

*Band  
2*

Rüdiger Buchkremer (Hrsg.)

*Analyse sensorischer E-Commerce-Elemente  
mittels Big-Data-Methoden und Künstlicher  
Intelligenz*

~  
Hamacher / Blach / Kozlik / Muster / Nöllenburg / Ohletz / Franken / Hernes /  
Hinterding / Höveler / Huppertz / Leppkes / Lopez Rodriguez / Maucy /  
Petrov / Schäfer / Schneider / Spiegel / Stecker / Steinmann / Tembrink /  
Titze-Wolter / Vishnyakova / Zimmermann / Buchkremer

**ifid** Schriftenreihe

Beiträge zu IT-Management & Digitalisierung

**FOM**  
Hochschule

**ifid**

Institut für IT-Management &  
Digitalisierung  
der FOM University of Applied Sciences

**Hamacher / Blach / Kozlik / Muster / Nöllenburg / Ohletz / Franken / Hernes /  
Hinterding / Höveler / Huppertz / Leppkes / Lopez Rodriguez / Maucy / Petrov /  
Schäfer / Schneider / Spiegel / Stecker / Steinmann / Tembrink / Titze-Wolter /  
Vishnyakova / Zimmermann / Buchkremer**

*Analyse sensorischer E-Commerce-Elemente mittels Big-Data-Methoden und Künstlicher  
Intelligenz – Automatisierung sensorischer Bewertungen von E-Commerce- und  
Social-Media-Plattformen auf Basis des Online Sensory Marketing Index*

ifid Schriftenreihe der FOM, Band 2  
Beiträge zu IT-Management & Digitalisierung

Essen 2023

ISBN (Print) 978-3-89275-320-9      ISSN (Print) 2699-562X

ISBN (eBook) 978-3-89275-321-6      ISSN (eBook) 2699-5638

Dieses Werk wird herausgegeben vom ifid Institut für IT-Management & Digitalisierung  
der FOM Hochschule für Oekonomie & Management gGmbH

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliographie;  
detaillierte bibliographische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

© 2023 by



**MA Akademie  
Verlags- und Druck-  
Gesellschaft mbH**

MA Akademie Verlags-  
und Druck-Gesellschaft mbH  
Leimkugelstraße 6, 45141 Essen  
[info@mav-verlag.de](mailto:info@mav-verlag.de)

Das Werk einschließlich seiner  
Teile ist urheberrechtlich geschützt.  
Jede Verwertung außerhalb der  
engen Grenzen des Urhebergeset-  
zes ist ohne Zustimmung der MA  
Akademie Verlags- und Druck-  
Gesellschaft mbH unzulässig und  
strafbar. Das gilt insbesondere für  
Vervielfältigungen, Übersetzungen,  
Mikroverfilmungen und die Ein-  
speicherung und Verarbeitung in  
elektronischen Systemen.

Die Wiedergabe von Gebrauchs-  
namen, Handelsnamen, Warenbe-  
zeichnungen usw. in diesem Werk  
berechtigt auch ohne besondere  
Kennzeichnung nicht zu der Annah-  
me, dass solche Namen im Sinne  
der Warenzeichen- und Marken-  
schutz-Gesetzgebung als frei zu  
betrachten wären und daher von  
jedermann benutzt werden dürfen.  
Oft handelt es sich um gesetzlich  
geschützte eingetragene Waren-  
zeichen, auch wenn sie nicht als  
solche gekennzeichnet sind.

Rüdiger Buchkremer (Hrsg.)

***Analyse sensorischer E-Commerce-Elemente mittels  
Big-Data-Methoden und Künstlicher Intelligenz –  
Automatisierung sensorischer Bewertungen von  
E-Commerce- und Social-Media-Plattformen auf  
Basis des Online Sensory Marketing Index***

Kevin Hamacher, Maike Blach, Jan Kozlik,  
Felix Muster, Pascal-Philipp Nöllenburg, Jan-Henrik Ohletz,  
Gereon Franken, Dominic Hernes, Martin Hinterding,  
Patrick Höveler, Michael Huppertz, Niklas Leppkes,  
Ariana Lopez Rodriguez, Kevin Maucy, Aleksandar Petrov,  
Dennis Schäfer, Robert Schneider, Bastian Spiegel,  
Rouven Stecker, Philipp Steinmann, Christoph Tembrink,  
Philipp Titze-Wolter, Liudmila Vishnyakova, Jürgen Zimmermann,  
Rüdiger Buchkremer

**Autorenkontakt:**

E-Mail: [kevin.hamacher@fom-net.de](mailto:kevin.hamacher@fom-net.de)

## **Vorwort**

Sehr geehrte Leserinnen und Leser,

die fortschreitende Digitalisierung unserer Welt hat dazu geführt, dass Kundinnen und Kunden zunehmend über digitale Plattformen Produkte und Dienstleistungen kaufen und sich über sie austauschen. Das sensorische Online-Erlebnis, das auf diesen Plattformen geboten wird, ist jedoch oft auf visuelle, akustische und schriftliche Elemente begrenzt. Wie in der wissenschaftlichen Literatur bereits belegt, ist eine verbesserte sensorische Erfahrung auf diesen Plattformen ein wichtiger Faktor, um z. B. die Verweildauer, den Warenkorbwert und letztlich die Kaufentscheidungen von Kundinnen und Kunden positiv zu beeinflussen. Es ist daher von hoher Bedeutung, Medientypen und Inhalte quantitativ zu analysieren, zu bewerten und sensorisch zu optimieren, um die digitale Customer Journey nachhaltig zu verbessern.

Im zweiten Band der ifid Schriftenreihe präsentieren die Autorinnen und Autoren Ansätze und Ergebnisse zur Entwicklung von alternativen Verfahren zur automatischen Erfassung und Bewertung sensorischer Elemente auf E-Commerce-Webseiten und auf Social-Media-Plattformen. Dazu setzen sie moderne Big-Data- und Künstliche-Intelligenz-Ansätze aus dem Bereich der Verarbeitung natürlicher Sprache und der Bildanalyse ein.

Die Ergebnisse zeigen, dass der Einsatz von KI-Methoden wie Worteinbettungstechniken und Transformer-Technologien wie der BERT-Algorithmus zur Untersuchung von über 200 Webseiten zu einem funktionalen und skalierbaren Ansatz zur Bestimmung von Elementen aus dem Themenfeld des sensorischen Marketings für E-Commerce-Plattformen und Social Media führt.

Die Autorinnen und Autoren hoffen, dass dieser Band einen Beitrag dazu leistet, die Bedeutung von sensorischem Marketing im digitalen Zeitalter zu unterstreichen und gleichzeitig die vielfältigen Möglichkeiten der eingesetzten Big-Data-Methoden zu demonstrieren, mit dem Ziel, weitere Forschungen auf diesem und weiteren Gebieten zu initiieren.

Düsseldorf, im Februar 2023

Kevin Hamacher, M.Sc.

Research Fellow ifid Institut für IT-Management & Digitalisierung



## **Zusammenfassung**

Kundinnen und Kunden kommunizieren, kaufen und teilen Produkte und Dienstleistungen über digitale Plattformen, in denen technische Möglichkeiten zur Präsentation sensorischer Elemente im Normalfall auf visuelle, akustische und schriftliche Elemente begrenzt sind. Um das sensorische Online-Erlebnis zu verbessern, ist es erforderlich, Medientypen und Inhalte quantitativ zu bewerten und sensorisch zu optimieren. Basierend auf dem Online Sensory Marketing Index (OSMI), einem Schema zur sensorischen Bewertung von E-Commerce-Webseiten, präsentiert das Paper Ansätze und Ergebnisse zur Entwicklung von alternativen Verfahren zur automatischen Erfassung und Bewertung sensorischer Elemente wie Texte und Bilder. In diesem Beitrag werden zu diesem Zweck moderne Big-Data- und Künstliche-Intelligenz-Ansätze aus dem Bereich der Verarbeitung natürlicher Sprache, „Natural Language Processing“ und „Natural Language Understanding“, angewendet. Dazu werden Worteinbettungstechniken mit GloVe und Word2Vec sowie Transformer-Technologien wie der BERT-Algorithmus zur Untersuchung von über 200 Webseiten eingesetzt. Die Ergebnisse demonstrieren einen funktionalen und skalierbaren KI-Ansatz zur Bestimmung von Elementen aus dem Themenfeld des sensorischen Marketings für E-Commerce-Plattformen und Social Media von verschiedenen Branchen und Unternehmen. Darüber hinaus sind die neuen KI-Auswahl- und Bewertungsverfahren deutlich effizienter und mit nur geringen Leistungseinbußen im Vergleich zu einer manuellen Bewertung verbunden.

## Inhalt

Vorwort .....	III
Zusammenfassung .....	IV
Abkürzungsverzeichnis .....	VII
Abbildungsverzeichnis .....	VIII
Tabellenverzeichnis .....	IX
Formelverzeichnis .....	X
1 Einleitung .....	1
1.1 Forschungsfrage und methodischer Aufbau .....	2
1.2 Verwandte Arbeiten .....	2
2 Online Sensory Marketing Index (OSMI) .....	5
2.1 OSMI-Parameter .....	5
2.2 OSMI-Indikatoren .....	5
2.2.1 Haptische Indikatoren .....	9
2.2.2 Olfaktorische Indikatoren .....	13
2.2.3 Auditiv Indikatoren .....	14
2.2.4 Gustatorische Indikatoren .....	16
2.2.5 Visuelle Indikatoren .....	17
2.3 Indizes .....	19
2.4 Projektprozess und Organisation .....	21
2.5 Allokation der Arbeitsschwerpunkte .....	22
3 Textdaten .....	24
3.1 Vorgehensweise .....	24
3.2 Datenbeschaffung .....	24
3.2.1 Web .....	24
3.2.2 Social Media .....	27
3.3 Datenanalyse .....	30
3.3.1 Architektur .....	30
3.3.2 Preprocessing .....	34
3.3.3 Word Embedding .....	36
3.3.4 Modellierung .....	39

4	Bilddaten.....	42
4.1	Vorgehensweise .....	42
4.2	Datenbeschaffung.....	42
4.2.1	Web .....	42
4.2.2	Social Media .....	46
4.3	Analyse .....	51
4.3.1	Ermittlung der Indikatoren .....	51
4.3.2	Objekterkennung .....	53
4.3.3	Weitere Bildanalysen.....	60
5	Visualisierung .....	68
5.1	Interaktive Visualisierung.....	68
5.1.1	Dynamische Dashboards .....	68
5.1.2	Menschliche Wahrnehmung .....	69
5.2	Vorbereitungen für die Visualisierung und explorative Datenanalyse..	70
5.2.1	Transfer und Vorbereitung der Daten.....	70
5.2.2	Toolauswahl.....	71
5.3	Dashboard Erstellung .....	72
5.3.1	Gesamtübersicht .....	73
5.3.2	Radar- und Word-Chart.....	77
5.3.3	Verteilungsdiagramm.....	80
5.3.4	K-Means-Algorithmus pro Sinn.....	82
5.3.5	Detailauswertungen der Textanalyseverfahren.....	84
5.3.6	Aggregierte Gesamtübersicht nach Branchen unterteilt .....	86
6	Schlussbetrachtung .....	88
6.1	Diskussion.....	88
6.1.1	Interpretation der Ergebnisse .....	91
6.1.2	Limitationen der Untersuchung.....	92
6.1.3	Empfehlung für weiterführende Forschung .....	93
6.2	Fazit .....	94
	Literatur.....	97

## **Abkürzungsverzeichnis**

API	Application Programming Interface
CSV	Character-separated Values / Comma-separated Values
EFA	explorative Faktorenanalyse
HTML	Hypertext Markup Language
NLP	Natural Language Processing
OSMI	Online Sensory Marketing Index
SQL	Structured Query Language
TF-IDF	Term-Frequency-Inverse Document Frequency
WQI	Web Quality Index

## Abbildungsverzeichnis

Abbildung 1:	Projektprozess und Struktur.....	22
Abbildung 2:	Prozessarchitektur .....	31
Abbildung 3:	Berechnung der textbasierten OSMI-Parameter .....	33
Abbildung 4:	Aufbau der Textdatenanalyse .....	35
Abbildung 5:	TF-IDF-Berechnung .....	39
Abbildung 6:	Web-Scraping-Prozess .....	44
Abbildung 7:	Instagram-Account von BMW .....	48
Abbildung 8:	Funktionsweise des YOLO-Modells.....	56
Abbildung 9:	Schematische Darstellung Intersection over Unit.....	57
Abbildung 10:	Beispiel einer Objekterkennung mit YOLOv5 .....	59
Abbildung 11:	Beispielbild für eine Label-Analyse.....	61
Abbildung 12:	Beispielbild für eine Gesichtserkennung.....	63
Abbildung 13:	Beispielbild für die Texterkennung.....	65
Abbildung 14:	Kennzahlen und Filter des Dashboards.....	72
Abbildung 15:	Ausschnitt des OSMI in der Gesamtübersicht.....	74
Abbildung 16:	Balkendiagramm zum OSMI pro Branche und zu den Top 10 Firmen.....	75
Abbildung 17:	Balkendiagramm zum OSMI nach Medium und Verfahren .....	76
Abbildung 18:	Radar-Chart (oben) und Word-Cloud (unten) für Weleda .....	78
Abbildung 19:	Radar-Chart und Word-Cloud für GhdHair .....	79
Abbildung 20:	Verteilungsdiagramm und Lasso-Filter .....	81
Abbildung 21:	K-Means-Algorithmus: eine Iteration .....	82
Abbildung 22:	K-Means-Algorithmus: drei Iterationen .....	83
Abbildung 23:	Geschmackssinn und Gesamt-OSMI als Firmenübersicht.....	84
Abbildung 24:	Gesamt-OSMI-Algorithmen-Vergleich .....	85
Abbildung 25:	OSMI pro Branche .....	86
Abbildung 26:	OSMI pro Branche nach Bereinigung der Ausreißer .....	86

## **Tabellenverzeichnis**

Tabelle 1:	Übersicht der OSMI-Parameter .....	5
Tabelle 2:	Übersicht der Online-Sensory-Marketing-Index-Indikatoren .....	6
Tabelle 3:	Ergebnis der Label-Analyse.....	62
Tabelle 4:	Analyseergebnis der Gesichtserkennung .....	64
Tabelle 5:	Analyseergebnis Gesichtserkennung – Emotionen.....	64
Tabelle 6:	Analyseergebnis der Texterkennung .....	66
Tabelle 7:	Vorgeschlagene neue OSMI-Indikatoren .....	89
Tabelle 8:	Vorgeschlagene neue OSMI-Indikatoren .....	90
Tabelle 9:	Indikatoren der Bildanalyse .....	90
Tabelle 10:	Branchenbezogene relative Wichtigkeit der Sinnesmodalitäten .....	91
Tabelle 11:	Weitere Forschungsfelder.....	93

## **Formelverzeichnis**

Formel 1:	TF-IDF-Score .....	40
Formel 2:	Term Frequency.....	40
Formel 3:	Inverse Document Frequency.....	40
Formel 4:	Faltung .....	54
Formel 5:	Konfidenzwert .....	58
Formel 6:	Klassenzugehörigkeitswahrscheinlichkeit.....	58

## 1 Einleitung

Gemäß Towers werden Kaufentscheidungen von Konsumentinnen und Konsumenten durch alle fünf menschlichen Sinne markant beeinflusst.<sup>1</sup> Dazu existieren verschiedene Studien, die in Kapitel 1.2 näher erläutert werden. Das menschliche Gehirn nimmt Eindrücke über das Sehen, Hören, Fühlen, Schmecken und Riechen auf und entschlüsselt diese. Erst nach dieser Entschlüsselung reagiert der Mensch auf seine Umwelt. Wenn dementsprechend ein Produkt oder eine Marke einen hohen Wiedererkennungswert oder eine Differenzierungsmöglichkeit gegenüber anderen Wettbewerbern erreichen soll, ist die Ansprache der menschlichen Sinne zentral. Dabei besteht das Risiko eines Overflows, wenn zu viele Sinne stimuliert werden oder zu stark stimuliert wird.<sup>2</sup>

Die Beeinflussung der menschlichen Sinne lässt sich im Alltag beobachten.<sup>3</sup> Beispiele hierfür sind das Abspielen von Musik in Geschäften, gewachste Äpfel, die dem Auge schmeicheln sollen, oder der Geruch von Parfüm. Das Ziel, eine Kaufreaktion auszulösen, soll durch die Beeinflussung der Sinne der Konsumentinnen und Konsumenten erreicht werden.<sup>4</sup> Marketingstrategien, die mit der Beeinflussung der Sinne arbeiten, werden auch als „multisensorisches Marketing“ beschrieben. Das „multisensorische Marketing“ beschreibt das Zusammenwirken mehrerer Reizmodalitäten und zeigt dabei die systematische Abstimmung aller sinnlich wahrnehmbaren Marketingmaßnahmen auf.<sup>5</sup>

Die forcierte Digitalisierung, die derzeit durch die Covid-19-Pandemie noch beschleunigt wird, führt zu immer mehr online durchgeführten Einkäufen. Dementsprechend werden auch Marketingstrategien kontinuierlich weiter digitalisiert. Da die onlinekaufende Kundschaft die Produkte nicht tatsächlich sehen, riechen, schmecken, hören oder anfassen kann, soll dies auf eine indirekte Weise geschehen. Dieses Paper baut auf vorhandenen Forschungsergebnissen auf und untersucht mithilfe von Big-Data-Methoden die Unternehmenskommunikation zu den Endkundinnen und -kunden auf sensorische Inhalte.<sup>6</sup>

Im Rahmen dieser Untersuchung wird der „Online Sensory Marketing Index“, kurz OSMI, verwendet. Dieser wird in Kapitel 2 erläutert.

---

<sup>1</sup> Vgl. Towers 2014, S. 107–117.

<sup>2</sup> Vgl. Foscht et al. 2010, S. 18.

<sup>3</sup> Vgl. Petit et al. 2015, S. 1.

<sup>4</sup> Vgl. Elder & Krishna 2010, S. 755; Hartmann & Haupt 2016, S. 29.

<sup>5</sup> Vgl. Krishna 2012, S. 332.; Peck & 2008, o.S.

<sup>6</sup> Vgl. IfD Allensbach 2016; Statistisches Bundesamt 2017.



## **1.1 Forschungsfrage und methodischer Aufbau**

Im Verlauf dieser Arbeit sollen folgende Forschungsfragen beleuchtet werden:

1. Welche Unterschiede sind hinsichtlich branchenspezifischer Besonderheiten im sensorischen Marketing erkennbar?
  - In Bezug auf die Gewichtung der Sinne
  - In Bezug auf eine mögliche Implementierung
  - In Bezug auf die Grenzen der Implementierung
  - In Bezug auf das Design der sensorischen Kommunikation
2. Können branchenspezifische Gewichtungen anhand des OSMI festgestellt werden?

Um die Forschungsfragen zu beantworten, wird zu Beginn auf die theoretischen Grundlagen eingegangen. Im Anschluss wird die praktische Durchführung vorgestellt. Diese gliedert sich in drei Kapitel. Zuerst soll die Datenbeschaffung und -analyse von Texten aufgezeigt werden. Im nächsten Kapitel wird das Vorgehen der Datenbeschaffung und -analyse von Bildern erläutert, um im Anschluss die genutzte Datenvisualisierung vorzustellen. Zum Schluss sollen die Analyseergebnisse kritisch beleuchtet und in einen Sachzusammenhang gebracht werden.

## **1.2 Verwandte Arbeiten**

Auch wenn der im Rahmen dieser Arbeit entwickelte OSMI an sich eine Neuentwicklung ist, haben sich bereits einige Arbeiten internationaler Wissenschaftlerinnen und Wissenschaftler mit der zugrundeliegenden Thematik befasst. In der Literatur finden sich zahlreiche Publikationen zu Evaluierungsmodellen, die eine Bewertungsgrundlage für Webseiten und weitere Inhalte des Internets zur Verfügung stellen. Die Motivation bei der Entwicklung dieser Modelle variiert dabei stark, je nach Natur der Forschungsfrage.

Obwohl die initialen Fragestellungen aus verschiedenen Branchen stammen, teilen sie jedoch ihre grundlegende Zielsetzung. Die Entwicklung eines zuverlässigen Bewertungsinstruments für die Qualität von Webseiten als Basis für Online-Marketing-Maßnahmen.

Ein Beispiel liefert die Arbeit von Fernández-Cavia et al. mit dem Web Quality Index (WQI), der die Qualität offizieller spanischer und internationaler Tourismus-

Webseiten misst. Zur Ermittlung des WQI wird das Zusammenspiel aus technischen, kommunikativen, beziehungstechnischen und überzeugungstechnischen Aspekten gemessen, womit eine Ausgewogenheit aus quantitativen und qualitativen Daten erreicht werden soll. Diese vier Hauptaspekte werden jeweils in drei ebenfalls zu bewertende Unterkategorien untergliedert, die von den Autoren als wesentliche Gütekriterien für die Qualität von Webseiten angesehen werden. Der Wert des WQI insgesamt und die Wertungen der Unterkategorien bewerten die Stärken und Schwächen und können somit die Verbesserungspotenziale der Tourismus-Webseiten offenlegen. Obwohl der WQI als Vergleichswerkzeug für Tourismus-Webseiten beworben wird, findet auch die Einsicht, dass individuelle Präferenzen einzelner Touristinnen und Touristen nicht berücksichtigt werden können, Ausdruck im abschließenden Urteil der Autoren.<sup>7</sup>

Aus dem Bereich des Online-Marketings und der E-Stores ist die Arbeit von Abdallah und Jaleel anzuführen. Diese befasst sich mit der quantitativen Bewertung der Attraktivität von über achtzig nahöstlichen E-Store-Webseiten und der Wirksamkeit von Online-Marketing-Maßnahmen. Als Grundlage für die Bewertung sollten die Attribute von Webseiten identifiziert werden, die Kaufentscheidungen positiv beeinflussen. Der hier zentrale Begriff des „website appeals“, frei übersetzt als Webseitenattraktivität, beschreibt den Grad, mit dem eine Webseite von Nutzerinnen und Nutzern als angenehm, ansprechend und angemessen wahrgenommen wird. Ein Hauptanliegen der Autoren war es, diesen Begriff näher zu definieren und dessen Bedeutung hervorzuheben. Das entwickelte Bewertungsinstrument basiert auf diesem Konzept.<sup>8</sup> Weiterhin wird eine Einteilung in allgemeine Webseiten-Elemente und Marketing-Mix-Elemente mit fünf bzw. acht Unterkategorien vorgenommen, die als fundamentale Bewertungskriterien herangezogen wurden. Zur Wertermittlung wird die explorative Faktoranalyse (EFA) angewandt. Die ermittelten Hauptfaktoren Anziehungskraft, persönliche Einbeziehung, Konkurrenzfähigkeit und Kundenbindungsfähigkeit werden anschließend als Quadranten in einem zweidimensionalen Modell eingeordnet. Die allgemeine Webseitenattraktivität, die von einem guten ersten Eindruck bis zu einer langfristig beeindruckenden Erfahrung reicht und die Marketingperspektive, die mit der klassischen Verkaufsphilosophie beginnt und mit einem ausgereiften Kundenbeziehungsdenken abschließt, werden als Dimensionen genutzt. Langfristig soll der Einsatz dieser Methode Verbesserungspotenziale der E-Stores offenlegen und so das Erreichen der angestrebten Marketing-Ziele ermöglichen.<sup>9</sup> Die Autoren

---

<sup>7</sup> Vgl. Fernández-Cavia et al. 2014, S. 5–12.

<sup>8</sup> Vgl. Abdallah & Jaleel 2015, S. 45–47.

<sup>9</sup> Vgl. Abdallah & Jaleel 2015, S. 49–55.

wollen mit dem entwickelten EFA-basierten Modell ein vollumfängliches Bewertungsinstrument schaffen, das über oberflächliche visuelle Reize und allgemeine Ästhetik hinausgeht. Der möglicherweise gewonnene gute erste Eindruck sollte durch die ganzheitliche Konzeption der Webseite auf verschiedenen Ebenen zu einem langfristig bindenden Erlebnis weiterentwickelt werden. So soll das Marketingziel der Schaffung einer loyalen Kundenbasis flankiert werden. Weiterhin soll es als Vergleichsinstrument für E-Stores dienen. Die verwendeten Parameter wurden jedoch besonders auf die kaufenden Schichten des arabischen Sprachraums ausgerichtet. Zu den bekannten Problembereichen bei der Arbeit mit Webseiten, wie deren dynamische Entwicklung und der Beständigkeit von Inhalten, gaben die Autoren zu bedenken, dass bei den verwendeten Bewertungskategorien durchaus Interdependenzen bestehen könnten.<sup>10</sup>

---

<sup>10</sup> Vgl. Abdallah & Jaleel 2015, S. 56–57.

## 2 Online Sensory Marketing Index (OSMI)

### 2.1 OSMI-Parameter

Die sensorischen Parameter des Online Sensory Marketing Index (OSMI) basieren auf und richten sich nach den menschlichen Sinnen und sind somit in fünf Parameter definiert, die wiederum durch Indikatoren charakterisiert werden.<sup>11</sup> Die nachfolgende Tabelle illustriert dabei zusammenfassend den schematischen Aufbau des OSMI:<sup>12</sup>

**Tabelle 1:** Übersicht der OSMI-Parameter

Parameter	Anzahl Indikatoren	Kürzel
Haptik	12	H1 – H12
Olfaktorik	2	O1 – O2
Akustik	4	A1 – A7
Gustatorik	5	G1 – G5
Visualität	6	V1 – V6

### 2.2 OSMI-Indikatoren

Der OSMI soll ein Bewertungsinstrument darstellen, das Kriterien der sensorischen Gestaltung von E-Commerce-Webseiten konkludiert und infolgedessen einen raschen Überblick über die sensorische Kommunikationsqualität gewährt und Optimierungspotenziale aufdecken kann.<sup>13</sup> Im Folgenden werden daher Indikatoren auf Basis der Fachliteratur vorgestellt und in einer Übersicht zusammengefasst. Diese Indikatoren sollen demonstrieren können, ob zu untersuchende E-Commerce-Webseiten aus multisensorischer Perspektive erfolgreich kommunizieren. Analog der Studie von Fernandez et al. (2013) basiert der OSMI als Evaluierungsmodell auf Indikatoren, die unterschiedliche Skalen zur Bewertung von Merkmalsausprägungen verwenden: Die Skala 0–1 in Fällen, wenn eine spezifische Charakteristik vorhanden ist oder fehlt. Im Kontrast dazu gewährt die Skala 0–2 respektive 0–3 eine konkretere Bewertung, die gewährt, spezifischere

<sup>11</sup> Vgl. Hamacher & Buchkremer 2022a; Hamacher & Buchkremer 2022b.

<sup>12</sup> Der nachfolgend beschriebene und dem der Untersuchung zugrundeliegende OSMI entspricht einer frühen, unveröffentlichten Version, die nicht identisch mit der von Hamacher und Buchkremer (2022a) veröffentlichten Version ist, jedoch bereits alle Inhalte bzw. Indikatoren enthält.

<sup>13</sup> Vgl. Hamacher & Buchkremer 2022a.

Aussagen hinsichtlich der Qualität einer Charakteristik zu treffen, wie etwa: Schwach (0); Standard (1); Gut (2); Exzellent (3).

Folglich wird in Tabelle 4 je Sinn zusammenfassend dargestellt, aus welchen Indikatoren je Parameter das Evaluierungsmodell OSMI aufgebaut ist. Zusätzlich werden in der linken Spalte neben dem Kürzel (z. B. O für Olfaktorik) ferner auch die möglichen numerischen Ausprägungen je Indikator inkludiert. Im Anschluss werden diese Ausprägungsoptionen durch Indizes näher umschrieben, so dass verdeutlicht wird, unter welchen inhaltlichen Voraussetzungen einer E-Commerce-Webseite die jeweilige Bewertung gegeben werden kann.

**Tabelle 2:** Übersicht der Online-Sensory-Marketing-Index-Indikatoren

Parameter Haptik (H)		
H1 0-3	2-D-Bilder	0 – Schwach: keine oder kleine Bilder 1 – Standard: Super-Zoom-Bilder 2 – Gut: Darstellung aus verschiedenen Blickwinkeln 3 – Exzellent: Verwendung Ich-Perspektive
H2 0-1	3-D-Bilder	0 – Schwach: nicht vorhanden 1 – Standard: vorhanden
H3 0-2	Video (Bewegtbilder)	0 – Schwach: nicht vorhanden 1 – Standard: Produkt in Aktion zu sehen 2 – Gut: Produkt in Aktion und durch Sprecher erklärt
H4 0-3	Handlungsbezogene Formulierungen	0 – Schwach: nicht vorhanden 1 – Standard: konkreter, produktbezogener Kontext 2 – Gut: Verwendung aktiver Verben (Imperativ o. ä.) 3 – Exzellent: narrative Bereitstellung von Content
H5 0-3	Endowment-Effekt	0 – Schwach: nicht erkennbar 1 – Standard: Wecken des Besitzgefühls mit textlichen Hinweisen (ggf. mit Hinweis auf Verlustaversion) 2 – Gut: (1) ergänzt um (hochauflösende) emotionale Bilder

		3 – Exzellent: (2) ergänzt um Bilder mit Ich-Perspektive und/oder 3-D-Bilder
H6 0-1	Bilder, die die Dimensionen/Abmessungen des Produktes zeigen	0 – Schwach: nicht vorhanden 1 – Standard: vorhanden
H7 0-1	Darstellung interpersoneller Berührungen	0 – Schwach: nicht vorhanden 1 – Standard: vorhanden
H8 0-2	Marke (Brand)	0 – Schwach: nicht vorhanden 1 – Standard: vorhanden 2 – Gut: Marke mit hoher Bekanntheit
H9 0-1	Online-Forum	0 – Schwach: nicht vorhanden 1 – Standard: vorhanden
H10 0-3	Positive Stimmung (gute Laune)	0 – Schwach: schlechte Stimmung 1 – Standard: neutral 2 – Gut: enthusiastische Bilder von Personen/Tieren 3 – Exzellent: Bilder aus (2) kombiniert mit Texten
H11 0-1	Interaktiver Chat mit Mitarbeitenden	0 – Schwach: nicht vorhanden 1 – Standard: vorhanden
H12 0-1	Rabatte (price promotion)	0 – Schwach: nicht vorhanden 1 – Standard: vorhanden
<b>Parameter Olfaktorik (O)</b>		
O1 0-3	Mentale Simulation von Düften	0 – Schwach: nur Bild oder nur Text 1 – Standard: Bild und Text vorhanden, aber sensorisch nicht gut aufeinander abgestimmt 2 – Gut: indirekte Formulierungen & Bild 3 – Exzellent: schriftliche Aufforderung & Bild
O2 0-2	Schlüsselbegriffe für Düfte	0 – Schwach: nicht vorhanden 1 – Standard: vorhanden 2 – Gut: kombiniert mit sensorischen Bildern

Parameter Akustik (A)		
A1 0-3	Töne / Musik (u. a. in Verbindung mit Oberflächen und Formen)	<p>0 – Schwach: keine Akustik vorhanden</p> <p>1 – Standard: Töne vorhanden, aber nicht kongruent zu anderen Sinneseindrücken</p> <p>2 – Gut: Einzelne Töne auf andere Sinne abgestimmt</p> <p>3 – Exzellente: Musik sinneskongruent abgestimmt, per Klick abspielbar</p>
A2 0-1	Schlüsselbegriffe für Akustik	<p>0 – Schwach: nicht vorhanden</p> <p>1 – Standard: vorhanden</p>
A3 0-1	Wahrgenommene Klänge über gelesene Wörter	<p>0 – Schwach: Wörter sind gemäß ihrem Klang nicht auf Produktkategorie sinnlich angepasst</p> <p>1 – Standard: Wörter sind gemäß ihrem Klang auf Produkt(-kategorie) sinnlich angepasst</p>
A4 0-3	Sprecher / Stimme	<p>0 – Schwach: Kein Sprecher / Stimme integriert</p> <p>1 – Standard: Sprecher / Stimme integriert</p> <p>2 – Gut: Sprecher / Stimme ist sinnlich abgestimmt</p> <p>3 – Exzellente: Sprecher / Stimme sinneskongruent abgestimmt und verbreitet positive Stimmung</p>
Parameter Gustatorik (G)		
G1 0-3	Mentale Simulation des Schmeckens	<p>0 – Schwach: keine mentale Simulation möglich</p> <p>1 – Standard: Bild und Text vorhanden, aber nicht aufeinander abgestimmt</p> <p>2 – Gut: Bild und indirekte Formulierungen aufeinander abgestimmt</p> <p>3 – Exzellente: Bild und schriftliche Aufforderung zur Vorstellung des Geschmacks</p>
G2 0-1	Verwendung von Farbschemata	<p>0 – Schwach: nicht vorhanden bzw. in Graustufe oder zu viele, bunt gemischte Farben</p> <p>1 – Standard: Eine Farbe sinnlich kongruent angewandt</p>
G3 0-1	Anpassung / Gestaltung des Marken- / Produktnamens	<p>0 – Schwach: Marke und oder Produktname nicht auf Produktkategorie sinnlich angepasst</p> <p>1 – Standard: Produktkategorie und Name der Marke bzw. des Produkts sinnlich abgestimmt</p>

G4 0-1	Typografie der Produktbeschreibung	0 – Schwach: Typografie ist nicht auf Produktgeschmack abgestimmt 1 – Standard: Typografie ist auf Produktgeschmack abgestimmt
G5 0-1	Produktpolitik (inkl. Verpackung) und Oberflächen	0 – Schwach: Produktlayout (Gesamt) ist nicht auf Produktgeschmack abgestimmt 1 – Standard: Produktlayout (Gesamt) ist auf Produktgeschmack abgestimmt
<b>Parameter Visualität (V)</b>		
V1 0-3	Farbgebung	0 – Schwach: nicht abgestimmt 1 – Standard: abgestimmt auf Produkt 2 – Gut: abgestimmt auf Produkt und Marke 3 – Exzellente: geschlechterspezifische Farbschemata
V2 0-1	Schlüsselbilder	0 – Schwach: nicht vorhanden 1 – Standard: vorhanden
V3 0-1	Dynamische Bilder	0 – Schwach: nicht vorhanden 1 – Standard: vorhanden
V4 0-1	Kontrast der Bilder / Webseite	0 – Schwach: nicht abgestimmt 1 – Standard: starker Kontrast
V5 0-1	Oberfläche (Glanz vs. Matt)	0 – Schwach: nicht aufeinander abgestimmt 1 – Standard: vorhanden
V6 0-1	Handschriftliche Informationen	0 – Schwach: nicht vorhanden 1 – Standard: vorhanden

### 2.2.1 Haptische Indikatoren

Die Haptik als ein Parameter des OSMI weist eine Dominanz in Bezug auf die Quantität sowie der Ausprägungen der haptischen Elemente gegenüber den anderen Parametern auf. Insgesamt wird der Parameter Haptik durch zwölf Indikatoren definiert. Hierbei basieren die haptischen Indikatoren auf zahlreichen Studien und den darin festgehaltenen Erkenntnissen. Die Forschung suggeriert im Bereich der Haptik, dass insbesondere Bilder für Evaluierungen der Haptik einen entscheidenden Stellenwert innehaben. Insofern sind die Indikatoren H1 und H2 konkret auf visuelle Inhalte von E-Commerce-Webseiten ausgelegt, die eine Bewertung der haptischen Eigenschaften von Produkten ermöglichen sollen. H1 umfasst 2-D-Bilder, die heute dem gängigen Webseiten-Standard zurechenbar



sind, sich jedoch auch heute noch qualitativ stark in der sensorischen Ansprache von Konsumentinnen und Konsumenten unterscheiden können. Um eine gute Bewertung für sensorisch gutes Bildmaterial vergeben zu können, sollte die Darstellung des Produkts aus diversen Blickwinkeln ebenso wie eine hochauflösende Strukturablichtung mittels (2-D-) Super-Zoom-Bildern vorhanden sein. Insbesondere solche Bilder bieten die Option, taktile Informationen in Bezug auf das Gewicht, die Dicke (z. B. von Stoffen eines Kleidungsstücks), das Material sowie die Textur – z. B. Weichheit und Oberflächenstruktur wie etwa Glätte/Rauheit – detailliert veranschaulichen zu können. Über Beschreibungen hingegen können in der Regel deutlich weniger Informationen zu Produkten transportiert werden. Aus diesem Grund sind 2-D-Bilder als effektiver zu erachten. Dabei stellt die höchste Stufe für eine virtuelle Produkterfahrung mit statischen Bildern jene dar, die eine „Ich-Perspektive“ beinhaltet und auf diese Weise eine lebhaftere mentale Imagination erzeugt. Dies bedeutet, dass die Betrachterinnen und Betrachter der Webseite anhand der Darstellungsweise eine „quasi-sinnliche Erfahrung“ erleben können, die häufig auf bereits erlebten Erfahrungen beruht und eine mentale Simulation der Verwendung eines illustrierten Produkts erzeugt. Hierzu sind spezifische Assoziationsoptionen wie etwa zur Händigkeit zu nennen, d. h. die Art und Weise, wie ein Produkt dargestellt ist. Insbesondere die individuelle Präferenz der Konsumentinnen und Konsumenten beispielsweise bezüglich der Rechts- oder Linkshändigkeit hat eine entsprechende Bedeutung für die Positionierung von Produkten auf Werbebildern. Daher kann eine Übereinstimmung zwischen der Objektorientierung und der Händigkeit die mentale Simulation einer Interaktion mit einem Produkt maßgeblich beeinflussen und infolgedessen zu einer Erhöhung der Kaufabsicht führen.<sup>14</sup>

Ferner kann Drehen, Bewegen sowie Zoomen von Produkten in 3-D-Qualität haptische, wenn auch begrenzte Vorstellungen der Verwendung ermöglichen, die durch Indikator H2 abgedeckt sind. Durch 3-D-Bilder und Animationen können Interaktivität und dadurch lebhaftere mentale Bilder der Verwendung von Produkten hervorgerufen werden, die zu höheren Kaufabsichten führen können, als wenn die gleichen Informationen lediglich passiv (beispielsweise über Bilder und/oder Texte) zur Verfügung gestellt werden. Auch Bilder, die Abmessungen von Produkten zeigen, unterstützen die Vorstellungskraft. Der Indikator H6 bezieht sich folglich auf solche Bildmaterialien, die z. B. in einer Simulation zeigen, wie ein Mensch das Produkt in den Händen hält, um so die Dimensionen zu ver-

---

<sup>14</sup> Vgl. Krishna et al. 2014, S. 142.

anschaulichen. Dies ist jedoch nicht gleichzusetzen mit reinen Maßzeichnungen, die im Sinne der Sensorik schlechter zu bewerten wären.

