



**Hochschule
Bonn-Rhein-Sieg**
University of Applied Sciences

Fachbereich Informatik
Department of Computer Science

Bachelorarbeit

im Bachelor-Studiengang Wirtschaftsinformatik

**Entwicklung einer Schnittstelle für die Anbindung von
austauschbaren Datenquellen an KI-Algorithmen**

von

Laurenz Anton Dilba

Erstprüfer: Prof. Dr. Matthias Bertram
Zweitprüfer: Prof. Dr. Wolfgang Heiden
Unternehmen: CONET Solutions GmbH

Eingereicht am: 30. November 2022

Inhaltsverzeichnis

1	Problemstellung	1
2	Fragestellung	1
3	Ziele	2
4	Methode	3
4.1	Anforderungen definieren	3
4.2	Konzept entwickeln	3
4.3	Prototypisch umsetzen	3
4.4	Umsetzung evaluieren	3
4.5	Anpassen	4
5	Evaluierungsstrategie	4
6	Literaturverzeichnis	5
7	Gliederung der Bachelorarbeit	5

1 Problemstellung

Jeden Tag werden riesige Mengen an Daten produziert. Im Jahr 2020 wurden weltweit 64,2 Zettabyte produziert.¹ Dies entspricht 64.000.000.000.000 Gigabyte. Doch aus reinen Daten kann nicht direkt Wissen abgeleitet werden. Mithilfe von Datenmanagement und Datenanalyse wird versucht, die Daten soweit aufzubereiten, dass sie durch Menschen und Computer ausgewertet werden können. Je nach Datenmenge und Abweichung der Daten untereinander kann dies ein aufwändiger, langwieriger und damit teurer Prozess sein. Bei größerer Komplexität oder Menge der Daten wird es für Menschen schwerer, Zusammenhänge, Abweichungen und Auffälligkeiten zu erkennen. Dies liegt unter anderem daran, dass Muster von neu erfassten Daten nur aus Erinnerungen aus dem Kurzzeitgedächtnis abgeleitet werden können.² Um dieses Problem zu lösen, wurden Algorithmen entwickelt, die mit großen Datenmengen trainiert werden können um allgemeine Aussagen über die eingegebenen Daten treffen zu können. Je nach Datenquelle und Art der Aussage, die über diese Daten getroffen werden soll, werden unterschiedliche Algorithmen aus dem Bereich der künstlichen Intelligenz benötigt. Bei dem Ansatz, eine bestimmte Art von Datenquelle an eine KI anzubinden, entsteht eine feste Verdrahtung zwischen dem Datenerhebungsalgorithmus und dem KI-gestützten Datenverarbeitungsalgorithmus. Sollte sich entweder die Datenerhebung oder die Auswertung verändern, muss in der Regel der gesamte Prozess überarbeitet werden. Dies kann nur von jemandem durchgeführt werden, der sich mit den Daten, der eingesetzten künstlichen Intelligenz und den dazu programmierten Schnittstellen auskennt.

2 Fragestellung

Das Thema der Bachelorarbeit soll die Entwicklung und Erarbeitung einer Methode sein, mit der die durch die Datenerhebung gesammelten Daten leichter an die KI-gestützte Datenverarbeitung angeschlossen werden können. Jede künstliche Intelligenz braucht als Input Daten in einem bestimmten Datenformat. Dieses kann sich von Algorithmus zu Algorithmus ändern. Beispielsweise brauchen bestimmte Spracherkennungsalgorithmen reinen ASCII-Text oder Bilderkennungsalgorithmen JPEG Dateien mit einer festgelegten Auflösung. Ein Entwickler, der eine KI mit gesammelten Daten trainieren möchte, muss diese Daten vorher genau auf das Format bringen, welches die KI benötigt. Sollte die KI oder der Datenerhebungsalgorithmus ausgetauscht werden, muss der Entwickler darauf achten, dass die produzierten Daten auch weiterhin kompatibel sind und das gewollte Ergebnis liefern. Die Frage ist demnach: „Wie kann ein Entwickler nach Einrichtung der KI die Daten austauschen ohne dabei den gesamten Anschluss neu programmieren zu müssen?“ Ebenso ist die andere Richtung eine zentrale Frage in der Bachelorarbeit: „Wie kann ein Entwickler eine bereits angeschlossene KI mit einer anderen austauschen, ohne die Daten verändern zu müssen?“

¹Statista, [o. D.], Volumen der jährlich generierten digitalen Datenmenge.

