit eine **Verbesserung der Conversion** um immerhin 15 Prozentpunkte erreicht, sofern die KI-Bildanalyse genutzt wurde.

Es ist im Projektzeitraum allerdings nicht gelungen, die durch KI zu extrahierenden 34 Kennwerte der HKA in hinreichender Qualität zu erkennen und so die **Erfordernisse der manuellen Dateneingabe** für Nutzer*innen auf ein Minimum zu reduzieren. Zum Projektabschluss sind weitere Hypothesen zur Optimierung der Erkennungsqualität vorhanden, die im Zuge erneuter Trainingsiterationen erprobt werden könnten. Um dieser Herausforderung zu begegnen, erhalten Nutzer*innen im Prototypen **umfassende, dynamische Hilfestellungen**, die sie bei Eingabe der Kennwerte aus der HKA unterstützen. Die Entwicklung dieses dynamischen Hilfsangebot entspringt ebenfalls der Nutzerzentrierung und stellt eine große Weiterentwicklung zum ursprünglichen Angebot des HeizChecks dar.

Für eine zukünftige **Verbesserung der Erkennungsqualität** ist die **Erweiterung des Datenstamms zum Training des Neuronalen Netzes** von zentraler Bedeutung. Die

Datengrundlage zur Bildanalyse kann vorrangig dadurch erweitert werden, dass zusätzliche HKAs in der Betriebsphase des entwickelten smarten HeizChecks durch Nutzer*innen hochgeladen werden. Der so erweiterte Datenstamm könnte dann erneut analysiert und für weitere Trainingsiterationen aufbereitet werden. In jedem Fall ist von einem **hohen jährlichen Wartungs- und Pflegeaufwand für die KI-Komponente** der entwickelten (prototypischen) Anwendung auszugehen.

Ein weiteres Ziel und Qualitätskriterium des Forschungsvorhabens war es, **die Heizkostenanalyse für Verbraucher*innen zu beschleunigen**. Nach aktuellem Stand ist davon auszugehen, dass der entwickelte Prototyp diesem Anspruch nicht ausreichend gerecht wird. Das liegt zum einen daran, dass die Nutzer*innen noch immer viel Zeit damit verbringen, die nicht richtig erkannten Verbrauchsinformationen ihrer HKA manuell in die Eingabefelder der Anwendung zu übertragen. Zum anderen geht die KI-Analyse mit erheblichen **Rechenkapazitäten und diesbezüglichen Wartezeiten auf Nutzerseite** einher. So ist die Verarbeitungsdauer der Bilderkennung zum Berichtszeitpunkt mit durchschnittlich 20 - 40 Sekunden pro verarbeiteter Seite noch recht hoch. Die Ergebnisse der Nutzertests in der Beta-Phase zeigen zwar, dass die Wartezeit für viele kein Abbruchkriterium darstellt. Allerdings bedarf die erforderliche Rechenkapazität ein hohes Maß **an kostenintensiver Hardwareressourcen** (idealerweise GPU-Kerne).

In Hinblick auf die Kosten der Wartung und Hardware für die KI-Komponente stellt die **Refinanzierung KI-gestützter Dienstleistungen für die Zwecke des Verbraucherschutzes** eine ungelöste Herausforderung dar. Im Zuge der Verwertung der Projektergebnisse ist weiter zu prüfen, wie ein gemeinwohlorientierter Service trotz dieses Kostendrucks refinanziert und für Verbraucher*innen kostengünstig zur Verfügung gestellt werden kann. Hierbei muss ein Kompromiss zwischen zwei legitimen Anliegen gefunden werden: Auf der einen Seite der Notwendigkeit, **Einnahmen für die Kostendeckung** zu generieren. Und auf der anderen Seite, die **Unabhängigkeit** des Informationsangebots nicht zu untergraben oder durch zu hohe **Selbstkostenbeiträge** vulnerable Verbraucher*innen vom Angebot auszuschließen.

5.2. Nutzerzentrierung bei KI-Entwicklungen

Es ist gängige Praxis, dass die Güte einer KI derzeit fast ausschließlich daran bemessen wird, wie gut und wie schnell sie bestimmte Daten erkennt und weiterverarbeitet. Die Frage danach, ob sie ein bestimmtes Problem von Verbraucher*innen löst oder eine hohe Nutzerfreundlichkeit aufweist, ist zumindest in der KI-Forschung bislang unzureichend beleuchtet bzw. gewinnt dort erst in jüngster Zeit an Bedeutung (Auer & Felderer, 2018, Beede et al. 2020).

