

Me-Commerce: Eine explorative Analyse der Filternutzung im Online Shop zur Personalisierung am Beispiel von ottoversand.at

Masterarbeit

eingereicht von: **Edith Holzner, BA**
Matrikelnummer: 51826800

im Fachhochschul-Masterstudiengang Wirtschaftsinformatik
der Ferdinand Porsche FernFH GmbH

zur Erlangung des akademischen Grades

Master of Arts in Business

Betreuung und Beurteilung: Christoph Jungbauer, BA MA MA

Zweitgutachten: Dipl.-Ing.ⁱⁿ Eszter Geresics-Földi, BSc MSc

Wiener Neustadt, September 2020

Ehrenwörtliche Erklärung

Ich versichere hiermit,

1. dass ich die vorliegende Masterarbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Alle Inhalte, die direkt oder indirekt aus fremden Quellen entnommen sind, sind durch entsprechende Quellenangaben gekennzeichnet.
2. dass ich diese Masterarbeit bisher weder im Inland noch im Ausland in irgendeiner Form als Prüfungsarbeit zur Beurteilung vorgelegt oder veröffentlicht habe.
3. dass die vorliegende Fassung der Arbeit mit der eingereichten elektronischen Version in allen Teilen übereinstimmt.

Wiener Neustadt, 16.09.2020

Unterschrift

Kurzzusammenfassung: Me-Commerce: Eine explorative Analyse der
Filternutzung im Online Shop zur Personalisierung am Beispiel von
ottoversand.at

Diese Arbeit beleuchtet die Filternutzung auf Suchergebnisseiten im OTTO Online Shop, und beantwortet die wissenschaftliche Frage, welche Personalisierungsmaßnahmen sich basierend auf der Filternutzung ableiten lassen. Dabei wurde in einer hermeneutischen Forschung versucht, die vorherrschende Literaturlücke in Bezug auf Personalisierung basierend auf der Filternutzung auf Suchergebnisseiten im Online Shop zu schließen. Es wurde festgestellt, dass der Kunde der neuen 4. Generation großen Wert auf relevanten und perfekt zugeschnittenen Inhalt und Produktergebnisse legt und dabei erwartet, dass die personenbezogenen Daten zu jeder Zeit sicher und geschützt sind. Herausforderungen bei der Produktsuche sind ein möglicher Information Overload sowie das Auswahl-Paradoxon. Genau dem kann durch personalisierte Suchergebnisseiten basierend auf der Filternutzung entgegengesteuert werden. In Bezug auf die Ergebnisse der explorativen Analyse sollten Informationen zu den relevantesten Filtern (Farbe, Größe, Preis, Marke) zum einen explizit – z.B. über Eingabeformulare von Lieblingsmarken und getragenen Größen – und zum anderen implizit durch das Web-Tracking der Klickdaten gesammelt werden. Die Speicherung und Verarbeitung dieser Daten ermöglicht es den Online-Retailern, den Kunden bei der Produktsuche perfekt zu unterstützen, den richtigen Content zu liefern und so die Kaufabschlüsse zu steigern. Ein DSGVO-konformer Umgang mit den personenbezogenen Daten ist dabei eine Grundvoraussetzung.

Schlagwörter:

Personalisierung; Onsite Suche; Filternutzung; E-Commerce; Online Shop; Filter; Produktsuche; Produktfilterung;

Abstract: Me-Commerce: An exploratory analysis of filter usage in the online shop for personalization using the example of ottoversand.at

This thesis examines the filter usage on search results pages in the OTTO online shop and answers the research question of which personalization actions can be taken based on the filter usage. Hermeneutic research attempted to close the prevalent literature gap with regard to personalization based on the use of filters on search result pages in the online shop. It was found that the new 4th generation of customer appreciates relevant and perfectly tailored content and product results and expects that personal data are secure and protected at all times. The challenges in the product search are a possible information overload and the paradox of choice. This can be counteracted through personalized search results pages based on filter usage. With regard to the results of the exploratory analysis, information on the most relevant filters (colour, size, price, brand) should be explicitly – e.g. via input forms of favourite brands and sizes worn – and on the other hand, are collected implicitly through web tracking of the click data. The storage and processing of this data enables online retailers to perfectly support customers in their product search, to deliver the right content and thus to increase sales. A GDPR-compliant handling of personal data is a basic requirement.

Keywords:

personalization; on-site search; filter usage; e-commerce; online shop; filter; search results; product filtering

Inhaltsverzeichnis

1. Einleitung	1
1.1 Problemstellung, Relevanz & Zielsetzung.....	1
1.2 Aktueller Forschungsstand & konkrete Forschungsfrage.....	3
1.3 Aufbau der Arbeit & methodisches Vorgehen	4
2. Theoretische Grundlagen	5
2.1 Definitionen und Begriffserklärungen.....	5
2.1.1 E-Commerce.....	5
2.1.2 Customer Journey und Touchpoints	5
2.1.3 Personalisierung.....	7
2.1.4 Onsite Search und Filter.....	8
2.1.5 Web Analytics	10
2.1.6 Big Data	11
2.1.7 Data Mining.....	12
2.1.8 Clusteranalyse.....	12
2.2 Reizüberflutung	15
2.3 Paradox of Choice.....	16
2.4 Filter-Bubble	18
3. Hermeneutische Forschung	20
3.1 Die Entwicklung des E-Commerce hin zum Ubiquitous Commerce.....	20
3.2 Der Kunde 4.0 im Jahr 2020	22
3.3 Die Customer Journey 4.0 im E-Commerce	24
3.4 Onsite Search am Beispiel von OTTO Österreich.....	30
3.5 Die Herausforderungen und Probleme bei der Produktsuche.....	32
3.6 Die Herausforderungen der Personalisierung hinsichtlich Datenschutz..	35
3.7 Die Herausforderungen der Personalisierung hinsichtlich Filter-Bubbles	38
3.8 Die Chancen und Möglichkeiten der Personalisierung im Online Shop ...	41

4. Durchführung der Datensammlung im Online Shop.....	45
4.1 Implementierung der Filter-Funktion im OTTO Online Shop.....	45
4.2 Erhebung und Aufzeichnung von Filternutzungs-Trackingdaten	46
4.3 Beschreibung der Auswertung der Filternutzungs-Daten	49
5. Ergebnisse und Interpretation der empirischen Forschung.....	52
5.1 Wie hoch ist die Filternutzungsrate allgemein?	52
5.2 Sind Sortimente erkennbar, bei denen Filter relevanter bzw. weniger relevant sind?	52
5.3 Sind Suchbegriffe erkennbar, bei denen Filter relevanter bzw. weniger relevant sind?	54
5.4 Welche Erkenntnisse lassen sich aus dem Filternutzungsverhalten der User*innen ableiten?	56
5.4.1 Markenfilter.....	57
5.4.2 Preisfilter	64
5.4.3 Größenfilter	72
5.4.4 Farbfilter.....	78
5.4.5 Geschlechtsfilter & Verfügbarkeitsfilter.....	84
5.5 Gibt es User*innen, die Suchergebnisse häufig mit Filter einschränken und in mehreren Sessions untersuchte Filter setzen?	91
6. Diskussion	94
6.1 Konkrete Handlungsempfehlungen & Implikationen	94
6.2 Zusammenfassung und Fazit	101
6.3 Limitationen der Arbeit und Ausblick.....	103
Literaturverzeichnis	106
Abbildungsverzeichnis	116
Tabellenverzeichnis	118
A. Quellcode	121
A.1. SQL-Abfrage zur Aggregation der Filternutzung	121
A.2. Ergänzung der einzelnen Filterwerte der untersuchten Filter.....	125

1. Einleitung

1.1 Problemstellung, Relevanz & Zielsetzung

Das Internet hat zweifelslos eine enorme Welle an Veränderungen eingeleitet. Bereits 2000 schrieb Watson über Geschäfte, die in einer Welt gemacht werden, welche allgegenwärtig, universell, einzigartig und unisono ist (Richard T. Watson, 2000). 20 Jahre später ist der Begriff „Ubiquitous Computing“ – die allgegenwärtige Datenverarbeitung – Teil der gängigen Volkssprache (Krumm, 2018). Mit der Weiterentwicklung des E-Commerce haben die Fortschritte in der Informations- & Kommunikationstechnologie (IKT) das Aufkeimen einer neuen Art des Handelns beschleunigt. Dieser allgegenwertige Handel tritt als kontinuierlicher, nahtloser Strom, bestehend aus Kommunikation, Content und Dienstleistungen auf, der zwischen Unternehmen, Lieferant*innen, Mitarbeiter*innen, Kund*innen und Produkten ausgetauscht wird (Kumar, Joshi & Saquib, 2015).