Aufbauend auf den zuvor erläuterten statischen Bildquellen lassen sich mithilfe von Videomaterial (H3) Produktinformationen noch besser interpretierbar darstellen. Die Fachliteratur suggeriert dabei, dass gerade die Produktnutzung und/oder die Produktherstellung sich besonders anbieten, um per Videoformat vermittelt zu werden. Inhaltlich sollten dabei spezifische Informationen (Qualität, Ursprungsland, Marke, Garantie etc.) im Videomaterial verarbeitet werden.<sup>15</sup> Somit kann das teilweise bei den Betrachtenden vorhandene Bedürfnis zur Berührung aus Gründen der Qualitätsverifizierung des gezeigten Produkts gemindert werden.

Neben Bildern und Videos können auch rein textliche Elemente Informationen vermitteln, wie z. B. die Materialspezifikationen, aber auch sensorische Beschreibungen. Aufgrund dessen werden über den Indikator H4 handlungsbezogene Formulierungen als sensorische Beschreibung in den OSMI inkludiert. Diese wirken ähnlich wie Bilder und können die Aktivität motorischer Hirnareale anregen. Das Ausmaß dessen kann derart ausgestaltet sein, als würden die Betrachterinnen und Betrachter die Tätigkeit, wie etwa in die Hände klatschen, tatsächlich im Moment des Lesens ausführen. Hierbei sollten die angewandten Begriffe allerdings in einem konkreten, sachbezogenen Kontext verfasst sein.<sup>16</sup> Auch aktive Verben anstelle von Substantivierungen sind aus den zuvor genannten Gründen von Vorteil.<sup>17</sup> Besondere Bedeutung kommt hierbei dem Motorcortex zu, der neben der realen Ausführung von Bewegungen, ebenso deren Vorstellung erleichtern kann. Es ist folglich empfehlenswert, eher Formulierungen wie die Folgenden zu verfassen: „Riechen Sie den frischen Kaffee?“ anstelle von „frischer Kaffee“. In Ergänzung dazu sollten nicht eindeutig definierte Begrifflichkeiten wie z. B. „geschmeidig“ vermieden werden, da sie in der Regel individuell interpretiert werden und somit abstrakt sind. Unterstützt werden die bisherigen Erläuterungen durch eine Untersuchung von McCabe und Nowlis (2003), in der diese aufzeigen konnten, dass Produktbeschreibungen ebenso wie Werbebotschaften stärker wahrgenommen werden, wenn diese eine konkrete und bildhafte Beschreibung eines sensorischen Erlebnisses beinhalten. Diese Erkenntnisse suggerieren, dass Unternehmen inhaltlich spezifischere haptische Informationen darbieten müssen als die reine Aufzählung von Produktattributen.<sup>18</sup> Ebenso weisen Lee et

---

<sup>15</sup> Vgl. San-Martín et al. 2017, S. 962.

<sup>16</sup> Vgl. Raposo et al. 2009, S. 394.

<sup>17</sup> Vgl. Bower 2004, S. 83.

<sup>18</sup> Vgl. McCabe & Nowlis 2003, S. 438.

al. (2018) darauf hin, dass sensorisch formulierte Werbetexte im direkten Vergleich zu üblichen, nicht-sensorischen Texten eine deutlich positivere Resonanz bei den Konsumentinnen und Konsumenten hervorrufen.<sup>19</sup> Hiernach ist es von Vorteil, eine sensorisch basierte Erzählform einer reinen Produktbeschreibung vorzuziehen, denn insbesondere das Vorstellen von Produkten und Leistungen in ansprechender Erzählform trägt dazu bei, dass sich Konsumentinnen und Konsumenten vorstellen können, ein Teil jener Erzählung zu sein.<sup>20</sup> Daher kann eine narrative Bereitstellung von Inhalten die Fähigkeit der Konsumierenden verstärken, wichtige Inhalte abzurufen und zukünftiges Verhalten zu beeinflussen<sup>21</sup> und wird deshalb mit der höchsten Bewertungsstufe (Exzellent) ausgestattet.

Der Endowment-Effekt (H5) symbolisiert als Indikator die Dauer/Zeitspanne der vorgestellten Berührung und ist ebenfalls von hoher Bedeutung in der Customer Journey.<sup>22</sup> Dies müsste zwar grundsätzlich zunächst spezifisch gemessen werden, jedoch kann im Rahmen der Evaluierung eine bessere Bewertung vorgenommen werden, wenn 3-D-Inhalte auf der E-Commerce-Webseite verfügbar sind. Denn es konnte nachgewiesen werden, dass diese Inhalte zu einem längeren Betrachtungsintervall beitragen und einen positiven Einfluss auf die Kaufabsicht haben.<sup>23</sup> Gleichwohl ist ein Endowment-Effekt nicht in einem kausalen Zusammenhang zur Haptik zu setzen, sondern in umgekehrter Wirkungsweise. Demnach können Berührungen von Produkten zu einer Verstärkung des Endowment-Effekts beitragen.<sup>24</sup> Der Indikator H5 ist folglich primär an das Vorhandensein von Indikator H2 geknüpft und wird als nicht kausaler Indikator unter dem Parameter der Haptik untergeordnet.

Faktoren hinsichtlich der wahrgenommenen Sicherheit wie etwa Produktbewertungen anderer Nutzerinnen und Nutzer (z. B. bei Adidas oder Otto), kostenlose Retouren etc. sind ebenfalls in weit ausgelegter Interpretation für die Haptik im E-Commerce-Segment relevant. Ferner sind Faktoren wie Online-Foren respektive Chats von Bedeutung, welche die Unsicherheit der Konsumentinnen und Konsumenten lindern können. Die Studie von Gonzalez et al. (2015) zeigt in diesem Kontext, dass führende Marken, unabhängig von objektiven Produkteigenschaften, einen Vorteil bei der Qualitätsbewertung der Verbraucherinnen und Verbraucher genießen. Auf E-Commerce-Webseiten kann dieser Vorteil größer

---

<sup>19</sup> Vgl. Lee et al. 2018, S. 74.

<sup>20</sup> Vgl. Lee et al. 2018, S. 69.

<sup>21</sup> Vgl. Lee et al. 2018, S. 74.

<sup>22</sup> Vgl. Peck et al. 2013; Shu & Peck 2007.

<sup>23</sup> Vgl. Brasel & Gips 2015, S. 232.

<sup>24</sup> Vgl. Peck & Shu 2009.

sein, allerdings nur für Produktkategorien, bei denen der fehlende physische Kontakt mit dem Produkt während des Kaufprozesses eine erhebliche Einschränkung darstellt.<sup>25</sup> In solchen Fällen können den Autoren zufolge Markenassoziationen respektive das Vertrauen in die Marke und die Qualität die fehlende Option zur haptischen Beurteilung vor dem Kauf kompensieren.

Die Haptik sollte schließlich auch Freude bereiten oder neutrale respektive positive Assoziationen wecken. Ein solcher Effekt konnte bereits durch Studien mit Printmaterialien gezeigt werden.<sup>26</sup> Dieser Faktor wird infolgedessen durch den Indikator H10 aufgegriffen. Die Argumentation wird unterstützt durch die Studie von Yazdanparast und Spears (2013), die untersucht haben, ob Verbraucherinnen und Verbraucher im Kontext von Onlinekäufen darauf verzichten können, Produkte zu berühren. Sie schilderten dabei nicht-haptische Situationsfaktoren und untersuchten unter anderem, ob ein positiver Gemütszustand („gute Laune“) die fehlenden Berührungsoptionen im Rahmen von E-Commerce-Käufen kompensieren kann.<sup>27</sup> Sie konnten dabei aufzeigen, dass dies bei Personen mit hohem Drang zur Berührung von Objekten der Fall ist. Dies bedeutet, dass eine gute Laune folglich Frustration kompensieren kann, die aufgrund mangelnder Haptik zuvor entstanden ist. So können unter anderem ein ansprechendes Design des Online-Shops, positive Produktbeschreibungen respektive -bewertungen, Chats mit sympathischen Beraterinnen und Beratern, reichlich Humor und/oder anregende Bilder Personen mit hohem „Need For Touch“ für das Online-Shopping gewinnen. Ebenfalls kann Fantasie eine Möglichkeit der Kompensation fehlender Berührungsoptionen sein. Dies funktioniert allerdings nicht bei Personen mit einem hohen Drang zur Berührung von Objekten. Schließlich kann Fantasie allein nicht die faktische Berührung ausgleichen.<sup>28</sup>

### 2.2.2 Olfaktorische Indikatoren

Der Parameter Olfaktorik ist zwar grundsätzlich von enormer Bedeutung für sinnliche Erfahrungen, jedoch ist dessen Abbild in Indikatoren zur Messung des sensorischen Kommunikationserfolgs auf E-Commerce-Webseiten im direkten Vergleich zu den anderen Parametern aufgrund technischer Limitationen stark eingeschränkt und umfasst lediglich die folgenden zwei Indikatoren:

---

<sup>25</sup> Vgl. González-Benito et al. 2015, S. 122.

<sup>26</sup> Vgl. Peck & Wiggins 2006, S. 65.

<sup>27</sup> Vgl. Yazdanparast & Spears 2013, S. 59

<sup>28</sup> Vgl. Spears & Yazdanparast 2014, S. 370.

Gerüche sind zeitlich äußerst stabil. Sie rufen durch enge Verknüpfungen mit Erinnerungen zahlreiche sensorische Assoziationen hervor. Umso anspruchsvoller ist es, einen Geruch in der digitalen Welt „riechbar zu machen“, ohne dass dieser tatsächlich über die Nase der Konsumentinnen und Konsumenten wahrzunehmen ist. Krishna et al. (2014) befassen sich jedoch genau mit dieser Problematik. Die Autoren untersuchten in ihrer Studie, auf welche Weise Werbung in Form von Werbeanzeigen Menschen dazu ermutigen kann, sich Gerüche vorzustellen und ferner wie imaginierte Gerüche ähnlich wie bei tatsächlichen Gerüchen wirken können.<sup>29</sup> Sie untersuchten folglich, inwiefern Gerüche ohne eine visuelle Referenz vorstellbar sind. Zusammenfassend lässt sich konstatieren, dass die reine Aufforderung, sich vorzustellen, wie ein beworbenes Produkt (im Falle der Studie ein Gericht) riechen könnte, dazu führen kann, dass das Verlangen nach dem beworbenen Produkt wächst. Allerdings funktionierte dies in der Studie nur, sofern die Anzeige ein Bild des Essens beinhaltete. Der Indikator O1 des OSMI bildet diese Untersuchungsergebnisse ab und entspricht in seiner höchsten Evaluierungsstufe einer Kombination aus Bild und schriftlicher Aufforderung zur Geruchsvorstellung. Etwas seltener, aber dennoch als vorteilhaft zu beurteilen, sind indirekte Formulierungen. Dies ist der Fall, wenn die Betrachterinnen und Betrachter wie nachfolgend per Text beispielhaft gefragt, aber nicht konkret zu etwas aufgefordert werden: „Kennen Sie den Geruch von frischem Kaffee am Morgen?“

Von hoher olfaktorischer Bedeutung sind neben handlungsbezogenen Formulierungen auch duftrelevante Schlüsselbegriffe, denn solche Begriffe können (z. B. das Wort „Basilikum“) nicht selten nur anhand des einzelnen geschriebenen Wortes bereits umfangreiche Assoziationen hervorrufen und die Betrachterinnen und Betrachter können sich nicht nur den Geruch, sondern auch alle damit verbundenen Erinnerungen gedanklich ins Bewusstsein rufen. Im Sinne der höchsten Bewertungsstufe des Indikators O2 kann dieser Effekt darüber hinaus mit sensorisch stimmigen Bildern nochmals verstärkt werden.

### **2.2.3 Auditive Indikatoren**

Der Parameter der Akustik ist im Vergleich zur Gustatorik wiederum mit mehr Indikatoren versehen. Die Ursache hierfür liegt jedoch u. a. daran, dass die Akustik grundsätzlich über Lautsprecher im Bereich des E-Commerce bereits seit jeher Anwendung finden kann. Allerdings ist anzumerken, dass zwischen Nutz-

---

<sup>29</sup> Vgl. Krishna et al. 2014, S. 26.

und Störschall unterschieden werden sollte. In der Regel fühlen sich Betrachterinnen und Betrachter einer E-Commerce-Webseite von nicht aktiv angeklickten akustischen Elementen eher gestört. Aufgrund dessen sind die akustischen Indikatoren als solche zu verstehen, die ausschließlich auf klickbaren und damit seitens der Betrachterinnen und Betrachter bewusst gewollten akustischen Signalen beruhen. Vier Indikatoren hinsichtlich dieses Parameters konnten anhand der Literatur identifiziert werden, die wie folgt in das allgemeine OSMI-Evaluierungsmodell Einzug finden.

Der Indikator (A1) bezieht sich auf per Klick abspielbare Musik bzw. Töne. Auch hier fließen Erkenntnisse aus der Fachliteratur in die Ausprägungsmodalitäten ein. So gilt als erwiesen, dass weiche Musik im Kontrast zu harter oder gänzlich fehlender Musik einen positiven Einfluss auf die wahrgenommene Weichheit von Produkten ausüben kann, da diese signifikant steigt.<sup>30</sup> Jedoch muss in diesem Zusammenhang eine affektive Komponente berücksichtigt werden, da dieser Effekt revidiert wird, wenn sich die Betrachterinnen und Betrachter aktiv und kognitiv über die gehörte Akustik Gedanken machen.<sup>31</sup>

Analog zu den olfaktorischen Schlüsselbegriffen existieren auch zahlreiche akustische Schlüsselbegriffe (A2), wie z. B. der Begriff Klavier, bei deren Nennung und Erkennung durch die Besucherinnen und Besucher auf der jeweiligen Webseite bereits detaillierte Assoziationen zu den erzeugten akustischen Elementen mental hervorgerufen werden können. Interessanterweise haben Wörter allgemein zum Teil spezifische Bedeutungen hinsichtlich der Aussprache sowie der damit verbundenen erzeugten Klänge und werden im Indikator A3 des OSMI berücksichtigt. Dieser deckt die sinnliche Kongruenz bezüglich der Wortwahl ab. Worte sollten folglich so gewählt werden, dass die Vokale darin und die damit assoziierten sensorischen Eigenschaften hinsichtlich des Geschmacks, der Optik sowie der Akustik kongruent sind.

Ebenso können Stimmen in den akustischen Inhalten von E-Commerce-Webseiten einen positiven Effekt auf die multisensorische Ansprache von Konsumentinnen und Konsumenten ausüben. Zumal es per se von Vorteil sein kann, wenn eine Stimme oder eine sprechende Person vorhanden ist, wodurch eine menschlichere Kommunikationsebene dargeboten wird. Auch sollte das kommunizierte Informationsmaterial sinnlich auf das zu bewerbende Produkt designet sein. Damit ist gemeint, dass die Sprechenden insbesondere ihre Lautstärke und das Tempo der Sprache daran ausrichten sollten. Nichtsdestotrotz spielt auch die

---

<sup>30</sup> Vgl. Stout & Leckenby 1988, S. 232.

<sup>31</sup> Vgl. Homburg et al. 2013, S. 12.

Tonlage der verwendeten Stimme eine immense Rolle, die sich darin widerspiegelt, dass eine tiefe Stimme nicht selten als angenehmer wahrgenommen wird und deshalb mit der höchsten Evaluierungsstufe innerhalb des Indikators A4 berücksichtigt wird.

#### **2.2.4 Gustatorische Indikatoren**

Insgesamt determinieren fünf gustatorische Indikatoren den Parameter Gustatorik. Der Indikator G1 fußt, wie auch schon bei den zuvor erläuterten Indikatoren anderer Sinnesreize, auf die mentale Simulation von gustatorischen Reizen. Daher können, analog olfaktorischer Reize, ebenso Reize des Schmeckens effektiv simuliert werden. Hierzu ist jedoch eine detailliertere Evaluierung notwendig, die primär untersucht, ob sowohl Text und Bild auf einer Webseite verwendet werden und ob diese zum Vorstellen eines Geschmacksvorgangs animieren. Darüber hinaus wird dahingehend differenziert, inwiefern die textliche Komponente der betrachteten E-Commerce-Webseite dazu anregt, den avisierten Geschmack mental zu simulieren. Hierzu tragen indirekte Formulierungen bei wie z. B. „Der Geschmack von diesem frisch gebackenen Brot am Morgen ist ein Genuss für jeden Gaumen“. Diese Formulierung ist zwar grundsätzlich von Vorteil und trägt zu einer Imagination bei den lesenden Betrachterinnen und Betrachtern der Webseite bei. Eine direkte Aufforderung in Schriftform wird allerdings als noch effektiver bewertet, sofern darin aufgefordert wird, sich vorzustellen, wie der Geschmack wahrnehmbar sein könnte. Hierfür wird die höchste Evaluierungsstufe (3) des Indikators G1 vergeben. In einem solchen Fall sind Formulierungen wie z. B. die folgende zu wählen: „Stellen Sie sich vor, wie das frisch gebackene Brot am Morgen Ihren Gaumen berührt und Sie die Frische herauschmecken können“.

In puncto Geschmack ist auch die Verwendung von Farben, neben der mentalen Simulation, auf der E-Commerce-Webseite zu evaluieren und wird in Indikator G2 definiert. Eine sinneskongruente Farbkombination hinsichtlich des zu bewerbenden Produkts kann einen hohen Stellenwert einnehmen. Der Evaluierungslogik für Farben wird die Übersicht von Spence et al. (2016) zugrunde gelegt, die die allgemeine Farbassoziationen der vier Grundgeschmäcker salzig, bitter, sauer und süß beinhaltet.<sup>32</sup> Der Indikator G2 besitzt wissentlich nur eine Skala

---

<sup>32</sup> Vgl. Spence et al. 2016.

zwischen 0 und 1, denn eine Bildverwendung, gemäß den Forschungserkenntnissen, kann heutzutage als Standard betrachtet werden, eine spezifischere Bewertung erscheint ferner als nicht sinnvoll.

Ebenso kann das Design eines Marken- bzw. Produktnamens mitunter einen großen Einfluss auf geschmackliche Assoziationen auslösen. Dies ist in der Regel gegeben, wenn solche Begriffe gelesen und ausgesprochen werden. Für würzige und mild schmeckende Produkte ist es daher ratsam, eine andere Formulierung zu wählen als für knusprige sowie scharf schmeckende Produkte. Auch wird mit null oder eins im Rahmen des Indikators G3 klassifiziert, da eine grundsätzliche Ausrichtung dessen zu erwarten ist.

Ferner können Formen einen Einfluss auf Geschmacksassoziationen ausüben. Darunter ist zu verstehen, dass z. B. milde Produkte auch mit rundlichen Formen aus dem Verpackungsdesign sowie der Schriftart versehen werden sollten. Schließlich sollten neben der Typografie (G4) ebenso das Produktdesign (G5) auf sinnliche Kongruenz hin entwickelt werden. Innerhalb des Bewertungsmodells des OSMI ergibt sich eine gute Bewertung, sofern scharfe, schnelle Töne mit rauen Oberflächen oder sanfte, rund-klingende Töne mit weichen Oberflächen sinnlich kongruent kombiniert werden. Es wird positiv bewertet, wenn kantige und spitze Formen zum angebotenen Produkt passen, die entweder Bitterkeit, Säure, Härte oder Knusprigkeit beinhalten. Alternativ wäre es ebenso möglich, bei cremigen und süßen sowie weichen Produkten runde Formen für das Produkt zu verwenden, um innerhalb des OSMI eine gute Bewertung zu erhalten.

### **2.2.5 Visuelle Indikatoren**

Der Sehsinn ist zusammen mit der Akustik der zweite und damit aktuell letzte Sinn, der unmittelbar über E-Commerce-Webseiten tangiert werden kann. Der visuelle Sinn ist jener, der prozentual gesehen am intensivsten angesprochen wird, da zum einen die anderen Sinne nicht direkt ansprechbar sind (nur über Imagination) und akustische Webseiteninhalte – wie zuvor erläutert – nur per bewusstem Klick und daher nicht automatisiert wahrnehmbar sind. Die Dominanz visueller sensorischer Webseitenkomponenten ist offensichtlich und bietet die Möglichkeit, als eine Art Vermittler der anderen Sinnesmodalitäten zu fungieren. Im Rahmen des Online-Kaufenerlebnisses nehmen Betrachterinnen und Betrachter einer Webseite folglich primär das wahr, was ihre Augen sehen. Demnach werden die folgenden Indikatoren beschrieben, die eine Erfolgsmessung im Rahmen der multisensorischen, visuellen Ansprache der Konsumentinnen und Konsumenten initiieren können:



V1 als erster Indikator des Sehens beruht auf dem Farbdesign von E-Commerce-Webseiten. Ein besonderes Augenmerk sollte im Rahmen von V1 auf eine generelle Optik und deren Auswirkung auf die wahrgenommene Grundstimmung bei den Betrachtenden der Webseite gerichtet sein. Hierbei muss differenziert werden, dass die Farbgebung für gustatorische Assoziationen mit V1 nicht überprüft wird, da dies bereits separat durch G2 abgedeckt ist. Farben dienen in diesem rein visuellen Kontext vielmehr dazu, zu verstehen, dass beispielsweise die Farbe Rot den menschlichen Puls erhöhen kann und bei den Betrachtenden gewissermaßen Anspannung erzeugt sowie auf vermeintliche Schnäppchen hinweist. Die Farbe Blau wiederum löst in der Regel das genaue Gegenteil aus. In Bezug auf das Farbdesign und eine damit verbundene Farbanalyse muss konstatiert werden, dass die Variable V1 ausschließlich auf einer gängigen, europäischen Farbassoziation beruht. Insofern kann beispielsweise angeführt werden, dass die Farbe Weiß in Asien mit negativen Gedanken verknüpft wird (z. B. in der Bedeutung des Todes).

Auch im visuellen Sinnesbereich spielen Schlüsselbilder eine erhebliche Rolle und werden durch V2 in der Evaluierung berücksichtigt. Gemeint sind in erster Instanz Bilder, auch „Key Visuals“ genannt, die mehrere Sinne simultan ansprechen können. Derartige Schlüsselbilder können als Großbildaufnahmen verschiedener Szenarien in Erscheinung treten, indem die Natur oder Ähnliches illustriert und Assoziationen wie Freiheit, Frische, ein positives Lebensgefühl etc. vermittelt werden. Darüber hinaus ist auch die Dynamik von Bildern nicht außer Acht zu lassen. Sie wird über den Indikator V3 abgedeckt. Roggeveen et al. kommen in ihrer Studie zu dem Ergebnis, dass ein dynamisches Präsentationsformat in Bildform die Betrachtenden automatisch veranlasst, sich die Produkterfahrung mental vorzustellen. Hierdurch wird das Involvement gestärkt.<sup>33</sup> Auch Cian et al. (2014) argumentieren in diese Richtung, da sie sich auf die Fähigkeit eines statischen visuellen Bildes anhand von Logos konzentrierten, das Bewegung vermitteln soll, ohne sich tatsächlich im Sinne von statischen, aber dynamischen Bildern zu bewegen.<sup>34</sup>

Zusätzlich zum gewählten Kontrast (V4) auf einer E-Commerce-Webseite wird durch den OSMI ferner auch der abgebildete Glanz in die Bewertung aufgenommen (V5). Der Glanz hat tendenziell sensorisch eine Bedeutung und Meert et al. (2014) zufolge werden glänzende Produkte durch die Betrachterinnen und Betrachter generell besser bewertet. Der Grund dafür kann darin gesehen werden,

---

<sup>33</sup> Vgl. Roggeveen et al. 2015, S. 43.

<sup>34</sup> Vgl. Cian et al. 2014, S. 188.

dass Wasser seit jeher als essenzielle Ressource zum Überleben fungiert. Meert et al. legen indes nahe, dass die Verwendung von Wasser in der werblichen Kommunikation durch z. B. Bilder in der Tat ein positives Gefühl auslösen kann und somit auch nachfolgende (Kauf-)Entscheidungen beeinflussen kann.<sup>35</sup>

Ergänzend können handschriftliche Informationen im Sinne eines sensorischen Reizes das Gefühl eines persönlicheren, visuellen Erlebnisses fördern.<sup>36</sup> Ein solches sensorisches Kommunikationselement ist daher nützlich zur sensorischen Inszenierung einer E-Commerce-Webseite und wird durch den Indikator V6 ausgedrückt.

### **2.3 Indizes**

Die in Kapitel 2.1.2 erläuterten Indikatoren der fünf Parameter des OSMI-Evaluierungsmodells müssen zum Zwecke der Operationalisierung und für eine spätere Vergleichbarkeit skaliert werden. Die Qualität des zu messenden sensorischen Inhalts von E-Commerce-Webseiten soll schließlich durch eine zusammenfassende Kennziffer – den Online Sensory Marketing Index – ausgedrückt werden. Diese Kennziffer ist durch eine Skalierung im Bereich von null bis eins operationalisiert und soll auf diese Weise eine Vergleichbarkeit zwischen Messergebnissen ermöglichen. Üblicherweise sind Indikatoren in der Regel auch zu gewichten. Jeder Indikator bekäme demnach ein relatives Gewicht entsprechend seiner Bedeutung innerhalb des betrachteten Parameters zugewiesen. Eine solche Gewichtung von Indikatoren entsteht zumeist durch Diskussion und Verhandlung des Wertes. In der verwandten Arbeit von Fernandez et al. wurde der Web Quality Index auf Basis von Meinungen von Expertinnen und Experten innerhalb eines Forschungsteams und anhand der vorhandenen Fachliteratur durchgeführt. Der OSMI jedoch verzichtet derzeit gänzlich auf eine solche Gewichtung. Die Gründe dafür sind breit gefächert. Zwar legt die Vielzahl der eruierten wissenschaftlichen Beiträge nahe, dass z. B. der Parameter Haptik höher zu gewichten ist, jedoch variiert das Bedürfnis nach Berührung von Mensch zu Mensch und gemäß der betrachteten Branche. Dies könnte schließlich dazu führen, dass E-Commerce-Webseiten in der Evaluierung einen guten OSMI-Score zugewiesen bekommen, wenn sie haptische Elemente in den Fokus rücken, aber andere Sinne nur sekundär behandeln, welche aber mittels einer hohen Gewich-

---

<sup>35</sup> Vgl. Meert et al. 2014, S. 203.

<sup>36</sup> Vgl. Hultén 2011, S. 268.

tung mathematisch ausgeglichen werden. Auch ist nicht zweifelsfrei zu argumentieren, ob eine höhere Gewichtung einzelner Indikatoren der Parameter angemessen ist oder nicht. Der verwendete OSMI wird kontinuierlich weiterentwickelt, so auch die Abgrenzung der zahlreichen Indikatoren und deren jeweilige Relevanz, um künftig eine Gewichtung zu ermöglichen, welche in dieser Arbeit jedoch nicht zum Tragen kommt.

Die Bewertung der einzelnen Indikatoren mündet in einen zwischen null und eins basierten Index für jeden Parameter und wird wie jede Dezimalzahl in dieser Arbeit auf zwei Nachkommastellen gerundet. Die eruierten Werte auf Basis der vorgefundenen Charakteristika einer E-Commerce-Webseite liegen somit stets zwischen null und eins für jeden sinnlichen Parameter und jede der analysierten Webseiten. Abschließend wird der Gesamt-Index (OSMI) für jede Webseite berechnet, ohne Gewichtung und durch die Anwendung des arithmetischen Mittels. Dieser ist ebenso zwischen null und eins normiert. Der kombinierte OSMI-Wert kann als Gesamtbewertung der sensorischen Kommunikationsqualität der analysierten E-Commerce-Webseite verstanden werden. Der Logik folgend ist diese Kennziffer so zu verstehen, dass je näher der Wert des OSMI bei 1 liegt, desto erfolgreicher kommuniziert die analysierte E-Commerce-Webseite aus sensorischer Sicht. Eine Punktzahl nahe null lässt wiederum erkennen, dass die Webseite wichtige sensorische Elemente nicht enthält und/oder die Indikatoren für erfolgreiche multisensorische Ansprache der Konsumentinnen und Konsumenten nicht erfüllt, die für eine E-Commerce-Webseite jedoch nötig sowie relevant erscheinen. Die Interpretation des OSMI-Wertes sollte jedoch stets vor dem Hintergrund eines möglichen Sensory Overloads getätigt werden, denn ein Wert nahe eins muss nicht zwingend für Erfolg stehen. Vielmehr sind auch die einzelnen Werte pro Parameter von Interesse. Einzelne Studien auf dem noch rudimentären Forschungsgebiet des Sensory Overloads deuten bereits darauf hin, dass es nicht erforderlich ist, alle fünf menschlichen Sinne werblich gleichermaßen anzusprechen. Der generelle Vorteil der hier gewählten Methode lässt sich abschließend darin begründen, dass die Indizes des OSMI generelle Vergleiche zwischen den analysierten Webseiten möglich machen und diejenigen Webseiten identifiziert werden können, welche in der Branche sensorisch federführend kommunizieren. Es können durch den OSMI Beispiele für gute und weniger gute bis schlechte sensorische Gestaltungen identifiziert und bewertet sowie miteinander verglichen werden. Verbesserungspotentiale für multisensorische Optimierungen sind durch diese Vorgehensweise ebenso aufzeigbar.

## 2.4 Projektprozess und Organisation

Dieser Abschnitt befasst sich mit dem Prozess innerhalb des Projekts, der Projektorganisation und -leitung, die zur Anfertigung dieser Arbeit angewandt wurden. Der Projektauftrag mit den wissenschaftlichen Fragestellungen wurde von Kevin Hamacher schriftlich verfasst und von Prof. Dr. Rüdiger Buchkremer, dem Dozenten, in der ersten von insgesamt fünf vorgesehen Vorlesungen im Rahmen des Big-Data-Analyseprojekts den Studierenden vorgestellt. Der zeitliche Rahmen zur Fertigstellung war auf ein Semester – fünf Monate – beschränkt. Die zur Verfügung stehenden Personen-Ressourcen waren durch die Anzahl der Studierenden gegeben. Weiteres Projektbudget stand nicht zur Verfügung. Somit wurden alle Ergebnisse mithilfe von kostenfreier Software, Studierendenlizenzen und den Hard- und Software-Ressourcen der Studierenden erarbeitet.

