

# **Datenverarbeitung und Visualisierung von Umrichter- Testbench-Daten**

**Jon Feddersen**

22122017

Bachelorarbeit im Studiengang Elektrotechnik und

Informationstechnik bei

Prof. Dr. rer. nat. Kristina Schädler

Semesteranschrift  
Ochsendrift 31  
25853 Drelsdorf

Studienfach  
Elektrotechnik und  
Informationstechnik

Abgabetermin: 01.10.2025

Fachsemesterzahl: 9

## **Sperrvermerk**

Diese Arbeit enthält vertrauliche Daten und Informationen des Unternehmens, in dem die Bachelor-/Masterarbeit angefertigt wurde. Sie darf Dritten deshalb nicht zugänglich gemacht werden.

Die für die Prüfung notwendigen Exemplare verbleiben beim Prüfungsamt und beim betreuenden Hochschullehrer.

## Inhaltsverzeichnis

<b>Abbildungsverzeichnis</b>	<b>IV</b>
<b>Tabellenverzeichnis</b>	<b>V</b>
<b>Abkürzungsverzeichnis</b>	<b>VI</b>
<b>1. Einleitung</b>	<b>1</b>
<b>2. Grundlagen</b>	<b>3</b>
2.1. Überblick über den Umrichter-Prüfstand . . . . .	3
2.1.1. Aufbau des Teststandes . . . . .	3
2.1.2. Testmodule . . . . .	4
2.1.3. Testablauf . . . . .	5
2.2. Verarbeitung von XML-Daten . . . . .	6
2.2.1. XML-Strukturaufbau . . . . .	6
2.2.2. Verarbeiten von XML-Dateien . . . . .	8
2.2.2.1. XML-Parser . . . . .	8
2.2.2.2. Adressierung mit XPath . . . . .	9
2.2.2.3. Pythonbibliotheken für XML-Verarbeitung . . . . .	11
2.3. Datenbankentwurf und Normalisierung . . . . .	12
2.3.1. Datenbankstruktur . . . . .	13
2.3.2. Datenintegrität und Normalisierung . . . . .	13
2.3.3. Vorgehen zur Erstellung der Datenbank . . . . .	14
2.3.4. Methoden zum Datenbankzugriff in Flask . . . . .	15
2.4. Grundlagen der Datenvisualisierung . . . . .	17
2.4.1. Graphen erstellung mit JS . . . . .	17
2.5. Anforderungen an modulare Softwareentwicklung . . . . .	17
2.5.1. Klare Modulabgrenzung und Verantwortlichkeiten . . . . .	18
2.5.2. Geringe Kopplung und hohe Kohäsion . . . . .	18
2.5.3. Einheitliche Schnittstellen . . . . .	18
2.5.4. Wiederverwendbarkeit . . . . .	18
2.5.5. Erweiterbarkeit und Anpassungsfähigkeit . . . . .	19
2.5.6. Testbarkeit . . . . .	19
2.5.7. Konsistenz und Standardisierung . . . . .	19
<b>3. Analyse und Konzeption</b>	<b>20</b>
3.1. Definition der Anforderungen . . . . .	20
3.1.1. Funktionale Anforderungen . . . . .	20
3.1.2. Nicht-funktionale Anforderungen . . . . .	21
3.1.3. Anforderungen an die Datenbank . . . . .	22
3.2. Grundlegender Ablauf des Programmes . . . . .	23
3.3. Analyse der generierten XML-Berichte und bestehenden Strukturen . . . . .	24
3.3.1. XML-Berichtsstrukturschema . . . . .	24
3.4. Datenbankdesign und Strukturkonzeption . . . . .	29
3.5. Grundkonzept des Benutzeroberflächen-Design . . . . .	31
3.6. Entwurf der Applikationsarchitektur . . . . .	34
3.6.1. Architekturübersicht . . . . .	34
3.6.2. Kommunikation und Datenfluss . . . . .	37

<b>4. Implementierung</b>	<b>38</b>
4.1. Einlesen und Verarbeiten von XML-Daten . . . . .	38
4.2. Implementierung der Datenbank . . . . .	38
4.3. Entwicklung der Benutzeroberfläche . . . . .	38
4.4. Technische Details zur Visualisierung . . . . .	38
<b>5. Integration und Test</b>	<b>39</b>
5.1. Einbindung in die bestehende Systemlandschaft . . . . .	39
5.2. Testmethoden und Durchführung . . . . .	39
5.3. Ergebnisse der Testmethoden . . . . .	39
<b>6. Fazit und Ausblick</b>	<b>40</b>
6.1. Zusammenfassung der Ergebnisse . . . . .	40
6.2. Kritische Bewertung . . . . .	40
6.3. Möglichkeiten für zukünftige Erweiterungen . . . . .	40
<b>i. Literaturverzeichnis</b>	<b>i</b>
<b>ii. Anhangsverzeichnis</b>	<b>iii</b>
<b>iii. Erklärung</b>	<b>iv</b>

## Abbildungsverzeichnis

1.	Aufbau des Teststandes . . . . .	4
2.	XML Prolog Beispielcode . . . . .	7
3.	XML-Elemente Beispielcode . . . . .	7
4.	XML-Attribut Beispielcode . . . . .	7
5.	Beispiel XPath-Ausdruck . . . . .	11
6.	Grundlegendes Ablaufdiagramm der Nutzung der Web-Applikation . . . . .	24
7.	Aufbau unveränderlicher XML-Berichtstruktur . . . . .	25
8.	Elementaufbau aus XML-Bericht . . . . .	25
9.	Direkte Kinderelement unter Stammelement . . . . .	26
10.	XML-Strukturbeispiel Info-Element . . . . .	26
11.	XML-Strukturbeispiel Testbench-Element . . . . .	27
12.	XML-Strukturbeispiel Testmodulheader . . . . .	27
13.	Beispiel XPowerTest Floatblock-Elemente . . . . .	28
14.	Beispiel Fehler bei DriverConsumptionTest . . . . .	28
15.	ER-Modell Überlegung der Datenbankstruktur . . . . .	30
16.	Darstellung der Datenbanktabellen und ihrer Verbindungen . . . . .	31
17.	Benutzeroberflächenentwurf der Seite zum Einlesen der XML-Struktur . . . . .	32
18.	Benutzeroberflächenentwurf der Seite zum Ausgeben der Berichtstabelle . . . . .	33
19.	Benutzeroberflächenentwurf der Seite zum Ausgeben der Graphen . . . . .	34
20.	schematisches Architekturdiagramm der Applikation . . . . .	35
21.	Grundlegenden Struktur des Applikations-Ordners . . . . .	37

**Tabellenverzeichnis**

1.	Übersicht über XPath-Achsen . . . . .	10
2.	Vor- und Nachteile des direkten SQL-Zugriffs in Flask . . . . .	16
3.	Vor- und Nachteile der Verwendung von Flask-SQLAlchemy . . . . .	17

## Abkürzungsverzeichnis

<b>DUTs</b>	Devices-Under-Test
<b>XML</b>	Extensible Markup Language
<b>DTD</b>	Document Type Definition
<b>API</b>	Application Programming Interface
<b>DOM</b>	Document Object Model
<b>SAX</b>	Simple API for XML
<b>FSW</b>	Functional-Switching
<b>OCP</b>	Over-Current-Protection
<b>DSCP</b>	Dynamic-Short-Circuit-Protection
<b>BIT</b>	Burn-In-Test
<b>OTP</b>	Overtemperature-Protection-Test
<b>XSLT</b>	Extensible Stylesheet Language Transformations
<b>XSD</b>	XML Schema Definition
<b>XXE</b>	XML External Entity
<b>SQL</b>	Structured Query Language
<b>ORM</b>	Object-Relational Mapping

## 1. Einleitung

„Data is the new oil. Like oil, data is valuable, but if unrefined it cannot really be used.“ Clive Humby (2006)

Die Aufbereitung von Daten, sowohl in ihrer Struktur als auch in ihrer Darstellung, ist entscheidend, um technische Prozesse zu verstehen und effizient mit großen Datenmengen arbeiten zu können. Das Erfassen, Aufbereiten und Interpretieren von Messdaten ist heute aus der Industrie und Wissenschaft nicht mehr wegzudenken. In nahezu allen technischen Bereichen werden Daten erhoben, sei es in der Forschung, bei Qualitätsprüfungen oder in industriellen Testverfahren. Damit diese Daten nutzbar sind, müssen sie verständlich aufbereitet und übersichtlich dargestellt werden.

Die Verarbeitung und Visualisierung großer Datenmengen erfolgt heutzutage meist mithilfe spezialisierter Softwarelösungen. Viele dieser Systeme sind jedoch allgemein gehalten und nur begrenzt auf spezifische technische Anwendungen anpassbar. Das kann sich negativ auf Effizienz, Benutzerfreundlichkeit und Aussagekraft der Ergebnisse auswirken. Besonders in Bereichen, in denen die Datensätze komplex und individuell strukturiert sind, wie bei der Prüfung von Leistungselektronik aus Windenergieanlagen, besteht Bedarf an speziell angepassten Softwarelösungen.

Ziel dieser Arbeit ist es, eine Web-Applikation zu entwickeln, die automatisch erzeugte XML-Berichte eines Umrichter-Teststandes einliest, die enthaltenen Mess- und Gerätedaten in einer Datenbank speichert und diese anschließend grafisch aufbereitet. Auf diese Weise sollen die Daten besser ausgewertet und sowohl für interne Analysen als auch für externe Berichte genutzt werden können. Dabei liegt der Fokus auf einer übersichtlichen Benutzeroberfläche, einer modularen Erweiterbarkeit und einer einfachen Integration in die bestehende Systemlandschaft des Unternehmens.

Die Arbeit entstand in Zusammenarbeit mit einem Serviceanbieter für die Instandhaltung von Windenergieanlagen. Der zugrunde liegende Teststand wird zur Überprüfung von Umrichtern eingesetzt, die aus bestehenden Anlagen stammen. Mithilfe spezieller Prüfverfahren wird bestimmt, ob diese Geräte nach der Instandsetzung wiederverwendet werden können. Die dabei entstehenden XML-Dateien enthalten sämtliche Messwerte, Parameter und Prüfstandsinformationen und bilden die Grundlage für die zu entwickelnde Software.

Für die Umsetzung wurde ein iteratives Vorgehen gewählt. Das bedeutet, dass die Anwendung schrittweise entwickelt und nach jedem Zwischenschritt getestet und verbessert wird. Dieses Vorgehen orientiert sich an Prinzipien agiler Softwareentwicklung, um flexibel auf mögliche Anpassungen reagieren zu können. So entsteht eine funktionierende Anwendung, die sich im Laufe der Entwicklung immer weiter verfeinern lässt.

Die Arbeit gliedert sich in sieben Kapitel. Nach dieser Einleitung werden in Kapitel 2 die funktionalen und nicht-funktionalen Anforderungen beschrieben. Kapitel 3 behandelt die theoretischen und technologischen Grundlagen, die für die Entwicklung der Anwendung relevant sind. In Kapitel 4 folgen die Analyse der vorhandenen Strukturen und der Entwurf der Systemarchitektur. Kapitel



5 beschreibt die Implementierung der Applikation, während Kapitel 6 die Integration in die bestehende Systemumgebung und die Testdurchführung erläutert. Kapitel 7 fasst die Ergebnisse zusammen und gibt einen Ausblick auf mögliche Erweiterungen.

## 2. Grundlagen

In diesem Kapitel werden die theoretischen Grundlagen für das Entwickeln der Software, so wie das notwendige Verständnis des Teststandes und seine Abläufe behandelt.

### 2.1. Überblick über den Umrichter-Prüfstand

In diesem Abschnitt wird der Umrichter-Teststand, von dem die zu verarbeitenden Datensätze stammen, beschrieben, da dies für das generelle Verständnis der einzulesenden Datensatzstruktur unerlässlich ist. Die genaue Bezeichnung des Teststandes „USTB DWT Test Bench (XCT0006-1)“ wird im Folgenden als Test-Bench oder Teststand benannt. Diese Art Test-Bench wird im Allgemeinen für die End-of-Line-Prüfung von unterschiedlichen Umrichtern nach ihrer Herstellung genutzt, um die Produktqualität und -funktionalität sicherzustellen. [1]

In dem hier vorliegenden Fall wird der Teststand verwendet, um die aus dem Feld kommenden Umrichter auf ihre weitere Nutzungstauglichkeit zu testen. Die weitere Nutzungstauglichkeit wird ermittelt, indem die Messwerte mit Mittelwerten, die von mehreren fabrikneuen Umrichtern stammen, verglichen werden. Diese Messwerte müssen sich in einem vorher definierten Toleranzbereich befinden, um weiter im Feld verwendet zu werden.

Die Umrichter werden in der gegebenen Fachliteratur zur Test-Bench als Devices-Under-Test (DUTs) bezeichnet. Diese Bezeichnung kommt auch in den Berichten auf dem Teststand vor, daher hat der Autor diese Abkürzung übernommen.

#### 2.1.1. Aufbau des Teststandes

Die Test-Bench besteht aus mehreren Komponenten, die sich im Testraum in unterschiedlichen Schaltschränken befinden. Der Teststand besteht aus folgenden Hauptkomponenten: Die Test-Bench besteht aus mehreren Komponenten, die sich im Testraum in unterschiedlichen Schaltschränken befinden, und der Teststand umfasst folgende Hauptkomponenten:

- Das Netzteil wandelt die 400-V-Netzspannung in eine isolierte Gleichspannung für den Zwischenkreis um. Das Netzteil liefert maximal 80 kW mit 1200 V DC oder 800 V DC, welche Werte verwendet werden, kann vor Teststart bestimmt werden. In Abbildung 1 wird das Netzteil als PSU bezeichnet, was für „Power Supply Unit“ steht.
- Das Elektronik-Rack, auf dem die Mess- und Steuerkomponenten befestigt sind, ist ein wichtiger Bestandteil des Teststandes. Hier befindet sich auch der (XCS2100) System-Controller, der das ganze System mit dem PC, auf dem die Test-Bench-Software läuft, via Ethernet verbindet. In Abbildung 1 mit ER bezeichnet, für „Electronic Rack“.
- Der Testmatrix-Schrank, in dem die Sammelschienen für den Stromanschluss und die Schützen sitzen. In Abbildung 1 mit TM bezeichnet, für „Test Matrix cabine“.

- Der Schrank mit dem Kühlungssystem, da die Umrichter während des Betriebes mit Wasser oder Luft gekühlt werden müssen. In Abbildung 1 mit „Cool1“ bezeichnet.
- Der Carrier, auf dem die Umrichter befestigt werden, ist speziell für bestimmte Umrichter konstruiert. Dieser wird speziell für bestimmte Umrichter konstruiert. In Abbildung 1 wird der Carrier als Carrier1 bezeichnet.