Potentielle Kund*innen haben im Zeitalter des E-Commerce täglich, immer und überall die Möglichkeit, Inhalte zu konsumieren und Kaufentscheidungen zu treffen (Heinemann, 2019). Kund*innen und User*innen werden ständig mit vielen Marketing-Reizen konfrontiert. Die Informationen, Meldungen und Nachrichten werden jedoch häufig ignoriert, es sei denn, die User*innen erkennen einen Mehrwert darin (Orenga-Roglá & Chalmeta, 2016). Im besten Fall wird jede Kund*in und jeder User als Individuum betrachtet (Heinemann, 2019). Um den Individuen genau die Inhalte bieten zu können, die aufgrund von persönlichen Merkmalen, Vorlieben und Interessen einen Mehrwert liefern, ist es heute Pflicht, Daten zu personalisieren (Anshari, Almunawar, Lim & Al-Mudimigh, 2018). Denn die Fokussierung auf die unterschiedlichen Kundenwünsche durch Individualisierung und Personalisierung der angebotenen Inhalte und Services stellt eine wesentliche Vertrauensbasis im Onlinegeschäft dar (Stüber & Hudetz, 2017).

Geht man nach den Annahmen des Homo oeconomicus, müsste es für die User*in gut sein, möglichst viele Auswahlmöglichkeiten bei der Kaufentscheidung zu haben. So kann die Kund*in das für sie passendste Produkt wählen und die Bedürfnisse bestens befriedigen. Laut Befragungen gibt auch immer wieder die Mehrheit der Befragten an, dass sie eine größere Auswahl an Optionen besser findet. Doch bei genaueren Untersuchungen ergibt sich das differenzierte Bild, dass Auswahlzufriedenheit und Entscheidungsfreude nicht gleichläufig sind (Bauer & Koth, 2014). Schwartz (2005) nennt diesen Zusammenhang das „Paradox of Choice“. Geht man nach diesem Paradoxon, müssen Unternehmen der Kund*in die Kaufentscheidung so leicht wie möglich machen. Bei einer zu großen Auswahl sind die Menschen überfordert, bevor sie eine falsche Entscheidung treffen, treffen sie lieber gar keine. Im Zweifel müssen die Unternehmen den Kund*innen weniger statt mehr Optionen zur Auswahl anbieten (Schwartz, 2005).

Speziell für Online Shops ist dieses Phänomen in der Decision-Making-Phase eine enorme Herausforderung. Ein großes Sortiment ist auf den ersten Blick ein Alleinstellungsmerkmal, doch kann es die Kaufentscheidung extrem erschweren und sogar dazu führen, dass die Kund*in bzw. User*in den Kauf abbricht. Die Anzahl an angezeigten Produkt-Optionen kann beispielsweise durch den intensiven Einsatz von Filterkriterien und Sortieralgorithmen reduziert werden. (Spreer, 2018)

Auch eine von IfH Köln durchgeführte Studie im Jahr 2015 belegt, dass eine gute Filterfunktion in Online Shops extrem wichtig für die Österreicher ist. So behaupten 43,1 %, dass für sie eine gute Filterfunktion absolut wichtig ist, 35,2 % gaben an, dass sie eher schon wichtig ist (Statista Research Department, 2015). Denkt man einen Schritt weiter, könnte das User-Verhalten der Filternutzung im Online Shop zur Personalisierung des Shops genutzt werden (Mayer, 2017). Denn Personalisierung funktioniert dort am besten, wo die gläserne Kund*in ausreichend Informationen und Daten über sich selbst preisgibt (Volkens & Anderson, 2017. S. 82). Im ökonomischen Feld des E-Commerce kann laut Volkens & Anderson (2017) die Such- und Kaufhistorie von Kund*innen noch leichter verfolgt werden, die

Aggregation, Konsolidierung und Analyse der Daten erfolgt nahezu in Echtzeit. Umso relevanter ist die Gewährleistung des Datenschutzes und der Datensicherheit. Auch die Sicherstellung ethischer Grenzen ist sowohl für Unternehmen als auch für Kund*innen relevant (Volkens & Anderson, 2017, S. 82).

Diese Masterarbeit soll analysieren und erklären, welche Potentiale die Nutzung der Filter auf den Suchergebnisseiten hat, um den User*innen relevante und personalisierte Inhalte und Produkte anzubieten und den Personalisierungsgrad noch weiter zu steigern.

1.2 Aktueller Forschungsstand & konkrete Forschungsfrage

Es wurden bereits verschiedene Forschungen und wissenschaftliche Arbeiten zu verwandten und thematisch nahen Bereichen dieser Arbeit durchgeführt. So gibt es in Bezug auf die allgemein gesehene Personalisierung in Online Shops eine Vielzahl an Studien, die die Vor- und Nachteile beleuchten (Kwon & Kim, 2012; Lee & Cranage, 2011; Lewandowski, 2018; Pariser, 2011b; Stüber & Hudetz, 2017).

Auch mit Fokus auf die Personalisierung von Search Engines gibt es verschiedene Forschungsergebnisse (Jürgens et al., 2014; Lewandowski, 2018; Oulasvirta, Hukkinen & Schwartz, 2009; Riemer & Brüggemann, 2007; Storey, 2011).

Eine klar erkennbare Forschungslücke konnte in der durchgeführten Literaturanalyse jedoch in Bezug auf die Kombination der Filternutzung auf Suchergebnisseiten in Online Shops und die Weiterverarbeitung der gesammelten Nutzungsdaten dieser Filter in der übergreifenden Personalisierungsstrategie identifiziert werden. Basierend auf den Erkenntnissen der Problemstellung und der identifizierten Forschungslücke lässt sich folgende Forschungsfrage ableiten, die mit Hilfe der Masterarbeit beantwortet werden soll:

Welche Personalisierungs-Maßnahmen lassen sich durch die Filter-Nutzung auf Suchergebnisseiten im OTTO Online Shop ableiten?

1.3 Aufbau der Arbeit & methodisches Vorgehen

Diese Masterarbeit ist so aufgebaut, dass in Kapitel 1 die konkrete Problemstellung und Zielsetzung der Arbeit dargestellt werden. Der aktuelle Forschungsstand bzw. die Forschungslücke wird angeführt und eine Forschungsfrage ist definiert, auf die sich diese Arbeit stützt.