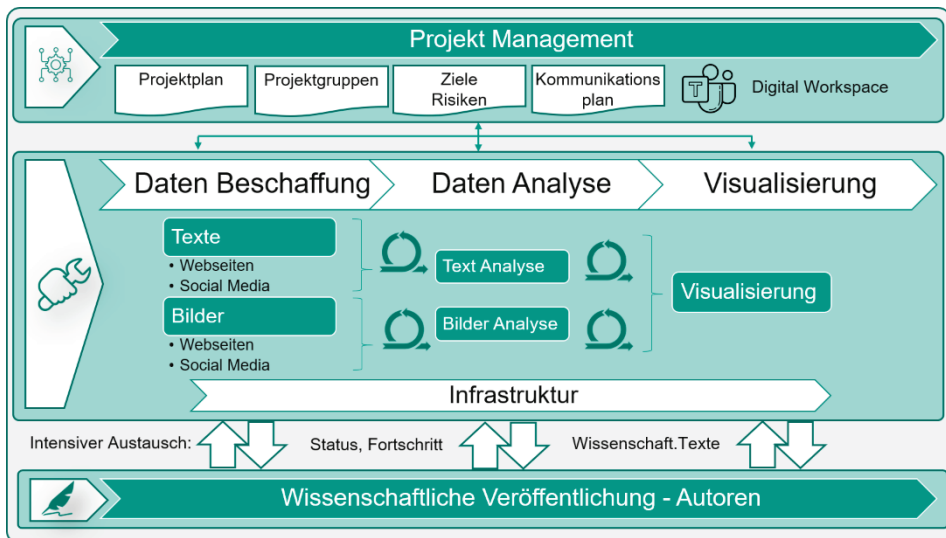
Insgesamt bildeten 33 Big Data & Analytics-Master-Studierende, verteilt auf zwölf Gruppen, sowie Kevin Hamacher und Rüdiger Buchkremer das Projektteam. Die Gruppen, wie auch eine erste Gruppeneinteilung der Studierenden, wurden bereits am Ende der ersten Vorlesung mithilfe eines Brainstormings mit den Ideen und Wünschen der Studierenden festgehalten. Das Projektmanagement bestehend aus Maike Blach und Jan Kozlik übernahm ab diesem Zeitpunkt die weitere Koordinierung und Durchführung des Projektes und der Vorlesungstermine. Zu den ersten Aufgaben gehörten, neben den Projektzieldefinitionen, eine erste Risikoanalyse, ein vorläufiger Projektplan sowie auch die Erstellung eines Kommunikationsplans mit dem dazugehörigen – auf „Microsoft Teams“ – basierten Digital Workspace. Dieser erwies sich mit den vielen praktischen Kollaborationsapplikationen, der hohen Verfügbarkeit und aufgrund der durch Covid-19-Pandemie bedingten Einschränkungen als besonders wichtiger Faktor für die erfolgreiche und terminliche Zielerreichung (siehe **Anhang 1**). Die Gruppeneinteilung stellte sich als besonders dynamisch heraus. In den ersten Wochen waren einige Studierende noch keinem Team zugewiesen. Andere arbeiteten in mehreren Gruppen und entschieden sich erst im weiteren Verlauf des Projektes für eine Arbeitsgruppe (siehe **Anhang 2**). Diese Dynamik sorgte an vielen Stellen für teamübergreifendes Verständnis und bewirkte eine effiziente Kommunikation zwischen den Gruppen.

Der gewählte und verfolgte Projektmanagementansatz setzte auf den agilen Prinzipien auf: Selbstorganisierte Gruppen mit motivierten Gruppenmitgliedern, die

späte Anforderungen akzeptieren, direkt kommunizieren und im wöchentlichen Zyklus ihre Ergebnisse vorstellen.<sup>37</sup>

Der Projektverlauf folgte dem Weg der Daten, von der Beschaffung über die Analyse hin zur Visualisierung der gewonnenen Informationen. Die Autorengruppe, bestehend aus vier Studierenden, stand über die ganze Projektlaufzeit in einem sehr engen Kontakt zu allen Untergruppen. Hier wurden der Status, Fortschritte und vor allem Erkenntnisse der Gruppen direkt weitergegeben. Abbildung 1 verdeutlicht dieses Vorgehen anhand der Gruppen sowie den Übergabepunkten zwischen den Gruppen:

**Abbildung 1:** Projektprozess und Struktur



## 2.5 Allokation der Arbeitsschwerpunkte

Mit den vorgegebenen Ressourcen, dem Projektzeitraum und der Teamstruktur mussten die Studierenden Schwerpunkte über die Auswahl der zu bearbeitenden Indikatoren setzen. Diese Schwerpunkte fielen der Studienrichtung nach auf Big-Data-Herausforderungen im Bereich der Text- und Bildanalyse. Damit die Forschungsfragen beantwortet werden können, einigten sich alle Teams auf eine gemeinsame Basis. Hierzu gehörten neben der abgestimmten Liste der zu untersuchenden Unternehmen, auch die grundlegende Datenstruktur, die ein späteres

<sup>37</sup> Vgl. Beck et al. 2001.

Zusammenführen der Analyseergebnisse ermöglichen sollte. Diese Datenstruktur besteht aus dem Unternehmen, der Branche, dem Kommunikationskanal (Web / Social Media) und der Webseite bzw. der Social-Media-Nachricht.

Die Bildung der zu untersuchenden Indikatoren wird in den folgenden beiden Kapiteln Textdaten und Bilddaten beschrieben. Die Zusammenführung der OSMI-Indikatoren zu den Parametern und dem Gesamt-OSMI wird im darauffolgenden Kapitel Visualisierung erläutert.

### **3 Textdaten**

#### **3.1 Vorgehensweise**

Aus organisatorischen Gründen und den aufwändigen Teilarbeitsschritten in diesem Bereich befassten sich insgesamt fünf Arbeitsgruppen des Analyseprojektes mit Textdaten. Innerhalb dieses Arbeitsbereichs wurde eine weitere Untergliederung in Datenbeschaffung und Datenanalyse vorgenommen. Im Bereich der Textdatenbeschaffung konzentrierte sich eine Arbeitsgruppe auf das World Wide Web im Allgemeinen, eine weitere auf soziale Medien im Speziellen. Die drei mit der Textdatenanalyse betrauten Arbeitsgruppen gliederten sich in eine Architekturarbeitsgruppe und zwei Analysearbeitsgruppen auf, um den Bereich gemeinsam zu bearbeiten. Obwohl im Laufe des Kapitels deutlich wird, dass die Indikatoren O2, A2 und G3 angesprochen werden, stellte die Ansprache der Indikatoren keine Leitlinie für die Arbeit der Gruppen dar. Bei der Analyse der Textdaten wurde vielmehr von einer indirekten Ansprache der Sinne durch das geschriebene Wort ausgegangen. Die Ergebnisse der verschiedenen Gruppen werden in den folgenden Unterkapiteln eingehend behandelt.

#### **3.2 Datenbeschaffung**

##### **3.2.1 Web**

Das Sammeln von Daten aus Webseiten unterschiedlichster Branchen bildet den Ausgangspunkt für die Datenanalyse. Die Speicherung dieser Daten, die Sortierung nach Merkmalen und deren Indexierung ist mithilfe von häufig vordefinierten Bibliotheken aus verschiedenen Programmiersprachen möglich.<sup>38</sup> Dieser Prozess wird auch als Web Scraping bezeichnet, dem das Web Crawling vorgelagert ist. Beim Web Crawling handelt es sich um einen automatisierten Prozess, der mithilfe von Computerprogrammen unter anderem Hypertext Markup Language (HTML)-Webseiten nach bestimmten vordefinierten Mustern durchsucht und Daten abspeichert. Die anschließende Extraktion dieser Inhalte erfolgt im Web-Scraping-Prozess. Eine der häufigsten Anwendungen von Webseiten-Crawlern ist unter anderem die Indexierung von Webseiten für Suchmaschinen. Hierbei ist zu erwähnen, dass nicht alle Daten von HTML-Webseiten durch den Web-Scraping-Prozess extrahiert werden können. Das Daten-Scraping bildet oft die Grundlage für die Erstellung von großen Datensätzen aus den Inhalten des Internets. Die dabei erlangten Daten können in einer Vielzahl von Formen erfasst werden,

---

<sup>38</sup> Vgl. Daugherty & Wilson 2018, S. 9–12.

so z. B. als Text-, Bild- oder Videodaten. Anwendung finden Webcrawler z. B. in der automatisierten Analyse spezieller Produktdaten und Preise durch Preisvergleichsportale. Diese versuchen so Produkte mit dem besten Preis-Leistungs-Verhältnis zu ermitteln. Auch können für Marketingmaßnahmen E-Mails oder postalische Adressen erfasst und gespeichert sowie anschließend für Werbezwecke oder Marketingkampagnen bereitgestellt werden. Dies dient z. B. der Kundengewinnung und letztlich der Umsatzsteigerung der Unternehmen.<sup>39</sup>

Im Zuge der Analyse von Webdaten hat sich eine Arbeitsgruppe des Big Data-Analyseprojekts mit dem Crawling der Webseiten von Unternehmen aus verschiedenen Branchen befasst. Dabei fanden die Python-Bibliotheken Scrapy und BeautifulSoup Verwendung, um einen automatisierten Crawling-Prozess abzubilden.

Ziel dabei war die Entwicklung eines allgemeingültigen Webseiten-Crawlers basierend auf der Programmiersprache Python. Dieser sollte in der Lage sein, Textdaten mit ihren zugehörigen HTML-Tags und Cascading-Style-Sheet-Klassen aus 166 zuvor festgelegten, englischsprachigen Webseiten unterschiedlicher Branchen zu erfassen und zu strukturieren. Die dabei gewonnenen Daten stellen eine Grundlage für den im Rahmen des Analyseprojektes zu entwickelnden OSMI dar. Die Speicherung der Daten sollte ebenfalls automatisiert auf einer MariaDB-Datenbank erfolgen. MariaDB ist ein Open-Source-Produkt, es handelt sich um ein voll kompatibles, relationales Datenbankverwaltungssystem.

Für die technische Umsetzung des Teilziels wurde ein virtueller Server des Typs VPS 2000 G9 des Unternehmens netcup gehostet. Dieser Server stellte als wesentliches Leistungsmerkmal 80 TB Datentransfer monatlich zur Verfügung sowie eine Festplattenspeicherkapazität von 320 GB auf einer SSD. Der Arbeitsspeicher RAM betrug 16 GB mit vier CPUs auf einem Sockel. Als Betriebssystem wurde Linux Debian 10 Buster verwendet. Zudem wurde die Programmiersprache Python und eine Datenbank auf dem oben beschriebenen Linux-Server installiert.

Der darauffolgend entwickelte Crawling-Prozess kennzeichnete sich durch das Zusammenspiel der bereits erwähnten Python-Bibliotheken. Der Einsatz von Scrapy ermöglichte das automatisierte Anstoßen der vordefinierten Webseiten und im Anschluss das Auslesen aller dort vorhandenen Textelemente. BeautifulSoup wurde innerhalb des Prozesses genutzt, um im Vorfeld der eigentlichen

---

<sup>39</sup> Vgl. Hejing et al. 2020, S. 64–68.



Aufspaltung der HTML-Strukturen der einzelnen Webseiten, was als Parsen verstanden wird, die relevanten Klassen und Tags auszulesen, welche die benötigten Textelemente enthalten. Zur Verarbeitung von JavaScript-basierten Webseiten wurde zusätzlich die Software Splash eingesetzt. Splash wurde zu Beginn des Prozesses für einige Sekunden ausgeführt, um alle JavaScript-Elemente zu laden. Danach erfolgte die Übergabe des vollständigen HTML-Codes an BeautifulSoup zur weiteren Verarbeitung. Sobald das Crawl-Limit von 500 Unterseiten erreicht war, wurde der Prozess beendet.

Nach Abschluss des Crawlings der Webseiten erfolgte die Übergabe der ausgelesenen Textdaten an die Datenbank. Diese Übergabe geschah mittels eines weiteren Python-Skripts, der sogenannten Pipeline, das die Daten aufbereitete und in die vordefinierte Struktur der Datenbank schrieb.

Nach der Übertragung der Daten an die MariaDB-Datenbank wurden die vorhandenen Texte einer Überprüfung unterzogen und in verschiedene Kategorien eingeordnet, die auch als Ausschlusskriterien dienten. Dies kam einer Bereinigung der gewonnenen Daten nahe, wobei die Daten selbst nicht verändert wurden, sondern lediglich ein Löschkennzeichen in der jeweiligen Datenbankzeile ergänzt wurde. Auf diese Weise konnte der Ursprungszustand der Daten zum Zeitpunkt des Crawling-Vorganges beibehalten werden. Dieser Umstand ermöglichte eine spätere Filterung auf Grundlage der Kennzeichen und die Möglichkeit der Weiterverarbeitung in der nachfolgenden Analyse.

Die Einteilung erfolgte in die folgenden fünf Kategorien:

- 1 = Text von einer falschen Domain.
- 2 = Text ist nicht in englischer Sprache.
- 3 = Text enthält JavaScript Object Notation (JSON)/HTML-Tags.
- 4 = Text wurde aufgrund fest definierter Regex-Befehle herausgefiltert (zur Erkennung von irrelevanten Wort- und Zeichenkombinationen).
- 0 = Text kann für die Analyse verwendet werden.

Der erzeugte Datensatz wurde unter Einsatz der Beekeeper-Studio-Software weiterverarbeitet, die eine Steuerung über Structured Query Language (SQL)-Befehle ermöglichte. Die erzeugte Tabellenstruktur umfasste die Spalten BRANCHE, TOP\_URL, SITE\_URL, DOWNLOADED\_TS, TEXT und CLASS. Die Spalte BRANCHE diente der Beschreibung des Wirtschaftsbereichs der extrahierten Webseite. Hierzu wurden vorab die Branchen Automobile, Cosmetics, Fashion, Food, Healthcare, Household, Interior, Leisure, Lifestyle, Technology

und Vacation definiert. Die TOP\_URL beschrieb die URL der Webseite. Durch die SITE\_URL wurden alle URLs beschrieben, die von der TOP\_URL aus erreicht wurden. Die Spalte DOWNLOAD\_TS enthielt das Crawling-Datum der Textdaten. In der Spalte TEXT wurden alle Texte der jeweiligen URL angezeigt, die dem Scraping-Prozess unterzogen wurde. Die Spalte CLASS beinhaltete alle Klassen oder HTML-Tags aus dem HTML-Code. Die Spalte ALLOWED\_DOMAIN wies die Domain aus, die dem Crawler übermittelte, welche Domain bearbeitet werden sollte.

Der Einsatz der beschriebenen Open-Source-Software ermöglichte in Kombination mit den Python-Bibliotheken Scrapy, BeautifulSoup etc. die Erstellung eines standardisierten Quellcodes, der Textdaten von rund 166 Webseiten extrahieren und diese in einer vordefinierten Struktur unter Einsatz von SQL in einer Maria DB-Datenbank erfassen konnte. Eine besondere Herausforderung bestand in der Extraktion der relevanten Textdaten der verschiedenen Webseiten. Dabei fiel die große Variation in der Bezeichnung der verwendeten Klassen und Tags zwischen den Webseiten auf. Diese bewirkte, dass keine fest definierten Bezeichnungen zum Auslesen genutzt werden konnten. In der Folge mussten Cascading-Style-Sheets-Klassen aus jeder einzelnen Webseite vor dem eigentlichen Crawling dynamisch ausgelesen werden, sodass nur diese für den Crawling-Vorgang verwendet wurden. So konnte sichergestellt werden, dass jedes Textelement berücksichtigt wird. Weiterhin muss erwähnt werden, dass insgesamt 39 von 205 Webseiten nicht durch Crawling erfasst werden konnten, da z. B. bei der Webseite des Unternehmens Ford die Fehlermeldung 403 auftrat, was darauf hinwies, dass die Webseite kein Crawling zuließ.

Des Weiteren fiel während des Crawling-Vorgangs auf, dass einige Webseiten nicht konsequent der HTML5-Struktur folgen. Beispielsweise wurde der HTML-Tag Paragraph `<p>` nach der Eröffnung nicht mit einem schließenden HTML-Tag `</p>` versehen. Im Ergebnis wurden daher stellenweise ganze HTML-Blöcke durch das Scraping erfasst. Dieser Umstand machte es notwendig, das Crawling- und Scraping-Ergebnis mit einem eigens geschriebenen Quellcode zu bereinigen.

### **3.2.2 Social Media**

Neben dem zuvor beschriebenen, allgemein gehaltenen Vorgang des Web-Crawlings von Textdaten beschäftigte sich eine weitere Arbeitsgruppe in erster Linie mit Daten aus sozialen Medien. Soziale Medien gelten im weiteren Sinne als technologiezentriertes, aber nicht rein technologisches Ökosystem, in dem

Akteurinnen und Akteure über vielfältige Interaktionen und Austauschformen miteinander verbundenen sind.<sup>40</sup> Dieser Bereich gilt als besonders zukunftsreich und die Möglichkeit, gezielt auf diesen Plattformen zu werben, wird für Unternehmen aufgrund hoher Nutzungszahlen immer attraktiver. Das große Potenzial zur Unternehmensdarstellung und die Nähe zu den Konsumentinnen und Konsumenten gilt als weiterer Anreiz.<sup>41</sup>

Bei der Bearbeitung dieser Thematik stellte sich die Arbeitsgruppe zwei grundlegende Fragen:

1. Welche sozialen Medien eignen sich grundsätzlich zur Beschaffung von textuellen Informationen?
2. Mit welchen Methoden können diese Texte aus den sozialen Medien extrahiert werden?

Bei Auswahl der zu analysierenden Medien wurden fünf Ausschlusskriterien definiert. Sobald eines dieser Kriterien als erfüllt galt, wurde die Analyse der betrachteten Plattform verworfen. Außerdem musste berücksichtigt werden, dass nicht alle Plattformen öffentlich zugänglich sind. Die Kriterien lauteten wie folgt:

1. Keine textuellen Informationen (sondern nur Bild und/oder Video).
2. Messenger-Dienst (1:1- oder Gruppen-Konversationen).
3. Plattformen, die das Extrahieren der Daten nicht erlauben.
4. Plattformen, deren Popularität in der Vergangenheit stark abgenommen haben.
5. Plattformen, auf denen sich Unternehmen nicht präsentieren (können).

Dieser eng gefasste Anforderungsrahmen hatte zur Folge, dass lediglich ein soziales Medium in Betracht gezogen werden konnte. Nur der Kurznachrichtendienst Twitter erfüllte die aufgestellten Anforderungen.

Bei der Beschaffung der Twitter-Daten wurden zwei Ansätze gewählt. Zum einen das bereits erwähnte Web-Scraping und zum anderen die Verwendung des Twitter-API (Application Programming Interface). Dieser kombinierte Ansatz sollte die vorhandenen Vor- und Nachteile der einzelnen Lösungen ausbalancieren. Dazu zählten in erster Linie die grundsätzliche Flexibilität der Lösung, der Umfang, in

---

<sup>40</sup> Vgl. Appel et al. 2020, S. 80.

<sup>41</sup> Vgl. Köster-Düpreé 2020, S. 176.

dem Daten abgerufen werden konnten, möglicherweise anfallende Kosten gegenüber der Fehleranfälligkeit und der Komplexität des Aufbaus.

Im Datenbeschaffungsprozess sollten die Twitter-Kanäle von Unternehmen verschiedener Branchen systematisch untersucht werden. Diese Untersuchung musste dabei auf die, sofern vorhanden, aktiven Hauptkanäle der Unternehmen begrenzt werden, deren Twitter-Nachrichten in englischer Sprache gepostet werden.

Zur Verwendung des twittereigenen APIs der Version 1.1 wurde die Python-Bibliothek Tweepy verwendet. Nach der Erstellung eines Twitter-Developer-Accounts können die im Anschluss zugewiesenen Consumer Keys und Authentifizierungstokens über Tweepy eingebunden werden, was letztlich den automatisierten Abruf der Tweets zuvor definierter Unternehmen ermöglicht. Die abgerufenen Twitter-Nachrichten enthielten im Wesentlichen die folgenden Daten:

- eindeutiger Name des Twitteraccounts,
- eindeutige Tweet-ID,
- Erstellungsdatum des Tweets,
- Textinhalt des Tweets.

Die über diesen iterativen Prozess gewonnenen Daten wurden tageweise für einzelne Unternehmen in einer Character-/Comma-separated values (CSV)-Datei gesichert. Mit der beschriebenen Methode wurden während einer Laufzeit von 14 Stunden über 500.000 Tweets gesammelt.

Der verfolgte Crawling-Ansatz bedurfte einer Eigenentwicklung der Arbeitsgruppe. So wurde mit einer Kombination des Selenium-Frameworks zur Simulation menschlicher Eingaben zum Nachladen von Twitter-Inhalten und des bereits erwähnten BeautifulSoup-Crawlers gearbeitet. Die entwickelte Methode bot eine, von den offiziellen APIs unabhängige Möglichkeit, Daten von Twitter zu extrahieren. Hierbei war keine Registrierung oder Bewerbung für Entwicklerzugänge notwendig. Es bestand keine Begrenzung in der Gesamtanzahl der Tweets, die in einem bestimmten Zeitintervall abgerufen werden konnten. Die Limitation der maximal 800 Tweets pro Konto schränkte diese Form der Extraktion jedoch ein. Dies wurde vor allem bei Twitter-Kanälen ersichtlich, die Tweets in großer Zahl publizieren.

Nach Abschluss der Rohdatenerfassung wurden die Ergebnisse beider Herangehensweisen in einer Tabelle zusammengeführt und auf Dubletten kontrolliert. In der Folge wurde daraus ein einheitlicher Datensatz generiert. Dieser neue Datensatz wurde nach eingehender Bereinigung unter Hinzuziehung der Analyse-

Arbeitsgruppen mit den Daten der Web-Arbeitsgruppe in der von dieser Gruppe erstellten MariaDB-Datenbank zusammengeführt, um die anschließende Analyse zu vereinfachen.

Bezüglich der anfänglichen Fragestellungen der Social-Media-Arbeitsgruppe muss festgestellt werden, dass aufgrund der eng gefassten Rahmenbedingungen keine großflächige Analyse verschiedener sozialer Medien stattfinden konnte. Dies scheint jedoch durchaus im Interesse der Betreiber der Plattformen zu liegen. Insbesondere die kostenneutrale Bereitstellung von Daten ist nicht als üblich zu erachten. Auch die zur Analyse verbliebene Plattform Twitter reglementiert nicht registrierte Zugriffe mittels Web-Scraping – Crawling zunehmend. Das grundsätzlich sehr nutzerfreundliche und kostenlose nutzbare API beschränkt wiederum den Umfang der Nutzung.

Die zweite, für den Erfolg des Projektes insgesamt relevantere Zielsetzung, einen strukturierten, aufbereiteten textuellen Datensatz zur Verfügung zu haben, konnte erreicht werden. Die parallele Verwendung zweier Ansätze sollte dabei für eine Ausgewogenheit der Stärken und Schwächen des jeweils anderen Ansatzes sorgen. In Zusammenarbeit mit der Web-Arbeitsgruppe wurde die Grundlage für die nachfolgende Arbeit der Analysegruppen geschaffen.

### **3.3 Datenanalyse**

#### **3.3.1 Architektur**

Die Größe der Arbeitsgruppe und thematische Tiefe der Textanalyse erforderte eine separat aufgebaute Analysearchitektur, von der aus die zuvor beschafften Daten untersucht und verarbeitet werden konnten. Bedeutende Vorzüge dieses Vorgehens waren die verstärkte Standardisierung der Arbeitsweise der Analysegruppen und die durch diese Arbeitsteilung erreichte größere Fokussierung der einzelnen Analysearbeitsgruppen. Weiterhin wurde durch die Entkopplung der Architektur von der Analysearbeit eine größere Anpassungsfähigkeit erreicht, die das nachträgliche Hinzunehmen und Entfernen von Plug-ins vereinfachte.<sup>42</sup> Die Anpassungsfähigkeit der Architektur einer Software gilt als Schlüsselkriterium.<sup>43</sup>

Der Aufbau der Architektur orientierte sich an der Vorgehensweise von Kim et al. aus der Veröffentlichung „Automated Text Analysis Based on Skip-Gram Model

---

<sup>42</sup> Vgl. Hayes-Roth et al. 1995, S. 288.

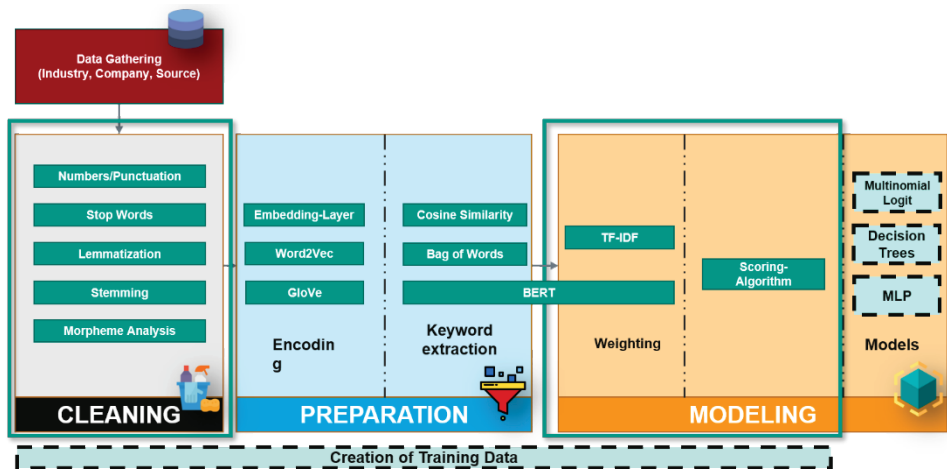
<sup>43</sup> Vgl. Ahmadiyah et al. 2020, S. 1.

for Food Evaluation in Predicting Consumer Acceptance“. Dort wurde ein auf Sinnesausprägung ausgerichtetes Modell zu der Bewertung von Nudelgerichten anhand von Online-Reviews entwickelt.<sup>44</sup>

Die Struktur der Textanalyse folgte in der vorliegenden Arbeit ebenfalls einem klassischen Preprocessing mit der Entfernung von wenig aussagekräftigen Stopwords und der Rückführung von Worten auf ihren Wortstamm. Weiterhin wurde eine Morphemanalyse durchgeführt, mit der Wörtern ihre grammatikalisch kleinste Eigenschaft (das Morphem) zugewiesen und bestimmt wird, ob es sich um ein Nomen, Adjektiv oder Verb handelt. Adjektive wurden als beschreibende Worte für den weiteren Gang der Analyse als relevanteste Wortgruppe bestimmt. Diese Vorarbeit diente der anschließenden Vektorisierung der Worte mittels Word Embedding. Ein Hauptunterschied bestand in der Modifizierung des Modells von dem zuvor dominierenden Geschmackssinn hin zu einer alle fünf Sinne berücksichtigenden Methode im Sinne des OSMI. Weiterhin sollten nicht nur Online-Reviews verarbeitet werden.<sup>45</sup>

Die folgende Abbildung 2 stellt den adaptierten Prozess grafisch dar:

**Abbildung 2:** Prozessarchitektur



Die Architektur zur OSMI-Ermittlung sollte drei Hauptarbeitsschritte umfassen: die Datenbereinigung, die Datenvorbereitung und die Modellierung.

<sup>44</sup> Vgl. Kim et al. 2018, S.1.

<sup>45</sup> Vgl. Kim et al. 2018, S. 1–3.

Die Datenbereinigung bzw. das Data-Cleaning diente der Aufbereitung der zuvor durch die in Kapitel 3.2 beschriebenen Prozesse gewonnenen Daten. Wichtige Bereiche bildeten hier die Entfernung sogenannter Stopwords oder Füllwörter, die Sätze lediglich verlängern oder der grammatikalischen Korrektheit dienen, allerdings für den eigentlichen Inhalt einer Aussage und somit auch für die Analyse irrelevant sind. Die bereits erwähnte Rückführung auf Wortstämme wurde hauptsächlich mit der Lemmatization-Methode erreicht. Zu erwähnen ist in diesem Zusammenhang allerdings auch die Stemming-Methode, durch deren Einsatz beispielsweise für die folgende Analyse irrelevante Konjugationen von Verben auf den Wortursprung zurückgeführt werden. Zuletzt erfolgte die eingangs beschriebene Morphemanalyse.

In der Datenvorbereitung wurden verschiedene Algorithmen zur Vektorisierung der Textdaten verwendet. Der modulare Aufbau der Architektur erlaubte dabei die Verfolgung verschiedener Ansätze. Die Wahl fiel hier auf den Einsatz der Word2Vec- und GloVe-Algorithmen sowie der BERT-Technologie. Einen weiteren Teil der Datenvorbereitung bildete die Herstellung von Zusammenhängen zwischen bestimmten Worten und den menschlichen Sinnen. Dazu wurde der Bag of Words-Ansatz in Verbindung mit der Kosinusähnlichkeit eingesetzt. Dieser Arbeitsschritt wird im Kapitel 3.3.3 der Word-Embedding-Arbeitsgruppe näher beschrieben.

Den letzten Schritt der Architektur bildete die Vorgabe der Modellierung des OSMI. Hier kam der TF-IDF (Term-Frequency-Inverse Document Frequency)-Algorithmus zum Einsatz, der zur Erfassung der Häufigkeit bestimmter Schlüsselwörter in Dokumenten konzipiert wurde.<sup>46</sup> Neben der Häufigkeitsbestimmung in Zusammenhang mit den menschlichen Sinnen wurde mittels TF-IDF eine weitere Gewichtung vorgenommen. Diese Gewichtung ermöglichte eine Kategorisierung der Wörter beginnend bei den speziellsten, am wenigsten vorkommenden und damit auch aussagekräftigsten. Diesen Wörtern wurde eine hohe Gewichtung zugewiesen, während häufig vorkommende Wörter ein geringeres Gewicht erhielten.<sup>47</sup>

Durch die Werte der TF-IDF-Methode konnte mittels des Scoring-Algorithmus ein OSMI ermittelt werden. Dieser konnte durch die Beibehaltung der vorgegebenen Datenstruktur auf einzelne Webseiten, einzelne Unternehmen und hochaggregiert auch auf ganze Branchen zurückgeführt werden.

---

<sup>46</sup> Vgl. Ramos 2003, o. S.

<sup>47</sup> Vgl. Ao et al. 2020, S. 1365–1359.

Die Umwandlung des TF-IDF-Outputs sollte in einen OSMI zwischen null und drei münden. Zuerst mussten dafür alle Werte auf den Wertebereich null bis eins normalisiert werden. Um eine zu große Konzentration des OSMI um null zu verhindern, wurden die TF-IDF-Werte mittels eines Min-Max-Scalers entfernt. Nach der Wertberechnung für jeden Sinn konnte ein Skalenniveau gewählt werden. Dazu wurde zuerst über die Liste für jeden Sinn pro Branche das arithmetische Mittel gebildet. Der mittlere Wert der Worte eines Sinns entsprach dem OSMI des Sinns. Die ermittelten Werte lagen jedoch weiterhin zwischen null und eins, weshalb diese mit drei multipliziert wurden und der finale OSMI-Score berechnet werden konnte. Anschließend wurde das arithmetische Mittel über alle Sinne einer Branche bestimmt, um den OSMI pro Branche zu erhalten.

Diese Berechnung der textbasierten Indikatoren der OSMI-Parameter erfolgte mit dem dargestellten Algorithmus:

**Abbildung 3:** Berechnung der textbasierten OSMI-Parameter

```

1 file := 'file with tf idf values per website'
2 Function calculateOsmi(file)
3   osmis_col := []
4   For website in file
5     osmis := {}
6     For sense in website
7       If NOT Len(sense)
8         # no word for sense present
9         normalized_osmi := 0
10      Else If Len(sense) == 1
11        # one word for sense present
12        normalized_osmi := sigmoid(sense)
13      Else
14        # more words for sense present
15        normalized_osmi := MinMaxScaler(sense)
16        osmis[sense] := Mean(normalized_osmi) * 3
17      Endfor
18      osmis['total_osmi'] := Mean([osmis[mean_sense] For mean_sense in osmis])
19      osmis_col.append(osmis)
20    Endfor
21    file['Osmi_Scores'] := osmis_col
22 Endfunction

```

Der entstandene Datensatz besaß damit erneut einen Eintrag pro Unternehmen, mit einem OSMI pro Sinn und einem Gesamtwert des OSMI.

Zur Verfolgung des BERT-Ansatzes, der sich im Wesentlichen durch die Verarbeitung ganzer Sätze anstatt einzelner Wörter von den anderen Methoden unterscheidet, musste eine modifizierte OSMI-Berechnung erfolgen.

Die Ausgaben des BERT-Modells waren Kosinusähnlichkeiten für jeden Sinn und für jeden gefundenen Satz auf einer Webseite. Da diese bereits einen Wertebereich von null bis eins aufwiesen, konnte wie im ersten Verfahren vorgegangen



werden, ohne dass eine Normalisierung der Daten erfolgen musste. Notwendig war lediglich die Multiplikation jedes Wertes mit drei, gefolgt vom Bilden der arithmetischen Mittel der Sinne. Eine Abweichung vom TF-IDF-Ansatz stellte die notwendige Konsolidierung der Werte aller Sätze einer Webseite dar. Vor der Berechnung wurden alle Tupel eines Datensatzes nach den Attributen „Company“, „Source“ und „Branche“ gruppiert. Das entstandene Datenset besaß damit erneut einen Eintrag pro Unternehmen, mit einem OSMI pro Sinn und einem OSMI-Gesamtwert.

### **3.3.2 Preprocessing**

Die Arbeitsgruppen der Textanalyse widmeten sich der Frage, wie der Textinhalt von E-Commerce-Webseiten und Twitter-Konten im Sinne des sensorischen Marketings durch die Anwendung verschiedener Natural Language Processing (NLP)-Techniken bewertet werden kann.

Das Hauptziel lag in der Entwicklung eines Frameworks, mit dem sensorische Texte automatisch ausgewertet werden können, einschließlich der Schritte der Datenbereinigung und -transformation bis hin zu ihrer Bewertung mithilfe der in Kapitel 3.3.1 beschriebenen TF-IDF-Methode. Weiterhin waren Schlüsselbereiche, wie jener der Bedeutung verschiedener sensorischer Wörter in verschiedenen Branchen und der Auswirkungen der Größe und Qualität eines Textes auf die TF-IDF-Ergebnisse, von besonderem Interesse.

Der Fokus der ersten Textanalysearbeitsgruppe lag daher im Besonderen auf der Vorverarbeitung der Textdaten und der Analyse unter Einsatz der TF-IDF-Methode, deren Durchführung im Kapitel 3.3.4 beschrieben wird.