Das Vorhaben Smart_HEC demonstriert jedoch die Notwendigkeit, diese Perspektive von Vorhabenbeginn an in den Blick zu nehmen, da verschiedene Erhebungsschritte gezeigt haben, dass selbst die beste KI ihr Ziel für die Verbraucher*innen verfehlt, wenn sie nicht in eine Anwendung integriert ist, die den **Anforderungen an hohe Nutzerfreundlichkeit, einfache Bedienbarkeit und gutem Datenschutz** genügt. Das übergeordnete Ziel, den Heizenergieverbrauch zu senken und so das Klima zu schonen, kann somit nur erreicht werden, wenn KI-Entwicklung und Nutzerzentrierung eng verzahnt und gemeinsam gedacht werden. Die konsequente Sicherstellung der Nutzerperspektive sowie die Einbeziehung von Nutzer*innen im Entwicklungsprozess sollte daher auch bei zukünftigen verbraucherorientierten Anwendungen berücksichtigt werden.

5.3. Strategische Steuerung des Einsatzes von KI-Verfahren

Der Einsatz von **KI-Verfahren zur Dokumentenverarbeitung** und Informationsextraktion ist ein **sehr dynamisches Forschungsfeld**, an dem sich international eine Reihe namhafter Institutionen und auch große Konzerne mit beachtlichen Ressourcen beteiligen. Es erscheinen in schneller Folge neue Ansätze und KI-Architekturen (und Produkte), die im Rahmen des laufenden Projekts für die Umsetzung nicht durchgängig berücksichtigt werden konnten. Die zu Projektbeginn gewählten State-of-the-Art Modellarchitekturen und insbesondere die zugehörigen Quellcodebibliotheken, auf denen die Projektarbeiten aufsetzen, sind heute in Teilen bereits veraltet bzw. werden nicht mehr gepflegt oder weiterentwickelt – was den Anpassungsaufwand für ihre Nutzung im Rahmen des Prototyps bzw. eines späteren Produktes stark erhöht. Umso wichtiger ist es, **KI-basierte Systeme nur dort einzusetzen**, wo das zu lösende Problem **nicht auf besser geeignete Weise** – beispielsweise durch Gesetzgebung, Regulierung oder Standardisierung – gelöst werden kann.

Bei der Umsetzung KI-basierter Forschungs- und Entwicklungsvorhaben ist eine **stark modularisierte Systemarchitektur** und hinsichtlich der sich wiederholenden Test- und Evaluationsabläufe **automatisierte Herangehensweise** zu verfolgen. Dies gilt auch bezüglich der einzelnen Komponenten innerhalb der KI-Module. Die Pflege aller genutzten Komponenten übersteigt jedoch die **Betriebs- und Wartungskapazitäten** eines einzelnen Betreibers einer solchen Webanwendung. Daher wurde mit OCR-D eine **modulare und quelloffene technologische Grundlage** gewählt, die von einer Entwicklergemeinschaft in Deutschland sowie hier ansässigen Institutionen (z. B. Landes- und Universitätsbibliotheken) getragen wird, und innerhalb derer man sich (perspektivisch) vernetzen, engagieren und mitentscheiden kann (Engl, 2020).

Sofern die OCR-D-Entwicklergemeinschaft die von ihr bereitgestellten Komponenten erfolgreich pflegen und weiterentwickeln kann, können Betreiber darauf aufsetzender Produkte

diese Weiterentwicklungen **lizenzkostenfrei** übernehmen. Darüber hinaus sind die Betreiber in der **Wahl der Dienstleister für Wartung und Pflege** der Produkte grundsätzlich **weniger eingeschränkt** als dies bei proprietären Technologien der Fall wäre. Der **Spezialisierungsgrad** sowie der **Zugang zu bezahlbaren Rechenkapazitäten** für das Trainieren der KI-Komponenten verbleiben als Einschränkungen bzw. Hürden bei der weiteren Pflege und Wartung der KI-Module.

