

Open Source Lakehouse Container (mittels DuckDB)

Seminararbeit

vorgelegt am 17. Januar 2025

Fakultät Wirtschaft und Gesundheit

Studiengang Wirtschaftsinformatik

Kurs WWI2022F

von

DAVID KREISMANN

DHBW Stuttgart:

Andreas Buckenhofer
Dozent für Data Management

Inhaltsverzeichnis

Abkürzungsverzeichnis	III
Abbildungsverzeichnis	IV
Tabellenverzeichnis	V
1 Grundlagen	1
1.1 Definition von einem Data Lakehouse	1
1.2 Beschreibung von einem Data Lakehouse	1
2 Installation	2
2.1 Werkzeuge	2
3 Beispiel	3
3.1 Data Lakehouse	3
Literaturverzeichnis	4

Abkürzungsverzeichnis

CRM	Customer Relationship Management
DHBW	Duale Hochschule Baden-Württemberg
IEEE	Institute of Electrical and Electronics Engineers
ITIL	IT Infrastructure Library
RoI	Return-On-Invest
UCS	Universal Character Set
UTF-8	8-Bit UCS Transformation Format

Abbildungsverzeichnis

1	DHBW-Logo 2cm hoch	3
---	------------------------------	---

Tabellenverzeichnis

1 Grundlagen

In diesem Kapitel werden die grundlegenden Aspekte des Themas umfassend analysiert und beschrieben. Zunächst erfolgt eine detaillierte Darstellung des Themas, um ein solides Verständnis zu gewährleisten. Anschließend wird die eigene Umsetzung erläutert und in den Gesamtkontext eingebettet.

1.1 Definition von einem Data Lakehouse

In diesem Abschnitt gibt die Abbildungen 1 und ??, die beide das Logo der DHBW zeigen.

1.2 Beschreibung von einem Data Lakehouse

In diesem Abschnitt gibt die Abbildungen 1 und ??, die beide das Logo der DHBW zeigen.

2 Installation

Bald kann nun der Text Ihrer Projekt- oder Bachelorarbeit beginnen. Dank \LaTeX wird Ihre Arbeit garantiert professionell aussehen. Für den Inhalt sind Sie aber weiterhin selbst verantwortlich ;-)

Natürlich ist es schwer, sich vorzustellen, wie das Dokument aussieht, wenn die Vorlage doch gar keinen Text enthält. Aus diesem Grund wird mit Hilfe des Pakets „blindtext“ so genannter Blindtext erzeugt. Mit dem Befehl `\blinddocument` wird nachfolgend ein ganzes Kapitel sinnfreier Blindtext eingefügt.¹

In Abschnitt 2.1 werden die benötigten Werkzeuge erklärt, bevor dann die Verwendung der Vorlage beschrieben wird. Abschnitt ?? gibt Hilfestellungen für bestimmte Fehler. In Kapitel ?? finden sich Beispiele, wie Sie Quellen korrekt zitieren können. In Kapitel ?? werden Abbildungen, Tabellen, ein Code-Listing und auch mathematische Formeln in den Text eingebunden. Ab Seite 4 finden Sie das Literaturverzeichnis.

2.1 Werkzeuge

¹Beachten Sie, dass Sie in Ihrer Arbeit eine Strukturierung wie in Abschnitt 2.1 vermeiden sollten: Dort gibt es einen Abschnitt 2.1.1, aber keinen Abschnitt 2.1.2.

3 Beispiel

In diesem Kapitel werden die grundlegenden Aspekte des Themas umfassend analysiert und beschrieben. Zunächst erfolgt eine detaillierte Darstellung des Themas, um ein solides Verständnis zu gewährleisten. Anschließend wird die eigene Umsetzung erläutert und in den Gesamtkontext eingebettet.

3.1 Data Lakehouse

In diesem Abschnitt gibt die Abbildungen 1 und ??, die beide das Logo der DHBW zeigen.



Abb. 1: DHBW-Logo 2cm hoch.²

²Mit Änderungen entnommen aus: o.V. o.J.

Literaturverzeichnis

o.V. (o.J.): UnbekannterAutor, unbekannterTitel.

Erklärung zur Verwendung generativer KI-Systeme

Bei der Erstellung der eingereichten Arbeit habe ich die nachfolgend aufgeführten auf künstlicher Intelligenz (KI) basierten Systeme benutzt:

1. Consensus
2. ChatGPT-4o

Ich erkläre, dass ich

- mich aktiv über die Leistungsfähigkeit und Beschränkungen der oben genannten KI-Systeme informiert habe,³
- die aus den oben angegebenen KI-Systemen direkt oder sinngemäß übernommenen Passagen gekennzeichnet habe,
- überprüft habe, dass die mithilfe der oben genannten KI-Systeme generierten und von mir übernommenen Inhalte faktisch richtig sind,
- mir bewusst bin, dass ich als Autorin bzw. Autor dieser Arbeit die Verantwortung für die in ihr gemachten Angaben und Aussagen trage.

Die oben genannten KI-Systeme habe ich wie im Folgenden dargestellt eingesetzt:


Arbeitsschritt in der wissenschaftlichen Arbeit	Eingesetzte(s) KI-System(e)	Beschreibung der Verwendungsweise
Literaturrecherche	Consensus	Unterstützung bei der Suche nach wissenschaftlicher Literatur.
Korrektur der Arbeit	ChatGPT-4	Einzelne Kapitel der Arbeit wurden ChatGPT zum Korrigieren gegeben.
Fehleranalyse von Programmcode	ChatGPT-4	Einige Programmzeilen wurden an ChatGPT übergeben, um sie auf Fehler zu prüfen.

³U.a. gilt es hierbei zu beachten, dass an KI weitergegebene Inhalte ggf. als Trainingsdaten genutzt und wiederverwendet werden. Dies ist insb. für betriebliche Aspekte als kritisch einzustufen.

Erklärung

Ich versichere hiermit, dass ich die vorliegende Arbeit mit dem Thema: *Open Source Lakehouse Container* (mittels *DuckDB*) selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Ich versichere zudem, dass die eingereichte elektronische Fassung mit der gedruckten Fassung übereinstimmt.

Stuttgart, 17. Januar 2025
(Ort, Datum)



(Unterschrift)