Neben den Hauptkomponenten befinden sich außerhalb des Sicherheitsbereiches, der während des Betriebes nicht betreten werden darf, ein PC mit einer Software zum Steuern der Testeinrichtung sowie eine Betriebsanzeige und ein Notaus. [1]

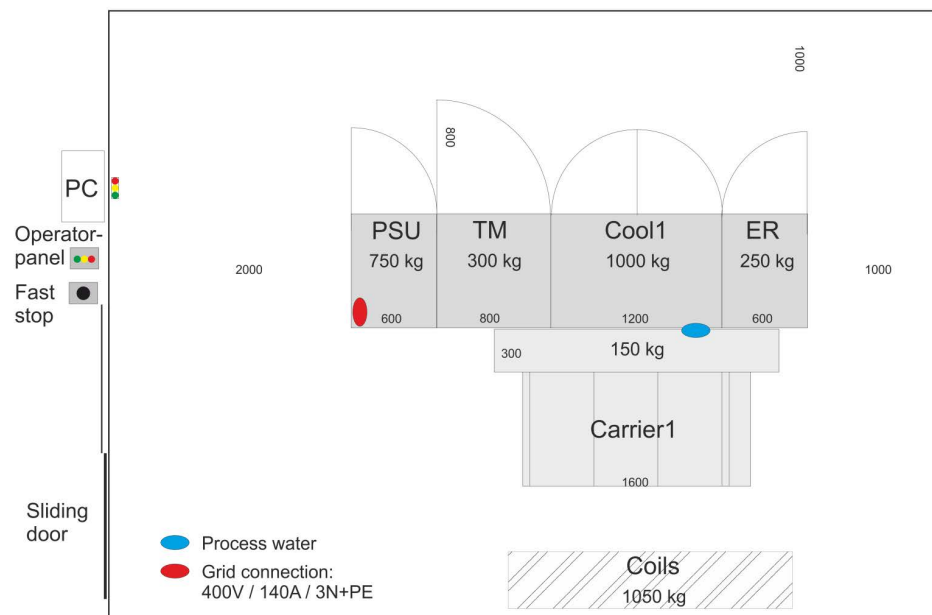


Abbildung 1: Aufbau des Teststandes

Quelle: [1, S. 7]

### 2.1.2. Testmodule

Es gibt mehrere Testmodule, die auf dem Teststand laufen und verschiedene Funktionen der Umrichter testen. Einige der Funktionen eines DUT können mit dem gleichen Modus eines Testmoduls überprüft werden, indem die entsprechenden Parameter ausgewählt werden. Jeder Testablauf ist autonom und kann mehrmals ausgeführt werden, auch mit unterschiedlichen Parametern.

Im Folgenden wird eine kurze Beschreibung der Funktionen der für diese Arbeit relevanten Testmodule gegeben.

**Driver Consumption Test:** Der Driver Consumption Test überprüft den Stromverbrauch des Treibers im Leerlauf und während Pulssprüngen.

**Pulse Test:** Der Impulstest verfügt über drei Funktionsmodi.

- Im Modus Functional-Switching (FSW) kann überprüft werden, ob die Halbleiter, die sich in den DUTs befinden, generell schalten.

- Im Modus Over-Current-Protection (OCP) kann die Überstromüberwachung weiche Kurzschlüsse überprüfen. Ein weicher Kurzschluss ist, wenn der Stromfluss nicht sofort und vollkommen unterbrochen wird.
- Im Modus Dynamic-Short-Circuit-Protection (DSCP) wird ein harter Kurzschluss, also ein vollständiger und sofortiger Kurzschluss, überprüft.

Power Test: Mit diesem Test werden zwei verschiedene Funktionen getestet werden:

- Burn-In-Test (BIT) dient dazu, die DUTs zyklisch zu betreiben und so reale Betriebszustände zu simulieren. Zudem kann mit Hilfe dieser Funktion die Kühltemperatur überprüft werden, um die korrekte Wärmeübertragung der Halbleiter sicherzustellen.
- Während des Overtemperature-Protection-Test (OTP) wird ein DUT, ähnlich wie beim BIT, nur mit reduzierter Kühlung betrieben, bis die maximal zulässige Kühlkörpertemperatur erreicht ist und die Temperaturschutzschaltung auslöst.[1]

### 2.1.3. Testablauf

Die Schrittfolge des Testablaufes mit Montage der DUTs ist klar festgelegt und vor jedem Durchlauf gleich. Die Montage läuft wie folgt ab:

1. Die Umrichter werden auf dem Carrier befestigt. Es werden meist 3 dieser Geräte gleichzeitig getestet, Abweichungen je nach Bauform. Die Reihenfolge der elektrischen Phasen ist bei Draufsicht des Carriers von links nach rechts U-V-W. Dies ist für das Layout des Berichtes relevant.
2. Der Kühlkreislauf wird angeschlossen, je nach Umrichtertyp Lüftungs- oder Wasserkühlung.
3. Die DUTs werden mit den AC- und DC-Link-Kontakten verbunden, um das Gerät zu betreiben und die DC-Spannung wieder abzuleiten.
4. Das Anbringen von Signalkabeln zwischen DUT und Merkurbox. Die Merkurbox dient als Schnittstelle zwischen DUTs und Testbench.
5. Das Montieren von VCE-Klemmen an die AC-Kontakte der DUTs.

Vor Beginn des Testablaufs benötigt die Test-Bench noch weitere Informationen zu den DUTs, Testsequenzen und der Hardware-Topologie. Diese Daten können über das Scannen von vorliegenden QR- oder Barcodes eingefügt werden. Alternativ können diese auch über ein Eingabefenster im Teststandprogramm eingetragen werden. Je nach Topologiekonfiguration kann auch ein zweiter Träger mit DUTs gescannt werden, Dies kommt bei dem Test, den das Unternehmen durchführt, nicht vor. [1]

Das eigentliche Testen mit den verschiedenen Testmodulen wird autonom durchgeführt. Im Unternehmen läuft der Test regulär wie folgt ab:

1. Durchführung des Driver Consumption Testes, um die Stromversorgung der Treiberplatine auf den Umrichter zu testen.
2. Durchführung des Pulse Testes im Modus Functional-Switching, dies prüft die Umrichter.
3. Durchführung des Power Testes, in den Berichten auch XPower Test genannt. Dieser Test simuliert die Belastung der Umrichter im Feld und überprüft, ob sich die Werte in einem bestimmten Toleranzbereich befinden.

Nach dem Durchlaufen eines Tests wird automatisch ein Extensible Markup Language (XML)-Datenfile mit den erhobenen Messdaten generiert, die enthaltenen Messwerte und alle vorher bestimmten Einstellungen erstellt und in eine Datei eingefügt. Bei der Demontage nach dem Testdurchlauf werden alle Schritte in umgekehrter Reihenfolge durchgeführt.

## 2.2. Verarbeitung von XML-Daten

Dieses Kapitel behandelt einige grundlegende und für diese Arbeit relevante Aspekte von dem Dokumententyp XML. Da die Messwerte und Berichte, die der Teststand generiert, in diesem Format vorliegen, ist es relevant, dieses Kapitel zu behandeln.

Bei XML handelt es sich um eine Auszeichnungssprache, also eine formale Sprache, die verwendet wird, um die Struktur und Darstellung von Daten oder Texten zu beschreiben [2]. XML wurde entwickelt, um Informationen in einem maschinenlesbaren und strukturierten Format zu speichern und zu übermitteln. Sie wird hauptsächlich in Bereichen wie Webdiensten, Datenbanken und Konfigurationsdateien eingesetzt. XML ermöglicht die hierarchische Organisation von Informationen in einem strukturierten Aufbau und kann sowohl von Menschen als auch von Maschinen interpretiert werden. [3, S. 4]

Das Grundkonzept hinter XML war, eine universell einsetzbare und erweiterbare Sprache zu erschaffen, die von verschiedenen Systemen unabhängig von deren fundamentalen Technologieansatz genutzt werden kann. Hierbei wäre das angestrebte Ziel, Daten in einem einheitlichen Standard zwischen verschiedenen Anwendungen und Plattformen zu speichern und auszutauschen. zu kommen. [3, S. 3–5]

### 2.2.1. XML-Strukturaufbau

Eine XML-Datei beginnt mit einem Prolog, der die XML-Version und die verwendete Zeichencodierung definiert. In Abbildung 2 ist ein häufig genutzter Prolog dargestellt, der auch in den Teststand-Berichten zum Einsatz kommt. Die erste Zeile des Prologs ist die sogenannte XML-Deklaration. Die XML-Deklaration enthält häufig die Attribute „version“ und „encoding“, jedoch ist nur das Attribut „version“ Pflicht. Werden auch die anderen notiert, müssen sie in der angegebenen Reihenfolge deklariert werden. Attribut „version“: Mit „version“ wird die verwendete XML-Version angegeben. Das Attribut „encoding“ gibt die im Dokument verwendete Zeichencodierung an, d. h., mit welcher Codierung die Datei gespeichert wird. Fehlt die Angabe, wird als Vorgabe UTF-8 (8-Bit Unicode Transformation Format) verwendet. Neben der XML-Deklaration

können im Prolog auch noch Verarbeitungsanweisungen und Verweise auf eine Document Type Definition (DTD) deklariert werden. Diese sind jedoch optional und für diese Arbeit nicht weiter relevant. [4, S. 8, 9]

```
1 <?xml version="1.0" encoding="UTF-8"?>
2   <root>
3     ....
4   </root>
```

Abbildung 2: XML Prolog Beispielcode

Quelle: eigene Darstellung

Der Hauptteil eines XML-Dokuments besteht aus einer Reihe von Elementen, die durch Tags markiert sind. Für jedes Element gibt es ein Starttag und ein Endtag, welche das Element beginnen und beenden. Ein Starttag kann beispielsweise so aussehen: „<NamedesTags>“ So aus, dann würde der dazugehörige Endtag „</NamedesTags>“ so aussehen. Der entscheidende Unterschied ist hierbei der Schrägstrich beim Endtag. Der Name des Elementes wird durch den Inhalt der Keiler- und Großer-Zeichen bestimmt. Bei diesem Beispiel wäre der Name „NamedesTags“. Elemente haben einen Inhalt, der aus Text, weiteren Elementen oder aus beidem bestehen kann. Wenn Elemente andere Elemente beinhalten, werden sie als Elternelemente und die enthaltenen Elemente oft als Kindelemente bezeichnet. Diese Eigenschaft der Elemente sorgt dafür, dass XML-Dateien einer hierarchischen Baumstruktur folgen. Hierbei wird das oberste Element als Wurzelement bezeichnet, im Englischen „root“. [4, S. 10–14]

```
1 <buch><!-- Das Element "Buch" enthält 2 Kinderelement-->
2   <titel>XML-Grundlagen</titel>
3   <autor>Max Mustermann</autor>
4 </buch>
```

Abbildung 3: XML-Elemente Beispielcode

Quelle: eigene Darstellung

Jedes Element kann neben Inhalt auch mit beliebig vielen Attributen ausgestattet sein, die zusätzliche Informationen enthalten. Attribute werden im Starttag eines Elements definiert. Diese bestehen immer aus einem Attributnamen und einem Wert. Der Wert wird dabei mit Anführungszeichen deklariert, wie in Abbildung 4 gezeigt. [4, S. 10–14]

```
1 <buch genre="Lehrbuch">ES</buch>
2 <!--Hier ist das Attribut "genre" mit dem Inhalt "Lehrbuch"-->
```

Abbildung 4: XML-Attribut Beispielcode

Quelle: eigene Darstellung

Kommentare werden mit den Tags „<!--“ und „-->“ eingefügt und dienen der Dokumentation oder

dem Hinweis auf bestimmte Teile des Codes [4, S. 10–14]. Dies ist für die automatisch generierten Berichte irrelevant, jedoch für das Beschreiben der Beispiele hilfreich. In Abbildung 3 und Abbildung 4 werden diese zur Beschreibung verwendet. Kommentare werden beim Parsen des Dokuments ignoriert. Parsen und Parser werden im nachfolgenden Kapitel behandelt.

### 2.2.2. Verarbeiten von XML-Dateien

In diesem Abschnitt wird eine Zusammenfassung der grundlegenden Methoden und Techniken gegeben, um XML-Dateien mithilfe verschiedener Tools unabhängig von der verwendeten Programmiersprache zu verarbeiten. Es wird beschrieben, wie man XML-Dateien analysiert, modifiziert und überprüft, um sie für verschiedene Zwecke einsatzbereit zu machen. Am Ende des Abschnittes wird jedoch etwas genauer auf Methoden, die in Python genutzt werden, eingegangen.

**2.2.2.1. XML-Parser** Der erste Schritt beim Verarbeiten einer XML-Datei ist das Parsen. Hierbei wird die XML-Datei in ein Programm geladen und in ein Format umgewandelt, das dieses interpretieren kann. XML-Parser prüfen hierbei auch die XML-Daten auf Korrektheit, also ob die Vorlagen eingehalten werden und das Dokument vollständig ist. Dabei wird zwischen nicht-validierenden und validierenden Parsern differenziert. Der Unterschied besteht darin, dass validierte Parser neben der korrekten Schachtelung und Bezeichnung der Strukturelemente wie die nicht validierten Parser auch noch auf eine Vorgabe einer Dokumenttypdefinition oder eines Schemas prüfen. [4, S. 10]

Parser werden verwendet, um einer Applikation über eine Application Programming Interface (API) eine Schnittstelle zu einem XML-Dokument zu geben. Bei APIs wird in diesem Bereich zwischen zwei Grundtypen unterschieden: [4, S. 405]

#### **Baumbasiert**

Die baumbasierten APIs lesen über den XML-Parser das XML-Dokument ein, parst es und erzeugt ein Modell als Baum von Knoten im Arbeitsspeicher. Auf Grund der im XML-Dokument vorkommenden Informationseinheiten wird in verschiedene Knotentypen unterschieden. Das generierte Modell dient der Applikation für die weitere anwendungsspezifische Verarbeitung. Ein Beispiel für eine baumbasierte API ist Document Object Model (DOM).

DOM ist ein objektorientiertes Modell, das die Struktur eines XML-Dokuments abbildet. In der Baumstruktur wird das gesamte Dokument abgebildet, indem jedes Element, jedes Attribut und jeder Text bzw. Inhalt als Knoten gilt. Das DOM hat den Vorteil, dass es das XML-Dokument vollständig im Arbeitsspeicher darstellt, was das Durchsuchen und Bearbeiten des Dokuments erleichtert. Zudem bietet es eine einfache Schnittstelle, um XML-Daten zu erreichen und zu verändern. Allerdings benötigt diese Herangehensweise viel Speicherplatz, da es das gesamte Dokument im Arbeitsspeicher ablegt, was bei umfangreichen XML-Dokumenten ein Problem darstellen kann. [4, S. 413, 414]

### Ereignisbasiert

Die ereignisbasierten APIs lesen XML-Dokumente sequenziell von Beginn an. Ein und meldet während des Lesens jedes Ereignis durch sogenannte Callbacks an die aufrufende Applikation zurück. Ein Ereignis ist ein Signal, das Änderungen in dem Markup-Status anzeigt. Das bedeutet, Ereignisse traten bei Element-Tags, Zeichendaten, Kommentaren und Verarbeitungsanweisungen sowie bei den Grenzen des Dokumentes auf. Der Parser sendet hierbei durch die Callbacks eine Mitteilung an die aufrufende Applikation, welche Ereignisse eingetreten sind. Das Programm, das den Parser aufgerufen hat, muss nun das Ereignis interpretieren und entsprechend reagieren. Das XML-Dokument wird hierbei nicht vollständig im Arbeitsspeicher gespeichert, sondern in Teilen gelesen und bearbeitet. Deshalb eignen sich ereignisbasierte APIs besonders gut für die Verarbeitung großer XML-Dateien, da dies für eine geringe Arbeitsspeicherbelastung sorgt. Trotz der Effizienz hinsichtlich des Speicherverbrauchs sind ereignisbasierte APIs für die Datenmanipulation nicht sonderlich gut geeignet. Durch das teilweise Einlesen wird nämlich keine umfassende Dokumentstruktur wie beim vollständigen Aufbauen erzeugt. Die baumbasierten APIs, welches die Manipulation bzw. Bearbeitung erschwert. Ein Beispiel für eine bedeutende ereignisbasierte API ist Simple API for XML (SAX). Diese API arbeitet nach dem oben beschriebenen Prinzip und wurde ursprünglich als Java-API entwickelt. Inzwischen gibt es SAX auch für weitere Sprachen wie z. B. C++, Perl und Python. [4, S. 405]

**2.2.2.2. Adressierung mit XPath** Neben dem Einlesen der XML-Datei muss zum Verwenden der Daten in dem XML-Dokument navigiert werden, um bestimmte Knoten zu adressieren. Aus diesem Grund wurde die Abfragesprache XML Path Language, kurz XPath, entwickelt. XPath wird hauptsächlich in der Transformationssprache Extensible Stylesheet Language Transformations (XSLT) eingesetzt, welche XML-Dateien in andere Datenformate umwandeln soll. Zudem wird XPath in anderen Programmiersprachen wie z. B. JavaScript, C oder einigen Pythonbibliotheken für die Adressierung von Bestandteilen des XML-Baumes verwendet.