Die im Forschungsprojekt entwickelte KI-Komponente des Prototyps zeigt, in welchem Rahmen die automatische Extraktion von Kennwerten aus HKAs derzeit **allein mit Open Source-Komponenten** möglich ist. Um den entwickelten Anwendungsprototypen dauerhaft zu betreiben, sind **erwartungsgemäß** weitere **Anpassungen** (z. B. an neuartige oder veränderte HKA) sowie turnusmäßige **Wartungsarbeiten** (z. B. Aktualisierung von Komponenten etwa mit Blick auf IT-Sicherheit) erforderlich.⁴⁷

5.4. Weiterentwicklung und Digitalisierung der Verbrauchsinformationen der HKA

Ausgangspunkt des Forschungsvorhabens war, dass die zu Vorhabenbeginn gängige Form von HKAs die **Chancen für Kostenersparnis und Klimaschutz im Verbraucher- und Klimaschutz verfehlten**. Die Gründe dafür sind vielfältig: Die Abrechnungen sind für Fachfremde oft **nicht nachvollziehbar und nicht standardisiert**. So verwendet jeder MDL unterschiedliche Formate und Begrifflichkeiten. Zudem existiert keine Norm, die den Aufbau und die Form von HKAs festlegt. Wichtige Werte und Bezugsgrößen sind nicht herauslesbar (z. B. Energieverbrauch pro m²/pro Wohnung) und somit auch nicht vergleichbar. HKAs tragen daher bisher wenig dazu bei, Transparenz bei Verbrauch und Kosten herzustellen und insbesondere die Ursachen für hohe Kosten zu benennen, wie etwa einen schlechten energetischen Zustand des Gebäudes und/oder das eigene Verbrauchsverhalten. Bisherige Arbeiten des Verbundpartners co2online zeigen das enorme Einsparpotenzial von nachvollziehbaren Verbrauchsinformationen im Bereich des Heizens auf. So wurde in dem Forschungsbericht „Informative und transparent Heizkostenabrechnung als Beitrag für den Klimaschutz“ für das Umweltbundesamt bereits dargelegt, wie sich die HKA zu einem informativeren und transparenteren Instrument weiterentwickeln lässt und es wurde eine Musterabrechnung hierfür erarbeitet (Keimeyer et al., 2016).

⁴⁷ Das Thema „IT-Sicherheit“ wurde im Rahmen des Forschungsvorhabens nicht berücksichtigt, weswegen der entwickelte Prototyp nur in einer (vorhandenen) abgesicherten Umgebung betrieben und sein Betrieb kontinuierlich überwacht werden sollte.

Durch die zum 01.12.2021 in Kraft getretene **Novelle der HeizkostenV**⁴⁸ wurden einige der skizzierten Probleme bereits durch gesetzliche Vorgaben verbessert. So ist in § 6a (3) geregelt, in welcher Form Abrechnungs- und Verbrauchsinformationen durch Gebäudeeigentümer*innen für Abrechnungszeiträume ab dem 01.12.2021 zur Verfügung gestellt werden müssen. Die Informationen werden u. a. ergänzt durch Angaben zu Steuern, Abgaben und Zöllen (mithin dem CO₂-Preis), Kontaktinformationen zu weiterführenden Beratungsangeboten sowie einem Vergleich des Verbrauchs des/der Nutzer*in mit einem Durchschnittsnutzer. Auch der geplante Vergleich des witterungsbereinigten Energieverbrauchs zum vorherigen Abrechnungszeitraum dürfte **für mehr Transparenz bei Verbraucher*innen sorgen**. Diese Entwicklungen werden vom Verbund grundsätzlich positiv bewertet und begrüßt. Zeitgleich bleibt die Novelle der HeizkostenV in einigen Bereichen hinter den Erwartungen zurück. So wurden nach wie vor **keine Standards** für die Abrechnungslayouts und verwendeten Begrifflichkeiten definiert. Auch liegt der Fokus der Verbrauchsinformationen und dazugehöriger Vergleiche stark auf dem/der einzelnen Nutzer*in bzw. dem einzelnen Haushalt, während der **Gebäudeverbrauch** nicht bewertet und anhand verständlicher Vergleichswerte (sog. Benchmarks) eingeordnet wird. Die so definierten Mindestanforderungen sind ergo **nicht geeignet, eine umfassende und unabhängige Heizkostenanalyse, wie die des smarten HeizChecks, zu ersetzen**.