Im nachfolgenden zweiten Abschnitt werden die wichtigsten Begrifflichkeiten definiert und verschiedene Modelle bzw. Theorien genauer erläutert. Zu den Definitionen zählen beispielsweise Begriffe wie E-Commerce, die Customer Journey, Personalisierung oder Web Analytics, Big Data und Data Mining. Zudem wird in Kapitel 2 die Clusteranalyse beschrieben. Theorien, auf die sich diese Arbeit bezieht, sind die Reizüberflutung (= Information Overload), Paradox of Choice und das Prinzip der Filter-Bubble. Kapitel 3 umfasst die Literaturanalyse, welche die Verbindung zwischen E-Commerce, den sich verändernden Bedürfnissen der 4. Kunden-Generation, der Onsite-Suche inkl. der Relevanz der Produktfilter auf Suchergebnisseiten und der Personalisierung herstellt. In dem Abschnitt wird insbesondere auch versucht die Forschungslücke literaturseitig zu schließen und verschiedene Ansichten und Blickwinkel zu vereinen.

Nachfolgend geht es über in den empirischen Part der Arbeit. Dabei wird in Kapitel 4 erläutert, wie die Datensammlung im Online Shop durchgeführt und die Filternutzungsdaten ausgewertet wurden. Abschnitt 5 stellt sehr ausführlich die Ergebnisse der verschiedenen untersuchten Filter dar, zudem werden business-seitig relevante Fragen beantwortet. Zu den beleuchteten Filtern zählen der Marken-, Preis-, Größen-, Farb-, Verfügbarkeits- und Geschlechtsfilter. Als führende Methodik der empirischen Forschung zählt neben der Aggregation von Daten insbesondere die Clusteranalyse.

In Kapitel 6 werden konkrete Handlungsempfehlungen abgegeben, die die Markenverantwortlichen von OTTO Österreich umsetzen können. Außerdem wird ein Fazit gezogen und die Limitationen der Arbeit sowie ein Ausblick gegeben.

2. Theoretische Grundlagen

2.1 Definitionen und Begriffserklärungen

2.1.1 E-Commerce

Unter E-Commerce wird nach einer allgemein anerkannten Definition der Austausch von Geschäftsinformationen, die Pflege von Geschäftsbeziehungen und die Abwicklung von Geschäftsvorgängen über Telekommunikationsnetze verstanden (Zwass, 1996). Das ist aber nur einer von vielen Versuchen, E-Commerce zu definieren. Wirtschaftswissenschaftliche Beiträge, die das vergleichsweise junge Themenfeld in der Betriebswirtschaft behandeln, beleuchten unterschiedliche Charakteristika des Begriffs E-Commerce (Schmeken, 2008, S. 11). Verschiedene Forscher*innen stellen dabei Eigenschaften wie eine neue digitale Marktform, digitale Transaktion, elektronische Verbindung oder elektronische Abwicklung in den Fokus (Adam & Yesha, 1996; Kotler & Bliemel, 2001; Zwass, 1996).

Der elektronische Handel bzw. E-Commerce umfasst Kauf- und Verkaufsprozesse, wobei elektronische Mittel – hauptsächlich das Internet – die Prozesse unterstützen (Kotler, Armstrong, Harris, Piercy, 2010, S. 963). Bei den vier zentralen Bereichen des E-Commerce handelt es sich laut Kotler et al. (2010, S. 963) um Business-to-Consumer (B2C), Business-to-Business (B2B), Consumer-to-Consumer (C2C) und Consumer-to-Business (C2B).

Im Rahmen dieser Arbeit wird der Fokus ausschließlich auf den elektronischen Handel im B2C-Bereich gelegt, dem Online-Verkauf von Waren an die Endverbraucher.

2.1.2 Customer Journey und Touchpoints

Customer Journey bedeutet wörtlich übersetzt „Die Reise des Kunden“ und beschreibt im E-Commerce Kontext als Metapher den Prozess der Kaufentscheidung von potentiellen Kund*innen über verschiedene Kontaktpunkte (Touchpoints) mit

einem Online Shop (Holland & Flocke, 2014, S. 827). Der Bundesverband Digitale Wirtschaft (BVDW) definiert die Customer Journey als *„alle messbaren Kontaktpunkte eines Nutzers auf dem Weg zu einer definierten Aktion. Hierbei werden alle Marketingkanäle berücksichtigt, mit denen ein Konsument im Rahmen dieser Aktion in Berührung kommt, wobei sowohl Sicht- als auch Klickkontakte einbezogen werden.“* (Bundesverband Digitale Wirtschaft (BVDW) e. V., 2012).

In Anbetracht dieser Definitionen bezieht sich die digitale Customer Journey daher auf die Erfahrung eines Menschen mit der Interaktion einer Website, eines Online Shops oder einer anderen digitalen Plattform. Diese geht von verschiedenen Touchpoints wie der ersten Suche nach einer Brand, über den ersten Seitenaufruf, Klicks, Sessions, Einsprünge und Transaktionen aus (Wright, 2019).

Unter Touchpoints versteht man Berührungspunkte zwischen Nachfragern und Anbietern, sie stellen die Basis für die Kommunikation sowie den Aufbau einer Handlung dar (Kawohl, 2017). Vor Einzug der digitalen Entwicklung waren Offline-Kanäle wie TV- oder Print-Werbung die wichtigsten Kontaktpunkte, durch die Digitalisierung fand jedoch eine Transformation des Handels in die Welt des Internets statt (Esch & Knörle, 2015, S.132-135).

Abbildung 1 von Kawohl (2017) stellt den Vergleich von Touchpoints der klassischen (offline) und der modernen (online) Customer Journey mit den Phasen „Consideration“, „Evaluation“, „Moment of Purchase“ und „Post Purchase“ dar.

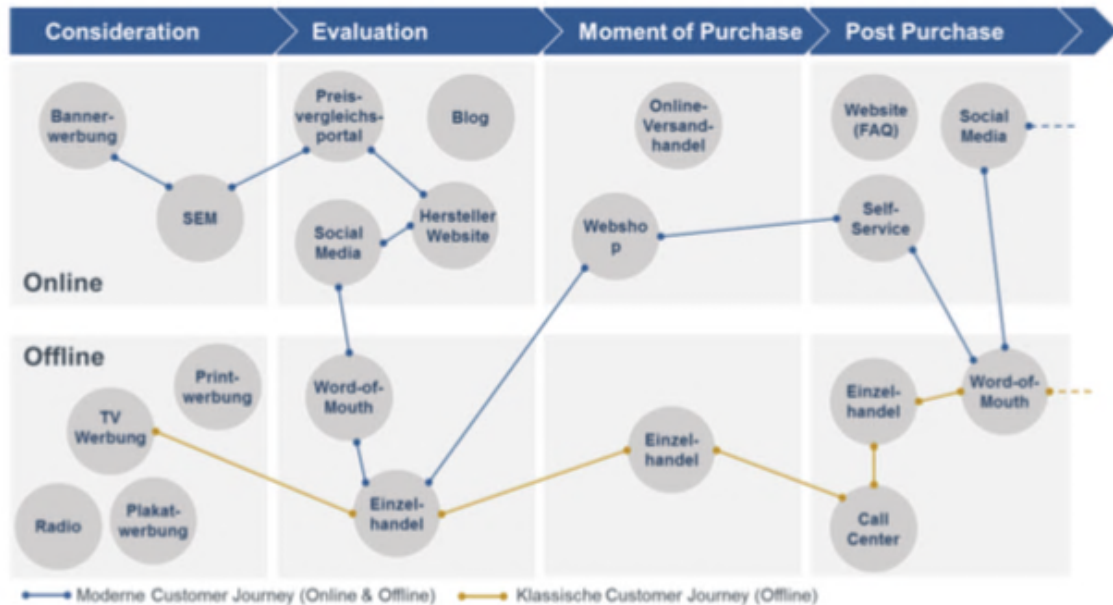


Abbildung 1: Vergleich der modernen und der klassischen Customer Journey.
Quelle: Kawohl, 2017

2.1.3 Personalisierung

Personalisierung ist ein strategisches Instrument zur Differenzierung von Produkten und Dienstleistungen, insbesondere dann, wenn der Wettbewerb am Markt stark umkämpft ist (Kwon & Kim, 2012, S. 101). Bereits 1997 beschrieben Peppers & Rogers Personalisierung – oder das 1:1-Marketing – als den Prozess der Verwendung von Kundeninformationen, um für die Kund*in eine gezielte Lösung bzw. Angebot bereitzustellen. Die Grundidee der Personalisierung besteht also darin, jede Kund*in als ein eigenes Individuum zu behandeln und maßgeschneiderte Empfehlungsnachrichten basierend auf den individuellen Präferenzen zu gestalten (Kalyanaraman & Sundar, 2006).

