

# **Effektive Informationsvisualisierung in digitalen Informationssystemen bei hoch-heterogener Daten- und Datenquellendichte**

## **Thesis**

zur Erlangung des Grades

### **Master of Science**

im Studiengang Business Application Architectures

an der Fakultät Wirtschaftsinformatik

der Hochschule Furtwangen University

vorgelegt von

**Tobias Straub**

---

Referenten: Prof. Dr. Stefan Noll  
Prof. Dr. Christoph Meinel

Eingereicht am: 27. Februar 2018



## **Eidesstattliche Erklärung**

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Thesis selbständig und ohne unzulässige fremde Hilfe angefertigt habe.

Die verwendeten Quellen sind vollständig zitiert.

---

Ort, Datum

---

Unterschrift



## **Abstract**

Die vorliegende Masterarbeit analysiert die Problematiken effektiver grafischer Repräsentationen in digitalen Informationssystemen, die mit besonders hoher und dynamischer Datendichte und Datenquellen einhergehen. Anschließend an die Analyse der Problematiken erarbeitet der Autor dieser Forschungsarbeit ein konzeptuelles Modell zur Bewältigung der geschilderten Problematiken, auf Basis von semantisch beschriebenen, wiederverwendbaren grafischen Visualisierungselementen und den ebenfalls semantisch beschriebenen in die Visualisierung zu überführenden Daten.

Besonders hervorzuhebende Erkenntnisse dieser Masterarbeit sind die Identifikation von Qualitätskriterien zur Zielführung einer effektiven Visualisierung gemäß der visuellen Wahrnehmung des menschlichen kognitiven Systems, die Notwendigkeit zur Erweiterung der Vokabularmenge der schema.org-Ontologie zur Anwendung der identifizierten Qualitätskriterien und das Auffinden geeigneter Visualisierungselemente sowie das Zuordnen der Daten zu entsprechenden Visualisierungselementen über den Aufbau und Vergleich einer Baumstruktur für sowohl die Daten als auch die der Visualisierungselemente.

Diese Forschungsarbeit ist von besonderer Relevanz für Entscheider, Projektmanager und Softwareentwickler, die digitale Informationssysteme mit einer hohen Anzahl an heterogenen Datensätzen und Datenquellen entwickeln.



## Inhaltsverzeichnis

<b>Abstract</b> .....	<b>I</b>
<b>Inhaltsverzeichnis</b> .....	<b>III</b>
<b>Abbildungsverzeichnis</b> .....	<b>V</b>
<b>Tabellenverzeichnis</b> .....	<b>VII</b>
<b>1 Einleitung</b> .....	<b>1</b>
1.1 Problemstellung und Forschungsfrage . . . . .	2
1.2 Zielsetzung und Abgrenzung . . . . .	6
1.3 Aufbau der Arbeit . . . . .	7
<b>2 Kontextbetrachtung</b> .....	<b>11</b>
<b>3 Analyse</b> .....	<b>17</b>
3.1 Visualisierung . . . . .	18
3.1.1 Prozess der Wahrnehmung . . . . .	19
3.1.2 Bedeutsamkeit und Zielsetzungen von Visualisierungen .	21
3.1.3 Qualitätskriterien an Visualisierungen . . . . .	23
3.2 Visualisierungsprozess in digitalen Informationssystemen . . . .	27
3.2.1 Visualisierungspipeline . . . . .	27
3.2.2 Statische Informationsvisualisierung . . . . .	29
3.2.3 Voraussetzungen statischer Informationsvisualisierung . .	31
3.2.4 Problematik statischer Informationsvisualisierung . . . . .	32
3.2.5 Statisch-generische Informationsvisualisierung . . . . .	34
3.3 Dynamische Informationsvisualisierung durch vollständige auto- matisierte Generierung grafischer Benutzeroberflächen . . . . .	35

---

<b>4 Entwurf und Design des konzeptuellen Modells</b> .....	<b>37</b>
4.1 Strukturen und Prozesse . . . . .	39
4.1.1 Anatomie grafischer Benutzeroberflächen . . . . .	40
4.1.2 Grundlegender Aufbau der Strukturen und Prozesse . . . . .	44
4.1.3 Konkretisierter Aufbau der Strukturen und Prozesse . . . . .	57
4.1.4 Integration in den Anwendungskontext . . . . .	60
4.2 Spezifikation des Semantisierungsvokabulars . . . . .	65
4.2.1 Grundlagen der Semantik . . . . .	66
4.2.2 Auswahl des Basisvokabulars . . . . .	69
4.2.3 Art und Struktur der Daten . . . . .	78
4.2.4 Zielerreichung des Anwenders . . . . .	84
4.2.5 Vorwissen des Anwenders . . . . .	93
4.2.6 Visuelle Fähigkeiten und Vorlieben des Anwenders . . . . .	99
4.2.7 Konventionen des Anwendungskontexts . . . . .	108
4.2.8 Charakteristika des Darstellungsmediums . . . . .	109
4.2.9 Web Components . . . . .	110
<b>5 Schlussbetrachtung</b> .....	<b>113</b>
<b>Literaturverzeichnis</b> .....	<b>117</b>

## Abbildungsverzeichnis

2.1	Verhältnis zwischen Informationsüberflutung und Entscheidungsqualität . . . . .	12
2.2	Exemplarische Ursachen für Informationsüberflutung . . . . .	13
3.1	Klassische Sinne neben der visuellen Wahrnehmung . . . . .	19
3.2	Prozess der visuellen Wahrnehmung . . . . .	20
3.3	Determinaten der Informationsüberflutung . . . . .	21
3.4	Exemplarische Visualisierung zur Verdeutlichung des Effektivitätskriterium der Visualisierungsqualität . . . . .	24
3.5	Visualisierungspipeline . . . . .	28
3.6	Typischer Visualisierungsprozess in Softwaresystemen . . . . .	30
3.7	Visualisierungsprozess mit unbekanntem und nicht-formalisierten Daten . . . . .	33
3.8	Visualisierungsprozess mit generischer Datenrepräsentation . . . . .	34
4.1	Anatomie grafischer Benutzeroberflächen . . . . .	41
4.2	Exemplarische Darstellung eines Suchformulars assoziiert als Molekül-Atom-Verbund gemäß Atomic design . . . . .	42
4.3	Grundstruktur des konzeptuellen Modells . . . . .	45
4.4	Exemplarische Darstellung eines digitalen Arbeitsblattes . . . . .	49
4.5	Taxonomie von Art der Datenwerte . . . . .	52
4.6	Exemplarische hierarchische Darstellung einer Struktur und eines Visualisierungselements . . . . .	54
4.7	Exemplarische hierarchische Darstellung einer Struktur und eines Visualisierungselements mit Transformation des Datenbaums aufgrund des Visualisierungsziels . . . . .	56
4.8	Prozess von den Daten zur Visualisierung aus Sicht des Akteurs . . . . .	61

4.9 Anwendungsfalldiagramm für die Softwarekomponente der Schul- Cloud . . . . .	63
4.10 Evaluation von Ontologien zur Auswahl eines Basisvokabulars .	75
4.11 LRMI Typen und Eigenschaften integriert in die schema.org Ty- penhierarchie . . . . .	77
5.1 Spezifizierte Vokabularmenge zur Abbildung identifizierter Ein- flussfaktoren der Qualitätskriterien von Visualisierungen . . . . .	115

## Tabellenverzeichnis

4.1	Semantisierungsspezifikation des Begriffs 'Visualization' . . . . .	80
4.2	Semantisierungsspezifikation des Begriffs 'preferredVisualization' . . . . .	81
4.3	Semantisierungsspezifikation des Begriffs 'hasPart' . . . . .	81
4.4	Semantisierungsspezifikation des Begriffs 'isPartOf' . . . . .	82
4.5	Semantisierungsspezifikation des Begriffs 'isNotPart' . . . . .	83
4.6	Semantisierungsspezifikation des Begriffs 'ItemVisualizationGoal' . . . . .	85
4.7	Semantisierungsspezifikation des Begriffs 'Identify' . . . . .	86
4.8	Semantisierungsspezifikation des Begriffs 'Locate' . . . . .	86
4.9	Semantisierungsspezifikation des Begriffs 'Differ' . . . . .	87
4.10	Semantisierungsspezifikation des Begriffs 'Categorize' . . . . .	88
4.11	Semantisierungsspezifikation des Begriffs 'Aggregate' . . . . .	89
4.12	Semantisierungsspezifikation des Begriffs 'Organize' . . . . .	89
4.13	Semantisierungsspezifikation des Begriffs 'Compare' . . . . .	90
4.14	Semantisierungsspezifikation des Begriffs 'IdentifyRelations' . . . . .	91
4.15	Semantisierungsspezifikation des Begriffs 'Associate' . . . . .	91
4.16	Semantisierungsspezifikation des Begriffs 'Correlate' . . . . .	92
4.17	Semantisierungsspezifikation des Begriffs 'visualizationGoal' . . . . .	93
4.18	Semantisierungsspezifikation des Begriffs 'ItemEducationalRoleClassified' . . . . .	95
4.19	Semantisierungsspezifikation des Begriffs 'educationalRoleClassified' . . . . .	95
4.20	Semantisierungsspezifikation des Begriffs 'GeneralEducated' . . . . .	96
4.21	Semantisierungsspezifikation des Begriffs 'SchoolEducated' . . . . .	97
4.22	Semantisierungsspezifikation des Begriffs 'OccupationEducated' . . . . .	97
4.23	Semantisierungsspezifikation des Begriffs 'HigherEducated' . . . . .	98
4.24	Semantisierungsspezifikation des Begriffs 'ScholarlyPerson' . . . . .	98

---

4.25 Semantisierungsspezifikation des Begriffs 'ItemVisualPerceptionDisorder' . . . . .	100
4.26 Semantisierungsspezifikation des Begriffs 'visualPerceptionDisorder' . . . . .	100
4.27 Semantisierungsspezifikation des Begriffs 'SpatialPositionDetection' . . . . .	101
4.28 Semantisierungsspezifikation des Begriffs 'AcquisitionAndDifferentiation' . . . . .	102
4.29 Semantisierungsspezifikation des Begriffs 'FigureBasicPerception'	102
4.30 Semantisierungsspezifikation des Begriffs 'VisuomotorCoordinationDisorder' . . . . .	103
4.31 Semantisierungsspezifikation des Begriffs 'preferredHexColor' . . . . .	104
4.32 Semantisierungsspezifikation des Begriffs 'disfavoredHexColor' . . . . .	105
4.33 Semantisierungsspezifikation des Begriffs 'exclusiveHexColor' . . . . .	105
4.34 Semantisierungsspezifikation des Begriffs 'ItemInteractivityType' . . . . .	106
4.35 Semantisierungsspezifikation des Begriffs 'interactivityType' . . . . .	106
4.36 Semantisierungsspezifikation des Begriffs 'Active' . . . . .	107
4.37 Semantisierungsspezifikation des Begriffs 'Expositive' . . . . .	107
4.38 Semantisierungsspezifikation des Begriffs 'Mixed' . . . . .	108
4.39 Semantisierungsspezifikation des Begriffs 'WebComponent' . . . . .	110
4.40 Semantisierungsspezifikation des Begriffs 'Tag' . . . . .	110
4.41 Semantisierungsspezifikation des Begriffs 'Attribute' . . . . .	111
4.42 Semantisierungsspezifikation des Begriffs 'value' . . . . .	111
4.43 Semantisierungsspezifikation des Begriffs 'startTag' . . . . .	111

## 1 Einleitung

Als Rohstoffe des 21. Jahrhunderts bezeichnete die Bundeskanzlerin Dr. Angela Merkel der Bundesrepublik Deutschland einmal die heutige Masse an Daten [1]. Als am 20. Dezember 1990 an der europäischen Organisation für Kernforschung CERN die weltweit erste Webseite online ging, glaubte noch niemand an einen durch diese Technologie verursachten Wandel der Gesellschaft. Mit dem Jahr 1993 hatten aber bereits 14 Millionen Anwender Zugriff auf 130 verschiedene Webseiten. Damit waren die gesellschaftlichen Veränderungen auf soziologischer Ebene, die durch das von Sir Tim Berners-Lee entwickelten und innovativen World Wide Webs, nicht mehr zu ignorieren [2].

Aus heutiger Sicht werden zunehmend mehr und mehr Lebensbereiche in das World Wide Web verlagert. Einige Aktivitäten substituieren oder ergänzen hierbei sogar Bereiche unseres Alltags und machen diesen für uns noch komfortabler. So können wir heute bereits unseren Einkauf mithilfe des World Wide Webs bis zu unserer Haustüre delegieren lassen, bequem mit Menschen auf der ganzen Welt vom heimischen Sofa aus kommunizieren oder unser kognitives Bewusstsein durch Autodidaktik vom Bett aus, anstatt aus dem Hörsaal erweitern [3].

Bei allen Aktionen die wir im World Wide Web tätigen hinterlassen wir stets Spuren in Form von Daten. Hierbei ist es irrelevant, ob es sich nur um die Eingabe eines Suchbegriffs in einer Suchmaschine handelt, um das Verfassen einer Rezension nach dem Einkauf vom Sofa aus oder um eine anders gelagerte Transaktion [3]. Die hinter den Aktionen stehenden Prozesse erfassen systematisch eine hohe Dichte an quantitativen Daten. Alleine die Suchmaschine Google verarbeitet täglich vierundzwanzig Petabyte an Daten [4]. Hierbei ist Google keine Ausnahme. Soziale Netzwerke wie Facebook oder Twitter haben ähnliche Datenverarbeitungsdaten [5].

Den massiven Anstieg an Daten beobachten Experten bereits seit Jahren. Neunzig Prozent an Daten der gesamten Weltgeschichte entstanden innerhalb von nur zwei Jahren im Zeitraum von 2010 bis 2012. In der Literatur wird ein exponentielles Wachstum des Datenvolumens sowie der Datentypen prognostiziert. Die Autoren Bughin, Chui und Manyika sprechen von einer kontinuierlichen Verdoppelung der Datenmenge nach jeweils achtzehn Monaten [3, 6, 7, 8, 9]. Zusätzliche sozio-technologische Veränderungen wie Big Data und das Internet der Dinge verstärken den Anstieg quantitativer Daten und rücken damit die Limitierungen des menschlichen kognitiven Systems zur Aufnahme von Informationen in den Fokus [7].

In der Fachliteratur wird allgemein davon ausgegangen, dass achtundneunzig Prozent der Informationen, mit denen der Mensch konfrontiert wird, vom Menschen nicht wahrgenommen und verarbeitet werden [7, 10]. Im Kontext des exponentiellen Datenanstiegs hat sich der Begriff der 'Informationsüberflutung' etabliert. Die Begrifflichkeit beschreibt das Phänomen der genannten kognitiven Beschränktheit der Informationsverarbeitungskapazität beim menschlichen Entscheidungsfindungsprozess [7, 11]. Die Forschungsaktivitäten, die sich auf die Informationsüberflutung beziehen, zeigen hierbei nicht ausschließlich die Beschränkungen der kognitiven Fähigkeiten von Menschen auf. Die Aktivitätsresultate definieren exakte Auslöser, die zur menschlichen Beschränkung der Informationsverarbeitungskapazität führen. Zwei der Hauptdeterminanten, die zur Beschränkung der Informationsverarbeitungskapazität führen, sind die Anzahl sowie die Komplexität von Informationen [7, 8, 9].

## **1.1 Problemstellung und Forschungsfrage**

Die im vorangegangenen Kapitel 1 beschriebene Evolution der Verlagerung menschlicher Aktivitäten in das World Wide Web und den damit verbundenen Anstieg quantitativer Daten führt zu der in dieser Forschungsarbeit abzuhandelnden Problemstellung.

Die visuelle Wahrnehmung des kognitiven Systems eines jeden Menschen besitzt eine hohe Aktivierungsdichte an Rezeptoren, Neuronen und dem Cortex.

Dies macht die visuelle Wahrnehmung zum besten Übertragungsmedium für eine hohe Bandbreite an Daten, die zum Gehirn und in das Bewusstsein gelangen sollen [12, 13, 14]. Die Akquirierung von Daten durch die Aufnahme und Verarbeitung visueller Reize zu einem bestimmten Moment durch das unterbewusste Erkennen von Strukturen, Mustern, Trends und Beziehungen ist durch die visuelle Wahrnehmung in deutlich höherem Umfang möglich, als mit allen anderen Sinnen in Kombination [15, 16].

Für die Ableitung relevanter Informationen aus einer großen Datenmenge, auf Grundlage des in Kapitel 1 beschriebenen exponentiellen Anstiegs an Daten, bedarf es eines Mittels zur Verringerung beziehungsweise Vermeidung der kognitiven Beschränktheit der Informationsverarbeitungskapazität. Ein besonders effektives Mittel ist angelehnt an die visuelle Wahrnehmung, die grafische Repräsentation der vom menschlichen kognitiven System zu verarbeitenden Daten. Das Mittel der grafischen Repräsentation hilft insbesondere beim Erkennen von Trends und Mustern in Daten, die sonst in Zeilen und Spalten von Tabellen verborgen geblieben wären. Durch die effektive grafische Repräsentation von Informationen können relevante Daten nicht nur identifiziert, sondern auch verstanden und kommuniziert werden [14].

Eine Überführung von Rohdaten in grafische Repräsentationen wird in digitalen Informationssystemen durch den Einsatz spezifischer Computeralgorithmen realisiert [15, 17]. Bei der Entwicklung eines solchen Computeralgorithmus müssen dem Softwareentwickler bei der Programmierung zwei elementare Faktoren bekannt sein:

- a) **Die Struktur der Daten.** Dem Softwareentwickler muss mindestens bekannt sein, welche Struktur die Daten aufweisen. Nur dann kann er die Daten entsprechend extrahieren und in die Struktur der grafischen Repräsentation überführen [15, 17].
- b) **Die Struktur der grafischen Repräsentation.** Wenn dem Softwareentwickler die Struktur der grafischen Repräsentation unbekannt ist, kann er dementsprechend keine Abbildung von Daten zur grafischen Repräsentation im Algorithmus codieren [15, 17].

Die strukturelle Bekanntheit der Daten zur Zeit der Entwicklung des digitalen Informationssystems ermöglicht den Designern von grafischen Benutzeroberflächen eine spezifische für die Daten und den Informationstypen effektive grafische Repräsentation zu entwickeln. Die Weitergabe des Wissens der Struktur der spezifisch entwickelten grafischen Repräsentation an den Softwareentwickler ermöglicht die Programmierung des Algorithmus zur Zuordnung der Daten zur grafischen Repräsentation. Sind die zwei genannten elementaren Faktoren bekannt, können somit für gegebene Daten die effektivsten Formen der grafischen Repräsentationen entwickelt werden [15, 17, 18].

Für digitale Informationssysteme, die mit Daten arbeiten, die sich aufgrund bestimmter Einflüsse zur Zeit der Entwicklung des Informationssystems nicht statisch abbilden lassen, sind die genannten Faktoren nicht bestimmbar. Exemplarisch hierfür sind externe Datenquellen, deren Konsistenz hinsichtlich der Struktur von Daten nicht dauerhaft gewährleistet werden kann. Ein anderes Beispiel sind Informationssysteme, die Daten aus einer sehr hohen Anzahl an heterogenen Datenquellen beziehen. Oft ist der Aufwand zur Entwicklung spezifischer grafischer Repräsentationen für jede Art differierender Daten sowie die Anpassung der Logik zur Zuordnung der Daten zur grafischen Repräsentation bei einem solchen Szenario zu hoch [15, 17, 18].

Mit der steigenden Tendenz der Generierung quantitativer Daten ist davon auszugehen, dass die in verschiedenen Instanzen entstehenden Datensätze immer seltener eine strukturelle Homogenität aufweisen. Schon heute können Webservices aus einer großen Anzahl unterschiedlicher Formate zur Strukturierung ihrer Inhalte auswählen und damit die Beantwortung der Anfragen flexibel gestalten, bis hin zur Flexibilität der Beantwortung einzelner Anfragen durch unterschiedliche Datenformate [5]. Neben der Heterogenität der Struktur von Daten werden digitale Informationssysteme immer komplexer und beziehen ihre Daten aus unterschiedlichen Quellen. Eine effektive grafische Repräsentation wird damit deutlich erschwert und das Phänomen der Informationsüberflutung begünstigt. Das daraus resultierende Ergebnis ist eine verlangsamte Informationsaufnahme durch den Betrachter.

Aus der beschriebenen Problematik, die den Gegenstand dieser Forschungs-

arbeit repräsentiert, kann entsprechend die Folgende in der Forschungsarbeit zu beantwortende Forschungsfrage abgeleitet werden: Wie können effektive grafische Repräsentationen für Daten gewährleistet werden, die zur Entwicklungszeit der Anwendung entweder nicht bekannt oder aufgrund der hohen Menge heterogener Datenstrukturen sowie Datenquellen nicht betrachtet werden können?

Die Forschungsfrage lässt sich durch die Erkenntnisse der folgenden Leitfragen beantworten:

1. Was wird unter einer effektiven Informationsvisualisierung verstanden und wie kann eine solche effektive Informationsvisualisierung gewährleistet werden?
2. Wie wird die Informationsvisualisierung in heutigen digitalen Informationssystemen realisiert und welche Problematiken hinsichtlich des Effektivitätskriteriums existieren derzeit bei der Informationsvisualisierung in heutigen digitalen Informationssystemen?
3. Welche Problematiken existieren bei der dynamischen Informationsvisualisierung durch vollständige, automatisierte Generierung von Benutzeroberflächen und warum ist der Einsatz dieses Verfahrens ungeeignet zur Beantwortung der Forschungsfrage?
4. Welche architektonischen Strukturen und Prozesse muss ein konzeptuelles Modell aufweisen, um die erst zur Laufzeit bekannt werdenden heterogenen Datenstrukturen und Datenquellen gemäß dem Effektivitätskriterium grafisch zu repräsentieren und wie wird eine mögliche Integration der grafischen Repräsentation in den Anwendungskontext konzeptionell realisiert?
5. Wie muss eine Vokabularmenge zur Semantisierung spezifiziert werden, sodass Daten und grafische Repräsentationselemente semantisch beschrieben werden können und ein Auffinden von grafischen Repräsentationselementen in Abhängigkeit zum Effektivitätskriterium für entsprechende Daten als auch das Zuordnen der entsprechenden Daten zum

gefundenen grafischen Repräsentationselement, gemäß dem konzeptuellen Modells ermöglicht wird?

## 1.2 Zielsetzung und Abgrenzung

Ziel dieser Forschungsarbeit ist die Beantwortung der in vorangegangenen Kapitel 1.1 definierten Forschungsfrage durch die Bewältigung der erläuterten Problemstellung. Die ebenfalls im vorangegangenen Kapitel 1.1 aufgestellten Leitfragen zur Beantwortung der Forschungsfrage spezifizieren die Zielsetzung dieser Arbeit.

Zusammenfassend lässt sich die Zielsetzung durch die Analyse des Begriffs der Effektivität mit entsprechenden Voraussetzungen im Kontext grafischer Repräsentationen beschreiben, sowie die detaillierte Analyse der Problematiken beim grafischen statischen Repräsentationsprozess mit einer hohen Dichte heterogener Datenquellen und Datenstrukturen bei der Softwareentwicklung eines digitalen Informationssystems. Mit den Ergebnissen der Analyse soll ein konzeptuelles Modell entwickelt werden, dass die beschriebene Problematik durch das Ermöglichen effektiver grafischer Repräsentationen bewältigt, trotz einer gegebenen hohen Dichte an Datenquellen sowie Datenstrukturen, die unter Umständen zur Softwareentwicklungszeit des Informationssystems unbekannt sind.

Zur Konkretisierung des konzeptuellen Modells dient die vom Hasso-Plattner-Institut [19] entwickelte Schul-Cloud [20]. Die Schul-Cloud ermöglicht Schülerinnen und Schülern einen niederschweligen Zugang zu digitalen Unterrichtsmaterialien. Hierbei bezieht die Software entsprechende Daten aus einer Vielzahl von Datenquellen mit unterschiedlichen Strukturen der Daten. Die Datenquellen sind hierbei technischer Natur (beispielsweise durch Programmierschnittstellen wie die von Wikipedia [21]) als auch menschlicher Natur durch das Zuführen von Daten durch beispielsweise Lehrerinnen und Lehrer. Die Schul-Cloud als Szenario für das konzeptuelle Modell wird in Kapitel 2 detailliert betrachtet.

Das in dieser Forschungsarbeit entwickelte konzeptuelle Modell ist generell auch für Domänen außerhalb des E-Learnings verwendbar. Dennoch wird an einigen Stellen dieser Arbeit der Fokus primär auf den Bereich des E-Learnings gelegt, gemäß dem Anwendungskontext der Schul-Cloud. Im Rahmen des konzeptuellen Modells dieser Arbeit wird ein semantisches Modell entwickelt. Diese Arbeit behandelt dennoch nicht die Auszeichnung von Inhalten durch Semantik, genau so wenig wie die Transformation von Rohdaten in das semantische Modell.

### **1.3 Aufbau der Arbeit**

Der Aufbau dieser Forschungsarbeit gliedert sich auf oberster Ebene in die fünf folgenden Abschnitte: Einleitung, Kontextbetrachtung, Analyse, Entwurf und Design des konzeptuellen Modells sowie die Schlussbetrachtung.

Im ersten Abschnitt ermöglicht die Masterarbeit einen Einblick in den Forschungsgegenstand, definiert die Problemstellung und Forschungsfrage und deklariert die Zielsetzung sowie Abgrenzung dieser Arbeit zu anderen Forschungsarbeiten als auch zur Eingrenzung des Forschungsgegenstands selbst. Zusätzlich erfolgt in diesem Abschnitt die Erläuterung zum Aufbau dieser Forschungsarbeit.

Der zweite Abschnitt betrachtet und erläutert den Kontext dieser Forschungsarbeit hinsichtlich der Problemstellung und Forschungsfrage. Hierbei wird für den Entwurf und das Design des konzeptuellen Modells, das in Abschnitt vier erläutert wird, eine spezifische Domäne und daraus eine spezifische Anwendung zur Veranschaulichung gewählt.

Anschließend zum zweiten Abschnitt erfolgt der dritte Abschnitt, der den Forschungsgegenstand anhand der im ersten Abschnitt definierten Problemstellung analysiert. Der Abschnitt gliedert sich in drei Teile. Der erste Teil untersucht gemäß der Problemstellung den Bereich der Visualisierung. Hierzu wird die Aufnahme von Informationen durch den menschlichen Prozess der Perception analysiert, die Bedeutsamkeit und Zielsetzung von Visualisierungen

für die Perzeption verdeutlicht und Qualitätskriterien an Visualisierungen zur Unterstützung des Perzeptionsprozesses identifiziert. Im zweiten Teil des Abschnitts wird der in den heutigen digitalen Informationssystemen verwendeter, Visualisierungsprozess skizziert. Die Skizzierung des statischen Visualisierungsprozesses in heutigen digitalen Informationssystemen wird hierbei durch die Einführung der Visualisierungspipeline erläutert. Daran anschließend werden die Voraussetzungen und die hierdurch verursachten Problematiken statischer Informationsvisualisierung identifiziert und gemäß Problemstellung analysiert, sowie die Alternative zur Bewältigung der Problematiken anhand eines statisch-generischen Modells aufgezeigt. Weiterhin wird in Teil drei des Abschnitts die vermeintliche Bewältigung von Problematiken statischer Informationsvisualisierung hinsichtlich der Effektivität von Visualisierungen zur Unterstützung der Perzeption anhand der dynamischen Informationsvisualisierung durch die vollständig-automatisierte Generierung grafischer Benutzeroberflächen aufgezeigt.

Der vierte Abschnitt dieser Forschungsarbeit stellt den Entwurf und das Design eines konzeptuellen Modells vor, das die Bewältigung der definierten Problemstellung zur Folge hat. Zum Entwurf und Design des konzeptuellen Modells werden die Erkenntnisse der Analyse aus Abschnitt drei genutzt. Der Abschnitt gliedert sich in zwei Teile. Im ersten Teil des Abschnitts werden die Strukturen und Prozesse definiert, die an ein solches konzeptuelles Modell zur Bewältigung der Problemstellung in Aussicht werden. Hierzu wird die Anatomie grafischer Benutzeroberflächen untersucht, der grundlegende Aufbau der Strukturen und Prozesse konzipiert sowie anschließend konkretisiert und verfeinert. Weiterhin wird das Integrationsdesign des konzeptuellen Modells in den im Abschnitt zwei erläuterten, Anwendungskontext entworfen. Der zweite Teil des Abschnitts spezifiziert das Vokabular zur Semantisierung, das für das konzeptuelle Modell von primärer Relevanz ist. Hierzu werden beginnend die Grundlagen der Semantik exkursiert und anschließend das Basisvokabular ausgewählt auf deren Grundlage die Spezifikation des Vokabulars zur Bewältigung der Problemstellung durchgeführt wird. Das zu spezifizierende Vokabular, aufbauend auf gewähltem Basisvokabular wird auf Ebene der Erkenntnisse identifizierter Einflussfaktoren der Qualitätskriterien an eine Visualisierung

durchgeführt. Dies sind die Einflussfaktoren Art und Struktur der Daten, Zielerreichung des Anwenders, Vorwissen des Anwenders, Visuelle Fähigkeiten und Vorlieben des Anwenders, Konventionen des Anwendungskontexts sowie Charakteristika des Darstellungsmediums. Zuletzt wird das Vokabular für die Visualisierung selbst auf Betrachtungsgrundlage von Web Components spezifiziert.

Im fünften Abschnitt wird eine Retrospektive der Forschungsarbeit durchgeführt. Insbesondere die Erkenntnisse der einzelnen Abschnittsbereiche die mit deren definierten Teilfragen zur Beantwortung der Forschungsfrage führen stehen im Fokus. Der Abschnitt deklariert außerdem, aufbauend auf dieser Masterarbeit, ein wertendes Fazit und einen Ausblick für weitere Forschungsaktivitäten.

In den Anhängen zu dieser Masterarbeit sind zusätzliche Informationen definiert, die von entsprechenden Kapiteln dieser Arbeit referenziert werden.



## 2 Kontextbetrachtung

Die fortschreitende Digitalisierung sorgt für eine ununterbrochene Akkumulation von Daten [3, 12]. Eine Vielzahl dieser sich ansammelnden Daten generieren wir als Menschen durch ein Pro-aktives Verhalten. Beispielsweise durch die Nutzung der Fotokamera unseres Smartphones oder das Versenden einer Text- und Sprachnachricht. Viele Daten im heutigen Informationszeitalter werden allerdings zusätzlich angehäuft ohne ein unmittelbares menschliches Zutun. Ein Beispiel hierfür ist das Internet der Dinge. Im Internet der Dinge werden Alltagsgegenstände mit Sensorik versehen, die stetig und rund um die Uhr, Impulse erfassen und aus diesen Impulsen Daten generieren [12]. Die Digitalisierung ermöglicht die Messung und Vermessung unserer Welt und erzeugt dabei immense Daten.

Der durch die zunehmende Digitalisierung verursachte Anstieg an Daten erschwert die Identifikation relevanter Informationen im Gegensatz zur Vergangenheit massiv. War die Identifikation relevanter Informationen aus einer Menge von Daten in der Vergangenheit mittels einer einfachen Suche durch eine Suchmaschine durch den Abgleich der Suchzeichenkette mit den Vorkommen in der zu durchsuchenden Datenmenge ausreichend, ist dies als alleiniges Relevanzkriterium nicht weiter ausreichend [22]. Eine zu hohe Anzahl an Informationen machen es schwierig, die für den Informationsentscheider relevanten Informationen zu filtern [23]. Die Betrachtung einer zu hohen, irrelevanten oder unpräzisen Datendichte hingegen verursacht das in Kapitel 1 betrachtete Problem der Informationsüberflutung. Die kognitiven Fähigkeiten des Menschen sind somit nicht ausreichend, für eine adäquate Verarbeitung einer zu hohen, irrelevanten oder unpräzisen Datendichte [7, 11, 24]. Wie bereits in Kapitel 1 erwähnt, gehen Kognitionswissenschaftler davon aus, dass nur zwei Prozent der Informationen mit denen ein Mensch konfrontiert wird,

von diesem auch wahrgenommen und verarbeitet werden kann. Bedeutend ist der Umkehrschluss, bei dem achtundneunzig Prozent der aufzunehmenden Informationen vom Menschen eben nicht wahrgenommen und verarbeitet werden können, selbst wenn bereits die Relevanz und Präzision der Informationen sichergestellt wurde [7, 10].

Die Limitierung der kognitiven Verarbeitungsprozesse des Menschen wurde bereits durch kognitionspsychologische Erkenntnisse im vorherigen Jahrhundert nachgewiesen [7, 25, 26]. Abbildung 2.1 illustriert diese kognitiven Verarbeitungsprozesslimitierungen anhand dem von Schroder, Driver und Streufert aufgestellten Modell, welches das Verhältnis zwischen Informationsüberflutung und der daraus resultierenden Entscheidungsqualität zeigt [7, 27, 28, 29].