Um dem Anspruch zu genügen, dass Verbraucher*innen selbst oder über (unabhängige) Dritte eine über die gesetzlichen Anforderungen hinausgehende Analyse ihrer Heizkosten/ihrer Heizenergieverbrauchs durchführen können, sollten **die Daten, die der HKA zu Grunde liegen, auf Basis verpflichtender, offener Standards den Nutzer*innen zur Verfügung gestellt werden. Dies gilt ausdrücklich auch für die Kostendaten**. Da Mieter*innen und Wohnungseigentümer*innen für die Erstellung der HKA bezahlen, haben Sie auch einen Anspruch darauf, diese Daten selbst weternutzen zu können – in etwa für eine individuelle Energieberatung oder ähnliche Dienstleistungen.

Mit der Novelle der HeizkostenV ist vorgesehen, dass HKAs künftig in elektronischer Form – also als PDF-Datei – vorgelegt, statt in Papierform per Post verschickt werden. Dieser erste Schritt der Digitalisierung der Abrechnungen erleichtert zwar die Extraktion der für eine Heizkostenanalyse nötigen Kennwerte, bleibt allerdings hinter der Forderung nach einer offenen Datenschnittstelle zurück. Der Verbund sieht in der **Digitalisierung der HKA** einen notwendigen Schritt, um verhältnismäßig aufwändige Verfahren der Datenextraktion mit Hilfe

⁴⁸ Vgl.: https://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBI&jumpTo=bgbl121s4964.pdf#_bgbl_%2F%2F%5B%40attr_id%3D%27bgbl121s4964.pdf%27%5D_1639152568180

von KI obsolet zu machen. Die diesbezüglichen Empfehlungen sind mit einem absteigenden Grad an Digitalisierung wie folgt:

- **API-Schnittstelle:** Die Daten der HKA sollten über eine API-Schnittstelle verfügbar gemacht werden, sodass von dem/der Nutzer*in bevollmächtigte Dritte direkt auf die Daten zugreifen können. Dieser Schritt würde die Datenextraktion durch KI mittels Bildanalyse überflüssig machen.
- **Maschinenlesbares Format:** Die den Heizkosten zugrunde liegenden Verbrauchsinformationen sollen in einem leicht maschinenlesbaren Format (z. B. CSV, JSON, XML) mit einheitlicher Vorgabe der Kennwertbezeichnung verfügbar gemacht werden. Auch dieser Schritt würde die Datenextraktion durch KI mittels Bildanalyse überflüssig und die Datenverarbeitung sehr viel einfacher machen. Ein auf jeder HKA aufgedruckter QR-Code könnte analog zum digitalen Impfausweis zum Abruf der Daten führen.

Alternativ könnten auch Anforderungen an PDF-Dateien gestellt werden, in deren Form die HKA übermittelt wird:

- **PDF-Datei mit eingebetteten Daten:** Eine geeignete Möglichkeit besteht darin, die HKA als PDF-Datei zu erzeugen, wobei die erforderlichen Kennwerte zusätzlich als Datenobjekte eingebettet werden und direkt ausgelesen werden können.⁴⁹ Für das Auslesen der Daten wäre dann keinerlei KI oder maschinelles Lernen erforderlich, insbesondere würde die Bildanalyse von HKA-Dokumenten entfallen. Ein Einsatz von OCR wäre ebenfalls nicht mehr erforderlich.
- **Digital native PDF-Datei (Preprint):** Soweit die HKA wie jetzt geplant nur als digital native PDF-Datei verfügbar gemacht werden kann (vergleichbar mit heutigen verbraucherorientierten digitalen Rechnungen etwa im Telekommunikationssektor) ist die im Forschungsvorhaben entwickelte Datenerfassung mit Hilfe von KI allerdings noch nahezu in vollem Umfang nötig (computer-visuelle Analyse der HKA-Dokumente, inkl. Layoutanalyse, Segmentierung, Erkennung und Klassifikation von Kontext- und Kennwertregionen). Dies gilt insbesondere, sofern keine einheitlichen Tabellenstrukturen für die Kennwerte vorgegeben werden und solange die einzelnen MDL weiterhin individuelle Layouts für die HKAs nutzen. Die Analyse könnte allerdings dadurch vereinfacht werden, dass die Qualität der Dokumente einheitlich ist und eine OCR auf den