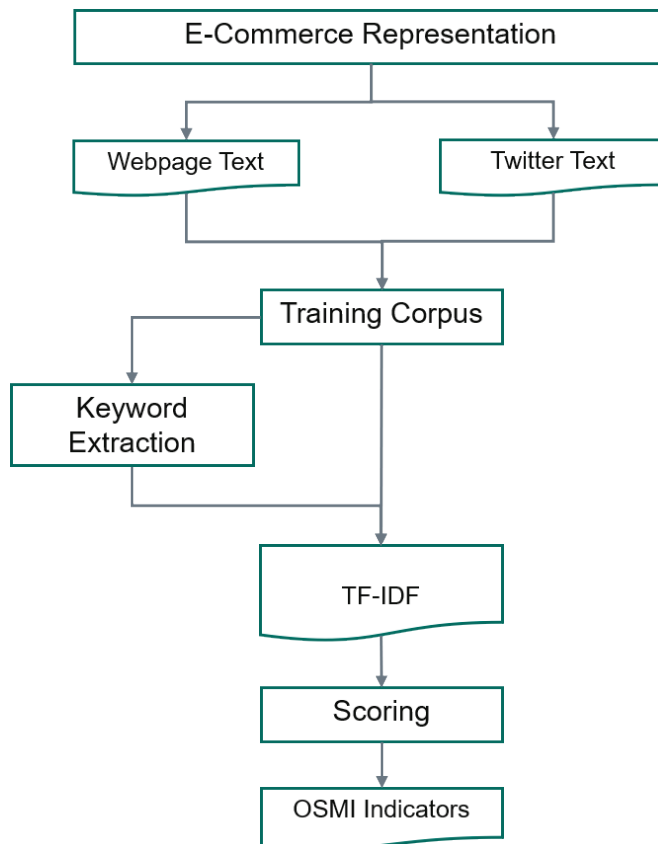
Die Notwendigkeit der Datenbereinigung bzw. des Preprocessings wurde, trotz bereits erfolgter Bereinigung in der Textbeschaffung, weiterhin als gegeben betrachtet, da das Rohdatenformat für die Datenanalyse nicht ausreichte. Es wurde befürchtet, dass insbesondere irrelevante Daten die Analyseergebnisse verfälschen könnten. Zur Einschätzung des Einsatzes sensorischer Beschreibungen auf Unternehmenswebseiten wurden demnach vorerst die vordefinierten Prozessschritte der Analysearbeitsgruppe angewandt.

Der Analyseprozess sollte daher von der Vorverarbeitung der bereitgestellten Textdaten aus verschiedenen Quellen über die Extraktion von Schlüsselbegriffen bis hin zu Bewertung reichen. Das Trainieren verschiedener Modelle mit den vorverarbeiteten Daten sollte in der Aufstellung einer Liste der einhundert Worte mit

der jeweils größten visuellen, akustischen, haptischen, gustatorischen und olfaktorischen Repräsentation münden. Die TF-IDF-Werte wurden anhand der Häufigkeit im Trainingskorpus gebildet und anschließend zum OSMI umgewandelt.

Der Prozess wurde wie in der folgenden Abbildung aufgebaut:

**Abbildung 4:** Aufbau der Textdatenanalyse



Die Arbeitsgruppe wendete für die Vorverarbeitung der von den Textbeschaffungsarbeitsgruppen beschafften Daten unter anderem die Python-Bibliotheken Twitter Preprocessor, Natural Language Toolkit und SpaCy an. Diese dienen beispielsweise der Entfernung von nicht verarbeitbaren Textteilen wie Hashtags oder Emojis sowie der Anwendung von Techniken wie Tokenisierung, Stemming, Lemmatization, Tagging oder Parsing. Zu erwähnen ist jedoch, dass gerade die Anwendung der Lemmatization auf den Textkorpus eine hervorgehobene Rolle bei der Schlüsselwortextraktion einnahm. Schließlich bedarf es für die weitere

Analyse real existierende Begriffe. Durch die Rückführung auf die Grundform eines jeden Wortes wird sichergestellt, dass die inhaltlich sensorische Bedeutung nicht verloren geht. Überdies wird durch Lemmatization vermieden, verschiedene Formen desselben Wortes als verschiedene Wörter zu behandeln. Das erleichtert die Ausgangsbasis für eine spätere Bewertung mittels weiterer Big-Data-Verfahren.

Ein weiterer Schritt der Vorverarbeitung war die Zusammenführung der Textdaten aus Quellen zu einem allgemeinen Schema. Der so entstandene Textkorpus wurde für jedes Unternehmen auf den Texten basierend erstellt, die zuvor aus Webseiten und Twitter-Nachrichten extrahiert werden konnten. Es wurden elf Datentabellen mit Webseitentexten erstellt, eine „für“ jede Branche: „WP\_Automobile“, „WP\_Cosmetics“, „WP\_Fashion“, „WP\_Food“, „WP\_Healthcare“, „WP\_Household“, „WP\_Interior“, „WP\_Leisure“, „WP\_Lifestyle“, „WP\_Technology“, „WP\_Vacation“.

Die von den Textbeschaffungsgruppen zur Verfügung gestellten Daten umfassten 2.873.067 Einträge von Unternehmenswebseiten. Hinzu kamen über 303.683 Twitter-Datensätze. So konnten insgesamt 3.176.750 individuelle Textpassagen für eine erste Verarbeitung berücksichtigt werden. Die Durchführung der beschriebenen Vorverarbeitungsschritte und das Zusammenführen der Datensätze hatte jedoch zur Folge, dass lediglich 425.818 Textpassagen zur weiteren Analyse verblieben.

### 3.3.3 Word Embedding

Die zweite Arbeitsgruppe der Textanalyse befasste sich im Rahmen der Datenvorbereitung mit der Technik des Word Embedding. Ziel war dabei die Extraktion von Schlüsselworten in Bezug auf die menschlichen Sinne. Dazu wurde eine Reihe verschiedener Algorithmen eingesetzt, die im Folgenden vorgestellt werden.

Unter Word Embedding ist die Einbettung diskreter Variablen in einen Vektorraum zu verstehen.<sup>48</sup> Dabei wird jedem Wert eine Position in einem multidimensionalen Raum zugewiesen, sodass dieser durch einen n-dimensionalen Vektor repräsentiert wird.<sup>49</sup> Jede Dimension des Vektors beschreibt eine Worteigen-

---

<sup>48</sup> Vgl. Rao & McMahan 2019, S. 121.

<sup>49</sup> Vgl. Kuang & Davison 2020, S. 8265–8269.

schaft in Bezug auf Bedeutung, Semantik oder die Klasse eines Wortes. Die Vorzüge dieser Darstellungsform liegen zum einen in ihrer Effizienz in Bezug auf den Umfang eines Vokabulars, da die Anzahl der Dimensionen limitiert ist, und zum anderen in der Möglichkeit aus der Einbettung des Wortes Informationen ablesen zu können, die durch die Werte eines Wortvektors repräsentiert werden.<sup>50</sup>

Aus der Repräsentation von Worten in Vektoren lässt sich in der Folge die Beziehung zwischen zwei Wörtern berechnen. Eine Wortähnlichkeit lässt sich danach in zwei verschiedenen Maßen ermitteln. Zum einen ist dies die Distanz zwischen zwei Vektoren im Vektorraum. Wörter mit ähnlichen Bedeutungen liegen durch ihre Vektoren näher beieinander als Wörter, die gänzlich verschiedene Bedeutungen aufweisen. Daher bildet die Distanz zwischen zwei Punkten ein geeignetes Maß zur Erkennung von Beziehungen zwischen Wörtern in einem Vektorraum. Dieser Abstand wird auch als euklidische Distanz bezeichnet. Wichtig ist hier, dass die Beziehung zwischen den beiden Wörtern umso stärker wird, je kleiner der Wert für die Entfernung ausfällt. Die zweite Möglichkeit zur Bemessung der Beziehung von zwei Wortvektoren ist die Ermittlung des Kosinuswinkels. Dieses Maß wird auch als Kosinusähnlichkeit bezeichnet und stellt die Beziehung zwischen zwei Vektoren in Richtung und Länge dar.<sup>51</sup>

Auf dieser theoretischen Grundlage aufbauend, kamen unterschiedliche Techniken zum Einsatz, die im Folgenden kurz erläutert werden. Die vorverarbeiteten Textdaten wurden mit den Modellen BERT, GloVe, Keras und Word2Vec trainiert, um mit Wortvektoren die Beziehung zu einem sinnesverwandten Wort wie „Geruch“ oder „Geschmack“ zu extrahieren, die in diesem Zusammenhang besonders aussagekräftig sind. Die zuvor beschriebenen Methoden der euklidischen Distanz und Kosinusähnlichkeit wurden genutzt, um bedeutende Worte im Umfeld der menschlichen Sinne herauszustellen.

Nach dem Trainieren der Modelle konnte über den ermittelten Index der generierte Vektor mit den einzelnen Wörtern des Vokabulars im Textkorpus in Verbindung gesetzt werden. Somit sind die in den Modellen ermittelten Embedding-Vektoren für die Wörter der Vokabulare bekannt. Mit dem Einsatz der euklidischen Distanz lassen sich anschließend die nächstgelegenen Wörter der Suchbegriffe im Vektorraum bestimmen. Über die Kosinusähnlichkeit können die Wörter mit der größten Vektorähnlichkeit bestimmt werden.

---

<sup>50</sup> Vgl. Kedia & Rasu 2020, S. 4.

<sup>51</sup> Vgl. Trask 2020, S. 230–231.

Nach der Erarbeitung der verschiedenen Methoden und NLP-Modelle bot sich ein Vergleich der Ähnlichkeiten und Unterschiede der ermittelten Schlüsselworte an. Dabei konnten die Methoden Keras, Word2Vec sowie GloVe nicht direkt mit dem BERT-Ansatz verglichen werden, da dieser sich aus den bereits in Kapitel 3.3.1 genannten Gründen, vor allem der nicht analogen Herangehensweise, zu stark von den anderen unterscheidet.

Der Vergleich der Ergebnisse legte den Schluss nahe, dass die GloVe- und Word2Vec-Methoden deutlich passendere Worte mit Sinnesbezug aus dem Textkorpus extrahieren konnten als dies beim Keras-Modell der Fall war. Grundsätzlich konnte festgestellt werden, dass sich die Herangehensweise des Keras-Modells weniger zur Extraktion optimaler Wortvektoren aus einem Textkorpus eignete. In der Folge wurden die Ergebnisse der Keras-Methode nicht weiterverwendet. Die über GloVe und Word2Vec gewonnenen Ergebnisse ähnelten sich in weiten Teilen. Die GloVe-Methode wies im Bereich des Hörsinns besonders gute Werte auf, während der Einsatz von Word2Vec sich besonders im Bereich des Geschmackssinns hervortat. Bei der detaillierten Betrachtung einiger Sinne fiel auf, dass der verfolgte Ansatz der Kosinusähnlichkeit robustere, plausiblere bzw. passendere Ergebnisse lieferte, als dies bei der euklidischen Distanz der Fall war.

Die spezifischen Anforderungen an die Datenbeschaffenheit und -qualität durch BERT hatten zur Folge, dass dieser Ansatz separat bewertet werden musste. Grundsätzlich waren die Ergebnisse jedoch so zufriedenstellend, dass auch die mit dieser Technik ermittelten Schlüsselworte Eingang in die OSMI-Berechnung fanden. Den niedrigsten Wert wies das BERT-Modell für den Hörsinn aus, der höchste Wert konnte im Bereich des Geschmackssinns erzielt werden.

Als Ergebnis der Schlüsselwortextraktion, oder auch Key Word Extraction, wurden die wichtigsten sensorischen Worte durch Anwendung der Word2Vec-, GloVe- und BERT-Methoden sowie verschiedener Entfernungsmetriken wie die Kosinusähnlichkeit und euklidische Entfernung identifiziert. Basierend auf den einhundert relevantesten Wörtern pro Sinn wurden durch die erste Gruppe der Textanalyse die TF-IDF-Werte berechnet. Die Grundlage dazu bildeten insgesamt 126.051 Werte für 1.013 eindeutige Worte aus elf Wirtschaftsbranchen, die durch den Einsatz der verschiedenen Word-Embedding-Techniken ermittelt wurden.

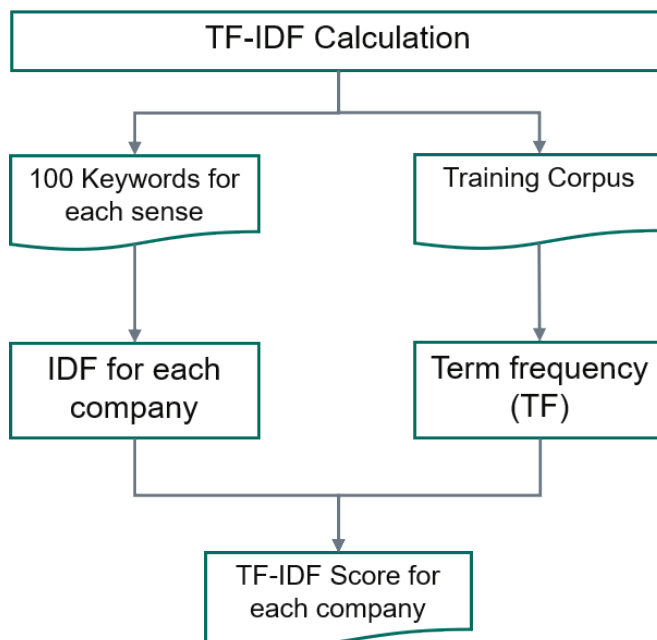
### 3.3.4 Modellierung

Die aufgestellten Listen von Schlüsselwörtern bildeten die Eingabewerte für die anschließende TF-IDF-Berechnung. Zur Identifikation der Schlüsselwörter in einer begrenzten Anzahl von Texten wurden Word-Embedding-Techniken auf die gesamten vorverarbeiteten Daten angewendet. Ziel war es, die nächsten oder ähnlichsten sensorischen Wörter im Vektorraum zu ermitteln. Worteinbettungstechniken ermöglichen die Darstellung von Wörtern in mehrdimensionalen Vektoren, was Maschinen bei der Interpretation von Wortbedeutungen unterstützt. Beziehungen zwischen Wörtern können durch die euklidische Distanz oder die Kosinusähnlichkeit bestimmt werden.

Bei der Analyse wurde der Zeitraum von 2019 an bis zum Ende des Jahres 2020 gewählt. Für die Analyse der Schlüsselworte je Sinn wurden die Techniken GloVe und Word2Vec angewandt. Die Ergebnisse der BERT-Technik wurden separat berechnet.

Der technische Ablauf wird in der folgenden Abbildung 5 noch einmal grafisch verdeutlicht:

**Abbildung 5:** TF-IDF-Berechnung



Der auf Grundlage der Vorarbeiten berechnete TF-IDF-Wert ließ sich wie folgt berechnen:

- t – term: Wort aus der Liste der Schlüsselworte
- d – document: Datensatz von Wörtern pro Unternehmen
- n – Anzahl der Dokumente, die Wort t enthalten
- corpus – alle Dokumente (Anzahl der Unternehmen in einer Branche)

**Formel 1:** TF-IDF-Score<sup>52</sup>

$$\begin{aligned} & \mathbf{TF - IDF\ Score} \\ & = \mathbf{Term\ Frequency\ (TF)} \\ & * \mathbf{Inverse\ Document\ Frequency\ (IDF)} \end{aligned}$$

**Formel 2:** Term Frequency<sup>53</sup>

$$TF(t, d) = \frac{\mathit{count}(t\ \mathit{in}\ d)}{\mathit{len}(d)}$$

**Formel 3:** Inverse Document Frequency<sup>54</sup>

$$IDF(\mathit{corpus}, n) = \log_e \left( \frac{\mathit{len}(\mathit{corpus})}{1 + n} \right)$$

Die höchsten TF-IDF-Werte wurden mit Word2Vec-Cosinus erzielt, gefolgt von GloVe-Cosinus.

Die Ergebnisse des GloVe-Cosinus-Einsatzes legten eine starke Verbindung von Firmen- und Markennamen zu den menschlichen Sinnen offen. Weiterhin waren für die Ergebnisse im Bereich der Automobilindustrie die Wörter „sharin“ und „inconvenient“ mit dem Tastsinn verbunden. In der Kosmetikindustrie war eine Verbindung von „axemobile“ mit dem Geruchssinn festzustellen. Bei Lebensmittelmarken wie „budweiser“, „whatevertarian“, „lipton“, „iced“ und „decaf“ wurde ein Zusammenhang mit dem Geschmackssinn erkannt. Im Bereich der Gesundheitsbranche stachen keine besonderen Wörter heraus. Die höchsten Punktzahlen wurden mit den Worten „swiffer“ und „febreze“ erzielt, die jeweils eine Verbindung

---

<sup>52</sup> Vgl. Ao et al. 2020, S. 1366.

<sup>53</sup> Vgl. Ao et al. 2020, S. 1366.

<sup>54</sup> Vgl. Ao et al. 2020, S. 1366.

zum Geruchssinn aufwiesen. Die Innen- und Freizeitbranche hatte keine Ausreißer. Bei den Ergebnissen der Lifestyle-Branche stach der einzelne Begriff „snackskin“ hervor und war mit dem Geschmackssinn verbunden. In der Technologiebranche gab es mit dem Wort „sony“ einen Bezug auf den Hörsinn.

Zum Vergleich wurden auch die Word2Vec-Cosinus-Ergebnisse überprüft. Für die Automobilindustrie waren keine Ausreißer vorhanden. Die Kosmetikindustrie wies erwartbare Ergebnisse mit Wörtern wie „deodorant“, „parfum“ und „moisturize“ auf. Außerdem erreichten einige Markennamen wie „chanel“ und „ritual“ hohe Punktzahlen. Für die Modebranche gab es nur eine „crease“, die sich auf den Geruchssinn bezog. Für die Lebensmittelbranche gab es mehrere gute Ergebnisse wie „sour“, „garnish“, „flour“, „onion“ usw., die allerdings keine Überraschung darstellten. Darüber hinaus standen sie alle in Beziehung zum Geschmackssinn. Die Ergebnisse für die Gesundheitsbranche waren auffälliger, da sie zum Teil keine echten Wörter darstellten, wie „btes“ und „evk“, sowie Wörter wie „earplug“ und „inearz“ lieferten. Alle Wörter der Gesundheitsbranche waren mit dem Hörsinn verbunden. Die Ergebnisse für den Haushalt wurden durch „febreze“ dominiert (abgesehen von einem Zeichen, das von Swiffer stammt) und wiesen eine Verbindung zum Geruchssinn auf. Beispiele sind „freshener“, „refresher“, „stink“ und „freshspiced“. Für die Einrichtungsbranche gab es nur ein Wort: „hansgrohe“, das sich auf den haptischen Sinn bezieht. Die Branchen Freizeit und Lifestyle / Schmuck hatten keine Ausreißer. Innerhalb der Technologiebranche gab es Ergebnisse für „Bang & Olufsen“, „Bose“ und „Sony“. 20 von 21 Wörtern bezogen sich auf den Hörsinn. Viele der Wörter waren Markennamen wie „Beoplay“, „Beovision“, „Bose“ oder „Sony“.

Eine zusammenfassende Betrachtung der Analyseergebnisse legte den Schluss nahe, dass in den einzelnen Branchen bestimmten Schlüsselbegriffen eine besondere Bedeutung zukommt. In Bezug auf den Ursprung der Textdaten konnte festgestellt werden, dass diese Schlüsselwörter insbesondere auf der Kurznachrichtenplattform Twitter größere Verwendung finden als es auf den E-Commerce-Webseiten der Unternehmen der Fall ist. Weiterhin konnte mit der TF-IDF-Berechnung ein Zusammenhang zwischen bestimmten Worten und einzelnen Sinnen hergestellt und deren Bedeutung gemessen werden. Eine tiefergehende Betrachtung der ermittelten OSMI-Werte erfolgt im Kapitel 5.



## **4 Bilddaten**

### **4.1 Vorgehensweise**

Mit der Beschaffung und Analyse von Bilddaten befassten sich insgesamt drei Arbeitsgruppen. Die Beschaffung wurde von zwei Arbeitsgruppen durchgeführt, die sich jeweils mit der Datenbeschaffung für die E-Commerce-Plattformen sowie Unternehmenswebseiten befassten. Die aggregierten Daten dieser beiden Gruppen wurden anschließend von einer Analysegruppe untersucht. Die Ergebnisse der Analysegruppe ermittelten zwei relevante OSMI-Indikatoren. Ferner konnten weitere Methoden als Ausblick der OSMI-bezogenen Analyse entwickelt werden.

### **4.2 Datenbeschaffung**

#### **4.2.1 Web**

Im Rahmen des Analyseprojekts sollten Bilddaten von insgesamt 235 Unternehmenswebseiten aus elf unterschiedlichen Branchen extrahiert werden.

Die Beschaffung dieser Bilddaten oblag einer weiteren Arbeitsgruppe und wurde ebenfalls mithilfe des in Kapitel 3.2.1 beschriebenen Web-Scraping-Prozesses und des in Python geschriebenen Frameworks Scrapy realisiert.

Das Speicherkonzept unterschied sich allerdings grundlegend von den Speicherkonzepten bei der Textdatenbeschaffung. Die deutlich größeren Datenmengen der Bilddateien erforderten einen skalierbaren Objektspeicher in Form des Cloud-Speichers Amazon S3. Für die Speicherung der zu den Bilddateien zugehörigen Metadaten entschied man sich, aufgrund der Schemafreiheit, für eine nicht-relationale Datenbank in Form des Datenbankmanagementsystems MongoDB.

Der Durchführung des Web-Scraping-Prozesses ging die Definition einiger Anforderungen voraus. So befanden sich auf den Webseiten, neben den für den OSMI relevanten Bildern, auch Grafiken, die keine inhaltliche Relevanz für die spätere Analyse hatten. Diese lagen in Form von Logos, Bannern, Buttons, Icons oder Gütesiegeln vor und mussten aussortiert werden. Ferner wurde das Dateiformat der Speicherung auf JPEG festgelegt. Die Attribute der Metadaten wurden, in Abstimmung mit der für die Analyse der Bilddaten zuständigen Arbeitsgruppe, wie folgt beschrieben:

- Laufende Nummer: eindeutiger Identifier
- Dateiname: Name der Bilddatei im Cloud-Speicher
- Unternehmen: untersuchtes Unternehmen
- Branche: Branche des Unternehmens
- URL: URL des angefragten Bildes
- Größe: Größe des Bildes in Pixel

Den Ausgangspunkt der Datenextraktion stellten die englischsprachigen Versionen der zu verarbeitenden Unternehmenswebseiten dar. Abhängig von der Verfügbarkeit wurde die US-amerikanische oder britische Version der Webseite zur Verarbeitung herangezogen. Stand keine dieser beiden Versionen zur Verfügung, wurde auf die deutschsprachige Version ausgewichen.

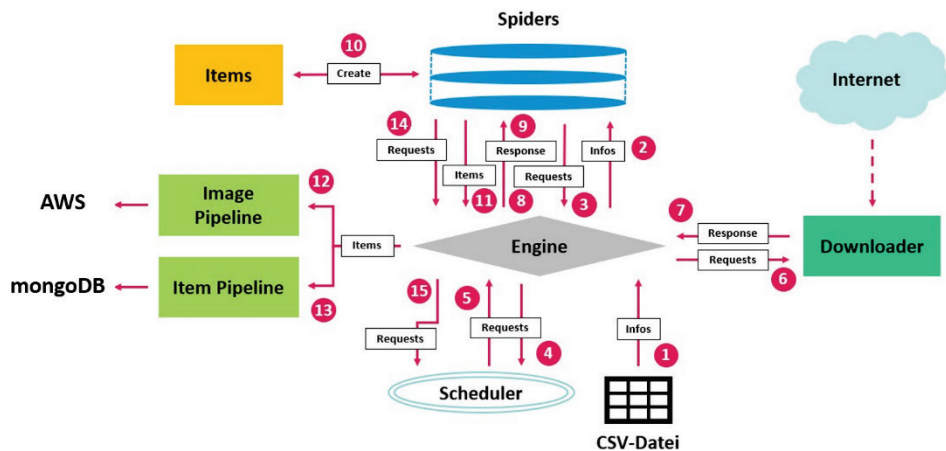
Um der Arbeitsgruppe für die Datenanalyse eine möglichst breite Datenbasis zur Verfügung zu stellen, wurde eine hohe Anzahl von 50 zu untersuchenden Webseiten pro Unternehmen angestrebt. Die Unternehmenswebseite durfte dabei nicht verlassen werden, beispielsweise durch das Folgen verlinkter Social-Media-Kanäle.

Die Zuordnung zwischen Bilddateien und Metadaten erfolgte über den global eindeutigen Dateinamen.

Nach der Definition der Anforderungen konnte die Extraktion der Bilddaten anhand des Web-Scraping-Prozesses durchgeführt werden.

Zu Beginn wurde die Basiskonfiguration des Frameworks um die Anbindung mit dem Cloud-Speicher ergänzt. Anschließend wurde ein erster Filter festgelegt, der eine Mindestgröße von 95 x 95 Pixel für ein Bild vorgab. So konnten Buttons, Icons und Gütesiegel im Vorfeld aussortiert werden. Des Weiteren wurde ein Timeout von 15 Sekunden für Webseitenanfragen konfiguriert. Das anschließende Vorgehen im Web-Scraping-Prozess wird anhand der schematisch dargestellten Architektur des Frameworks in Abbildung 6 dargestellt:

**Abbildung 6:** Web-Scraping-Prozess



Quelle: Eigene Darstellung in Anlehnung an <https://docs.scrapy.org/en/latest/topics/architecture.html>, Zugriff am: 09.01.2023.

Als Erstes werden Anfragen zu den Webseiten der Unternehmenswebseiten, sogenannte Requests, in einem Spider<sup>55</sup> erzeugt und einer Engine übergeben (3). Den initialen Request stellt hierbei die Anfrage zur Startseite des jeweiligen Unternehmens dar. Diese Requests werden, über die Engine, an den Scheduler weitergegeben (4), der die sequenzielle Abarbeitung aller Requests ermöglicht. Anschließend leitet der Scheduler die gesammelten Requests geordnet zurück an die Engine (5), die wiederum den Downloader aufruft (6). Der Downloader führt die Requests aus und lädt somit den gesamten HTML-Code der angefragten Webseite herunter. Diese Daten werden über die Engine (7) an den Spider zurückgegeben (8) und durch einen Parser innerhalb dieses Spiders verarbeitet (9). Der Parser stellt eine Eigenentwicklung abweichend vom Scrapy-Framework dar und wird im Folgenden genauer erläutert.

Die implementierte Parser-Logik versucht, die URLs von verlinkten Bildern und Grafiken innerhalb des HTML-Codes zu sammeln (10), um anhand dieser Adressen die Bilddateien und Metadaten herunterzuladen (12 & 13). Ferner werden weitere URLs auf der jeweiligen Webseite identifiziert, die als Requests dem Scheduler übergeben (14 & 15) und im Scraping-Prozess verarbeitet werden. Die Parser-Logik wird, unter Umsetzung der zuvor erläuterten Anforderungen, um die

<sup>55</sup> Spider stellen Python-Klassen dar, in denen das Durchlaufen der Webseiten (crawling) und die Extraktion deren Inhalte (scraping) beschrieben ist.

Webseite und festgelegte Sprachversion nicht zu verlassen, im Weiteren beschrieben. In einem ersten Analyseschritt werden korrespondierende HTML-Tags und -Klassen als Suchkriterien definiert. Anschließend können die URLs, mittels XML Path Language und Cascading-Style-Sheet-Selektoren, aus dem HTML-Code extrahiert werden. Erschwerend kommt hinzu, dass URLs zu Bildern responsiver Webseiten überwiegend in Sets gespeichert werden. Diese Sets stellen einzelne Strings dar, in denen mehrere URLs zum gleichen Bild in unterschiedlichen Auflösungen hinterlegt sind. Der jeweils letzte Eintrag dieses Strings beinhaltet das Bild in der höchsten verfügbaren Auflösung und wird extrahiert.

Nach der Extraktion der URLs folgt deren Verarbeitung. Queries<sup>56</sup> und Duplikate werden zu Beginn entfernt. Anschließend wird, gemäß Anforderung, ein weiterer Filter definiert, der Bilddateinamen anhand von Schlüsselwörtern aussortiert, die auf eine Verwendung des Bildes als Logo, Banner oder Button hindeuten. Des Weiteren werden der syntaktisch korrekte Aufbau und die Erreichbarkeit der URL sowie der Typ der verlinkten Datei geprüft. URLs, die nicht erreichbar sind oder nicht zu Bilddateien führen, werden vorerst aussortiert, indem der entsprechende Hypertext Transfer Protocol (HTTP)-Request gesendet und der Header der HTTP-Response auf den Content-Type sowie den Status-Code geprüft wird. Die Status-Codes 200, 301 und 302 sowie die Multipurpose Internet Mail Extensions (MIME)-Types JPEG und Portable Network Graphics (PNG) im Content-Type werden als gültig definiert. Ein ungültiger HTTP-Status-Code sorgt – mit Ausnahme des Codes 403, der sofort verworfen wird – für eine erneute Anfrage der URL mit einem User Agent<sup>57</sup>, die nach einem zweiten erfolglosen Versuch endgültig aussortiert wird.

Im Anschluss können die URLs erneut per HTTP-Request aufgerufen, die verlinkten Bilder heruntergeladen und im Amazon-S3-Bucket abgelegt werden. Sollte ein Bild heruntergeladen werden, das bereits im S3-Bucket gespeichert ist, wird es überschrieben. Die Metadaten zu diesem Duplikat werden dennoch gespeichert. Bilddateien, die im PNG-Format vorliegen, werden, gemäß Anforderung, in das JPEG-Format konvertiert. Die zuvor definierten Attribute der Metadaten werden innerhalb der Spider-Klasse für jedes Bild gespeichert und im JavaScript-Object-Notation-Format in der MongoDB-Datenbank abgelegt.

---

<sup>56</sup> Queries sind Teil einer URL und folgen nach einem Fragezeichen. Sie bestehen aus Parametern und geben beispielsweise Filter oder Sortierungen an, anhand derer die Darstellung der Webseite angepasst werden kann.

<sup>57</sup> Ein User Agent ist ein spezieller Header, der den Namen und die Version eines Browsers übermittelt.

Wie eingangs erwähnt, werden analog zur Verarbeitung der Bilddatei-URLs weitere URLs auf der jeweiligen Webseite identifiziert. Allerdings wird in diesem Zusammenhang der MIME-Type HTML als gültig definiert. Nach der Verarbeitung der URLs werden diese an den Scheduler übergeben und für den weiteren Scraping-Prozess vorgemerkt. Ein individuell implementierter Queuing-Prozess sorgt dafür, dass nicht mehr als 50 verschiedene Webseiten pro Unternehmenswebseite untersucht werden.

Als Ergebnis konnten 207 der geforderten 235 Unternehmenswebseiten erfolgreich mithilfe des Web-Scraping-Prozesses unter Einhaltung der definierten Anforderungen untersucht werden. Die ursprünglich 235 Unternehmenswebseiten wurden um Duplikate bereinigt und so auf 226 Sites reduziert. Eine Untersuchung der verbleibenden 19 Webseiten konnte nicht durchgeführt werden. Mögliche Ursachen können Captcha-Prüfungen, Altersfreigaben, ein dynamischer Aufbau der Webseite zur Laufzeit mittels JavaScript oder generelle Sicherheitsmaßnahmen seitens der Unternehmen gewesen sein.

Von den Webseiten der 9.138 angefragten URLs konnten insgesamt 43.055 Bilder extrahiert werden. Von diesen Bildern wurden 32.607 als eindeutig erkannt und im Cloud-Speicher abgelegt.

Der Arbeitsgruppe für die Datenanalyse kann eine breite Datenbasis zur Verfügung gestellt werden.

#### **4.2.2 Social Media**

Eine weitere Arbeitsgruppe befasste sich mit der Bilddatenbeschaffung aus sozialen Medien.

Die für die Analyse notwendigen Bilddaten sollten von den Unternehmensaccounts verschiedener Social-Media-Plattformen extrahiert werden. Grundlage waren auch hier die 235 Unternehmen unterschiedlicher Branchen.

Diese Arbeitsgruppe entschied sich, aufgrund der einfachen Handhabung des Speicherkonzeptes, für die webbasierte Kollaborationsplattformen SharePoint als Dateispeicher der Bilddateien. Die Metadaten dieser Bilddateien wurden als CSV-Datei ebenfalls im SharePoint abgelegt.

Eine erste Auswahl der Plattformen wurde mit dem Ziel vorgenommen, möglichst viele sozialen Medien unterschiedlicher Nutzungsarten in die Analyse miteinzubeziehen. Zu diesem Zweck wurde die Kategorisierung der sozialen Medien nach

Kreutzer herangezogen. Kreutzer unterteilt soziale Medien in drei Nutzungskategorien:<sup>58</sup>

- Kommunikation: Die Kommunikation der Nutzenden untereinander steht im Vordergrund.
- Content-Sharing: Das Teilen von Inhalten wie Texte, Bilder, Videos oder Audioinhalte steht im Vordergrund.
- Kooperation: Die Kooperation zwischen den Nutzenden steht im Vordergrund.

Die erste Kategorie Kommunikation umfasst Plattformen wie Blogs, Micro-Blogs (bspw. Twitter), private und professionelle soziale Netzwerke (bspw. Facebook, LinkedIn, XING oder Yammer), Messenger-Dienste (bspw. Facebook Messenger, Snapchat, WeChat oder WhatsApp) sowie Foren und Community. Das Content-Sharing betreiben Media-Sharing-Plattformen wie Instagram, Pinterest, Flickr, Vimeo oder YouTube. Die Kategorie Kooperation umfasst Bewertungs- und Auskunftsplattformen, Kreativ- und Auskunftsportale sowie Wikis wie Wikipedia oder Wikileaks.<sup>59</sup>

Die Nutzungskategorien Kommunikation und Content-Sharing sollten durch die Auswahl der Micro-Blog-Plattform Twitter, der sozialen Netzwerke Facebook, LinkedIn und XING sowie der Media-Sharing-Plattform Instagram abgebildet werden. Die generell textbasierten Plattformen der Kategorie Kooperation waren für die Bilddatenbeschaffung ungeeignet und wurden demnach nicht ausgewählt.

Anschließend werden die einzelnen Plattformen beschrieben und inhaltlich auf ihre Eignung zur Beschaffung von Bilddaten geprüft.

Die Plattform Instagram ist auf das Teilen von Bildern und Videos ausgelegt. Dies ist anhand der dargestellten Account-Seite der Unternehmensmarke BMW erkennbar (Abbildung 7).