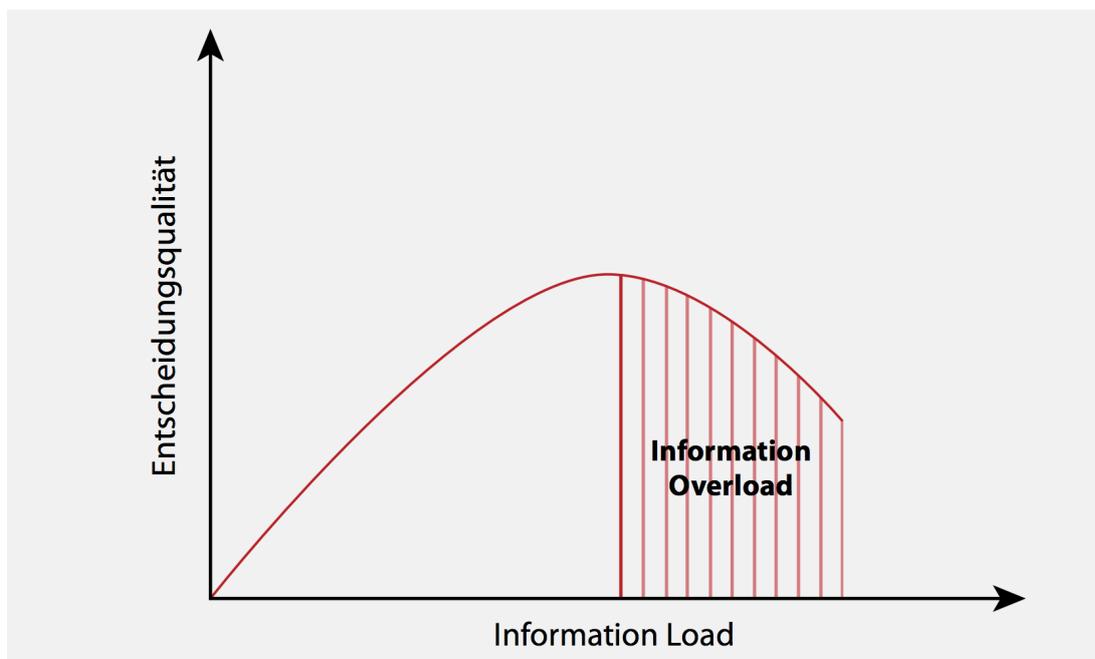


Abbildung 2.1: Das Verhältnis zwischen Informationsüberflutung und der daraus resultierenden Entscheidungsqualität [7].

Abbildung 2.1 zeigt die zunehmende Verbesserung der Entscheidungsqualität mit zunehmender Informationsmenge, bis zum Erreichen einer bestimmten Schwelle. Beim Erreichen dieser Schwelle nimmt die Qualität der Entscheidung anhand der Informationen stark ab. Eine Vielzahl empirischer Studien [7, 30, 31, 32, 33, 34] unterschiedlicher Forschungsdisziplinen bestätigen

diese Verschlechterung der Entscheidungsqualität durch Informationsüberflutung. Der Fortschritt in der Digitalisierung durch zunehmende Entwicklung von Informations- und Kommunikationstechnologie sorgt auch weiterhin für einen anhaltenden quantitativen Anstieg an Daten und gilt somit für die Ursache der Informationsüberflutung [7, 35].

Grundsätzlich sind alle Menschen von kognitiven Einschränkungen hinsichtlich der Informationsaufnahme beschränkt. Die Höhe der Schwelle zur Erreichung des Phänomens der Informationsüberflutung ist allerdings individuell und abhängig von definierten Ursachen. Beispielhafte Ursachen, die zur Informationsüberflutung führen, zeigt Abbildung 2.2.

Ursachen des Information Overload	Beispiele für Einflüsse auf die Variablen des Information Overload
Persönliche Faktoren	Begrenzungen der menschlichen Informationsverarbeitungskapazität, Motivation, persönliche Situation
Informationscharakteristika	Quantität, Häufigkeit, Intensität und Qualität der Information(en)
Aufgaben- und Prozessparameter	Komplexität der Aufgaben, Zeitdruck, Unterbrechungen bei komplexen Aufgaben, interdisziplinäre Aufgaben
Organisationales Design	Informationssuche des Endnutzers vs. des Experten, Gruppenheterogenität, Informationsansammlung zur „Machtdemonstration“
Informationstechnologie	Verschiedene Verteilungskanäle für gleichen Inhalt, große Speicherkapazitäten, geringe Vervielfältigungskosten

Abbildung 2.2: Exemplarische Ursachen, die zur Informationsüberflutung führen [7].

Ableitend aus den exemplarisch genannten Ursachen für eine Informationsüberflutung in Abbildung 2.2 ist der Bereich der Bildung einer von vielen besonders anfälligen Bereichen für eine Informationsüberflutung. Handelt es sich beispielsweise um junge Schülerinnen und Schüler sind die persönlichen Faktoren meist deutlich schwächer ausgeprägt als die von Erwachsenen. So ist die Informationsverarbeitungskapazität von Schülerinnen und Schüler in der Regel deutlich niedriger. Auch eine Motivation zum Lernen ist oft nicht so stark ausgeprägt wie beispielhaft in der Erwachsenenbildung. Die Qualität von Informationen die Lernende zum ersten Mal erhalten, können zusätzlich oft nicht eingeschätzt werden. Hier lassen sich noch weitere Faktoren finden.

Kapitel 1.2 beschreibt, dass zur Konkretisierung des konzeptuellen Modells die vom Hasso-Plattner-Institut entwickelte Schul-Cloud genutzt werden soll. Dabei handelt es sich um den genannten Bereich der Bildung. Präzisiert handelt es sich um einen Bereich der Bildung, der durch digitale Medien unterstützt wird. Die Integration digitaler Technologien in der Domäne der Bildung wird dabei häufig als E-Learning bezeichnet. Eine einheitliche Definition für die Begrifflichkeit des E-Learnings existiert in der Literatur allerdings nicht. In dieser Forschungsarbeit beschreibt der Begriff des E-Learnings zur Präzisierung des konzeptuellen Modells durch die Schul-Cloud die E-Learning-Plattform. Eine E-Learning-Plattform stellt hierbei einen virtuellen Bereich dar, der zur Aneignung von Bildung im digitalen Raum dient. Die Aneignung entsprechender Bildung erfolgt hierbei über unterschiedliche, speziell digitalisierte Lehr- und Lernmaterialien. Weiterhin ermöglicht eine E-Learning-Plattform in einem solchen virtuellen Raum auch die Nutzung administrative Fähigkeiten. Beispielsweise die Verwaltung von Lehrveranstaltungen, zu der sich Teilnehmer online registrierten können oder spezielle Kommunikationsmöglichkeiten zwischen den Teilnehmern oder zwischen Lehrpersonal und Lernenden durch Messaging. Die Vorteile des E-Learning liegen auf der Hand. Besonders die verschiedenen Lernmethoden und Anwendbarkeit verschiedener Lehrstile sind von besonderer Relevanz [36].

Insbesondere Schulen können von der Domäne des E-Learnings profitieren. Mit dem Zugang zum World Wide Web über Computer, Tablets und Smartphones wird der Zugang zu Informationen revolutioniert durch einen unbegrenzten Zugang zu Informationen. Gleichzeitig verlagert sich die Aufgabe des Lehrpersonals zunehmend vom Wissensvermittler zum Lernbegleiter. Die besondere Tätigkeit besteht darin, die Informationsflut für Schülerinnen und Schüler zu vermeiden [37]. Die Digitalisierung verändert damit nicht nur das Verhältnis von Lehrerinnen und Lehrern zur Wissensvermittlung, sondern auch das Verhältnis von Schülerinnen und Schüler zum Wissensempfang. Eine interaktivere Einbindung von Schülerinnen und Schülern wird durch E-Learning möglich. So werden Schüler zu Bildungspartnern anstatt zu reinen Wissensaufnahmeobjekten [38].

Als eine solche E-Learning-Plattform kann die Schul-Cloud des Hasso-Plattner-

Instituts bezeichnet werden. Die Schul-Cloud ist ein Pilotprojekt zur Modernisierung des Schulunterrichts mit Fokussierung auf Verbesserung der Lern- und Lehrinfrastruktur. Die Nutzung digitaler Medien und Lehrmethoden stellt heute viele Schulen vor große Herausforderungen, sowohl finanzieller als auch administrativer Art. Veraltete Hardware und die Separierung von digitalen und traditionellen Lehrangeboten durch räumliche Trennung in Klassen- und Computerräume sowie das fehlende Personal für eine fachgerechte Wartung der Systeme in Computerräumen sind nur einige wenige dieser Herausforderungen. Gleichzeitig ist es für Schüler unabdingbar moderne Lehr- und Lerninfrastruktur kennenzulernen als auch zu nutzen, um den Jugendlichen so eine optimale Vorbereitung für die digitale Gesellschaft zu ermöglichen, in die sie hereinwachsen [38].

Obwohl der Umgang mit digitalen Medien für die heutige Generation lebenswichtig ist, spielt der Umgang mit diesen in Schulen immer noch keine Rolle. Der Grund hierfür ist die fehlende Infrastruktur in Schulen, wie beispielsweise WLAN, Breitbandanbindung und Smart Devices. Schülerinnen und Schüler müssen heute immer noch separate Rechnerkabinette besuchen, um einen Zugang zur Informationswelt zu erhalten. Diese wiederum sind selten auf dem neusten Stand und professionell gewartet. Rechnerkabinette sind hingegen schlecht, denn Medien sollen in allen Fächern genutzt werden können, wie ein Schulbuch. Digitale Lerninhalte bieten im Gegensatz zu Schulbüchern denn Vorteil, dass sie individueller auf die besonderen Fähigkeiten des Schülers eingehen können, und zwar viel individueller als ein Lehrer mit X Kindern es hinbekommen könnte. Mit der Schul-Cloud verbessert sich die Möglichkeit der Zusammenarbeit, da man sich Nachrichten senden kann und es kann von verschiedenen Orten aus gearbeitet werden. Die Schul-Cloud besitzt bereits jetzt eine fünfstellige Zahl von digitalen Lerninhalten in der Cloud, von Landkarten bis hin zu hochkomplexen interaktiven Lernsystemen. Dem Lehrer steht durch die Schul-Cloud alles zur Verfügung, dient aber als Entscheider, was für den Unterricht genutzt wird. Die Schüler können die Materialien nicht nur im Unterricht nutzen, sondern auch zu Hause und in den Ferien. Diese können außerdem sicher abgelegt werden, sodass der Schüler sich sicher sein kann, dass seine personenbezogenen Daten die Schul-Cloud nicht verlassen [39].

Mit der Schul-Cloud wurde ein kostengünstiger und zukunftssicherer Ansatz entwickelt, um Jugendlichen den Zugriff auf modernste Lehr- und Lernmöglichkeiten zu ermöglichen. Durch diesen Ansatz werden Lehrer sowie private und institutionelle Inhaltsanbieter motiviert, neue digitale Lehrangebote zu entwickeln. Aber auch Schülern ist es mit der Schul-Cloud möglich, zu Lehrenden zu werden durch die Erstellung eigener digitaler Materialien für die Schul-Cloud. Diese Materialien können beispielsweise für die Nachhilfe anderer Schüler zur Verfügung gestellt werden. Somit sind Bildungsinhalte nicht weiter auf abgeschlossene Silos und Lehrbücher beschränkt, sondern durch die Cloud zentral und von überall verfügbar [38].

Die Schulen selbst benötigen hierbei nur noch einen Zugang zum Internet sowie webfähige Endgeräte, um den Schülern Zugang zu den modernen Lehrmaterialien zu ermöglichen. Auf den Endgeräten muss dabei keine Software installiert werden. Schulunabhängige Experten kümmern sich gleichwohl um die Aktualisierung von Hard- und Software der Schul-Cloud. Lehrkräfte können sich darauf besinnen, die Inhalte, welche über die Schul-Cloud zur Verfügung gestellt werden, passgenau in ihren Unterricht zu integrieren [38].

### 3 Analyse

In Abschnitt zwei Kontextbetrachtung wurde der Rahmen der zu bewältigenden Problemstellung skizziert und anhand der konkreten Anwendung der Schul-Cloud aus der Domäne des E-Learnings spezialisiert. Die Inhalte der folgenden Kapitel sollen sich der Beantwortung der Forschungsfrage annähern, durch die Beantwortung entsprechend-deklarerter Leitfragen. Die in den nachfolgenden Kapiteln erfassten Erkenntnisse dienen dem anschließenden Abschnitt als Initial zur Beantwortung weiterer Leitfragen und schließlich der als Endresultat zu beantwortenden Forschungsfrage.

Die folgenden Leitfragen soll dieser Abschnitt in den nachfolgenden Kapiteln beantworten:

1. Was wird unter einer effektiven Informationsvisualisierung verstanden und wie kann eine solche effektive Informationsvisualisierung gewährleistet werden?
2. Wie wird die Informationsvisualisierung in heutigen digitalen Informationssystemen realisiert und welche Problematiken hinsichtlich des Effektivitätskriteriums existieren derzeit bei der Informationsvisualisierung in heutigen digitalen Informationssystemen?
3. Welche Problematiken existieren bei der dynamischen Informationsvisualisierung durch vollständige, automatisierte Generierung von Benutzeroberflächen und warum ist der Einsatz dieses Verfahrens ungeeignet zur Beantwortung der Forschungsfrage?

Zur Beantwortung der Leitfragen gliedert sich dieser Abschnitt, orientiert an den für diesen Abschnitt definierten Leitfragen, in auf die Leitfragen abgestimmter chronologischer Ordnung in die drei Kapitel: 3.1 Visualisierung, 3.2

Visualisierungsprozess in digitalen Informationssystemen und 3.3 Dynamische Informationsvisualisierung durch vollständige automatisierte Generierung grafischer Benutzeroberflächen.

### 3.1 Visualisierung

Kapitel 2 erläuterte, wie in immer mehr und mehr unterschiedlichsten Anwendungsgebieten Daten gesammelt werden. Gleichzeitig lässt sich daraus ableiten, dass die gesammelten Daten erst an Wert für den Menschen gewinnen, wenn die Bedeutung dieser Daten verstanden wird. Eine Methode, um das Verständnis von Daten für das kognitive System des Menschen zu gewährleisten, bildet die Visualisierung entsprechender Daten. Ein einzelnes Bild kann im besten Fall, Antworten eine ganze Reihe von Fragen liefern. Damit gewinnt die Visualisierung in allen Domänen von der Forschung bis hin in die Industrie aufgrund des quantitativ steigenden Datenvolumens in der Welt an hoher Bedeutung [13, 40, 41]. Die visuelle Repräsentation von Daten kann entscheidend dazu beitragen, das Phänomen der Informationsüberflutung zu verringern oder gar zu vermeiden.

Die Literatur definiert die Begrifflichkeit der Visualisierung nicht einheitlich. Somit wird der Begriff *Visualisierung* in dieser Forschungsarbeit unabhängig davon verwendet, ob es sich bei einer Visualisierung um eine wissenschaftlich-technische, eine Informationsvisualisierung oder um eine ästhetische Visualisierung durch spezielle Benutzeroberflächenelemente handelt. Diese Arbeit orientiert sich beim Begriff der Visualisierung an der Definition des Autors *Card* aus dem Jahre 1990: "Die Verwendung Computerunterstützten, interaktiven, visuellen Repräsentationen von Daten zur Verstärkung der Kognition"[42].

Zur Beantwortung der Teilfrage "*Was wird unter einer effektiven Visualisierung verstanden?*" und damit im weiteren zur Beantwortung der Forschungsfrage, dienen die anknüpfenden Kapitel. Diese zeigen wiederum, wie visuelle Informationen vom kognitiven System des Menschen verarbeitet werden, welche Bedeutung und Zielsetzung die Visualisierung hat und schließlich, was eine effektive Visualisierung ausmacht.

### 3.1.1 Prozess der Wahrnehmung

Die Wahrnehmungspsychologie identifiziert verschiedene Sinne, die es dem Menschen ermöglichen, Informationen aufzunehmen. Die in unserem menschlichen Lebensraum täglich auf uns einströmenden Informationen werden dabei durch die identifizierten Sinne in erster Instanz erfasst und anschließend durch unser Gehirn verarbeitet. Dabei sind die auf uns einwirkenden Informationen zu jedem Zeitpunkt immens. Dies veranlasst unser kognitives System dazu, äußerst selektiv Informationen an das Gehirn weiterzuleiten, die im Anschluss unser Bewusstsein erreichen. Ein bewusstes Empfinden aller von der Sinneswahrnehmung erfassten Informationen ist hierbei nicht möglich [43].

Als eines der am weitesten entwickelten und gleichzeitig komplexesten Sinnesorgane gilt das Auge der visuellen Wahrnehmung [44]. Die visuelle Wahrnehmung ist für die Aufnahme visueller Reize zuständig, welche ausschließlich über das Sinnesorgan *Auge* möglich ist [45, 46]. Neben der visuellen Wahrnehmung kann der Mensch zusätzliche Reize über weitere Wahrnehmungen erfassen. Abbildung 3.1 illustriert diese. In der Literatur werden dem Menschen von einigen Experten weitere Sinne zugeordnet.



Abbildung 3.1: Die von der Wahrnehmungspsychologie klassischen dem Menschen zugeordneten Sinne neben der visuellen Wahrnehmung (eigene Darstellung).

Über das Sinnesorgan *Ohr* können Schallinformationen aufgenommen werden. In der Wahrnehmungspsychologie wird von der auditiven Wahrnehmung gesprochen. Beim Schall handelt es sich dabei um konkrete Instanzen bestehend aus Tönen, Geräuschen oder Klängen. Mit dem Sinnesorgan *Na-*

se hingegen können Geruchsinformationen wahrgenommen werden. Hierbei wird von der olfaktorischen Wahrnehmung gesprochen. Bei der gustatorischen Wahrnehmung handelt es sich um die Stimulation der Geschmacksknospen im Sinnesorgan *Mund*, die uns Menschen das Schmecken ermöglichen. Das Empfinden von Berührungen wird durch die Stimulation von Rezeptoren des Sinnesorgan *Haut* ermöglicht. In der Wahrnehmungspsychologie wird dies als taktile Wahrnehmung bezeichnet [46, 47, 48, 49, 50].

Selbst wenn der Mensch alle in der Abbildung 3.1 aufgeführten Sinne in Kombination einsetzen würde, kann dennoch nicht die gleiche Informationsaufnahmekapazität erreicht werden, wie es mit der visuellen Wahrnehmung möglich wäre.

Der Prozess der visuellen Wahrnehmung wird in Abbildung 3.2 skizziert. Er zeigt den Ablauf von der visuellen Aufnahme von Informationen bis zur Verarbeitung dieser Informationen durch das zentrale Nervensystem und insbesondere dem Gehirn.



Abbildung 3.2: Der Prozess der visuellen Wahrnehmung von der visuellen Aufnahme von Informationen bis zur Verarbeitung dieser Informationen durch das zentrale Nervensystem (eigene Darstellung).

Die sogenannte Photorezeption ermöglicht das Erfassen von Informationen aus der Umwelt durch Rezeptoren der Netzhaut des Auges. Hierbei werden die Rezeptoren durch einfallendes Licht aktiviert und so die Information erfasst. Die erfassten Informationen werden anschließend in elektrische Nervenimpulse umgewandelt, die im Gehirn entsprechend zusammengetragen werden. Eine Weiterleitung der Informationen zum visuellen Cortex - auch Sehirinde genannt - wird durchgeführt und dort zu einer einheitlich schlüssigen Wahrnehmung verarbeitet. Dem sogenannten Perzept [44].

### 3.1.2 Bedeutsamkeit und Zielsetzungen von Visualisierungen

Vier Milliarden Neuronen, siebzig Prozent aller Rezeptoren und über vierzig Prozent des Cortex beim Menschen sind bei der Aufnahme von Informationen durch die visuelle Wahrnehmung über das Auge als Sinnesorgan beteiligt. Dies macht den im vergangenen Kapitel 3.1.1 beschriebenen Prozess der visuellen Wahrnehmung zu einem primären Werkzeug zur Lösung des in Kapitel 2 beschriebenen Phänomens der Informationsüberflutung. Damit einhergehend wird die Visualisierung zum besten Medium, relevante Informations-Ausschnitte in hoher Bandbreite zum Gehirn und in das Bewusstsein zu fördern [13].

Die Bedeutsamkeit von Visualisierung als Informationsträger zum Bewusstsein kann zusätzlich verdeutlicht werden, durch die Betrachtung der Determinanten, die zu einer Informationsüberflutung führen. Abbildung 3.3 skizziert diese.

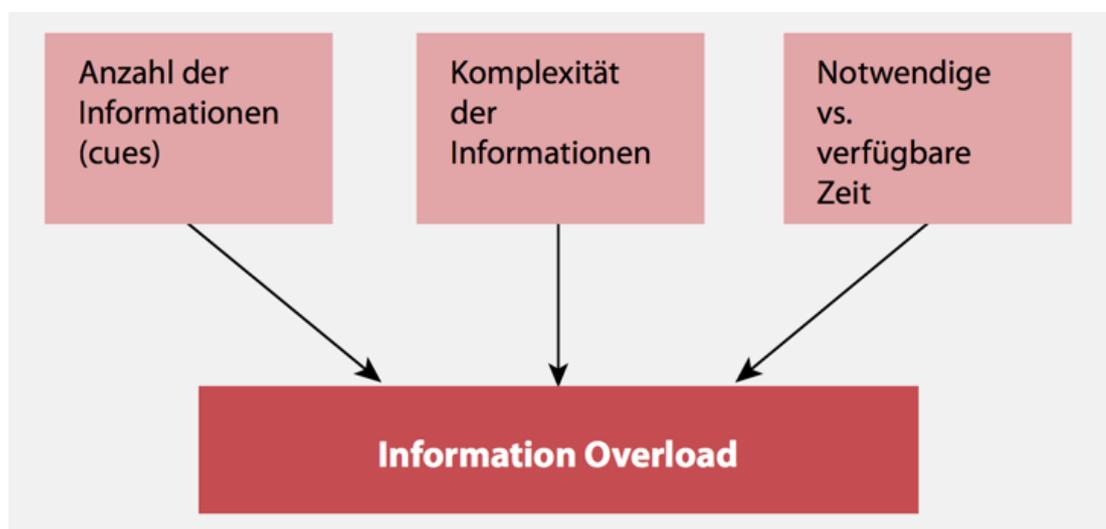


Abbildung 3.3: Die zum Phänomen der Informationsüberflutung führenden Determinanten [7].

Für die drei Hauptdeterminanten Informationsanzahl, Informationskomplexität sowie die Zeit [7, 8, 9] lassen sich die folgenden zu einer informatorischen Überlastung führenden Hauptursachen ableiten [7]:

- a) Die Information selbst
- b) Die Person, welche die Informationen empfängt, verarbeitet oder kommuniziert
- c) Den Aufgaben oder Prozessen, die von einer Person erfüllt werden müssen
- d) Die zur Verfügung stehenden Informationsverarbeitungstechnologien bzw. deren Art des Einsatzes

Dabei entsteht eine Informationsüberlastung gewöhnlich nicht durch einen einzelnen der aufgeführten Ursachen, sondern durch die Kombination der genannten [7, 51]. Die Vermeidung oder zumindest Reduktion der informativischen Überlastung kann durch die Zielsetzung der Visualisierung erreicht werden.

Auf oberster Ebene kann die Zielsetzung der Visualisierung beschrieben werden, als ein Verständnissgewinn des Betrachters für (große) Datenmengen sowie dem Erkennen von Mustern und Strukturen in diesen Daten. Dies wird erreicht durch die Transformation von symbolischen Daten, Informationen oder Wissen zu grafische Repräsentationen, durch die Abbildung von Datenattributen zu Visualisierungselementattributen wie beispielsweise Position, Größe, Form oder Farbe [13, 40, 41, 52].

Auf einer spezifischeren Ebene lassen sich die folgenden sechs elementaren Aufgaben ableiten, welche die Zielsetzung der Visualisierung definieren [13, 52]:

- a) Exploration, Analyse, Interpretation, Entscheidungsfindung sowie Erklärung und Kommunikation der Daten ermöglichen
- b) Aufdeckung verschiedener Strukturen in Daten ermöglichen
- c) Kohärenz (kompakte Repräsentation) großer Datensätze ermöglichen
- d) Daten- und Informationsaufbereitung von verschiedenen Gesichtspunkten ermöglichen

- e) Daten- und Informationsaufbereitung auf verschiedenen Detailstufen ermöglichen
- f) Bei visuellen Vergleichen unterstützen

Verschiedene Studien zeigen, dass sich die Visualisierung von Daten positiv auf die menschliche Gedächtnisleistung und Informationsaufnahme auswirkt. Zusammengefasst besteht das Ziel der Visualisierung also darin, abstrakte Informationen aus beliebigen, voranging textbasierten Informationssystemen strukturell aufzubereiten und grafisch darzustellen, um den kognitiven Zugang zu elektronisch gespeicherten Informationen zu erleichtern. Grafische Benutzeroberflächen als heutiger de facto Standard für die Mensch-Computer-Interaktion ermöglichen die Umsetzung der Zielsetzung von Visualisierungen [52].

### 3.1.3 Qualitätskriterien an Visualisierungen

Nachdem die Bedeutsamkeit und Zielsetzung von Visualisierungen im vergangenen Kapitel 3.1.2 erläutert wurde, stellt sich nun die Frage, welche Qualitätskriterien eine Visualisierung erfüllen muss, um die kognitive Gedächtnisleistung des Betrachters zur Aufnahme entsprechender Informationen zu steigern.

Primär lässt sich die Qualität einer Visualisierung durch den Wirkungsgrad definieren, durch welche die grafische Repräsentation die kommunikative Zielsetzung erreicht. Als Hauptqualitätskriterien lassen sich Expressivität, Effektivität und Angemessenheit der Visualisierung definieren. Die Expressivität konventionalisiert, dass alle zu visualisierenden Daten tatsächlich in die Visualisierung einfließen. Die Integration irrelevanter Daten oder den Verzicht von relevanten Daten wird ausgelassen. Dadurch wird eine unverfälschte Wiedergabe der Daten gewährleistet. Bei der Effektivität hingegen handelt es sich um ein Kriterium, das den Grad der intuitiven Vermittlung von Informationen gegenüber dem Betrachter angibt bezogen auf die gewählte Darstellungsform, welche die Informationen enthält. Das Kriterium der Angemessenheit stellt sicher, dass sich die Relation zwischen Aufwand und Nutzen im Gleichgewicht hält.

Beim Aufwand im Kontext der Angemessenheit handelt es sich dabei nicht ausschließlich um die Kosten der Technologie. Hier seien besonders auch die menschlichen Faktoren erwähnt. Exemplarisch hierfür sind zeitliche, kognitive, physische oder psychische Aufwendungen, den ein Betrachter investieren muss, um ein Verständnis für die Visualisierung zu erhalten. Eine Visualisierung, die das Kriterium der Angemessenheit dabei nicht erfüllt, widerspricht damit gleichzeitig auch dem Qualitätskriterium der Effektivität [53].

Während die Expressivität von Daten ein einfacher zu erfüllendes Qualitätskriterium ist, lässt sich die Effektivität und Angemessenheit von Visualisierungen nur schwer beschreiben. Der Grund hierfür ist die nicht einfache Messbarkeit des Effektivitätskriteriums und Angemessenheitskriterium. Anders als physikalische Größen lassen sich die beiden Kriterien nur durch direkte Tests beim Betrachter messen. Solche Tests werden auch Usability-Tests genannt. Insbesondere die Anwendung des Effektivitätskriteriums innerhalb dieser Forschungsarbeit beruht daher auf der Grundlage von in der Vergangenheit getestete Szenarien und deren Ergebnisse, die von verschiedenen Autoren als Leitlinien niedergeschrieben wurden [42]. Abbildung 3.4 zeigt die Umsetzung einer solchen Leitrichtlinie zur Effektivität der Visualisierung exemplarisch.

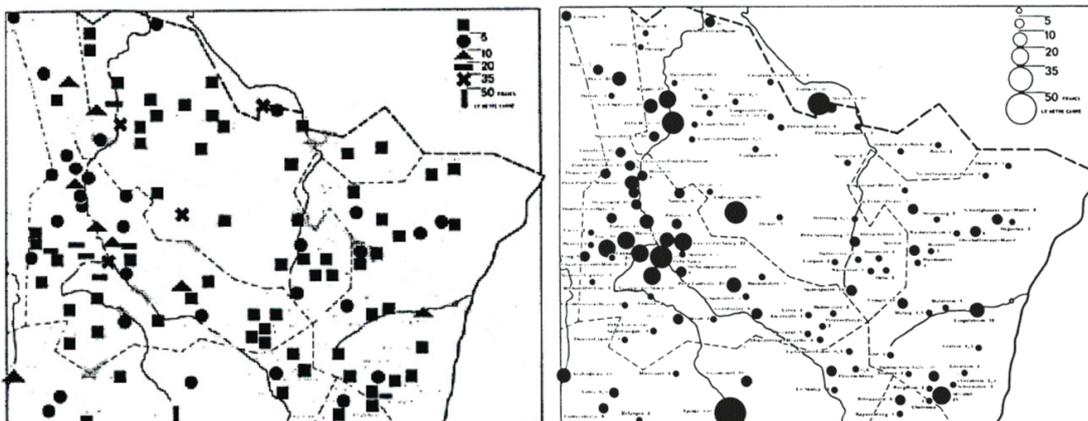


Abbildung 3.4: Exemplarische Visualisierung von Grundstückspreisen in einer Region zur Verdeutlichung der Umsetzung einer Richtlinie die das Effektivitätskriterium zur Bestimmung der Qualität von Visualisierungen erfüllt [42].

Beide Darstellungen der Abbildung 3.4 visualisieren jeweils die Grundstücks-

preise in einer Region. Die Darstellung auf der rechten Seite ist dabei als effektiver zu betrachten. Mit den unterschiedlichen Größenverhältnissen der Kreise wird beim Betrachter sofort die richtige Assoziation geweckt. Der Betrachter kann unmittelbar assoziieren, dass ein großer Kreis wohl einen teureren Grundstückspreis zur Folge hat, als ein kleinerer Kreis. Dadurch kann der Betrachter auch die Frage beantworten, in welcher Region die Grundstückspreise am teuersten sind.

Für die genannten Hauptqualitätskriterien Expressivität, Effektivität und Angemessenheit lassen sich verschiedene Einflussfaktoren ableiten, die dementsprechend direkten Einfluss auf die Qualität der Visualisierung haben. Entsprechend werden diese Faktoren genutzt, um eine effektive Visualisierungsform zu finden oder eine neue Visualisierungsform zu erzeugen [53]:

**Art und Struktur von Daten** Typ der Daten sowie Dimension und Struktur des Beobachtungsbereichs

**Zielerreichung des Anwenders** Die Zielerreichung ist maßgeblich verantwortlich dafür, welche Datenattribute Teil der Visualisierung werden, und wie diese Datenattribute dem Visualisierungselement hinsichtlich der Position zugeordnet werden. Der Faktor der Zielerreichung wird dabei durch den Anwender selbst gesteuert. So wird das exemplarische Ziel, einen Überblick über einen Datensatz zu erhalten zu einem anderen Visualisierungsergebnis führen, als das Ziel spezifische Informationen des Datensatzes detailliert analysieren zu wollen.

**Vorwissen des Anwenders** Das Vorwissen des Anwenders gilt ebenfalls als ein wichtiger Indikator zur Auswahl geeigneter Visualisierungselemente. Mit dem Wissen eines Mathematikers beispielsweise können mathematische Sachverhalte durch die Anwendung von mathematischer Notation deutlich kompakter visualisiert werden. Mit dem Wissen eines Schülers der Klasse fünf muss der gleiche Sachverhalt aber womöglich ohne mathematische Notation

visualisiert werden, da der Schüler nicht über entsprechendes Wissen mathematischer Notationen verfügt. Der Faktor des Anwendervorwissens steht in enger Harmonie mit der Zielerreichung des Anwenders.

**Visuelle Fähigkeiten und Vorlieben des Anwenders** Als Mensch besitzen wir verschiedene visuelle Fähigkeiten, die eine Äquivalenz zu den Funktionen der visuellen Wahrnehmung bilden. Grundfunktionen der visuellen Wahrnehmung sind beispielsweise das Erkennen und Differenzieren von Größe und Form, die Identifizierung einer Formkonstanz, die Wahrnehmung der Lage im Raum, die Wahrnehmung räumlicher Beziehungen, visuelle Mengenerfassung, Farbenerkennung und -zuordnung, Helligkeit oder Sehschärfe. Aufgrund von kognitiven oder körperlichen Einschränkungen besitzt aber nicht jedes menschliche Individuum die gleichen visuellen Fähigkeiten. Im Kontext der visuellen Wahrnehmung können bei Individuen Wahrnehmungsstörungen auftreten. Exemplarisch für eine solche Wahrnehmungsstörung ist die Achromatopsie (Farbenblindheit) aus der Kategorie der Wahrnehmungsstörungen. Diese verursacht, dass das störungsbelastende Individuum keine Farben, sondern nur Kontraste erkennen kann. Die visuellen Fähigkeiten sind somit nicht unerheblich für die Auswahl eines effektiven Visualisierungselements. Von nicht ganz so großer Bedeutung wie die visuellen Fähigkeiten sind die Vorlieben des Anwenders. Dennoch können diese einen erheblichen Einfluss auf die kognitive Gedächtnisfähigkeit bei der Informationsaufnahme haben. Anwendervorlieben können beispielsweise spezielle Farbkombinationen oder Anordnungen der Datenattribute sein.

**Konventionen des Anwendungskontexts** Beim Faktor der Konventionen im Anwendungskontext handelt es sich um spezifische Merkmale, welche die Anwendung im Umgang mit der Anwendung gewohnt sind. Dies können beispielsweise spezielle Positionsanordnungen der einzelnen Datenattribute im Visualisierungselement sein oder die Verwendung spezieller Symbole.

**Charakterisika des Darstellungsmediums** Die Darstellung verschiedener Visualisierungselemente wird heute nicht mehr nur auf einer begrenzten Anzahl an Endgeräten gerendert. Der durchschnittliche Mensch nutzt heute täglich eine Vielzahl unterschiedlicher Endgeräte. Diese Endgeräte besitzen alle eine differierende Kombination aus Eigenschaften. Beispielsweise kann ein mobiles Endgerät wie ein Smartphone ein vollständig anderes Farbspektrum zur Darstellung der Visualisierung besitzen als der stationäre Monitor, der am Desktop-PC angeschlossen wird. Weiterhin kann dieser stationäre Desktop-PC eine vollständig andere Rechenleistung und Bildschirmauflösung aufweisen als die des Smartphones oder eines smarten Fernsehgeräts. Die Reaktionsfreudigkeit der einzelnen Visualisierungselemente beziehungsweise Auswahlalgorithmen für Visualisierungselemente, auf unterschiedliche Endgerätkonfigurationen hinsichtlich der Mediumseigenschaften wird in der heutigen Gesellschaft eine besondere Bedeutung zuteil.