Andere Forscher sehen Personalisierung als den strategischen Punkt, an dem ein Unternehmen basierend auf zuvor gesammelten Kundendaten entscheidet, welcher Marketing-Mix für das Individuum am besten geeignet ist (Arora et al., 2008; Murthi & Sarkar, 2003). Einen weiteren interessanten Ansatz liefert Bruce Kasanoff (2009) in einem Interview mit Frank T. Piller (2009): „*Personal = Smarter. [...] the more a company customizes, the smarter it becomes. The smarter it gets, the more formidable*

a competitor it becomes.“ Folgt man Piller’s Definition, nutzt Personalisierung Technologie, um die Unterschiede zwischen Menschen auszugleichen. Richtig gemacht, ist es eine Win-Win-Strategie, wodurch sowohl Unternehmen als auch die beteiligten Individuen ein besseres Ergebnis erzielen (Piller, 2009). Denn so wie viele verschiedene Ansätze unterschiedlicher Definitionen von Personalisierung in der Literatur zu finden sind, wurden die Dimensionen der Personalisierung auf ihre interdisziplinären Merkmale in verschiedenen akademischen Bereichen wie Wirtschaft, Management, Marketing, Informationssysteme und Informatik untersucht. Das Finden einer einheitlichen Definition über die verschiedensten Forschungsbereiche hinweg, ist so schier unmöglich (Kwon & Kim, 2012; Stüber, 2013).

2.1.4 Onsite Search und Filter

Die Onsite Suche beschreibt das Suchfeld, über welches die meisten Websites verfügen. Das Ziel des Onsite-Suchschlitzes ist, den Besucher*innen bestimmte Inhalte anhand von eingegebenen Textzeichenfolgen auf der Website oder im Online Shop zu zeigen, die sie durch die Sucheingabe gesucht haben (A. Greco, 2012).

Die Suche stellt eine der relevantesten Shop-Funktionen dar. Denn durch die Onsite-Suche sollen Nutzende – abhängig von den individuellen Bedürfnissen – möglichst effizient die relevanten und weiterführenden Produkte bzw. Inhalte im Online Shop finden (Vogt & Quelle-Korting, 2017, S. 15). Die Onsite Search ist im analogen Handel vergleichbar mit dem Personal im stationären Shop, den die Kundschaft nach bestimmten Produkten und Merkmalen fragen würde (Graf & Schneider, 2016). Durch die Unterstützung im Selektionsprozess ist die Suche im Webshop maßgeblich für das Finden der relevanten Produkte auf effiziente Art und Weise verantwortlich und hat laut Vogt & Quelle-Korting (2017) einen großen Einfluss auf Umsatz- und Ergebnisströme eines Online Shops.

Die entscheidende Rolle, die die Suchfunktion spielt, wird auch durch die Positionierung im Shop deutlich. Am Beispiel von Amazon (Abbildung 2) ändert der

Suchschlitz je nach Device und Auflösung dynamisch die Größe. Zudem ist die Onsite Search durch die Platzierung im Header auf beinahe jedem Template im Shop nutzbar. (Graf & Schneider, 2016)

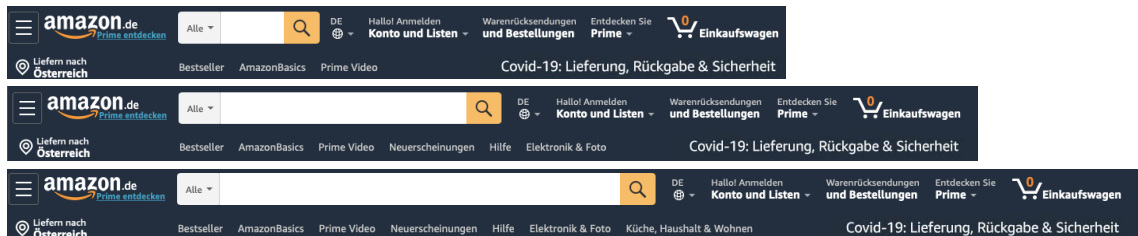


Abbildung 2: Dynamische Änderung des Suchfelds je nach Browser-Breite.
Quelle: amazon.de, 2020

Je generischer die Suchanfrage im Online Shop gestellt wird, desto größer ist der Recall, sprich die Anzahl der gefundenen Produkte (Scholz, 2019). Dabei stellen Filter eine beliebte Funktion im E-Commerce Umfeld dar, um Produktlisten und Suchergebnisse auf die persönlichen Bedürfnisse anzupassen und einzugrenzen (Graf & Schneider, 2016). Durch das Auswählen und Setzen von Filter können die Ergebnisse genau auf die Bedürfnisse des Nutzenden angepasst werden (Schneider & Graf, 2017) Abbildung 3 stellt den Suchschlitz – eingebettet im Header – und verschiedene Filtermöglichkeiten anhand des Suchbegriffs „Ballerina“ dar.

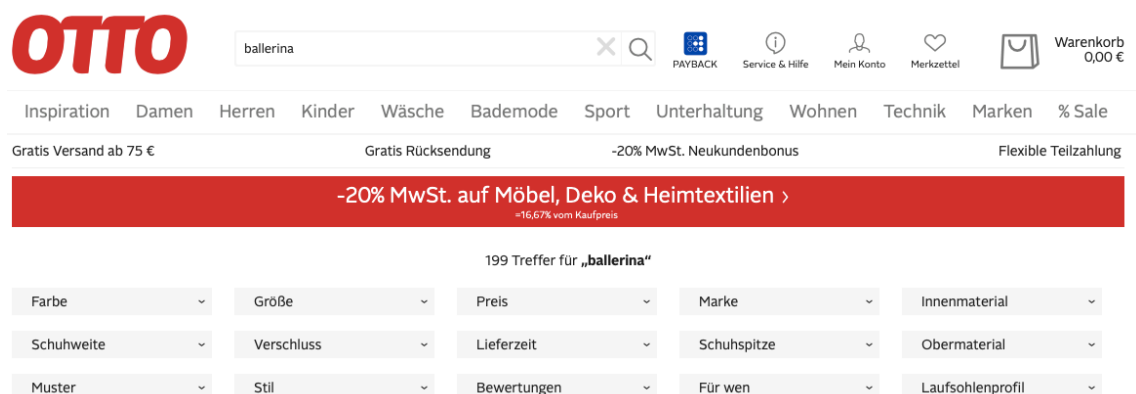


Abbildung 3: Such- & Filterfunktion bei ottoversand.at am Beispiel des Suchworts "Ballerina".
Quelle: ottoversand.at, 2020

2.1.5 Web Analytics

Im Web besteht im Gegensatz zu allen anderen Medien die Möglichkeit, jede einzelne User*in ganz genau zu beobachten und das Verhalten zu analysieren (Hassler, 2011, S. 27). Die Disziplin, in der jeder Klick, jede Abfolge von Seitenaufrufen, jeder Kaufabschluss und alle weiteren Interaktionen mit Webseiten und Online Shops erschließt, nennt sich laut Hassler (2011, S. 27) Web Analytics. Anhand dieser Disziplin können Unternehmen lernen und interpretieren, wie Kund*innen mit dem Online Shop oder Apps interagieren (Beasley, 2013).