Im Zusammenhang mit XML und XPath wird oft der Begriff „Knoten“ verwendet. Knoten können Teile des XML-Dokumentes beschreiben, aber auch über den Wurzelknoten das gesamte Dokument. Knoten können alle im XML-Dokument vorhandenen strukturellen Teile sein. Daher kann zwischen den verschiedenen Typen von Knoten im XML-Dokument unterschieden werden:

- Wurzelknoten: Auch Dokumentknoten oder root genannt, sind der Ursprung des XML-Dokuments.
- Elementknoten: ein beliebiges Element.
- Textknoten: ein Text, welcher einem Element untergeordnet ist.
- Attributknoten: ein beliebiges Attribut eines Elements.
- Kommentarknoten: ein beliebiger Kommentar.
- Namensraumknoten: eine beliebige Namensraumangabe eines Elements oder Attributs.
- Verarbeitungsanweisungsknoten: ein XML-Verarbeitungshinweis.



Über diesen Knoten kann man durch die XML-Struktur navigieren, aber um die Knoten zu adressieren, werden Achsen benötigt. Achsen spezifiziert die Beziehung zwischen den Knoten, um so die gewünschten Knoten zu selektieren. Es gibt Achsen, welche nur einen Knoten adressieren, so wie Achsen, welche mehrere Knoten gleichzeitig auswählen. In der folgenden Tabelle 1 werden die verschiedenen Achsen benannt und beschrieben. Für alle Achsen, die keine Kurz-Notation enthalten, muss im Code der Name ganz ausgeschrieben werden. [5]

Name	Kurz-Notation	selektierte Knoten
/	/	Wurzelknoten
child	(nicht notwendig)	direkt untergeordnete Knoten (Kindknoten)
self	.	aktuelle Knoten (Kontextknoten)
parent	..	direkt übergeordneter Knoten (Elternknoten)
descendant	./	alle untergeordnete Knoten
descendant-or-self		alle untergeordnete Knoten sowie der aktuelle Knoten
ancestor		alle übergeordnete Knoten
ancestor-or-self		alle übergeordneten Knoten sowie der aktuelle Knoten
following		alle nachfolgende Knoten (ohne Kindknoten)
following-sibling		alle nachfolgende Knoten (ohne Kindknoten), die den gleichen Elternknoten haben
preceding		alle vorangehende Knoten (ohne alle Elternknoten)
preceding-sibling		alle vorangehende Knoten (ohne alle Elternknoten), die den gleichen Elternknoten haben
attribute	@	Attributknoten
namespace		Namensraumknoten

Tabelle 1: Übersicht über XPath-Achsen

Quelle: eigene Darstellung nach Tabelle aus [5]

Um die Selektierung der Knoten noch weiter einzugrenzen, können sogenannte Prädikate verwendet werden. Prädikate befinden sich immer in eckigen Klammern und in diesen können verschiedene Operatoren genutzt werden. Dazu zählen die mathematischen Operatoren + (Addition), – (Subtraktion), × (Multiplikation), div (Division) und mod (Modulo, Divisionsrest), aber auch die logischen Operatoren and (und) und or (oder) sowie die Vergleichsoperatoren = (gleich) und != (ungleich), < (kleiner als), (kleiner-gleich), > (größer als) und (größer-gleich). Mit diesen Operatoren und Aneinanderreihungen von diesen kann eine Vielzahl von Sucheinschränkungen geschaffen werden, welche die Genauigkeit der Suchen stark erhöhen und variabel gestalten. [5]

Ein XPath-Ausdruck, also ein Path, welcher zur Adressierung von einem oder mehreren Knoten verwendet wird, besteht aus einem oder mehreren Lokalisierungsschritten. Ein Lokalisierungs-

schritt besteht aus weiteren 3 Teilen: der Achse, einem Knotentest und bei Bedarf aus einem oder mehreren Prädikaten. Der Knotentest ist hierbei nichts anderes als der Name des Knotens. Für die korrekte Form des Lokalisierungsschrittes muss der Achsenname durch zwei Doppelpunkte vom Knotentest getrennt werden. Bei Kurz-Notationen können die Doppelpunkte jedoch weggelassen werden, wenn anstelle des Knotentests bzw. dem Knotennamen das Wildcard-Zeichen in einem Lokalisierungsschritt verwendet wird, werden unabhängig vom Namen und von den Achsen alle Knoten selektiert, welche sich unter dem letzten Lokalisierungsschritt befinden. Die Lokalisierungsschritte werden mithilfe des Schrägstrichs „/“ voneinander getrennt. [5] In der folgenden Abbildung 5 ist ein Beispiel für einen XPath-Ausdruck. Dieser XPath würde in einem XML-Dokument mit Büchern das erste Buch auswählen, welches vor 1900 erschienen ist. Geht dann zu nächsten Geschwisterelement weiter und gibt den Titel zurück.

```
1 //buch[jahr<1900]/following-sibling::buch[1]/titel
```

Abbildung 5: Beispiel XPath-Ausdruck

Quelle: eigene Darstellung

**2.2.2.3. Pythonbibliotheken für XML-Verarbeitung** Für die Bearbeitung von XML-Dokumenten gibt es in Python einige Bibliotheken. Diese basieren auf den obengenannten Ansätzen, sind jedoch in dem Umfang deutlich kleiner gehalten und bieten weniger Funktionen. Im Folgenden werden einige wichtige Bibliotheken genannt und kurz erläutert:

#### **xml.etree.ElementTree**

Die Standardbibliothek für XML in Python. Sie sind einfach zu benutzen und es sind keine Zusatzinstallationen. Gut geeignet, um kleine bis mittlere XML-Dateien zu verarbeiten. Unterstützt grundlegendes Parsen, Schreiben und einfaches XPath, aber bietet keine XSLT oder komplexe Features für XML-Dokumenten. [6]

#### **xml.dom.minidom**

Eine DOM-basierte Implementierung in Python. Ermöglicht den Zugriff auf XML über das klassische Document Object Model. Bietet feine Kontrolle, wirkt aber oft sperrig und weniger performant. [7]

#### **xml.sax**

Ein eventbasiertes XML-Parsing. Liest XML zeilenweise und löst Ereignisse aus, wenn Elemente gefunden werden. Sehr speichersparend und ideal für sehr große XML-Dateien, aber auch komplizierter in der Anwendung. [8]

#### **lxml**

Die leistungsfähigste XML/HTML-Bibliothek in Python. Basiert auf den C-Bibliotheken libxml2 und libxslt und ist daher sehr schnell. Unterstützt XPath, XSLT, Validierung und kann sowohl sauberes XML als auch „kaputtes“ HTML verarbeiten. Der wird oft als heutiger De-facto-Standard für komplexe XML-Verarbeitung in Python bezeichnet. [9]

**BeautifulSoup**

Eigentlich für HTML gedacht, funktioniert aber auch mit XML. Sehr tolerant gegenüber fehlerhaftem Code, was es beim Web-Scraping nützlich macht. Langsamer als lxml, aber sehr einfach in der Benutzung. [10]

**xmlschema**

Spezialisiert auf XML Schema Definition (XSD). XSD ist eine Sprache, die die Struktur und den Inhalt von XML-Dokumenten definiert. XML-Schema kann XML-Dokumente validieren und direkt in Python-Objekte umwandeln. Gut geeignet, wenn du sicherstellen musst, dass XML-Daten einer bestimmten Struktur entsprechen. [11]

**defusedxml**

Eine sichere Variante der Standardbibliotheken. Schützt vor bekannten XML-Sicherheitslücken wie „Billion Laughs“ oder XML External Entity (XXE). Billion Laughs beschreibt das Einsetzen von stark verschachtelten und sich wiederholenden XML-Strukturen, die, wenn Sie in den Parser geladen werden, wodurch es zu Speicher- oder CPU-Überlastung führen kann. Mit XXE können Befehle in der XML-Struktur genutzt werden, welche zu Sicherheitslücken beim XML-Parsing führen. So können Angreifer externe Entitäten einbinden, um Daten auszulesen oder sogar Befehle auszuführen. Wichtig, wenn XML aus unsicheren Quellen kommt. [12]

**untangle**

Sehr einfache Bibliothek, die XML in Python-Objekte übersetzt. Damit können XML-Strukturen fast wie normale Attribute angesprochen werden. Gut für kleine Projekte, aber eingeschränkt im Funktionsumfang. [13]

## 2.3. Datenbankentwurf und Normalisierung

Datenbanken sind strukturierte Zusammenstellungen von Daten, die elektronisch gespeichert und verwaltet werden. Ihr Hauptziel ist es, große Datenmengen strukturiert zu speichern, den Zugriff zu optimieren und die Integrität der Daten zu gewährleisten. Datenbanken haben im Vergleich zu anderen Dateisystemen Mechanismen, die mehreren Benutzern die parallele Nutzung ermöglichen, sowie redundante Datenhaltung vermeiden und die effiziente Abfragen über spezielle Sprachen wie die Structured Query Language (SQL) ermöglichen. [14, S. 6]

In der folgen werden die Hauptanforderungen an Datenbanken bzw. Datenbankmanagementsysteme kurz zusammengefasst: [15, S. 7],[14, S. 6]

- Datenunabhängigkeit: Speicherung unabhängig von Programmen und Plattformen
- Benutzerfreundlichkeit: Einfache Sprachen und grafische Oberflächen
- Mehrfachzugriff: Gleichzeitiger Zugriff für autorisierte Benutzer
- Flexibilität: Wahlfreier Zugriff und fortlaufende Verarbeitung
- Effizienz: Kurze Zeiten für Abfragen, Änderungen und Ergänzungen
- Datenschutz: Zugriff nach Benutzergruppen beschränkt

- Datensicherheit: Schutz vor Fehlern und Ausfällen
- Datenintegrität: Vollständige, korrekte und widerspruchsfreie Speicherung
- Redundanzfreiheit: Daten nur einmal speichern, Redundanz vermeiden

### 2.3.1. Datenbankstruktur

Die Struktur der Datenbank legt fest, wie das Datenbanksystem organisiert ist und wie die Datenelemente angeordnet sind. Das Grundprinzip relationaler Datenbanken sind im Grunde Tabellen und die Beziehungen zwischen diesen. Die Hauptbestandteile sind: [16, S. 35]

- Tabellen – das grundlegende Element, welches Daten in Zeilen (Tupeln) und Spalten (Attributen) strukturiert.
- Ein Datensatz in einer Tabelle wird durch Primärschlüssel (Primary Keys) eindeutig identifiziert.
- Schlüssel aus anderen Tabellen (Foreign Keys): Tabellenverknüpfung, um die referenzielle Integrität zu gewährleisten.
- Indizes: Datenstrukturen, die das Beschleunigen von Abfragen ermöglichen.
- Views: Sie sind virtuelle Tabellen und beruhen auf den Ergebnissen von Abfragen.
- Constraints: Vorgaben zur Gewährleistung der Datenintegrität (z. B. Wertebereiche, Pflichtfelder oder Dopplungen).

### 2.3.2. Datenintegrität und Normalisierung

Ein methodischer Ansatz zur Reduzierung von Redundanzen und Anomalien ist die Normalisierung. Der Prozess erfolgt schrittweise durch die Normalformen, die durch Indizes (1NF, 2NF, 3NF, BCNF) definiert sind. Alle Normalformen haben spezifische Arten von Datenanomalien zum Ziel: [16, S. 38–41]

- 1. Normalform (1NF): Beseitigung mehrfacher Werte innerhalb einer Zelle.
- 2. Normalform (2NF): Beseitigung partieller Abhängigkeiten.
- 3. Normalform (3NF): Beseitigung transitiver Abhängigkeiten.

Die Gewährleistung, dass Daten korrekt, konsistent und vollständig sind, fällt unter das Konzept der Datenintegrität. Sie wird erzielt durch: [16, S. 38–41]

- Entity-Integrität (eindeutige Primärschlüssel).
- Referentielle Integrität (gültige Fremdschlüsselverweise).
- Domänenintegrität (gültige Wertebereiche und Datentypen).

- Entity-Integrität (eindeutige Primärschlüssel).
- Referentielle Integrität (gültige Fremdschlüsselverweise).
- Integrität der Domäne (gültige Wertebereiche und Datentypen).

### 2.3.3. Vorgehen zur Erstellung der Datenbank

Bei der Erstellung einer Datenbank sollte nach folgendem Schema vorgegangen werden. Bei diesem Schema gilt es, sowohl technische als auch konzeptionelle Aspekte zu berücksichtigen: [15, S. 9–11]

#### 1. Anforderungsanalyse

Zuerst wird im Hinblick auf das Geschäftsumfeld bestimmt, welchen konkreten Zweck die Datenbank erfüllen soll. Hierbei ist die Zweckbestimmung entscheidend, da sie festlegt, welche Daten als relevant gelten. Der Konzeptvorschlag, der das Projektziel definiert und die Vorgehensweise umreißt, ist das Ergebnis.

- Festlegung der fachlichen und technischen Voraussetzungen.
- Bestimmung der relevanten Datenquellen und -formate.
- Definition von Integritäts- und Sicherheitsanforderungen.

#### 2. Konzeptionelles Datenmodell

Es soll das geschäftliche Umfeld betrachtet werden. Die bestehenden Objekte (z. B. Reports mit Materialnummer, Datum, Version), deren Attribute sowie die Beziehungen und Einschränkungen zwischen diesen Objekten. Das Entity-Relationship-Modell (ERM) nach Chen oder das PrecisedERM (PERM) werden häufig zur Modellierung verwendet. Der konzeptionelle Entwurf ist die Grundlage für die nächste Phase und dient als Diskussionsbasis.

- Die Entwicklung eines Entity-Relationship-Modells (ERM), das die reale Welt in Entitäten, Attribute und Beziehungen abbildet.
- Die Einbeziehung von Kardinalitäten (1:1, 1:n, n:m).