²Snyder, 2000, Music and memory: An introduction.

3 Ziele

Der dynamische Austausch der Daten oder der KI soll durch eine Schnittstelle zwischen Datenerhebung und Datenverarbeitung gelöst werden. Die Betreiber der KI sollen lediglich eine Konfigurationsdatei anpassen müssen, um die Daten für die KI vorzubereiten. Ziel der Bachelorarbeit wird es sein, ein Programm zu entwickeln, welches mithilfe dieser Konfigurationsdatei Daten aus einem Datenerhebungsalgorithmus mittels einer REST-API Schnittstelle annimmt und diese automatisch für eine vorgesehene KI präpariert. Im Anschluss soll das Programm mit den vorbereiteten Daten eine Anfrage an die KI senden und das Ergebnis dem Benutzer wieder zurücksenden.

Wichtig bei der Entwicklung ist es, dass sowohl Datenerhebung als auch Datenverarbeitung modular entwickelt werden. Da die Schnittstelle als REST-API entwickelt wird, können zukünftige Entwickler die Schnittstelle nutzen, auch wenn sich die Daten oder die KI's verändern sollten.

Die REST-API wird mithilfe des Python Web-Frameworks Flask entwickelt. Dies bietet eine einfache und schnelle Möglichkeit einen Backend-Server aufzusetzen. Des Weiteren ist Python eine umfangreiche und einfach zu benutzende Programmiersprache für Datenverarbeitung.

Die Konfiguration der Datenverarbeitung wird mit dem Web-Frontend Framework Angular entwickelt. Dort soll der Entwickler eine Schritt für Schritt Anleitung für die Konfiguration der Schnittstelle präsentiert bekommen. Dort kann die Art der Daten, deren Format und die anzuschließende KI ausgewählt werden.

Mithilfe dieser Schnittstelle soll sich der Arbeitsaufwand verringern, der durch die Anbindung neuer KI's entsteht. Damit können erfahrene Entwickler entlastet werden und Personalkosten gespart werden. Ebenso können schnell und ohne großen Mehraufwand alternative Aufbereitungen der Daten oder neue KI's getestet werden.

4 Methode

Die Bachelorarbeit wird mithilfe einer Variante des Wasserfallmodells bearbeitet werden.