⁴⁹ Im Bereich der elektronischen Rechnungen verfolgt der Standard ZUGFeRD ebenfalls den Ansatz einer hybriden Rechnung. Das bedeutet, strukturierte maschinenlesbare Daten werden (zumeist als XML-Datenobjekte) in PDF-Dokumente eingebettet. Herausforderung hierbei ist allerdings, die Konsistenz des menschenlesbaren PDF-Dokuments mit den enthaltenen maschinenlesbaren Daten zu gewährleisten bzw. (automatisch) prüfen zu können. Weitere Informationen: <https://www.ferd-net.de/standards/was-ist-zugferd/was-ist-zugferd.html>

Dokumenten entfallen kann, da die textuellen Inhalte und numerischen Daten (in Textform) durchgängig in den PDF-Dokumenten enthalten wären.

Es kann zum Berichtszeitpunkt noch nicht abschließend bewertet werden, wie die MDL die neuen Vorgaben in die Praxis umsetzen. Die skizzierten Forderungen nach weiterer Transparenz in den Verbrauchsinformationen sowie der Digitalisierung der Abrechnungen sollten künftig auf politischer Ebene diskutiert werden. Dadurch ließen sich die **Ziele des Verbraucher- und Klimaschutzes gerade hinsichtlich der aktuellen Klimaschutzherausforderungen verbinden.**

6. Literaturverzeichnis

- Abdulla, W. (2017). Mask R-CNN for object detection and instance segmentation on Keras and TensorFlow. In GitHub repository. Github.
- Adena, M. (2018). Nudging und soziales Engagement: wie kann Nudging dazu beitragen, die Spendenbereitschaft zu erhöhen?. *Vierteljahrshefte zur Wirtschaftsforschung*, 87(2), 139-152.
- Adena, M., & Huck, S. (2020). Online fundraising, self-image, and the long-term impact of ask avoidance. *Management Science*, 66(2), 722-743.
- Altmann, S., Falk, A., Heidhues, P., Jayaraman, R., & Teirlinck, M. (2019). Defaults and donations: Evidence from a field experiment. *Review of Economics and Statistics*, 101(5), 808-826.
- Auer, F., & Felderer, M. (2018). Shifting Quality Assurance of Machine Learning Algorithms to Live Systems. In M. Tichy, E. Bodden, M. Kuhrmann, S. Wagner, & J.-P. Steghöfer, *Software Engineering und Software Management 2018* (pp. 211–212). Bonn: Gesellschaft für Informatik. Retrieved from <https://dl.gi.de/handle/20.500.12116/21162>
- Baur, N., & Blasius, J. (2014). *Handbuch Methoden der empirischen Sozialforschung* (2014th ed.). Springer VS.
- Beede, E., Baylor, E., Hersch, F., Iurchenko, A., Wilcox, L., Ruamviboonsuk, P., & Vardoulakis, L. M. (2020). A Human-Centered Evaluation of a Deep Learning System Deployed in Clinics for the Detection of Diabetic Retinopathy. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–12. Honolulu, HI, USA: ACM. doi: 10.1145/3313831.3376718
- Bekkers, R., & Wiepking, P. (2010). A Literature Review of Empirical Studies of Philanthropy. *Nonprofit and Voluntary Sector Quarterly*, 40, 924 - 973.
- Dietrich, T., Trischler, J., Schuster, L., & Rundle-Thiele, S. (2017). Co-designing services with vulnerable consumers. *Journal of Service Theory and Practice*.
- Dinter, B., Funk, L., & Pagel, S. (2014). Der verletzte Verbraucher im E-Commerce, in: Bala, C. & Müller, K. (2014). *Der verletzte Verbraucher. Die sozialpolitische Dimension der Verbraucherpolitik, Beiträge zur Verbraucherschutzforschung, Band 2*, S. 123–145; Verbraucherzentrale NRW, Düsseldorf
- Ellermann, E., Noll, C., Weiß, U., Kraus, A., Pehnt, M., Hermann, L., Fleischmann, C. Ifflaender, D. (2020). *Innovative Konzepte zur Stärkung der Energieberatung für Wohngebäude*

Engl, E. (2020). OCR-D kompakt: Ergebnisse und Stand der Forschung in der Förderinitiative. *Bibliothek Forschung Und Praxis*, 44(2), 218–230. doi: doi:10.1515/bfp-2020-0024

European Commission (2016). Consumer vulnerability across key markets in the European Union.