#### 3. Logisches Datenmodell

Der konzeptionelle Entwurf wird in einen logischen Entwurf umgewandelt, der die fachlichen Konzepte in ein datenbanktechnisches Format überführt. In der Regel kommt das Relationenmodell zum Einsatz, welches die Daten in Form von Tabellen organisiert. Transformationsregeln garantieren, dass Beziehungen und Integritätsbedingungen richtig umgesetzt werden.

- Transformation des konzeptionellen Modells in ein relationales Schema.

- Festlegung von Tabellen, Spalten (Attributen), Primärschlüsseln und Fremdschlüsseln.
- Definition von Datentypen und Zellkonfigurationen (z. B. NOT NULL, UNIQUE, CHECK).

#### 4. Physisches Datenmodell

Eine physische Datenbankstruktur wird durch SQL erstellt, basierend auf dem logischen Entwurf. Die konkrete Festlegung von Tabellen, Indizes, Constraints usw. erfolgt dabei durch das implementierte System, welches immer wieder getestet und zusammen mit den Nutzerinnen auf fachliche Richtigkeit überprüft wird.

- Realisierung des logischen Modells in einer spezifischen Datenbankmanagementsoftware (z. B. MySQL, MariaDB oder Microsoft SQL Server).
- Die Verbesserung der Speicherstrukturen, der Indexierung und der Partitionierung.

#### 5. Implementierung und Prüfung

Nach der erfolgreichen Umsetzung wird das System vom Kunden gemäß einem vorher festgelegten Abnahmeplan freigegeben. Ein Wartungsplan kümmert sich anschließend um die Betreuung, schult die Endbenutzer und überwacht die IT-Umgebung kontinuierlich.

- Der Aufbau der Tabellen und Relationen erfolgt entsprechend dem Datenbankschema.
- Die Testdaten implementieren.
- Die Kontrolle der Funktionalität und Leistung.

### 2.3.4. Methoden zum Datenbankzugriff in Flask

In diesem Abschnitt werden die maßgeblichen Methoden für die Arbeit mit SQL-Datenbanken in Flask vorgestellt, ihre Charakteristika erläutert und ein Vergleich gezogen. Hier werden zwei der meistgenutzten Ansätze vorgestellt und Vor-/Nachteilen aufgezeigt.

#### Direkter SQL-Zugriff via DB-Treiber

Bei dieser Methode wird direkt mit einem Datenbanktreiber (z. B. psycopg2, mariaDB oder mysql-connector) gearbeitet. SQL-Statements werden als Strings formuliert und über einen Cursor ausgeführt. Ein Cursor ist ein Steuerobjekt, das SQL-Befehle ausführt und Ergebnisse verwaltet. Diese Methode eignet sich vorwiegend für kleine Anwendungen oder für hochoptimierte Spezialfälle, bei denen abstrakte Ein- und Auslesen-Befehle vorgenommen werden. Ein typischer Ablauf für diese Methode läuft wie folgt ab: [17][18][19]

1. Der Verbindungsaufbau zu einer Datenbank.
2. Ausführung eines SQL-Statements durch einen Cursor (z. B. SELECT, INSERT, UPDATE, DELETE).

3. Abrufen der Resultate und Überführen in eine geeignete Datenstruktur (z. B. eine Liste oder Dictionary).
4. Schließen der Verbindung nach fertiger Ausführung.

In der folgenden Tabelle 2 werden einige Vor- und Nachteile der Methode aufgeführt.

Vorteile	Nachteile
Maximale Kontrolle über das ausgeführte SQL	Gefahr für SQL-Injections, wenn Parameter nicht sicher gebunden werden
Kein Overhead durch zusätzliche Abstraktionsschichten	Manuelle Fehleranfälligkeit (z. B. bei Verbindungshandling oder Transaktionen)
Nützlich bei komplexen, datenbankspezifischen Abfragen, die ein ORM möglicherweise schwer modellieren kann	Schwer zu skalieren bei komplexer Domänenlogik
—	Geringe Lesbarkeit, insbesondere bei vielen Joins oder Abhängigkeiten

Tabelle 2: Vor- und Nachteile des direkten SQL-Zugriffs in Flask

Quelle: eigene Darstellung

## Object-Relational Mapping

Object-Relational Mapping (ORM) eine Methode in der Softwareentwicklung, die es ermöglicht, Objekte aus einer objektorientierten Programmiersprache auf Tabellen in einer relationalen Datenbank abzubilden, um die Interaktion mit Datenbanken zu erleichtern. Entwickler müssen keine SQL-Abfragen mehr manuell erstellen. Sie interagieren einfach mit Objekten in ihrer Programmiersprache, und das ORM-Tool wandelt dies in die passenden Datenbankoperationen um. [20] Eine der am meisten verwendete ist Python-Bibliothek für diese Methode ist SQLAlchemy.

SQLAlchemy ist ein umfangreiches Toolkit und ORM für relationale Datenbanken. In Flask nutzt man normalerweise die Erweiterung Flask-SQLAlchemy, die SQLAlchemy in die Kontextstruktur von Flask integriert und den typischen Boilerplate-Aufwand reduziert.[21]

Mit Flask-SQLAlchemy definiert man Modelle als Python-Klassen, die automatisch auf Datenbanktabellen abgebildet werden (Declarative Mapping). Die Objekte dieser Klassen repräsentieren Datensätze, und Operationen daran (z. B. `session.add()`, `session.commit()`) generieren hinter den Kulissen passende SQL-Befehle.[21]

In der folgenden Tabelle 3 werden einige Vor- und Nachteile der Methode bzw. SQLAlchemy aufgeführt. Neben SQLAlchemy gibt es noch einige andere Python-Bibliotheken die ORM- nutzen z. B. .

Vorteile	Nachteile
ORM-Abstraktion: Arbeiten mit Objekten statt SQL	Leicht höhere Komplexität als direkter SQL-Zugriff
Automatisches Erstellen und Aktualisieren von Tabellen	Performance bei sehr großen Datenmengen leicht geringer
Integriert gut in Flask	—
Kompatibel mit vielen Datenbanken (SQLite, MySQL, PostgreSQL, etc.)	—

Tabelle 3: Vor- und Nachteile der Verwendung von Flask-SQLAlchemy

Quelle: eigene Darstellung

## 2.4. Grundlagen der Datenvisualisierung

### 2.4.1. Graphen erstellung mit JS

**Chart.js** Eine einfache und beliebte Bibliothek für Standarddiagramme wie Balken, Linien oder Kreisdiagramme. Schnell eingebunden, gute Standardoptik, ideal für kleinere Projekte oder schnelle Visualisierungen.

**D3.js** Eine sehr mächtige Low-Level-Bibliothek, die mit SVG, Canvas und HTML arbeitet. Extrem flexibel für individuelle und interaktive Visualisierungen, aber mit einer steilen Lernkurve.

**Plotly.js** Bietet viele interaktive Diagramme (Zoom, Hover, Export als PNG). Unterstützt auch 3D- und wissenschaftliche Charts. Gut geeignet für Dashboards und Datenanalyse.

**ECharts** Eine leistungsstarke Open-Source-Bibliothek von Apache. Unterstützt viele Diagrammtypen (inkl. Heatmaps, Maps, Candlesticks) mit Animationen und Interaktivität. Besonders beliebt für komplexe Enterprise-Web-Apps.

**Highcharts** Eine professionelle, kommerzielle Lösung (kostenlos für private Nutzung). Sehr einfach konfigurierbar, liefert Business-taugliche Diagramme mit vielen Optionen. Oft in Unternehmen eingesetzt.

**vis.js** Spezialisiert auf Netzwerk- und Zeitachsenvisualisierungen. Eignet sich besonders für Knoten-Graphen, Beziehungen oder Prozessdiagramme mit Interaktivität.

**Recharts** Eine React-spezifische Bibliothek, die auf D3.js basiert. Bietet eine einfache Komponenten-API für gängige Diagramme. Ideal, wenn deine Web-App mit React entwickelt ist.

## 2.5. Anforderungen an modulare Softwareentwicklung

Die modulare Softwareentwicklung ist ein etabliertes Denkmuster, das darauf abzielt, komplexe Softwaresysteme in klar abgegrenzte, wiederverwendbare und unabhängig voneinander entwickelbare Einheiten zu zerlegen. Ziel ist es, sowohl die Wartbarkeit, Erweiterbarkeit als auch die



Qualität der Software zu erhöhen. Um diese Ziele zu erreichen, müssen spezifische Anforderungen an die Gestaltung und Umsetzung modularer Systeme beachtet werden.

### **2.5.1. Klare Modulabgrenzung und Verantwortlichkeiten**

Ein Modul sollte eine klar definierte Aufgabe erfüllen und über eine eindeutig abgegrenzte Funktionalität verfügen. Diese Trennung wird in der Literatur häufig als *Separation of Concerns* (SoC) bezeichnet. Durch eine eindeutige Abgrenzung lassen sich Abhängigkeiten reduzieren, wodurch Änderungen in einem Modul nur minimale Auswirkungen auf andere Systemkomponenten haben.

### **2.5.2. Geringe Kopplung und hohe Kohäsion**

Zwei zentrale Qualitätsmerkmale modularer Systeme sind niedrige Kopplung und hohe Kohäsion.

- Kohäsion beschreibt, wie eng die Elemente eines Moduls zusammenarbeiten, um eine spezifische Aufgabe zu erfüllen.
- Kopplung hingegen beschreibt den Grad der Abhängigkeit zwischen einzelnen Modulen. Ein hoher Kohäsionsgrad bei gleichzeitig niedriger Kopplung erleichtert sowohl die Wiederverwendung als auch die Wartung.

### **2.5.3. Einheitliche Schnittstellen**

Module interagieren über klar definierte und stabile Schnittstellen. Diese sollten standardisiert, dokumentiert und möglichst unabhängig von internen Implementierungsdetails sein. Eine wohldefinierte API (Application Programming Interface) ermöglicht die parallele Entwicklung mehrerer Module und erleichtert zukünftige Erweiterungen.

### **2.5.4. Wiederverwendbarkeit**

Ein wesentliches Ziel der Modularisierung ist die Wiederverwendbarkeit von Modulen in unterschiedlichen Projekten oder Kontexten. Wiederverwendbare Module reduzieren Entwicklungsaufwand, erhöhen die Konsistenz und tragen zur Qualitätssteigerung bei. Hierfür ist eine generische und konfigurierbare Implementierung erforderlich.

### 2.5.5. Erweiterbarkeit und Anpassungsfähigkeit

Modulare Software muss so konzipiert sein, dass neue Funktionen ohne weitreichende Änderungen am bestehenden System integriert werden können. Dieses Prinzip wird oft mit dem *Open/Closed Principle* aus den SOLID-Prinzipien beschrieben: „Software entities should be open for extension, but closed for modification“.

### 2.5.6. Testbarkeit

Durch die Abgrenzung von Modulen wird eine gezielte und unabhängige Prüfung einzelner Systemkomponenten möglich. Unit-Tests, Integrationstests und Mocking-Strategien profitieren stark von einer modularen Struktur. Testbare Module führen zu einer höheren Softwarequalität und geringeren Fehlerquoten im Betrieb.

### 2.5.7. Konsistenz und Standardisierung

Die Implementierung sollte einheitliche Konventionen in Bezug auf Namensgebung, Dokumentationsstandards und Code-Struktur aufweisen. Konsistenz erleichtert die Zusammenarbeit in Entwicklerteams und minimiert Missverständnisse.

### 3. Analyse und Konzeption

In diesem Kapitel werden die grundlegenden Anforderungen und konzeptionellen Überlegungen für die Entwicklung der Webapplikation beschrieben. Ziel ist es, eine fundierte Basis für die spätere Implementierung zu schaffen, indem sowohl die funktionalen als auch die nicht-funktionalen Anforderungen analysiert und geeignete technische Konzepte erarbeitet werden. Dazu werden zunächst die Ausgangssituation und die zu verarbeitenden Datenquellen untersucht. Die erarbeiteten Konzepte bilden die Grundlage für das Design und die Realisierung der Software im weiteren Verlauf der Arbeit.

#### 3.1. Definition der Anforderungen

In diesem Abschnitt werden die funktionalen und nicht-funktionalen Anforderungen der zu entwickelnden Web-Applikation beschrieben. Sie wurden in Abstimmung mit den späteren Nutzerinnen und Nutzern sowie den betreuenden Personen im Unternehmen festgelegt. Die Anforderungen bilden die Grundlage für den Entwurf und die Umsetzung der Software und dienen dazu, deren Zielsetzung, Funktionsumfang und technische Rahmenbedingungen klar zu definieren.

##### 3.1.1. Funktionale Anforderungen

Die funktionalen Anforderungen beschreiben die vorgesehenen Funktionen und Abläufe der Applikation. Sie definieren, was die Anwendung leisten soll und wie die Nutzerinnen und Nutzer mit ihr interagieren können.

###### 1. Navigation und Benutzerführung

Die Applikation soll eine einfache und übersichtliche Navigation ermöglichen. Der Wechsel zwischen den Hauptfunktionen erfolgt über eine feststehende Menüleiste im oberen Bereich der Benutzeroberfläche. Dadurch können die Bereiche *Upload*, *Berichte* und *Diagramme* direkt aufgerufen werden.

###### 2. Einlesen von XML-Dateien

Das Einlesen der automatisch generierten XML-Berichte erfolgt über eine Upload-Seite. Die Anwendung soll sowohl ältere als auch neuere XML-Strukturen verarbeiten können, da der Teststand unterschiedliche Versionen erzeugt.

###### 3. Datenverarbeitung und Speicherung

Nach dem Upload sollen die XML-Dateien automatisch geparkt, validiert und in die Datenbank überführt werden. Doppelte Einträge sollen erkannt und vermieden werden.

###### 4. Anzeige der Berichte

Die gespeicherten Berichte sollen tabellarisch dargestellt werden. Jede Zeile repräsentiert einen Testbericht und zeigt die wichtigsten Informationen wie Datum, Seriennummer, Materialnummer und Testergebnis an.

**5. Filter- und Suchfunktionen**

Die Berichtstabelle soll über Filteroptionen verfügen, um gezielt nach bestimmten Kriterien (z. B. Datum, Seriennummer, Testmodul oder Ergebnisstatus) zu suchen.

**6. Grafische Darstellung der Testdaten**

Die Anwendung soll die Messergebnisse der einzelnen Module grafisch darstellen. Die Diagramme sind nach Modulen sortiert und enthalten eine klare Achsenbeschriftung sowie Legenden. Das Layout der Diagramme orientiert sich an den Graphen des Teststandprogramms, um den Nutzerinnen und Nutzern die Interpretation zu erleichtern.

**7. Systemmeldungen und Fehlermanagement**

Nach erfolgreichen oder fehlerhaften Aktionen (z. B. Upload, Datenbankeintrag, Filterabfrage) sollen entsprechende Systemmeldungen angezeigt werden, um den Benutzerstatus transparent zu machen.

**8. Exportfunktionen (NEU)**

Die Anwendung soll die Möglichkeit bieten, ausgewählte Datensätze oder Berichte in ein externes Format zu exportieren (z. B. CSV oder Excel). Dadurch können die Daten auch außerhalb der Anwendung weiterverarbeitet werden.