Meier & Zumstein (2012, S. 7) beschreiben den Begriff Web Analytics als die *„Evaluation, Definition, Messung und Auswertung von Webkennzahlen, um die Inhaltsnutzung der Website und das Benutzerverhalten analysieren zu können“*.

Um das Verhalten von Benutzer*innen im Online Shop oder auf der Webseite messen, analysieren und auswerten zu können, müssen Leistungs- bzw. Erfolgskennzahlen – sogenannte Key Performance Indikatoren (KPIs) – vom Unternehmen festgelegt werden. Nur so ist es möglich, die Erreichung strategischer Ziele zu messen (Zumstein & Gächter, 2016, S. 378).

Web-KPIs können im E-Commerce-Umfeld beispielsweise sein (Zumstein & Gächter, 2016, S. 384):

- Impressionen/Views
- Klicks bzw. Click Through Rate (CTR)
- Conversion/Kaufabschlüsse bzw. Conversion Rate (CVR)
- Warenkorbwert
- Anzahl der Produktseitenaufrufe
- Add-to-Basket Rate
- Filternutzungsrate auf Produktlisten
- ...

2.1.6 Big Data

Big Data beschreibt im Wesentlichen, dass Datensätze, die bisher für herkömmliche Datenverarbeitungssysteme zu groß sind, neue Verarbeitungstechnologien erfordern (Fawcett & Provost, 2013, S. 7). Da es laut De Mauro, Greco & Grimaldi (2016, S. 126) keine einvernehmlich anerkannte Definition von Big Data gibt, haben diverse Forschende implizite Definitionen durch Anekdoten, Erfolgsgeschichten, technologische Merkmale, Trends und deren Auswirkungen auf Gesellschaft, Unternehmen und Geschäftsprozesse übernommen. Aus diesem Grund haben De Mauro et al. (2016) in ihrer Arbeit basierend auf den bestehenden Big-Data-Definitionen und den wichtigsten Forschungsthemen analysiert, welche Aspekte der Kern des Big-Data-Konzepts umfasst:

- Volumen (Volume), Geschwindigkeit (Velocity) und Vielfalt (Variety), um die Merkmale von Informationen zu beschreiben;
- Technologie und Analysemethoden, um die Anforderungen zu beschreiben, die für die ordnungsgemäße Nutzung dieser Informationen erforderlich sind;
- Den Wert (Value), um die Umwandlung von Informationen in Erkenntnisse zu beschreiben, die wirtschaftliche Auswirkungen haben können.

Die Forschenden empfehlen aufgrund ihres Forschungsergebnisses folgende einvernehmliche Definition von Big Data: *„Big Data is the Information asset characterized by such a High Volume, Velocity and Variety to require specific Technology and Analytical Methods for its transformation into Value.“* (De Mauro et al., 2016, S. 127)

Big-Data-Technologien werden für viele Aufgaben in vielen verschiedenen Bereichen und Branchen eingesetzt: Gesundheitswesen, Biowissenschaften, Medien- und Unterhaltungselektronik, Logistik, Telekommunikation oder auch im Handel (Villars, Olofson & Eastwood, 2011).

2.1.7 Data Mining

Data Mining definiert das Identifizieren einer interessanten Struktur von Daten (Fayyad & Uthurusamy, 2002, S. 28). Laut Grabmeier & Rudolph (2002, S. 303) beschreibt der Begriff alle Methoden und Techniken, mit denen sehr große Datenmengen analysiert werden können, um bisher unbekannte Strukturen und Beziehungen aus riesigen Informations-Bergen zu extrahieren und zu entdecken. Diese Informationen werden gefiltert, aufbereitet und klassifiziert, sodass sie eine wertvolle Hilfe für Entscheidungen und Strategien darstellen (Grabmeier & Rudolph, 2002).

Data-Mining-Techniken werden für einen ganz bestimmten Zweck verwendet – normalerweise, um Einblicke und Verbesserungen in Geschäftsfunktionen zu erhalten (de Ville, 2001, S. 2). Sie können wesentlich zu einer datengetriebenen Entscheidungsfindung (Data-driven Decision-making) beitragen (Fawcett & Provost, 2013, S. 2), indem beispielsweise Verhaltensmuster (Behavior Patterns) extrahiert bzw. erkannt werden oder das Kundenverhalten aus großen Datenbanken prognostiziert wird (Fan, Lau & Zhao, 2015, S. 29). Einen wesentlichen Beitrag der Extrahierung von verborgenem „nützlichen“ Wissen in Form von Datensätzen leistet die Segmentierung. Dabei unterteilt sie Daten in kleinere, homogene und praktisch zweckmäßige Teilmengen (Grimmer & Mucha, 2013, S. 109).

2.1.8 Clusteranalyse

Die Clusteranalyse (Lloyd, 1982; MacQueen, 1967) fasst Verfahren aus dem Bereich der multivariaten Statistik zusammen, und setzt voraus, dass eine untersuchte Stichprobe eine bestimmte Gruppenstruktur aufweist (Wiedenbeck & Züll, 2010, S. 525). Diese Struktur ermöglicht, dass die – im allgemeinen ungeordnete und umfangreiche – Stichprobe automatisiert und objektiv in kleinere, homogene Teilmengen bzw. Cluster eingeteilt wird (Grimmer & Mucha, 2013, S. 109). Die Einheiten der Cluster sind innerhalb der Gruppen sehr ähnlich zueinander, weisen aber zwischen verschiedenen Cluster deutliche Unterschiede auf (Wiedenbeck &

Züll, 2010, S. 525). Im Zuge dieser Masterarbeit wurde die Visualisierungs-Software *Tableau* verwendet, welche den k-Means-Algorithmus für das Clustering verwendet. Daher wird dieser nachfolgend genauer beleuchtet.

Im Allgemeinen wird der k-Means-Algorithmus eingesetzt, um Datenpunkte eines Datensets basierend auf den nächsten Mittelwerten in Cluster oder Klassen zu unterteilen, wobei das k die Anzahl an Cluster in den Daten beschreibt. Diese Menge an Klassen wird im Vorfeld vom Untersuchenden angegeben (Pierson, 2015, S. 80). Das *Means* beschreibt den Mittelwert – in Form der arithmetischen Mittel der Werte entlang jeder Dimension für die Einheiten in den jeweiligen Clustern – und wird durch den Zentroid bzw. den Clusterschwerpunkt dargestellt, wobei dieser Schwerpunkt der Durchschnitt aller Einheiten je Cluster ist (Fawcett & Provost, 2013, S. 169).

Der Algorithmus beginnt mit der Erstellung von k anfänglichen Clusterzentren, welche zufällig gewählt oder durch die Angabe spezifischer Datenpunkte bestimmt werden können (Arthur & Vassilvitskii, 2007). Die Zentroide je Cluster werden gebildet, indem bestimmt wird, welches Zentrum jedem Punkt am nächsten liegt (Fawcett & Provost, 2013, S. 171). Als nächstes wird für jeden dieser Cluster das Clusterzentrum neu berechnet, indem der tatsächliche Schwerpunkt basierend auf jedem einzelnen Datenpunkt im Cluster ermittelt wird. Dadurch verschieben sich die Schwerpunkte normalerweise. Dieser Vorgang iteriert so lange, bis sich die Cluster-Zentren nicht mehr ändern (Fawcett & Provost, 2013, S.171).