Keimeyer, F. Hennig, P., Jank, S., Metzger, S., Lück, M. & Seidensal, O. (2016). Informative und transparente Heizkostenabrechnung als Beitrag für den Klimaschutz. *Climate Change*, 01/2016. <https://www.umweltbundesamt.de/publikationen/informative-transparente-heizkostenabrechnung-als>

Gerpott, T. (2017). Pay-What-You-Want pricing: An integrative review of the empirical research literature. *Management Science Letters* , 7(1), 35-62.

Goette, L., & Tripodi, E. (2021). Social influence in prosocial behavior: evidence from a large-scale experiment. *Journal of the European Economic Association*, 19(4), 2373-2398.

He, K., Zhang, X., Ren, S., Sun, J., He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770–778. doi: 10.1109/CVPR.2016.90

He, K., Gkioxari, G., Dollár, P., & Girshick, R. B. (2017). Mask R-CNN. 2017 IEEE International Conference on Computer Vision (ICCV). doi: 10.1109/ICCV.2017.322

Kettner, S.E. & Thorun, C. (2019). Verbraucherstudie 2019: Wie erreicht man Verbraucherinnen und Verbraucher im Zeitalter digitaler Informationsangebote. Endbericht.

Lampe, B., Ziemann, K., & Ullrich, A. (2015). *Praxishandbuch Online-Fundraising: wie man im Internet und mit Social Media erfolgreich Spenden sammelt* (p. 188). transcript Verlag.

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars, *Computer Vision – ECCV 2014* (pp. 740–755). Springer International Publishing. https://doi.org/10.1007/978-3-319-10602-1_48

Lumbres (2021). 7 Ways To Increase Online Donations. <https://www.elevateclicks.com/the-nonprofit-marketing-blog/7-ways-to-increase-online-donations>.

Machado, F., & Sinha, R. K. (2013). The viability of pay what you want pricing. *Management Science Working Paper*.

Miller, S., Howard, J., Adams, P., Schwan, M., & Slater, R. (2021). Multi-Modal Classification Using Images and Text. *SMU Data Science Review*, 3(3), Article 3. <https://scholar.smu.edu/datasciencereview/vol3/iss3/6/>

Moser, C. (2013). User experience design. In User experience design (pp. 1-22). Springer, Berlin, Heidelberg.

Neudecker, C., Baierer, K., Federbusch, M., Boenig, M., Würzner, K.-M., Hartmann, V., & Herrmann, E. (2019). OCR-D: An End-to-End Open Source OCR Framework for Historical Printed Documents. Proceedings of the 3rd International Conference on Digital Access to Textual Cultural Heritage, 53–58. <https://doi.org/10.1145/3322905.3322917>

Pollach, I., Treiblmaier, H., & Floh, A. (2005). Online fundraising for nonprofit organizations. Transactions on Advanced Research, 1(2), 56-62.

Tan, M., Le, Q. V., Tan, M., & Le, Q. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In K. Chaudhuri & R. Salakhutdinov, Proceedings of the 36th International Conference on Machine Learning (pp. 6105–6114). PMLR. Retrieved from <https://proceedings.mlr.press/v97/tan19a.html>

Tan, M., Pang, R., & Le, Q. V. (2020). EfficientDet: Scalable and Efficient Object Detection. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 10778–10787. doi: 10.1109/CVPR42600.2020.01079

Von Kossak, S., Albrecht, O., Seidensal, O. & Wettengel, J. (2018). ZIELGRUPPEN ONLINE ERREICHEN. Wie Sie über Google & Co die digitale Reichweite steigern.

Wick, C., Reul, C., & Puppe, F. (2020). Calamari - A High-Performance Tensorflow-based Deep Learning Package for Optical Character Recognition. Digital Humanities Quarterly, 14(1), Article 1.