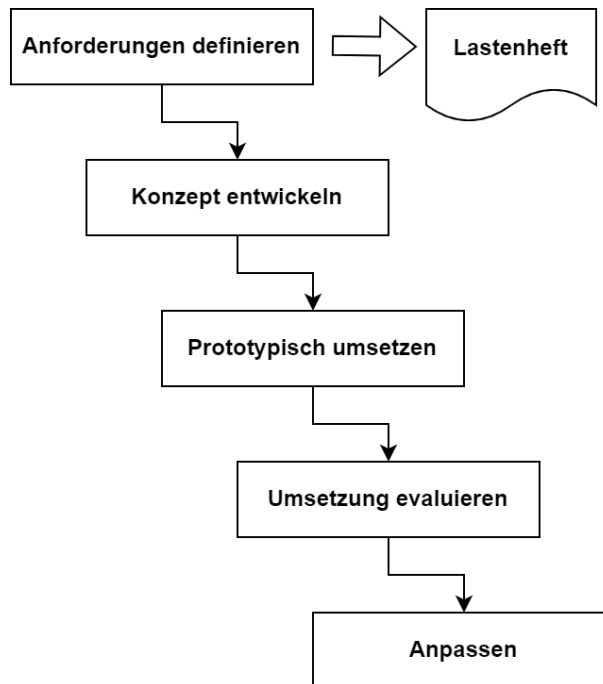


Abbildung 1: Wasserfallmodell Bachelorarbeit

4.1 Anforderungen definieren

Im ersten Schritt werden alle Anforderungen an die zu entwickelnde Software erhoben. Dabei werden sowohl geplante Features, erwünschte Ergebnisse als auch erforderliche Dokumentationen festgehalten. Anschließend wird aus den gesammelten Anforderungen ein Lastenheft erstellt.

4.2 Konzept entwickeln

Im zweiten Schritt des Wasserfallmodells wird das Konzept für die Software erstellt. Das Konzept besteht aus der Softwarearchitektur und dem Programmablauf. Ebenso werden hier alle verwendeten Technologien festgehalten.

4.3 Prototypisch umsetzen

Die anschließend erfolgende prototypische Umsetzung wird den Großteil der Bachelorarbeit ausmachen. Dort wird die Schnittstelle zwischen Datenerhebungs- und Datenverarbeitungsalgorithmen entwickelt und dokumentiert. Ebenso wird die Benutzerschnittstelle für die Einrichtung der Schnittstelle entwickelt.

4.4 Umsetzung evaluieren

Die Umsetzung wird nach der prototypischen Implementierung anhand des Lastenhefts evaluiert. Des Weiteren werden Probanden zur Bedienbarkeit und Verständlichkeit der Schnittstelle befragt. Dazu werden beispielhafte Daten zur Implementation vorbereitet und

eine künstliche Intelligenz zur Auswertung der Daten bereitgestellt. Zur Befragung der Probanden wird ein SUS-Fragebogen verwendet.

4.5 Anpassen

Zuletzt wird in der „Anpassen“ Phase das Feedback der Probanden genutzt, um eventuelle Schwachstellen, Unklarheiten oder Fehler der Software zu beheben.

5 Evaluierungsstrategie

Die Evaluierung der innerhalb der Bachelorarbeit entwickelten Schnittstelle wird zunächst mit dem Lastenheft durchgeführt. Hierbei wird festgehalten, welche der Funktionalitäten implementiert wurden und welche noch nicht entwickelt werden konnten. Anhand der noch fehlenden Features kann auch eine Zukunftsprognose für die Schnittstelle entworfen werden. Nachdem die interne Evaluierung abgeschlossen ist, werden Probanden gebeten, eine testweise Implementation einer KI mithilfe der Schnittstelle durchzuführen. Dazu werden Daten und eine künstliche Intelligenz zur Datenverarbeitung bereitgestellt. Anschließend werden die Probanden mithilfe des System Usability Scale Fragebogens³ nach einer Einschätzung der Bedienbarkeit der Software befragt. Dazu werden mehrere Fragen definiert, die ein Proband auf einer Skala von „Strongly Agree“ bis „Strongly Disagree“ in fünf Abstufungen beantworten kann.

³Administration, [o. D.], System Usability Scale (SUS).

6 Literaturverzeichnis

ADMINISTRATION, U.S. General Services, [o. D.]. *System Usability Scale (SUS)* [online]. [besucht am 2022-07-25]. Abger. unter: <https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html>.

SNYDER, B., 2000. *Music and memory: An introduction*. (MIT press).

STATISTA, [o. D.]. *Volumen der jährlich generierten digitalen Datenmenge* [online]. [besucht am 2022-07-25]. Abger. unter: <https://de.statista.com/statistik/daten/studie/267974/umfrage/prognose-zum-weltweit-generierten-datenvolumen/>.

7 Gliederung der Bachelorarbeit

1. Einleitung
2. Grundlagen
3. Vorgehen im Projekt
4. Projektergebnisse
5. Fazit