Im oberen Bereich der Seite steht der Markenname gefolgt von einer kurzen Beschreibung. Anschließend sind ausschließlich Bilder oder Videos in tabellarischer Form hinterlegt. Beschreibungen zu den Medien können durch deren Anklicken aufgerufen werden. Diese fallen meist sehr kurz aus und spielen eine untergeordnete Rolle. Demnach ordnet Instagram der visuellen Präsenz in Form von Bil-

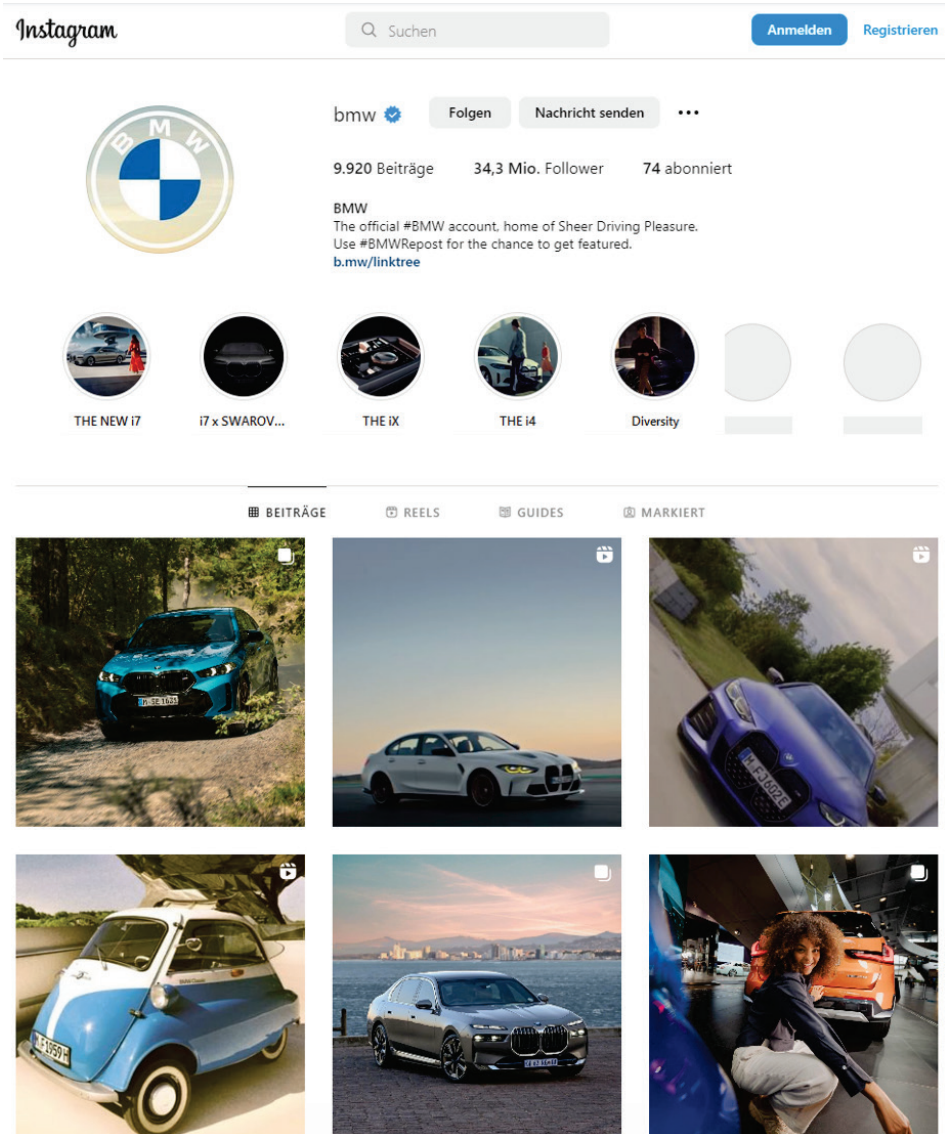
---

<sup>58</sup> Vgl. Kreutzer et al. 2020, S. 240.

<sup>59</sup> Vgl. Kreutzer et al. 2020, S. 240.

dem eine hohe Bedeutung zu und machte die Plattform für die Beschaffung von Bilddaten interessant.

**Abbildung 7:** Instagram-Account von BMW



Quelle: Instagram-Profil „bmw“, <https://www.instagram.com/bmw/>, Zugriff am 26.02.2023.

Eine weitere, für die Analyse relevante Plattform war Twitter. Die Micro-Blog-Plattform ermöglicht das Verfassen und Veröffentlichen von kurzen Nachrichten in begrenzter Länge. Diesen Nachrichten können Bilder angehängt werden, die im Rahmen der Datenbeschaffung extrahiert werden sollten. Der Fokus der Kommunikation liegt allerdings auf textueller Information.

Das soziale Netzwerk Facebook fokussiert sich bei der Kommunikation mit den Nutzenden ebenfalls auf Text. Diese Kommunikation findet über das Veröffentlichen von Nachrichten über die Account-Seiten der privaten und geschäftlichen Akteurinnen und Akteure statt. Den Nachrichten können Bilder optional angehängt werden, die extrahiert werden sollten.

Die beiden professionellen sozialen Netzwerke LinkedIn und XING bieten eine Plattform für die Kommunikation im beruflichen Kontext. Die Plattform XING wurde, nach näherer Betrachtung, aufgrund der begrenzten nationalen Ausrichtung von der Datenbeschaffung ausgeschlossen. Die Kommunikation auf LinkedIn findet, wie bei Twitter und Facebook, über das Veröffentlichen von Nachrichten statt. Auch hier liegt der Fokus auf dem verfassten Text.

Zu jeder der genannten Plattformen sollte die Extraktion der Bilddaten mittels API erfolgen. Da eine Prüfung der technischen Umsetzung einer solchen Extraktion bei der Auswahl der Plattformen nicht erfolgte, wurden einige Plattformen aufgrund von technischen Schwierigkeiten während der Extraktion von der Datenbeschaffung ausgeschlossen.

Vorerst wurden die Attribute der Metadaten festgelegt, die zu jedem Bild gespeichert werden sollten. Dies geschah erneut in Abstimmung mit der Arbeitsgruppe für die Bilddatenanalyse:

- Laufende Nummer: eindeutiger Identifier
- Dateiname: Name der Bilddatei im SharePoint
- Unternehmen: untersuchtes Unternehmen
- Branche: Branche des Unternehmens
- URL: URL des angefragten Bildes
- Größe: Größe des Bildes in Pixel

Anschließend konnte die technische Umsetzung der Extraktion mithilfe der den Plattformen zugehörigen APIs erfolgen.

Die Bilddaten der Plattform Instagram wurden mithilfe des Python-Packages Instaloader extrahiert, indem alle Nachrichten pro Account heruntergeladen wurden.



Dazu wurde vorerst ein Filter definiert, der solche Nachrichten aussortierte, die aus Videos oder Geotags bestehen. Des Weiteren wurden Kommentare und Metadaten zu den einzelnen Nachrichten ausgeschlossen. Anschließend wurde eine Liste mit Instagram-Accounts der jeweiligen Unternehmen erstellt und für jeden Eintrag dieser Liste der Account auf Existenz geprüft, um mit sämtlichen Nachrichten heruntergeladen werden zu können. Hierbei unterlag das Instagram-API der Einschränkung, dass nur eine bestimmte Anzahl an Anfragen in einer gewissen Zeit durchgeführt werden konnte. Dieser Einschränkung wurde insofern begegnet, als die Anfrage nach einer bestimmten Verzögerung erneut gestellt wurde. Zuletzt wurden die Bilddateien und Metadaten aus den Nachrichten extrahiert und im SharePoint abgelegt. So konnten von der Plattform Instagram insgesamt 334.803 Bilddatensätze extrahiert werden.

Im Anschluss sollte die Datenbeschaffung der Plattform Facebook erfolgen. Dies war allerdings aufgrund einer fehlenden Berechtigung nicht möglich.

Im Folgenden wird die Bilddatenbeschaffung von der Plattform Twitter erläutert. Die Extraktion der Bilddaten wurde mithilfe des Python-Packages `python-twitter` durchgeführt. Das API begrenzte hierbei die Anfrage der Nachrichten auf die letzten 200 eines Accounts. Diese bis zu 200 Nachrichten wurden pro Account angefragt und als Objekt zurückgegeben. Aus diesen Objekten wurden die URLs der Bilddateien extrahiert und aufgerufen. Das Ergebnis dieser Extraktion waren 16.652 erfolgreich von der Plattform Twitter beschaffte Bilddatensätze. Die geringe Anzahl an extrahierten Bildern lässt sich durch die Ausrichtung der Kommunikation der Plattform auf Textnachrichten und die Begrenzung der Anfrage auf 200 Nachrichten pro Account erklären.

Als Nächstes sollte die Bilddatenbeschaffung der Plattform LinkedIn erfolgen. Dies konnte allerdings, wie bei der Plattform Facebook zuvor, nicht durchgeführt werden, da die Betreiber von LinkedIn die Nutzung ihres API nicht zu Forschungszwecken gestatten. Einzig die Nutzung des APIs in Verbindung mit einer LinkedIn-App ist erlaubt.

Als Ergebnis konnten von den beiden Plattformen Instagram und Twitter 351.455 Bilddatensätze inklusive Metadaten von 220 verschiedenen Unternehmensaccounts heruntergeladen werden. Insgesamt wurden 157 Unternehmensaccounts auf Twitter und 158 Unternehmensaccounts auf Instagram untersucht. Die große Datenmenge der Bilddaten von Instagram führte zu Verarbeitungsproblemen bei der Bildanalyse. Die Arbeitsgruppe Datenbeschaffung Bilder grenzte daraufhin die Daten der Plattform Instagram auf den Zeitraum vom 1.9.2020 bis zum 30.9.2020 ein und reduzierte somit die Datenmenge von 334.803 auf 6.514 Da-

tensätze. Die verbleibenden 23.166 Bilddatensätze wurden der Arbeitsgruppe für die Datenanalyse übergeben.

## **4.3 Analyse**

### **4.3.1 Ermittlung der Indikatoren**

Die für die Bilddatenanalyse zuständige Arbeitsgruppe befasste sich mit der Ermittlung OSMI-relevanter Indikatoren aus den aggregierten Datensätzen der beiden Arbeitsgruppen für die Bilddatenbeschaffung.

Der Indikator G2 gibt Auskunft über die Verwendung von Farbschemata in Bildern der untersuchten Webseiten und Social-Media-Plattformen. Dieser kann durch die zwei Ausprägungen Schwach (Wert des Indikators liegt bei 0) und Standard (Wert des Indikators liegt bei 1) unterschieden werden. Eine schwache Ausprägung beschreibt die Abwesenheit von Farben bzw. Graustufen oder auch das Vorhandensein von vielen, bunt gemischten Farben. Eine Standardausprägung beschreibt hingegen die sinnlich kongruente Anwendung einer Farbe.

Die Arbeitsgruppe untersuchte die zur Verfügung stehenden Bilder vorerst auf eine Abwesenheit von Farben. Zu diesem Zweck wurden die beiden Bildtypen Graustufe und RGB (Red Green Blue) eingeführt. Die Bildinhalte bei Graustufenbildern werden gespeichert, indem jedem Pixel innerhalb eines zweidimensionalen Arrays ein numerischer Wert von 0 bis 255 zugewiesen wird. Dieser Wert gibt die Farbintensität des Pixels wieder, wobei der Wert 0 der Farbe Schwarz und der Wert 255 der Farbe Weiß entspricht. Die Indizes der Spalten und Zeilen des zweidimensionalen Arrays bilden dabei die Koordinaten des Pixels im Bild ab.

Bildinhalte von RGB-Bildern werden mithilfe dreidimensionaler Arrays gespeichert. Dabei werden drei zweidimensionale Arrays identischer Größe gespeichert. Jedes Array repräsentiert einen Farbkanal des Bildes. Die Farbintensitäten der einzelnen Primärfarben Rot, Grün und Blau werden über die entsprechenden Farbkanäle wiedergegeben.

Die Anzahl der Dimensionen des Bildes gibt demnach Aufschluss über eine schwache Ausprägung des Indikators G2. Ein Bild mit zwei Dimensionen entspricht dem Wert 0.

Ein Bild mit drei Dimensionen muss, zur korrekten Bestimmung von G2, einer genaueren Untersuchung unterzogen werden.

In diesem Zusammenhang sollte eine Überrepräsentation der Farben in Form von vielen, bunt gemischten Farben ermittelt werden. Diese Überrepräsentation

der Farben wurde anhand der Anzahl der dominanten Farben in den zur Verfügung gestellten Bildern festgestellt. Dazu wurde vorerst ein Verständnis für das additive Farbsystem des RGB-Farbmodells hergestellt.

Die Farbeindrücke<sup>60</sup> im additiven Farbsystem werden durch das Mischen der Primärfarben in unterschiedlichen Farbintensitäten generiert. Werden zwei Primärfarben mit gleicher Intensität vermischt, entstehen die Sekundärfarben Gelb, Cyan und Magenta. Die Intensität der Farben wird durch die Helligkeit des Lichtes bestimmt. Werden die Primärfarben mit jeweils einer Helligkeit von 0 (also kein Licht) vermischt, entsteht die Farbe Schwarz. Die Mischung aller drei Primärfarben in ihrer maximalen Intensität ergibt die Farbe Weiß.

Werden nun die drei Farbkanäle aus dem dreidimensionalen Array des Bildes vermischt, entsteht das eigentliche Farbbild.<sup>61</sup>

Die Farbdominanz kann folglich ermittelt werden, indem für jeden Pixel die entsprechende Farbe, wie beschrieben, festgestellt und eine passende Farbbezeichnung gespeichert wird. Die Aggregationen dieser Bezeichnungen werden ins Verhältnis gesetzt und diese Verhältnisse anschließend ausgewertet. Eine schwache Ausprägung des Indikators G2 mit dem diskreten Wert 0 wird bei einer Farbdominanz von über 10 Prozent bei mehr als drei Farben erreicht. Wird diese Farbdominanz bei drei oder weniger Farben erreicht, wird dem Indikator G2 eine Standardausprägung mit dem diskreten Wert eins zugewiesen.

Die Analyseergebnisse dieses Indikators waren ungenügend, insofern entschied sich die Arbeitsgruppe dafür, das Attribut aus dem Gesamtergebnis zu entfernen. Als Grund konnte hier eine unzureichende Datenqualität festgestellt werden, so dass das beschriebene Verfahren zur Ermittlung des Indikators G2 weiterhin gültig ist.

Der Indikator V4 gibt Auskunft über den Kontrast der untersuchten Bilder. Dieser kann durch die zwei Ausprägungen Schwach (Wert des Indikators liegt bei 0) und Standard (Wert des Indikators liegt bei 1) unterschieden werden. Eine schwache Ausprägung beschreibt einen Kontrast, der nicht auf das Produkt oder die Marke abgestimmt ist. Eine Standardausprägung hingegen beschreibt einen starken Kontrast im untersuchten Bild.

Zunächst wurde ein Verständnis für den Bildkontrast hergestellt. Mithilfe der Histogram Equalization kann der Kontrast eines Bildes verbessert werden. Hierbei

---

<sup>60</sup> Man spricht von Farbeindruck, da es sich hierbei um einen subjektiv empfundenen Sinneseindruck des Menschen handelt.

<sup>61</sup> Vgl. Trussel et al. 2005, S. 17–18.

wird die Verteilungsdichte der Farbintensitäten der Pixel des Bildes abgeflacht.<sup>62</sup> Diese Verteilung über die komplette Breite des Histogramms bewirkt einen besseren Kontrast. Daraus folgt, dass eine geringe Reichweite der Farbintensitäten der Pixel sowie steile Spitzen in der Verteilung einen geringen Kontrast bedeuten und vice versa. Hierbei ist die Platzierung der Verteilungsspitzen nicht relevant, da ein schwacher Kontrast im hellen sowie auch im dunklen Farbspektrum vorhanden sein kann.

Aufgrund dieses Verständnisses konnte eine Möglichkeit der Messung des Kontrastes mit einem Algorithmus erarbeitet werden, der die zu untersuchenden Bilder in Graustufenbilder umwandelte, die entsprechenden Histogramme bildete und die Farbintensitäten der Histogramme auf die oben genannten Eigenschaften untersuchte. Auf Basis dieser Untersuchung erfolgte die Bewertung des Indikators V4.

Die Indikatoren sowie weitere, während der Analyse erworbene Erkenntnisse über die Bilddatensätze werden den Attributen der Metadaten hinzugefügt.

Die bereits von den Arbeitsgruppen der Bilddatenbeschaffung festgelegten Attribute wurden um folgende Attribute ergänzt:

- height: Höhe des Arrays
- width: Breite des Arrays
- chanel: Anzahl der Farbkanäle
- indikator\_v4: Wert des Indikators V4
- dominante\_farben: Dominante Farben des Bildes

### **4.3.2 Objekterkennung**

Im Folgenden können mithilfe der Objekterkennung weitere Erkenntnisse über die Bilder erworben werden.

Die Objekterkennung ist eine Computer-Vision-Technik, welche die Identifizierung und Lokalisierung von Objekten in Bildern ermöglicht. Der gewählte Ansatz mittels Deep Learning erkennt erst die Objekte unterschiedlicher Klassen im Bild und lokalisiert diese anschließend. Die Arbeitsgruppe entschied sich bei der Bearbeitung dieser Problemstellung für ein Deep-Learning-Modell in Form eines

---

<sup>62</sup> Vgl. Kim 1997, S. 1.

Convolutional Neural Networks (CNN), dessen Relevanz als Problemlöser im Folgenden verdeutlicht wird.

CNNs stellen eine spezielle Form neuronaler Netze zur Verarbeitung von Daten mit einer bekannten rasterähnlichen Topologie dar. Als Beispiele seien hier Zeitreihendaten, die ein eindimensionales Raster mit Stichproben in regelmäßigen Zeitabständen darstellen, oder auch Bilddaten, die ein zweidimensionales Raster aus Pixeln darstellen, genannt.

Diese Netze besitzen zwei besondere Merkmale, die sie von anderen neuronalen Netzen unterscheiden. Zum einen die Verwendung der namensgebenden mathematischen Operation der Faltung und zum anderen das nachgelagerte Pooling.

Die Faltung (siehe Formel 4) beschreibt einen mathematischen Operator, der für zwei Funktionen  $f$  und  $g$  eine dritte Funktion  $f * g$  berechnet. Diese Faltung  $f * g$  kann auch als gewichteter Mittelwert der Funktion  $f$  mit der Gewichtung von  $g$  dargestellt werden. Dabei wird der Funktionswert  $f(x)$  mit dem Wert  $g(x - y)$  gewichtet, wodurch der Wert  $x$  aus  $f$  einen anderen gewichteten Mittelwert erhält.<sup>63</sup>

**Formel 4:** Faltung

$$(f * g)(x) = \int_{R^n} f(x - y)g(y)dy = \int_{R^n} f(y)g(x - y)dy = (g * f)(x)$$

Quelle: Yosida 1965: chaps. 3, Eq 1

Die Relevanz der Faltung für die Objekterkennung in Bildern lässt sich über den Anwendungsfall der Kanten- sowie Mustererkennung verdeutlichen. Hierbei handelt es sich um eine Faltung im zweidimensionalen Raum. Das Bild entspricht dabei der Funktion  $f$  und die Funktion  $g$ , hier als Kernel bezeichnet, beschreibt eine Matrix in Form eines Arrays.<sup>64</sup> Der Kernel der Größe  $2 \times 2$  durchläuft dabei elementweise den Eingabetensor von der Größe  $4 \times 3$  und führt die Matrixmultiplikation durch. Die Ausgabe wird dabei auf die Position beschränkt, in welcher der Kernel vollständig liegt. Folglich durchläuft der Kernel den Eingabetensor in sechs Bereichen, sodass nach der Faltung ein Ausgabebtensor der Größe  $3 \times 2$  entsteht.

In der Regel folgt auf die Faltung das Pooling. Die Pooling-Funktion ersetzt die Ausgabe des Netzes an einer bestimmten Stelle durch eine zusammengefasste

---

<sup>63</sup> Vgl. Yosida 1965, Kapitel 3.

<sup>64</sup> Vgl. Goodfellow et al. 2016: n. Abbildung 9.1.

statistische Größe der nächstgelegenen Ausgaben.<sup>65</sup> So liefern unterschiedliche Pooling-Varianten wie Maximum Pooling oder Average Pooling Ausgaben mit maximalen oder durchschnittlichen Werten der Umgebung. Die Ergebnisse des jeweiligen Kernels werden jeweils in der Erkennungsphase errechnet und die maximalen Werte in die Pooling-Phase übertragen.<sup>66</sup> Das Pooling dient schließlich zum einen dem Downsampling und bringt zum anderen die Eigenschaft der Invarianz mit sich. Downsampling bezeichnet ein Verfahren zur Reduktion von Einheiten. Bei der Bildverarbeitung werden so die Pixel der Ein- und Ausgaben reduziert. In den meisten Fällen wird hierbei ein  $2 \times 2$  Kernel mit  $\text{stride}^{67} = 2$  eingesetzt, was eine Reduktion der Ausgabe um den Faktor zwei zur Folge hat. Beispielsweise kann so ein Ausgabebtensor der Größe  $6 \times 6$  mit einer Gesamtanzahl von 36 Pixel auf eine Größe von  $3 \times 3$  mit 9 Pixel reduziert werden.

Ferner hilft das Pooling dabei, die zusammengefassten Ausgaben annähernd invariant gegenüber kleineren Verschiebungen der Eingabe zu machen. Dies ist für die Merkmalerkennung von Objekten in Bildern relevant, wenn die Lage des Merkmals keine Rolle spielt. Als Beispiel kann hier ein Bild mit einer Katze angeführt werden. Ist die Katze das zu erkennende Objekt, ist die genaue Position des Merkmals „Ohren“ nicht von Bedeutung, da diese sich auf dem Kopf der Katze befinden.

Das Zusammenspiel von Faltung und Pooling wird anhand eines der ersten CNNs, dem LeNet-5, verdeutlicht. Das LeNet-5 wurde 1989 von Yann LeCun et al. für die Bildklassifikation von handgeschriebenen Zahlen entworfen und gilt als Vorreiter in diesem Bereich. Das neuronale Netz besteht aus sieben Layern. Die beiden Convolutional Layer, gefolgt von zwei Pooling Layern, dienen der Objektidentifizierung. Anschließend folgt der Klassifikationsblock mit drei Fully-connected Layern. Für die Objekterkennung existieren nun zwei Lösungsansätze, die sich jeweils auf ein CNN als Deep-Learning-Modell stützen.

Im Region-Proposal-Ansatz wird die Objekterkennung in verschiedenen sich gegenseitig beeinflussenden Schritten durchgeführt. Diese sind die Ermittlung von Regionsvorschlägen, die Merkmalsextraktion mithilfe eines CNNs sowie die Identifikation und Lokalisierung von Objekten mittels Umrandungsrechtecken. Dieser mehrstufige Ansatz erfordert viel Rechenzeit und führt insbesondere in Echtzeitanwendungen zu Engpässen.<sup>68</sup> Infolgedessen wird ein einstufiger Ansatz eingeführt.

---

<sup>65</sup> Vgl. Goodfellow et. al 2016, S. 330.

<sup>66</sup> Vgl. Goodfellow et al. 2016: n. Abbildung 9.8.

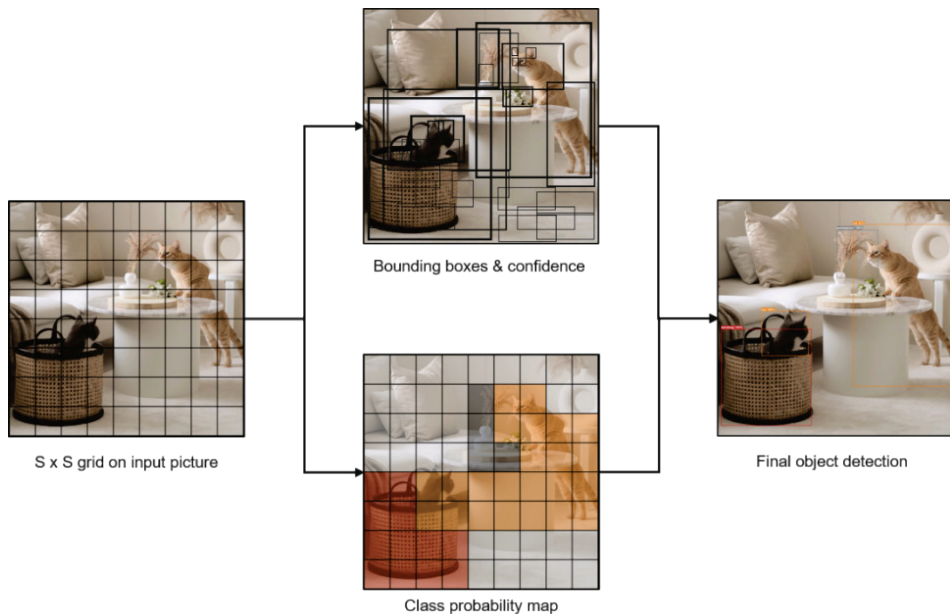
<sup>67</sup> Anzahl der Einheiten, um die der Kernel verschoben wird.

<sup>68</sup> Vgl. Zhao et al. 2019, S. 3219.

Im Regressions-/ Klassifikationsansatz werden Objektidentifizierung und -lokalisierung in einem Schritt durchgeführt. Ein bekanntes Modell dieses Ansatzes ist das You Only Look Once (YOLO)-Modell. Es wurde 2016 von Redmon et al. als Open-Source-Projekt entwickelt und ist in der Lage die Identifikation und Lokalisierung der auf einem Bild dargestellten Objekte in einer Regression durchzuführen. Dieser namensgebende einstufige Ansatz bringt entscheidende Vorteile wie Geschwindigkeit,<sup>69</sup> geringe Anzahl an Hintergrundfehlern<sup>70</sup> sowie Robustheit gegenüber unbekanntem Bildern<sup>71</sup> mit sich. Allerdings schwächt das Modell bei der Erkennung von kleinen oder dicht gedrängten Objekten.<sup>72</sup>

Das YOLO-Modell wird im Folgenden anhand der Abbildung 8 vorgestellt:

**Abbildung 8:** Funktionsweise des YOLO-Modells



Quelle: Eigene Darstellung in Anlehnung an Redmon et al. 2016, S. 780. Bild: Instagram-Profil „hmhome“, Posting vom 19.09.2020, <https://www.instagram.com/CFUEJTisQSE/?igshid=MDJmNzVkMjY%3D>, Zugriff am 09.01.2023.

<sup>69</sup> Vgl. Redmon et al. 2016, S. 779.

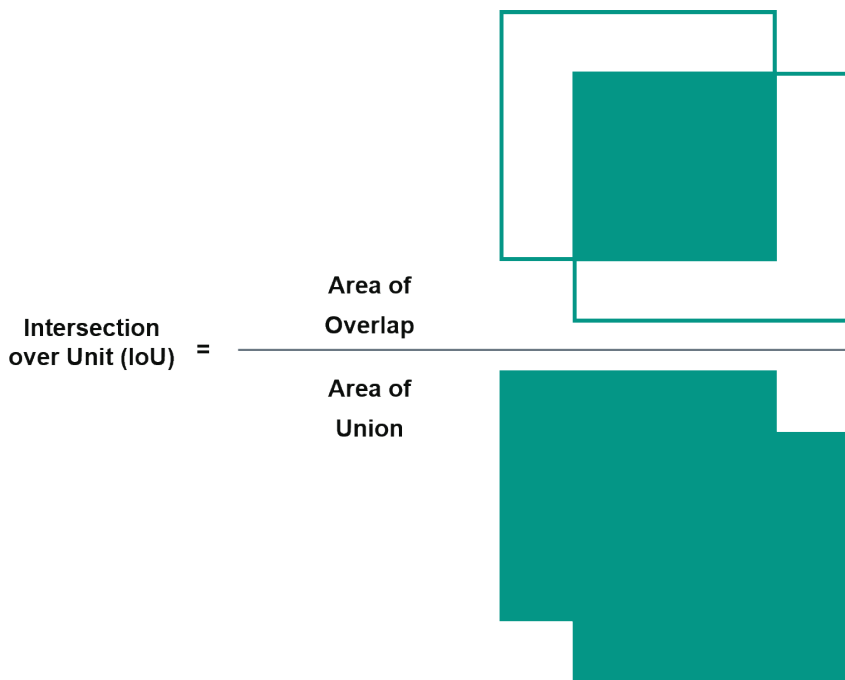
<sup>70</sup> Vgl. Redmon et al. 2016, S. 780.

<sup>71</sup> Vgl. Redmon et al. 2016, S. 780.

<sup>72</sup> Vgl. Redmon et al. 2016, S. 780.

Das zu analysierende Bild wird in ein  $S \times S$ -Raster aufgeteilt und die abgebildeten Objekte von der Zelle aus ermittelt, in welcher der Mittelpunkt des Objekts liegt. Anschließend werden für jede Zelle  $B$  Umrandungsrechtecke sowie die jeweiligen Konfidenzwerte ermittelt. Die Konfidenzwerte werden als Produkt aus der Sicherheit, dass ein Objekt in einer Zelle vorhanden ist, und der Intersection over Unit (IOU) berechnet (siehe Formel 4). Die IOU gibt das Verhältnis von der Schnittfläche des wahren und des berechneten Umrandungsrechtecks zu ihrer Einheit an und wird in nachfolgender Abbildung 9 schematisch dargestellt:

**Abbildung 9:** Schematische Darstellung Intersection over Unit



Quelle: Eigene Darstellung in Anlehnung an Rosebrock 2022.

Ferner wird für jede Zelle die Klassenzugehörigkeitswahrscheinlichkeit je Objektklasse  $\text{Pr}(\text{Class}|\text{Object})$  ermittelt. Während des Trainingslaufs wird diese mit dem jeweiligen Konfidenzwert multipliziert und ergibt den klassenspezifischen Konfidenzwert (siehe Formel 5). Die Wahrscheinlichkeit, mit der es sich bei einem Objekt um ein Objekt der jeweiligen Klasse handelt, sowie die Präzision, mit der das Umrandungsrechteck ermittelt werden konnte, werden durch diesen Wert beschrieben.



**Formel 5:** Konfidenzwert

$$\Pr(\text{Object}) * IOU_{pred}^{truth}$$

Quelle: Redmon et al. 2016: 780

**Formel 6:** Klassenzugehörigkeitswahrscheinlichkeit

$$\Pr(\text{Class}_i|\text{Object}) * \Pr(\text{Object}) * IOU_{pred}^{truth} = \Pr(\text{Class}_i) * IOU_{pred}^{truth} \quad (4)$$

Quelle: Redmon et al. 2016: 780

Das zugrunde liegende CNN des YOLO-Modells besteht aus 24 Convolution Layern für die Objektidentifizierung und zwei Fully-connected Layern, mit denen die Wahrscheinlichkeiten und Umrandungsrechtecke ermittelt werden.<sup>73</sup>

Die Arbeitsgruppe der Bildanalyse setzte für die Objekterkennung das YOLO-Modell der aktuellsten Entwicklungsstufe ein. YOLOv5 wurde im Juni 2020 von Jocher, ohne begleitendes Paper, veröffentlicht. Es existieren verschiedene Modellvarianten mit unterschiedlichen Leistungsmerkmalen. Eingesetzt wurde das Modell in der schnellsten, allerdings auch ungenaueren Variante YOLOv5s<sup>74</sup>. Aufgrund des Einsatzes lokaler Rechenleistung und der zu verarbeitenden Menge an Daten entschied sich die Arbeitsgruppe für diese YOLO-Variante. Der Programmcode wurde für die Analyse angepasst, sodass je erkanntem Objekt folgende Attribute in einer CSV-Datei abgelegt werden konnten:

- Dauer der Analyse,
- Index der zugewiesenen Klasse,
- Name der zugewiesenen Klasse,
- Wahrscheinlichkeit der Klassenzugehörigkeit,
- Koordinaten des Umrandungsrechtecks.

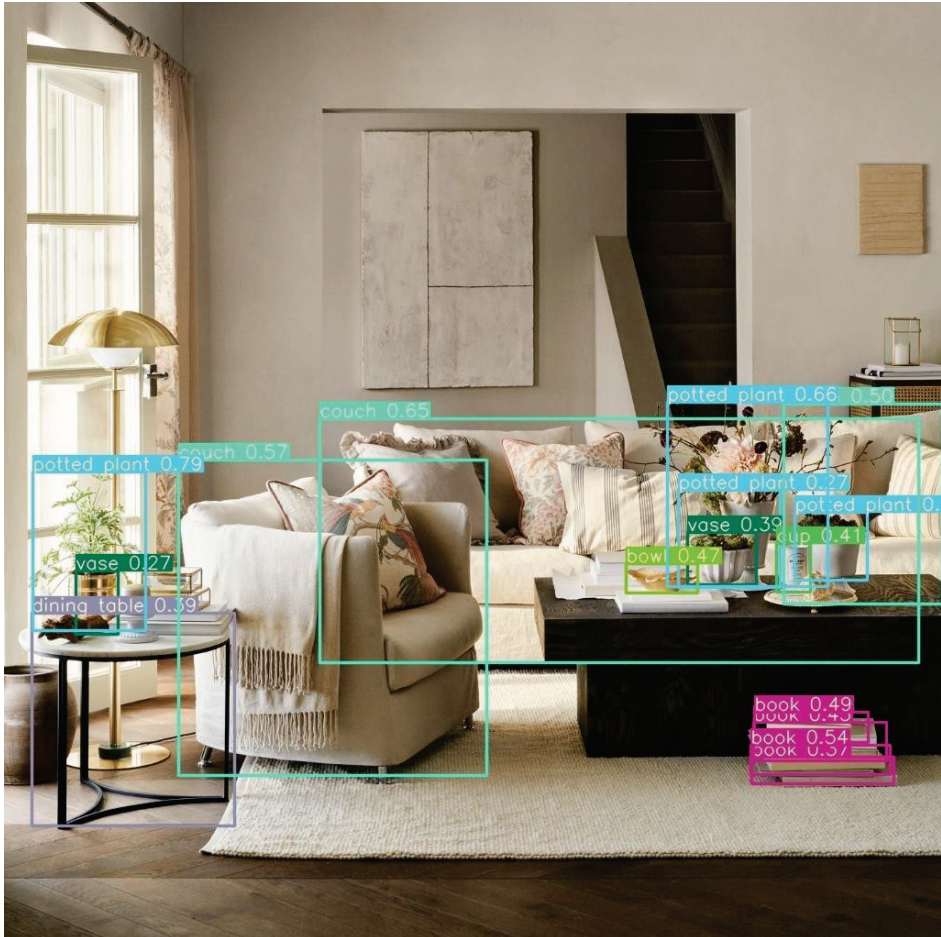
Eine visuell aufbereitete Ausgabe der Ergebnisse der Objekterkennung zeigt Abbildung 10:

---

<sup>73</sup> Vgl. Redmon et al. 2016, S. 781.

<sup>74</sup> Vgl. <https://github.com/ultralytics/yolov5>, Zugriff am 10.02.2021.