Abbildung 4 zeigt einen Beispiellauf von $k = 3$ Zentroiden mit 90 Dateneinheiten. Dabei werden im oberen Streudiagramm die anfänglichen Datenpunkte vom Clustering gezeigt. Im unteren Diagramm sind die Endergebnisse des Algorithmus nach 16 Iterationen dargestellt. Bei jeder Wiederholung wird neu berechnet, welche Punkte zu welchem Cluster gehören. Sobald diese neu zugewiesen sind, werden ggf. auch die Zentroide erneut verschoben. Die drei (unregelmäßigen) Linien zeigen den Pfad von der anfänglichen (zufälligen) Position jedes Schwerpunkts hin zu seiner

endgültigen Position. Die Datenpunkte in den drei Clustern werden durch verschiedene Symbole gekennzeichnet. (Fawcett & Provost, 2013, S. 172)

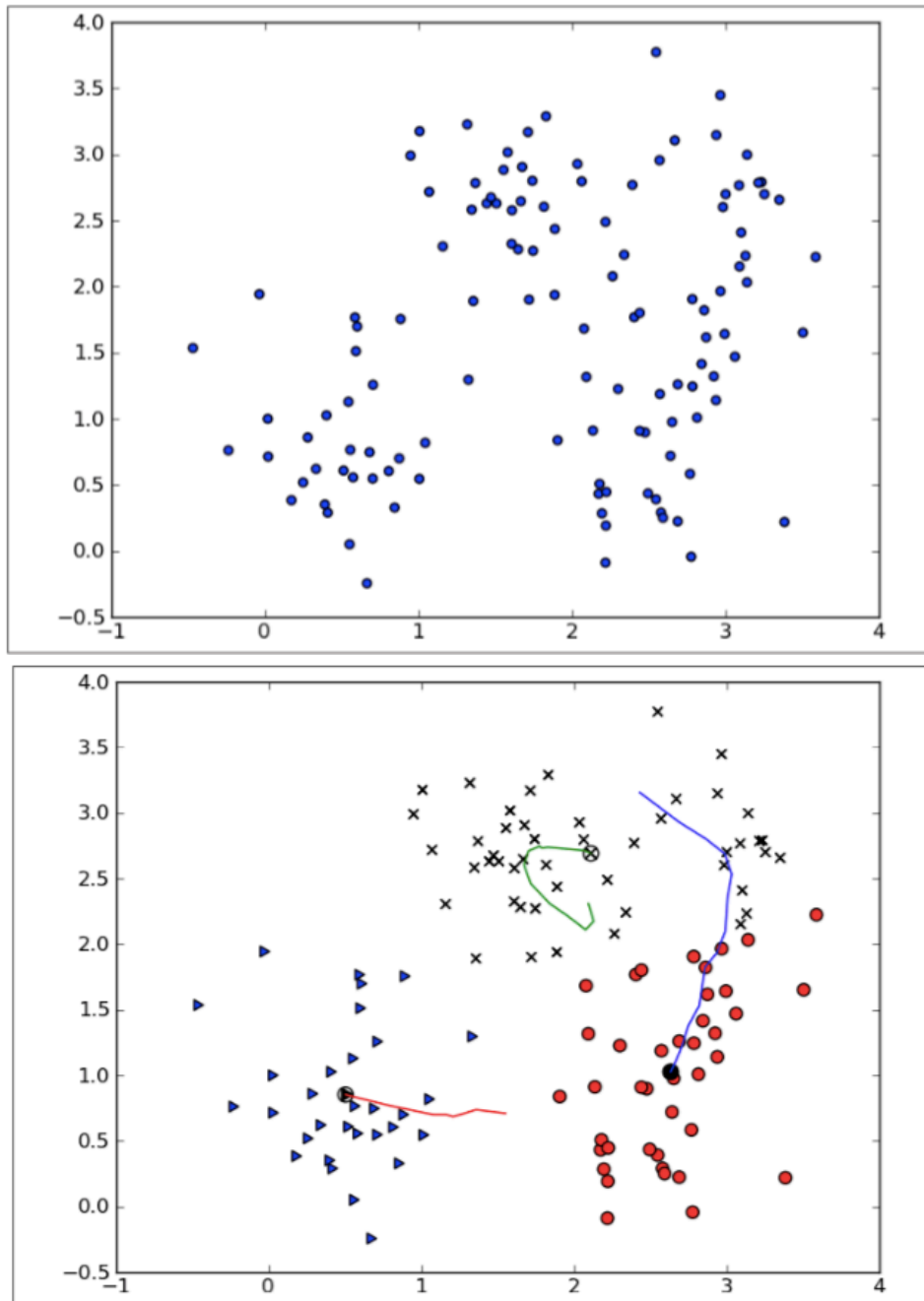


Abbildung 4: Ein k -Means-Clustering mit 90 Punkten in einer Ebene und $k=3$ Zentroiden. Oben wird der anfängliche Satz an Punkten dargestellt, unten sind die Bewegungspfade der Zentroiden zu sehen.

Quelle: Fawcett & Provost, 2013, S. 173

2.2 Reizüberflutung

Durch das Internet und insbesondere auch im Bereich des Online Shoppings sind Kund*innen einer Vielzahl an Produktinformationen ausgesetzt (Gross, 2014). Zwar tragen laut Alba et al. (1997, S. 42) reichlich vorhandene Informationen positiv zur Verbesserung der Einkaufsentscheidung bei. Basierend auf der Theorie der Informationsüberflutung müssen Verbraucher jedoch möglicherweise mehr Anstrengungen aufwenden und können schlechter Entscheidungen treffen, wenn die Last der Informationen über einen Schwellwert hinaus steigt (Jacoby, Speller & Kohn, 1974). Denn die Fähigkeit des Menschen, Informationen während einer bestimmten Zeiteinheit aufzunehmen und zu verarbeiten, hat Grenzen. Werden diese Grenzen überschritten, neigen die Menschen dazu, verwirrt und dysfunktional zu werden (Jacoby et al., 1974, S. 33).

Frühere Forscher wollten zwar eine universelle Schwelle für die Informationslast finden, an der sich die Entscheidung der Verbraucher verschlechtert (Chen, Shang & Kao, 2009, S. 48). Konsument*innen unterscheiden sich jedoch in ihrer Fähigkeit zur Informationsverarbeitung (Henry, 1980), daher kann auch dieser Schwellwert zur Überlastung von Person zu Person variieren (Chen et al., 2009, S. 48). Denn die menschliche Entscheidungsfindung ist ein hoch komplexer und dynamischer Prozess, der sehr stark von den Erfahrungen einer Person in der Vergangenheit beeinflusst wird (Tversky, 1972). Diese Erfahrungen können eine Art „interner Filtermechanismus“ sein, der sich darauf auswirken kann, welchen Informationen sich diese Person aussetzt und von ihr selektiv wahrgenommen werden (McGuire, 1976).