## 3.2 Visualisierungsprozess in digitalen Informationssystemen

Die in Kapitel 1.1 definierte Forschungsfrage impliziert, dass in den meisten Informationssystemen der Visualisierungsprozess nur Ergebnisse in Form von Daten zugeordneten Visualisierungen liefert, wenn die Daten selbst zur Entwicklungszeit bekannt sind. Zur Beantwortung der Forschungsfrage stellt sich die Teilfrage, wie in klassischen Informationssystemen der Prozess von den Daten zur Visualisierung realisiert wird.

### 3.2.1 Visualisierungspipeline

In abstrakter Weise kann die typische Vorgehensweise zur Erzeugung einer Visualisierung anhand der in Abbildung 3.5 skizzierten Visualisierungspipeline beschrieben werden.

Der Prozess besteht wie die Abbildung 3.5 zeigt, aus drei Vorgängen: Filterung, Mapping und Rendering. Jeder dieser drei Vorgänge transformiert eine

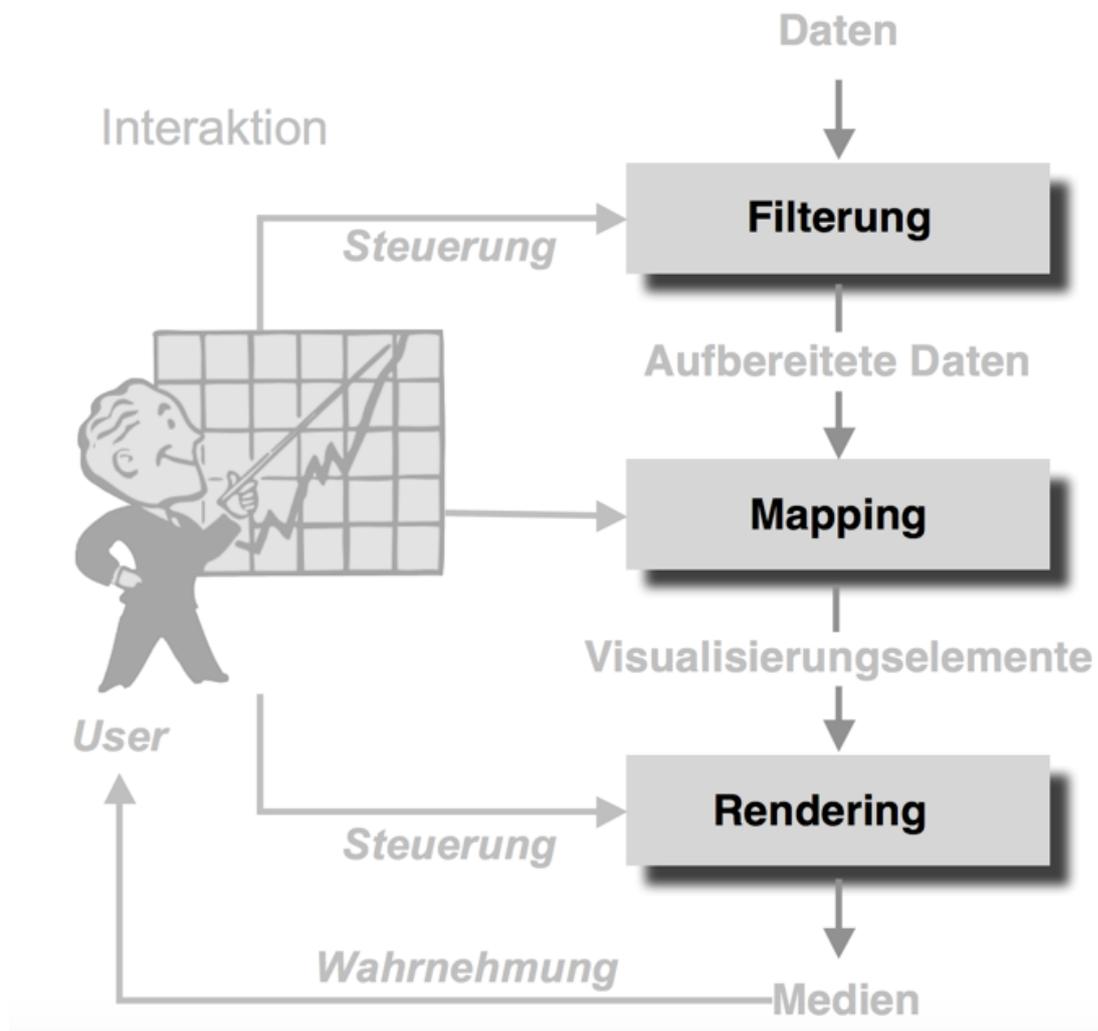


Abbildung 3.5: Die verschiedenen Vorgänge und Zwischenergebnisse der Visualisierungspipeline, welche die typische Vorgehensweise zur Erzeugung einer Visualisierung für vorgegebene Daten und ein vorgegebenes Visualisierungsziel repräsentiert [42].

Eingabe zu einem innerhalb des Vorgangs produzierten Ergebnis. Diese werden als Zwischenergebnisse bezeichnet. Daten stellen die Eingabe für den Vorgang der Filterung dar. Gleichzeitig sind die Daten Ausgangspunkt für die Visualisierung. In diesem Stadium wird von Rohdaten gesprochen. Sie stammen aus unterschiedlichsten Quellen. Beispielsweise aus manuellen Eingaben, Datenbanken oder Programmierschnittstellen (API). Der Vorgang der Filterung bereitet diese Rohdaten auf. Eine Aufbereitung von Daten kann beispielsweise durch Interpolation (Vervollständigung), Reduktion, Selektion, Projektion oder Ergänzung durch weitere Datensätze erfolgen. Das produzierte Ergebnis dieses Vorgangs sind die aufbereiteten Daten, die gleichzeitig die Eingabe für den Vorgang des Mappings darstellen. Die Aufgabe des Mapping ist die Zuordnung von Datenattributen zu entsprechenden Visualisierungselementattributen. Das Ergebnis dieses Vorgangs sind die mit Daten unteretzten Visualisierungselemente. Der Vorgang des Renderings sorgt abschließend für die Darstellung der Visualisierungselemente auf dem Anzeigegerät (Medium) des Betrachters. Der Betrachter kann die Darstellung nun durch seine visuelle Wahrnehmung interpretieren. Einfluss auf die Visualisierungspipeline kann vom Anwender durch bewusste Interaktionssteuerung der Vorgänge Filterung und Rendering genommen werden. Getriggert wird dies in der Regel durch Änderung des Visualisierungsziels [42].

### 3.2.2 Statische Informationsvisualisierung

Nachfolgend wird der in Kapitel 3.2.1 abstrakt skizzierte Visualisierungsprozess auf Basis der Abbildung 3.6 im Detail aufgezeigt. Er bildet die Grundlage für die meisten in Softwaresystemen verwendeten Visualisierungsengeines.

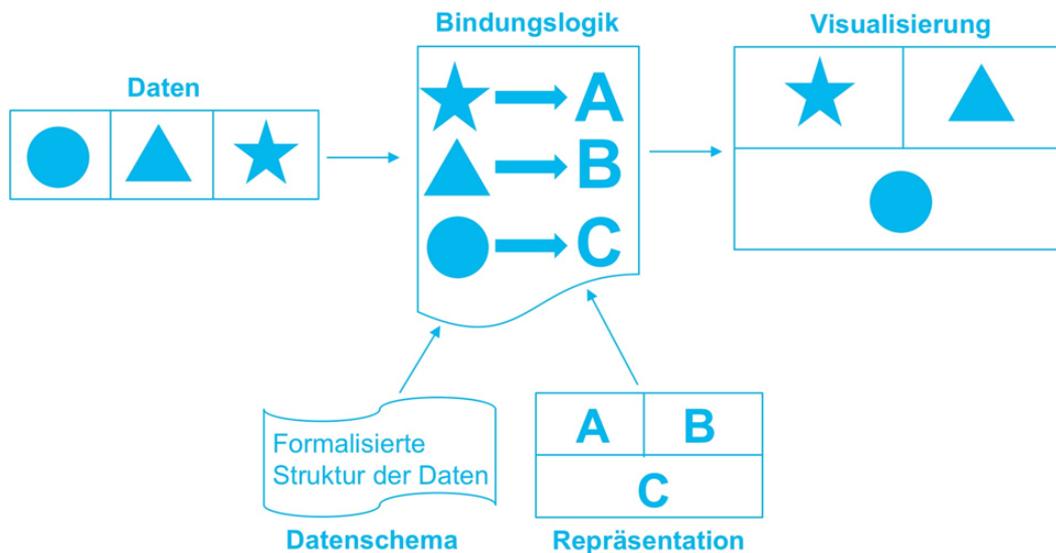


Abbildung 3.6: Typischer Visualisierungsprozess, wie er für die meisten Softwaresysteme die Grundlage der Visualisierungseingabe bildet (eigene Darstellung).

Das Szenario in Abbildung 3.6 zeigt einen Datensatz mit drei Datenattributen, dargestellt durch den Kreis, das Rechteck und den Stern. Dem Softwareentwickler ist zur Entwicklungszeit der Anwendung sowohl die Art als auch die Struktur des Datensatzes bekannt. Weiterhin kennt der Softwareentwickler die Bedeutung (Semantik) der zu visualisierenden Daten. Gleichzeitig ist dem Softwareentwickler auch der Kontext (Pragmatik) bekannt, in welchem die Daten visualisiert werden (Anwendungskontext). Durch die zur Entwicklungszeit bekannte Art und Struktur des Datensatzes kann der Softwareentwickler ein Datenschema formalisieren. Ein solches Datenschema spezifiziert, welche Art und Struktur von Daten innerhalb der Anwendung verarbeitet werden kann - insbesondere durch die Visualisierungseingabe, also jene Komponente der Anwendung, die eine Zuordnung von Datenattributen zu Visualisierungselementattributen durchführt - und wie diese Daten in Beziehung zueinanderstehen. Dadurch, dass der Softwareentwickler die Semantik und Pragmatik der Daten kennt, kann eine im Sinne der in Kapitel 3.1.3 beschriebenen Qualitätskriterien, effektives Visualisierungselement für die Daten entwickelt werden. Die

in Abbildung 3.6 aufgezeigte Bindungslogik führt dabei das Mapping der Datenattribute zu den Visualisierungselementattributen durch. Auch diese muss durch den Softwareentwickler zur Entwicklungszeit formalisiert werden. Durch das Wissen über die Daten und das Wissen der entwickelten Visualisierung ist die Formalisierung der Bindungslogik ein einfaches. Durch den anschließenden Renderingprozess erhält der Betrachter eine den Qualitätskriterien aus Kapitel 3.1.3 entsprechende, effektive Visualisierung.

### 3.2.3 Voraussetzungen statischer Informationsvisualisierung

Betrachtet man den im Detail beschriebenen Visualisierungsprozess aus Kapitel 3.2.1 und 3.2.2 genauer, so lassen sich diverse Voraussetzungen ableiten, die zwingend notwendig für die Funktionalität des beschriebenen Visualisierungsprozesses sind:

**Struktur und Art der Daten** Dem Softwareentwickler muss zur Entwicklungszeit bekannt sein, wie die Struktur und Art der Datensätze ist, die in die Anwendung gespeist werden. Erst dann existiert die Möglichkeit zum Aufbereiten und Formalisieren der Daten in einem Datenschema, die für den Visualisierungsprozess nötig sind, um anschließend das Mapping von Datenattributen zu Visualisierungselementattributen durchzuführen.

**Semantik der Daten** Dem Softwareentwickler muss zur Entwicklungszeit bekannt sein, welche Bedeutungen die Daten haben. Das Kennen der Bedeutung der zu visualisierenden Daten ermöglicht dem Softwareentwickler eine nach Kapitel 3.1.3 definierte effektive Visualisierung zu erzeugen. Weiterhin schränkt das Kennen des Anwendungskontexts die Menge der möglichen Visualisierungsziele ein.

**Pragmatik der Daten** Dem Softwareentwickler muss zur Entwicklungszeit der Kontext bekannt sein, in welchem die Daten zu visualisieren sind. Das Kennen des Anwendungskontexts ermöglicht ebenfalls wie die Semantik eine nach dem Kapitel 3.1.3 definierte effektive Visualisierung zu erzeugen. Gleichzeitig schränkt das Kennen des Anwendungskontexts die Menge der möglichen Visualisierungsziele ein.

### 3.2.4 Problematik statischer Informationsvisualisierung

Die in Kapitel 3.2.3 beschriebenen drei Voraussetzungen für den definierten Visualisierungsprozess limitieren diesen dadurch, dass diese Voraussetzungen bereits zur Entwicklungszeit dem Softwareentwickler bekannt sein müssen. Sollte nur eine dieser Voraussetzungen nicht zur Entwicklungszeit bekannt sein, führt dies zu einem nicht mehr realisierbaren Visualisierungsprozess.

Exemplarisch sei das in Abbildung 3.7 repräsentierte Szenario veranschaulicht. Man stelle sich vor, eine Anwendung arbeitet mit einer Vielzahl an heterogenen Datenquellen. Beispielsweise Datenbanken, Programmierschnittstellen, Endgerätschnittstellen etc. Jede dieser Datenquellen stellt ihre Daten dabei in einem anderen Format bereit. Bedeutet, dass die Art und Struktur der bereitgestellten Daten sich für jede Datenquelle unterscheiden (heterogene Datenformate). Bei Programmierschnittstellen, die in der Regel von externen Anbietern bereitgestellt werden - also anders als Datenbanken, nicht im eigenen System verwaltet werden - tritt zusätzlich das Risiko auf, dass sich die Datenformate rasant ändern können. Werden externe Programmierschnittstellen weiterhin über ein Web Service Discovery Mechanismus (vergleiche Universal Description, Discovery and Integration) automatisiert gefunden um diese anschließend zu konsumieren, ist es nicht unüblich, dass Informationssysteme mit einer unüberschaubaren Anzahl an Datenquellen und Datenformaten arbeiten.

Zur Entwicklungszeit wäre der Aufwand zur Definition von formalisierten Datenschemata, zur Entwicklung effektiver Visualisierungselemente und Codierung entsprechender Bindungslogiken zum Mapping von Datenattributen nach

Visualisierungselementattributen nicht zu bewältigen. Doch selbst bei Realisierung der beschriebenen aufwandstreibenden Aktionen zur Realisierung eines funktionsfähigen Visualisierungsprozesses ist die Chance hoch, dass sich rasant Änderungen an einer oder mehrerer Datenquellen ergeben, die dadurch eine ständige Anpassung der Anwendung erfordern.

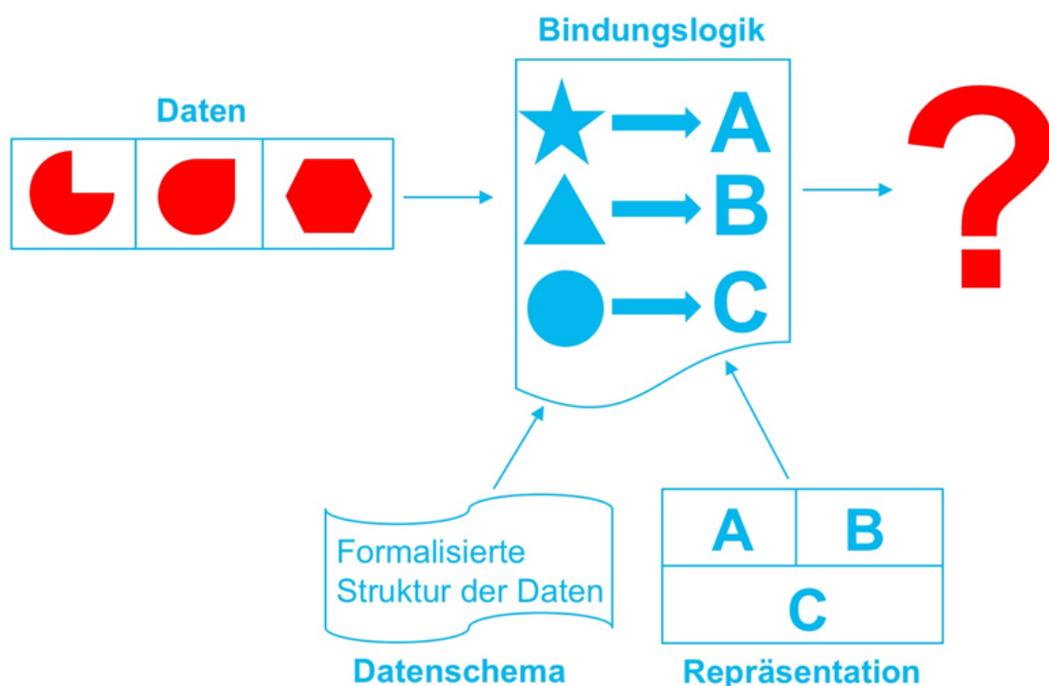


Abbildung 3.7: Die Darstellung eines Visualisierungsprozesses mit unbekanntem und nicht in einem Datenschema formalisiertem Datensatz (eigene Darstellung).

Abbildung 3.7 visualisiert den Vorgang, eines unbekanntem und damit nicht in einem Datenschema formalisierten Datensatz, dargestellt durch die roten Elemente. Die Bindungslogik kann die nicht bekannten Daten dementsprechend nicht Attribut-basiert den Visualisierungselementattributen zuordnen. Daraus resultiert, dass die Visualisierung nicht möglich ist. Der Visualisierungsprozess ist nicht funktionsfähig.

### 3.2.5 Statisch-generische Informationsvisualisierung

Die in Kapitel 3.2.4 beschriebene Problematik ist Gegenstand vieler Informationssysteme die eine hohen Anzahl heterogener Datenquellen und Datenformaten verarbeiten müssen. Dennoch wird in solchen Informationssystemen eine Visualisierung entsprechender Daten gewährleistet.

Abbildung 3.8 illustriert den Vorgang in Informationssystemen, die trotz hoher Dichte heterogener Datenquellen und Datenformate eine Visualisierung rendern.

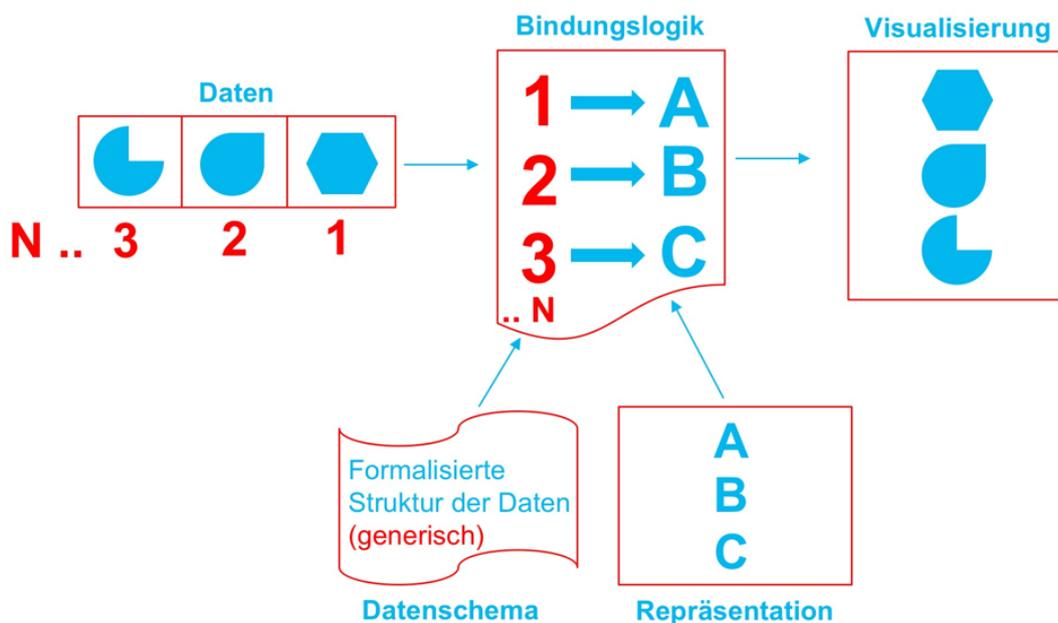


Abbildung 3.8: Die Darstellung eines Visualisierungsprozesses mit generischer Repräsentation der Daten (eigene Darstellung).

Die Lösung in solchen Systemen wird durch eine generische Visualisierung erreicht. Exemplarisch für eine solche generische Visualisierung ist die einfache textuelle Ausgabe jedes Datenattributs in einer Zeile. Neben der Visualisierung müssen bei diesem Visualisierungsprozess ebenfalls das Datenschema

und die Bindungslogik generisch entwickelt werden. Dies stellt sicher, dass jede Art von Daten, unabhängig der Struktur repräsentiert werden kann.

Eine generische Visualisierung ist in vielen Informationssystemen die Lösung trotz hoher Datenquellen- und Datenformatdichte, Daten zu visualisieren. Generische Visualisierungen widersprechen dabei aber den in Kapitel 3.1.3 definierten Qualitätsanforderungen an eine effektive Visualisierung. Gleichzeitig tragen diese damit nicht dazu bei, das in Kapitel 2 beschriebene Phänomen der Informationsüberflutung zu eliminieren oder wenigstens zu reduzieren.

### **3.3 Dynamische Informationsvisualisierung durch vollständige automatisierte Generierung grafischer Benutzeroberflächen**

Um die Problematik statischer Informationsvisualisierung, wie in Kapitel 3.2.4 beschrieben zu bewältigen und gleichzeitig eine spezifische und damit den Qualitätsanforderungen entsprechende effektive Visualisierung in Negation zur in Kapitel 3.2.5 beschriebenen generischen Visualisierung zu gewährleisten, bietet sich die automatisierte Generierung von Benutzeroberflächen an.

Bei der automatisierten Generierung von Benutzeroberflächen wird anhand der in ein Schema formalisierten Informationen eine vollständige Benutzeroberfläche generiert. Basis hierfür sind in der Regel die nicht mehr weiter unterteilbaren Benutzeroberflächenelemente wie beispielsweise Textfelder, Eingabefelder oder Buttons.

Die Fachliteratur zeigt, dass in der Vergangenheit die Forschung in diesem Bereich sehr kontrovers diskutiert wird. Die Ergebnisse verschiedener Studien zeigen weiterhin, dass eine automatisierte Generierung von Benutzeroberflächen für die meisten Anwendungen im Sinne der in Kapitel 3.1.3 definierten Qualitätskriterien nicht erfolgreich ist. Eine effektive Visualisierung durch automatisierte Generierung kann somit nicht gewährleistet werden. Der Grund hierfür ist die hohe Anzahl an Kombinationen zur Zusammenstellung geeigneter Teilvisualisierungen und die Anordnung der Teilvisualisierungen in einem ebenfalls für die Benutzeroberfläche generierten Layout. Eine Effektivität der

Visualisierung kann somit nie gewährleistet werden. Dennoch ist die Möglichkeit zur vollständig automatisierten Generierung mit Einhaltung der Effektivität möglich. Hierzu müssen allerdings diverse Voraussetzungen erfüllt sein. Exemplarisch hierfür sei die hohe Anzahl an Metadaten zu erwähnen. Je mehr Daten zur Generierung der Benutzeroberfläche bekannt sind, desto spezifischer kann die Benutzeroberfläche in Abstimmung auf die zu visualisierenden Daten automatisiert generiert werden. Metadaten können dabei die genaue Beschreibung der Benutzeroberfläche enthalten oder die Definition der Fähigkeiten, welche die Benutzeroberfläche erfüllen soll. Insgesamt kann festgehalten werden, dass je dynamischer die Daten sind und damit je dynamischer die Benutzeroberfläche sein soll, desto schwieriger ist es ein geeignetes Modell zu formalisieren zur Abbildung der Metainformationen und Daten für die Benutzeroberfläche. Somit können beispielsweise formularbasierte Benutzeroberflächen oder Benutzeroberflächen, die sehr datengetrieben sind, aber keine Manipulation der Daten benötigen, sehr einfach generiert werden, während zu manipulierende Daten Schwierigkeiten bei der Generierung machen. Direktmanipulation von Daten ist deshalb so schwierig, da auch hier die Vielzahl unterschiedlicher Abstufungen zur Visualisierung bei der Manipulation des Datenobjekts zu beachten ist [54, 55, 56].

Eine vollständige automatisierte Generierung ist in der Praxis schwierig und die Effektivität der Visualisierung gemäß den in Kapitel 3.1.3 beschriebenen Qualitätskriterien ist nicht zu gewährleisten. Für das konzeptuelle Modell muss in diesem Fall eine Alternative zwischen statischer und dynamischer Informationsvisualisierung gefunden werden.

## 4 Entwurf und Design des konzeptuellen Modells

Die Analyseergebnisse des Abschnitts drei beantworten bereits einige der in Kapitel 1.1 definierten Leitfragen. Hierbei wurde unter anderem die Wichtigkeit der Perzeption durch effektive Visualisierung hervorgehoben und die Gewährleistung einer effektiven Visualisierung durch die Identifikation von Qualitätskriterien beschrieben. Weiterhin wurde der in heutigen digitalen Informationssystemen angewandte Visualisierungsprozess anhand der Visualisierungspipeline analysiert, die statische Informationsvisualisierung präsentiert sowie die Voraussetzungen und Problematiken, verursacht durch statische Informationsvisualisierung identifiziert. Alternativen durch ein statisch-generisches Modell zur Informationsvisualisierung und die vollständige und automatisierte dynamische Informationsvisualisierung durch Generierung grafischer Benutzeroberflächen wurden illustriert.

Die im Abschnitt der Analyse gewonnenen Erkenntnisse beantworten die definierten Leitfragen für den Abschnitt der Analyse zusammengefasst wie folgt:

1. **Was wird unter einer effektiven Informationsvisualisierung verstanden und wie kann eine solche effektive Informationsvisualisierung gewährleistet werden?** Eine effektive Visualisierung unterstützt den menschlichen Perzeptionsprozess durch die visuelle Repräsentation der elementar zu verarbeitenden Informationen, sodass die Aufnahme der Informationen durch beispielsweise das Erkennen von Mustern und Strukturen vereinfacht wird. Die Gewährleistung effektiver Informationsvisualisierung wird hierbei durch Qualitätskriterien an Visualisierungen sichergestellt, die durch verschiedenste identifizierte Einflussfaktoren bestimmt werden.
2. **Wie wird die Informationsvisualisierung in heutigen digitalen Informationssystemen realisiert und welche Problematiken hinsichtlich**

**des Effektivitätskriteriums existieren derzeit bei der Informationsvisualisierung in heutigen digitalen Informationssystemen?** In heutigen digitalen Informationssystemen wird die effektive Informationsvisualisierung durch statische Codierungen von Visualisierungen für definierte Daten realisiert. Die statischen Codierungen von Visualisierungen sind hierbei nur zu realisieren, wenn die zu visualisierenden Daten zur Zeit der Entwicklung der Visualisierungen bekannt sind. Bei hoch-heterogenen Daten- und Datenquellendichten sind in der Regel vielerlei Daten unterschiedlicher Struktur involviert, die zur Entwicklungszeit nicht bekannt sind oder aufgrund der Höhe der Differenzierungen, zeitlich nicht zur Entwicklung realisierbar sind. Eine statische Realisierung der Visualisierung ist in diesem Fall nur durch die generische Repräsentation der Daten möglich, die das Effektivitätskriterium nicht erfüllt. Eine effektive Visualisierung gemäß dem in dieser Forschungsarbeit definierten Begriff der Effektivität bei benannter Problematik hinsichtlich hoch-heterogener Daten- und Datenquellendichte, durch den in heutigen digitalen Informationssystemen verwendeten statischen Informationsprozess ist somit nicht möglich.

3. **Welche Problematiken existieren bei der dynamischen Informationsvisualisierung durch vollständige, automatisierte Generierung von Benutzeroberflächen und warum ist der Einsatz dieses Verfahrens ungeeignet zur Beantwortung der Forschungsfrage?** Die dynamische Informationsvisualisierung durch vollständig-automatisierte Generierung grafischer Benutzeroberflächen besitzen die Problematik, dass grundlegend eine Visualisierung gemäß dem Effektivitätskriterium ermöglicht werden kann, eine Gewährleistung dieser Effektivität aber nicht sichergestellt werden kann. Grund hierfür ist die zu hohe Anzahl der dynamisch eingesetzten Faktoren, die bei der vollständigen Generierung grafischer Benutzeroberflächen angewendet werden. Eine effektive Visualisierung gemäß dem in dieser Forschungsarbeit definierten Begriff der Effektivität durch vollständig automatisierte dynamische Generierung grafischer Benutzeroberflächen ist somit grundlegend möglich, die Effektivität aber nicht zu gewährleisten.

Aufbauend auf den Erkenntnissen der Analyse sollen die folgenden Kapitel dieses Abschnitts die nachfolgenden Leitfragen beantworten, sodass anschließend mit den beantworteten Leitfragen dieses Abschnitts und des Abschnitts der Analyse die Beantwortung der Forschungsfrage das Endresultat bildet.

Die folgenden Leitfragen werden durch diesen Abschnitt definiert:

1. Welche architektonischen Strukturen und Prozesse muss ein konzeptuelles Modell aufweisen, um die erst zur Laufzeit bekannt werdenden heterogenen Datenstrukturen und Datenquellen gemäß dem Effektivitätskriterium grafisch zu repräsentieren und wie wird eine mögliche Integration der grafischen Repräsentation in den Anwendungskontext konzeptionell realisiert?
2. Wie muss eine Vokabularmenge zur Semantisierung spezifiziert werden, sodass Daten und grafische Repräsentationselemente semantisch beschrieben werden können und ein Auffinden von grafischen Repräsentationselementen in Abhängigkeit zum Effektivitätskriterium für entsprechende Daten als auch das Zuordnen der entsprechenden Daten zum gefundenen grafischen Repräsentationselement, gemäß dem konzeptuellen Modells ermöglicht wird?

Zur Beantwortung der deklarierten Leitfragen orientiert sich auch dieser Abschnitt an die für den Abschnitt definierten Leitfragen, in auf die Leitfragen abgestimmter chronologischer Ordnung in die zwei Kapitel: 4.1 Strukturen und Prozesse und 4.2 Spezifikation des Semantisierungsvokabulars.

#### **4.1 Strukturen und Prozesse**

Dieses Kapitel mit seinen Unterkapiteln stellt die folgende Leitfrage in den Fokus: 'Welche architektonischen Strukturen und Prozesse muss ein konzeptuelles Modell aufweisen, um die erst zur Laufzeit bekannt werdenden heterogenen Datenstrukturen und Datenquellen gemäß dem Effektivitätskriterium grafisch zu repräsentieren und wie wird eine mögliche Integration der grafischen Repräsentation in den Anwendungskontext konzeptionell realisiert?'

Zur Beantwortung, der für dieses Kapitel und entsprechende Unterkapitel definierte Leitfrage ist im ersten Schritt ein grundlegendes Verständnis für grafische Repräsentationen im Gesamtkontext grafischer Benutzeroberflächen von Nöten. Durch das aufgebaute Wissen der Anatomie von grafischen Benutzeroberflächen ist es in Kombination mit den Ergebnissen aus dem Analysekapitel 3 möglich, ein konzeptuelles Modell aufzustellen, dass die Vorteile der statischen Informationsvisualisierung, mit der Notwendigkeit der dynamischen Informationsvisualisierung zur Beantwortung der gestellten Forschungsfrage kombiniert. Der Entwurf des Modells wird entsprechend in einen Grobentwurf und einen Feinentwurf gegliedert. Weiterhin wird Thema dieses Kapitels sein, die Integration des entwickelten konzeptuellen Modells in den Anwendungskontext der Schul-Cloud zu überführen.

#### 4.1.1 Anatomie grafischer Benutzeroberflächen

Ein Verständnis für die Anatomie grafischer Benutzeroberflächen ist von elementarer Bedeutung um ein Verständnis für das im folgenden vorgestellten konzeptuellen Modell zu erhalten. Der Aufbau solcher grafischer Benutzeroberflächen wird daher entsprechend an der Methodologie 'Atomic design' erläutert. Atomic design ist eine von dem Webdesigner Brad Frost vorgestellte Methodologie welche die hierarchische Gliederung grafischer Benutzeroberflächen anhand atomisierter grafischer Elemente in einem Gedankenmodell skizziert.

Frost definiert in seinem Gedankenmodell fünf Bestandteile zur Beschreibung der Anatomie grafischer Benutzeroberflächen: Atome, Moleküle, Organismen, Templates und Seiten. Abbildung 4.1 illustriert diese Bereiche [57].

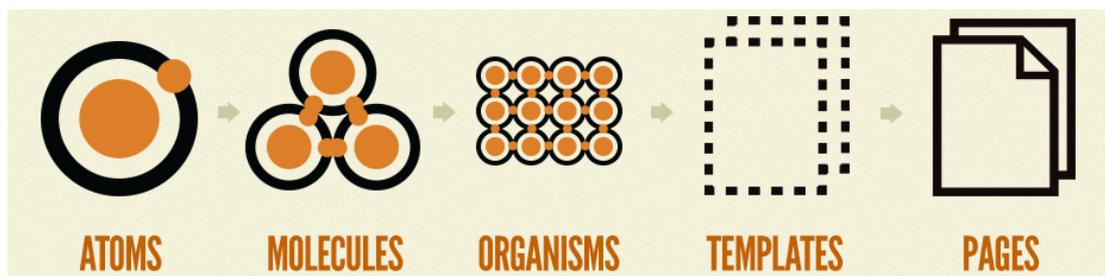


Abbildung 4.1: Die fünf Bestandteile der Atomic design Methodologie zur Illustration der Anatomie grafischer Benutzeroberflächen: Atome, Moleküle, Organismen, Templates und Seiten [57].