**Abbildung 10:** Beispiel einer Objekterkennung mit YOLOv5



Quelle: Instagram-Profil „hmhome“, Posting vom 01.09.2020, <https://www.instagram.com/p/CCshpFJspX7/?igshid=MDJmNzVkmjY%3D>, Zugriff am 09.01.2023.

Die Objekte werden den Metadaten der Bilder als Attribut hinzugefügt:

- yolo\_objects: Attribute aus der Objekterkennung mit dem YOLO-Modell

### 4.3.3 Weitere Bildanalysen

Es folgten weitere Bildanalysen, die das Erkennen von Objekten, Ereignissen, Motiven, Gesichtern sowie Texten auf Bildern zum Ziel hatten. Die Arbeitsgruppe setzte die Computer-Vision-Plattform Amazon Rekognition für diese Analysen ein. Amazon Rekognition ist ein kostenpflichtiger Cloud-Service der Amazon-Web-Services-Produktfamilie. Für die Analysen werden zwar Deep-Learning-Modelle verwendet, eine Nutzung der Plattform setzt allerdings keine Machine-Learning-Kenntnisse voraus.

Die Nutzung von Amazon Rekognition erfolgte mittels API über die Funktionen des Python-Software-Development-Kit Boto3. Das API wurde über die Instanz eines Boto3-Clients angesprochen. Diese Instanz ermöglichte den Aufruf verschiedener Bildanalysefunktionen, denen die Bilder übergeben wurden. Anschließend wurden die Analyseergebnisse als Rückgabewerte der Funktionen in einer CSV-Datei abgelegt.

Die eingangs genannten Analysen konnten durch den Aufruf von drei Funktionen durchgeführt werden. Diese Funktionen werden im Folgenden erläutert.

Amazon Rekognition bietet mit der Label-Analyse die Möglichkeit Objekte, wie Bäume oder Tische, Ereignisse wie Hochzeiten oder Geburtstagsfeiern, sowie Motive, wie Landschaften, in Bildern zu erkennen.<sup>75</sup> Diese Lösung kann, aufgrund des proprietären Cloud-Ansatzes, nicht die Geschwindigkeit und individuelle Anpassungsmöglichkeit bieten, die mit dem YOLO-Modell möglich sind. Allerdings können mit dem Erkennen von Ereignissen und Motiven weitere Erkenntnisse aus den Bildern erworben werden. Die zu untersuchenden Bilder werden dafür der Funktion `detect_labels` übergeben. Diese gibt anschließend als Rückgabewert ein Array mit dem entsprechenden Label sowie dem jeweiligen Konfidenzwert zurück. Ferner werden hierarchische Taxonomien der zugewiesenen Labels bis in die oberste Ebene zurückgegeben.<sup>76</sup> Dies erfolgt durch die Zuordnung entsprechender Parent-Labels und wird anhand der Analyse des Bildes in Abbildung 11 verdeutlicht.

---

<sup>75</sup> Vgl. Mishra 2019, S. 4.

<sup>76</sup> Vgl. Mishra 2019, S. 370.

**Abbildung 11:** Beispielbeispiel für eine Label-Analyse



Quelle: Instagram-Profil „landmann.germany“, Posting vom 16.07.2020, <https://www.instagram.com/p/CEmVhOVC4XQ/?igshid=MDJmNzVkJmY=>, Zugriff am 09.01.2023.

Tabelle 3 stellt das Analyseergebnis des Bildes dar. Mit einem Konfidenzwert von über 90 Prozent wurden die Objekte Person, Food und Meal sowie die Ereignisse Leisure Activities und Vacation korrekt erkannt. Die genannten Elemente sind Parent-Labels des Labels Picnic, das ebenfalls dem Bild korrekt zugeordnet und mit einer Wahrscheinlichkeit von rund 92 Prozent zurückgegeben wurde.

**Tabelle 3:** Ergebnis der Label-Analyse

Label	Konfidenzwert in %
Food	99,86
Meal	99,86
Human	99,60
Person	99,60
Person	99,52
Person	99,52
Vacation	99,37
Leisure Activities	94,21
Picnic	91,87
Tent	81,04
Shorts	63,36
Apparel	63,36
Clothing	63,36
Transportation	62,05
Vehicle	56,22
Plant	55,33
People	55,27

Eine weitere Möglichkeit der Bildanalyse mit Amazon Rekognition ist die Gesichtserkennung. Mit der Funktion `detect_faces` können Merkmale und Emotionen aus maximal hundert auf dem Bild abgebildeten Gesichtern sowie Koordinaten markanter Gesichtspunkte wie Mundwinkel oder Auge ermittelt werden.<sup>77</sup> Diese werden in einem Array mit dem entsprechenden Konfidenzwert zurückgegeben. Eine Übersicht der zur Verfügung stehenden Merkmale und Emotionen ist anhand des Ergebnisses der Analyse des Bildes in Abbildung 12 mit den Tabellen 4 und 5 gegeben.

---

<sup>77</sup> Vgl. Mishra 2019, S. 366–368.



**Abbildung 12:** Beispielbild für eine Gesichtserkennung



Quelle: Instagram-Profil „douglas\_cosmetics“, Posting 13.09.2020, <https://www.instagram.com/p/CFFLqz2Dkj0/?igshid=MDJmNzVkmjY=>; Zugriff am 10.01.2023.

**Tabelle 4:** Analyseergebnis der Gesichtserkennung

Merkmal	Wert	Konfidenzwert
Altersbereich	17–29	
Lächeln	Ja	99,90
Brille	Nein	99,58
Sonnenbrille	Nein	99,89
Geschlecht	Weiblich	99,79
Bart	Nein	99,79
Schnurbart	Nein	99,93
Augen geöffnet	Ja	96,17
Mund geöffnet	Ja	99,76

**Tabelle 5:** Analyseergebnis Gesichtserkennung – Emotionen

Emotion	Konfidenzwert
Glücklich	99,83
Überrascht	0,08
Verärgert	0,03
Traurig	0,01
Gelassen	0,01
Verwirrt	0,01
Entsetzt	0,01

Die Texterkennung von Amazon Rekognition ermöglicht das Erkennen von Texten in Bildern. Es können bis zu fünfzig Wörter in mehreren Zeilen mit einem Winkel der Schrift von +/- 90° zur horizontalen Achse ausgelesen werden.<sup>78</sup> Der Texterkennung werden mit der Funktion `detect_text` die zu untersuchenden Bilder übergeben. Anschließend werden diese analysiert und das Ergebnis als Array zurückgegeben. Dieses Array enthält den erkannten Text, den Konfidenzwert, die Koordinaten des Umrandungsrechtecks, das Umrandungspolygon aus vier Punkten für

<sup>78</sup> Vgl. Mishra 2019, S. 257–260.

die Umrandung von schräg ausgerichtetem Text sowie den Typ des Textes. Der Texttyp beschreibt dabei entweder eine Zeile oder ein Wort. Als Beispiel sind hier die Analyseergebnisse des Bildes in Abbildung 13 mit Tabelle 6 genannt.

**Abbildung 13:** Beispielbild für die Texterkennung



Quelle: Instagram-Profil „iglo\_deutschland“, Posting 17.09.2020, <https://www.instagram.com/p/CFPbXyEquXU/?igshid=YmMyMTA2M2Y=>, Zugriff am 10.01.2023.



**Tabelle 6:** Analyseergebnis der Texterkennung

Textzeile	Konfidenzwert
Mach mit bei der	99,94
VEGGIELUTION	99,90
Jetzt im	99,09
TIEFKUHLREGAL	99,88
NEU	99,50
iglo	99,79
GReeN	96,09
CuisiNe	95,10
100% Veggie Power	98,30

Die Arrays der Label-, Gesichts- sowie Texterkennung wurden den Metadaten der Bilder als Attribute hinzugefügt:

- aws\_label: Array mit den Analyseergebnissen der Label-Analyse
- aws\_faces: Array mit den Analyseergebnissen der Gesichtserkennung
- aws\_text: Array mit den Analyseergebnissen der Texterkennung

Der Arbeitsgruppe für die Visualisierung wurde abschließend eine CSV-Datei mit den folgenden Attributen übergeben:

- Laufende Nummer: eindeutiger Identifier
- Dateiname: Name der Bilddatei im SharePoint
- Unternehmen: untersuchtes Unternehmen
- Branche: Branche des Unternehmens
- URL: URL des angefragten Bildes
- Größe: Größe des Bildes in Pixel
- height: Höhe des Arrays
- width: Breite des Arrays
- chanel: Anzahl der Farbkanäle
- indikator\_v4: Wert des Indikators V4

- dominante\_farben: Dominante Farben des Bildes
- yolo\_objects: Attribute aus der Objekterkennung mit dem YOLO-Modell
- aws\_label: Array mit den Analyseergebnissen der Label-Analyse
- aws\_faces: Array mit den Analyseergebnissen der Gesichtserkennung
- aws\_text: Array mit den Analyseergebnissen der Texterkennung t

## 5 Visualisierung

Die Visualisierung von Daten hat im Rahmen von Big Data stark an Bedeutung gewonnen. Es können komplexe Zusammenhänge innerhalb riesiger Datenmengen mithilfe unterschiedlicher Visualisierungsinstrumente verständlich und kompakt dargestellt werden. Die Datenvisualisierung mit Business-Intelligence-Dashboards ermöglicht einer breiten Zielgruppe die eigenständige Datenanalyse und Erkenntnisgewinnung.<sup>79</sup>

Die im Rahmen dieses Beitrags separat untersuchten und ausgewerteten Bild- und Textdaten bilden die Datengrundlage für die Visualisierung. Diese Datengrundlage wurde in Bezug auf den OSMI analysiert. Die Visualisierung wurde dabei mit Qlik Sense, einem cloudbasierten Produkt der Firma QlikTech, umgesetzt.

### 5.1 Interaktive Visualisierung

#### 5.1.1 Dynamische Dashboards

Das Ziel der Visualisierung war, die Ergebnisse der Text- und Bildanalyse mithilfe eines dynamischen Business-Intelligence-Dashboards visuell darzustellen und aufzubereiten. Ein Business-Intelligence-Dashboard basiert in der Regel auf strukturierten Daten in Form einer oder mehrerer Tabellen. Beispiele für Tabellen sind unter anderem Excel-Tabellen, Tabellen aus CSV-Dateien und relationale Datenbanktabellen eines Data Warehouses. Diese sollen den Endnutzerinnen und Endnutzern ermöglichen, eine ansprechende, übersichtliche und verständnisfördernde Datenvisualisierung zu erzeugen.<sup>80</sup>

Wesentliche Vorteile eines dynamischen Dashboards sind z. B. ein konsistenter Überblick über alle Kennzahlen oder der Zugriff auf alle Daten in Echtzeit, wodurch Unternehmen wertvolle Zeit und Ressourcen sparen können. In diesem Zusammenhang werden zudem sogenannte Online-Analytical-Processing-Dashboards eingesetzt, welche die gleichzeitige Visualisierung verschiedener Datenquellen ermöglichen.<sup>81</sup>

---

<sup>79</sup> Vgl. Tufle 1983, S. 177–190.

<sup>80</sup> Vgl. Fels et al. 2015, S. 260.

<sup>81</sup> Vgl. Kohlhammer et al. 2018, S. 107–108.

### 5.1.2 Menschliche Wahrnehmung

Damit ein Business-Intelligence-Dashboard durch den Menschen effizient wahrgenommen und schnell Informationen verarbeiten werden können, ist neben dem dynamischen Interaktionsmechanismus die visuelle Konstruktion von zentraler Bedeutung.<sup>82</sup>

Um die Informationsverarbeitung des Menschen genauer zu verstehen, spielt die evolutionsbiologische Aufnahmekapazität des Menschen eine tragende Rolle. Ware hat hierzu untersucht, wie das menschliche Gehirn bei visuellen Darstellungen am schnellsten zu einer Entscheidungsfindung kommt. Ware definiert die folgenden drei Stufen der Wahrnehmung.

Die erste Stufe ist die parallele Verarbeitung von Informationen durch das menschliche Auge. Einzelne Gehirnneuronen nehmen unterschiedliche Informationen wahr. Im menschlichen Gehirn arbeiten Millionen von Neuronen parallel. Durch diese Neuronen können beispielsweise Kanten, Farben und Umrisse innerhalb von Millisekunden verarbeitet werden. Diese Informationsverarbeitung ist weitestgehend unabhängig von dem, was von Menschen bewusst wahrgenommen wird.<sup>83</sup>

Die zweite Stufe kennzeichnet die Mustererkennung. Das Gehirn verarbeitet kleine Regionen derselben Farbe, der gleichen Kontur sowie einfache Muster. Diese Phase ist äußerst flexibel und wird durch die Menge an Informationen aus der ersten Stufe beeinflusst. Auch der Schwerpunkt der eigenen Aufmerksamkeit spielt eine wichtige Rolle. Der Mensch ist deutlich schneller in der Lage Muster zu erkennen, wenn er bewusst danach sucht. Ware beschreibt zudem, dass der Prozess der Mustererkennung sowohl im Kurzzeit- als auch im Langzeitgedächtnis stattfindet.<sup>84</sup>

Die Berliner Schule der Gestaltpsychologie hat die wesentlichen Regeln für die Mustererkennung schriftlich festgehalten. Vor allem drei Gesetze sind für die visuelle Reporterstellung hierbei von zentraler Bedeutung: das Gesetz der Nähe, das Gesetz der Ähnlichkeit und das Gesetz der Verbundenheit.

Das Gesetz der Nähe beschreibt, dass Elemente, die sich in direkter Umgebung zueinander befinden, als zusammenhängend erkannt werden. Bei dem Gesetz der Ähnlichkeit werden Elemente mit gleichen Merkmalen eher als Gruppe wahr-

---

<sup>82</sup> Vgl. Taschnerx 2014, S. 79.

<sup>83</sup> Vgl. Ware 2004, S. 20–22.

<sup>84</sup> Vgl. Ware 2004, S. 20–22.

genommen als Elemente mit unterschiedlichen Merkmalen. Das Gesetz der Verbundenheit besagt, dass Elemente, die eine direkte Verbindung zueinander aufweisen, beispielsweise Punkte, die durch Linien verbunden wurden, als Einheit wahrgenommen werden. Insgesamt wirkt das Gesetz der Verbundenheit am stärksten, gefolgt vom Gesetz der Ähnlichkeit. Die schwächste Wirkung hat das Gesetz der Nähe.<sup>85</sup>

Die dritte und letzte Stufe wird von Ware als sequenzielle und zielgerichtete Verarbeitung gekennzeichnet. Es können nur wenige Objekte gleichzeitig bei der zielgerichteten Suche verfolgt werden. Der Suchprozess ist mit dem sprachlichen Gedächtnis verbunden, damit jedem Bild ein Wort zugeordnet werden kann. Wenn beispielsweise in einer Gewinn- und Verlustrechnung nach einer starken Budgetüberschreitung gesucht wird, so würde das gefundene Konto mit dem passenden Kontonamen in unserem Gehirn verarbeitet werden.<sup>86</sup>

Zusammenfassend sollte ein interaktives Dashboard alle drei Wahrnehmungsstufen berücksichtigen, wobei ebenfalls eine effiziente Gestaltung relevant ist. Um dies zu erreichen, werden Symbole, Farben und räumliche Anordnung von Informationen gewählt, die den kognitiven Wahrnehmungsprozess erleichtern und die Anwenderinnen und Anwender befähigen, Muster, Zusammenhänge oder Unregelmäßigkeiten in den Daten zu erkennen.<sup>87</sup>

Es wurden sowohl Kennzahlen, Grafiken als auch Tabellen verwendet, die in ihrer Kombination den Anwenderinnen und Anwendern einen möglichst guten Überblick über die Daten geben. Zudem wurden Selektionsmöglichkeiten eingebaut, die den Benutzerinnen und Benutzern die Möglichkeit geben, spezielle Teilansichten in den Dimensionen auszuwählen.

## **5.2 Vorbereitungen für die Visualisierung und explorative Datenanalyse**

### **5.2.1 Transfer und Vorbereitung der Daten**

Die Bild- und Textbeschaffung sowie Analyse erfolgten in zwei unterschiedlichen Ansätzen. Die aus diesen Ansätzen gewonnenen Daten mussten für die Visualisierung zusammengeführt werden. Um die Daten dieser beiden Ansätze zusammenzuführen, mussten die Textdaten zunächst transformiert bzw. standardisiert werden. Zunächst wurde die Min-Max-Normalisierung auf die OSMI-Indikatoren

---

<sup>85</sup> Vgl. Kohlhammer et al. 2018, S. 117.

<sup>86</sup> Vgl. Ware 2004, S. 22.

<sup>87</sup> Vgl. Jacobs & Hensel-Börner 2020, S. 52.

der Ergebnisse der Textanalyse angewandt. Dabei wurden die Werte zwischen null und eins skaliert. Im Anschluss wurden die aus der Bild- und Textanalyse gelieferten Daten zusammengeführt. Die Ergebnisse der Bildanalyse beinhalten, wie bereits in Kapitel 4.2.1 erläutert, ausschließlich einen Indikator des Sehsinns. Zur Ermittlung eines einheitlichen Sehsinns wurden die Daten aus der Textanalyse mit fünf Sechstel und aus der Bildanalyse mit einem Sechstel gemäß der OSMI-Definition gewichtet. Die Addition der gewichteten Werte ergab den neuen Seh Sinn.

Für die Dashboard-Erstellung wurden die Attribute Firma, Branche und Medium (Twitter bzw. Webseiten) verwendet. Die aus der Bildanalyse extrahierten Instagram-Daten wurden für die Visualisierungsergebnisse durch die zusätzliche Reduktion der Datenmenge, aus zeitlichen Gründen, nicht verwendet. Die extrahierten und von den Textanalysegruppen bereinigten Wörter wurden für die Word-Cloud in dem Dashboard inkludiert.<sup>88</sup>

## 5.2.2 Toolauswahl

Qlik Sense ist ein Business-Intelligence-Cloud-Service der Firma QlikTech. Diese erlaubt es, Datensätze zu verarbeiten und in einem Dashboard zu visualisieren. Qlik Sense kann explorativ genutzt werden, um unmittelbar Erkenntnisse über Datenstrukturen zu erhalten.<sup>89</sup>

QlikTech führte im Jahre 2014 mit Qlik Sense eine Self-Service-BI-Lösung ein, die es den Anwenderinnen und Anwendern ermöglicht, durch Drag- und Drop-Funktionen, Erkenntnisse zu gewinnen und Dashboards zu erstellen. Hierfür greift Qlik Sense wie schon sein Vorgänger Tool QlikView auf die In-Memory-Technologie zurück.<sup>90</sup>

Neben der Verwendung der Drag- und Drop-Funktion können die Entwickler auch auf Skriptfunktionalitäten zurückgreifen, um Datentabellen auf unterschiedlichste Weise miteinander zu kombinieren. So können Left-, Right-, Inner- und Outer-Join-Befehle durchgeführt werden. Zudem können Keep- und Concatenate-Statements verwendet werden. Diese Operationen helfen das benötigte Datenmodell zu konstruieren. Insgesamt bietet Qlik Sense aufgrund der vielen Funktionalitäten eine hohe Freiheitsdichte in der Gestaltung des Datenmodells.<sup>91</sup>

---

<sup>88</sup> Vgl. Larose & Larose 2015, S. 102.

<sup>89</sup> Vgl. Ilacqua et al. 2017, S. 4.

<sup>90</sup> Vgl. Labbe et al. 2018, o. S.

<sup>91</sup> Vgl. Harmsen & Garcia 2012, S. 227–231.

Dank der Qlik-Sense-Konnektivität können unterschiedliche Datenquellen und -formate verwendet werden. Die Konnektoren ermöglichen neben der direkten Verbindung zu verschiedenen Datenbanken auch das Einlesen externer Daten, beispielsweise von Google Analytics.<sup>92</sup>

Zudem bietet Qlik Sense R- und Python-Konnektivitäten. Mithilfe dieser Konnektivitäten können die Berechnungen in diesen Umgebungen ermöglicht und die Ergebnisse im Anschluss über Qlik Sense visualisiert werden. Dies birgt den Vorteil, dass unstrukturierte Daten zunächst mit Python verarbeitet und in eine strukturierte Form transformiert werden können.<sup>93</sup>

Qlik Sense kann entweder als On-Premise- oder in der Software-as-a-Service-Lösung genutzt werden. Für das OSMI-Dashboard wird auf die Software-as-a-Service-Lösung von QlikTech mit einer akademischen Lizenz zurückgegriffen.<sup>94</sup>

Ein Qlik-Sense-Dashboard bildet eine eigenständige Applikation auf dem Qlik-Server ab. Hierbei kann die Qlik-Applikation mehrere Arbeitsblätter beinhalten, die wiederum mehrere Elemente und Diagramme kombinieren. Zusätzlich können unterschiedliche Filterfunktionen auf den einzelnen Arbeitsblättern bereitgestellt werden.<sup>95</sup>

### 5.3 Dashboard Erstellung

Das für diesen Beitrag erstellte Dashboard setzt sich aus 15 Übersichtsblättern zusammen, die einem ähnlichen Aufbau folgen. Jedes Übersichtsblatt beinhaltet in der oberen Leiste acht Kennzahlen und vier Filtermöglichkeiten, die in Abbildung 14 dargestellt sind:

**Abbildung 14:** Kennzahlen und Filter des Dashboards



<sup>92</sup> Vgl. Rostalski, Wiegard & Kahl 2020, S. 101.

<sup>93</sup> Vgl. Labbe et al. 2019, S. 382–389.

<sup>94</sup> Vgl. Labbe et al. 2018, o. S.

<sup>95</sup> Vgl. Rostalski, Wiegard & Kahl 2020, S. 101.

Bei ausgewählten Filterkriterien erfolgt die simultane Anpassung der Kennzahlen und Graphen auf allen Übersichtsblättern. Das Entfernen oder Ändern dieser Filter kann an beliebigen Stellen im Dashboard erfolgen, womit eine hohe Benutzerfreundlichkeit gewährleistet ist.<sup>96</sup>

Das erste Übersichtsblatt bietet eine komprimierte Darstellung des finalen OSMI-Ergebnisses bezogen auf die vorhandenen Ebenen: Unternehmen, Branche, Medium sowie die Ansicht der verwendeten Algorithmen der Textanalyse.

Das zweite Blatt liefert die Word-Cloud der bereinigten Inputdaten von der Bild- und Textanalyse und die direkte Branchengegenüberstellung der fünf OSMI-Sinne in einem Radar-Chart.

Das dritte Blatt beinhaltet ein auf dem OSMI basierendes Verteilungsdiagramm über die Unternehmen je Branche.

Die Blätter vier bis acht beinhalten einen k-Means-Algorithmus je Branche. Die menschlichen Sinne werden auf den jeweiligen Seiten separat dargestellt

Die Blätter neun bis 14 beinhalten detaillierte Ergebnisse der verwendeten Textanalyse-Algorithmen in tabellarischer Form. Die Seiten neun bis 13 stellen diese Informationen separat je von den Analysegruppen verwendeten Verfahren zur Verfügung. Blatt 14 veranschaulicht die Gesamtwerte pro Verfahren.

Das letzte Blatt (15) bietet eine Erweiterung des ersten Blatts, indem es die Werte des OSMI pro Branche darstellt.

### **5.3.1 Gesamtübersicht**

Das erste Übersichtsblatt beinhaltet neben den wichtigsten Kennzahlen fünf Bereiche. Es werden die Branchen mit den dazugehörigen Unternehmen in tabellarischer Form mit entsprechenden OSMI pro Sinn dargestellt.<sup>97</sup> Die Werte werden mithilfe des arithmetischen Mittels über die verschiedenen Algorithmen gebildet. Qlik Sense bietet die Möglichkeit die Tabelle nach Bedarf zu sortieren, zu filtern und zu exportieren.

In Abbildung 15 ist zu erkennen, dass der höchste OSMI beim Sehsinn erreicht wurde. Dies legt die Vermutung nahe, dass im Onlinemarketing der Fokus in erster Linie auf der Präsentation visueller Reize liegt.

---

<sup>96</sup> Vgl. Labbe et al. 2018, S. 31–32.

<sup>97</sup> Vgl. Solanki & Hegde 2017, S. 211–212.



Wird die Tabelle hingegen absteigend nach dem Geschmackssinn sortiert, erscheinen vermehrt Unternehmen aus dem Branchencluster Food. Der zweite und dritte Bereich beinhaltet jeweils zwei Balkendiagramme. Es werden der Gesamt-OSMI je Branche sowie die Unternehmen mit den zehn höchsten OSMI-Werten dargestellt.

**Abbildung 15:** Ausschnitt des OSMI in der Gesamtübersicht

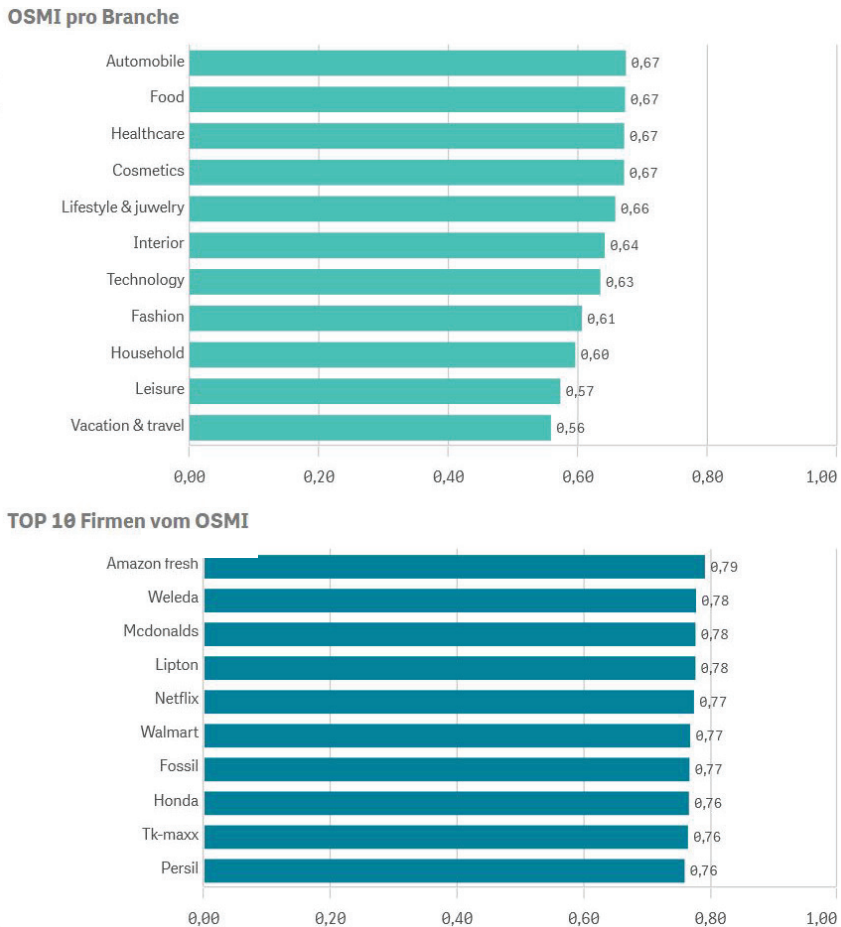
Übersicht OSMI

Firma	Branche	Ø-Sehsinn	Ø-Geruchssinn	Ø-Hörsinn	Ø-Geschmackssinn	Ø-Tastsinn	Ø-Gesamt
<b>Gesamtwerte</b>		<b>0,70</b>	<b>0,66</b>	<b>0,62</b>	<b>0,61</b>	<b>0,63</b>	<b>0,64</b>
28black	Food	0,76	0,69	0,60	0,74	0,61	0,68
Abercrombie & fitch	Fashion	0,82	0,74	0,73	0,75	0,70	0,75
Acer	Technology	0,68	0,72	0,63	0,72	0,64	0,68
Adidas	Cosmetics	0,75	0,68	0,74	0,61	0,67	0,69
Adidas	Fashion	0,75	0,72	0,69	0,66	0,68	0,70
After eight	Food	0,74	0,70	0,73	0,71	0,70	0,72
Aida	Vacation & travel	0,60	0,48	0,58	0,54	0,46	0,53
Alfa romeo	Automobile	0,67	0,65	0,61	0,63	0,64	0,64
Alltours	Vacation & travel	0,13	0,52	0,26	0,52	0,16	0,32
Alpina	Automobile	0,76	0,66	0,66	0,61	0,67	0,67
Always	Cosmetics	0,69	0,66	0,66	0,57	0,59	0,63
Amazon fresh	Food	0,75	0,82	0,82	0,83	0,73	0,79
American airlines	Vacation & travel	0,80	0,75	0,68	0,59	0,66	0,69
American standard	Interior	0,72	0,74	0,64	0,71	0,69	0,70
Amg	Automobile	0,72	0,66	0,66	0,65	0,72	0,68
Apple	Technology	0,65	0,63	0,65	0,56	0,65	0,63
Ariel	Household	0,50	0,53	0,42	0,11	0,43	0,40
Armani	Fashion	0,73	0,75	0,72	0,55	0,59	0,67
Asos	Fashion	0,77	0,69	0,76	0,80	0,73	0,75
Aston martin	Automobile	0,73	0,65	0,69	0,62	0,70	0,68
Audi	Automobile	0,77	0,69	0,59	0,56	0,65	0,65
Axe	Cosmetics	0,69	0,66	0,61	0,75	0,69	0,68
Bang & olufsen	Technology	0,71	0,68	0,65	0,63	0,63	0,66
Bayer	Healthcare	0,73	0,68	0,59	0,58	0,61	0,64
Bebe	Cosmetics	0,72	0,57	0,50	0,32	0,47	0,52
Ben jerrys	Food	0,76	0,67	0,76	0,82	0,74	0,75
Bentley	Automobile	0,72	0,75	0,60	0,66	0,68	0,68

Das Format des Balkendiagramms ist die bekannteste Visualisierungsmethode für direkte Größen- bzw. Mengenvergleiche unterschiedlicher Kategorien.<sup>98</sup> In Abbildung 16 werden zwei Beispiele aufgezeigt.

<sup>98</sup> Vgl. Garcia Pagans 2015, S. 161–163.

**Abbildung 16:** Balkendiagramm zum OSMI pro Branche und zu den Top 10 Firmen

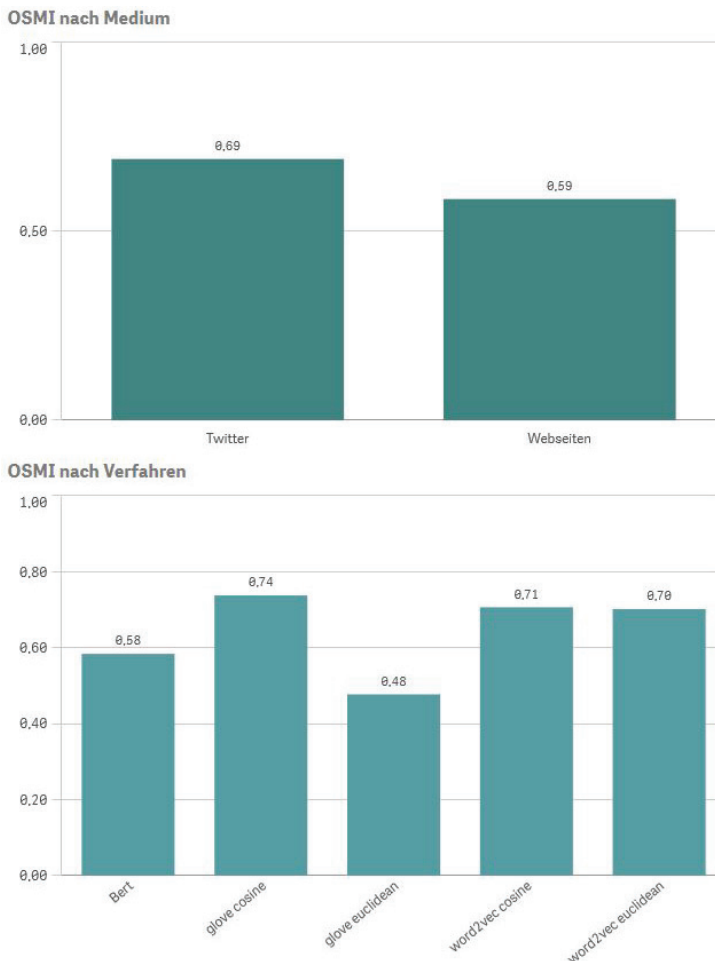


Mit der Auswahl einer Branche werden jeweils die Top-10-Unternehmen der jeweiligen Branchen angezeigt. In diesem Zusammenhang ist die ungleiche Verteilung der Unternehmen je Branche aufgefallen. Die Branchen Household und Healthcare setzen sich aus sechs Firmen zusammen, wohingegen in der Food-Branche 44 Firmen abgebildet werden und diese somit am stärksten repräsentiert ist. Aufgrund der geringen Stichprobengröße sind die Kategorien Household und Healthcare weniger aussagekräftig, da das Risiko einer Verzerrung durch Extremwerte besteht.

Der vierte und fünfte Bereich beinhaltet Darstellungen, die den OSMI nach Kommunikationsmedium und angewandtem Verfahren abbilden. Insgesamt lässt sich an der Abbildung 17 ablesen, dass Twitter als Kanal einen höheren OSMI aufweist

als Webseiten. Twitter-Nachrichten sind auf maximal 280 Zeichen begrenzt. Dies lässt erwarten, dass bei dem Einsatz von sensorischen Marketinginstrumenten eine höhere Dichte der markanten Wörter erreicht wird. Bei der Betrachtung des angewandten Textanalyseverfahrens lässt sich erkennen, dass bei dem BERT-Algorithmus die Twitter-Daten keinen höheren OSMI erreichen. Dies ist dadurch begründet, dass der BERT-Algorithmus keine Worteinbettung mit der anschließenden Bewertung vornimmt, sondern Sätze als Ganzes auswertet.<sup>99</sup> Die Machine-Learning-Word-Embedding-Verfahren GloVe und Word2Vec hingegen betrachten die einzelnen Wörter als Vektoren eines Vektorraums (Textkorpus).<sup>100</sup>

**Abbildung 17:** Balkendiagramm zum OSMI nach Medium und Verfahren



<sup>99</sup> Vgl. Sun et al. 2019, S. 194–206.

<sup>100</sup> Vgl. Pennington et al. 2014, S. 1532–1543.

Die Filterung nach den unterschiedlichen Verfahren hat Auswirkungen auf die tabellarische Darstellung der Gesamtergebnisse. Die Reihenfolge der jeweiligen Branchen und Firmen ist ebenfalls durch die Filterung betroffen. Es ist zu erkennen, dass innerhalb eines Verfahrens (GloVe vs. Word2Vec) die Ergebnisse eine hohe Ähnlichkeit aufweisen, auch wenn die absoluten Werte zum Teil große Unterschiede zeigen. In Abbildung 22 lässt sich dies gut an den Werten für Word2Vec cosine und Word2Vec euclidean einsehen. Insgesamt bewerten alle Verfahren Branchen wie Food und Cosmetics relativ hoch, wohingegen andere Kategorien wie Leisure und Vacation & Travel niedrige Bewertungen erhalten.