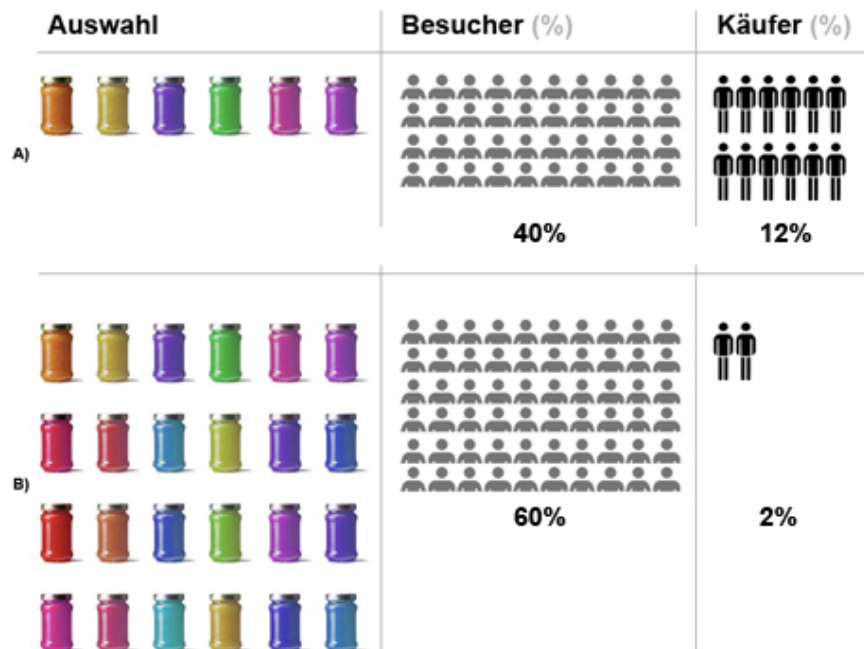
Auch für den Bereich des Online Einkaufserlebnisses gilt, dass User*innen einen internen Mechanismus zum Filtern irrelevanter Informationen unbewusst nutzen, um Informationen wahrzunehmen und zu selektieren (Chen et al., 2009, S. 50). Und genau dieses Unbewusste schützt Kund*innen auch im Online Shop vor einer Reizüberflutung, denn es werden nicht mehr wie in einem Trichtermodell die Anzahl der Informationen (z.B. Marken, Produkte, Sortimente etc.) reduziert, sondern

durch den kognitiven Verarbeitungsprozess direkt begrenzt (Menthe & Sieg, 2017, S. 9). Da in Online Shops typischerweise eine gewisse Vielzahl an Produkten mit unterschiedlichen Produkteigenschaften angeboten werden – bei ottoversand.at sind es mehr als 400.000 unterschiedliche Artikel – ist die Gefahr eines Information Overloads bei Shop-Besucher*innen groß (Groissberger & Riedl, 2018). Um Unzufriedenheit und zu niedrige Abschlussraten seitens der Online-Kundschaften zu verringern, sind beispielsweise Shop-Funktionen, um das Produkt-Angebot zu filtern, Sortieren oder Vergleichen ein hilfreiches Werkzeug. So helfen laut Groissberger & Riedl (2018, S. 95) Produktfilter einer potentiellen Kund*in dabei, die Auswahl durch das Festlegen gewünschter Grenzwerte der wichtigsten Attribute einzuschränken.

2.3 Paradox of Choice

Ein weiteres Phänomen, dass sich auf die Reizüberflutung und auf die Qualität der getroffenen Entscheidung bezieht, ist das Auswahl-Paradoxon. Je mehr Auswahl, desto besser – das ist in der modernen Gesellschaft eine weit verbreitete Annahme (Iyengar & Lepper, 2000). Die Forscherin Iyengar und der Forscher Lepper führten basierend auf dieser Hypothese im Jahr 2000 mehrere spannende Experimente in Kalifornien durch und können sie widerlegen. Im Rahmen eines dieser Experimente konnten Kund*innen in einem Delikatessengeschäft Toastscheiben, bestrichen mit verschiedenen Marmeladesorten, kosten und wenn es schmeckte, die jeweilige Sorte auch kaufen. Dabei entwickelten die beiden Forschenden eine Art A/B-Testsituation. Bei Variante A wurden den Interessierten lediglich sechs Marmeladesorten angeboten, bei Variante B 24 Sorten. Das Ergebnis war im ersten Moment erwartungskonform: die kleinere Auswahl lockte nur 40 % aller Besucher*innen an den Probiertisch; die große Auswahl mit 24 unterschiedlichen Sorten lockte 60 % an den Stand – es scheint sich die Annahme zu bestätigen, dass eine große Auswahl von mehr Menschen bevorzugt wird. Überrascht hat allerdings das Endergebnis, dass bei der Variante B (24 Marmeladensorten) nur 2 % der Besuchenden auch zuschlugen und eine Sorte kauften. Bei der kleinen Auswahl mit

sechs verschiedenen Marmeladensorten griffen allerdings 12 % der Experiment-Teilnehmer zum Glas und kauften ein (Abbildung 5).



*Abbildung 5: Das Konfitüren-Experiment von Iyengar & Lepper.
Quelle: Henrici, 2012*

Schwartz (2005) bezeichnet diese Phänomen als das Auswahl-Paradoxon bzw. Paradox of Choice. Dieses besagt, dass je mehr Optionen angeboten werden, desto schlechter die getroffenen Entscheidungen sind. Eine hohe Anzahl an Wahlmöglichkeiten erhöht nicht die Kapazitäten, die unser Gehirn für Entscheidungen zur Verfügung stellt, sondern führt zu Unsicherheit und Angst (Schwartz, 2005).

Der Effekt des Paradox of Choice hat mehrere Gründe, wie im Buch PsyConversion von Spreer (2018) dargestellt wird. Zum einen ist es so, dass bei vielen Alternativen, die zur Verfügung stehen, auch großer Aufwand für das Bewerten und Abwägen investiert werden muss. Das steht im Konflikt mit dem Gehirn, welches so programmiert ist, dass mit Heuristiken Aufwände reduziert werden. Viele

Wahlmöglichkeiten blasen zudem die Erwartungen gegenüber dem Produkt auf, was dazu führt, dass Konsument*innen im Nachhinein nicht zufrieden sein werden.

Im Ökonomischen Sinne bedeutet das Auswahlparadox für Unternehmen, dass sie versuchen sollten, den Kundschaften den Decision-Making-Prozess so einfach wie möglich zu gestalten und im Zweifelsfall weniger statt mehr Auswahlmöglichkeiten bieten (Bauer & Koth, 2014).

2.4 Filter-Bubble

Pariser (2011) prägte den Begriff „Filter-Bubble“ und beschreibt in seinem gleichnamigen Buch, wie Personalisierung den ursprünglichen Zweck des Internets als offene Plattform für die Verbreitung von Ideen untergräbt und uns alle in einer isolierten, hallenden Welt zurücklassen könnte.

Die meisten User*innen gehen nämlich davon aus, dass beim Googlen eines Begriffs jeder dieselben Ergebnisse sieht – diejenigen, die Google vorschlägt, sind aufgrund verschiedener Suchmaschinenoptimierungs-Faktoren die relevantesten (Pariser, 2011b). Seit Dezember 2009 ist das allerdings nicht mehr der Fall. Pariser (2011) bezieht sich dabei in seinem Buch auf einen Post von Google, der am 4.12.2009 auf dem Unternehmens-Blog veröffentlicht wurde: „Personalized Search for everyone“ (Official Google Blog, 2009). Dabei verlautbarte Google, dass Suchergebnisse für jede Nutzer*in individualisiert werden. Anstatt ihnen das allgemein beliebteste Ergebnis zu liefern, versucht Google seit 2009 vorherzusagen, auf welches Suchergebnis die User*innen am wahrscheinlichsten klicken werden. Jetzt erhalten Search Engine User das Ergebnis, das der Google-Algorithmus für sie als am besten vorschlägt – und jemand anders sieht möglicherweise etwas ganz anderes. Laut Pariser (2011) gibt es kein Standard-Google mehr.