Die als Atome bezeichneten Bestandteile der Atomic design Methodologie repräsentieren die kleinsten Einheiten innerhalb grafischer Benutzeroberflächen. Bei der Auflösung der Analogie der Atome aus der Chemie hin zur Informationsvisualisierung, würde man bei grafischen Benutzeroberflächen einfache grafische Bestandteile wie exemplarisch Textfelder, Eingabefelder oder Buttons als Atome bezeichnen. Die als Atome bezeichneten Elemente innerhalb grafischer Benutzeroberflächen erfüllen hierbei die Bedingungen, nicht weiter zerlegbar zu sein, ohne einen entsprechenden Verlust der für das Element definierten Funktionalität aufzuweisen. Die exemplarisch genannten atomisierten Elemente Textfelder, Eingabefelder und Buttons erfüllen entsprechend diese Bedingung zur Geltung als nicht weiter atomisierbare Elemente. Die als Atome bezeichneten Elemente repräsentieren dementsprechend die Bausteine zur Realisierung grafischer Benutzeroberflächen. Jedes dieser Atome besitzt eine einzigartige Kombination von Eigenschaften. Dies können in grafischen Benutzeroberflächen etwa die Dimension eines Bildes oder die Schriftgröße einer Überschrift sein. Die für das Atom definierten Eigenschaften referenzieren sich hierbei immer auf den Dateninhalt des jeweiligen Atoms. Die deklarierten und initialisierten Eigenschaften für ein definiertes Atom X können somit nicht auf den Dateninhalt eines definierten Atom Y angewendet werden. Exemplarisch hierfür ist die Definition eines Textfeldes als Atom mit dem Dateninhalt 'Schul-Cloud' und der Eigenschaft 'Schriftgröße' mit dem Wert achtzehn Pixel. Die definierte Eigenschaft 'Schriftgröße' würde dementsprechend auf den Dateninhalt 'Schul-Cloud' angewendet werden, welches wiederum mit der Schriftgröße von achtzehn Pixeln dargestellt werden würde. Die Menge der Eigenschaften

aller in einer grafischen Benutzeroberfläche repräsentierten Atome wirken sich elementar auf die Eingliederung der entsprechenden Atome in das Layout der grafischen Benutzeroberfläche aus [57].

Neben den Atomen führt Frost die Ebene der Moleküle zur Beschreibung der Anatomie grafischer Benutzeroberflächen ein. Moleküle stellen nach den Atomen die nächste größere Einheit dar bei Betrachtung der grafischen Benutzeroberfläche als Hierarchiestruktur. Betrachtet man Moleküle im Kontext der Domäne der Chemie, gelten die Moleküle als eine Gruppe zusammengesetzter Atome. Wie auch Atome können Moleküle Eigenschaften definierten und damit repräsentieren, die für die gesamte Menge der Atome innerhalb des Moleküls gelten. In grafischen Benutzeroberflächen sind Moleküle dementsprechend ein Verbund grafischer Elemente [57].

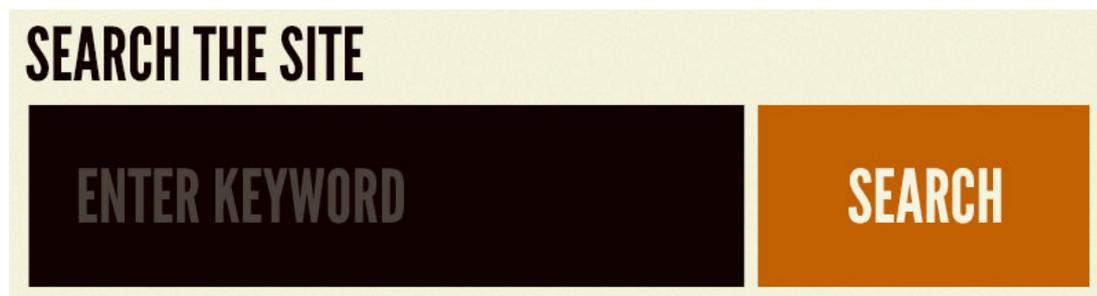


Abbildung 4.2: Exemplarisch als Molekül assoziiertes Suchformular, mit den grafischen Elementen Textfeld, Eingabefeld und Button, assoziiert als Verbund von Atomen innerhalb des Moleküls [57].

Ein solcher exemplarischer Verbund grafischer Elemente assoziiert mit einer Menge von Atomen, gebunden in einem Molekül illustriert die Abbildung 4.2. Hierbei werden die grafischen Elemente Textfeld, Sucheingabe und Button die im Sinne der Atomic design Methodologie als Atome assoziiert werden, zu einem Verbund von grafischen Elementen zusammengesetzt, dass in Atomic design als Molekül bezeichnet wird. Wie bereits bei den Atomen beschrieben, sind die grafischen Elemente, die als Atome bezeichnet werden, soweit zerlegt, dass diese gerade noch deren bestimmende Funktionalität besitzen. Die Zusammensetzung der einzelnen grafischen Elemente als Verbund realisiert einen Kontext - im Beispiel den Kontext einer Suche - und definiert dabei eine

für den Betrachter auszuführende Funktionalität, auf höherer Ebene. Im Beispiel der Suche wäre die übergeordnete Funktionalität, die für den Betrachter durchführbar ist, die Suche nach Informationen. Die Funktionalität auf höherer Ebene, durch den Verbund grafischer Elemente mit einer minimalen Funktionalität machen die in Atomic design bezeichneten Moleküle zu einer einfachen, portablen und wiederverwendbaren Komponente, die in allen grafischen Benutzeroberflächen verwendet werden, in der die gleiche übergeordnete Funktionalität dem Betrachter angeboten werden soll. Konkretisiert für das Beispiel, kann die Komponente durch den Verbund grafischer Elemente zu einer Suche in allen grafischen Benutzeroberflächen wiederverwendet werden, in denen eine Suchfunktionalität benötigt wird [57].

Organismen stellen die in der Atomic design Methodologie nächste größere Einheit dar. Die in der von Frost verwendeten Analogie zur Chemie als Organismus bezeichnetes Element, dient der Gruppierung von sowohl Atomen, Molekülen als auch anderen Organismen. Beim Verlassen der Analogie hin zu grafischen Benutzeroberflächen, skizziert ein solches Element die unterschiedlichen Abschnitte einer grafischen Benutzeroberfläche. Exemplarisch hierfür wäre der Verbund des Suchformulars aus vorherigem Beispiel in Kombination mit einer Navigation als Verbund atomisierter grafischer Elemente in weiterer Kombination mit einem atomisierten Element welches ein Logo repräsentiert. Dieses Beispiel könnte dementsprechend den Kopfbereichsabschnitt einer grafischen Benutzeroberfläche darstellen [57].

Die nächste Abstufung innerhalb der Atomic design Methodologie stellen Templates dar. Templates arrangieren die Anordnung einzelner grafischer Elemente in einem definierten Layout und bildet damit die zugrunde liegende Inhaltsstruktur ab. Template repräsentieren entsprechend den Gesamtkontext einzelner grafischer Elemente. Eine weitere Abstufung von Atomic design nach Template sind Seiten, welche die höchste Hierarchiestufe innerhalb grafischer Benutzeroberflächen darstellen und lediglich die konkreten Instanzen von Templates repräsentieren [57].

#### 4.1.2 Grundlegender Aufbau der Strukturen und Prozesse

Das Ziel dieses Kapitels ist der grundlegende Aufbau eines konzeptuellen Modells zur Abbildung einer Harmonie zwischen statischer und dynamischer Informationsvisualisierung mit der Gewährleistung der Effektivität für das menschliche kognitive System gemäß der in Kapitel 3.1.3 beschriebenen Qualitätskriterien an eine Visualisierung. Die Harmonie zwischen statischer und dynamischer Informationsvisualisierung soll gewährleistet werden durch die notwendige Nutzung des in Kapitel 3.3 beschriebenen Modells zur dynamischen Informationsvisualisierung. Das Modell soll entsprechend so revolutioniert werden, dass eine effektive Informationsvisualisierung gewährleistet werden kann, so wie sie bei der statischen Informationsvisualisierung aus Kapitel 3.2.2 bekannt ist. Für den grundlegenden Entwurf des konzeptuellen Modells soll das Wissen von der Anatomie grafischer Benutzeroberflächen genutzt werden, so wie im vergangenen Kapitel 4.1.1 beschrieben.

Die bereits in Kapitel 3.3 erläuterte dynamische Visualisierung durch vollständige automatische Generierung von Benutzeroberflächen ermöglicht zwar gemäß der in Kapitel 1.1 definierten Problemstellung die grafische Repräsentation von Daten, die zur Entwicklungszeit der Software unbekannt sind, behält allerdings die Problematik, dass grafische Repräsentationen nicht effektiv zu gewährleisten sind. Anders hingegen die statische Informationsvisualisierung, die aufgrund der zur Entwicklungszeit bekannten Daten, manuell effektiv grafische Repräsentationen der Anwendung hinterlegen kann. Würde das dynamische Informationsvisualisierungsverfahren eine effektive Visualisierung gewährleisten können, wäre die Problematik, die in der Problemstellung dieser Forschungsarbeit definiert wurde, bewältigt.

Die Effektivität im dynamischen Informationsvisualisierungsprozess kann nicht gewährleistet werden, da der automatisierte Aggregationsprozess einzelner nicht mehr zerlegbarer grafischer Elemente zu einem Verbund eines grafischen Elements mit einer definierten Funktionalität nach aktueller Studienlage nicht zuverlässig im Sinne zur Unterstützung der visuellen Wahrnehmung durchgeführt werden kann. In der Analogie der Atomic design Methodologie

bedeutet dies, dass eine Aggregation von definierten Atomen zu einem Molekül nicht automatisiert werden kann, ohne dabei eine Gewährleistung dafür zu erhalten, dass die Zusammenstellung einzelner Atome zu einem Molekül den visuellen Wahrnehmungsprozess des Menschen unterstützt und so die Informationsaufnahme vereinfacht.

Zur Sicherstellung der Effektivität bedeutet dies, dass nach aktuellen Erkenntnissen eine Sicherstellung der Unterstützung des kognitiven Wahrnehmungsprozesses durch Visualisierung generell nur sichergestellt werden kann, wenn die nicht mehr zerlegbaren grafischen Elemente manuell zu einem Verbund zusammengesetzt werden die eine entsprechende Funktionalität aufweisen und gezielt den kognitiven Wahrnehmungsprozess unterstützen. Abbildung 4.3 zeigt das hierfür entwickelte grundlegende konzeptuelle Modell.

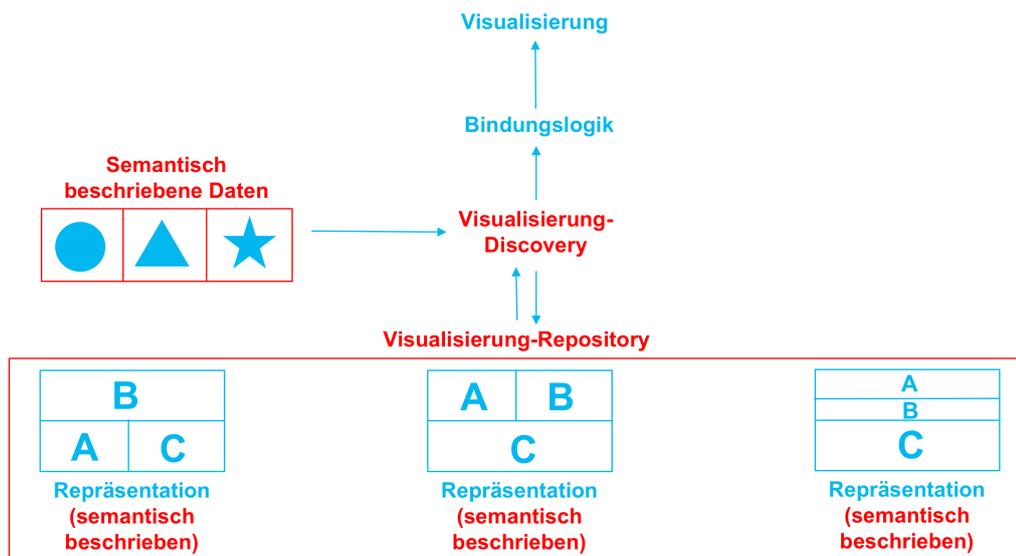


Abbildung 4.3: Die Grundstruktur des konzeptuellen Modells charakterisiert durch die drei Kernkomponenten: Semantisch beschriebene Daten, Visualisierungs-Repository mit semantisch beschriebenen grafischen Repräsentationselementen sowie Discovery- und Bindungslogik (eigene Darstellung).

Die Grundstruktur des konzeptuellen Modells ist charakterisiert durch drei Kernkomponenten:

- a) Semantisch beschriebene Daten
- b) Visualisierungs-Repository mit semantisch beschriebenen grafischen Repräsentationselementen
- c) Discovery- und Bindungslogik

Die Kernkomponente 'Visualisierungs-Repository' enthält einzelne grafische Repräsentationselemente, die im Sinne der Atomic design Methodologie als Moleküle bezeichnet werden. Das Repository ist hierbei als ein abgewandeltes Konzept eines Software-Repository anzusehen. Während in einem Software-Repository digitale Artefakte in Form von Quellcode abgelegt werden, handelt es sich beim Visualisierungs-Repository um die Ablage digitaler Artefakte in Form von Visualisierungselementen. Ein Repository kann dementsprechend abstrakt als eine Art Verzeichnis bezeichnet werden, das eine Menge digitaler Artefakte persistiert. Als eine solches soll das Visualisierungs-Repository dienen. Die im Repository abgelegten Visualisierungselemente erfüllen hierbei die in Kapitel 4.1.1 beschriebenen Eigenschaften, die gemäß der Atomic design Methodologie die Moleküle erfüllen. Zusammenfassend bedeutet dies, die im Repository abgelegten Visualisierungselemente enthalten jeweils eine für den kognitiven Wahrnehmungsprozess effektive Zuordnung von nicht mehr zerlegbaren Visualisierungselementen (Atome), die entsprechende Eigenschaften wie beispielsweise Farbe oder Schriftgröße annehmen und auf den Dateninhalt anwenden kann. Die einzelnen Visualisierungselemente (Moleküle) besitzen eine übergeordnete Funktion (beispielsweise Suche von Informationen) und sind wiederverwendbar, unabhängig des Kontextes, in der sie eingebunden werden. Die im Repository abgelegten Visualisierungselemente sind manuell durch ausgebildete User Interface Designer entwickelt und können so sicherstellen, dass eine effektive Visualisierung nach dem kognitiven Wahrnehmungsprozess gewährleistet werden kann. Die einzelnen Visualisierungselemente sind hierbei semantisch beschrieben. Die Semantisierung ermöglicht die Interpretation durch entsprechende Algorithmen der Kernkomponente 'Discovery- und Bindungslogik'.

Die Kernkomponente 'Daten' repräsentiert die heterogenen Daten, die von heterogenen Datenquellen stammen und gemäß der Problemstellung erst zur

Laufzeit ermittelt werden. Die Daten liegen dabei, ebenfalls wie die Visualisierungselemente semantisch beschrieben vor. Der Prozess des Transfers der unsemantisierten Daten in ein semantisches Modell ist hierbei nicht Gegenstand des konzeptuellen Modells.

Die Aufgabe der Kernkomponente 'Discovery- und Bindungslogik' besteht darin, anhand der Daten das effektivste Visualisierungselement aus dem 'Visualisierungs-Repository' zu finden und die Datenattribute den Visualisierungselementattributen zuzuordnen. Das Resultat hierbei ist ein für spezifische Daten effektives Visualisierungselement, das bereits die Daten zugeordnet hat und so in den Gesamtkontext der grafischen Benutzeroberfläche eingebunden werden kann. Das Auffinden eines Visualisierungselements für definierte Daten und das anschließende Zuordnen der definierten Daten zum gefundenen effektiven Visualisierungselement wird hierbei durch ein gemeinsames semantisches Modell mit einem gemeinsamen Semantikvokabular realisiert.

Ein semantisches Modell zur Beschreibung von Visualisierungselementen und Daten, sodass ein Auffinden von Visualisierungselementen und Zuordnen von Daten ermöglicht wird, ist somit elementar für die Realisierbarkeit des konzeptuellen Modells. Erst durch die Semantisierung von Daten und Visualisierungselementen können die Informationen maschinenlesbar interpretiert werden. Eine solche Interpretation findet hierbei durch die Algorithmik der Discovery- und Bindungslogik statt. Eine alternative Möglichkeit zur maschinellen Interpretation wäre die automatisierte Gewinnung der notwendigen Informationen aus Daten und Visualisierungselementen um ein Auffinden und Zuordnen zu gewährleisten. Der Autor dieser Forschungsarbeit ist allerdings der Meinung, dass die Gewinnung der entsprechenden Informationen über Informationsgewinnungsverfahren (Information Retrieval) ein ähnliches Gefahrenpotenzial in Bezug auf die Sicherstellung des Effektivitätskriteriums besitzt wie die dynamische Informationsvisualisierung durch vollständig automatisierte Generierung grafischer Benutzeroberflächen. Dieses Gefahrenpotenzial existiert bei der Semantisierung nicht, vorausgesetzt die Informationen werden korrekt in das semantische Modell transferiert. Das semantische Modell stellt die Abbildung entsprechender Qualitätskriterien sicher gemäß der in Kapitel 3.1.3 beschriebenen.

Zum Auffinden von effektiven Visualisierungselementen für spezifische Daten und Zuordnen der Daten zum Visualisierungselement muss ein semantisches Modell mit entsprechendem Vokabular spezifiziert werden, welches die in Kapitel 3.1.3 beschriebenen Hauptqualitätskriterien der Expressivität, Effektivität und Angemessenheit abbilden. Zur Konkretisierung hierfür wurde bereits im selbigen Kapitel 3.1.3 analysiert, welche Einflussfaktoren sich auf genannte Hauptqualitätskriterien auswirken. Dabei wurden die folgenden sechs Einflussfaktoren identifiziert:

1. Art und Struktur der Daten
2. Zielerreichung des Anwenders
3. Vorwissen des Anwenders
4. Visuelle Fähigkeiten und Vorlieben des Anwenders
5. Konventionen des Anwendungskontexts
6. Charakteristika des Darstellungsmediums

Viele Experimente des Autors mit verschiedenem definierten Semantisierungsvokabular erzielte dabei die Erkenntnis, dass die beiden Einflussfaktoren 'Art und Struktur der Daten' als auch 'Zielerreichung des Anwenders' existenziell sind für das Auffinden von Visualisierungselementen für definierte Daten und Zuordnen der definierten Daten zum Visualisierungselement. Die anderen genannten Einflussfaktoren sind hinsichtlich der Qualitätskriterien von hoher aber nicht von existenzieller Bedeutung zum Auffinden und Zuordnen. Die Begründung hierfür ergibt sich bei genauerer Betrachtung der beiden existenziell notwendigen Einflussfaktoren.

Inhaltsverzeichnis

- Wetter
- Witterung
- Klima
- Video
- Aufgaben

Sektion F  
 Wetter – Witterung – Klima

Sektion G

### Wetter

Das **Wetter** ändert sich oft täglich oder sogar mehrmals am Tag und ist selten überall gleich. Während es in Neuenburg regnet, kann es in Freiburg schön sein. Unter Wetter versteht man die *zeitlich veränderlichen, spürbaren Zustand der unteren Atmosphäre (Troposphäre)* zu einer bestimmten Zeit an einem bestimmten Ort. Das Wetter kann durch die Wetterelemente wie Temperatur, Luftdruck, Wind, Strahlung, Luftfeuchtigkeit, Bewölkung und Niederschlag beschrieben werden. (siehe auch: [Wetterelemente und ihre Messung](#))

### Witterung

Im Laufe eines Jahres kann das Wetter auch einige Tage bis Wochen anhalten und sich nicht ändern. **Witterung** stellt sich bei uns fast in jedem Jahr, gegen Ende September, warmes und trockenes Wetter mit ersten Nachfröhen dar. Die typischen Merkmale eines über mehrere Tage oder sogar Wochen gleich bleibenden Wetters nennen wir **Witterung**. Von der Witterung eines einzelnen Jahres hängt es zum Beispiel ab, wann die Landwirte ihre Felder bestellen und wie viel Sonne der Wein abbekommt und ob er dementsprechend süß oder sauer wird.

### Klima

Wenn aber ein neues Haus die richtige Dämmung und Wärmedämmung erhalten soll, spielen Wetter und Witterung eines einzelnen Jahres keine Rolle. Hierbei muss man das **Klima** beachten. Das sind die *durchschnittlichen* Verhältnisse von Wetter und Witterung über einen längeren Zeitraum von mehreren Jahrzehnten (man berechnet meist den Mittelwert von 30 Jahren). Diese werden jeden Tag aufgezeichnet, so dass man nach vielen Jahren die durchschnittlichen Werte daraus berechnen kann.

### Video

Sektion D

### Aufgaben

Sektion E

**AUFGABEN**

1. Beschreibe kurz anhand des Textes den Unterschied zwischen den Begriffen Wetter, Witterung und Klima?
2. Das Wetter kann man beobachten und messen, das Klima kann man nur berechnen. Erkläre diese Aussage.

Abbildung 4.4: Die exemplarische Darstellung eines digitalen Arbeitsblattes zur Nutzung als Lehrmaterial im Hypertext-Format [58] (modifizierte Darstellung).

Beim Einflussfaktor 'Art und Struktur der Daten' betrachtete der Autor dieser Forschungsarbeit eine Vielzahl an Dokumenten, um die Art und Struktur von Daten generisch allgemeingültig definieren zu können. Die generische Definition von Art und Struktur der Daten lässt sich dabei anhand eines Beispiels erläutern. Abbildung 4.4 zeigt ein digitales Arbeitsblatt zur Nutzung als Lehrmaterial im Hypertext-Format. Die visuelle Repräsentation des digitalen Arbeitsblattes ist hierbei ein gelungenes Exemplar einer effektiven grafischen Repräsentation von Daten. Die visuelle Wahrnehmung eines Menschen kann anhand der grafischen Repräsentation des Beispiels sofort identifizieren, welche Daten zusammenhängen: Textuelle Überschrift und beschreibender Text der Sektion A, textuelle Überschrift und beschreibender Text der Sektion B, textuelle Überschrift und beschreibender Text der Sektion C, textuelle Überschrift und Videomultimediaminhalt der Sektion D sowie textuelle Überschrift, Icongrafik und beschreibender Text der Sektion E. Zusätzlich identifizieren wir als Mensch auch die für das digitale Arbeitsblatt verfügbaren Metainformationen in Form der Dokumentenüberschrift innerhalb der Sektion F sowie die Metainformationen in Form der Dokumentenstrukturangabe innerhalb der Sektion G. Aus diesem Beispiel, das stellvertretend für eine Vielzahl von analysierten Dokumenten steht, lässt sich ableiten, dass Daten und damit auch die grafische Repräsentation Hierarchien aufweisen können.

Damit kann ein Datensatz mit dem Charakteristikum der Dimensionalität dessen beschrieben werden. Die Dimension wird dabei durch die Kombination multipler einzelner Daten charakterisiert. Die Anzahl differierender singulärer Datenwerte ergibt hierbei die Dimension des Datensatzes. Datensätze können hierbei wie in Sektion F des Beispiels eindimensional, wie in Sektion A, B, C oder D des Beispiels zweidimensional oder auch mehrdimensional sein wie in Sektion E des Beispiels. Eindimensionale Datensätze werden dabei oft nicht mit dem Begriff der Dimensionalität assoziiert und stellen den Standard da, wenn von einem Datensatz gesprochen wird. Bei der Beschreibung eines Datensatzes durch das Dimensionskriterium ist weiterhin zu beachten, dass die Dimensionalität ebenfalls in Abhängigkeit zum Datentyp des Datenwertes stehen kann. Bei zwei- und mehrdimensionalen Datensätzen entsprechend von den Datentypen aller Datenwerte innerhalb des Dimensionsverbunds. Beim

vorgestellten Beispiel könnte die Sektion D mit dem Videomultimediainhalt als komplexer Typ betrachtet werden, der zur Repräsentation des Videos Daten in mehreren Dimensionen bereitstellt. Generell lässt sich die Aussage treffen, dass je mehr Dimensionen ein Datensatz aufweist, desto schwerer ist die Realisierung einer effektiven grafischen Repräsentation zur Abbildung der Daten des Datensatzes [12, 59, 60].

Aus der Erkenntnis, dass die analysierten Dokumente Hierarchien von Daten und damit Hierarchien von Visualisierungselementen aufweisen, lässt sich neben der Dimensionalität eines Datensatzes auch anhand des Charakteristikums der Abhängigkeit von Daten untereinander definieren. Generell sind Daten dann untereinander abhängig, wenn ein oder mehrere (bei mehrdimensionalen Daten) Datenwerte eines Typs innerhalb eines Datensatzes von einem oder mehreren Datenwerten eines anderen Typs innerhalb des Datensatzes nicht unabhängig voneinander beschrieben werden können [60, 61].

Ein weiteres Charakteristikum zur generischen Beschreibung von Datensätzen wurde bei der Dokumentenanalyse im gültigen Wertebereich gefunden. Dieses Charakteristikum definiert einen Bereich von Werten, innerhalb dessen der zu deklarierende Wert für das Datum liegen muss. Ein solcher gültiger Wertebereich kann künstlich definiert werden oder ist durch technische Gegebenheiten bereits definiert. Die technischen Gegebenheiten beschränken sich aber in der Regel auf den Wertebereich, aus dessen ein Datentyp der für das Datum deklariert wurde, einen Wert für sich deklarieren kann. Exemplarisch hierfür wäre die konkrete Instanz eines primitiven skalaren Datentyps. In den Programmiersprachen der Softwareentwicklung gibt es hierfür verschiedene primitive Instanztypen. Einer dieser skalaren Datentypen in der Instanz ist Integer. Dieser Datentyp repräsentiert Ganzzahlen und ist im Wertebereich auf einem 32-Bit-Computersystem beschränkt auf maximal zehn Dezimalstellen ohne Vorzeichen. Die künstliche Auswahl des gültigen Wertebereichs ist damit durch die technische Gegebenheit beschränkt und kann nur aus diesem Bereich gewählt werden [59, 61].

Das letzte generische Charakteristikum zur Beschreibung von Datensätzen ist die Art der Datenwerte. Die Art der Datenwerte lässt sich dabei weiter unter-

gliedern in die Charakteristiken: skalare Daten, nominale Daten, ordinale Daten, kontinuierliche Daten, vektorielle Daten und tensorielle Daten. Abbildung 4.5 skizziert die Gliederung von Art der Datenwerte.

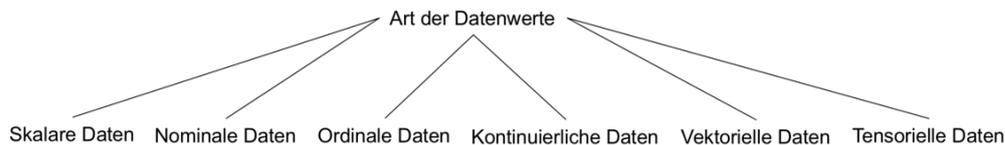


Abbildung 4.5: Taxonomie des Charakteristikums 'Art der Datenwerte' des Einflussfaktors Art und Struktur der Daten (eigene Darstellung).

Bei den skalaren Daten handelt es sich um Daten, denen einen eindeutigen Wert zugewiesen werden kann. Skalare Daten werden oft synonym zur Bezeichnung der quantitativen Daten verwendet. Eine Variation der skalaren Daten sind die skalaren Felder. Bei den skalaren Feldern handelt es sich um die Repräsentation von zwei- oder mehrdimensionalen Feldern skalarer Daten. Exemplarisch für skalare Daten in Anlehnung an das Beispiel wäre die Temperatur mit einem eindeutigen Wert von 10 Grad Celsius. Ein anderes Beispiel wäre die Abbildung einer Fließkommazahl mit dem eindeutigen Wert 3,652 [61, 62, 63].

Bei den nominalen Daten handelt es sich um Daten, deren Werte mit anderen Datenwerten kategorisiert werden können. Die als nominal bezeichnete Daten werden oft synonym zur Bezeichnung der qualitativen Daten verwendet. Kategorisierte Datenwerte nominalen Ursprungs weisen keine natürlich-definierte Ordnung auf. Trotz der nicht definierten Ordnung von nominalen Werten können die Datenwerte innerhalb einer nominalen Kategorisierung voneinander unterschieden werden. Exemplarisch für nominale Daten wäre das Geschlecht mit den beispielhaften Wertemöglichkeiten männlich und weiblich oder die Nutzung von Dicotomie mit den beispielhaften Wertemöglichkeiten Ja und Nein [59, 60, 61, 62].

Bei den ordinalen Daten handelt es sich Daten, deren Werte wie die der nominalen Daten kategorisiert werden können. Die als ordinal bezeichnete Daten

werden oft synonym zur Bezeichnung der qualitativen Daten verwendet. Im Unterschied zu den nominalen Werten weisen kategorisierte Datenwerte ordinalen Ursprungs aber eine natürlich-definierte Ordnung auf. Exemplarisch für ordinale Daten wären Größencharakteristiken mit den natürlich-geordneten Wertemöglichkeiten klein, mittel und groß oder Qualitätseinschätzungen mit den geordneten Wertemöglichkeiten sehr gut, gut, mittel, schlecht, sehr schlecht [59, 60, 61, 62].

Kontinuierliche Daten sind Daten, deren Werte innerhalb eines definierten Wertebereichs alle Werte annehmen können. Die als kontinuierliche bezeichnete Daten werden oft synonym zur Bezeichnung der quantitativen Daten verwendet. Hierbei sind die Werte nur beschränkt durch die Genauigkeit der Werterfassung. Exemplarisch für kontinuierliche Daten wäre menschliche Gewichtscharakteristiken mit einem definierten Wertebereich von 4,5 Kilogramm als Minimalgewicht und 596 Kilogramm als Maximalgewicht mit einer Werterfassungsgenauigkeit von 0,1 Kilogramm [64].

Vektorielle Daten sind Daten, die genau wie die skalaren Daten einen eindeutigen Wert zugewiesen haben. Zusätzlich zum Datenwert und damit in Unterscheidung zu den skalaren Daten werden den vektoriellen Daten zusätzlich eine Richtung zugewiesen. Das Richtungsattribut ist notwendig für Daten im zwei- oder hördimensionalen Raum. Exemplarisch für vektorielle Daten ist die Bewegungsrichtung oder Kraft [61, 62, 63, 64].

Tensorielle Daten sind eine Verallgemeinerung von skalaren und vektoriellen Daten auf höherer Dimension. Die Ordnung eines Tensors ist hierbei gleichbedeutend mit der Dimension, die ein Feld besitzen muss, um selbigen zu beschreiben [63, 64].

Durch das Aufweisen von Hierarchien in Daten und entsprechenden grafischen Repräsentationen und die Bestätigung durch die gewonnenen Erkenntnisse durch Aufführen der einzelnen generischen Charakteristiken zur Beschreibung von Daten lässt sich somit die existenzielle Bedeutung von der Art und Struktur der Daten nachweisen. Erst wenn die Struktur der Daten mit der Struktur des Visualisierungselements übereinstimmt, ist eine Zuordnung von Datenattributen zu Visualisierungselementattributen generell möglich. Abbildung 4.6

zeigt exemplarisch die Struktur eines Datensatzes und die Struktur eines Visualisierungselements die eine äquivalente Struktur aufweisen und damit ein Zuordnen von Daten ermöglichen.

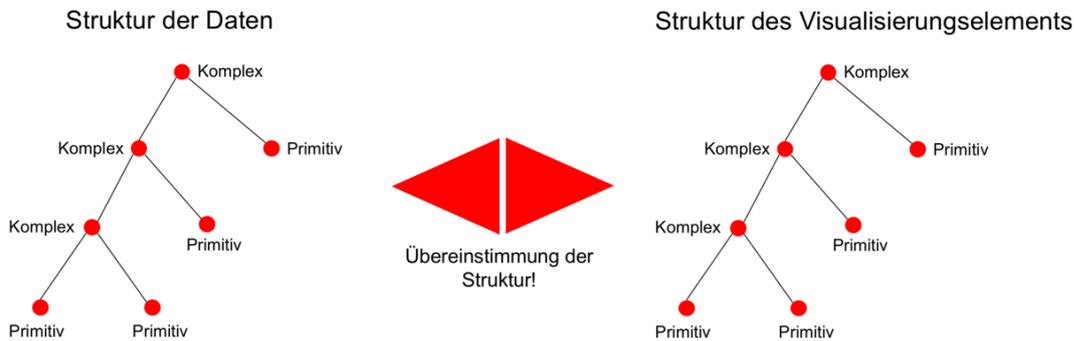


Abbildung 4.6: Exemplarische hierarchische Darstellung der Struktur eines Datensatzes und die Struktur eines Visualisierungselements als Baum die eine Äquivalenz hinsichtlich der Struktur aufweisen und damit die Zuordnung von Daten zum Visualisierungselement ermöglichen (eigene Darstellung).

Sowohl Daten als auch Visualisierungselement sind in der Abbildung 4.6 als Baum dargestellt. Die korrekte Semantisierung von Daten und Visualisierungsmodell in einem zu entwickelnden semantischen Modell soll den Aufbau solcher Bäume ermöglichen. Die verschiedenen Knoten innerhalb der Bäume sind entweder komplexer und primitiver Natur. Bei komplexen Knoten handelt es sich um tensorielle, vektorielle oder im Allgemeinen um mehrdimensionale Daten. Die primitiven Knoten stellen wiederum einen primitiven Datentypen wie zum Beispiel ein Skalar dar. Zum Zuordnen von Daten zum Visualisierungselement durch die 'Discovery- und Bindungslogik' ist es erforderlich, dass ein die Struktur von Daten und Visualisierungselement immer vollständig aufgebaut wird. Dies bedeutet, dass auf unterster Ebene der Bäume immer primitive Knoten repräsentiert werden, zu denen letztlich die Zuordnung ermöglicht werden kann. Das bedeutet, dass eine Zuordnung von Datenattributen zu Visualisierungselementattributen immer auf primitiver Ebene stattfindet. In der exemplarischen Darstellung 4.6, sind Struktur von Daten und Visualisierungselement jeweils äquivalent. Dies bedeutet, dass wenn die Art und Struktur der Daten

das alleinige Einflusskriterium wäre, bereits jetzt eine effektive Visualisierungsform für die entsprechenden Daten vorlege. Die Komponente 'Discovery- und Bindungslogik' würde beim Prozess des Auffindens also immer einen Baum für die Daten aufbauen, und diese mit allen Bäumen, der existierenden Visualisierungselementen im 'Visualisierungs-Repository' abgleichen. Existiert ein strukturell äquivalenter Baum für die Daten, wäre eine effektive Visualisierungsform für die Daten gefunden. Eine Zuordnung der Daten zum Visualisierungselement kann nun auf den jeweils einzelnen Ebenen des Baumes stattfinden.