Die Analyse der drei durchgeführten Verfahren von Word2Vec, GloVe und BERT hat zudem gezeigt, dass das Word2Vec-Verfahren die höchsten OSMI-Werte liefert.

### 5.3.2 Radar- und Word-Chart

Auf dem zweiten Übersichtsblatt befinden sich zum einem die Inputwörter in Form einer Word-Cloud und zum anderen ein Radar-Chart mit den daraus resultierenden OSMI-Parameterwerten. Das Radar- oder auch Netzdiagramm ist eine Darstellungsform, die häufig für Vergleichszwecke eingesetzt wird. Dabei werden auf der äußeren Kreislinie die jeweiligen Messgrößen, in diesem Fall die Sinne, aufgeführt. Die y-Achse verläuft beginnend vom Mittelpunkt aufsteigend zum Rand, der die Maximalwerte repräsentiert.<sup>101</sup>

Die Word-Cloud basiert auf den von der Textanalyse zur Verfügung gestellten Daten, sodass keine Unterscheidung nach den verwendeten Analyseverfahren vorgenommen wurde. Es handelt sich hierbei um einen konsolidierten Datensatz mit dem gesamten OSMI für alle Verfahren.<sup>102</sup>

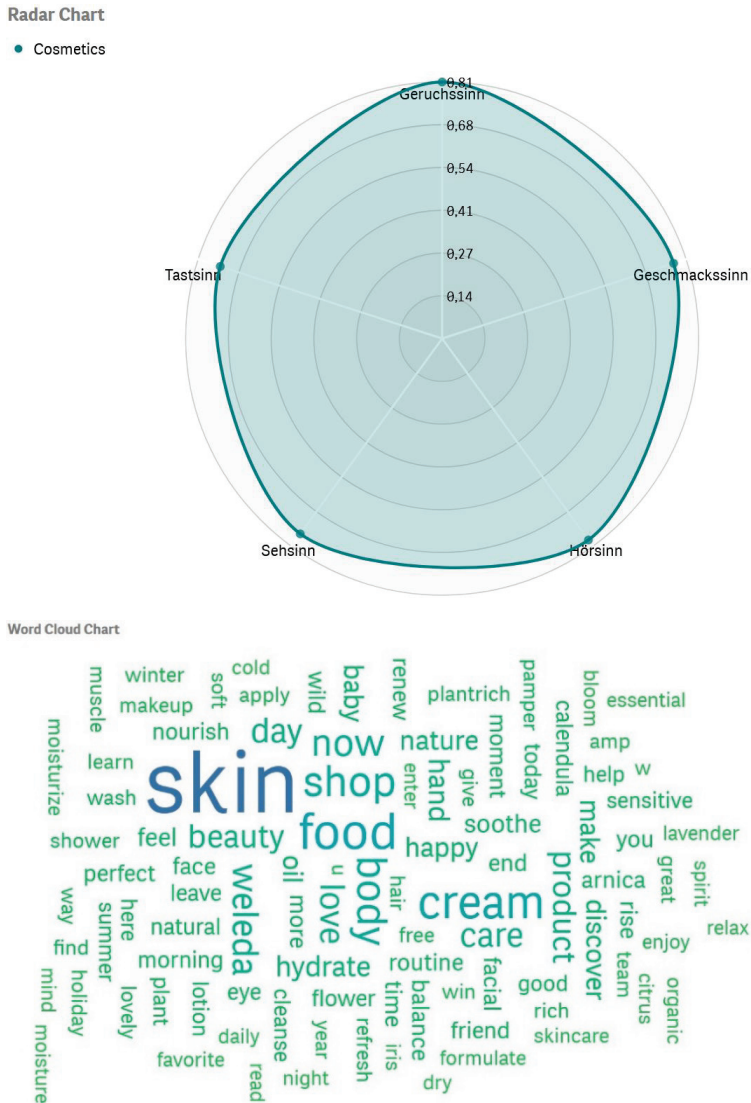
Abbildung 18 verdeutlicht am Beispiel Weleda, dass auch einzelne Unternehmen selektiert und dargestellt werden können. Weleda ist der Branche Cosmetics zugeordnet und innerhalb dieser das Unternehmen mit dem höchsten OSMI. Bei der Betrachtung der Ergebnisse fällt auf, dass alle Sinne gleichmäßig hohe Werte aufweisen.

---

<sup>101</sup> Vgl. Liu et al. 2008, S. 857–862.

<sup>102</sup> Vgl. Cui et al. 2010, S. 121–122.

**Abbildung 18:** Radar-Chart (oben) und Word-Cloud (unten) für Weleda



Die simultane Ansprache aller menschlichen Sinne liefert dementsprechend einen hohen Gesamtwert. Die in der Word-Cloud beinhalteten Wörter wie skin, hydrate, calendula, care und nature spiegeln das Image und die Werte des Unternehmens wider. Weleda steht in erster Linie für Nachhaltigkeit, Biodiversität sowie eine faire Produktion.<sup>103</sup>

<sup>103</sup> Vgl. Heidinger et al. 2019, S. 731–735.



### 5.3.3 Verteilungsdiagramm

Auf dem dritten Übersichtsblatt werden die Verteilungen der Unternehmen innerhalb der Branchen anhand des OSMI auf einer dynamischen Skala in Form eines Verteilungsdiagramms dargestellt. Die Skaleneinstellung kann zwischen statisch und dynamisch gewählt werden. Der Vorteil der dynamischen Darstellung ist, dass die zu untersuchenden Daten immer perfekt passend positioniert werden, dies birgt allerdings auch das Risiko einer Fehlinterpretation.

Bei der Betrachtung der Verteilungsdiagramme wird ein wesentlicher Vorteil der visuellen Datenanalyse deutlich. Diese bietet eine schnelle Möglichkeit, Ausreißer zu identifizieren.<sup>104</sup> In Abbildung 20 lässt sich erkennen, dass die zuvor erwähnten unterrepräsentierten Kategorien Household und Healthcare eine unterschiedliche Streuung des OSMI aufweisen. Healthcare liefert kompaktere, wenig streuende Ergebnisse. Im Gegenteil dazu weist der Bereich Household eine hohe Streuung auf, womit hier das Gesamtergebnis als volatil zu verstehen ist.

Mit der eingebauten Lasso-Funktion können Ausreißer herausgefiltert und die so bereinigten Ergebnisse pro Branche erneut betrachtet werden.

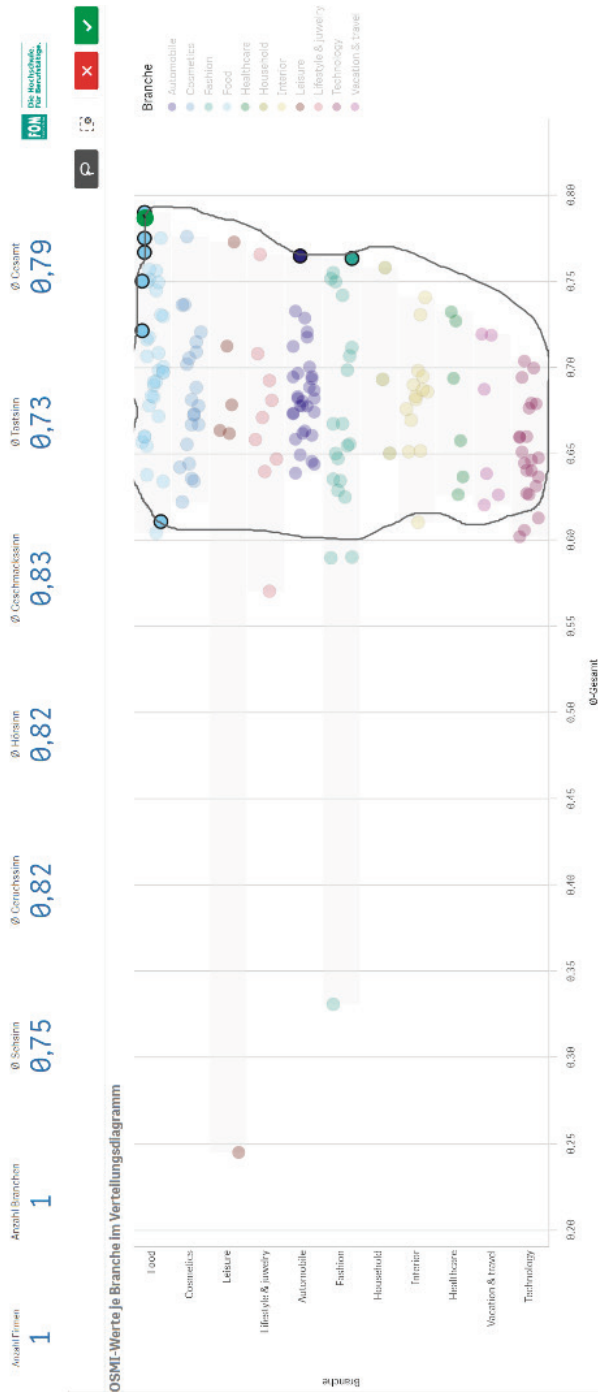
Die stärksten Ausreißer sind die Unternehmen Alltours, Dertour und Its-Reisen. Bei der Betrachtung der Webseiten dieser Unternehmen fällt auf, dass diese keine englischsprachige Internetpräsenz anbieten und somit fast ausschließlich deutsche Texte beinhalten. Der geringe OSMI dieser Webseiten lässt sich mithilfe der Datenbereinigung, -vorbereitung und schließlich den Analyseverfahren erklären, die allesamt auf englischsprachige Texte ausgelegt sind.

Vacation & Travel mit einem OSMI-Wert von 0,56 sowie Leisure mit 0,57 sind die Branchen, welche die niedrigsten Bewertungen erhalten haben. Vacation & Travel beinhaltet 13 Firmen und Leisure ist mit zehn Firmen vertreten. Der Abbildung 25 kann entnommen werden, dass diese Unternehmen insgesamt eine hohe Streuung innerhalb ihrer Branchen aufweisen. Im Vergleich dazu sind z. B. Automobile oder Technology mit einigen wenigen Ausreißern im Gesamtbild deutlich konsistenter. Wie zuvor beschrieben, treten vor allem bei den Kategorien mit kleinem Stichprobenumfang diese negativen Effekte auf. In den folgenden Unterkapiteln wird die Thematik erneut aufgegriffen sowie mögliche Gegenmaßnahmen erläutert.

---

<sup>104</sup> Vgl. Keim 2002, S. 1–6.

Abbildung 20: Verteilungsdiagramm und Lasso-Filter





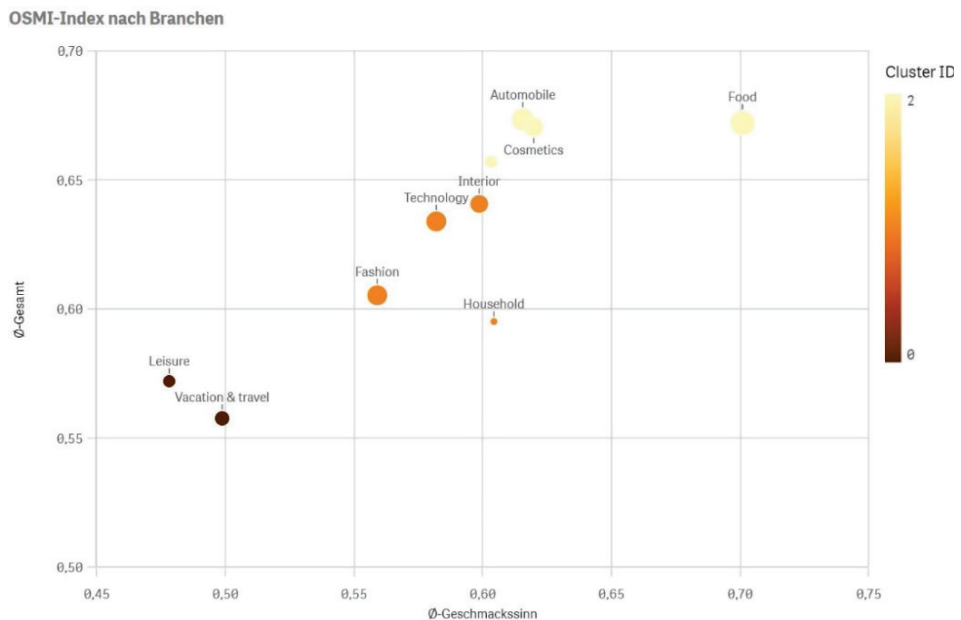
### 5.3.4 K-Means-Algorithmus pro Sinn

Auf den Seiten vier bis acht wird der k-Means-Algorithmus für die jeweiligen Sinne in Bezug auf den Gesamt-OSMI-Wert dargestellt. Der k-Means-Algorithmus ist ein Clusterverfahren. Mit diesem Algorithmus werden Datenpunkte in einem mehrdimensionalen Vektorraum in eine bestimmte Anzahl an Cluster gruppiert. Das Verfahren basiert auf der Minimierung des euklidischen Abstands jedes Punktes zu dem nächstgelegenen Cluster-Zentrum. Der Algorithmus verfolgt einen approximativen Ansatz und wird iterativ, mit dem Ziel ein lokales Minimum zu bestimmen, eingesetzt.<sup>105</sup>

Abbildung 21 verdeutlicht dies für den Geschmackssinn auf Basis von drei Clustern und einer Iteration. Es fällt der lineare Zusammenhang zwischen dem Geschmackssinn und dem Gesamt-OSMI auf. Die gebildeten Cluster sind:

- Dunkelbraun: Vacation & Travel, Leisure (Schwarz in Printversion)
- Orange: Household, Fashion, Technology, Interior (Dunkelgrau in Printversion)
- Gelb: Food, Cosmetics, Automobile, Healthcare und Lifestyle & Jewelry (Hellgrau in Printversion)

**Abbildung 21:** K-Means-Algorithmus: eine Iteration

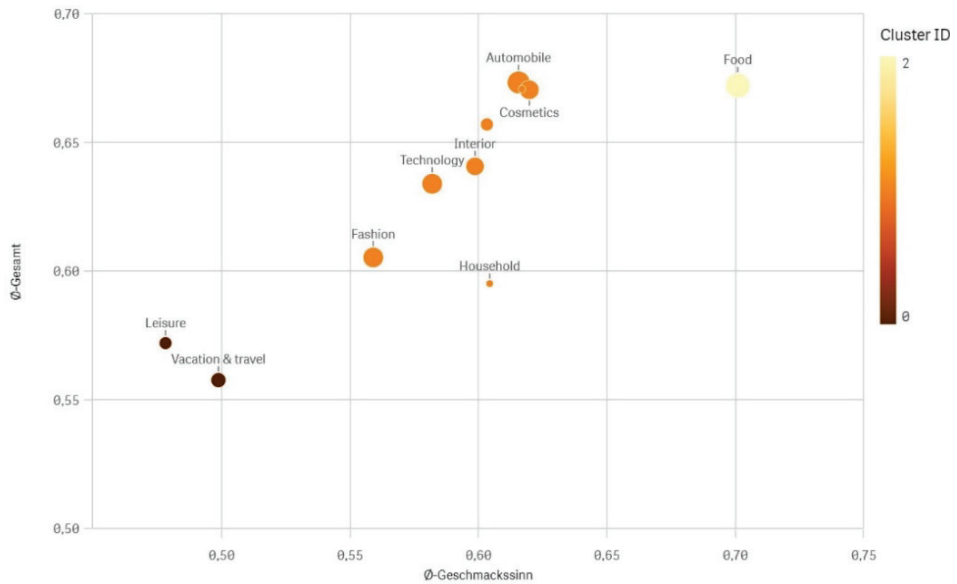


<sup>105</sup> Vgl. Kanungo et al. 2002, S. 881–892.

Wird die Anzahl der Iterationen auf drei erhöht, lässt sich eine Veränderung der Gruppierungen erkennen. Dies wird in Abbildung 22 dargestellt. Die erste Gruppe bleibt identisch, wohingegen die dritte Gruppe nur noch die Branche Food enthält. Die restlichen Branchen werden nun einem großen „Mittel-Cluster“ zugeordnet.

**Abbildung 22:** K-Means-Algorithmus: drei Iterationen

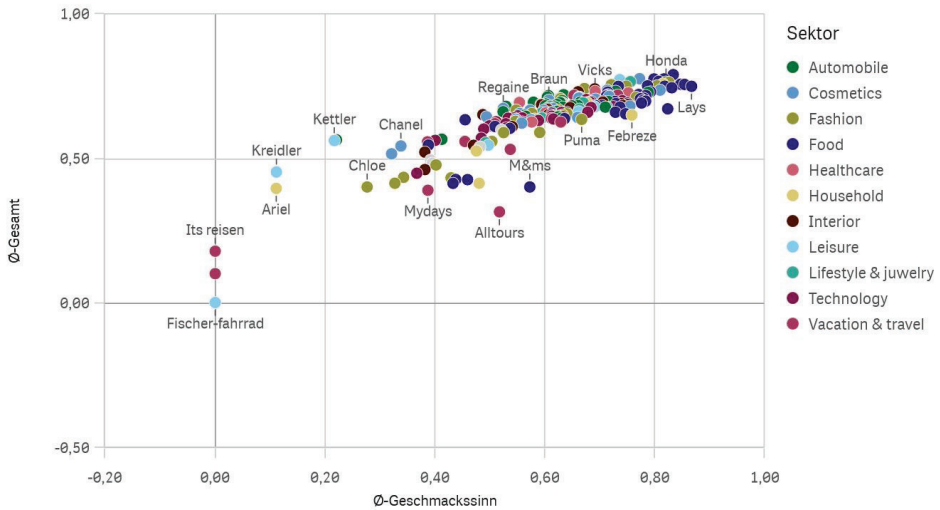
OSMI-Index nach Branchen



Es findet sich auf diesem Blatt ein detailliertes Punktdiagramm, das die jeweiligen Unternehmen anhand des Sinnes- und Gesamt-OSMI gegenüberstellt. Dies wird in Abbildung 23 verdeutlicht. Ausreißer sowie der lineare Zusammenhang zwischen Geschmackssinn und OSMI können ausgemacht werden.

**Abbildung 23:** Geschmackssinn und Gesamt-OSMI als Firmenübersicht

**OSMI nach Firmen**



### 5.3.5 Detailauswertungen der Textanalyseverfahren

Die Übersichtsblätter neun bis 13 des Dashboards beinhalten die detaillierten Ergebnisse der einzelnen Analyseverfahren für die jeweiligen Sinne in tabellarischer Form. Der OSMI wird jeweils pro Firma, Branche und Medium separat abgebildet und kann gut für konkrete Vergleichsauswertungen der Textanalyseverfahren verwendet werden.

Abbildung 24 zeigt die verfügbaren Tabellen des 14. Blatts. Es werden die OSMI-Ergebnisse der verschiedenen Verfahren abgebildet.

Die zuvor erläuterten Erkenntnisse bezüglich Twitter werden hier ebenfalls deutlich.

**Abbildung 24: Gesamt-OSMI-Algorithmen-Vergleich**

**OSMI-Algorithmen je Firma und Branche**

Firma	Q	Branche	Q	Ø-OSMI (glove cosine)	Ø-OSMI (glove euclidean)	Ø-OSMI (word2vec cosine)	Ø-OSMI (word2vec euclidean)	Ø-OSMI (Bert)	Ø-OSMI mit Bildern
<b>Gesamtwerte</b>				<b>0,75</b>	<b>0,48</b>	<b>0,71</b>	<b>0,71</b>	<b>0,58</b>	<b>0,64</b>
Langnese		Food		0,90	0,63	0,87	0,83	0,64	0,76
Honda		Automobile		0,90	0,61	0,85	0,87	0,58	0,76
Netflix		Leisure		0,90	0,66	0,87	0,87	0,57	0,77
Porsche		Automobile		0,90	0,48	0,76	0,75	0,56	0,68
Lays		Food		0,90	0,58	0,87	0,86	0,59	0,75
Persil		Household		0,90	0,64	0,83	0,81	0,61	0,76
Fossil		Lifestyle & jewelry		0,90	0,59	0,85	0,83	0,58	0,77
Nature valley		Food		0,89	0,57	0,79	0,80	0,62	0,74
Weleda		Cosmetics		0,89	0,66	0,88	0,88	0,60	0,78
Victorias secret		Fashion		0,89	0,60	0,81	0,76	0,60	0,74
Hollister		Fashion		0,89	0,66	0,84	0,84	0,58	0,76
Opel & vauxhall		Automobile		0,89	0,54	0,84	0,83	0,57	0,73
Asos		Fashion		0,88	0,62	0,85	0,86	0,57	0,75
Dominos		Food		0,88	0,70	0,83	0,82	0,59	0,75
Tk-maxx		Fashion		0,87	0,74	0,83	0,78	0,61	0,76
Mcdonalds		Food		0,87	0,69	0,83	0,82	0,59	0,78
Amazon fresh		Food		0,87	0,71	0,84	0,87	0,63	0,79
Abercrombie & fitch		Fashion		0,87	0,58	0,84	0,86	0,58	0,75
Dove		Cosmetics		0,87	0,47	0,85	0,85	0,59	0,72

**OSMI-Algorithmen je Branche**

Branche	Q	Ø-OSMI (glove cosine)	Ø-OSMI (glove euclidean)	Ø-OSMI (word2vec cosine)	Ø-OSMI (word2vec euclidean)	Ø-OSMI (Bert)	Ø-OSMI mit Bildern
<b>Gesamtwerte</b>		<b>0,75</b>	<b>0,48</b>	<b>0,71</b>	<b>0,71</b>	<b>0,58</b>	<b>0,64</b>
Healthcare		0,80	0,54	0,74	0,75	0,57	0,67
Automobile		0,80	0,54	0,75	0,75	0,56	0,67
Cosmetics		0,78	0,51	0,75	0,76	0,58	0,67
Lifestyle & jewelry		0,78	0,45	0,75	0,73	0,58	0,66
Food		0,76	0,51	0,74	0,74	0,60	0,67
Technology		0,76	0,47	0,72	0,71	0,56	0,63

**OSMI-Algorithmen je Medium**

Medium	Q	Ø-OSMI (glove cosine)	Ø-OSMI (glove euclidean)	Ø-OSMI (word2vec cosine)	Ø-OSMI (word2vec euclidean)	Ø-OSMI (Bert)	Ø-OSMI mit Bildern
<b>Gesamtwerte</b>		<b>0,75</b>	<b>0,48</b>	<b>0,71</b>	<b>0,71</b>	<b>0,58</b>	<b>0,64</b>
Twitter		0,81	0,56	0,76	0,76	0,58	0,69
Webseiten		0,67	0,39	0,66	0,65	0,58	0,59

### 5.3.6 Aggregierte Gesamtübersicht nach Branchen unterteilt

Wie bereits in Kapitel 5.3.3 erläutert, werden die OSMI-Ergebnisse teilweise stark von Ausreißern beeinflusst. Diese sollten eliminiert und damit die Ergebnisse bereinigt werden. Die Unternehmen wurden mittels des Lasso-Filters (Abbildung 20) von ursprünglich 213 auf 169 Unternehmen reduziert. Die eliminierten Ausreißer befinden sich alle im unteren OSMI-Wertebereich.

Für weiterführende Auswertungen lassen sich die Daten aus Qlik Sense in eine Excel-Tabelle überführen. Mithilfe einer bedingten Formatierung können die jeweiligen Sinne ihrem OSMI-Wert nach farblich hervorgehoben werden.

**Abbildung 25:** OSMI pro Branche

Rang	Branche	Ø-Sehsinn	Ø-Geruchssinn	Ø-Hörsinn	Ø-Geschmackssinn	Ø-Tastsinn	Ø-Gesamt
1	Automobile	0,74	0,70	0,65	0,62	0,67	0,67
2	Food	0,70	0,66	0,65	0,70	0,65	0,67
3	Healthcare	0,75	0,70	0,64	0,62	0,65	0,67
4	Cosmetics	0,69	0,70	0,68	0,62	0,66	0,67
5	Lifestyle & jewelry	0,72	0,71	0,63	0,60	0,63	0,66
6	Interior	0,71	0,66	0,61	0,60	0,63	0,64
7	Technology	0,71	0,65	0,61	0,58	0,62	0,63
8	Fashion	0,68	0,65	0,56	0,56	0,58	0,61
9	Household	0,66	0,59	0,56	0,60	0,57	0,60
10	Leisure	0,67	0,59	0,54	0,48	0,57	0,57
11	Vacation & travel	0,65	0,59	0,51	0,50	0,54	0,56

**Abbildung 26:** OSMI pro Branche nach Bereinigung der Ausreißer

Rang	Branche	Ø-Sehsinn	Ø-Geruchssinn	Ø-Hörsinn	Ø-Geschmackssinn	Ø-Tastsinn	Ø-Gesamt
1	Food	0,73	0,69	0,68	0,74	0,68	0,70
2	Household	0,74	0,66	0,67	0,75	0,69	0,70
3	Leisure	0,74	0,71	0,67	0,67	0,69	0,70
4	Interior	0,74	0,71	0,65	0,66	0,68	0,69
5	Automobile	0,75	0,70	0,66	0,63	0,68	0,68
6	Lifestyle & jewelry	0,73	0,75	0,65	0,64	0,64	0,68
7	Cosmetics	0,70	0,72	0,69	0,64	0,67	0,68
8	Fashion	0,73	0,71	0,65	0,65	0,65	0,68
9	Healthcare	0,75	0,70	0,64	0,62	0,65	0,67
10	Vacation & travel	0,76	0,70	0,61	0,62	0,65	0,67
11	Technology	0,71	0,67	0,62	0,60	0,64	0,65

Werden die Abbildung 25 und Abbildung 26 miteinander verglichen, wird deutlich, dass einige Branchen wie z. B. Automobile und Cosmetics im Gesamt-OSMI nur eine marginale Änderung (von 0,67 auf 0,68) durch die Ausreißer-Eliminierung erzielen. Dies bestätigt eine konsistente Datengrundlage mit nur wenigen, niedrig bewerteten Vertretern der Branche.

Im Vergleich dazu wird für die Branche Vacation & Travel eine deutliche Verbesserung sichtbar (von 0,56 auf 0,67). Dadurch ändert sich das Ranking der Branche von Platz elf auf Platz zehn.

Die Branchen Leisure und Household können sich ebenfalls verbessern. Jedoch hatten diese Branchen bereits vor der Eliminierung eine geringe Repräsentation von Unternehmen, die anschließend nochmals verringert wurde und daher als nicht repräsentativ betrachtet werden können.

## **6 Schlussbetrachtung**

### **6.1 Diskussion**

Das Ziel des Papers, sensorische Marketingunterschiede hinsichtlich branchenspezifischer Besonderheiten in Bezug auf die Sinne, eine Implementierung und das Design der Kommunikation zu erkennen sowie anhand des OSMI festzustellen, ob branchenspezifische Gewichtungen existieren, konnte mit dieser Arbeit nicht vollständig beantwortet werden. Jedoch können mit den Erkenntnissen dieser Arbeit sehr wohl Teilaussagen getroffen, Tendenzen abgelesen und neue Forschungsfragen und Hypothesen gebildet werden. Vor allem zeigt die Arbeit einen funktionierenden und skalierbaren Big-Data-Ansatz zur Determination eines OSMI für Branchen, für Unternehmen oder sogar einzelne Webseiten oder Inhalte der sozialen Medien der Unternehmen auf.

Der gesetzte Zeitrahmen und die gegebenen Ressourcen machten den Projektumfang zur einzig flexiblen Größe und führten somit zu einem nicht vollständigen OSMI. Es wurden alle vorgegebenen Unternehmen und Branchen untersucht, jedoch wurden viele der vorgestellten Indikatoren nicht bearbeitet. Diese lassen sich aber in Zukunft in die vorhandene OSMI-Architektur integrieren. Die erforschten Indikatoren dieser Arbeit basieren auf den Inhalten der Marketingkommunikation in Text- und Bildform. Erweiternd zum Projektauftrag wurden auch Inhalte aus den zukunftssträchtigen Social-Media-Kanälen extrahiert und untersucht. Einige der untersuchten Webseiten unterbinden die Extraktion der Inhalte durch verschiedene Mechanismen. Ein Umgehen der Mechanismen ist durch zusätzliche Programmierung möglich, wurde jedoch im Rahmen dieser Arbeit aus zeitlichen Gründen nicht weiterverfolgt. Die Extraktion der Web-Texte und Bilder wurde mithilfe der Scraping-Technik umgesetzt. Erweiternd zu den voreingestellten Extraktionskonditionen, die sicherstellen, dass die Inhalte nur von der Webseiten-Domäne stammen und Bilder eine bestimmte Größe haben, wurden im zweiten Schritt die Inhalte einer Qualitätsprüfung und Klassifizierung unterzogen. Die Social-Media-Inhalte wurden über die bereitgestellten, jedoch restriktiven API-Schnittstellen, wie auch durch erweitertes Scraping extrahiert. Diese Erweiterungen beinhalten benötigte Funktionalitäten wie das Aufschlagen der Webseiten durch ein automatisiertes Drücken des „Next Buttons“. Die im letzten Schritt geschaffene gemeinsame Datenbasis mit qualifizierten Social-Media- und Webseitentexten präsentiert eine skalierbare Input-Architektur. Der eindeutige Schlüssel dieser Datenbasis besteht aus der Branche, der Firma, dem Kommunikationskanal (Web bzw. Social-Media) und der Quelle (URL bzw. Social-Media-

ID). Der zugeordnete Inhalt besteht aus dem extrahierten Text bzw. dem Bildernamen, der wiederum den Schlüssel für das extrahierte Bild repräsentiert.

Aus der Input-Datenbasis lassen sich bereits erste Erkenntnisse aus Häufigkeiten ablesen. Als Beispiel werden die Branchen Haushalt und Healthcare genannt, die weit weniger Bilddaten in dem sozialen Netzwerk Instagram posten als Unternehmen anderer Branchen. Dies ist insbesondere dahingehend interessant, da viele der OSMI-Indikatoren erst eine stärkere Ausprägung vorweisen, wenn Bild und Text kombiniert werden und so das gesamtensorische Erlebnis maximieren. Die aufgestellte Hypothese, dass bestimmte Firmen oder ganze Branchen mit einem nicht materiellen Produktportfolio systematisch benachteiligt werden, sollte weiter untersucht werden, so dass der OSMI dahingehend durch Gewichtungen optimiert werden kann.

Erste Ergebnisse wurden jedoch erst durch die darauffolgenden, teils vielschichtigen Analysen geliefert. Die Textanalyse bildet den Schwerpunkt dieser Arbeit und trägt maßgeblich zu den kalkulierten OSMI-Werten bei. Der Ansatz verfolgt hierbei die Annahme, dass jeder der fünf Sinne indirekt durch das geschriebene bzw. gelesene Wort angesprochen werden kann. Dieser Ansatz kann, mit Einschränkungen, drei der beschriebenen Indikatoren zugeordnet werden (vgl. Tabelle 7). Die haptischen, visuellen und gustatorischen Ergebnisse lassen sich jedoch keinem der beschriebenen Indikatoren zuweisen – haben aber ihre Gültigkeit. Die Hinzunahme dreier textbasierter Indikatoren HT\_NEU, VT\_NEU und GT\_Neu (Tabelle 7) erlaubt die Berücksichtigung der Ergebnisse und schlägt dies als Verbesserungspotential für die OSMI-Definition vor.

**Tabelle 7:** Vorgeschlagene neue OSMI-Indikatoren

Indikator	Beschreibung	Ausprägung
VT_NEU 0 bis 3	Textbasierte Stimulation der visuellen Vorstellung	0 – Keine Stimulation 3 – Maximale Stimulation
HT_NEU 0 bis 3	Textbasierte Stimulation von haptischen Vorstellungen	0 – Keine Stimulation 3 – Maximale Stimulation
GT_NEU 0 bis 3	Textbasierte Stimulation von gustatorischen Vorstellungen	0 – Keine Stimulation 3 – Maximale Stimulation

Basierend auf den Erkenntnissen wird für den vorhandenen Indikator O2 eine stärkere Trennung der Inhaltsformen Text und Bild zur Berechnung vorgeschlagen (Tabelle 8). Diese Vorgehensweise ermöglicht zudem bei der anschließenden



den Ermittlung des OSMI mehr Gewichtungen vorzunehmen, die auch branchenspezifische Besonderheiten abbilden können. Die Skalen dieser Indikatoren sind zudem kontinuierlich und geben somit die Möglichkeit sowohl Quantität wie auch Qualität der sensorischen Merkmale zu erfassen.

**Tabelle 8:** Vorgeschlagene neue OSMI-Indikatoren

Indikator	Beschreibung	Ausprägung
O2 0 bis 2	Schlüsselbegriffe für Düfte	0 – Schwach: nicht vorhanden 1 – Standard: vorhanden 2 – Gut: kombiniert mit sensorischen Bildern
A2 0 bis	Schlüsselbegriffe für Akustik	0 – Keine Stimulation 3 – Maximale Stimulation

Die Erforschung der verschiedenen Algorithmen sowie der beiden Distanzmessungen eröffnet eine Vielzahl an neuen und interessanten Untersuchungsmöglichkeiten und Forschungsfeldern, die durch die vorgenommene Entkopplung der Textanalyse-Architektur ermöglicht wird. Auch zukünftige Änderungen, Optimierungen und Erweiterungen mit zusätzlichen Indikatoren können so vorgenommen werden.

Die Auswertung der Bilder lieferte zwei relevante OSMI-Indikatoren: G2 und V4. Ersterer untersucht und bewertet die Verwendung von Farbschemata, der zweite Indikator die Kontraste (Tabelle 9). Der Wertebereich beider Indikatoren ist binär skaliert. Der G2-Farbschemata-Indikator bietet jedoch Potential für eine kontinuierliche Skala, die eine differenzierte Sicht erlaubt. Die in Kapitel 4.3.3 vorgestellten fortschrittlichen Methoden der Bildanalyse wie die Objekterkennung, die Label-Analyse, die Gesichts- und Emotionserkennung sowie die Texterkennung in Bildern liefern dieser Arbeit einen spannenden Ausblick, was mit bereits verfügbarer Technologie möglich ist. Jedoch müssen die benötigten Speicher- und Prozessorressourcen für die Bildanalyse in zukünftigen Forschungsprojekten mit Bedacht ausgewählt und eingeplant werden.