**9. Modulare Erweiterbarkeit (NEU)**

Neue Funktionen, etwa zusätzliche Testmodule oder Analysefunktionen, sollen über separate Module oder Blueprints ergänzt werden können, ohne bestehende Teile der Anwendung zu verändern.

**10. Benutzerverwaltung (optional) (NEU)**

Die Architektur der Applikation soll es ermöglichen, zukünftig eine einfache Benutzerverwaltung zu integrieren. Damit können verschiedene Rollen (z. B. Administrator, Nutzer, Gast) eingeführt werden.

**3.1.2. Nicht-funktionale Anforderungen**

Die nicht-funktionalen Anforderungen beschreiben die Qualitätsmerkmale der Software. Sie legen fest, unter welchen Bedingungen die Applikation betrieben werden soll und welche Eigenschaften sie erfüllen muss.

**1. Plattform und Laufzeitumgebung**

Die Web-Applikation wird auf einem unternehmenseigenen Apache2-Server betrieben. Sie basiert auf Python und dem Webframework Flask.

**2. Performance**

Das System muss auch bei größeren XML-Dateien zuverlässig arbeiten. Eine aktive Laufzeitoptimierung ist nicht erforderlich, solange der Arbeitsfluss nicht beeinträchtigt wird.

**3. Usability**

Die Benutzeroberfläche soll intuitiv bedienbar sein und eine klare Struktur aufweisen. Eingabefehler sollen vermieden und, wenn möglich, automatisch erkannt werden.

#### 4. Modularität und Wartbarkeit

Die Applikation ist modular aufgebaut, um eine einfache Pflege und spätere Erweiterung zu ermöglichen. Änderungen an einzelnen Komponenten sollen keine Anpassungen an anderen Modulen erfordern.

#### 5. Sicherheit und Datenintegrität (NEU)

XML-Dateien müssen vor dem Einlesen auf korrekte Struktur und mögliche Sicherheitsrisiken geprüft werden (z. B. Schutz vor fehlerhaften oder manipulierten Dateien).

#### 6. Fehler- und Ausnahmebehandlung (NEU)

Unerwartete Systemfehler sollen abgefangen und im Log gespeichert werden. Für Benutzerinnen und Benutzer werden stattdessen verständliche Fehlermeldungen ausgegeben.

#### 7. Skalierbarkeit (NEU)

Das System soll bei Bedarf mit minimalem Aufwand um zusätzliche Funktionen, Testmodule oder Datenquellen erweitert werden können.

#### 8. Kompatibilität (NEU)

Die Applikation soll auf allen gängigen Browsern lauffähig sein (z. B. Chrome, Edge, Firefox).

#### 9. Dokumentation und Nachvollziehbarkeit (NEU)

Der Code soll übersichtlich dokumentiert werden. Wichtige Abläufe (Upload, Datenverarbeitung, Visualisierung) werden nachvollziehbar in der Projektdokumentation beschrieben.

### 3.1.3. Anforderungen an die Datenbank

Die Datenbank ist ein zentraler Bestandteil der Anwendung. Sie speichert alle relevanten Testdaten, Parameter und Informationen aus den XML-Berichten.

1. Die Datenbank basiert auf dem relationalen Datenbanksystem **MariaDB**.
2. Sie muss in der Lage sein, sämtliche in den XML-Berichten enthaltenen Daten vollständig abzubilden, um die Wiederherstellung eines gesamten Testberichts theoretisch zu ermöglichen.
3. Datendopplungen sollen vermieden werden. Häufig vorkommende Informationen (z. B. Teststandname, Versionsnummern, DUT-Typen) werden ausgelagert und über Fremdschlüssel referenziert.
4. Für DUT-Typen und Seriennummern werden separate Informationstabellen angelegt, um Redundanzen zu vermeiden und die Datenstruktur klar zu halten.
5. Die Datenbankstruktur soll so gestaltet sein, dass zukünftige Teststände oder zusätzliche Testmodule problemlos integriert werden können.

#### 6. Verwendung von ORM-Technologie (NEU)

Für die Kommunikation zwischen Anwendung und Datenbank wird **SQLAlchemy** eingesetzt. Dadurch kann der Datenzugriff objektorientiert erfolgen, was die Wartbarkeit und Lesbarkeit des Codes verbessert.

#### 7. Versionierung und Migration (NEU)

Änderungen an der Datenbankstruktur werden mit **Alembic** verwaltet, um eine konsistente Weiterentwicklung der Datenbank sicherzustellen.

#### 8. Datenvalidierung und Konsistenzprüfung (NEU)

Beim Einfügen neuer Datensätze sollen Validierungen sicherstellen, dass die Werte logisch und formal korrekt sind (z. B. keine Nullwerte bei Pflichtfeldern, korrekte Datentypen).

Die beschriebenen Anforderungen bilden den Rahmen für die Entwicklung der Web-Applikation. Während die funktionalen Anforderungen die konkreten Aufgaben und Abläufe definieren, legen die nicht-funktionalen Anforderungen fest, wie die Anwendung qualitativ umgesetzt werden soll. Zusammen mit den Datenbankanforderungen bilden sie die Grundlage für den folgenden Entwurf der Systemarchitektur und die spätere Implementierung.

### 3.2. Grundlegender Ablauf des Programmes

Abbildung 6 zeigt den geplanten Ablauf der Datenverarbeitung innerhalb der entwickelten Web-Applikation. Der Prozess beginnt mit dem Einlesen der XML-Dateien, die die Prüf- und Messdaten enthalten. Anschließend erfolgt eine Validierung der Datenstruktur und des Formats, um sicherzustellen, dass die XML-Dateien den definierten Spezifikationen entsprechen. Bei fehlerhaften oder unvollständigen Dateien wird der Prozess abgebrochen. Sind die Daten gültig, werden sie mit den vorhandenen Datenbankeinträgen abgeglichen. Abhängig vom Ergebnis werden entweder fehlende Datensätze ergänzt oder neue Einträge vollständig eingefügt. Dadurch wird sichergestellt, dass die Datenbank stets konsistente und aktuelle Informationen enthält.

Nach dem erfolgreichen Datenimport erfolgt die Auswahl eines Datensatzes, dessen Inhalte grafisch dargestellt werden. Optional kann der erzeugte Graph als PNG-Datei gespeichert werden.

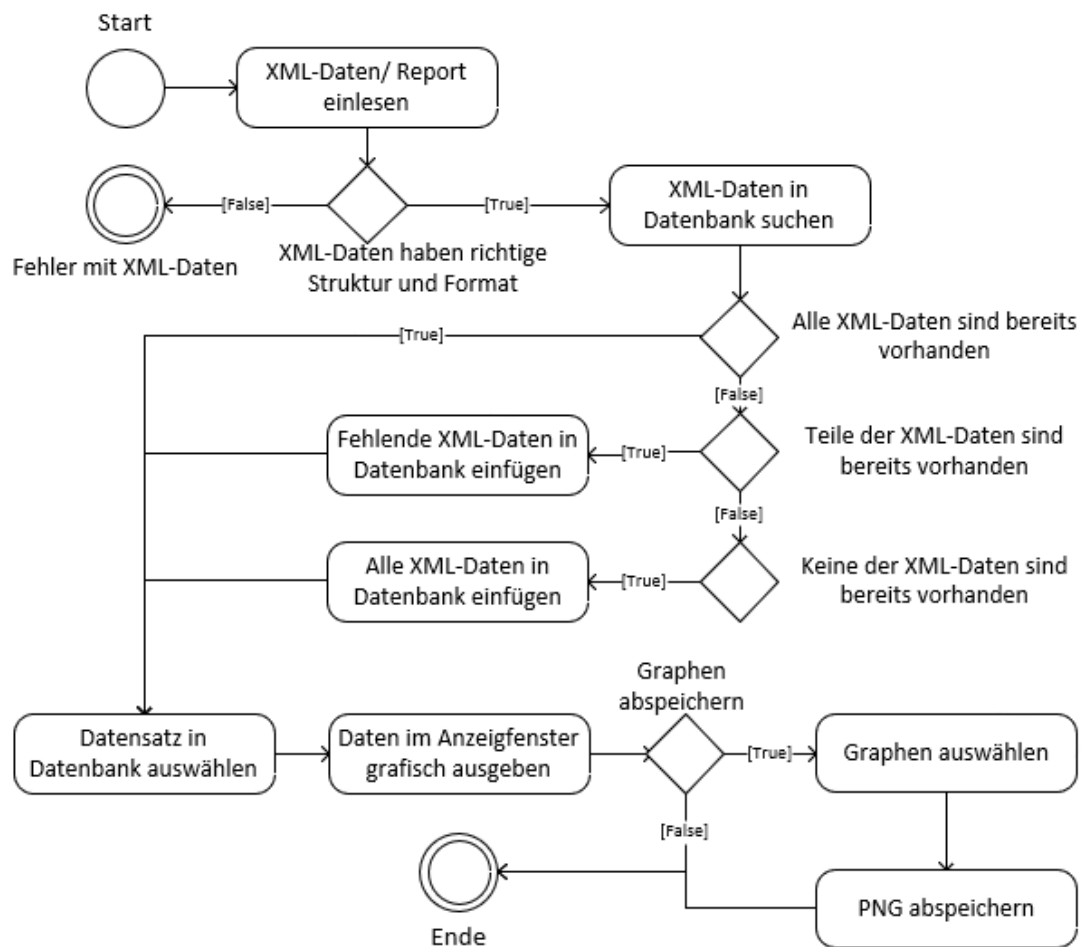


Abbildung 6: Grundlegendes Ablaufdiagramm der Nutzung der Web-Applikation

Quelle: Eigene Darstellung mit Microsoft Visio

### 3.3. Analyse der generierten XML-Berichte und bestehenden Strukturen

In diesem Unterkapitel wird der Aufbau der automatisch generierten Testberichte aus dem Teststand und der vorhandenen Ausgabe- und Speicherstruktur erläutert. Dies ist relevant für die Erstellung der Datenbank und das Verständnis der Ein- und Auslesefunktionen der Web-Applikation.

#### 3.3.1. XML-Berichtsstrukturschema

Die vom Teststand automatisch generierten Berichte folgen einem konstanten Strukturschema, welches sich bis zu einer gewissen Ebene der XM-Struktur in jedem Bericht wiederholt. Wie im Kapitel „Grundlagen“ beschrieben, kann ein Element Attribute, andere Elemente und einen Inhaltswert beinhalten.

Das Stammelement heißt in allen automatisch generierten Berichten „test“ und besitzt immer das Attribut „id“. Die Zahl in diesem Attribut beschreibt den Typ des DUTs die in diesem Test geprüft wurden.

Das Stammelement hat die Unter- bzw. Kinderelemente „info“, „testbench“, „string“ und dreimal „testmodule“. Diese Elemente haben wiederum alle weitere Unterelemente und Attribute. In der folgenden Abbildung 7 ist die unveränderliche Struktur eines Testberichtes abgebildet, welcher erfolgreich durchgeführt wurde.

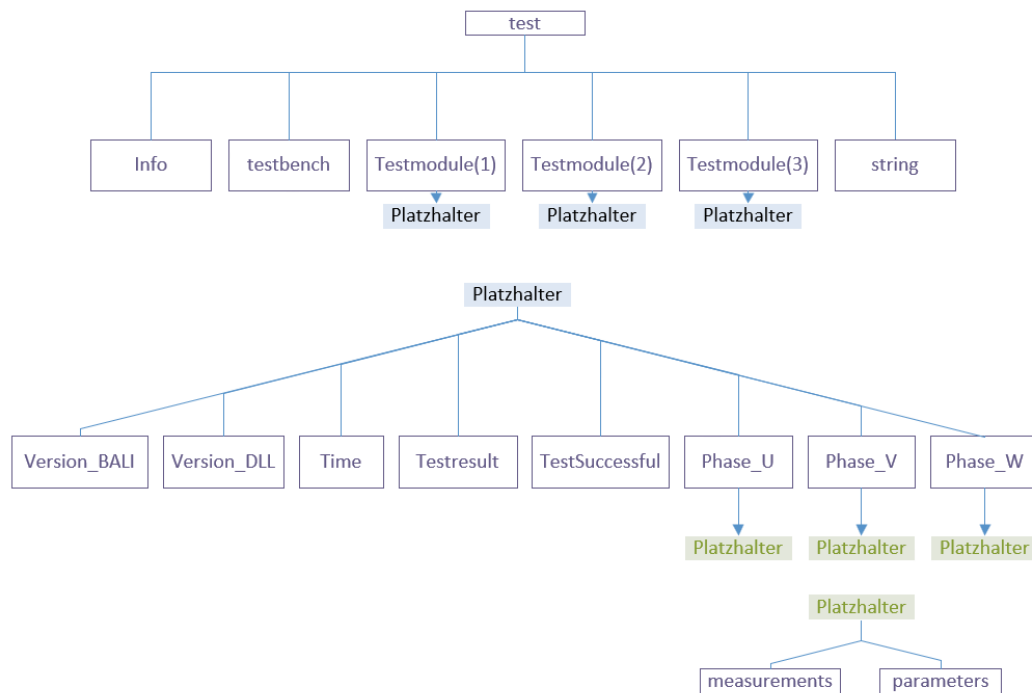


Abbildung 7: Aufbau unveränderlicher XML-Berichtstruktur

Quelle: Eigene Darstellung mit Microsoft Visio

Hierbei sollte angemerkt werden, dass in allen Testberichten die Elemente, die kein Unterelement besitzen gleich aufgebaut sind. Sie besitzen mindestens das Attribut „name“ und ihre Elementenbezeichnung ist immer nach dem Datentyp ihres Inhalts benannt. Für ein besseres Verständnis siehe Abbildung 8.