Betrachtet man die Sektion A der Abbildung 4.4 würde dort ein Baum entstehen mit einem Komplex, dass die Sektion repräsentiert und zwei primitiven Blattknoten, die jeweils die Überschrift und die textuelle Beschreibung repräsentieren. Der Baum für die Daten hätte ebenfalls ein Komplex und zwei Primitive. Die beiden Primitiven der Daten würden jeweils den Primitiven des Visualisierungselements zugeordnet werden.

Dies macht noch einmal die existenzielle Relevanz des Einflussfaktors 'Art und Struktur der Daten' klar. Nur wenn Art und Struktur von Daten als auch die des Visualisierungselements äquivalent sind, ist generell eine Zuordnung von Daten möglich. Das bedeutet, dass auch erst damit die Effektivität eines Visualisierungselements für spezifische Daten gewährleistet werden kann.

Zwei Ausnahmen ermöglichen eine Effektivität auch unabhängig der Äquivalenz der Bäume vollständiger Daten. Die erste Ausnahme ist begründet darin, dass Daten die eine hierarchische Struktur aufweisen nicht immer zwingend diese hierarchische Struktur beibehalten müssen. Wenn eine lineare Darstellung einer hierarchischen Struktur eines Datensatzes möglich ist, ohne dass der Sinn dieser Daten dabei verloren geht, kann der Datensatz unabhängig der hierarchischen Struktur mithilfe der 'Discovery- und Bindungslogik' ein effektives Visualisierungselement finden und Daten entsprechend zuordnen. Das zu entwickelnde semantische Modell muss deshalb klar definierbar ermöglichen, welche Daten unzertrennlich ohne funktionale Einschränkungen zusammengehörig sind und welche aufbrechbar sind. Nur dann ist ein Auffinden und Zuordnen durch die Komponente 'Discovery- und Bindungslogik' möglich. Die

zweite Ausnahme bildet ebenfalls die Transformation von Daten und die damit einhergehende Veränderung des Baumes. Hierbei ist der zweite existenziell genannte Einflussfaktor gemeint: Zielerreichung des Anwenders.

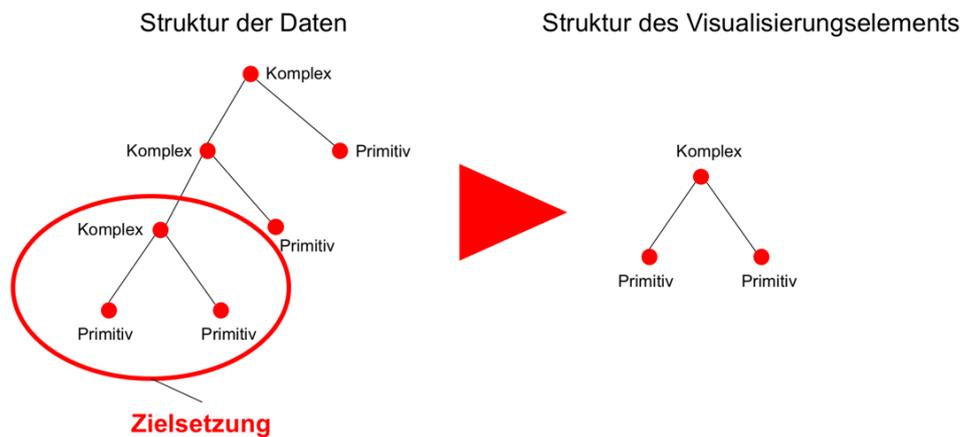


Abbildung 4.7: Exemplarische hierarchische Darstellung der Struktur eines Datensatzes und die Struktur eines Visualisierungselements als Baum mit einem Visualisierungsziel, das eine Transformation des Datenbaums zur Folge hat (eigene Darstellung).

Wie die erste Ausnahme bildet auch die Zielerreichung des Anwenders ein ähnliches Szenario. In diesem Fall ist das Ziel des Anwenders, ausschließlich die Darstellung der Sektion A des Beispiels in Abbildung 4.4. In diesem Fall würde der Baum wie in Abbildung 4.7 exemplarisch dargestellt, transformiert werden und die 'Discovery- und Bindungslogik' würde ein passendes Visualisierungselement für den transformierten Baum aufsuchen. Da auch die 'Zielerreichung des Anwenders' als Einflussfaktor unmittelbar Einfluss auf die 'Art und Struktur des Baumes' hat, ist auch dieser Einflussfaktor existenziell.

Die fünf weiteren Einflussfaktoren gemäß den Qualitätskriterien sind den Einflussfaktoren 'Art und Struktur der Daten' sowie 'Zielerreichung des Anwenders' untergeordnet. Das 'Vorwissen des Anwenders' als Einflussfaktor kann, nachdem ein Zuordnen von Daten grundsätzlich möglich ist, als Entscheidungskriterium für die Effektivität herangezogen werden, ebenfalls die 'Visuellen Fähigkeiten und Vorlieben des Anwenders', 'Konventionen des Anwendungskontexts' als auch 'Charakteristika des Darstellungsmediums'.

Alle Einflussfaktoren sind somit von Relevanz für das Auffinden und Zuordnen von Daten. Die beiden Einflussfaktoren 'Art und Struktur der Daten' und 'Zielerreichung des Anwenders' existenziell, die anderen Einflussfaktoren als weitere Entscheidungskriterien für die Effektivität eines Visualisierungselements. Dementsprechend müssen bei der Spezifikation eines Vokabulars für das semantische Modell alle Einflussfaktoren beachtet werden.

#### 4.1.3 Konkretisierter Aufbau der Strukturen und Prozesse

Dieses Kapitel konkretisiert den konzeptuellen Aufbau des vorherigen Kapitels 4.1.2. Zielsetzung ist die Konkretisierung der Strukturen und Prozesse, sodass eine Implementierung des konzeptuellen Modells ermöglicht wird.

Die Kernkomponente der semantisch beschriebenen Daten kann implementierungstechnisch durch das Artefakt von Hypertextdokumenten repräsentiert werden. Die Auszeichnung der Daten erfolgt mit dem spezifizierten Vokabular, dass in den nachfolgenden Kapiteln näher betrachtet wird. Die Auszeichnung von Daten grundsätzlich und der Transformationsprozess von Rohdaten in das semantische Modell ist nicht Gegenstand dieser Forschungsarbeit und wird deshalb nicht detailliert betrachtet. Es soll an dieser Stelle nur erwähnt werden, dass die Auszeichnung von Daten durch Semantik in der Regel durch einen manuellen Vorgang mit menschlichem Zutun erfolgt. Möglich wäre allerdings auch die automatisierte Transformation von Rohdaten in das semantische Modell, durch Algorithmen im Bereich der künstlichen Intelligenz, kombiniert mit Informationsretrievalverfahren. Die Integration des konzeptuellen Modells in den Anwendungskontext der Schul-Cloud wird in Kapitel 4.1.4 näher betrachtet. Die Vorarbeit zur Integration sollte durch die Zentralisierung der semantisierten Hypertextdokumente erfolgen. Eine solche Zentralisierung kann durch Ablage der Dokumente in einem Repository erreicht werden. Hierzu eignen sich private Repositories des Codehosting-Dienstleisters GitHub [65]. Bei privaten Repositories handelt es sich, anders als bei den öffentlichen Repositories um einen nur für den Eigentümer des Repository und deren Bevollmächtigten zugängliches Repository. GitHub als Plattform für das Hosting von Softwareartefakten ermöglicht in diesem Zuge den Zugriff auf Repositories

über Programmierschnittstellen [66], sodass ein maschineller Zugriff ermöglicht werden kann. Dies ist für die Integration in den Anwendungskontext der Schul-Cloud von besonderer Relevanz.

Bei der Konkretisierung der Kernkomponente 'Visualisierungs-Repository mit semantisch beschriebenen grafischen Repräsentationselementen' durch eine Instanz, bedarf es zweierlei grundlegender Überlegungen. Zum einen muss entschieden werden, wie die grafischen Repräsentationselemente gemäß den Beschreibungen der vorherigen Kapitel formalisiert werden können. Zum anderen muss entschieden werden, wie die grafischen Repräsentationselemente persistent gehalten werden können. Zur persistenten Haltung von Visualisierungselementen bieten sich ebenfalls Repositories beim Codehostingdienstleister GitHub [65] an.

Die erste Fragestellung kann beantwortet werden, durch die Nutzung sogenannter Web Components. Web Components sind wiederverwendbare Visualisierungselemente, die genau dies widerspiegeln welches der Autor Frost, wie in Kapitel 4.1.1 beschrieben, als Moleküle bezeichnet [67]. In vielen Fällen wird werden Web Components mit React Components verglichen. Beide deklarieren aber jeweils einen eigenständigen Problembereich für sich.

Bei den Web Components handelt es sich um den Gegenstandsbereich der wiederverwendbaren und strikten Entkoppelung vom Quellcode des Kontexts, in der die Component eingebunden wird [67].

React Components hingegen deklarieren den Gegenstandsbereich der stetigen Synchronität des Document Object Model (DOM) mit den Datenwerten. Aus diesem Grund empfiehlt sich die Verwendung von Web Components.

Für Web Components existiert zusätzlich die Plattform [WebComponents.org](https://webcomponents.org). Die Plattform ermöglicht das Teilen und Finden wiederverwendbarer Visualisierungselemente auf Basis von Web Components. Zur Erreichung der Mission von [WebComponents.org](https://webcomponents.org) durch Teilen und Finden von Visualisierungselementen, ermöglicht die Plattform das Veröffentlichen eigener Web Components sowie das Durchsuchen und auch gezielte Suchen nach allen auf der

Plattform veröffentlichten Web Components [68]. Statt der Nutzung eines einfachen Repository, bietet sich die Nutzung der Plattform WebComponents.org für das konzeptuelle Modell an. Die Auswahl der Plattform als Repräsentation des Visualisierungs-Repository ist hierbei begründet durch

- a) die bereits große Sammlung an wiederverwendbaren Visualisierungselementen.
- b) die grundsätzliche Bereitstellung von Infrastruktur um Visualisierungselemente auf der Plattform zu Veröffentlichen, zu Durchsuchen und gezielt zu Suchen.
- c) die Bereitstellung einer Programmierschnittstelle zur Nutzung der Infrastruktur der Plattform.
- d) die Installation und Aktualisation der meisten auf der Plattform veröffentlichten Visualisierungselemente über das Paketverwaltungstool Bower.
- e) das Hosting des Visualisierungselements auf der etablierten Plattform GitHub in einem öffentlich zugänglichen Repository.

WebComponents.org kann abstrakt als Verzeichnis betrachtet werden, denn die eigentlichen Visualisierungselemente liegen in öffentlichen Repositories des Codehostingdienstleisters GitHub [65]. Die Plattform bietet außerdem eine rudimentäre Programmierschnittstelle, welche das Durchsuchen und Suchen nach Web Components maschinell ermöglicht. Zur Nutzung der Plattform im konzeptuellen Modell, bietet sich an, die Programmierschnittstelle um ein Publish/Subscribe-Verfahren zu erweitern. Die Idee hierbei ist, dass sobald eine Web Component auf der Plattform WebComponents.org veröffentlicht wird, die Plattform das konzeptuelle Modell über die Veröffentlichung informiert. Die Softwarekomponente des konzeptuellen Modells wiederum prüft, ob die gerade veröffentlichte Web Component mit dem in dieser Forschungsarbeit spezifizierten Vokabular ausgezeichnet wurde. Hierzu kann der Entwickler der Web Component die Auszeichnungen in einer separaten Datei im Repository ablegen. Ist die entsprechende Web Component ausgezeichnet, kann der Visualisierungsparser als Softwarekomponente des konzeptuellen Modells einen Baum für diese Web Component aufbauen und den aufgebauten Baum

in einer Datenbank persistieren. Die Persistierung kann zur einfachen Formattierung als Entscheidungstabelle realisiert werden. Wird nun für ein semantisiertes Datendokument eine effektive Visualisierung gesucht, so kann für das Datendokument entsprechend ein Baum aufgebaut werden, und weiterhin mit allen formatiert gespeicherten Visualisierungselementbäumen auf Äquivalenz verglichen werden. Sobald eine Äquivalenz zwischen Daten und einem Visualisierungselement auftritt, kann der Datensatz mithilfe der semantisierten Auszeichnung des Visualisierungselements die Daten dem Visualisierungselement zuordnen. Hierzu wird die Web Component über den Paketmanager 'Bower' installiert, die Daten zugeordnet und das Visualisierungselement inklusive zugeordneter Daten als statisches Element per URL abrufbar gemacht. Das statische Element kann entsprechend über ein iFrame in den Anwendungskontext eingegliedert werden.

Das Anstoßen des Prozesses kann durch die Integrationskomponente per RESTful-Schnittstelle erfolgen. Die Übergabe der Integrationskomponente wäre primär das Artefakt der semantisierten Daten. Die Rückgabe wäre die URL der gefundenen Web Component beziehungsweise bei mehreren passenden gefundenen Web Components die URLs dieser.

#### 4.1.4 Integration in den Anwendungskontext

Die Integration des konzeptuellen Modells in den Anwendungskontext der Schul-Cloud wird in diesem Kapitel näher betrachtet.

Schüler analysieren im Schulunterricht als auch in der investierten Zeit für die Schule große und komplexe Datenmengen. Im Kontext der Bildung nimmt die effektive Visualisierung solcher Datenmengen für das Aneignen von neuem Wissen durch entsprechende Daten einen besonders hohen Stellenwert ein. Bilder und Grafiken als gerenderte Artefakte solcher Visualisierungen erheben beim Lernen eine substantielle Rolle, indem sie dem Lernenden dabei helfen, die bereitgestellten Informationen gemäß der in Kapitel 3.1.2 erläuterten Zielsetzungen von Visualisierungen zu verstehen. Zugleich ermöglichen die effektiven Visualisierungen im Bildungsbereich, Ideen leichter zu vermitteln als

mit rein textuellen Beschreibungen. Zahlreiche Studien belegen ein steigendes Engagement und Verständnis von Informationen beim Lernenden durch Einsatz visueller Bilder und Bewegtbilder. Ganz allgemein sorgen Design und Ästhetik so für einen tief greifenden Einfluss in die Wahrnehmung, den Lernprozess sowie in den Informationsglaubwürdigkeitsprozess des Betrachters.

Sowohl Lehrer als auch Schüler sind in der Regel nicht ausreichend geschult beziehungsweise verfügen nicht über entsprechendes Wissen, um qualitativ hochwertige Daten effektiv zu visualisieren, sodass ein entsprechend entwickeltes virtuelles Arbeitsblatt den Lernprozess unterstützt. Zusätzlich beeinflussen die durchzuführenden kognitiven Bemühungen zur Findung einer effektiven Darstellungsform maßgeblich die Effizienz von Lehrer und Schüler.

Die umzusetzende Softwarekomponente soll den beiden Akteuren *Lehrer* und *Schüler* der Schul-Cloud die Möglichkeit bieten, virtuelle Arbeitsblätter mit Nutzung multimedialer Inhalte zu erstellen. Die von den Akteuren entwickelten virtuellen Arbeitsblätter können anschließend persistent durch die Infrastruktur der Schul-Cloud abgelegt und sowohl von Lehrern als auch Schülern als Lehrmaterial genutzt werden. Exemplarisch hierfür könnte der Lehrer die virtuellen Arbeitsblätter als Lehrmaterial für seinen Schulunterricht nutzen. Schüler hingegen als Lehrmaterial für den Nachhilfeunterricht.

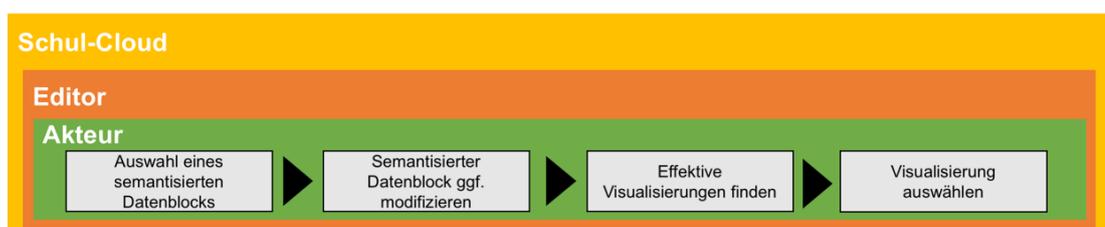


Abbildung 4.8: Der Prozess von schemantisierten Datenblöcken zur Visualisierung für virtuelle Arbeitsblätter des Anwendungskontexts der Schul-Cloud aus Sicht des Akteurs (eigene Darstellung).

Abbildung 4.8 zeigt den zu umsetzenden Prozess der Softwarekomponente zur Erzeugung eines virtuellen Arbeitsblattes. Im ersten Schritt wählt der Akteur aus einer Sammlung von semantisierten Datenblöcken, den auf dem virtuellen Arbeitsblatt darzustellenden Lehrinhalt. Der Transformationsvorgang

der Rohdaten zu semantisierten Datenblöcken ist wie in der Abgrenzung in Kapitel 1.2 beschrieben nicht Gegenstand der Softwarekomponente. Die semantisierten Datenblöcke sind in einem Materialrepository abgelegt. Der durch den Akteur ausgewählten Datenblock wird zur Modifizierung im What-You-See-Is-What-You-Get (WYSIWYG)-Editor zur Erstellung des virtuellen Arbeitsblattes angezeigt. Nach Aufforderung des Akteurs sollen alle effektiven Visualisierungselemente für den entsprechenden Datenblock gefunden werden. Die Datenattribute im semantisierten Datenblock sollen weiterhin automatisiert den Visualisierungselementattributen zugeordnet werden. Der Akteur hat damit die Möglichkeit, aus einer der vorgeschlagenen Visualisierungen auszuwählen. Die ausgewählte Visualisierung wird dann entsprechend auf dem virtuellen Arbeitsblatt angezeigt

Die nachfolgenden Anwendungsfälle definieren die Softwarekomponente. Abbildung 4.9 illustriert als Anwendungsfalldiagramm den Überblick über die entsprechenden Anwendungsfälle.

### **Semantisierte Daten auswählen**

**Auslösendes Ereignis:** Der Akteur wählt im WYSIWYG-Editor den Button *Material*.

**Verhalten im Normalfall:** Dem Akteur wird eine Vorschau aller verfügbaren semantisierte Datenmaterialien in Form von Datenblöcken aus einem definierten Materialrepository angezeigt. Der Akteur wählt einen semantisierten Datenblock, der Teil des virtuellen Arbeitsblattes und dementsprechend visualisiert werden soll. Der Datenblock wird im Editor angezeigt.

**Verhalten im Fehlerfall:** Sind keine semantisierten Datenblöcke im Materialrepository vorhanden, ist die Auswahl und Visualisierung eines Datenblocks nicht möglich. Dem Akteur wird in der Vorschau angezeigt, dass keine Materialien zur Verfügung stehen.

**Ergebnis:** Der zu visualisierende, semantisierte Datenblock wird im WYSIWYG-Editor angezeigt und kann vom Akteur modifiziert werden.

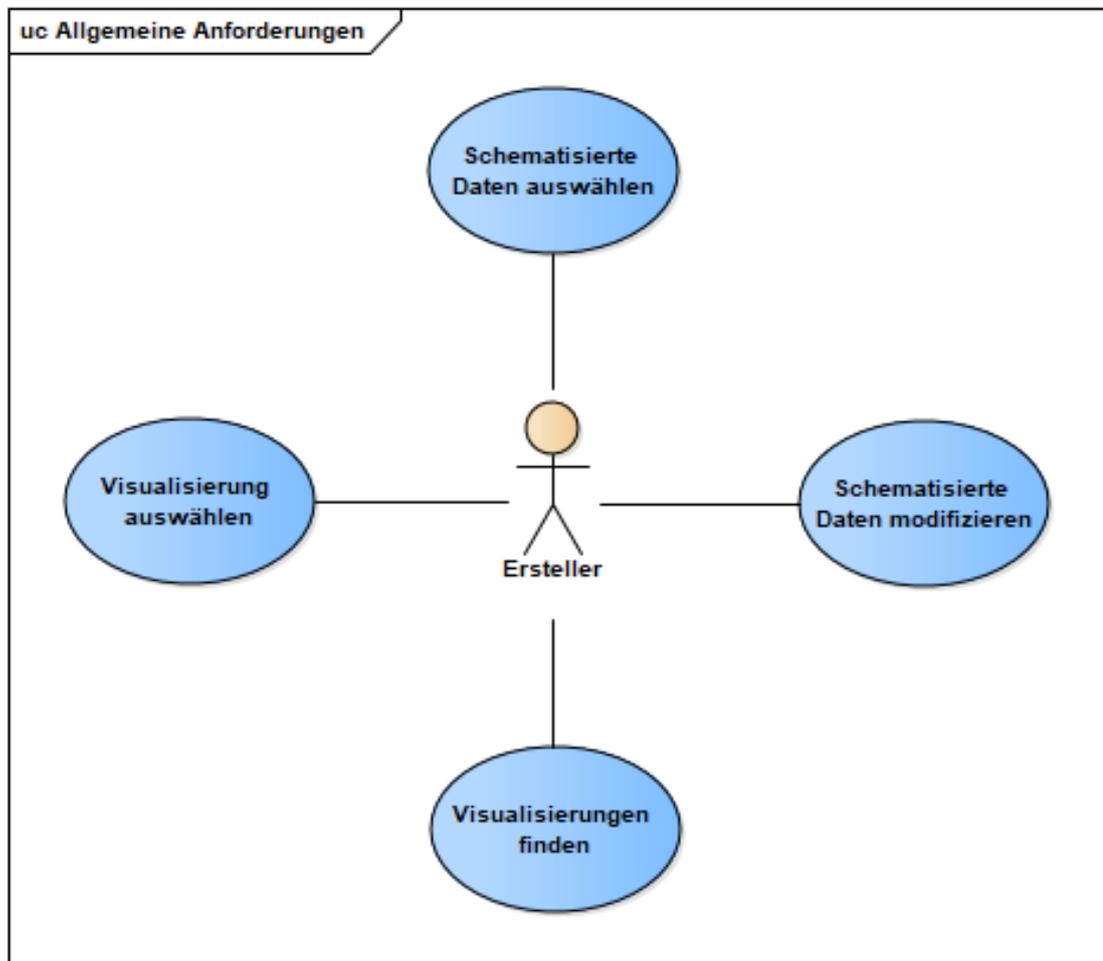


Abbildung 4.9: Anwendungsfalldiagramm für die allgemeinen Anforderungen welche an die Softwarekomponente gemäß der definierten Leitfrage gestellt wird (eigene Darstellung).

**Semantisierte Daten modifizieren**

**Auslösendes Ereignis:** Der Akteur möchte den im WYSIWYG-Editor angezeigten semantisierten Datenblock modifizieren.

**Verhalten im Normalfall:** Der Akteur modifiziert den semantisierten Datenblock durch die vom WYSIWYG-Editor standardmäßig bereitgestellten Funktionalitäten zur Datenmodifikation.

**Verhalten im Fehlerfall:** Wurde zuvor kein semantisierter Datenblock ausgewählt, ist eine Modifikation nicht möglich.

**Ergebnis:** Der zu visualisierende, semantisierte und vom Akteur modifizierte Datenblock wird im WYSIWYG-Editor angezeigt.

**Visualisierungen finden**

**Auslösendes Ereignis:** Der Akteur wählt im WYSIWYG-Editor den Button *Visualisieren*.

**Verhalten im Normalfall:** Die im WYSIWYG-Editor angezeigten und vom Akteur zuvor ausgewählten und ggf. modifizierten Datenblöcke werden über eine Programmierschnittstelle (API) zum zweiten Teil des konzeptuellen Modells gesendet. Dieser separate Service enthält die spezifische Programmlogik zur Auswahl der effektiven Visualisierungselemente auf Grundlage der gesendeten, semantisierten Datenblöcke und zur Zuordnung der Datenattribute zu den entsprechenden Visualisierungselementattributen, der ausgewählten Visualisierung. Die Rückgabe der Programmierschnittstelle sind entsprechende, effektive Visualisierungen für die gegebenen Datenblöcke. In einer Vorschau kann der Akteur alle für die Datenblöcke gefundenen Visualisierungen ansehen.

**Verhalten im Fehlerfall:** Wurde zuvor kein semantisierter Datenblock ausgewählt, ist eine Visualisierung nicht möglich. Dem Akteur wird ein entsprechender Fehler angezeigt.

Wurde bei der Modifikation, das Schema verletzt, ist eine Visualisierung ggf. nicht mehr oder nicht effektiv möglich.

Wird keine effektive Visualisierung für die semantisierten Datenblöcke gefunden, ist eine Visualisierung nicht möglich. Dem Akteur wird ein entsprechender

Fehler angezeigt.

**Ergebnis:** Alle gefundenen effektiven Visualisierungselemente mit entsprechend zugeordneten Datenattributen zu Visualisierungselementattributen werden dem Akteur in einer Vorschau angezeigt.

### **Visualisierung auswählen**

**Auslösendes Ereignis:** Der Akteur wählt eine Visualisierung aus der Vorschau aller für einen Datenblock gefundenen, effektiven Visualisierungen aus, die auf dem virtuellen Arbeitsblatt angezeigt werden soll.

**Verhalten im Normalfall:** Die ausgewählte und bereits mit Daten zugeordnete Visualisierung wird auf dem virtuellen Arbeitsblatt angezeigt.

**Verhalten im Fehlerfall:** keines

**Ergebnis:** Die mit Daten versetzte Visualisierung wird auf dem virtuellen Arbeitsblatt angezeigt.

## **4.2 Spezifikation des Semantisierungsvokabulars**

Dieses Kapitel mit seinen Unterkapiteln stellt die folgende Leitfrage in den Fokus: 'Wie muss eine Vokabularmenge zur Semantisierung spezifiziert werden, sodass Daten und grafische Repräsentationselemente semantisch beschrieben werden können und ein Auffinden von grafischen Repräsentationselementen in Abhängigkeit zum Effektivitätskriterium für entsprechende Daten als auch das Zuordnen der entsprechenden Daten zum gefundenen grafischen Repräsentationselement, gemäß dem konzeptuellen Modells ermöglicht wird?'

Zur Beantwortung, der für dieses Kapitel und entsprechenden Unterkapiteln definierte Leitfrage ist, im ersten Schritt ein grundlegendes Verständnis für die Semantik notwendig. Im Anschluss daran steht die Auswahl eines Basisschemas im Mittelpunkt zur Unterstützung der in dieser Forschungsarbeit nicht im Fokus stehenden Aspekte wie beispielsweise die grundlegende Verwendung generischer Aspekte der Semantisierung oder die spezifische Unterstützung der Domäne des E-Learnings die gemäß der Kontextbetrachtung in Kapitel

2 durch die Schul-Cloud definiert wurde. Daran anschließend wird systematisch anhand der Einflussfaktoren auf die in Kapitel 3.1.3 definierten Qualitätskriterien das Vokabular spezifiziert, sodass am Ende des Kapitels sowohl Daten als auch Visualisierungselemente semantisch so beschrieben werden können, dass die Komponente 'Discovery- und Zuordnungslogik' ein Auffinden von effektiven Visualisierungselementen für spezifische Daten und das Zuordnen dieser Daten zum Visualisierungselement ermöglichen kann. Zur Realisierung der Spezifikation des Vokabulars werden dabei für jeden Einflussfaktor generische Charakteristiken ermittelt, welche den Einflussfaktor beschreiben. Auf Grundlage dieser Charakteristiken wird entsprechendes Vokabular spezifiziert.

#### 4.2.1 Grundlagen der Semantik

Vor der Spezifikation des semantischen Vokabulars bedarf es einiger Grundlagen des Themenbereichs der Semantik, die nachfolgend in einem kleinen Exkurs erläutert werden.

Wird von der Begrifflichkeit der Semantik gesprochen, wird die 'Bedeutung' eines Objektes fokussiert. Im Kontext von Daten definiert die Semantik entsprechend die Bedeutung der Daten. In Kombination ermöglicht die Bedeutung von Daten (Semantik) mit der Struktur von Daten (Syntax) eine Repräsentation von Daten bereitzustellen, die maschinenlesbar und -interpretierbar ist. Bei Betrachtung von algorithmischer Automatisierung auf Ebene der Bedeutung durch Maschinenlesbarer und -interpretierbarer Daten im World Wide Web, wird auch von Semantic Web gesprochen. Die grundlegende Idee des Semantic Web ist hierbei die in der Regel von Menschen manuell durchgeführte Anreicherung von Inhalten im World Wide Web, insbesondere von Hypertext-Dokumenten, sodass ein Verständnis der Dokumente nicht nur von Menschen erfasst werden kann, sondern auch durch Maschinen. Es bietet damit Lösungsansätze für beispielsweise das Auffinden relevanter Informationen und die Integration von Informationen aus heterogenen Quellen und stellen damit das primäre Konzept dieser Forschungsarbeit zur Bewältigung der in der

Problemstellung definierten Problematik und Beantwortung der spezifizierten Forschungsfrage dar [69, 70, 71].

Eine besondere Rolle bei der Semantisierung von Daten nimmt der Standard Resource Description Framework (RDF) ein. Bei RDF handelt es sich hierbei um eine speziell für das Semantic Web entwickelte formale Sprache, welche die Beschreibung strukturierter Informationen in Form von meist Dokumenten ermöglicht. Ein mit RDF semantisch ausgezeichnetes Dokument enthält nach der Auszeichnung dementsprechend modelliertes Wissen einer entsprechenden Anwendungsdomäne. Das Wissen selbst ist in sogenannten Ontologien modelliert, die begrifflich äquivalent zu Wissensbasen sind. Anders als in Hypertext Markup Language (HTML) fokussiert sich RDF nicht ausschließlich auf die konkrete Repräsentation von Dokumenten, sondern fokussiert sich auf die Kombination sowie Weiterverarbeitung der ausgezeichneten Informationen. Die Auszeichnung von Informationen durch RDF wird im Framework durch Graphen realisiert, die bei RDF meist durch den sogenannten RDF-Triple beschrieben wird. Beim RDF-Triple handelt es sich um die Beschreibung einzelner Ressourcen im Graphen durch Subjekt, Prädikat und Objekt, wobei das Subjekt den Anfangspunkt, das Objekt den Endpunkt und das Prädikat die Beziehung von Subjekt und Objekt repräsentiert. Die Ressourcen im Graphen werden durch Uniform Resource Identifier (URI)s bezeichnet, da URIs die Eigenschaften besitzen, eindeutig und einmalig zu sein, sodass es keine Verwechslung von Ressourcen geben kann. Literale ermöglichen im Graphen die Beschreibung von Datenwerten. Literale bestehen hierbei aus den beiden Attributen der Wertebezeichnung und eines entsprechend für den Wert gültigen Datentyps [69, 71, 72]. Resource Description Framework Schema (RDFS) ist neben RDF eine Sprache, mit der das Vokabular spezifiziert werden kann, dass dann entsprechend zur Auszeichnung in RDF-Dokumenten verwendet wird [73].

Die Darstellung von RDF findet wie erläutert durch Graphen statt. Kapitel 4.1.2 hingegen erläuterte die Notwendigkeit zur Erzeugung von Bäumen der Daten und Visualisierungselemente um strukturelle Äquivalenzen zwischen Daten und Visualisierungselemente gemäß den ermittelten Einflussfaktoren der Qualitätskriterien von Visualisierungen zu identifizieren. Solche streng-hierarchischen

Aufbauten wie Bäume es ermöglichen sind mit Graphen nicht abzubilden. Dementsprechend sind die in RDF getätigten Aussagen, die durch gerichtete Graphen Beziehungen zwischen den Aussagen modellieren alle gleichwertig [72]. Die Realisierung zum Aufbau von Bäumen muss entsprechend durch Semantisierungsvokabular realisiert werden, mit dem entsprechende Hierarchien auszeichnungstechnisch aufgebaut werden können.

Die nachfolgenden Spezifikationen der einzelnen Vokabularien werden auf Basis von RDF realisiert. Die Repräsentation dieser Vokabularien wird in den folgenden Kapiteln in Tabellenform angegeben.

Die einzelnen Tabellen enthalten dabei die folgende Mindestmenge an Attributen zur Beschreibung der Vokabularien:

- a) **Name:** Token dass an die URI des Namespaces angehängt wird, um das Vokabularium zu erzeugen.
- b) **Label:** Für den Begriff zugewiesener und von Menschen verständlicher Name des Vokabulariums.
- c) **URI:** Der Uniform Resource Identifier zur eindeutigen Identifikation des Vokabulariums.
- d) **Beschreibung:** Aussage, die das Konzept und die Art des Vokabulariums beschreibt.
- e) **Typ des Ausdrucks:** Typ des Vokabulariums - Klasse, Eigenschaft oder Member.

Die folgenden Attribute können gegebenenfalls zusätzliche Informationen zum Vokabularium definieren:

- a) **Ist Unterklasse von:** Klasse, zu welcher das beschriebene Vokabularium eine Unterklasse ist.
- b) **Inkludiert in Klasse:** Klasse, zu welcher das beschriebene Vokabularium zugehörig ist.

- c) **Erwarteter Typ:** Datentyp, welcher das beschriebene Vokabularium erwartet.

#### 4.2.2 Auswahl des Basisvokabulars

Die Zielsetzung dieses Kapitels ist die Auswahl einer Menge von Basisvokabularien, anhand deren das spezifische zu definierende Vokabular zur Beschreibung der in Kapitel 3.1.3 identifizierten Einflussfaktoren der Qualitätskriterien von Visualisierungen ermöglicht wird. Auf Basis des spezifisch zu entwickelnden Vokabulars sollen entsprechend Daten und Visualisierungselemente semantisch ausgezeichnet werden können, sodass diese von der 'Discovery- und Bindungslogik' als maschinelle Einheit gelesen und interpretiert werden können und ein Auffinden effektiver Visualisierungselemente für definierte Daten und das Zuordnen der definierter Daten zum Visualisierungselement ermöglicht wird.