**Tabelle 9:** Indikatoren der Bildanalyse

Indikator	Beschreibung	Ausprägung
G2 0-1	Verwendung von Farbschemata	0 – Schwach: nicht vorhanden bzw. schwarz, weiß oder zu viele, bunt gemischte Farben

		1 – Standard: Eine Farbe sinnlich kongruent angewandt
V4 0-1	Kontrast der Bilder / Webseite	0 – Schwach: nicht abgestimmt 1 – Standard: starker Kontrast

### 6.1.1 Interpretation der Ergebnisse

Aufbauend auf den Erkenntnissen und Interpretationen der Datenbeschaffung, Datenanalyse und vor allem der Visualisierung müssen diese auch mit der verfügbaren Literatur abgeglichen werden. Hierfür bietet sich die von Killian (2010a) vorgestellte branchenbezogene relative Wichtigkeit der Sinnesmodalitäten an (Tabelle 10).

**Tabelle 10:** Branchenbezogene relative Wichtigkeit der Sinnesmodalitäten<sup>106</sup>

Branche	Beispiel	Haptik	Geruch	Geschmack	Sound	Optik
<b>Automobile</b>	Auto	●●●●	●●●●	●	●●●●	●●●●●●
<b>Technologie</b>	TV	●●●●	●	●	●●●●●●	●●●●●●
<b>Haushalt</b>	Microwelle	●●●●	●●●●	●	●●●●	●●●●
<b>Einrichtung</b>	Couch	●●●●●●	●●●●	●	●●●●	●●●●●●
<b>Mode</b>	T-Shirt	●●●●●●	●●●●	●	●●	●●●●●●
<b>Büroartikel</b>	Stift	●●●●	●●	●	●●●●	●●●●
<b>Hygiene</b>	Parfüm	●●●●●●	●●●●●●	●	●●	●●●●
<b>Lebensmittel</b>	Chips	●●●●	●●●●●●	●●●●●●	●●●●	●●●●●●
<b>Reisen</b>	Flüge	●●●●	●●●●	●●	●●●●	●●●●●●
<b>Gesundheit*</b>	Vitamine	●●●●	●●●●	●●●●	●●●●	●●●●●●
<b>Freizeit*</b>	Streaming	●●●●	●	●	●●●●	●●●●●●
<b>Lifestyle*</b>	Uhr	●●●●	●●●●	●●	●●●●	●●●●●●

Legende:

- extrem wichtig / ● extrem unwichtig (Basis)
- Ursprüngliche Gewichtung / ● adaptiert auf digitales Marketing/
- \* zusätzliche Branchen (nicht in Killian (2010) enthalten)

<sup>106</sup> Tabelle in Anlehnung an Hamacher & Buchkremer 2022a.

Der im Rahmen dieser Arbeit determinierte OSMI spiegelt einen Großteil der von Killian aufgeführten Wichtigkeiten in Form eines hohen Wertes wider. Die folgende Liste beinhaltet eine unvollständige Auswahl von Übereinstimmungen:

- Der Sehsinn hat den insgesamt höchsten Stellenwert sowie OSMI-Wert.
- Der Geschmack hat den niedrigsten Stellenwert sowie OSMI-Wert.
- Schmecken, Riechen und Sehen haben den höchsten Stellenwert innerhalb der Food-Branche.
- Der Sehsinn hat den höchsten Stellenwert für die Automobil-, Fashion-, Elektronik- bzw. Technologiebranche.
- Der Geruchssinn hat den höchsten Stellenwert für die Hygiene- bzw. Kosmetikbranche.

Dieser Vergleich und die daraus abgelesene Übereinstimmung beweisen, dass es branchenspezifische OSMI-Wichtigkeiten gibt. Die Forschungsfrage dieser Arbeit ist eingeschränkt beantwortet und es werden darauf aufbauende Forschungsfelder aufgezeigt. Die in der Literatur aufgestellten Thesen sind damit bewiesen. Zusätzlich zeigt das Ergebnis, das Unternehmen für ihre sensorischen Marketinginhalte bevorzugt und mit höherer Dichte soziale Medien als Kommunikationskanäle nutzen.

### **6.1.2 Limitationen der Untersuchung**

Die Ergebnisse dieser Arbeit stellen eine Momentaufnahme dar, da die verwendeten Texte und Bilder allesamt auf extrahierten Internetinhalten basieren. Diese Inhalte sind dynamisch und wurden bereits im Jahr 2000 von McMillan als „Moving Target“ bezeichnet.<sup>107</sup> Diese Dynamisierung des Webcontents, die zunächst durch den Publikationsstil getrieben wurde, beschleunigte sich zunehmend über die letzten Jahre hin zu einer immer stärkeren Personalisierung des Internets. Zunächst wurde dies durch Marketing-Automation und heute durch die Anwendung von Künstlicher Intelligenz getrieben.<sup>108</sup>

Der in dieser Arbeit berechnete und angewandte OSMI beinhaltet nicht alle relevanten Inhalte und kann daher nicht als vollständig verstanden werden. Audio- und Videoinhalte sowie der Aufbau und Umfang der Unternehmenswebseiten

---

<sup>107</sup> Vgl. McMillan 2000.

<sup>108</sup> Vgl. Piyush Jain 2020.

fehlen gänzlich. Die beiden vorhandenen Inhaltstypen Texte und Bilder bilden lediglich sieben OSMI-Indikatoren ab.

### 6.1.3 Empfehlung für weiterführende Forschung

Tabelle 11 gliedert die weiterführende Forschungsempfehlung in die folgenden drei Bereiche: neue Inhalte, Vertiefung und Optimierung. Zudem stehen die offenen Forschungsfelder in einem Bezug zueinander. Zum Beispiel lassen sich Videoinhalte in einen auditiven und visuellen Teil unterteilen. Aus diesen Teilen können gesprochene oder visuelle Textinhalte extrahiert werden und dem in dieser Arbeit vorgestellten Textanalyse-Korpus zur Bewertung zugeführt werden. Die neu gewonnenen Kennzahlen können anschließend in Form vorhandener oder neuer OSMI-Indikatoren hinterlegt werden. Zudem stehen die Gewichtungen der Indikatoren zueinander und der Parameter unter Berücksichtigung der jeweiligen Branche aus.

**Tabelle 11:** Weitere Forschungsfelder

Neue Inhalte	
Audio	Audioinhalte wurden in dieser Arbeit nicht berücksichtigt. Sie stellen ein Element der Grundanforderungen zur Determination des OSM-Index dar.
Video	Videoinhalte wurden in dieser Arbeit nicht berücksichtigt. Sie stellen ein Element der Grundanforderungen zur Determination des OSM-Index dar.
Webseitenumfang	Der Umfang der Webseiten mit Prüfung, ob Forum, Chat und andere Online-Features angeboten werden, wurde in dieser Arbeit nicht berücksichtigt, wird jedoch für die vollständige Ermittlung des OSMI benötigt und bedarf somit einer zusätzlichen Erarbeitung.
Virtual Reality	Die Erforschung, Bewertung und Integration von VR-Inhalten stellt eine weitere vielversprechende Erweiterung des OSMI dar.
Vertiefung der Forschung	
Bildanalyse	Die in Kapitel 4.3.3 Weitere Bildanalysen beschriebenen Möglichkeiten stellen ein enormes Potenzial zur OSMI-Berechnung dar.
Textanalyse Methoden	Evaluierung der drei genutzten Algorithmen Glove, Word2vec und BERT mit anschließender Bestimmung des oder der optimalen Algorithmen.
Textanalyse Methoden	Identifikation und Bewertung von aktiven Formulierungen: „Erleben Sie die Frische“.

OSMI-Optimierungen	
Evaluierungsebenen	Die Bereitstellung der OSMI-relevanten Ergebnisse auf einer detaillierteren Ebene ermöglicht eine bessere Nachvollziehbarkeit und bereitet die Basis für die effektive Überprüfung und Optimierung der Algorithmen und der Gewichtungen.
Indikatoren nach Inhaltstyp	Eine weitere Aufspaltung der Indikatoren ermöglicht die separate Speicherung und Gewichtung der aus den Analysen gewonnenen Ergebnisse. Dies unterstützt die Erstellung von branchen- oder unternehmensspezifischen Gewichtungen, die eine höhere Genauigkeit und Vergleichbarkeit garantieren.

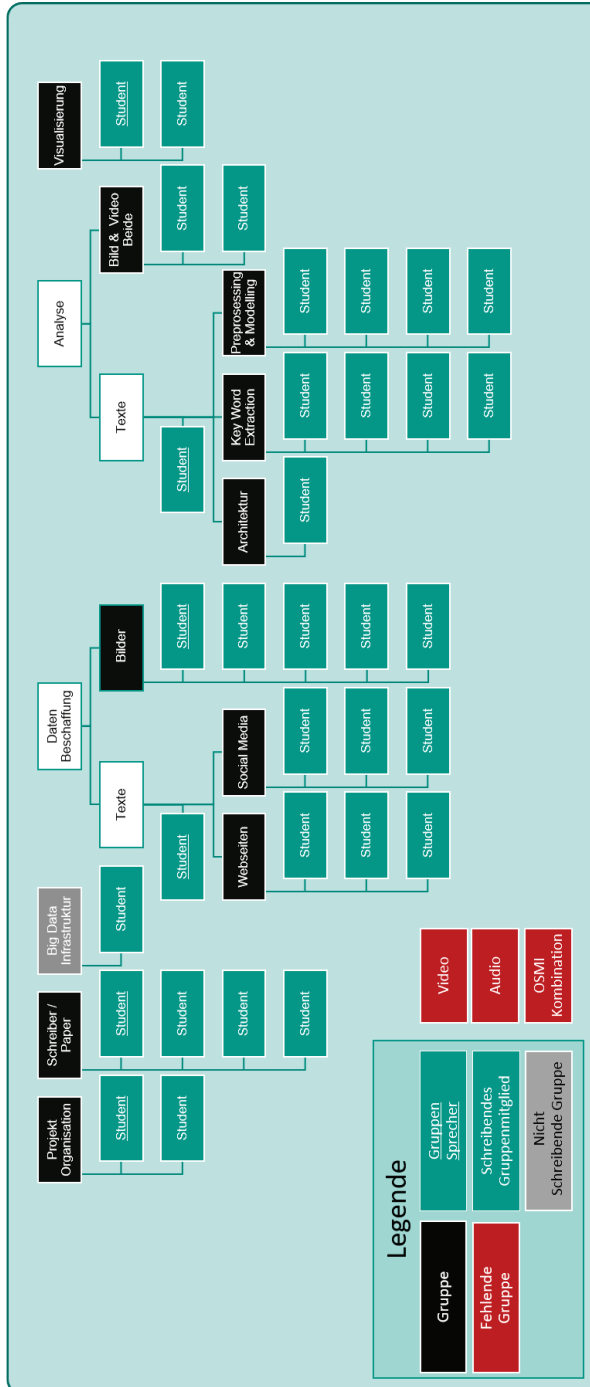
## 6.2 Fazit

Im Rahmen des Zeitplans, der begrenzten Ressourcen und der vorgegebenen Forschungsfragen ist das Ergebnis als erfolgreich zu bewerten. Die Arbeit beantwortet nicht alle Forschungsfragen, weist aber weitere Forschungsfelder aus, um tiefer gehende Erkenntnisse zu erlangen.

Die Zusammenarbeit von 33 Studierenden an diesem Paper war für alle Beteiligten eine praxisnahe und lehrreiche Erfahrung. Sich in einer Gruppe von über 30 Kommilitoninnen und Kommilitonen für eine wissenschaftlichen Arbeit, während der durch die Corona-Pandemie bedingten Einschränkungen zu koordinieren, abzustimmen und zu motivieren, ohne die Möglichkeit zu haben, sich persönlich zu treffen, stellte ein besondere Herausforderung dar, welche souverän gemeistert wurde.



## Anhang 2: Projektgruppenplan



## Literatur

- Abdallah, S. & Jaleel, B. (2015). *Website Appeal: Development of an Assessment Tool and Evaluation Framework of E-Marketing*. 10(3), 45–62. <https://doi.org/10.4067/S0718-18762015000300005>
- Ahmadiyah, A. S., Sarno, R. & Revindasari, F. (2020). Adopted topic modeling for business process and software component conformity checking. *Telkomnika (Telecommunication Computing Electronics and Control)*, 18(6). <https://doi.org/10.12928/TELKOMNIKA.v18i6.13381>
- Ao, X., Yu, X., Liu, D. & Tian, H. (2020). News keywords extraction algorithm based on TextRank and classified TF-IDF. *2020 International Wireless Communications and Mobile Computing, IWCMC 2020*, 1364–1369. <https://doi.org/10.1109/IWCMC48107.2020.9148491>
- Appel, G., Grewal, L., Hadi, R. & Stephen, A. T. (2020). The future of social media in marketing. *Journal of the Academy of Marketing Science*, 48(1). <https://doi.org/10.1007/s11747-019-00695-1>
- Bower, B. (2004). The Brain's Word Act. *Science News*, 165(6), 83. <https://doi.org/10.2307/4014782>
- Brasel, S. A. & Gips, J. (2015). Interface Psychology: Touchscreens Change Attribute Importance, Decision Criteria, and Behavior in Online Choice. *Cyberpsychology, Behavior, and Social Networking*, 18(9), 534–538. <https://doi.org/10.1089/cyber.2014.0546>
- Cian, L., Krishna, A. & Elder, R. S. (2014). This Logo Moves Me: Dynamic Imagery from Static Images. *Journal of Marketing Research*, 51(2), 184–197.
- Cui, W., Wu, Y., Liu, S., Wei, F., Zhou, M. X. & Qu, H. (2010). Context Preserving Dynamic Word Cloud Visualization. In *Pacific Visualisation Symposium 2010*. Taiwan: IEEE.
- Daugherty, P. R. & Wilson, H. J. (2018). *Human + Machine: Künstliche Intelligenz und die Zukunft der Arbeit*. dtv Sachbuch.
- Elder, R. S. & Krishna, A. (2010). The Effects of Advertising Copy on Sensory Thoughts and Perceived Taste. *Journal of Consumer Research*, 36(5), 748–756.
- Fels, G., Lanquillon, C., Mallow, H., Schinkel, F. & Schulmeyer, C. (2015). Technik. In Joachim Dorschel (Hg.), *Praxishandbuch Big Data*. Springer Fachmedien Wiesbaden [https://doi.org/10.1007/978-3-658-07289-6\\_4](https://doi.org/10.1007/978-3-658-07289-6_4)



- Fernández-Cavia, J., Rovira, C., Díaz-Luque, P. & Cavaller, V. (2014). Web quality index (WQI) for official tourist destination websites. proposal for an assessment system. *Tourism Management Perspectives*, 9, 5–13.
- Foscht, T., Brandstätter, M. & Sinha, I. (2010). Reverse Psychology Marketing – Konsequent falsch und doch richtig. *Marketing Review St. Gallen* 6 (2010), 18–25, <https://doi.org/10.1007/s11621-010-0094-6>
- Garcia Pagans, F. (2015). *Predictive Analytics Using Rattle and Qlik Sense*. Packt Publishing.
- González-Benito, Ó., Martos-Partal, M. & San Martín, S. (2015). Brands as substitutes for the need for touch in online shopping. *Journal of Retailing and Consumer Services*, 27, 121–125. <https://doi.org/10.1016/j.jretconser.2015.07.015>
- Goodfellow, I., Bengio, Y., Courville, A. & Bengio, Y. (2016). *Deep Learning*. The MIT Press.
- Hamacher, K. & Buchkremer, R. (2022a). Measuring Online Sensory Consumer Experience: Introducing the Online Sensory Marketing Index (OSMI) as a Structural Modeling Approach. In *Journal of Theoretical and Applied Electronic Commerce Research*, 17(2), 751–772. <https://doi.org/10.3390/jtaer17020039>
- Hamacher, K. & Buchkremer, R. (2022b). The Application of Artificial Intelligence to Automate Sensory Assessments Combining Pretrained Transformers with Word Embedding Based on the Online Sensory Marketing Index. *Computers*, 11, 1–17.
- Harmsen, B. & Garcia, M. (2012). *Qlikview 11 for developers*. Packt Publishing.
- Hartmann, O. & Haupt, S. (2016). *Touch: Der Haptik-Effekt im multisensorischen Marketing* (2. erweiterte Auflage). Haufe-Lexware GmbH & Co. KG.
- Hayes-Roth, B., Pflieger, K., Lalanda, P., Morignot, P. & Balabanovic, M. (1995). A Domain-Specific Software Architecture for Adaptive Intelligent Systems. *IEEE Transactions on Software Engineering*, 21(4). <https://doi.org/10.1109/32.385968>
- Heidinger, I., Lotz, S. & Wyciok, B. (2019). Nachhaltigkeit aus dem Blickwinkel der Vielfalt am Beispiel von Weleda. In Marco Englert und Anabel Ternès (Hg.) *Nachhaltiges Management*, 731–47. Springer, [https://doi.org/10.1007/978-3-662-57693-9\\_39](https://doi.org/10.1007/978-3-662-57693-9_39)

- Hejing, W., Fang, L., Long, Z., Yabin, S. & Ran, C. (2020). Application Research of Crawler and Data Analysis Based on Python. *International Journal of Advanced Network, Monitoring and Controls*, 5(2). <https://doi.org/10.21307/ijanmc-2020-018>
- Hultén, B. (2011). Sensory marketing: the multi-sensory brand-experience concept. *European Business Review*, 23(3), 256–273.
- IfD Allensbach. (2016). *Anteil der Online-Käufer an der Bevölkerung in Deutschland\* von 2000 bis 2016*. (Statista - Das Statistik-Portal (ed.)). <https://de.statista.com/statistik/daten/studie/2054/umfrage/anteil-der-online-kaeuer-in-deutschland/>, Zugriff am 09.01.2023.
- Ilacqua, C., Cronstrom, H., Richardson, J. & Hand, P. (2017). *Qlik Sense: Advanced Data Visualization for Your Organization: Create smart data visualizations and predictive analytics solutions*. Packt Publishing.
- Jacobs, L. & Hensel-Börner, S. (2020). Die Kraft effektiver Daten-Visualisierung – CLEAR(I): Ein Leitfaden zur wirkungsvollen Dashboard-Gestaltung. In Silvia Boßow-Thies, Christina Hofmann-Stölting und Heike Jochims (Hg.), *Data-driven Marketing*, 43–75. Springer Fachmedien Wiesbaden, [https://doi.org/10.1007/978-3-658-29995-8\\_3](https://doi.org/10.1007/978-3-658-29995-8_3)
- Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R. & Wu, A. Y. (2002). An Efficient k-Means Clustering Algorithm: Analysis and Implementation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7), 881–92, <https://doi.org/10.1109/TPAMI.2002.1017616>
- Kedia, A. & Rasu, M. (2020). *Hands-On Python Natural Language Processing*. Packt Publishing.
- Keim, D. A. (2002). Information Visualization and Visual Data Mining. *IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS*, 7(1), 1–8.
- Kim, A. Y., Ha, J. G., Choi, H. & Moon, H. (2018). Automated text analysis based on skip-gram model for food evaluation in predicting consumer acceptance. *Computational Intelligence and Neuroscience*, 2018. <https://doi.org/10.1155/2018/9293437>
- Kim, Y. T. (1997). Contrast enhancement using brightness preserving bi-histogram equalization. *IEEE Transactions on Consumer Electronics* 43(1), 1–8. <https://doi.org/10.1109/30.580378>

- Kohlhammer, J., Proff, D. U. & Wiener, A. (2018). *Visual Business Analytics* (2. Auflage). dpunkt.verlag.
- Köster-Düpre, F. (2020). Soziale Medien als Kundenbindungsinstrument. In Oliver D. Doleski (Hg.), *Realisierung Utility 4.0 Band 2. Praxis der digitalen Energiewirtschaft vom Vertrieb bis zu innovativen Energy Services*, 171–82. Springer Fachmedien Wiesbaden. [https://doi.org/10.1007/978-3-658-25589-3\\_14](https://doi.org/10.1007/978-3-658-25589-3_14)
- Kreutzer, R. T., Rumler, A. & Wille-Baumkauff, B. (2020). B2B-Online-Marketing und Social Media. Springer Fachmedien. <https://doi.org/10.1007/978-3-658-27675-1>
- Krishna, A. (2012). An integrative review of sensory marketing: Engaging the senses to affect perception, judgment and behavior. *Journal of Consumer Psychology*, 22(3), 332–351.
- Krishna, A., Morrin, M. & Sayin, E. (2014). Smellizing cookies and salivating: A focus on olfactory imagery. *Journal of Consumer Research*, 41(1), 18–34. <https://doi.org/10.1086/674664>
- Kuang, S. & Davison, B. D. (2020). Learning class-specific word embeddings. *Journal of Supercomputing*, 76(10), 8265–8292. <https://doi.org/10.1007/s11227-019-03024-z>
- Labbe, P., Anjos, C., Solanki, K. & DiMaso, J. (2019). *Hands-On Business Intelligence with Qlik Sense: Implement self-service data analytics with insights and guidance from Qlik Sense experts*. Packt Publishing.
- Labbe, P., Hand, P. & Kharpate, N. (2018). *Qlik Sense Cookbook*. Packt Publishing.
- Larose, D. T. & Larose, C. D. (2015). *Data Mining and Predictive Analytics* (Second Edition). Wiley.
- Lee, S. (Ally), Jeong, M. & Oh, H. (2018). Enhancing customers' positive responses: Applying sensory marketing to the hotel website. *Journal of Global Scholars of Marketing Science*, 28(1), 68–85. <https://doi.org/10.1080/21639159.2017.1410772>
- Liu, W.-Y., Wang, B.-W., Yu, J.-X., Li, F., Wang, S.-X. & Hong, W.-X. (2008). Visualization classification method of multi-dimensional data based on radar chart mapping. In *2008 International Conference on Machine Learning and Cybernetics*. Kunming: IEEE.

- McCabe, D. B. & Nowlis, S. M. (2003). The Effect of Examining Actual Products or Product Descriptions on Consumer Preference. *Journal of Consumer Psychology*, 13(4), 431–439. [https://doi.org/10.1207/S15327663JCP1304\\_10](https://doi.org/10.1207/S15327663JCP1304_10)
- McMillan, S. J. (2000). The microscope and the moving target: The challenge of applying content analysis to the World Wide Web. *Journalism and Mass Communication Quarterly*, 77(1), 80–98. <https://doi.org/10.1177/107769900007700107>
- Meert, K., Pandelaere, M. & Patrick, V. M. (2014). Taking a shine to it: How the preference for glossy stems from an innate need for water. *Journal of Consumer Psychology*, 24(2), 195–206. <https://doi.org/10.1016/j.jcps.2013.12.005>
- Mishra, A. (2019). Amazon Rekognition Developer Guide. In *Machine Learning in the AWS Cloud*. Wiley.
- Peck, J. & Childers, T. L. (2008). Effects of Sensory Factors on Consumer Behavior: If It Tastes, Smells, Sounds, and Feels Like a Duck, Then It Must Be A. In Curtis P. Haugtvedt, Paul M. Herr, Frank R. Kardes (Hg.), *Handbook of Consumer Psychology*, 193–219. Routledge.
- Peck, J., Barger, V. A. & Webb, A. (2013). In search of a surrogate for touch: The effect of haptic imagery on perceived ownership. *Journal of Consumer Psychology*, 23(2), 189–196.
- Peck, J. & Shu, S. B. (2009). The Effect of Mere Touch on Perceived Ownership. *Journal of Consumer Research*, 36(3), 434–447. <https://doi.org/10.1086/598614>
- Peck, J. & Wiggins, J. (2006). It Just Feels Good: Customers' Affective Response to Touch and Its Influence on Persuasion. *Journal of Marketing*, 70(4), 56–69. <https://doi.org/10.1509/jmkg.70.4.56>
- Pennington, J., Socher, R. & Manning, C. D. (2014). GloVe: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). *Association for Computational Linguistics*.
- Petit, O., Cheok, A. D., Spence, C., Velasco, C. & Karunanayaka, K. T. (2015). Sensory marketing in light of new technologies. *Proceedings of the 12th International Conference on Advances in Computer Entertainment Technology - ACE '15*, 1–4.

- Piyush Jain, K. A. (2020). Transforming marketing with artificial intelligence. *Irjet*, 3(4), 291–297. <https://doi.org/10.13140/RG.2.2.25848.67844>
- Ramos, J. (2003). Using TF-IDF to Determine Word Relevance in Document Queries. *Proceedings of the First Instructional Conference on Machine Learning*, 242(1), 29–48.
- Rao, D. & McMahan, B. (2019). *Natural Language Processing mit PyTorch*. O'Reilly Media, Inc.
- Raposo, A., Moss, H. E., Stamatakis, E. A. & Tyler, L. K. (2009). Modulation of motor and premotor cortices by actions, action words and action sentences. *Neuropsychologia*, 47(2), 388–396. <https://doi.org/10.1016/j.neuropsychologia.2008.09.017>
- Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. <https://doi.org/10.1109/CVPR.2016.91>
- Roggeveen, A. L., Grewal, D., Townsend, C. & Krishnan, R. (2015). The Impact of Dynamic Presentation Format on Consumer Preferences for Hedonic Products and Services. *Journal of Marketing*, 79(6), 34–49.
- Rostalski, S.-M., Wiegard, M. & Kahl, T. (2020). Entwicklung eines Einsteiger-Workshops zur Datenvisualisierung und -analyse mit dem Business Intelligence Tool Qlik Sense®. In: Timo Kahl and Frank Zimmer (Hg.) *Interaktive Datenvisualisierung in Wissenschaft und Unternehmenspraxis*, 95–118,. Springer Vieweg. <https://doi.org/10.1007/978-3-658-29562-2>
- San-Martín, S., González-Benito, Ó. & Martos-Partal, M. (2017). To what extent does need for touch affect online perceived quality? *International Journal of Retail and Distribution Management*, 45(9), 950–968. <https://doi.org/10.1108/IJRDM-04-2016-0054>
- Shu, S. & Peck, J. (2007). To Hold Me Is To Love Me : The Role of Touch in the Endowment Effect. *NA - Advances in Consumer Research*, 34, 513–514.
- Solanki, K. & Hegde, G. (2017). *Implementing Qlik Sense: Design, Develop, and Validate BI solutions for consultants*. Packt Publishing.
- Spears, N. & Yazdanparast, A. (2014). Revealing obstacles to the consumer imagination. *Journal of Consumer Psychology* (24)3, 363–372. <https://doi.org/10.1016/j.jcps.2014.01.003>

- Spence, C., Okajima, K., Cheok, A. D., Petit, O. & Michel, C. (2016). Eating with our eyes: From visual hunger to digital satiation. *Brain and Cognition*, 110, 53–63.
- Statistisches Bundesamt. (2017). *Fast jedes 4. Unternehmen verkauft online* (Statista - Das Statistik-Portal (ed.)). <https://de.statista.com/infografik/12188/anteil-der-unternehmen-in-deutschland-die-im-e-commerce-aktiv-sind/>, Zugriff am 20.12.2022.
- Stout, P. A. & Leckenby, J. D. (1988). Let the music play: Music as a nonverbal element in television commercials. In S. Hecker und D. W. Stewart (Hg.), *Nonverbal Communication in Advertising*, 207–233. Lexington Books.
- Sun, C., Qiu, X., Xu, Y. & Huang, X. (2019). How to Fine-Tune BERT for Text Classification? *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11856 LNAI. [https://doi.org/10.1007/978-3-030-32381-3\\_16](https://doi.org/10.1007/978-3-030-32381-3_16)
- Taschnerx, A. (2014). Wie Management Reporting 2020 aussehen könnte. *Controlling & Management Review*, 58(3). <https://doi.org/10.1365/s12176-014-0930-7>
- Towers, I. (2014). Sensory Marketing. In Anabel Ternès und Ian Towers (Hg.), *Internationale Trends in der Markenkommunikation*. 107–17. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-01517-6>
- Trask, A. W. (2020). *Neuronale Netze und Deep Learning kapieren: Der einfache Praxiseinstieg mit Beispielen in Python*. MITP-Verlags GmbH & Co. KG.
- Trussel, H. J., Saber, E. & Vrhel, M. (2005). Color image processing. In *IEEE Signal Processing Magazine*, 22(1). <https://doi.org/10.1109/MSP.2005.1407711>
- Tufle, E. (1983). *The visual display of quantitative information - Graphical Excellence*. CT Graphics, Cheshire.
- Ware, C. (2004). *Information Visualization: Perception for Design* (Second Edition). Morgan Kaufmann. <https://doi.org/10.1016/B978-1-55860-819-1.X5000-6>
- Yazdanparast, A. & Spears, N. (2013). Can Consumers Forgo the Need to Touch Products? An Investigation of Nonhaptic Situational Factors in an Online Context. *Psychology and Marketing* 30(1), 46–61). <https://doi.org/10.1002/mar.20588>

Yosida, K. (1965). *Functional Analysis*. (Die Grundlehren der mathematischen Wissenschaften, Band 123) Springer.

Zhao, Z. Q., Zheng, P., Xu, S. T. & Wu, X. (2019). Object Detection with Deep Learning: A Review. In *IEEE Transactions on Neural Networks and Learning Systems*. <https://doi.org/10.1109/TNNLS.2018.2876865>

### **Internetquellen:**

Beck, K., Beedle, M. & Bennekum, A. van (2001). *Principles behind the agile manifesto*. <https://agilemanifesto.org/principles.html>, Zugriff am 09.01.2023.

Homburg, C., Imschloß, M. & Kühnl, C. (2013). *Multisensorisches Marketing : Beeinflussung der haptischen Weichheitswahrnehmung durch Musik*. [https://www.bwl.uni-mannheim.de/media/Einrichtungen/imu/Research\\_Insights/2013/RI\\_016.pdf](https://www.bwl.uni-mannheim.de/media/Einrichtungen/imu/Research_Insights/2013/RI_016.pdf), Zugriff: 09.01.2023.

<https://docs.scrapy.org/en/latest/topics/architecture.html>, Zugriff am: 09.01.2023.

<https://github.com/ultralytics/yolov5>, Zugriff am 10.02.2021

Rosebrock, A. (2022). *Intersection over Union (IoU) for object detection*. <https://pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/>, Zugriff am 05.01.2023.

**Folgende Bände sind bisher in dieser Reihe erschienen:**

**Band 1 (2021)**

Bähren, T. / Braka, D. / Burchard, P. / Cyron, S. / Demary, M. / Dragieva, M. /  
Eis, L. / Farid, A. T. / Gomes, D. / Hacker, M. / Kaiser, J. / Krüger, R. / Luu, S. /  
Maasjosthusmann, R. / Marks, A. / Pachocki, C. / Pongratz, M. / Schade, J. C. /  
Urban, P. / Walter, A. / Winter, V. / Yesilyurt, E. / Buchkremer, R.

Der Einsatz von grünem Tee und anderen Polyphenolen in der Medizin – eine  
Big-Data-Analyse der medizinischen Fachliteratur

Essen 2021

ISBN (Print) 978-3-89275-119-9 – ISBN (eBook) 978-3-89275-120-5

ISSN (Print) 2699-562X – ISSN (eBook) 2699-5638





Forschungsstark und praxisnah:

# Deutschlands Hochschule für Berufstätige

Raphaela Schmaltz studiert den  
berufsbegleitenden Master-Studiengang  
Taxation am FOM Hochschulzentrum Köln.

Die FOM ist Deutschlands Hochschule für Berufstätige. Sie bietet über 40 Bachelor- und Master-Studiengänge, die im Tages- oder Abendstudium berufsbegleitend absolviert werden können und Studierende auf aktuelle und künftige Anforderungen der Arbeitswelt vorbereiten.

In einem großen Forschungsbereich mit hochschuleigenen Instituten und KompetenzCentren forschen Lehrende – auch mit ihren Studierenden – in den unterschiedlichen Themenfeldern der Hochschule, wie zum Beispiel Wirtschaft & Management, Wirtschaftspsychologie, IT-Management oder Gesundheit & Soziales. Sie entwickeln im Rahmen nationaler und internationaler Projekte gemeinsam mit Partnern aus Wissenschaft und Wirtschaft Lösungen für Problemstellungen der betrieblichen Praxis.

Damit ist die FOM eine der forschungsstärksten privaten Hochschulen Deutschlands. Mit ihren insgesamt über 2.000 Lehrenden bietet die FOM mit mehr als 50.000 Studierenden ein berufsbegleitendes Präsenzstudium im Hörsaal an einem der 36 FOM Hochschulzentren und ein digitales Live-Studium mit Vorlesungen aus den hochmodernen FOM Studios.

Alle Institute und KompetenzCentren unter  
[fom.de/forschung](https://www.fom.de/forschung)

Die Hochschule.  
Für Berufstätige.



**FOM**



Institut für IT-Management &  
Digitalisierung  
der FOM University of Applied Sciences

## FOM Hochschule

FOM. Die Hochschule. Für Berufstätige.

Mit über 50.000 Studierenden ist die FOM eine der größten Hochschulen Europas und führt seit 1993 Studiengänge für Berufstätige durch, die einen staatlich und international anerkannten Hochschulabschluss (Bachelor/Master) erlangen wollen.

Die FOM ist der anwendungsorientierten Forschung verpflichtet und verfolgt das Ziel, adaptionsfähige Lösungen für betriebliche bzw. wirtschaftsnahe oder gesellschaftliche Problemstellungen zu generieren. Dabei spielt die Verzahnung von Forschung und Lehre eine große Rolle: Kongruent zu den Masterprogrammen sind Institute und KompetenzCentren gegründet worden. Sie geben der Hochschule ein fachliches Profil und eröffnen sowohl Wissenschaftlerinnen und Wissenschaftlern als auch engagierten Studierenden die Gelegenheit, sich aktiv in den Forschungsdiskurs einzubringen.

Weitere Informationen finden Sie unter [fom.de](https://www.fom.de)

## ifid

Das ifid Institut für IT-Management & Digitalisierung bündelt Kompetenzen in den Forschungsbereichen Künstliche Intelligenz (KI), Systemwissenschaften, IT-Management und digitale Transformation.

Die Aufgaben des Instituts umfassen Forschung und Entwicklung, Wissenstransfer und Innovationsförderung an der Schnittstelle von Wissenschaft und Praxis. Auch der Transfer von Forschungserkenntnissen in die Lehre spielt eine große Rolle.

Um diese Aufgaben zu erfüllen, setzt die Forschergruppe auf den Einsatz modernster Big Data-Architekturen und KI-Analysesysteme. Es bestehen Kooperationen mit den großen Technologie-Unternehmen und Instituten der Branche.

Die Wissenschaftlerinnen und Wissenschaftler beschäftigen sich insbesondere mit folgenden Feldern:

- Künstliche Intelligenz / Machine Learning / Data Science / Big Data
- Natural Language Processing (NLP)
- Enterprise Architekturen (insbesondere Big Data)
- Einsatz von Blockchain-Technologien
- Digitalisierung von Prozessen
- Integration der Forschung in die Lehre

Weitere Informationen finden Sie unter [fom-ifid.de](https://www.fom-ifid.de)