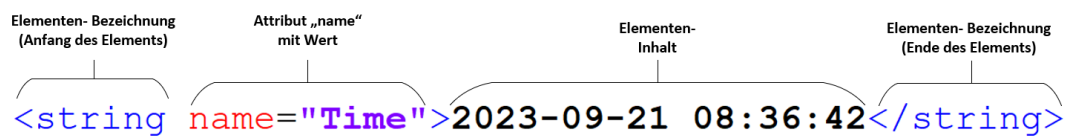


Abbildung 8: Elementaufbau aus XML-Bericht

Quelle: Eigene Darstellung mit Microsoft Visio

Jedes der Elemente „testmodule“ hat ein Attribut „name“, durch welches Sie unterschieden werden können. Das Element „string“ hat ein Namensattribut mit der Bezeichnung „test sequence status“. Der Inhalt dieses Elementes gibt an, ob der Test erfolgreich abgeschlossen wurde oder ob es einen Fehler bzw. eine Grenzüberschreitung von Messwerten gab. Die hier genannten eigenschaften sind in der Abbildung 9 zu erkennen.



```
1 <?xml version="1.0" encoding="UTF-8"?>
2   <test id="124">
3     <info></info>
4     <testbench></testbench>
5     <testmodule name="DriverConsumptionTest"></testmodule>
6     <testmodule name="PulseTest_FSW_L"></testmodule>
7     <testmodule name="XPowerTest"></testmodule>
8     <string name="test_sequence_status">
9       Test sequence successfully finished.</string>
10  </test>
```

Abbildung 9: Direkte Kinderelement unter Stammelement

Quelle: Eigene Darstellung

Das Element „info“ enthält in seinen Unterelementen die allgemeinen Testinformationen, wie die Startzeit des Tests, die Typen-ID, die Konfigurationsbezeichnung des Teststandes, die Bezeichnung des angeschlossenen Carriers und die Seriennummern des DUTs. Diese Elemente haben keine Unterelemente und sind alle mit einem Namensattribut und einem Inhalt definiert. Für die genauen Bezeichnungen und Struktur siehe Abbildung 10

```
1 <info>
2   <string name="Time">2023-09-21_08:36:33</string>
3   <string name="Material_number">124</string>
4   <string name="Configuration">L_3DUT_3P1Q</string>
5   <string name="ID_Carrier_left">Carrier 1</string>
6   <string name="ID_DUT_R">10-29550</string>
7   <string name="ID_DUT_S">10-29396</string>
8   <string name="ID_DUT_T">10-29482</string>
9 </info>
```

Abbildung 10: XML-Strukturbeispiel Info-Element

Quelle: Eigene Darstellung

Die Teststandbezeichnung und die Versionen der Hard- und Software werden im Element „testbench“ gespeichert. Diese Elemente enthalten auch keine weiteren Unterelemente und sind alle mit einem Namensattribut und einem Inhalt versehen. Die genauen Bezeichnungen und die Struktur sind in Abbildung 11 zu sehen.

```
1 <testbench>
2     <string name="Testbench_name">USTB_DtWind</string>
3     <string name="Testbench_HW_revision">V1.5</string>
4     <string name="Version_GUI">V1.4.1</string>
5     <string name="Version_controller">V1.3.2</string>
6 </testbench>
```

Abbildung 11: XML-Strukturbeispiel Testbench-Element

Quelle: Eigene Darstellung

Die Testmodule-Elemente sind alle nach dem gleichen Schema aufgebaut. Die ersten Unterelemente enthalten Grundinformationen zu dem Status des Testmoduls, der Tages- und Uhrzeit des Tests und welche Versionen davon verwendet wurden. Danach befinden sich noch drei weitere Elemente im Element „module“. Diese Elemente heißen „Phase U“, „Phase V“ und „Phase W“. Sie besitzen kein Namensattribut, siehe Abbildung 12.

```
1 <testmodule name="DriverConsumptionTest">
2     <integer name="TestSuccessful" min="1" max="1">1</integer>
3     <string name="Testresult">
4     Test finished successfully!</string>
5     <string name="Time">2023-09-21_08:36:42</string>
6     <string name="Version_DLL">V1.2.4</string>
7     <string name="Version_BALI">V1.2.4</string>
8     <Phase_U>...</Phase_U>
9     <Phase_V>...</Phase_V>
10    <Phase_W>...</Phase_W>
11 </testmodule>
```

Abbildung 12: XML-Strukturbeispiel Testmodulheader

Quelle: Eigene Darstellung

Diese Phasenelemente enthalten die Parameter und die Messdaten für die DUTs. Jede Phase enthält die Messwerte für ein DUT. Somit besitzen alle Module unter dem Element „Phase U“ die Messwerte und Parameter für den ersten DUT, dessen Seriennummer unter dem Element mit dem Namensattribut „ID DUT R“ in Abbildung 10 zu finden ist.

In dem Element „parameters“, welches in jedem Phasenelement vorkommt, befinden sich die voreingestellten Rahmbedingungen und Anforderungen für das Testmodul. Diese enthalten alle ein Namensattribut und einen Inhalt.

Die Elemente unter dem Eintrag „measurements“ enthalten die Messdaten der Testmodule. Sie haben oft noch extra Attribute wie „min“ oder „max“, welche die Toleranzgrenzen beschreiben. Diese Toleranzgrenzen zeigen, in welchem Bereich der jeweilige Elementinhalt angesiedelt sein muss, um kein positives bzw. fehlerfreies Ergebnis zu erzeugen. Die meisten Attributswerte finden sich schon in den Parametern wieder und sind somit doppelt im Bericht zu finden. Nur bei

den Floatblock-Elementen, also Elementen mit der Bezeichnung Floatblock, sind besondere Attribute zu finden, welche nicht gedoppelt vorkommen. Diese Elemente umfassen zusätzlich die Attribute „size“ und „duration“, welche die Anzahl der in dem Inhalt enthaltenen Floatwerte und die Dauer der Wertaufnahme angeben dies ist in Abbildung 13 zu erkennen. Diese sind für die grafische Darstellung besonders relevant.

```

1 <floatblock name="Time" size="698" duration="959.952" unit="s">
2 0.71365,2.1370,3.5323,...</floatblock>

```

Abbildung 13: Beispiel XPowerTest Floatblock-Elemente

Quelle: Eigene Darstellung

Auf der Phasenebene treten keine Veränderungen in der XML-Struktur auf, variieren lediglich die Elemente unter den Phasenelementen von Bericht zu Bericht etwas. Die beträchtliche Ausnahme von dieser Regel bilden die Berichte, bei denen ein Fehler aufgetreten ist oder sogar der Testvorgang abgebrochen wurde. Bei diesen Berichten hört die XML-Struktur bei den fehlgeschlagenen Testmodulen schon auf. Der Testbericht ist aber in sich geschlossen. Das bedeutet, die XML-Struktur ist vollkommen und kann von einem XML-Parser eingelesen werden. Es befinden sich nur weniger Testmodul-Elemente in dem Bericht und im Element für den Testsequenzstatus ist folgender Inhalt enthalten „Test sequence finished with errors“. Zudem wird bei dem Testmodul, bei dem der Fehler aufgetreten ist, unter dem Element mit dem Namen „Testresult“ ein Fehlercode und seine Bedeutung ausgegeben. In Abbildung 14 ist ein Beispiel für einen Bericht, welcher beim ersten Testmodul einen Fehler ausgegeben hatte. Bei einem Ausfall eines der zu testenden DUTs gibt es die Besonderheit, dass die anderen noch funktionsfähigen DUTs keine weiteren Datenaufnahmen mehr machen können. Denn das System benötigt alle drei Phasen, um den Test durchzuführen. Die XML-Struktur ist in sich geschlossen, aber es werden keine Daten von den anderen DUTs mehr aufgenommen.

```

1 <testmodule name="DriverConsumptionTest">
2   <integer name="TestSuccessful" min="1" max="1">0</integer>
3   <string name="Testresult">Idrv not within limits on Phase U
4 0x30010019 !
5 0x30010013</string>
6   <string name="Time">2021-03-10_09:10:37</string>
7   <string name="Version_DLL">V1.2.4</string>
8   <string name="Version_BALI">V1.2.4</string>
9   <Phase_U>...</Phase_U>
10  <Phase_V>...</Phase_V>
11  <Phase_W>...</Phase_W>
12 </testmodule>

```

Abbildung 14: Beispiel Fehler bei DriverConsumptionTest

Quelle: Eigene Darstellung

### 3.4. Datenbankdesign und Strukturkonzeption

In diesem Unterkapitel soll der Verlauf der Erstellung der Datenbank beschrieben und nachvollzogen werden. Es wurde in den Anforderungen festgelegt, dass MariaDB zu erstellen genutzt werden soll. MariaDB ist ein kostenloses, relationales Open-Source-Datenbankmanagementsystem, das aus einer Abspaltung von MySQL entwickelt wurde. MySQLs früherer Hauptentwickler Michael Widenius hat das Projekt ins Leben gerufen.

#### Betrachtung der Anforderungen

Für das genaue Interpretieren und Umsetzen der Anforderungen wurde festgelegt, dass die Ersteller/-in der Datenbank erst ein Konzept nach den festgelegten Anforderungen erarbeiten und umsetzen. Diese Umsetzung wird dann im Laufe der weiteren Erstellung der Applikation, falls nötig, angepasst.

#### Konzeptionelles Datenmodell

Für das Erstellen der Konzeption des Datenbankmodelles wurde die XML-Struktur analysiert, um die Kenntnisse und Muster aus Unterkapitel ?? abzuleiten. Aus diesen Erkenntnissen wurde dann ein ER-Modell erstellt, siehe Abbildung 15. Im Folgenden wird das ER-Modell beschrieben.

Durch das Durchführen eines Testes wird ein Bericht erstellt. Dieser Bericht enthält einige Attribute, welche Testinformationen, darunter die Seriennummern, das Datum und die Uhrzeit des Tests, die Materialnummer (DUT-ID), die Konfigurations-Version und die Carrier-Bezeichnung enthalten. Zudem besitzt der Bericht eine Unterentität Testbenchinformationen“, die die Hard- und Softwareversionen des Teststandes enthält. In der Unterentität Testmodelle befinden sich Versions-, Zeit- und Ergebnisinformationen sowie drei Datensätze, welche die Testparameter und die Messdaten enthalten. Diese Datensätze können den Phasen des Stromnetzes des Teststandes und somit den DUTs zugeordnet werden. Zusätzlich zu dem Aufbau wurden hier schon Schlüsselattribute hinzugefügt, die später die primären Schlüssel der Datenbanktabellen dienen sollen, orange markiert. Die Attribute sind hier nicht genau nach dem Inhalt der XML-Dateien benannt aber an ihre genauen Bezeichnungen angelehnt.

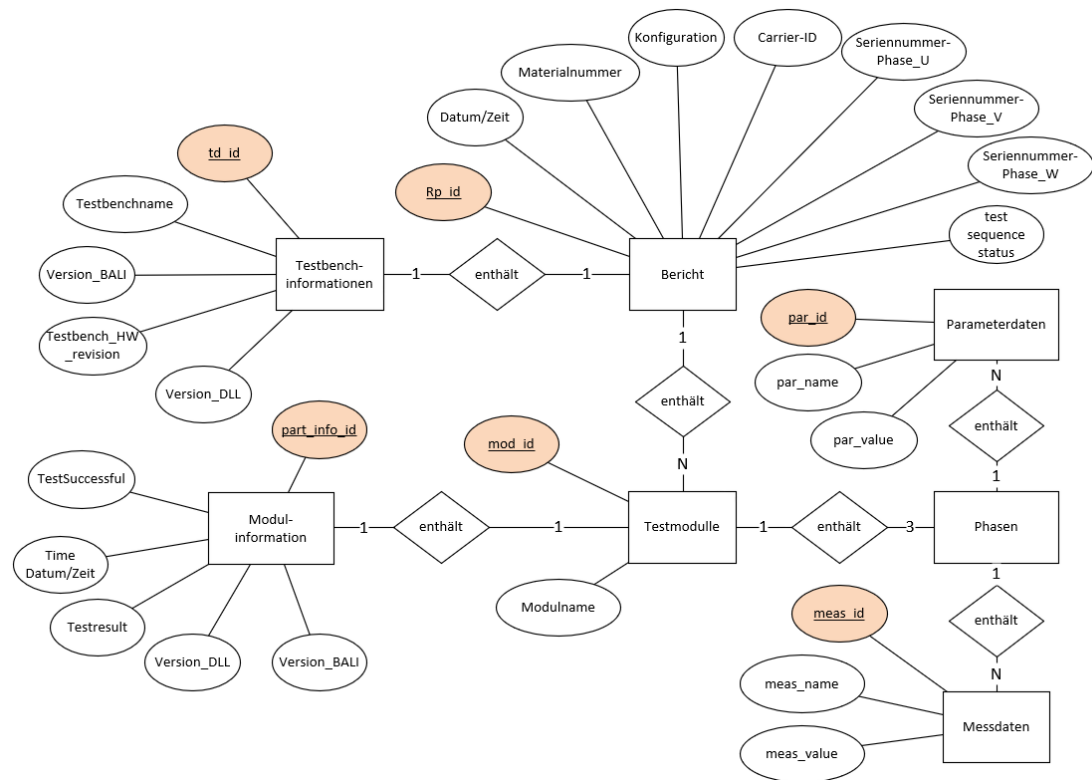


Abbildung 15: ER-Modell Überlegung der Datenbankstruktur

Quelle: Eigene Darstellung mit Microsoft Visio

Aus diesem ER-Modell wurde die Datenbanktabelle und ihre Verbindungen über Fremdschlüssel abgeleitet. In Abbildung 16 werden die Datenbanktabellen, ihre genauen Namensbezeichnungen der Tabelle, dem Datentyp in der Datenbank und ihr Inhalt, sowie ihre Verbindungen abgebildet.

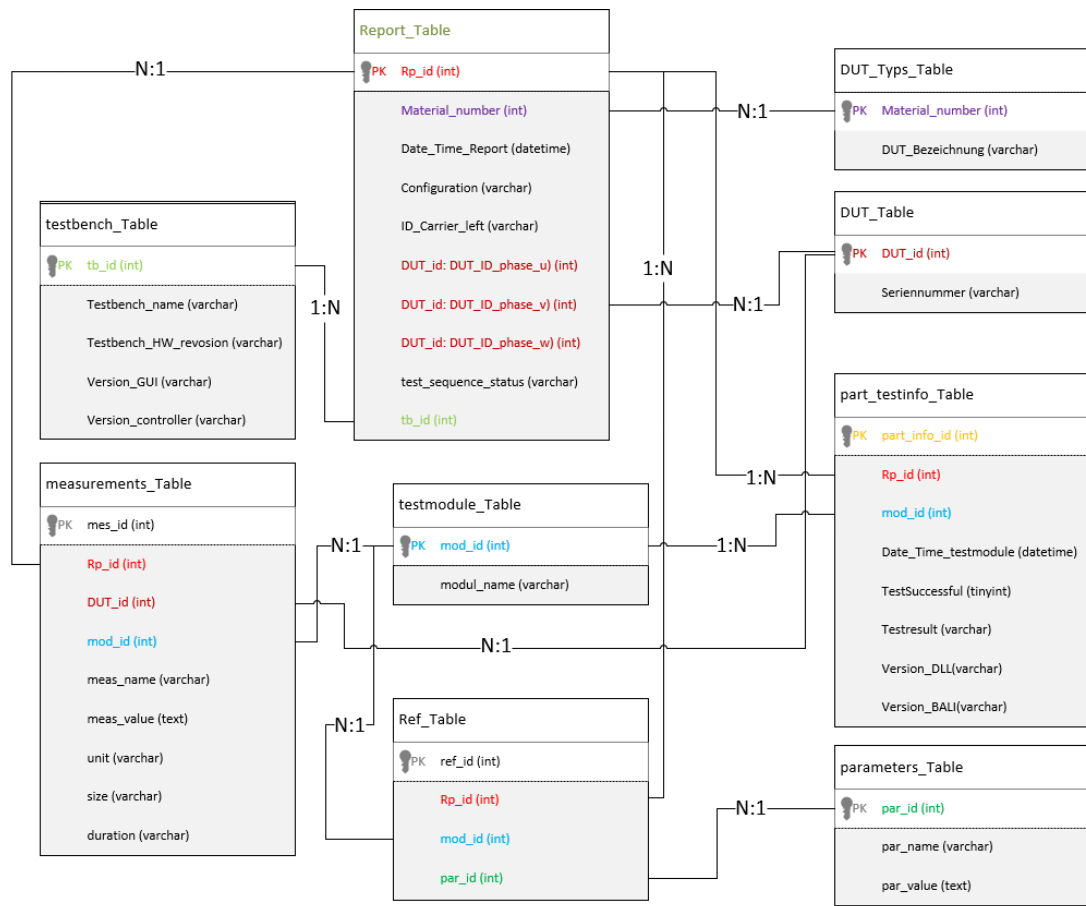


Abbildung 16: Darstellung der Datenbanktabellen und ihrer Verbindungen

Quelle: Eigene Darstellung mit Microsoft Visio

Zur persistenten Datenspeicherung wird das ORM-Framework *Flask-SQLAlchemy* verwendet, das auf *SQLAlchemy* basiert. Es ermöglicht die objektorientierte Abbildung relationaler Datenbanken und vereinfacht den Zugriff auf Daten über Python-Klassen. Dadurch wird eine saubere Trennung von Anwendungslogik und Datenzugriffsschicht erreicht.