Im Rahmen des Semantic Webs existieren bereits Vielzahlen an Ontologien, die eine grundlegende Basis für weiter zu spezifizierende Vokabularien schaffen. Der Begriff der Ontologie definiert hierbei wie bereits in Kapitel 4.2.1 beschrieben, eine definierte Menge an Begrifflichkeiten aus einem definierten Gegenstandsbereich sowie Beziehungen der Begrifflichkeiten des Gegenstandsbereichs zueinander. Die Gesamtheit der Begrifflichkeitsmenge, die für den Gegenstandsbereich definiert wurde, wird entsprechend als Vokabular bezeichnet.

Die Integration des in dieser Forschungsarbeit entwickelten konzeptuellen Modells sowie die Integration in den Anwendungskontext der Schul-Cloud definiert bereits den Gegenstandsbereich der Domäne des E-Learnings. Die Abbildung der Einflussfaktoren identifizierter Qualitätskriterien von Visualisierungen zu einer semantischen Vokabularmenge ist daher durch die Unterstützung eines Basisvokabulars aus dem Gegenstandsbereich des E-Learnings, insbesondere durch die Abbildung von Daten, als sinnvoll zu erachten. Die generische Natur der Einflussfaktoren definierter Qualitätskriterien von Visualisierungen und die daraus ableitbaren Charakteristiken bedürfen zusätzlich zur

spezifischen Vokabularmenge der Domäne des E-Learnings, die Bereitstellung einer generischer Vokabularmenge. Die Auswahl von Vokabularmengen aus bereits existierenden Ontologien weißt den Vorteil der Bewährtheit sowie der guten Testabdeckung der Vokabularien aus entnommenen Ontologien auf. Eine Auswahl von Basisvokabularien aus bewährten Ontologien sollte zur vollumgänglichen Unterstützung des konzeptuellen Modells entsprechend den spezifischen Domänenaspekt des E-Learnings als auch den generischen Aspekt der Charakteristiken definierter Einflussfaktoren der Qualitätskriterien von Visualisierungen entsprechen.

Zur Auswahl geeigneter Vokabularmengen definiert die nachfolgende Auflistung einen Satz an Anforderungen, welche zur Auswahl entsprechender spezifischer als auch generischer Vokabularmengen gestellt wird:

1. Die Vokabularmenge wird von einer nicht unerheblichen Anzahl von Anwendern akzeptiert und verwendet, sodass der Grad an Schwierigkeit durch gegebenenfalls entstehende Auszeichnungsproblematiken bei Daten, durch die gewählte Vokabularmenge gering bis nicht existent ist.
2. Die Vokabularmenge ist leichtgewichtig und deklariert eine geringe Anzahl an Abhängigkeiten zueinander, sodass eine Auszeichnung von Informationen durch die gewählte Vokabularmenge nicht ausschließlich Experten der semantischen Auszeichnung vorbehalten ist.
3. Die gewählte Vokabularmenge ist durch einfache Mittel um weitere Begrifflichkeiten erweiterbar.

Eine besonders verbreitete Ontologie zur Semantisierung von Informationen im Rahmen des Semantic Webs ist die der Initiative [schema.org](http://schema.org). Bei dieser Initiative handelt es sich um den Verbund der drei größten Suchmaschinen der Welt Google, Bing und Yahoo mit der Zielsetzung entsprechender Bereitstellung eines leicht-verständlichen Vokabularsatzes zur Semantisierung einer Vielzahl an Webinhalten. Der primäre Fokus von [schema.org](http://schema.org) ist damit die Einfachheit des Vokabularsatzes, sodass die Auszeichnung mit den spezifizierten Vokabularien auch Nicht-Experten semantischer Auszeichnung ermöglicht wird. Die Initiative um [schema.org](http://schema.org) ist in der Domäne der Ontologien als eine

Ausnahme zu erachten. Eine Vielzahl existierender Ontologien definierten Vokabularien für einen klar abgegrenzten Gegenstandsbereich. Die Vokabularien von schema.org hingegen sind als allgemeingültige gesamtbeschreibende und somit als generische Vokabularmenge zu erachten. Die Initiative spricht sich selbst gegen den Begriff einer weltbeschreibenden Ontologie aus. Dennoch ist nicht von der Hand zu weisen, dass schema.org eine Vielzahl an weltlichen Gegenstandsbereichen abdeckt durch die Spezifikation entsprechender Vokabularien. In Abgrenzung zu anderen Ontologien besitzt die Ontologie von schema.org einen weiteren Ausnahmefaktor. Bei diesem Ausnahmefaktor handelt es sich um die Modularität der Ontologie. Hierzu besteht die Ontologie der Initiative aus einer Vielzahl atomisierter Bestandteile, die in deren Gesamtheit die Ontologie von schema.org repräsentiert. Die einzelnen atomisierten Bestandteile, welche im Aufbau in der schema.org Ontologie resultieren, werden von der Initiative 'Schema' genannt. Entsprechend zur Atomisierung ist das der Ontologie zugrunde liegende Datenmodell generisch und strukturell abgeleitet aus RDF. In Anlehnung an die Beschreibung der Grundlagen von Semantik in Kapitel 4.2.1, ist ein RDF-Schema existent durch die Beschreibung eines Satzes an Klassen, welche durch ausgewählte Eigenschaften repräsentiert werden und damit entsprechend die Ontologie beschreibend repräsentieren. Die Einzelnen durch ihre Atomisierung existierenden Schemata der schema.org Initiative sind in der Ontologiedefinition als mehrfach vererbte Typenhierarchie arrangiert. Der Begriff des Typs und der Klasse wird im Rahmen der Semantisierung entsprechend synonym verwendet. Jeder Typ kann innerhalb der Typenhierarchie von schema.org, Unterklasse mehrerer anderer Typen sein. Mit der Entsprechung zu RDF, besitzt jeder Typ innerhalb der Typenhierarchie eine Menge an Eigenschaften, deren Datentypen wiederum aus einem oder mehreren Typen bestehen dürfen. Ein Schema innerhalb der Initiative besteht aus mindestens einem Typ bis zu einer unbegrenzten Anzahl an Typen in Kombination. Startpunkt und damit Wurzelknoten des Typenhierarchiebaums von schema.org ist der Typ 'Thing' [74, 75, 76, 77, 78]. Entsprechend sind alle anderen in schema.org definierten Typen abgeleitet aus dem Typ 'Thing' [74]. Der Wurzelknoten repräsentiert die generische Form aller mit Semantik auszuzeichnenden Informationen. Exemplarisch für diese generische Form sind beim Typ 'Thing' [74] die beschreibenden Eigenschaften: Name des Ele-

ments, Beschreibung des Elements und URL des Elements. Der modulare Aufbau der schema.org Ontologie durch Atomisierung selbiger ermöglicht den Vorteil zur einfachen Erweiterung der Ontologie um weitere Vokabularien in atomisierten Schemata. Das Kernvokabular der Initiative besteht derzeit aus 597 Typen mit insgesamt 867 Eigenschaften. Eine Vielzahl an Organisationen nutzen die schema.org Ontologie und erweitern das Kernvokabular entsprechend um weitere Vokabularien für einen spezifischen Gegenstandsbereich [75, 76, 77, 78].

Die Ontologie der Initiative schema.org ist aufgrund der generischen Natur bester Kandidat um die erwähnte Abbildung der generischen Charakteristiken definierter Einflussfaktoren der Qualitätskriterien von Visualisierungen abzubilden. Aufgrund des generischen Modells und der einfachen Erweiterung des Vokabulars der Initiative, sowie die Erfüllung aller definierten Anforderungen, wird schema.org für diese Forschungsarbeit als Basisvokabular zur Abbildung generischer Informationen verwendet.

Zur Abbildung des spezifischen Gegenstandsbereichs der Domäne des E-Learnings ist die Kernvokabularmenge der schema.org Ontologie nicht ausreichend. Es bedarf somit die Auswahl einer zusätzlichen Vokabularmenge zur Abbildung der Domäne des E-Learnings. Hierbei soll die Auswahl einer geeigneten Ontologie für den Gegenstandsbereich zum einen die definierten Anforderungen an ein Basisvokabular erfüllen und zum anderen nach Möglichkeit eine direkte Integration in die schema.org Ontologie ermöglichen, sodass der Aufwand zur Auszeichnung von Informationen auf Basis der schema.org Auszeichnungen getätigt werden können.

Nach Analyse einer Menge von Ontologien wurden die nachfolgenden Ontologien des Gegenstandsbereichs des E-Learnings zur Evaluierung als Basisvokabular für den Gegenstandsbereich des E-Learnings identifiziert. Die identifizierten Ontologien erfüllen bereits die Anforderung der hinreichenden Akzeptanz und Verbreitung von Standards [79]:

**Learning Object Metadata (LOM)** ist ein Standard aus der Domäne des E-Learnings, der definiert welche Aspekte eines Lernobjektes beschrieben wer-

den und welches Vokabular für die Beschreibung genutzt werden sollte. Ein Lernobjekt ist bei LOM als Entität definiert, welches digital oder analog für das Lernen beziehungsweise für die Bildung im Allgemeinen genutzt wird. Zusätzlich zur Aspekte- und Vokabulardeklaration von Lernobjekten, beschreibt LOM wie das hierfür genutzte Datenmodell geändert werden kann, um eine Erweiterung des Datenmodells herbeizuführen. Weiterhin beschreibt der Standard, wie Regeln genutzt werden, um das Datenmodell einzuschränken und so eine strikte Spezifikation anzuweisen. LOM ermöglicht mit seinem strukturellen Design die Erzeugung gut strukturierter Beschreibungen von Lerninhalten, welche das Discovery, Lokalisieren, Evaluieren und Erwerben von Lerninhalten für Studenten, Lehrer und automatisierte Softwareprozesse vereinfacht. Neben der Akzeptanz und Verbreitung von LOM erfüllt der Standard auch die Anforderung der einfachen Vokabularerweiterung. Die Anforderung der Leichtigkeit des von LOM spezifizierten Vokabulars selbst wird durch den Standard nicht erfüllt, genau so wenig wie das optionale Evaluierungskriterium der Integration des LOM Vokabulars in die schema.org Ontologie [80, 81].

**Dublin Core** ist ein Standard, der die Auszeichnung von Lerninhalten auf Metadatenebene ermöglicht. Die ursprüngliche Entwicklung von Dublin Core verfolgte das Ziel eines besseren Discoveries von digitalen Inhalten jeglicher Art. Das Kernvokabular von Dublin Core besteht hierbei aus fünfzehn Kerneigenschaften, welche die Charakteristiken besitzen weitgehendst generisch zu sein und eine breite Fächerung zulassen. Das spezifizierte Vokabular auf Metadatenebene eignet sich somit nicht ausschließlich zur Beschreibung der Domäne des E-Learnings, sondern ermöglicht ebenfalls die Auszeichnung generischer Szenarien außerhalb des expliziten Bereichs des E-Learnings. Für die explizite Auszeichnung von Dokumenten aus pädagogischer Sicht spezifiziert Dublin Core keine Eigenschaftsattribute. Dennoch ermöglicht der Standard eine einfache Vokabularerweiterung, welche durch die Definition eigener Typen sowie Eigenschaften eine pädagogische Dokumentenauszeichnung ermöglichen. Dublin Core erfüllt neben der Anforderung der Akzeptanz und Verbreitung sowie der einfachen Vokabularerweiterung zusätzlich die Leichtigkeit des Vokabulars, durch die leicht verständliche Spezifikation der Elemente

und die lose Kopplung zwischen den Elementen. Das optionale Evaluierungskriterium der Integration des Vokabulars von Dublin Core in die schema.org Ontologie wird von Dublin Core nicht erfüllt [80, 82, 83, 84].

**Sharable Content Object Reference Model (SCORM)** ist ein Standard aus der Domäne des E-Learnings, der sich auf die Spezifikation sowie vereinfachte Portabilität von Lerninhalten zwischen Lernplattformen fokussiert. Hierzu integriert SCORM das Metadatenmodell des bereits eingeführten Standards LOM und unterstützt das Abrufen jeglicher Lernobjekte in verschiedensten Konstellationen. Die Integration des Metadatenmodells von LOM ist nicht die einzige Integration nach SCORM. Der Standard integriert eine Reihe verschiedener Standards verschiedenster Organisationen. SCORM erfüllt neben der Anforderung der Akzeptanz und Verbreitung zusätzlich die einfache Vokabularerweiterung. Durch die Integration verschiedenster Standards ist die Leichtgewichtigkeit des Vokabulars gemäß der definierten Anforderung nicht gegeben. Ebenfalls unterstützt SCORM keine direkte Integration in die schema.org Ontologie [85, 86].

**IMS Question and Test Interoperability (QTI)** ist ein Standard aus der Domäne des E-Learnings, der sich auf die Spezifikation von Assessments fokussiert. Die Form der Assessments und damit dem Datenmodell von QTI ist nach Fragen und Antworten strukturiert. Weiterhin ermöglicht QTI die Deklaration von Triggern, die das eventbasierte Reagieren von entsprechenden Antworten auf Fragen ermöglichen. QTI erweitert die eigene Vokabularmenge durch die Metadatenvokabularmenge von LOM. Somit erfüllt QTI wie bereits LOM und SCORM die Anforderung der Akzeptanz und Verbreitung sowie der einfachen Vokabularerweiterung, nicht aber die Leichtgewichtigkeit des Vokabulars und die optionale Integration nach schema.org [87, 88].

**Learning Resource Metadata Initiative (LRMI)** ist ein Standard aus der Domäne des E-Learnings mit dem Fokus auf Suchmöglichkeiten von Bindungsinhalten für Lehrende und Lernende durch spezielle Suchmaschinen und Re-

source Discovery Services. LRMI ist einer der jüngsten Ontologiestandards in der Domäne des E-Learnings und setzt den Fokus nicht wie viele andere Standards in der Domäne auf die Katalogisierung und Klassifizierung der Lehrmaterialien. Primärer Ansatz der Initiative ist die leichtgewichtige Vokabularerweiterung und direkte Integration in die schema.org Ontologie. Neben der Anforderungserfüllung von Akzeptanz und Verbreitung erfüllt LRMI somit auch alle anderen definierten Anforderungen und zusätzlich das optionale Evaluationskriterium der Integration in die schema.org Ontologie. Die direkte Integration von LRMI in die Ontologie der schema.org Initiative ermöglicht LRMI zahlreiche Vorteile. Abgesehen vom Vorteil selbst, Bestandteil der schema.org Ontologie zu sein, kann sich das Standardisierungsgremium von LRMI vollständig und einzig auf die Spezifikation von Bildungsmerkmalen konzentrieren. Der Grund hierfür ist, dass die Auszeichnung der Ressourcen in erster Linie mit der Auszeichnungssyntax und dem Kernvokabular von schema.org realisiert werden kann. Ein anderer Vorteil ist der, dass die LRMI Initiative durch die Integration von allen schema.org spezifischen Vorteilen in der Gesamtheit profitiert, wie beispielsweise die von der schema.org Community bereitgestellten Entwicklungstools [78, 79].

<b>Anforderung / Standard</b>	LOM	Dublin Core	SCORM	QTI	LRMI
Akzeptiert & Verbreitet					
Leichtgewichtiges Vokabular					
Einfache Vokabular-Erweiterung					
optional: schema.org-Integration					

Abbildung 4.10: Evaluation von Ontologien, welche für die Domäne des E-Learning grundsätzlich geeignet sind und damit Kandidat zur Auswahl von Basisvokabular sind (eigene Darstellung).

Abbildung 4.10 illustriert zusammenfassend alle identifizierten und evaluierten Ontologien, die sich als Basisvokabular für die Domäne des E-Learnings eignen. Hierbei stellt die Evaluationstabelle der Abbildung den Ontologien die definierten Anforderungen gegenüber. Die mit der Farbe grün markierten Felder

zeigen an, dass die entsprechende Ontologie die entsprechende Anforderung erfüllt. Analog hierzu zeigen die Felder, welche mit der Farbe rot markiert sind, dass die entsprechende Ontologie die entsprechende Anforderung nicht erfüllt. Aus der Abbildung lässt sich lesen, dass sowohl der Standard Dublin Core als auch der Standard LRMI die einzigen Kandidaten sind, welche die nicht-optionalen Anforderungen vollumgänglich erfüllen. LRMI erfüllt zusätzlich das optionale Evaluationskriterium der Integration in die schema.org Ontologie.

Durch die Erfüllung aller nicht-optionalen Anforderungen und zusätzlich die Abdeckung der optionalen Anforderung, wird LRMI in dieser Forschungsarbeit als Standard zur Abbildung der domänenspezifischen Daten gewählt, welche nicht mit dem generischen Vokabular abgebildet werden können. Als weitere Begründung zur Wahl von LRMI als Basisvokabular für domänenspezifische Daten, liegt auf dem Fokus von LRMI. Der Fokus von LRMI auf der Katalogisierung und Klassifizierung von Bildungsinhalten ermöglicht eine weit- aus freiere Auszeichnung von Informationen im Bereich der Bildung, als es mit den vielen anderen streng-typisierten Ontologiestandards möglich wäre. Dies macht eine einfache Adoption der generisch durch schema.org Vokabularien und spezifischen durch LRMI Vokabularien ausgezeichneten Informationen in andere Kontextszenarien und Softwareanwendungen ohne größere Anpassungen vorzunehmen möglich. Eine Anpassung von beispielsweise Vokabularparser auf spezifische Kontextszenarien beziehungsweise Softwareanwendungen entfällt.

Abbildung 4.11 illustriert die von LRMI spezifizierten Typen und Eigenschaften (mit blauer Farbe markiert) und die Integration in die schema.org Typenhierarchie (mit roter Farbe markiert). LRMI spezifiziert zwei eigene Typen: 'EducationalAudience' [89, 90] und 'AlignmentObject' [90, 91] mit entsprechenden Eigenschaften. Integriert sind beide Typen in der schema.org Ontologiehierarchie. Der von LRMI spezifizierte Typ 'EducationalAudience' [89, 90] ist hierbei abgeleitet aus dem schema.org Typ 'Audience' [92]. Dieser wiederum ist abgeleitet aus weiteren Typen. Der LRMI Typ 'AlignmentObject' [90, 91] hingegen ist abgeleitet aus dem Typ 'Intangible' [93], welcher wiederum ebenfalls von anderen Typen abgeleitet ist. Den größten spezifizierten Anteil durch LRMI

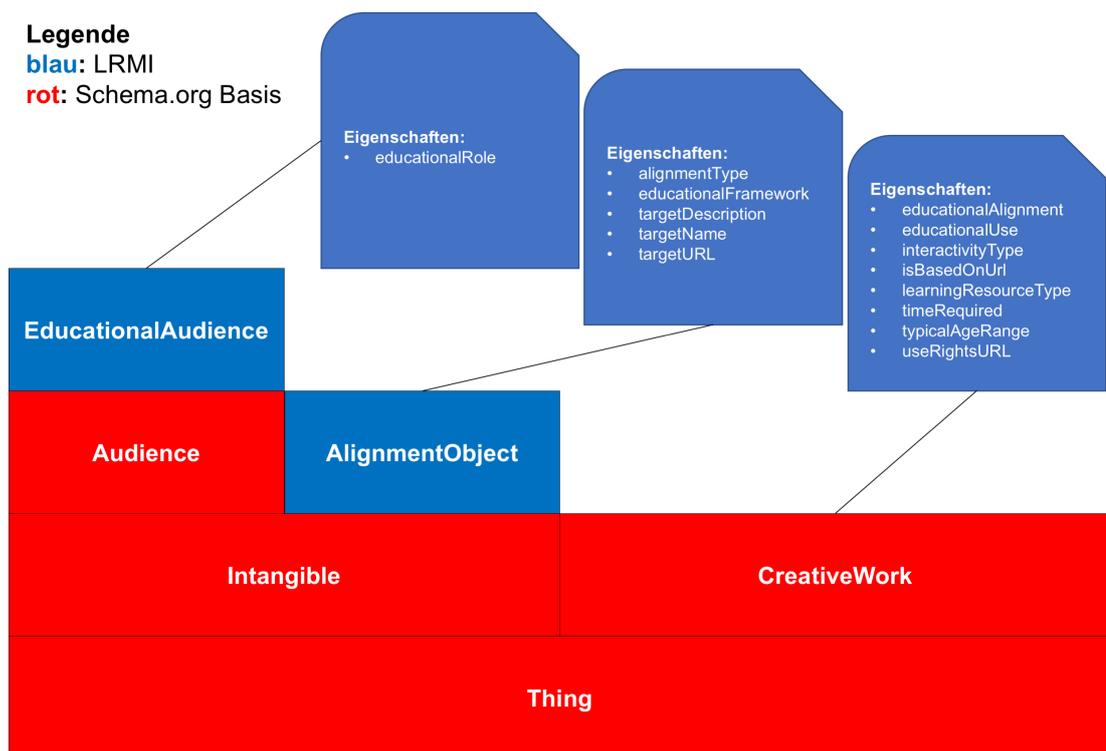


Abbildung 4.11: Die Typen und zugehörigen Eigenschaften, die durch LRMI spezifiziert wurden (mit der Farbe blau markiert) sowie die hierarchische Integration von LRMI spezifizierten Eigenschaften in die schema.org Basistypenhierarchie (mit der Farbe rot markiert). Der schema.org Basisvokabulartyp 'CreativeWork' wird durch LRMI zusätzlich um Eigenschaften erweitert (eigene Darstellung).

fällt auf Eigenschaften, welche den direkten Typ 'CreativeWork' [94] der schema.org Ontologie erweitern.

#### 4.2.3 Art und Struktur der Daten

Die im Fokus stehende Komponente bei grafischen Repräsentationen sind die Daten. Die Relevanz der Art und Struktur von Daten zum Auffinden eines effektiven Visualisierungselements für definierte Daten und die Zuordnung dieser Daten zum Visualisierungselement erläuterten bereits die Beschreibungen des Kapitels 4.1.2. In selbigem Kapitel wurde ebenfalls erläutert, dass die Klassifikation von Daten durch semantisch beschriebene Datencharakteristiken vorgenommen werden kann, sodass eine Maschinenlesbare und -interpretierbare Datenvariante existiert [59]. Hierzu wurden in Kapitel 4.1.2 folgende Datencharakteristiken identifiziert: Art der Datenwerte, Dimensionalität des Datensatzes, gültiger Wertebereich sowie Abhängigkeiten der Daten untereinander.

Das Charakteristikum 'Art der Datenwerte' kann wiederum generisch wie folgt charakterisiert werden:

- a) Skalare Daten
- b) Nominale Daten
- c) Ordinale Daten
- d) Kontinuierliche Daten
- e) Vektorielle Daten
- f) Tensorielle Daten

Das Auffinden und Zuordnen durch die 'Discovery- und Bindungslogik' ist in diesem Zusammenhang gemäß dem Effektivitätskriterium nur dann möglich, wenn die einzelnen Datencharakteristiken von der Algorithmik unterschieden werden können. Die Charakteristiken Art der Datenwerte, Dimensionalität des Datensatzes und Abhängigkeiten der Daten untereinander bilden hierbei den Verbund, der eine Hierarchisierung von Informationen ermöglicht. Durch die

Identifikation der Hierarchisierung durch Semantisierung fokussiert beschriebener Charakteristiken ist die Aufstellung der Bäume für die Daten als auch für Visualisierungselemente möglich.

Aufgrund der Abbildung von Dokumentenauszeichnungen in RDF als Graph, wie Kapitel 4.2.1 beschreibt, besteht die Notwendigkeit zur Spezifikation von Typen die eine Hierarchisierung durch Abbildung beschriebener Charakteristiken ermöglichen. Auf Basis der Spezifikationsdefinition soll dies durch die Deklaration eines generischen Typs realisiert werden, der als eine Art Container fungiert. Zusammenhängende Daten eines Dokuments werden durch Auszeichnung des Typs entsprechend dem Container zugeordnet.

Betrachtet man exemplarisch das Beispiel aus Abbildung 4.4 des Kapitels 4.1.2 würde somit jede Sektion des Dokuments einem Container zugeordnet werden. Jede Sektion, die jeweils einem eigenständigen Container zugeordnet wurde, bildet entsprechend eine Einheit und repräsentiert die Zugehörigkeit der Daten innerhalb des Containers in Abgrenzung zu Daten außerhalb des Containers. Die Abbildung des Charakteristikums der Dimensionalität wiederum kann durch die Verschachtelung mehrerer Container ineinander repräsentiert werden. Die Auszeichnung durch den Typen des Containers ermöglicht dabei gleichwohl die Repräsentation des Charakteristikums 'Art der Datenwerte'. So kann der Container selbst als tensorielle oder vektorielle Dateneinheit repräsentiert werden, während die spezifizierten Daten innerhalb des Containers ebenfalls große Dateneinheiten repräsentieren aber auch kleinste Einheiten auf exemplarischer Ebene der skalaren Daten.

Betrachtet man erneut das Beispiel aus Abbildung 4.4 würde Sektion A entsprechend jeweils beide Bestandteile textuelle Überschrift und textuelle Beschreibung durch den in schema.org deklarierten primitiven Datentyp 'Text' repräsentiert und auf höherer Ebene zur Hierarchisierung einem Container Element zugeordnet werden. Der Container ist mit dem Zeitpunkt der Zuordnung anderer Typen selbst komplexer Typ, der aus einem Satz primitiver Typen die nicht weiter zerlegbar sind oder aber durch weitere komplexe Typen besteht. Die Identifikation der Dimension eines Datensatzes ist hierbei abhängig vom Betrachtungspunkt. Wird das Dokumentenbeispiel aus Abbildung 4.4 vollstän-

dig betrachtet, so weißt das Dokument bis zu drei Dimensionen auf. Die erste Dimension repräsentiert die einzelnen Sektionen. Bei Betrachtung der Sektionen A bis einschließlich Sektion D offenbart sich eine zweite Dimension, die jeweils durch eine textuelle Überschrift sowie eine textuelle Beschreibung beziehungsweise bei Sektion D durch ein multimediales Element in Form eines Videos repräsentiert wird. Die Sektionen E und G weisen eine dritte Dimensionalität durch die erneute Untergliederung der Sektion auf, in textuelle Überschrift, Icon beziehungsweise Bullet Point sowie textuelle Beschreibung. Sektion F hingegen besteht wiederum aus nur einer Dimension. Wird der Betrachtungspunkt auf das Dokument verändert, wie es bei der Justierung der Zielsetzung, beschrieben in Kapitel 4.1.2 vorkommen kann, so verändert sich auch die grafisch zu repräsentierende Dimensionalität von Daten. In Anlehnung an das Beispiel der Abbildung bei Betrachtung einer einzelnen Sektion als Einstiegspunkt, beispielsweise Sektion A des Beispiels, so besteht die Dimension des Datensatzes nur noch aus einer Dimension statt zwei. Repräsentiert durch die textuelle Überschrift sowie die textuelle Beschreibung.

Der hier beschriebene Container wird durch den Typ 'Visualization' spezifiziert, der in Tabelle 4.1 formalisiert ist.

Tabelle 4.1: Zur Semantisierung spezifizierter RDF Klassenvokabularbegriff 'Visualization'

Name	Visualization
Label	Visualization
URI	<a href="http://ts.digital/msc/terms/Visualization">http://ts.digital/msc/terms/ Visualization</a>
Beschreibung	Buendelt Daten zu einer moeglichen Visualisierung
Typ des Ausdrucks	Klasse
Ist Unterklasse von	<a href="http://schema.org/Thing">http://schema.org/Thing</a>

Durch die spezifizierte und in Tabelle 4.2 formalisierte Eigenschaft 'preferredVisualization' kann ein Element und alle darin enthaltenen Unterelemente einer

Containerklasse des Typs 'Visualization' zugeordnet werden, um damit die Zusammengehörigkeit von Element und Unterelementen zu definieren.

Tabelle 4.2: Zur Semantisierung spezifizierter RDF Eigenschaftsvokabularbegriff 'preferredVisualization'

Name	preferredVisualization
Label	Preferred Visualization
URI	<a href="http://ts.digital/msc/terms/preferredVisualization">http://ts.digital/msc/terms/preferredVisualization</a>
Beschreibung	Spezifiziert ein Element als Visualisierung durch Angabe einer Instanz des Typs Visualization
Typ des Ausdrucks	Eigenschaft
Inkludiert in Klasse	<a href="http://schema.org/Thing">http://schema.org/Thing</a>
Erwarteter Typ	<a href="http://ts.digital/msc/terms/Visualization">http://ts.digital/msc/terms/Visualization</a>

Um eine Elementegruppe, die durch die Containerklasse des Typs 'Visualization' ausgezeichnet wurde neben der Visualisierung anderen semantischen Annotationen zuzuordnen, wurde die Eigenschaft 'hasPart' in Tabelle 4.3 formalisiert.

Tabelle 4.3: Zur Semantisierung spezifizierter RDF Eigenschaftsvokabularbegriff 'hasPart'

Name	hasPart
Label	Has part
URI	<a href="http://ts.digital/msc/terms/hasPart">http://ts.digital/msc/terms/hasPart</a>
Beschreibung	Spezifiziert welches Element Teil einer Instanz des Typs Visualization ist

*Fortsetzung auf nächster Seite*

Tabelle 4.3 – Fortgeführt von vorangegangener Seite

Typ des Ausdrucks	Eigenschaft
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> Visualization
Erwarteter Typ	<a href="http://schema.org/Thing">http://schema.org/Thing</a>

Weiterhin wird die Eigenschaft 'isPartOf' spezifiziert und in Tabelle 4.4 formalisiert, mit der ein einzelnes Element einer Containerklasse des Typs 'Visualization' zugeordnet werden kann, um die Zusammengehörigkeit des annotierten Elements mit anderen annotierten oder in der Instanz des Typs 'Visualization' befindlichen Elementen zu repräsentieren.

Tabelle 4.4: Zur Semantisierung spezifizierter RDF Eigenschaftsvokabularbegriff 'isPartOf'

Name	isPartOf
Label	Is part of
URI	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> isPartOf
Beschreibung	Spezifiziert ein Element als Teil einer Visualisierungsinstanz des Typs Visualization
Typ des Ausdrucks	Eigenschaft
Inkludiert in Klasse	<a href="http://schema.org/Thing">http://schema.org/Thing</a>
Erwarteter Typ	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> Visualization

Die spezifizierte und in Tabelle 4.5 formalisierte Eigenschaft 'isNotPart' ermöglicht ein Element innerhalb der Containerklasse des Typs 'Visualization' explizit als nicht Teil der Visualisierung auszuzeichnen.

Tabelle 4.5: Zur Semantisierung spezifizierter RDF Eigenschaftsvokabularbe-griff 'isNotPart'

Name	isNotPart
Label	Is not part
URI	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> isNot-Part
Beschreibung	Spezifiziert welches Element inner-halb des Typs Visualization nicht Teil der Instanz des Typs Visualiza-tion ist
Typ des Ausdrucks	Eigenschaft
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> Visuali-zation
Erwarteter Typ	<a href="http://schema.org/Thing">http://schema.org/Thing</a>

Zur weiteren Unterscheidung der Charakteristiken, die im Charakteristikum 'Art der Datenwerte' untergliedert sind, dient die von schema.org eingesetzte Typenhierarchie für Datentypen, die neben der Typenhierarchie der Semantisierungsklassen parallel durch die Initiative spezifiziert wurde.

Entsprechend der schema.org Typenhierarchie für Datentypen können skalarere Datenwerte durch die nachfolgenden schema.org Datentypen repräsentiert werden. Eine gesonderte Spezifikation von Datentypen zur Darstellung skalarer Daten ist somit nicht notwendig.

- a) Boolean für Wahrheitswerte (wahr und unwahr)
- b) Float für Fließkommazahlen
- c) Integer für Ganzzahlen
- d) Text für einzelne Zeichen oder Zeichenketten
- e) Date für die Darstellung eines Datums

- f) DateTime für die Darstellung eines Datums inklusive der Tagesuhrzeit
- g) URL für die Darstellung einer URL
- h) Time für die Darstellung einer Uhrzeit

Die Repräsentation nominaler Daten kann durch Typen erfolgen, die eine Merkmalausprägung ermöglichen, so beispielsweise der Datentyp Boolean auf unterster Ebene im schema.org Basisvokabular. Auf höherer Ebene kann die Repräsentation nominaler Daten durch den Typ Enumeration erfolgen, der ebenfalls im schema.org Grundvokabular verankert ist. Eine Vielzahl weiterer Typen spezifischer Natur im schema.org Grundvokabular ermöglichen ebenfalls eine Merkmalausprägung zur Repräsentation nominaler Daten. Eine gesonderte Spezifikation von Datentypen zur Darstellung nominaler Daten ist somit nicht notwendig. Bei der Repräsentation ordinaler Daten verhält sich die Auszeichnung auf ähnliche Weise wie bei nominalen Daten, mit der Unterscheidung des Ordnungskriteriums bei entsprechender Merkmalausprägung. Eine Sortierung kann durch den Typ 'ItemList' ?? des schema.org Vokabulars erreicht werden. Entsprechend ist für ordinale Daten eine Erweiterung des Vokabulars ebenfalls nicht notwendig. Vektorielle sowie tensorielle Daten können jeweils durch die unterschiedlichen Typen des schema.org Vokabulars sowie durch alle möglichen Erweiterungen - so auch durch Typen von LRMI - genutzt werden.

#### 4.2.4 Zielerreichung des Anwenders

Bereits in Kapitel 4.1.2 wurde die Relevanz der Zielerreichung des Anwenders für den strukturellen Aufbau in Form eines Baums von Daten und Visualisierungselementen erläutert. Die Zielerreichung des Anwenders ist maßgeblich dafür verantwortlich, welche Datenattribute Teil der Visualisierung werden und weitergehend wie diese Datenattribute dem Visualisierungselement zugeordnet werden.