### 3.5. Grundkonzept des Benutzeroberflächen-Design

In diesem Abschnitt soll das Grundkonzept der Benutzeroberfläche beschrieben und dargestellt werden.

Die Benutzeroberfläche wird in drei Seiten aufgeteilt, welche sich mit dem Einlesen der Daten, dem darstellen und auswählen zu visualisierenden Daten und dem Ausgeben des Ergebnis befasst. Für das Navigieren zwischen den Seiten soll sich auf jeder Seite eine Navigation-Menü im oberen Bereich der Seite befinden. Zudem sollen alle Seiten einen Bereich für Systemnachrichten wie Bestätigung oder

Fehlermeldungen und ein Label für die Seitenüberschrift besitzen. Diese soll direkt unter den Navigation- Menü angesiedelt werden, um die Grundstruktur identisch aufzubauen. In den nach-

folgenden Text werden die Konzepte der Benutzeroberflächen und ein grundlegendes Benutzungskonzept der einzelnen Seiten anhand der Abbildungen 17, 18 und 19 beschrieben.

Für das Einlesen der XML-Daten soll ein Eingabebox genutzt werden, in die die XML-Struktur kopiert wird. Die Eingaben soll über einen Button unter der Eingabebox bestätigt werden. Bei einem erfolgreichen Einlesen und einfügen in die Datenbank soll eine Bestätigungsnachricht im Bereich für Systemnachrichten erscheinen. Bei einem Fehler soll in diesem Bereich eine Fehlermeldung mit möglicher Lösung erscheinen. Der Aufbau dieser Seite ist in der Abbildung 17 dargestellt.

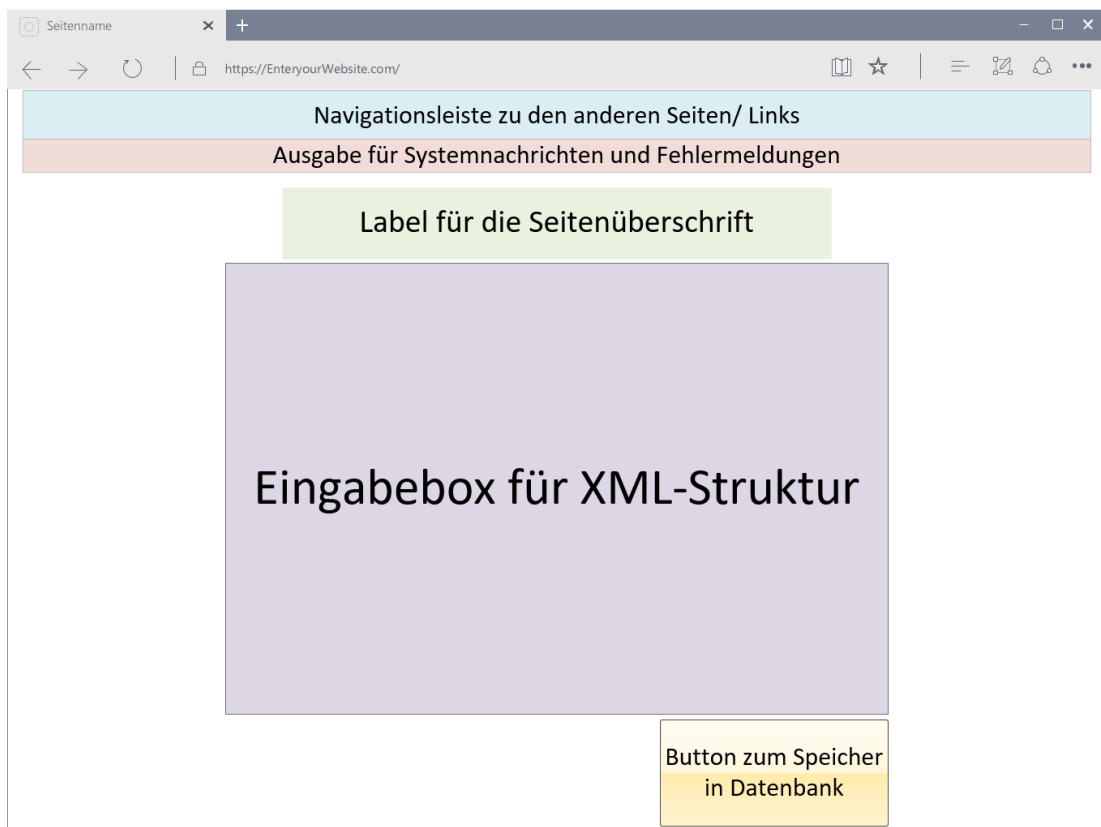


Abbildung 17: Benutzeroberflächenentwurf der Seite zum Einlesen der XML-Struktur

Quelle: Eigene Darstellung mit Microsoft Visio

Die Seite für das Darstellen und Auswählen zu visualisierenden Daten soll im oberen Abschnitt der Seite ein Bereich für Filtereinstellungen besitzen, mit möglichen Filteroption wie Datum oder Seriennummer. Neben dem Filtereinstellung soll zwei Button für das Bestätigen und Zurücksetzen der Filtereinstellungen sein. Bei einem Fehler bei der Filtereingaben soll in Bereich für Systemnachrichten eine Fehlermeldung mit möglicher Lösung erscheinen. Im unteren Bereich der Seite soll sich eine Tabelle, welche die eingelesenen Berichte in der Datenbank anzeigt. Diese soll durch die Filtereinstellungen angepasst werden. Hierbei muss auf eine Anzeige Möglichkeit für länge Tabellenstrukturen berücksichtigt werden, um die Benutzung effizient zu halten. Das Grundkonzept ist in Abbildung 18 dargestellt.

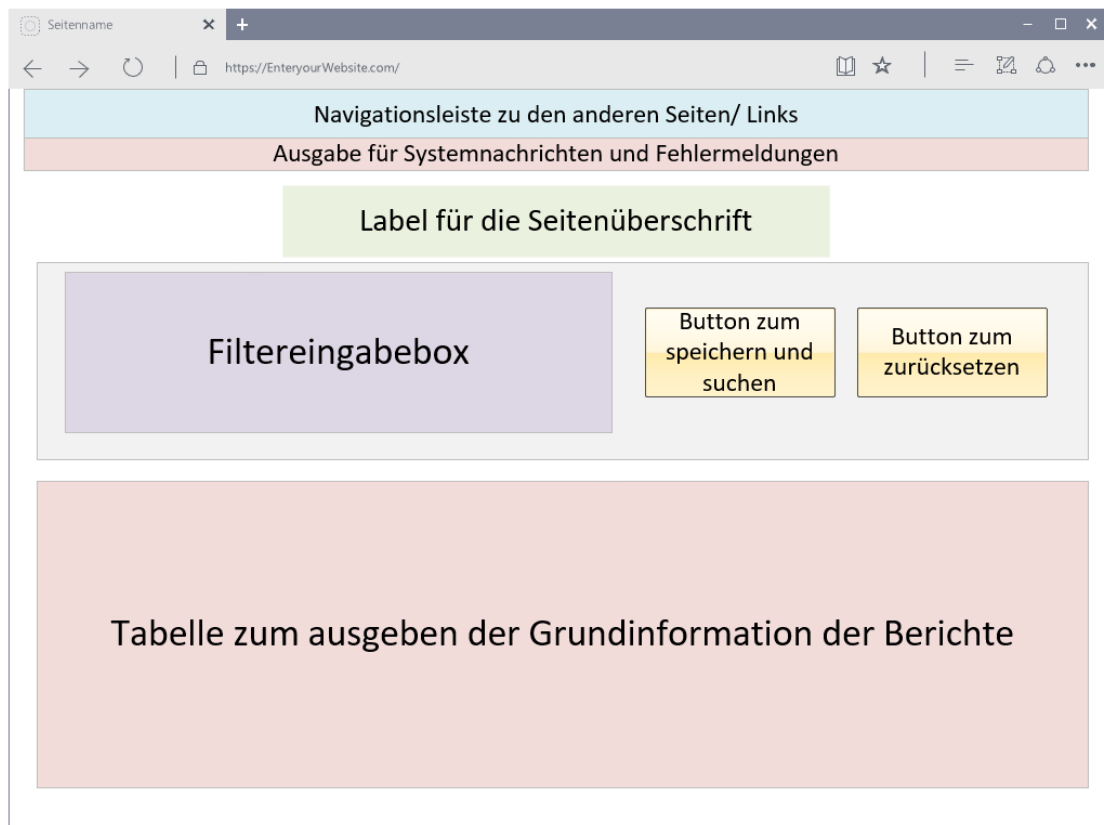


Abbildung 18: Benutzeroberflächenentwurf der Seite zum Ausgeben der Berichtstabelle

Quelle: Eigene Darstellung mit Microsoft Visio

Für das Ausgeben der ausgewählten Berichtsdaten soll im oberen Abschnitt der Seite eine Instanz eingefügt werden, die die Information zu dem Ausgewählten Berichtsdaten angezeigt werden. Im unteren Bereich soll ein Label für die Seriennummer der Ausgewählten Berichtsdaten eingefügt werden. Darunter werden Enthalten den Modulen mit Labeln benannt und unter den Modulname werden die Graphen und Werte für die visuelle Darstellung der Testdaten eingefügt. Die Anzahl der Darstellungen hängt von den in den Bericht Daten enthaltenden Modulen ab. Die Anzahl und der Aufbau des Unteren Bereiches ist variabel. Der grundlegende Aufbau dieser Seite ist in der Abbildung 19 dargestellt.



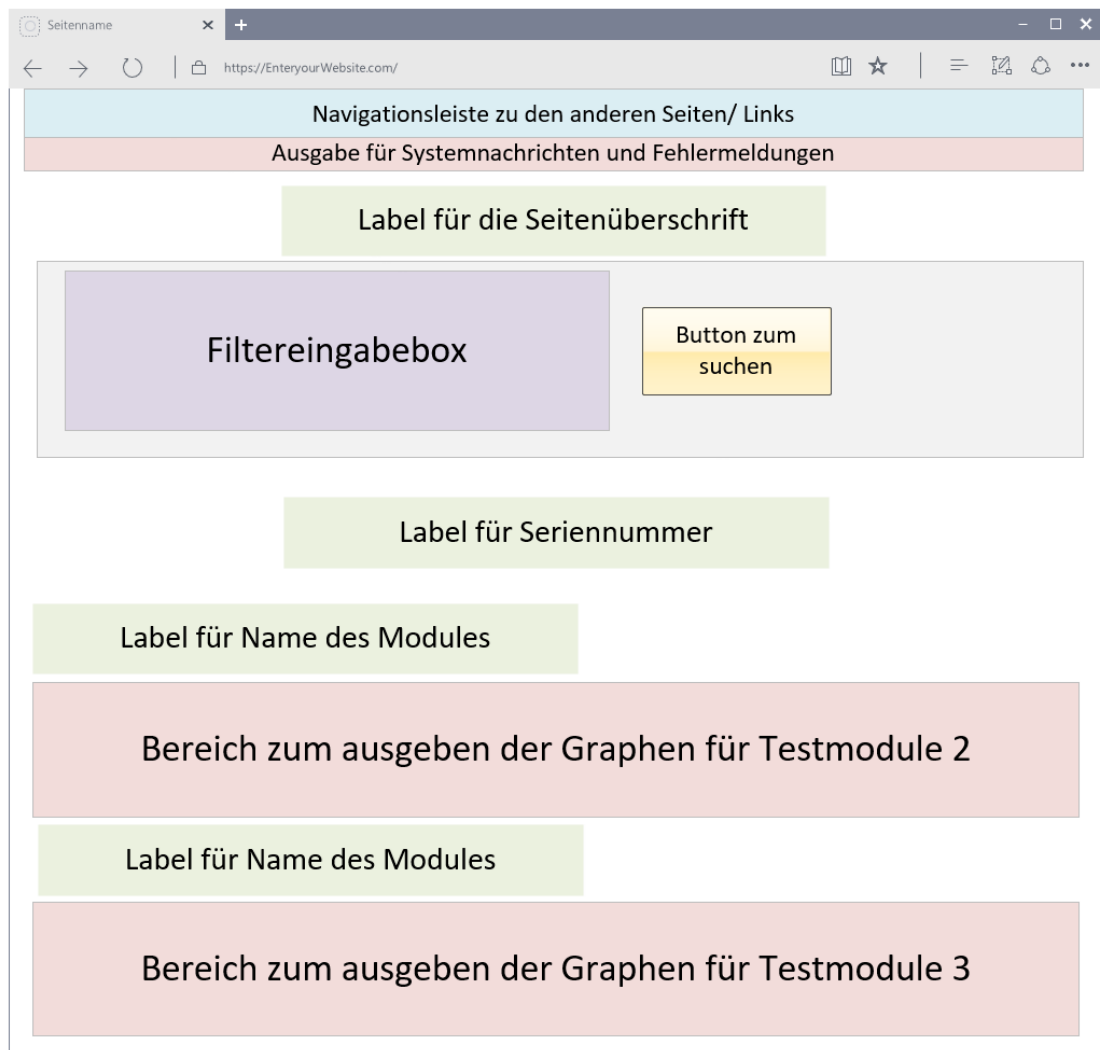


Abbildung 19: Benutzeroberflächenentwurf der Seite zum Ausgeben der Graphen

Quelle: Eigene Darstellung mit Microsoft Visio

### 3.6. Entwurf der Applikationsarchitektur

Der Entwurf der Applikationsarchitektur bildet die konzeptionelle Grundlage für die technische Umsetzung der entwickelten Web-Applikation. Das Ziel ist es, eine Struktur zu schaffen, die modular, wartbar und erweiterbar ist, und die gleichzeitig den funktionalen Anforderungen gerecht wird und sich in die bestehende Systemlandschaft integrieren lässt. Die Architektur der Anwendung ist schichten- und komponentenorientiert aufgebaut und nutzt das Webframework Flask, welches eine klare Trennung zwischen der Präsentations-, Logik- und Datenhaltungsschicht ermöglicht.

#### 3.6.1. Architekturübersicht

Die Applikation soll als serverbasierte Webanwendung im Client-Server-Modell funktionieren. Das bedeutet: Während der Server die Datenverarbeitung, -speicherung und -bereitstellung über-

nimmt, ermöglicht der Client über den Webbrowser die Darstellung und Interaktion. Die Anwendung ist in mehrere Schichten gegliedert, zu denen man die einzelnen Programmteile zuordnen kann. Die Anwendung wird in folgende Schichten gegliedert:

- Präsentationsschicht (Frontend): Bereitstellung der Benutzeroberfläche über HTML-Templates, CSS und JavaScript.
- Applikationslogik (Backend): Implementierung der Geschäftslogik, Steuerung des Datenflusses und Verarbeitung der XML-Dateien.
- Datenhaltungsschicht: Persistente Speicherung der extrahierten Mess- und Gerätedaten in einer relationalen Datenbank mittels ORM.