Eine ausgiebige Analyse von Zielsetzungen im Kontext der Domäne des E-Learnings und den Vergleich mit Zielsetzungen anderer Domänen im Gesamtkontext der Visualisierung zeigte, dass die Zielerreichung des Anwenders hin-

sichtlich der Visualisierung durch zehn generische Zielsetzungen repräsentiert werden kann. Durch die Beschränkung der Zielerreichung auf zehn Begrifflichkeiten bietet sich zur Spezifikation des Vokabulars eine Enumeration an. Diese strenge Typisierung verlangt, dass bei Auszeichnung der Eigenschaft zur Beschreibung der Zielerreichung der Wert der Eigenschaft aus der definierten Menge an Werten stammen muss, welche in der Enumeration deklariert wurde. Bei der Enumeration selbst handelt es sich um eine ungeordnete Liste der zehn definierten Visualisierungsziele.

Bei der Visualisierung sind zur Zielerreichung immer mindestens zwei Datenobjekte involviert. Beispielsweise können zwei Datenobjekte miteinander verglichen werden oder Werte zweier Datenobjekte können miteinander aggregiert werden. Bei der Auszeichnung durch entsprechendes Semantisierungsvokabular muss die Auszeichnung somit immer mindestens zwei Objekte im Rahmen der Zielerreichung auszeichnungstechnisch involvieren. Die beiden Objekte müssen hierbei primitiver Natur sein, sodass eine Anwendung der Zielsetzung auf die rohen Werte ermöglicht werden kann.

Die nachfolgenden Abschnitte formalisieren den spezifizierten Typ 'ItemVisualizationGoal' in Tabelle 4.6 als Enumeration zur Aufnahme der Zielsetzungen, sowie die einzelnen Enumerationswerte, welche auch Member genannt werden.

Tabelle 4.6: Zur Semantisierung spezifizierter RDF Klassenvokabularbegriff 'ItemVisualizationGoal'

Name	ItemVisualizationGoal
Label	Item Visualization goal
URI	<a href="http://ts.digital/msc/terms/ItemVisualizationGoal">http://ts.digital/msc/terms/ItemVisualizationGoal</a>
Beschreibung	Enumeration von Visualisierungszielen
Typ des Ausdrucks	Klasse
Ist Unterklasse von	<a href="http://schema.org/Enumeration">http://schema.org/Enumeration</a>

### Identifizieren

Bei der Zielsetzung der Identifikation sollen in der Visualisierung zwei oder mehrere Objekte eindeutig identifiziert werden können [95].

Tabelle 4.7: Zur Semantisierung spezifizierter RDF Membervokabularbegriff 'Identify'

Name	Identify
Label	Identify
URI	<a href="http://ts.digital/msc/terms/Identify">http://ts.digital/msc/terms/Identify</a>
Beschreibung	Indiziert 'Identifizieren' als Visualisierungsziel
Typ des Ausdrucks	Member
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/ItemVisualizationGoal">http://ts.digital/msc/terms/ItemVisualizationGoal</a>

### Lokalisieren

Bei der Zielsetzung der Lokalisierung sollen in der Visualisierung ein oder mehrere Objekte lokalisiert werden können. Exemplarisch für die Zielsetzung der Lokalisierung wäre das Bestimmen eines oder mehrerer Länder auf einer Landkarte. Eine effektive Visualisierung könnte dementsprechend das bzw. die zu lokalisierenden Länder farblich hervorheben und den visuellen Kognitionsprozess so unterstützen [95].

Tabelle 4.8: Zur Semantisierung spezifizierter RDF Membervokabularbegriff 'Locate'

Name	Locate
Label	Locate
URI	<a href="http://ts.digital/msc/terms/Locate">http://ts.digital/msc/terms/Locate</a>
Beschreibung	Indiziert 'Lokalisieren' als Visualisierungsziel

*Fortsetzung auf nächster Seite*

Tabelle 4.8 – Fortgeführt von vorangegangener Seite

Typ des Ausdrucks	Member
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> ItemVisualizationGoal

### Differieren

Bei der Zielsetzung der Differierung sollen in der Visualisierung zwei oder mehrere Objekte voneinander unterschieden werden können [95]. Exemplarisch für die Zielsetzung der Differierung wäre das Unterscheiden zweier Temperaturangaben gleichen Wertes, mit der spezifischen Zielsetzung der Unterscheidung selbiger in Abhängigkeit vom Standort an dem die Temperaturen gemessen wurden. Eine effektive Visualisierung könnte dementsprechend die beiden Temperaturangaben auf einer Standortkarte repräsentieren und so den visuellen Kognitionsprozess zur Differierung optimal unterstützen.

Tabelle 4.9: Zur Semantisierung spezifizierter RDF Membervokabularbegriff 'Differ'

Name	Differ
Label	Differ
URI	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> Differ
Beschreibung	Indiziert 'Differieren' als Visualisierungsziel
Typ des Ausdrucks	Member
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> ItemVisualizationGoal

### Kategorisieren

Bei der Zielsetzung der Kategorisierung sollen in der Visualisierung die Daten, die in einem definierten Zusammenhang stehen entsprechend kategorisiert

repräsentiert werden können. Exemplarisch für die Zielsetzung der Kategorisierung, ist das kategorisieren von Ländern nach deren klimatischen Bedingungen. Eine effektive Visualisierung könnte dementsprechend nach kontinentalem, mediterranem oder gemäßigttem Klima kategorisieren und die entsprechenden Länder in der jeweiligen Kategorie visuell repräsentieren [95].

Tabelle 4.10: Zur Semantisierung spezifizierter RDF Membervokabularbegriff 'Categorize'

Name	Categorize
Label	Categorize
URI	<a href="http://ts.digital/msc/terms/Categorize">http://ts.digital/msc/terms/Categorize</a>
Beschreibung	Indiziert 'Kategorisieren' als Visualisierungsziel
Typ des Ausdrucks	Member
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/ItemVisualizationGoal">http://ts.digital/msc/terms/ItemVisualizationGoal</a>

### Aggregieren

Bei der Zielsetzung der Aggregation sollen in der Visualisierung verschiedene Daten zu einem Ergebnis aggregiert werden. Hierbei werden in aller Regel solche Daten zusammengefasst, über die ein Überblick ermöglicht werden soll oder deren Entwicklung beobachtet werden möchte. Die Ursache für eine entsprechende Aggregation fußt im Normalfall aufgrund der hohen Datenpunktvelfalt, welche bei einzelner Datenpunktbeobachtung von entsprechenden Daten zu aufwendig wäre [95]. Exemplarisch für die Zielsetzung der Aggregation, wäre das aggregieren einzelner städtischer Temperaturen, um den Überblick einer Temperaturverteilung eines Landes zu repräsentieren.

Tabelle 4.11: Zur Semantisierung spezifizierter RDF Membervokabularbegriff 'Aggregate'

Name	Aggregate
Label	Aggregate
URI	http://ts.digital/msc/terms/ Aggregate
Beschreibung	Indiziert 'Aggregieren' als Visualisierungsziel
Typ des Ausdrucks	Member
Inkludiert in Klasse	http://ts.digital/msc/terms/ VisualizationGoal

### Ordnen

Bei der Zielsetzung der Ordnung sollen in der Visualisierung verschiedene Daten nach einer definierten Ordnung sortiert und dementsprechend visuell repräsentiert angeordnet werden [95]. Exemplarisch für die Zielsetzung der Ordnung wäre das Ordnen von Temperaturangaben aus mehreren Jahren eines Landes von der minimalsten Temperatur zur maximalen Temperatur. Damit könnte ein Überblick des Temperaturanstieges eines Landes über mehrere Jahre hinweg visuell repräsentiert werden.

Tabelle 4.12: Zur Semantisierung spezifizierter RDF Membervokabularbegriff 'Organize'

Name	Organize
Label	Organize
URI	http://ts.digital/msc/terms/ Organize
Beschreibung	Indiziert 'Ordnen' als Visualisierungsziel

*Fortsetzung auf nächster Seite*

Tabelle 4.12 – Fortgeführt von vorangegangener Seite

Typ des Ausdrucks	Member
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/ItemVisualizationGoal">http://ts.digital/msc/terms/ItemVisualizationGoal</a>

### Vergleichen

Bei der Zielsetzung des Vergleichens sollen in der Visualisierung zwei oder mehrere Objekte direkt miteinander verglichen werden können. Exemplarisch für die Zielsetzung des Vergleichens wäre das Anordnen von Temperaturangaben über den zeitlichen Verlauf der Messaufnahme entsprechender Temperaturen in einem Graphen für jedes Land. In der Visualisierung können die entsprechenden Graphen übereinandergelegt werden und so die Temperaturen über den zeitlichen Verlauf für jedes Land verglichen werden [95].

Tabelle 4.13: Zur Semantisierung spezifizierter RDF Membervokabularbegriff 'Compare'

Name	Compare
Label	Compare
URI	<a href="http://ts.digital/msc/terms/Compare">http://ts.digital/msc/terms/Compare</a>
Beschreibung	Indiziert 'Vergleichen' als Visualisierungsziel
Typ des Ausdrucks	Member
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/ItemVisualizationGoal">http://ts.digital/msc/terms/ItemVisualizationGoal</a>

### Relationen identifizieren

Bei der Zielsetzung der Identifikation von Beziehungen sollen in der Visualisierung die Relationen von untereinander abhängigen Datenattributen identi-

fiziert werden können [95]. Exemplarisch für die Zielsetzung der Identifikation von Relationen wäre die visuelle Relationsrepräsentation zwischen Temperaturanstieg und den Verkauf von Badeklamotten.

Tabelle 4.14: Zur Semantisierung spezifizierter RDF Membervokabularbegriff 'IdentifyRelations'

Name	IdentifyRelations
Label	Identify relations
URI	http://ts.digital/msc/terms/ Identify-Relations
Beschreibung	Indiziert 'Relationen identifizieren' als Visualisierungsziel
Typ des Ausdrucks	Member
Inkludiert in Klasse	http://ts.digital/msc/terms/ ItemVisualizationGoal

### Assoziieren

Bei der Zielsetzung der Assoziation sollen in der Visualisierung die Daten durch Assoziationen visualisiert werden. Exemplarisch für die Zielsetzung der Assoziation wäre die Repräsentation eines vektoriellen Datensatzes einer Windgeschwindigkeit, visuell repräsentiert durch einen Vektor [95].

Tabelle 4.15: Zur Semantisierung spezifizierter RDF Membervokabularbegriff 'Associate'

Name	Associate
Label	Associate
URI	http://ts.digital/msc/terms/ Associate

*Fortsetzung auf nächster Seite*

Tabelle 4.15 – Fortgeführt von vorangegangener Seite

Beschreibung	Indiziert 'Assoziieren' als Visualisierungsziel
Typ des Ausdrucks	Member
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> ItemVisualizationGoal

### Korrelieren

Bei der Zielsetzung der Korrelation sollen in der Visualisierung die Attribute korrelierend visuell repräsentiert werden, die sich gegenseitig beeinflussen, dabei aber keine erkennbare Abhängigkeit zueinander definieren [95].

Tabelle 4.16: Zur Semantisierung spezifizierter RDF Membervokabularbegriff 'Correlate'

Name	Correlate
Label	Correlate
URI	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> Correlate
Beschreibung	Indiziert 'Korrelieren' als Visualisierungsziel
Typ des Ausdrucks	Member
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> ItemVisualizationGoal

Die Anwendung der Zielsetzung wird durch die Eigenschaft 'visualizationGoal' realisiert, die in Tabelle 4.17 formalisiert ist. Hierbei ist der Typ direkt in den Basistyp 'Thing' der schema.org Ontologie integriert, sodass eine Anwendung der Zielerreichung für die Visualisierung auf Basis jedes Elements durchgeführt werden kann.

Tabelle 4.17: Zur Semantisierung spezifizierter RDF Eigenschaftsvokabularbegriff 'visualizationGoal'

Name	visualizationGoal
Label	Visualization goal
URI	<a href="http://ts.digital/msc/terms/visualizationGoal">http://ts.digital/msc/terms/ visualizationGoal</a>
Beschreibung	Gibt das Visualisierungsziel fuer die eingebetteten Elemente an
Typ des Ausdrucks	Eigenschaft
Inkludiert in Klasse	<a href="http://schema.org/Thing">http://schema.org/Thing</a>
Erwarteter Typ	<a href="http://ts.digital/msc/terms/ItemVisualizationGoal">http://ts.digital/msc/terms/ ItemVisualizationGoal</a>

#### 4.2.5 Vorwissen des Anwenders

Das Vorwissen des Anwenders gilt ebenfalls, wie in Kapitel 3.1.3 beschrieben, als wichtiger Indikator für die Effektivität einer Visualisierung. Das Vorwissen eines Anwenders ist spezifisch für eine Domäne. Oft lässt sich aus dem Vorwissen anderer Fachgebiete, Wissen für die spezifische Domäne ableiten. Aus den Kenntnissen (Wissen, dass) und Fertigkeiten (Wissen, wie) eines Anwenders lässt sich ebenfalls mögliches Vorwissen für eine bestimmte Domäne ableiten [96].

Das Vorwissen eines Anwenders lässt sich sowohl generisch als auch spezifisch durch das Domänenwissen des E-Learnings semantisieren. Mit dem Basisschema von LRMI lässt sich das Domänenwissen des E-Learnings bereits gut semantisieren. LRMI bietet hier die Eigenschaft 'educationalRole' innerhalb des Typs 'EducationalAudience' an. Die Eigenschaft ermöglicht das Vorwissen in Form von Bildungszielgruppen wie beispielsweise Lehrer, Rektor, Schüler, Auszubildender etc. anzugeben. Die Eigenschaft akzeptiert hierbei einen Freitext in Form eines Strings. Für die 'Discovery- und Bindungslogik'

sind solche Freitexte problematisch, da Freitexte eine große Variation äquivalenter Begrifflichkeiten zulassen. Eine bessere Variante ist hier die Codierung fester Werte in Form einer Enumeration anhand dessen Werte die 'Discovery- und Bindungslogik' spezifische Algorithmik implementieren kann. Ein weiterer Aspekt, der für die Typisierung der Eigenschaft spricht, ist, dass für das Auffinden und Zuordnen durch die 'Discovery- und Bindungslogik' eine zu feine Abstufung der Bildungszielgruppen nicht notwendig ist. Somit ist es zum Auffinden und Zuordnen irrelevant, ob die Zielgruppe Lehrer oder Rektoren sind. Beide Zielgruppen bringen etwa das gleiche Vorwissen hinsichtlich des Verständnisses einer Visualisierung mit. So können beide Zielgruppen exemplarisch statistische Visualisierungen verstehen, während die Zielgruppe Schüler der Klasse eine solche statistischen Visualisierungen in der Regel noch nicht verstehen. Eine domänenspezifische Abstufung des Vorwissens (z.B. Architekt oder Informatiker) wäre grundsätzlich möglich. Visualisierungen, die domänenspezifischer Natur sind, werden allerdings kaum verwendet, aus diesem Grund ist eine solche feine Abstufung auch aus Sicht der Domäne des E-Learnings nicht von Relevanz. Sollte es dennoch von Relevanz werden, kann das Anwenderwissen durch zusätzliche domänenspezifische Modelle erweitert werden. Das Basisschema LRMI ermöglicht mit dem Typ 'AlignmentObject' noch weitere Klassifizierungen des Anwenderwissens, die aufgrund der Feingliedrigkeit zum Auffinden und Zuordnen effektiver grafischer Repräsentationen durch die 'Discovery- und Bindungslogik' nicht benötigt werden. So definiert LRMI im Typ 'CreativeWork' die Eigenschaft 'typicalAgeRange', die ebenfalls genutzt werden kann zur Klassifizierung vom Vorwissen der Anwender. Dennoch ist der Autor der Forschungsarbeit der Meinung, dass die Klassifizierung des Vorwissens des Anwenders in der Domäne des E-Learnings über die Spezifikation von Bildungszielgruppen von größerer Relevanz ist. Durch Nutzung von schema.org könnten hier aber auch beide Eigenschaften genutzt werden, und die Bildungszielgruppe so um spezifische Altersgruppen detailliert werden.

### **Typisierung von Bildungszielgruppen**

Zur Typisierung der Bildungszielgruppen wird entsprechend die Enumeration 'ItemEducationalRoleClassified' spezifiziert und die Klasse 'EducationalAudi-

ence' um die Eigenschaft 'educationalRoleClassified' erweitert. Die von LRMI definierte Eigenschaft 'educationalRole' kann dann entsprechend genutzt werden, um besonders feingliedrige Abstufungen der Bildungszielgruppe zu repräsentieren, während die spezifizierte Eigenschaft 'educationalRoleClassified' genutzt wird, um die Bildungsgruppe typisiert zu deklarieren. Ebenfalls möglich und sinnvoll ist die Verwendung beider Eigenschaften, um die typisierte Deklaration mit der von LRMI spezifizierten Eigenschaft zu detaillieren.

Tabelle 4.18: Zur Semantisierung spezifizierter RDF Klassenvokabularbegriff 'ItemEducationalRoleClassified'

Name	ItemEducationalRoleClassified
Label	Item Educational role classified
URI	<a href="http://ts.digital/msc/terms/ItemEducationalRoleClassified">http://ts.digital/msc/terms/ItemEducationalRoleClassified</a>
Beschreibung	Enumeration von Bildungsrollen
Typ des Ausdrucks	Klasse
Ist Unterklasse von	<a href="http://schema.org/Enumeration">http://schema.org/Enumeration</a>

Tabelle 4.19: Zur Semantisierung spezifizierter RDF Eigenschaftsvokabularbegriff 'educationalRoleClassified'

Name	educationalRoleClassified
Label	Educational role classified
URI	<a href="http://ts.digital/msc/terms/educationalRoleClassified">http://ts.digital/msc/terms/educationalRoleClassified</a>
Beschreibung	Spezifiziert fuer ein annotiertes Element des Typs EducationalAudience eine spezifisch definierte Personengruppe aus dem Bereich der Bildung

*Fortsetzung auf nächster Seite*

Tabelle 4.19 – Fortgeführt von vorangegangener Seite

Typ des Ausdrucks	Eigenschaft
Inkludiert in Klasse	<a href="http://schema.org/">http://schema.org/</a> EducationalAudience
Erwarteter Typ	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> ItemEducationalRoleClassified

Die Enumeration zur Typisierung enthält die Mitglieder GeneralEducated (allgemein gebildete), SchoolEducated (Schulgebildete), OccupationEducated (Berufsbildete), HigherEducated (höher gebildete), ScholarlyPerson (Akademiker). Die Mitglieder der Enumeration orientieren sich hierbei an verschiedenen identifizierten Lebenssituationen von Menschen im Kontext der Bildung. So stellt die allgemeine Bildung das Enumerationsmitglied da, das von den meisten Menschen erfüllt wird. Die schulische Bildung stellt die Bildung dar, die ebenfalls von den meisten Menschen in Industrieländern erfüllt wird. Entsprechend trennt sich der Weg oft nach der schulischen Bildung. Die meisten Menschen eignen sich entsprechend berufliche Bildung oder höhere Bildung und akademische Bildung an. Auch ist eine Kombination möglich, so können Menschen erst beruflich gebildet werden und später sich für höhere oder akademische Bildung entscheiden. Die Eigenschaft 'educationalRoleClassified' repräsentiert hierbei immer die höchste Bildung einer Person.

Tabelle 4.20: Zur Semantisierung spezifizierter RDF Membervokabularbegriff 'GeneralEducated'

Name	GeneralEducated
Label	General educated
URI	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> GeneralEducated
Beschreibung	Indiziert 'allgemein gebildet' als Bildungsrolle

*Fortsetzung auf nächster Seite*

Tabelle 4.20 – Fortgeführt von vorangegangener Seite

Typ des Ausdrucks	Member
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> ItemE- educationalRoleClassified

Tabelle 4.21: Zur Semantisierung spezifizierter RDF Membervokabularbegriff 'SchoolEducated'

Name	SchoolEducated
Label	School educated
URI	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> Schoo- IEducated
Beschreibung	Indiziert 'schulisch gebildet' als Bil- dungsrolle
Typ des Ausdrucks	Member
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> ItemE- educationalRoleClassified

Tabelle 4.22: Zur Semantisierung spezifizierter RDF Membervokabularbegriff 'OccupationEducated'

Name	OccupationEducated
Label	Occupation educated
URI	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> Occupa- tionEducated
Beschreibung	Indiziert 'beruflich gebildet' als Bil- dungsrolle
Typ des Ausdrucks	Member

*Fortsetzung auf nächster Seite*

Tabelle 4.22 – Fortgeführt von vorangegangener Seite

Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> ItemE- educationalRoleClassified
----------------------	---

Tabelle 4.23: Zur Semantisierung spezifizierter RDF Membervokabularbegriff 'HigherEducated'

Name	HigherEducated
Label	Higher educated
URI	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> Highe- rEducated
Beschreibung	Indiziert 'hoeher gebildet' als Bil- dungsrolle
Typ des Ausdrucks	Member
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> ItemE- educationalRoleClassified

Tabelle 4.24: Zur Semantisierung spezifizierter RDF Membervokabularbegriff 'ScholarlyPerson'

Name	ScholarlyPerson
Label	Scholarly person
URI	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> Scholar- lyPerson
Beschreibung	Indiziert 'akademisch gebildet' als Bildungsrolle
Typ des Ausdrucks	Member
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> ItemE- educationalRoleClassified

#### 4.2.6 Visuelle Fähigkeiten und Vorlieben des Anwenders

Die visuellen Fähigkeiten eines Menschen lassen sich wie in Kapitel 3.1.3 beschrieben, als Äquivalent zu den Funktionen der visuellen Wahrnehmung definieren. Das semantisierte Modell könnte entsprechend alle visuellen Fähigkeiten von Menschen typisieren. Entsprechend können Daten durch die Typisierung visueller Fähigkeiten ausgezeichnet werden. Dadurch dass der überwiegende Teil der Weltbevölkerung nicht unter visuellen Wahrnehmungsstörungen leidet, ist eine semantische Klassifikation über die Negation visueller Fähigkeiten sinnvoller. Entsprechend müssen nur Daten hinsichtlich des Kriteriums der visuellen Fähigkeiten ausgezeichnet werden, die von bestimmten Personenzielgruppen mit spezifischen visuellen Wahrnehmungsstörungen nicht verarbeitet werden können. Daten, die auch von Personenzielgruppen visuell verarbeitet werden können, die unter visuellen Wahrnehmungsstörungen leiden, müssen entsprechend nicht ausgezeichnet werden. Der Autor dieser Forschungsarbeit geht davon aus, dass der überwiegende Teil an Daten auch von Personenzielgruppen mit visuellen Wahrnehmungsstörungen verarbeitet werden kann. Damit erleichtert die Negation der visuellen Fähigkeiten die Auszeichnung von Daten.

Zur Semantisierung von visuellen Wahrnehmungsstörungen müssen entsprechende Wahrnehmungsstörungen visueller Natur identifiziert werden. Es existiert eine große Anzahl äußerst spezifischer visueller Wahrnehmungsstörungen, die sich aber in die folgenden fünf Kategorien klassifizieren lassen [97]:

- a) Störung der visuellen Raumlageerfassung
- b) Störung der visuellen Erfassung und Differenzierung
- c) Störung der visuellen Figur-Grund-Wahrnehmung
- d) Visuomotorische Koordinationsstörung
- e) Andere Entwicklungsstörungen

Die Auszeichnung visueller Wahrnehmungsstörungen kann sowohl für die Daten als auch für das Visualisierungselement durchgeführt werden. Eine Annotation des Visualisierungselements würde entsprechend signalisieren, ob das annotierte Visualisierungselement für entsprechende Person geeignet ist. Die Annotation auf Basis des Datensatzes würde signalisieren, ob die entsprechende Person unter einer visuellen Wahrnehmungsstörung leidet.

### Typisierung von visuellen Wahrnehmungsstörungen

Zur Typisierung der visuellen Wahrnehmungsstörungen wird die Enumeration 'ItemVisualPerceptionDisorder' spezifiziert und der Typ 'Person' des schema.org-Grundvokabulars um die Eigenschaft 'visualPerceptionDisorder' erweitert.

Tabelle 4.25: Zur Semantisierung spezifizierter RDF Klassenvokabularbegriff 'ItemVisualPerceptionDisorder'

Name	ItemVisualPerceptionDisorder
Label	Item Visual perception disorder
URI	<a href="http://ts.digital/msc/terms/ItemVisualPerceptionDisorder">http://ts.digital/msc/terms/ItemVisualPerceptionDisorder</a>
Beschreibung	Enumeration visueller Wahrnehmungsstörungen
Typ des Ausdrucks	Klasse
Ist Unterklasse von	<a href="http://schema.org/Enumeration">http://schema.org/Enumeration</a>

Tabelle 4.26: Zur Semantisierung spezifizierter RDF Eigenschaftsvokabularbegriff 'visualPerceptionDisorder'

Name	visualPerceptionDisorder
Label	Visual perception disorder
URI	<a href="http://ts.digital/msc/terms/visualPerceptionDisorder">http://ts.digital/msc/terms/visualPerceptionDisorder</a>

*Fortsetzung auf nächster Seite*

Tabelle 4.26 – Fortgeführt von vorangegangener Seite

Beschreibung	Spezifiziert fuer ein annotiertes Element des Typs Person eine visuelle Wahrnehmungsstoerung
Typ des Ausdrucks	Eigenschaft
Inkludiert in Klasse	<a href="http://schema.org/Person">http://schema.org/Person</a>
Erwarteter Typ	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> ItemVisualPerceptionDisorder

Die Enumeration zur Typisierung der visuellen Wahrnehmungsstörung enthält die Mitglieder SpatialPositionDetection (Raumlageerfassung), AcquisitionAndDifferentiation (Erfassung und Differenzierung), FigureBasicPerception (Figur-Grund-Wahrnehmung), VisuomotorCoordinationDisorder (Visuomotorische Koordinationsstörung).

Tabelle 4.27: Zur Semantisierung spezifizierter RDF Membervokabularbegriff 'SpatialPositionDetection'

Name	SpatialPositionDetection
Label	Spatial position detection
URI	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> SpatialPositionDetection
Beschreibung	Indiziert 'visuelle Raumlageerfassung' als visuelle Wahrnehmungsstoerung
Typ des Ausdrucks	Member
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> ItemVisualPerceptionDisorder

Tabelle 4.28: Zur Semantisierung spezifizierter RDF Membervokabularbegriff 'AcquisitionAndDifferentiation'

Name	AcquisitionAndDifferentiation
Label	Acquisition and Differentiation
URI	<a href="http://ts.digital/msc/terms/AcquisitionAndDifferentiation">http://ts.digital/msc/terms/AcquisitionAndDifferentiation</a>
Beschreibung	Indiziert 'visuelle Erfassung und Differenzierung' als visuelle Wahrnehmungsstoerung
Typ des Ausdrucks	Member
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/ItemVisualPerceptionDisorder">http://ts.digital/msc/terms/ItemVisualPerceptionDisorder</a>

Tabelle 4.29: Zur Semantisierung spezifizierter RDF Membervokabularbegriff 'FigureBasicPerception'

Name	FigureBasicPerception
Label	Figure basic perception
URI	<a href="http://ts.digital/msc/terms/FigureBasicPerception">http://ts.digital/msc/terms/FigureBasicPerception</a>
Beschreibung	Indiziert 'visuelle Figur-Grund-Wahrnehmung' als visuelle Wahrnehmungsstoerung
Typ des Ausdrucks	Member
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/ItemVisualPerceptionDisorder">http://ts.digital/msc/terms/ItemVisualPerceptionDisorder</a>

Tabelle 4.30: Zur Semantisierung spezifizierter RDF Membervokabularbegriff 'VisuomotorCoordinationDisorder'

Name	VisuomotorCoordinationDisorder
Label	Visuomotor coordination disorder
URI	<a href="http://ts.digital/msc/terms/VisuomotorCoordinationDisorder">http://ts.digital/msc/terms/VisuomotorCoordinationDisorder</a>
Beschreibung	Indiziert 'visuomotorische Koordinationsstörung' als visuelle Wahrnehmungsstörung
Typ des Ausdrucks	Member
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/ItemVisualPerceptionDisorder">http://ts.digital/msc/terms/ItemVisualPerceptionDisorder</a>

Neben der Semantisierung von visuellen Fähigkeiten des Anwenders bedarf es ebenfalls der Semantisierung der Vorlieben des Anwenders. Wie Kapitel 3.1.3 hierzu ebenfalls erläutert, können die Vorlieben von Anwendern im Kontext grafischer Repräsentationen beispielsweise durch Farbgebungen oder bestimmte Objektanordnungen repräsentiert werden. Eine solche Objektanordnung kann beispielsweise so gestaltet werden, dass in Ländern in denen standardmäßig von links nach rechts gelesen wird, primäre Objekte auch links angeordnet werden, sekundäre hingegen rechts. In Ländern, in denen standardmäßig von rechts nach links gelesen wird, wird die Objektanordnung wie beschrieben entsprechend negiert repräsentiert.

Auch die Farbgebung grafischer Repräsentationen kann durch den kulturellen Hintergrund des Anwenders beeinflusst werden. In verschiedenen Ländern haben Farben so unterschiedliche gesellschaftliche Bedeutung. Die Farbe rot hat für die deutsche Gesellschaft beispielsweise die Bedeutung einer Alarmwirkung, während die chinesische Kultur die Farbe rot mit Freude und Heiterkeit verbindet [98]. Die Auszeichnung der Farbgebung durch das semantische Modell soll dabei den Typ 'Visualization' erweitern, durch die Spezifikation der

Eigenschaften 'preferredHexColor', 'disfavoredHexColor' und 'exclusiveHexColor'. Die Eigenschaft 'preferredHexColor' kann zur Auszeichnung entsprechend genutzt werden, wenn das Visualisierungselement die Annotation von Farben vollständig frei zulässt. Wenn das Visualisierungselement hingegen nur einen bestimmten Bereich an Farben zulässt, kann mit der Eigenschaft 'disfavoredHexColor' angegeben werden, welche Farben für das Visualisierungselement vermieden werden sollen, wenn die Farbe in 'preferredColor' nicht im spezifizierten Bereich liegt. Die Eigenschaft 'exclusiveHexColor' kann genutzt werden, wenn das Visualisierungselement die Farbauswahl nicht zulässt, sondern statisch im Visualisierungselement deklariert ist. In diesem Fall ist die in der Eigenschaft 'exclusiveColor' deklarierte Farbe das Ausschlusskriterium beim Auffinden eines effektiven Visualisierungselements.

Tabelle 4.31: Zur Semantisierung spezifizierter RDF Eigenschaftsvokabularbegriff 'preferredHexColor'

Name	preferredHexColor
Label	Preferred color (Hexadecimal Format)
URI	<a href="http://ts.digital/msc/terms/preferredHexColor">http://ts.digital/msc/terms/preferredHexColor</a>
Beschreibung	Spezifiziert eine Farbe im Hexadezimal-Format die moeglichst fuer die Elemente innerhalb des Typs Visualization genutzt werden soll
Typ des Ausdrucks	Eigenschaft
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/Visualization">http://ts.digital/msc/terms/Visualization</a>
Erwarteter Typ	<a href="http://schema.org/Text">http://schema.org/Text</a>

Tabelle 4.32: Zur Semantisierung spezifizierter RDF Eigenschaftsvokabularbegriff 'disfavoredHexColor'

Name	disfavoredHexColor
Label	Disfavored color (Hexadecimal Format)
URI	<a href="http://ts.digital/msc/terms/disfavoredHexColor">http://ts.digital/msc/terms/disfavoredHexColor</a>
Beschreibung	Spezifiziert eine Farbe im Hexadezimal-Format die nicht fuer die Elemente innerhalb des Typs Visualization genutzt werden soll
Typ des Ausdrucks	Eigenschaft
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/Visualization">http://ts.digital/msc/terms/Visualization</a>
Erwarteter Typ	<a href="http://schema.org/Text">http://schema.org/Text</a>

Tabelle 4.33: Zur Semantisierung spezifizierter RDF Klassenvokabularbegriff 'exclusiveHexColor'

Name	exclusiveHexColor
Label	Exclusive color (Hexadecimal Format)
URI	<a href="http://ts.digital/msc/terms/exclusiveHexColor">http://ts.digital/msc/terms/exclusiveHexColor</a>
Beschreibung	Spezifiziert eine Farbe im Hexadezimal-Format die moeglichst nicht fuer die Elemente innerhalb des Typs Visualization genutzt werden soll

*Fortsetzung auf nächster Seite*

Tabelle 4.33 – Fortgeführt von vorangegangener Seite

Typ des Ausdrucks	Eigenschaft
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> Visuali- zation
Erwarteter Typ	<a href="http://schema.org/Text">http://schema.org/Text</a>

### Typisierung der Interaktion mit einem Visualisierungselement

LRMI als Basisschema definiert ebenfalls die Eigenschaft 'interactivityType' im Typ 'CreativeWork', welches genutzt werden kann, um die Vorliebe des Anwenders hinsichtlich der Interaktionsmöglichkeiten eines Visualisierungselements zu spezifizieren (z.B. Quiz). LRMI gibt an, dass die Werte, die für diese Eigenschaft empfohlen werden, Active (Aktiv), Expositive (Erklärend) oder Mixed (Mix) sein sollten. Der Datentyp ist dementsprechend als Freitext in Form eines Strings deklariert. Besser wäre die Typisierung durch eine Enumeration der von LRMI empfohlenen Werte.

Tabelle 4.34: Zur Semantisierung spezifizierter RDF Klassenvokabularbegriff 'ItemInteractivityType'

Name	ItemInteractivityType
Label	Item Interactivity type
URI	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> ItemIn- teractivityType
Beschreibung	Enumeration von Interaktionen
Typ des Ausdrucks	Klasse
Ist Unterklasse von	<a href="http://schema.org/Enumeration">http://schema.org/Enumeration</a>

Tabelle 4.35: Zur Semantisierung spezifizierter RDF Klassenvokabularbegriff 'interactivityType'

Name	interactivityType
------	-------------------

Fortsetzung auf nächster Seite

Tabelle 4.35 – Fortgeführt von vorangegangener Seite

Label	Interactivity type
URI	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> InteractivityType
Beschreibung	Spezifiziert einen Interaktionstyp für das Element
Typ des Ausdrucks	Eigenschaft
Inkludiert in Klasse	<a href="http://schema.org/CreativeWork">http://schema.org/CreativeWork</a>
Erwarteter Typ	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> ItemInteractivityType

Tabelle 4.36: Zur Semantisierung spezifizierter RDF Membervokabularbegriff 'Active'

Name	Active
Label	Active
URI	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> Active
Beschreibung	Indiziert 'Aktiv' als Interaktionstyp
Typ des Ausdrucks	Member
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> ItemInteractivityType

Tabelle 4.37: Zur Semantisierung spezifizierter RDF Membervokabularbegriff 'Expositive'

Name	Expositive
Label	Expositive
URI	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> Expositive

Fortsetzung auf nächster Seite

Tabelle 4.37 – Fortgeführt von vorangegangener Seite

Beschreibung	Indiziert 'Erklaerend' als Interaktionstyp
Typ des Ausdrucks	Member
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> ItemInteractivityType

Tabelle 4.38: Zur Semantisierung spezifizierter RDF Membervokabularbegriff 'Mixed'

Name	Mixed
Label	Mixed
URI	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> Mixed
Beschreibung	Indiziert 'Mix' als Interaktionstyp
Typ des Ausdrucks	Member
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> ItemInteractivityType

#### 4.2.7 Konventionen des Anwendungskontexts

In Anlehnung an Kapitel 3.1.3 handelt es sich bei den Konventionen des Anwendungskontexts um spezifische Merkmale, die Anwender im Umgang mit einer Anwendung im selben Anwendungskontext gewohnt sind. Nach Analyse der Schul-Cloud, die als Anwendungskontext gemäß Kapitel 2 gilt, können hierbei die gleichen generischen Konventionen angewendet werden, wie bereits im vergangenen Kapitel 4.2.6 zu den Vorlieben des Anwenders gefunden wurden. Darunter Farbe, Formen unter Interaktionsmuster.

#### 4.2.8 Charakteristika des Darstellungsmediums

Eine Vielzahl unterschiedlicher Darstellungsmedien definieren wie in Kapitel 3.1.3 beschrieben auch eine Menge unterschiedlicher Eigenschaftskombinationen. Die nachfolgende Liste zeigt einige dieser Eigenschaften hinsichtlich des Darstellungsmediums. Auf eine Semantisierung des Darstellungsmediums wird hierbei verzichtet, da an die Visualisierungselemente generell die Anforderung gestellt wird, responsive sich zu verhalten und so unabhängig des Darstellungsmediums zu funktionieren.

Auf generischer Ebene lassen sich nach ausgiebiger Analyse verschiedener Endgeräte die folgenden Eigenschaften für Darstellungsmedien finden:

- a) Farbe
- b) Größe des Endgeräts (Breite, Höhe, Tiefe)
- c) Größe des Displays
- d) Bildschirmauflösung
- e) Kamera
- f) Sensoren
- g) Stromversorgung
- h) Betriebssystem
- i) Bedienungshilfen
- j) Unterstützte Sprachen
- k) Unterstützte Videoformate
- l) Anzeigbare Dokumenttypen

## 4.2.9 Web Components

Die Web Component spezifiziert und formalisiert nachfolgend eine Reihe an Typen und Eigenschaften, die eine Abbildung generischer Informationen ermöglichen. Hierbei orientiert sich die Spezifikation am Aufbau von Web Components die aus einem nicht unwesentlichen Teil von HTML-Quellcode bestehen.

Tabelle 4.39: Zur Semantisierung spezifizierter RDF Klassenvokabularbegriff 'WebComponent'

Name	WebComponent
Label	Web Component
URI	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> Web-Component
Beschreibung	Repraesentiert eine Web Komponente
Typ des Ausdrucks	Klasse
Ist Unterklasse von	<a href="http://schema.org/Thing">http://schema.org/Thing</a>

Tabelle 4.40: Zur Semantisierung spezifizierter RDF Klassenvokabularbegriff 'Tag'

Name	Tag
Label	Tag
URI	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> Tag
Beschreibung	Repraesentiert einen HTML-Tag einer Web Komponente
Typ des Ausdrucks	Klasse
Ist Unterklasse von	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> Web-Component

Tabelle 4.41: Zur Semantisierung spezifizierter RDF Klassenvokabularbegriff 'Attribute'

Name	Attribute
Label	Attribute
URI	<a href="http://ts.digital/msc/terms/Attribute">http://ts.digital/msc/terms/ Attribute</a>
Beschreibung	Repraesentiert ein Attribut eines HTML-Tag einer Web Komponente
Typ des Ausdrucks	Klasse
Ist Unterklasse von	<a href="http://ts.digital/msc/terms/Tag">http://ts.digital/msc/terms/ Tag</a>

Tabelle 4.42: Zur Semantisierung spezifizierter RDF Klassenvokabularbegriff 'value'

Name	value
Label	Value
URI	<a href="http://ts.digital/msc/terms/value">http://ts.digital/msc/terms/ value</a>
Beschreibung	Spezifiziert einen Wert fuer einen HTML-Tag oder ein HTML-Attribut
Typ des Ausdrucks	Eigenschaft
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/Tag">http://ts.digital/msc/terms/ Tag</a> und <a href="http://ts.digital/msc/terms/Attribute">http://ts.digital/msc/terms/ Attribute</a>
Erwarteter Typ	<a href="http://schema.org/Text">http://schema.org/Text</a>

Tabelle 4.43: Zur Semantisierung spezifizierter RDF Klassenvokabularbegriff 'startTag'

Name	startTag
Label	Start Tag
URI	<a href="http://ts.digital/msc/terms/startTag">http://ts.digital/msc/terms/ startTag</a>

*Fortsetzung auf nächster Seite*

Tabelle 4.43 – *Fortgeführt von vorangegangener Seite*

Beschreibung	Spezifiziert den umschliessenden Tag einer Web Komponente
Typ des Ausdrucks	Eigenschaft
Inkludiert in Klasse	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> Web-Component
Erwarteter Typ	<a href="http://ts.digital/msc/terms/">http://ts.digital/msc/terms/</a> Tag

## 5 Schlussbetrachtung

Mit Abschluss des vorherigen Abschnitts wurden alle Leitfragen beantwortet, die zur Unterstützung der Antwortfindung für die Forschungsfrage definiert wurden. Der vorherige Abschnitt illustrierte hierbei die Strukturen und Prozesse zur Bewältigung der Problemstellung und fokussierte sich auf die Semantisierung der Informationen als primäre Komponentenbestandteile des konzeptuellen Modells. Hierzu wurde entsprechendes Spezifikationsvokabular definiert, das eine Auszeichnung von Informationen zum Auffinden eines geeigneten Visualisierungselements für entsprechende Daten ermöglicht und die Zuordnung der entsprechenden Daten zum Visualisierungselement gewährleistet.

Die für den Abschnitt Entwurf und Design des konzeptuellen Modells aufgestellten und durch die Erkenntnisse beantworteten Leitfragen werden nachfolgend zusammengefasst:

1. **Welche architektonischen Strukturen und Prozesse muss ein konzeptuelles Modell aufweisen, um die erst zur Laufzeit bekannt werdenden heterogenen Datenstrukturen und Datenquellen gemäß dem Effektivitätskriterium grafisch zu repräsentieren und wie wird eine mögliche Integration der grafischen Repräsentation in den Anwendungskontext konzeptionell realisiert?** Um die zur Laufzeit bekannt werdenden heterogenen Datenstrukturen und Datenquellen gemäß dem in dieser Forschungsarbeit definierten Begriff der Effektivität grafisch zu repräsentieren, muss das konzeptuelle Modell aus den drei folgenden Komponenten bestehen: Semantisch beschriebene Daten, Visualisierungs-Repository mit semantisch beschriebenen grafischen Repräsentationselementen sowie Discovery- und Bindungslogik. Das Auffinden eines Visualisierungselements für spezifische Daten und das daran anschließende

de Zuordnen der Daten zum Visualisierungselement wird durch den algorithmischen Prozess des Abgleichs realisiert. Ein Abgleich wird ermöglicht, durch den Aufbau eines Baumes für die Daten und den Aufbau eines Baumes für zu untersuchende Visualisierungselemente. Der Aufbau der Bäume spiegelt dabei die Struktur der Daten beziehungsweise der Visualisierungselemente wieder. Existieren äquivalente Bäume zwischen Daten und Visualisierungselementen, existiert grundlegend eine Menge an Visualisierungselementen für die Daten, welche eine Zuordnung von Daten zum äquivalenten Visualisierungselement grundsätzlich ermöglichen. Im zweiten Schritt erfolgt auf der Ebene der Blätter der einzelnen Bäume die Evaluation von Äquivalenzen zwischen den Bäumen hinsichtlich der nicht strukturellen Einflussfaktoren identifizierter Qualitätskriterien von Visualisierungen. Das Resultat ist eine Menge an effektiven Visualisierungselementen für spezifische Daten. Die Integration des konzeptuellen Modells in den Anwendungskontext der Schul-Cloud erfolgt hierbei durch einen WYSIWYG-Editor.

- 2. Wie muss eine Vokabularmenge zur Semantisierung spezifiziert werden, sodass Daten und grafische Repräsentationselemente semantisch beschrieben werden können und ein Auffinden von grafischen Repräsentationselementen in Abhängigkeit zum Effektivitätskriterium für entsprechende Daten als auch das Zuordnen der entsprechenden Daten zum gefundenen grafischen Repräsentationselement, gemäß dem konzeptuellen Modells ermöglicht wird?** Die Spezifikation der Vokabularmenge zur Auszeichnung von Daten und grafischen Repräsentationselementen mit dem Ziel der Auffindbarkeit und Zuordenbarkeit durch die Discovery- und Bindungslogik des konzeptuellen Modells wird auf Basis der Ontologien schema.org für die generische Semantisierung und LRMI für die domänenspezifische Semantisierung realisiert. Aufbauend auf dem gewählten Basisvokabular, wurde das zusätzlich benötigte Vokabular spezifiziert, um alle Einflussfaktoren der Qualitätskriterien von Visualisierungen semantisch abbilden zu können. Die zusätzlich spezifizierten Vokabularien werden in Abbildung 5.1 illustriert.

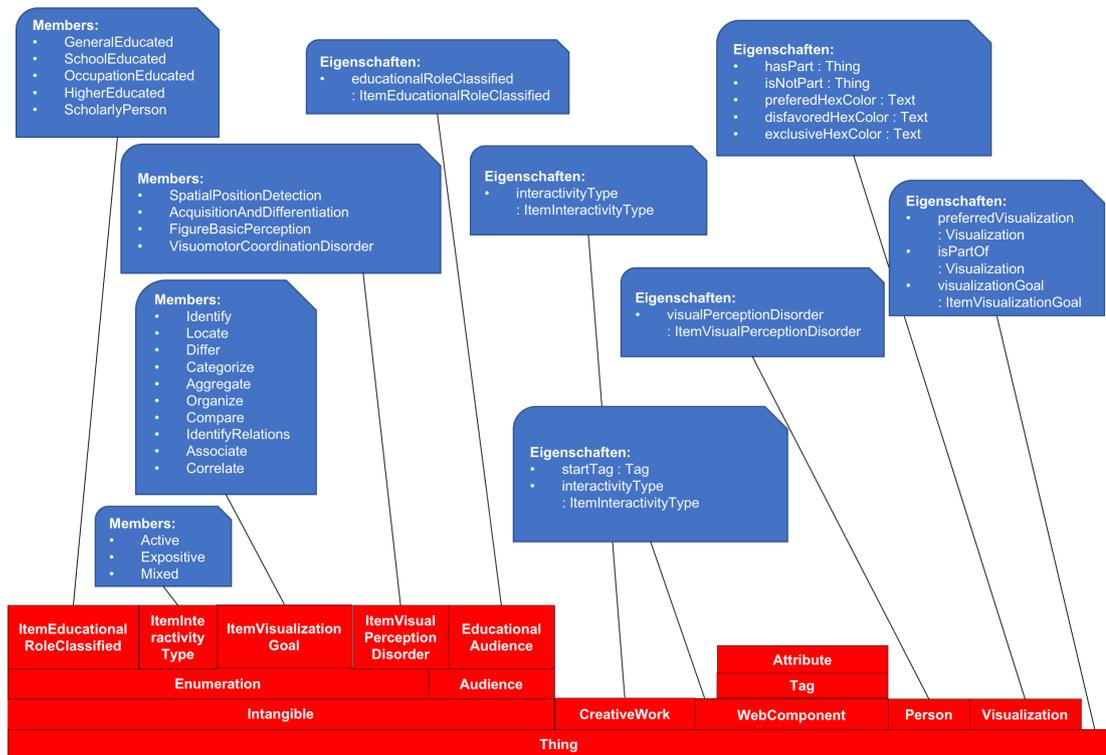


Abbildung 5.1: Grundlegende spezifizierte Vokabularmenge zur Abbildung identifizierter Einflussfaktoren der Qualitätskriterien von Visualisierungen (eigene Darstellung).

Mit den beantworteten Leitfragen des Abschnitts Entwurf und Design des konzeptuellen Modells sowie den beantworteten Leitfragen des Abschnitts Analyse lässt sich die definierte Forschungsfrage wie folgt beantworten: **Wie können effektive grafische Repräsentationen für Daten gewährleistet werden, die zur Entwicklungszeit der Anwendung entweder nicht bekannt oder aufgrund der hohen Menge heterogener Datenstrukturen sowie Datenquellen nicht betrachtet werden können?** Die Gewährleistung grafischer Repräsentationen gemäß dem Effektivitätskriterium für Daten, die zur Entwicklungszeit nicht bekannt oder aufgrund der hohen Menge heterogener Datenstrukturen sowie Datenquellen nicht betrachtet werden können, erfolgt durch Aufstellung eines konzeptuellen Modells, welches die Vorteile der statischen Informationsvisualisierung mit der Notwendigkeit der dynamischen Informationsvisualisierung kombiniert. Entscheidend hierfür sind die drei Kernkomponenten des konzeptuellen Modells - Semantisch beschriebene Daten, Visualisierungs-

Repository mit semantisch beschriebenen grafischen Repräsentationselementen sowie Discovery- und Bindungslogik - und die ganzheitliche Verwendung der Semantisierung durch spezifisch entwickeltes Vokabular.

Die Ergebnisse rund um die beantwortete Forschungsfrage dieser Masterarbeit zeigen, dass eine effektive Informationsvisualisierung in digitalen Informationssystemen auch bei hoch-heterogener Daten- und Datenquellendichte möglich ist. Der vielversprechende Ansatz der vollständigen und automatisierten Generierung ganzheitlicher grafischer Benutzeroberflächen ist hierbei nach Meinung des Autors aus heutiger Sicht noch ungeeignet, für Visualisierungen bei der die Effektivität der Visualisierung zur Unterstützung des menschlichen kognitiven Systems im Fokus steht. Bis zur Reife dieses Ansatzes ist das in dieser Masterarbeit vorgestellte Konzept für produktive Einsätze in Anwendungskontexten insofern geeignet, dass nur die Bereiche des Visualisierungsprozesses automatisiert werden, die aus heutiger Sicht unkritisch sind hinsichtlich der Gewährleistung von Funktionalität. Der in dieser Arbeit vorgestellte Ansatz verzichtet somit auf die nicht zu gewährleistende Automatisierung der Generierung effektiver Visualisierungen. Stattdessen wird die Erzeugung effektiver grafischer Repräsentationen bei den User-Interface-Designern belassen. Die größte Hürde bei Verwendung des Konzepts in Produktivanwendungen bleibt bei der Menge an Visualisierungselementen, aus denen ein effektives Visualisierungselement für spezifische Daten ausgewählt werden kann.

Das konzeptuelle Modell lässt sich zukünftig erweitern. Es ist beispielsweise denkbar, den in dieser Masterarbeit nicht betrachteten Prozess der Überführung von Rohdaten in das semantische Modell durch Algorithmen der künstlichen Intelligenz zu automatisieren. Weitere Aktivitäten können sich mit der grundlegenden Implementierung des konzeptuellen Modells beschäftigen.

Die Beantwortung der Forschungsfrage durch Aufstellung des konzeptuellen Modells lässt zukünftig auf viele Anwendungen hoffen, die dem Menschen zu übermittelnde Informationen dem kognitiven Prozess der visuellen Informationsverarbeitung entsprechend unterstützend grafisch repräsentieren.

## Literaturverzeichnis

- [1] A. Merkel and P. Mangold, "Video-Podcast der Bundeskanzlerin #10/2016 Transkription," Augsburg, Tech. Rep., 2016. [Online]. Available: <http://ts.digital/merkel2016>
- [2] J. Murphy and M. Roser, "Internet," p. 8, 2018. [Online]. Available: <http://ts.digital/murphy2018>
- [3] S. King, "Einleitung," in *Big Data*. Wiesbaden: Springer Fachmedien Wiesbaden, 2014, ch. 1, pp. 19–28. [Online]. Available: <http://ts.digital/king2014>
- [4] S. John Walker, "Big Data: A Revolution That Will Transform How We Live, Work, and Think," *International Journal of Advertising*, vol. 33, no. 1, pp. 181–183, jan 2014. [Online]. Available: <http://ts.digital/johnwalker2014>
- [5] D. Fasel and A. Meier, "Was versteht man unter Big Data und NoSQL?" in *Big Data*. Wiesbaden: Springer Vieweg, Wiesbaden, 2016, ch. 1, pp. 3–16. [Online]. Available: <http://ts.digital/fasel2016>
- [6] Osborne Clarke, "The data gold rush," Tech. Rep., 2012. [Online]. Available: <http://ts.digital/hutton1997>
- [7] M. Volnhals and B. Hirsch, "Information Overload und Controlling," *Controlling & Management*, vol. 52, no. S1, pp. 50–57, may 2008. [Online]. Available: <http://ts.digital/volnhals2008>
- [8] C. Speier, J. S. Valacich, and I. Vessey, "The Influence of Task Interruption on Individual Decision Making: An Information Overload Perspective," *Decision Sciences*, vol. 30, no. 2, pp. 337–360, mar 1999. [Online]. Available: <http://ts.digital/speier1999a>
- [9] R. Evaristo, C. Adams, and S. Curley, "Information Load Revisited: A Theoretical Model," p. 11, 1995. [Online]. Available: <http://ts.digital/evaristo1995>
- [10] W. Kroeber-Riel and F.-R. Esch, "Strategie und Technik der Werbung: verhaltenswissenschaftliche Ansätze, 5. Auflage, Stuttgart, Berlin, Köln," 2000.

- [11] J. T. Milord and R. P. Perry, "A Methodological Study of Overloadx," *The Journal of General Psychology*, vol. 97, no. 1, pp. 131–137, jul 1977. [Online]. Available: <http://ts.digital/milord1977>
- [12] D. Keim, "Information visualization and visual data mining," *IEEE Transactions on Visualization and Computer Graphics*, vol. 8, no. 1, pp. 1–8, 2002. [Online]. Available: <http://ts.digital/keim>
- [13] D. Krömker and W. Müller, "Visualisierung," pp. 1–23, 2007. [Online]. Available: <http://ts.digital/kromker2007>
- [14] N. Iliinsky, "Choosing visual properties for successful visualizations," p. 12, 2013. [Online]. Available: <http://ts.digital/iliinsky2013>
- [15] C. Ware, *Information visualization : perception for design*. Morgan Kaufmann, 2012. [Online]. Available: <http://ts.digital/ware>
- [16] U. M. Fayyad, A. Wierse, and G. G. Grinstein, *Information Visualization in Data Mining and Knowledge Discovery*, ser. The Morgan Kaufmann series in data management systems. Morgan Kaufmann, 2002. [Online]. Available: <http://ts.digital/fayyad2002information>
- [17] Microsoft, "Separate UI and app logic using the Model-View-ViewModel pattern," 2016. [Online]. Available: <http://ts.digital/microsoft1>
- [18] N. Gershon, S. G. Eick, and S. Card, "Information visualization," *interactions*, vol. 5, no. 2, pp. 9–15, mar 1998. [Online]. Available: <http://ts.digital/gershon1998>
- [19] Hasso-Plattner-Institut, "Das Hasso-Plattner-Institut," 2018. [Online]. Available: <http://ts.digital/hpi2>
- [20] Hasso Plattner Institut, "Schul-Cloud," 2018. [Online]. Available: <http://ts.digital/hassoplattnerinstitut1>
- [21] Hasso-Plattner-Institut, "Wikipedia Content Adapter," 2017. [Online]. Available: <http://ts.digital/hpi3>
- [22] U. Hengartner, "Semantische Suchverfahren in der Welt von Big Data," in *Big Data*. Springer Vieweg, Wiesbaden, 2016, ch. 13, pp. 269–287. [Online]. Available: <http://ts.digital/hengartner2016>
- [23] E. Portmann and A. Kuhn, "Extraktion und kartografische Visualisierung von Informationen aus Weblogs," *HMD Praxis der Wirtschaftsinformatik*, vol. 47, no. 1, pp. 81–90, feb 2010. [Online]. Available: <http://ts.digital/portmann2010>

- [24] C. H. Antoni and T. Ellwart, "Informationsüberlastung bei digitaler Zusammenarbeit – Ursachen, Folgen und Interventionsmöglichkeiten," *Gruppe. Interaktion. Organisation. Zeitschrift für Angewandte Organisationspsychologie (GIO)*, vol. 48, no. 4, pp. 305–315, dec 2017. [Online]. Available: <http://ts.digital/antoni2017>
- [25] G. A. Miller, "The magical number seven, plus or minus two: some limits on our capacity for processing information." *Psychological Review*, vol. 63, no. 2, pp. 81–97, 1956. [Online]. Available: <http://ts.digital/miller1956>
- [26] H. A. Simon, "Information processing models of cognition," p. 34, 1979. [Online]. Available: <http://ts.digital/simon1979>
- [27] H. M. Schroder, M. J. Driver, and S. Streufert, "Human information processing: individuals and groups functioning in complex social situations," p. 224, 1967. [Online]. Available: <http://ts.digital/schroder1967>
- [28] M. J. Driver, K. R. Brousseau, and P. L. Hunsaker, *The dynamic decisionmaker: five decision styles for executive and business success*. Harper & Row, 1990. [Online]. Available: <http://ts.digital/driver1990>
- [29] P. Suedfeld, "APA presidential addresses: The relation of integrative complexity to historical, professional, and personal factors." *Journal of Personality and Social Psychology*, vol. 49, no. 6, pp. 1643–1651, 1985. [Online]. Available: <http://ts.digital/suedfeld1985>
- [30] A. R. Abdel-Khalik, "The Effect of Aggregating Accounting Reports on the Quality of the Lending Decision: An Empirical Investigation," *Journal of Accounting Research*, vol. 11, pp. 104–138, 1973. [Online]. Available: <http://ts.digital/abdelkhalik1973>
- [31] E. G. Chewning and A. M. Harrell, "The effect of information load on decision makers' cue utilization levels and decision quality in a financial distress decision task," *Accounting, Organizations and Society*, vol. 15, no. 6, pp. 527–542, jan 1990. [Online]. Available: <http://ts.digital/chewning1990>
- [32] M. D. Shields, "Some effects on information load on search patterns used to analyze performance reports," *Accounting, Organizations and Society*, vol. 5, no. 4, pp. 429–442, jan 1980. [Online]. Available: <http://ts.digital/shields1980>
- [33] D. Snowball, "Some effects of accounting expertise and information load: An empirical study," *Accounting, Organizations and Society*, vol. 5, no. 3, pp. 323–338, jan 1980. [Online]. Available: <http://ts.digital/snowball1980>

- [34] F.-J. Hering, "Informationsüberlastung in Entscheidungsprozessen," 1986.
- [35] D. Bawden, "Information Overload," p. 15, 2001. [Online]. Available: <http://ts.digital/bawden2001>
- [36] S. Gerlach and I. Squarr, "E-Learning," in *Methodenhandbuch für Softwareschulungen*. Springer Vieweg, Berlin, Heidelberg, 2015, ch. 13, pp. 99–111. [Online]. Available: <http://ts.digital/gerlach2015a>
- [37] U. Herrmann, "Zwischenruf: Lernen im digitalen Zeitalter," 2013. [Online]. Available: <http://ts.digital/herrmann2013>
- [38] Hasso Plattner Institut, "Die Schul-Cloud," 2018. [Online]. Available: <http://ts.digital/hpisc>
- [39] T3n, "Digitales Klassenzimmer: „Den Schulen fehlt die Infrastruktur“,“ 2018. [Online]. Available: <http://ts.digital/t3n2018>
- [40] J. Heer and B. Shneiderman, "Interactive Dynamics for Visual Analysis," pp. 1–26, 2012. [Online]. Available: <http://ts.digital/heer2012>
- [41] S. K. Card, J. D. Mackinlay, and B. Shneiderman, *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann Publishers, 1999. [Online]. Available: <http://ts.digital/card1999>
- [42] D. Krömker and W. Müller, "Visualisierung," pp. 1–32, 2007. [Online]. Available: <http://ts.digital/kromker2007a>
- [43] T. Cornsweet, *Visual Perception*, 2012. [Online]. Available: <http://ts.digital/cornsweet2012>
- [44] C. Becker-Carus and M. Wendt, *Allgemeine Psychologie: Eine Einführung*, 2017. [Online]. Available: <http://ts.digital/beckercarus2017>
- [45] S. Frings and F. Müller, "Sehen," in *Biologie der Sinne*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 117–188. [Online]. Available: <http://ts.digital/frings2014b>
- [46] Universität Münster, "Sensorische Physiologie: Hintergrund," p. 11, 2018. [Online]. Available: <http://ts.digital/universitatMunster2018>
- [47] S. Frings and F. Müller, "Schmecken," in *Biologie der Sinne*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 77–93. [Online]. Available: <http://ts.digital/frings2014>

- [48] —, “Riechen,” in *Biologie der Sinne*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 95–116. [Online]. Available: <http://ts.digital/frings2014a>
- [49] —, “Hören,” in *Biologie der Sinne*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 189–226. [Online]. Available: <http://ts.digital/frings2014c>
- [50] —, “Tasten und Fühlen,” in *Biologie der Sinne*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 249–268. [Online]. Available: <http://ts.digital/frings2014d>
- [51] M. J. Eppler and J. Mengis, “The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines,” *The Information Society*, vol. 20, no. 5, pp. 325–344, 2004. [Online]. Available: <http://ts.digital/eppler2004>
- [52] R. Däßler, “Informationsvisualisierung: Stand, Kritik und Perspektiven,” p. 20, 1999. [Online]. Available: <http://fiz1.fh-potsdam.de/volltext/fhpotsdam/03021.pdf>
- [53] D. Krömker, “Visualisierung,” p. 9, 2000. [Online]. Available: <http://ts.digital/kromker2000>
- [54] J. Nichols, “Automatically generating high-quality user interfaces for appliances,” Ph.D. dissertation, Carnegie Mellon University, 2006. [Online]. Available: <http://ts.digital/nichols2006a>
- [55] K. Z. Gajos, “Automatically Generating Personalized User Interfaces,” Ph.D. dissertation, University of Washington, 2008. [Online]. Available: <http://ts.digital/gajos2008a>
- [56] J. Nichols and A. Faulring, “Automatic Interface Generation and Future User Interface Tools,” p. 2, 2005. [Online]. Available: <http://ts.digital/nichols2005>
- [57] B. Frost, “Atomic Design Methodology,” 2016. [Online]. Available: <http://ts.digital/frost>
- [58] A. Kalt, “Wetter – Witterung – Klima,” 2016. [Online]. Available: <http://ts.digital/kalt2016>
- [59] D. Krömker and W. Müller, “Visualisierung,” 2005. [Online]. Available: <http://ts.digital/muller2008>

- [60] R. Spence, “Representation,” in *Information Visualization*. Cham: Springer International Publishing, 2014, pp. 41–110. [Online]. Available: <http://ts.digital/spence>
- [61] O. Deussen, “Datenvisualisierung,” 2002. [Online]. Available: <http://ts.digital/deussen>
- [62] Universität Paderborn, “Attribute,” 2018. [Online]. Available: <http://ts.digital/paderbornattribute>
- [63] W. Kurth, “Charakterisierung von Datensätzen und Darstellungsformen,” 2018. [Online]. Available: <http://ts.digital/charakterisierung>
- [64] H. Leitte, “Datentypen, Datenrepräsentation und Visualisierungspipeline,” p. 70, 2012. [Online]. Available: <http://ts.digital/leitte2012>
- [65] GitHub, “Built for developers,” 2017. [Online]. Available: <http://ts.digital/github2017>
- [66] GitHub Developer, “GitHub Developer,” 2018. [Online]. Available: <http://ts.digital/gitHubDeveloper2018>
- [67] Mozilla Developer, “Web Components,” 2018. [Online]. Available: <http://ts.digital/mozillaDeveloper2018>
- [68] Webcomponents.org, “About webcomponents.org,” p. 1, 2018. [Online]. Available: <http://ts.digital/webcomponentsorg44>
- [69] P. Hitzler, M. Krötzsch, S. Rudolph, and Y. Sure, “Einleitung und Übersicht,” in *Semantic Web*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, ch. 1, pp. 1–5. [Online]. Available: <http://ts.digital/hitzler2008e>
- [70] K. Tochtermann and H. Maurer, “Semantic Web — Geschichte und Ausblick einer Vision,” in *Semantic Web*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 1–6. [Online]. Available: <http://ts.digital/tochtermann2006>
- [71] P. Hitzler, M. Krötzsch, S. Rudolph, and Y. Sure, “Die Idee des Semantic Web,” in *Semantic Web*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 7–14. [Online]. Available: <http://ts.digital/hitzler2008f>
- [72] ———, “Einfache Ontologien in RDF und RDF Schema,” in *Semantic Web*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, ch. 3, pp. 33–88. [Online]. Available: <http://ts.digital/hitzler2008g>

- [73] L. Yu, "RDFS and Ontology," in *A Developer's Guide to the Semantic Web*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 121–167. [Online]. Available: <http://ts.digital/yu2014>
- [74] Schema.org, "Thing," 2018. [Online]. Available: <http://ts.digital/schemaorg2018e1>
- [75] —, "Organization of Schemas," 2018. [Online]. Available: <http://ts.digital/schemaorg>
- [76] —, "Welcome to Schema.org," 2018. [Online]. Available: <http://ts.digital/schemaorga>
- [77] —, "Data Model," 2018. [Online]. Available: <http://ts.digital/schemaorgd>
- [78] P. Barker and L. M. Campbell, "LRMI, Learning Resource Metadata on the Web," in *Proceedings of the 24th International Conference on World Wide Web - WWW '15 Companion*. New York, New York, USA: ACM Press, 2015, pp. 687–687. [Online]. Available: <http://ts.digital/barker2015>
- [79] M. Al-Yahya, R. George, and A. Alfaries, "Ontologies in E-Learning: Review of the Literature," *International Journal of Software Engineering and Its Applications*, vol. 9, no. 2, pp. 67–84, 2015. [Online]. Available: <http://ts.digital/alyahya2015>
- [80] M. Haneefa K and M. Chembrakuzhi, "Metadata Standards for Open Educational Resources," in *59th ILA International Conference on Managing Libraries in the Changing Information World*, 2014, p. 11. [Online]. Available: <http://ts.digital/mcclelland2003>
- [81] P. Barker, "What is IEEE Learning Object Metadata / IMS Learning Resource Metadata?" 2011. [Online]. Available: <http://ts.digital/barker>
- [82] S. J. Miller, *Metadata for Digital Collections: A How-To-Do-It Manual*, 2011. [Online]. Available: <http://ts.digital/miller2011>
- [83] Network Working Group, "Dublin Core Metadata for Resource Discovery," 1998. [Online]. Available: <http://ts.digital/nwg1998>
- [84] X. Ma, "Metadata," in *Encyclopedia of Big Data*. Springer International Publishing, 2018, p. 6. [Online]. Available: <http://ts.digital/ma2018>
- [85] O. Bohl, J. Scheuhase, R. Sengler, and U. Winand, "The sharable content object reference model (SCORM) - a critical review," in *International Conference on Computers in Education, 2002. Proceedings*.

- IEEE Comput. Soc, 2002, pp. 950–951. [Online]. Available: <http://ts.digital/bohl2002>
- [86] C. Prpitsch and P. Veith, “Content and management standards: LOM, SCORM and Content Packaging,” in *Handbook on Quality and Standardisation in E-Learning*. Springer Berlin Heidelberg, 2006, ch. 15, pp. 209–223. [Online]. Available: <http://ts.digital/prpitsch2006>
- [87] Y. Miao, “IMS QTI Authoring,” in *Learning Network Services for Professional Development*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 389–396. [Online]. Available: <http://ts.digital/miao2009>
- [88] J. C. Ramalho, A. Simoes, and R. Queiros, *Innovations in XML Applications and Metadata Management: Advancing Technologies: Advancing Technologies*. IGI Global (701 E. Chocolate Avenue, Hershey, Pennsylvania, 17033, USA), 2012. [Online]. Available: <http://ts.digital/ramalho2012>
- [89] Schema.org, “EducationalAudience,” 2018. [Online]. Available: <http://ts.digital/schemaorg20181>
- [90] LRMI, “LRMI Metadata Terms (RDF),” 2015. [Online]. Available: <http://ts.digital/lrmi2015>
- [91] Schema.org, “AlignmentObject,” 2018. [Online]. Available: <http://ts.digital/schemaorg2018a1>
- [92] —, “Audience,” 2018. [Online]. Available: <http://ts.digital/schemaorg2018b1>
- [93] —, “Intangible,” 2018. [Online]. Available: <http://ts.digital/schemaorg2018c1>
- [94] —, “CreativeWork,” 2018. [Online]. Available: <http://ts.digital/schemaorg2018d1>
- [95] Universität Paderborn, “Operationen,” 2018. [Online]. Available: <http://ts.digital/operationen>
- [96] Spektrum, “Vorwissen,” 2000. [Online]. Available: <http://ts.digital/spektrum>
- [97] D. Irblich and G. Renner, *Diagnostik in der Klinischen Kinderpsychologie: Die ersten sieben Lebensjahre*. Hogrefe Verlag, 2009. [Online]. Available: <http://ts.digital/irblich2009>

- 
- [98] F. Gonzalez, "Andere Länder, andere Seiten – 6 Einblicke in interkulturelles Web-Design," 2014. [Online]. Available: <http://ts.digital/gonzalez2014>