Ein schematisches Architekturdiagramm dieser Struktur ist in Abbildung 20 dargestellt.

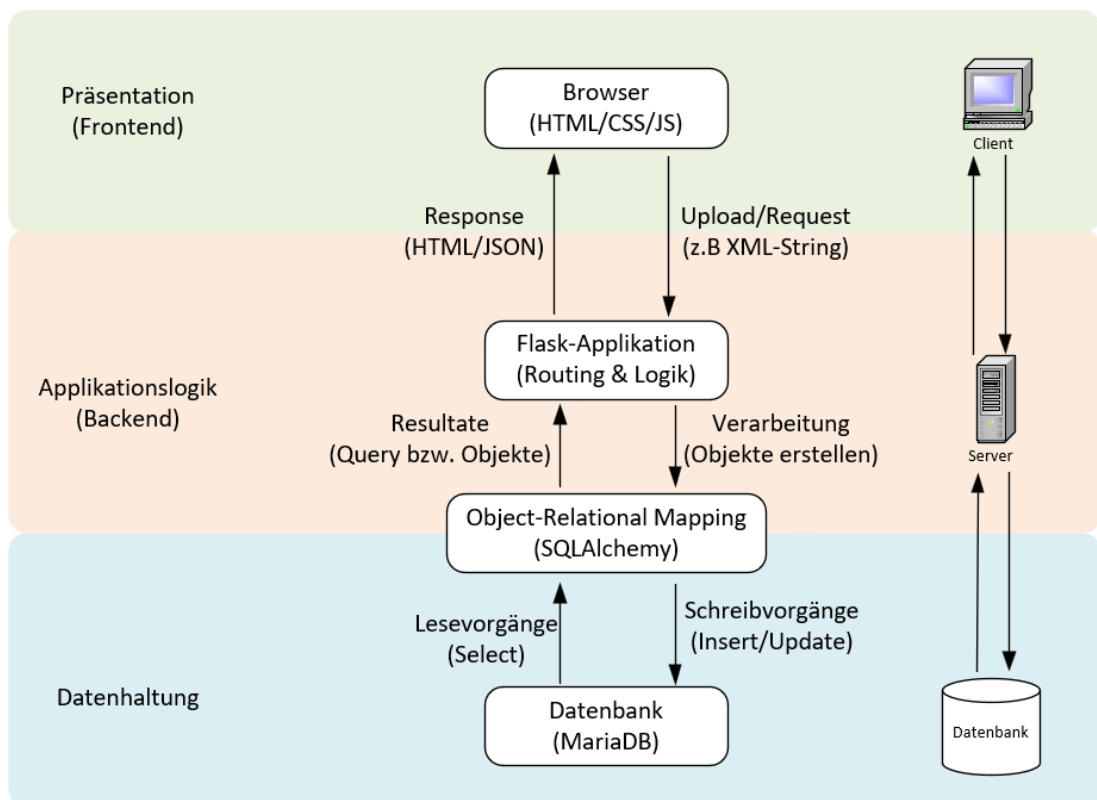


Abbildung 20: schematisches Architekturdiagramm der Applikation

Quelle: Eigene Darstellung mit Microsoft Visio

### Präsentationsschicht

Die Präsentationsschicht besteht aus einer Kombination aus HTML-Templates und statischen Ressourcen (CSS, JavaScript), die im Verzeichnis `templates/` bzw. `static/` abgelegt sind.

- Die Templates `upload.html`, `reports.html` und `dut_report.html` bilden die Kernseiten der Weboberfläche.

- Statische Dateien wie `style.css` und `dut_plots.js` dienen der Gestaltung und der interaktiven Darstellung von Messergebnissen.

Die Kommunikation mit der Logikschicht erfolgt über die definierten *Flask Blueprints*, die als modulare Controller agieren.

## Applikationslogik

Die zentrale Geschäftslogik ist in modularen *Blueprints* und *Service-Komponenten* realisiert.

Die drei Blueprints `upload`, `reports` und `dut` kapseln die jeweiligen Funktionsbereiche:

- `upload`: Einlesen und Validieren von XML-Dateien.
- `dut`: Verarbeitung und Anzeige der Daten eines „Device Under Test“.
- `reports`: Zusammenstellung und Ausgabe von Auswertungen.

Unterstützt werden diese durch das Modul `services/`, das die eigentliche Logik zur Datenverarbeitung bereitstellt:

- `xml_ingest.py` übernimmt das Parsen und Einlesen der XML-Daten.
- `queries.py` enthält vordefinierte Datenbankabfragen.
- `utils.py` und `utils_timeseries.py` stellen Hilfsfunktionen zur Verfügung, z. B. zur Zeitreihenanalyse und Datenaufbereitung.
- `report_builders.py` dient der Erstellung und Formatierung von Report-Daten für die grafische Darstellung.

Die Applikationslogik wird über die Datei `run.py` initialisiert, welche den Flask-Server startet und die Anwendungskonfiguration aus `config.py` einliest.

## Datenhaltungsschicht

Die Datenhaltung erfolgt über ein relationales Datenbanksystem, das über das Flask-eigene *SQLAlchemy ORM* angebunden ist.

Die Datenmodelle sind im Verzeichnis `models/` definiert und bilden die logischen Entitäten der Prüfanlage ab, darunter:

- `dut.py` (Device Under Test)
- `measurement.py` (Messdaten)
- `parameter.py` (Prüfparameter)
- `testbench.py`, `module.py` und `ref.py` (Referenzdaten und Prüfaufbau)

Die Beziehungen zwischen den Modellen ermöglichen eine strukturierte und relationale Abbildung der Prüfdaten, wodurch eine effiziente Abfrage und Analyse möglich ist.

Migrationen werden mit *Alembic* verwaltet, wie das Verzeichnis migrations/ zeigt.

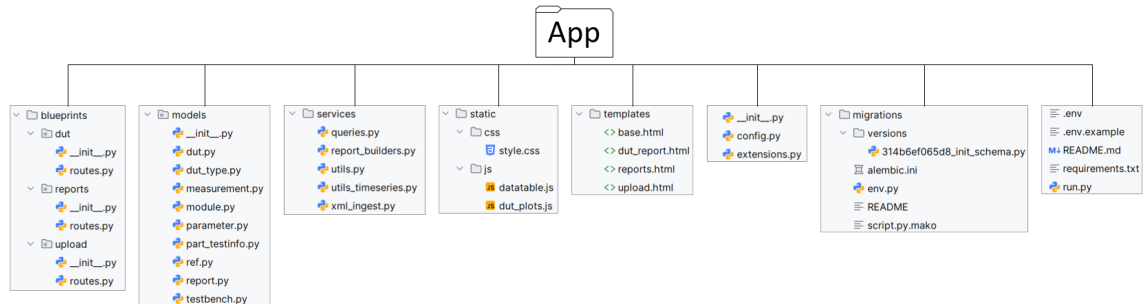


Abbildung 21: Grundlegenden Struktur des Applikations-Ordners

Quelle: Eigene Darstellung mit Microsoft Visio

### 3.6.2. Kommunikation und Datenfluss

Alles startet mit dem Hochladen einer XML-Datei über die Weboberfläche. Der Blueprint upload übergibt die Datei an das Modul `xml_ingest`, welches die XML-Struktur analysiert, relevante Informationen extrahiert und sie über die Datenmodelle in der Datenbank speichert. Die über den Blueprint reports oder dut gespeicherten Daten können anschließend abgerufen und visuell über die Templates dargestellt werden. Die Schichten kommunizieren ausschließlich über klar definierte Schnittstellen und ORM-Modelle, was eine lose Kopplung und eine hohe Wartbarkeit ermöglicht.

## **4. Implementierung**

In diesem Kapitel wird die technische Umsetzung der in Kapitel 4 beschriebenen Applikationsarchitektur behandelt. Hierbei werden die einzelnen Bausteine der erstellten Web-Applikation erklärt, die Implementierung beschrieben und die Beziehung zwischen den Modulen dargestellt.

Die Implementierung erfolgte schrittweise in iterativen Entwicklungszyklen. Zunächst wurde das Grundgerüst der Flask-Anwendung erstellt, anschließend die Datenbankbindung und zuletzt die Benutzeroberfläche mit den Auswertungs- und Visualisierungsfunktionen integriert. Dieser Aufbau erleichterte die fortlaufende Validierung und Testung einzelner Funktionsbereiche während der Entwicklung.

Die in diesem Kapitel dargestellten Ergebnisse entsprechen dem Stand zur Abgabe der Arbeit. Die Vollkommenheit und Funktionalität zu dem Abgabezeitpunkt wird in den späteren Kapiteln 6 und 7 behandelt.

### **4.1. Einlesen und Verarbeiten von XML-Daten**

### **4.2. Implementierung der Datenbank**

### **4.3. Entwicklung der Benutzeroberfläche**

### **4.4. Technische Details zur Visualisierung**

## **5. Integration und Test**

### **5.1. Einbindung in die bestehende Systemlandschaft**

### **5.2. Testmethoden und Durchführung**

### **5.3. Ergebnisse der Testmethoden**

## **6. Fazit und Ausblick**

### **6.1. Zusammenfassung der Ergebnisse**

### **6.2. Kritische Bewertung**

### **6.3. Möglichkeiten für zukünftige Erweiterungen**

## i. Literaturverzeichnis

### Literatur

- [1] A. GmbH, *USTB DWT (XCT0006-1) Main Manual V1.0*, Aachen: AixControl GmbH, 2018.
- [2] G. Neumann, „Auszeichnungssprache,“ in *Enzyklopädie der Wirtschaftsinformatik – Online-Lexikon*, Online; abgerufen am 18.08.2025, Berlin: GITO, 2019. besucht am 18. Aug. 2025. Adresse: <https://wi-lex.de/index.php/lexikon/technologische-und-methodische-grundlagen/sprache/auszeichnungssprache/>
- [3] P. Brezany, *XML und Datenbanken*, Vorlesungsskript, Online; abgerufen am 22.09.2025, 2003. besucht am 22. Sep. 2025. Adresse: <https://homepage.univie.ac.at/peter.brezany/teach/kfk/02ws-vo/skriptum/14-01-03.pdf>
- [4] M. Becher, „XML-Grundlagen,“ in *XML: DTD, XML-Schema, XPath, XQuery, XSL-FO, SAX, DOM*. Wiesbaden: Springer Fachmedien Wiesbaden, 2022, ISBN: 978-3-658-35435-0. DOI: 10.1007/978-3-658-35435-0\_1 Adresse: [https://doi.org/10.1007/978-3-658-35435-0\\_1](https://doi.org/10.1007/978-3-658-35435-0_1)
- [5] B. Jung. „Homepage Webhilfe: Grundlagen XPath.“ Online; abgerufen am 22.09.2025, besucht am 22. Sep. 2025. Adresse: <https://www.homepage-webhilfe.de/XML/XPath/>
- [6] Python Software Foundation. „xml.etree.ElementTree — The ElementTree XML API.“ Online; abgerufen am 21.08.2025, besucht am 21. Aug. 2025. Adresse: <https://docs.python.org/3/library/xml.etree.elementtree.html#module-xml.etree.ElementTree>
- [7] Python Software Foundation. „xml.dom.minidom — Minimal DOM implementation.“ Online; abgerufen am 22.09.2025, besucht am 22. Sep. 2025. Adresse: <https://docs.python.org/3/library/xml.dom.minidom.html>
- [8] Python Software Foundation. „xml.sax — Support for SAX2 parsers.“ Online; abgerufen am 22.09.2025, besucht am 22. Sep. 2025. Adresse: <https://docs.python.org/3/library/xml.sax.html>
- [9] S. Richter. „lxml — XML and HTML with Python.“ Online; abgerufen am 22.09.2025, besucht am 22. Sep. 2025. Adresse: <https://lxml.de/>
- [10] Python Software Foundation. „beautifulsoup4 4.13.5.“ Online; abgerufen am 22.09.2025, besucht am 22. Sep. 2025. Adresse: <https://pypi.org/project/beautifulsoup4/>
- [11] Python Software Foundation. „xmlschema 4.1.0.“ Online; abgerufen am 22.09.2025, besucht am 22. Sep. 2025. Adresse: <https://pypi.org/project/xmlschema/>
- [12] Python Software Foundation. „defusedxml 0.7.1.“ Online; abgerufen am 22.09.2025, besucht am 22. Sep. 2025. Adresse: <https://pypi.org/project/defusedxml/>
- [13] C. Stefanescu. „untangle: Convert XML to Python objects.“ Online; abgerufen am 22.09.2025, besucht am 22. Sep. 2025. Adresse: <https://untangle.readthedocs.io/en/latest/>



- [14] E. Fuchs, *SQL: Grundlagen und Datenbankdesign*, 1. Ausgabe. Bodenheim: HERDT-Verlag, Juli 2021, ISBN: 978-3-98569-009-1.
- [15] F. Herrmann, *Datenorganisation und Datenbanken: Praxisorientierte Übungen mit MS Access 2016*. Wiesbaden: Springer Vieweg, 2018, ISBN: 978-3-658-21330-5. DOI: 10.1007/978-3-658-21331-2 Adresse: <https://doi.org/10.1007/978-3-658-21331-2>
- [16] A. Gadatsch, *Datenmodellierung: Einführung in die Entity-Relationship-Modellierung und das Relationenmodell*, 2., aktualisierte Auflage. Wiesbaden: Springer Vieweg, 2019, ISBN: 978-3-658-25729-3. DOI: 10.1007/978-3-658-25730-9 Adresse: <https://doi.org/10.1007/978-3-658-25730-9>
- [17] Oracle Corporation. „MySQL Connector/Python.“ Version 9.1.0, zuletzt abgerufen am 09. Oktober 2025, Oracle Corporation. Adresse: <https://pypi.org/project/mysql-connector-python/>
- [18] Federico Di Gregorio und Daniele Varrazzo. „psycopg2 – PostgreSQL Database Adapter for Python.“ Version 2.9.10, zuletzt abgerufen am 09. Oktober 2025, Psycopg Project. Adresse: <https://pypi.org/project/psycopg2/>
- [19] MariaDB Foundation. „MariaDB Connector/Python Documentation.“ Zuletzt abgerufen am 09. Oktober 2025, MariaDB Foundation. Adresse: <https://mariadb.com/docs/connectors/mariadb-connector-python>
- [20] M. Grinberg. „The Flask Mega-Tutorial, Part IV: Database.“ Zugriff am 8. Oktober 2025. Adresse: <https://blog.miguelgrinberg.com/post/the-flask-mega-tutorial-part-iv-database>
- [21] Pallets Projects, *Flask-SQLAlchemy Documentation (Version 3.x)*, Zugriff am 8. Oktober 2025, 2025. Adresse: <https://flask-sqlalchemy.readthedocs.io/en/stable/quickstart/>

## **ii. Anhangsverzeichnis**

### **iii. Erklärung**

Hiermit erkläre ich, dass ich die von mir eingereichte Bachelor- / Masterarbeit ".....(Titel der Arbeit).....Belbständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt habe.

Ort und Datum

persönliche Unterschrift

(Name des Verfassers)