Er führt in seinem Buch ein einfaches Beispiel an, um die Auswirkungen der Filter-Bubble zu veranschaulichen. Und zwar bat er im Frühjahr 2010, während die Überreste der Bohrinself Deepwater Horizon (BP) Öl in den Golf von Mexico spuckte,

zwei Freundinnen, nach dem Suchbegriff „BP“ zu googlen. Die Ergebnisse, die sie sahen, waren ganz unterschiedlich (Abbildung 6). Die eine erhielt Investitionsinformationen über BP als Suchergebnis, die andere sah hauptsächlich Nachrichten und Neuigkeiten. Sogar die Anzahl der zurückgegebenen Ergebnisse war unterschiedlich. 180 Millionen für die eine Freundin, 139 Millionen für die andere.



Abbildung 6: Auswirkungen der Filter-Bubble anhand des Suchbegriffs „BP“ im Jahr 2010. Quelle: Storey, 2011

Mit dem Phänomen der Filter-Bubble versucht Pariser (2011) zu beschreiben, wie durch das Potential der Personalisierung Internet-User*innen effektiv von einer Vielzahl an Gesichtspunkten oder Inhalten isoliert werden können (Nguyen, Hui, Harper, Terveen & Konstan, 2014, S. 677). Denn im digitalen Zeitalter treffen inzwischen immer häufiger Algorithmen Entscheidungen, sie werden zu den neuen Gatekeepern in der Gesellschaft (Bozdog, 2013, S. 209). Diese Erkenntnis lässt sich nicht nur am Hand von beispielsweise dem BP-Google-Beispiel erklären. Auch Online Shops oder Webseiten sammeln fortlaufend Informationen und Daten über das User-Verhalten, wie Klicks oder Kaufabschlüsse. Eine systematische Auswertung dieser Daten veranschaulicht individuelle Vorlieben und ermöglicht sehr zielgerichtete Artikel- oder Serviceempfehlungen (Klug & Strang, 2018).

3. Hermeneutische Forschung

3.1 Die Entwicklung des E-Commerce hin zum Ubiquitous Commerce

Als Richard T. Watson (2000) über Geschäfte, die in einer Welt gemacht werden, welche allgegenwärtig, universell, einzigartig und unisono ist, geschrieben hat, ahnte er wohl nicht, dass 20 Jahre später der Begriff „Ubiquitous Computing“ – die allgegenwärtige Datenverarbeitung – Teil der gängigen Volkssprache sein wird (Krumm, 2018). Mit der Weiterentwicklung des E-Commerce haben die Fortschritte in der Informations- & Kommunikationstechnologie (IKT) eine aufkommende neue Art des Handelns beschleunigt. Durch Funktionen, die die Grenzen von Raum und Zeit überwinden, wird Mobile Commerce (M-Commerce) von Ubiquitous Commerce (U-Commerce) abgelöst (Junglas & Watson, 2003, S. 674).

Ubiquitous Computing hat laut Junglas & Watson (2003) ein neues Paradigma des Handelns ermöglicht, das über den traditionellen Handel hinausgeht. Diese Art und Weise des Geschäftemachens beschreibt die Fähigkeit mit „anything and anyone, anytime and anywhere“ zu interagieren und zu handeln (Richard T. Watson, 2000). Watson spricht dabei von der ultimativen Form des Handelsbetriebes. Dieser allgegenwärtige Handel tritt als kontinuierlicher, nahtloser Strom, bestehend aus Kommunikation, Content und Dienstleistungen auf, der zwischen Unternehmen, Lieferant*innen, Mitarbeiter*innen, Kund*innen und Produkten ausgetauscht wird (Kumar, Joshi & Saquib, 2015, S. 50). Franco, Rosa, Barbosa, Costa & Yamin (2011) beschreiben U-Commerce als die Integration von E-Commerce durch die elektronische Identifizierung von physischen Produkten, M-Commerce durch die Möglichkeit, überall und jederzeit einkaufen zu können und ubiquitäres Computing, indem Benutzer*innen mit Hilfe einer smarten Umgebung intelligent und intuitiv einkaufen können.

Junglas & Watson (2003, S. 670) identifizierten vier Faktoren, die den Übergang zwischen M-Commerce und U-Commerce beschreiben: mobile Anwendungen, mobile

Netzwerke, mobile Geräte und Datensynchronisierung. Dem Begriff U-Commerce werden neben „Ubiquity“ weitere Charakteristika zugewiesen: „Universality“, „Uniqueness“ und „Unison“. Diese vier übergeordneten Konstrukte setzen sich laut Junglas & Watson (2003, S. 671) wie folgt aus den oben beschriebenen Dimensionen zusammen:

- Ubiquity = äußerste Form von (Erreichbarkeit + Zugänglichkeit + Übertragbarkeit)
- Uniqueness = äußerste Form von (Lokalisierung + Identifikation + Übertragbarkeit)
- Universality = äußerste Verschmelzung von (mobilen Netzwerken + mobilen Geräten)
- Unison = äußerste Verschmelzung von (mobilen Anwendungen + Datensynchronisierung)

a) Ubiquity: Dies bedeutet, dass Computer überall vorhanden sind und jedes Gerät mit dem Internet verbunden ist. Diese Allgegenwart von Computerchips macht sie „unsichtbar“, da die Menschen sie nicht mehr bemerken werden (Watson, Pitt, Berthon & Zinkhan, 2002, S. 332).

b) Universality: Die Universalität beschreibt die Möglichkeit, in heterogenen Umgebungen arbeiten zu können. Die persönliche Sammlung an Geräten & Devices, die Individuen tagtäglich nutzen, ist universell einsetzbar und multifunktional. Aufgrund des Internets und der Satelliten stehen Desktop, Laptop oder Smartphone freie Mobilität und eine riesige Menge an Informationen zur Verfügung (Kumar et al., 2015, S. 51).

c) Uniqueness: Junglas & Watson (2003, S. 672) definieren die Einzigartigkeit als das Bestreben, die Merkmale und den Ort einer Person oder eines Dings genau zu kennen. Einzigartig bedeutet in diesem Sinn, dass die Informationen, die

User*innen zur Verfügung gestellt werden, leicht an ihren aktuellen Kontext und die besonderen Bedürfnisse jeder Person individuell angepasst werden können, beispielsweise regionale Nachrichtendienste oder Wetterberichte über den aktuellen Standort eines Devices.

d) Unison: Integrierte Daten werden über mehrere Anwendungen hinweg abgedeckt, sodass Benutzer*innen unabhängig von dem verwendeten Gerät eine konsistente Sicht auf ihre Informationen haben. Konsistenz bedeutet, dass beispielsweise Änderungen eines Kontakts im Smartphone auch gleichzeitig eine Änderung im Laptop und allen anderen vom User verwalteten elektronischen Telefonbüchern übernommen wird (Junglas & Watson, 2003, S. 673).

3.2 Der Kunde 4.0 im Jahr 2020

Das Zeitalter des Ubiquitous Commerce bringt massive Auswirkungen mit sich, indem Menschen ständig miteinander verbunden sind – jederzeit, überall und irgendwie. Jede Kund*in kann kontinuierlich mit Freund*innen, Familie, Kolleg*innen, Teams, Lieferant*innen oder anderen Kund*innen interagieren, während er sich über ein Produkt informiert, es auswählt oder kauft (Freeman, 2017). Zudem verändern die digitalen Möglichkeiten die Anforderungen bzw. Kundenwünsche, Kund*innen lassen sich nicht mehr kategorisieren – Max Mustermann hat quasi ausgedient (Weissman & Wegerer, 2018).

Diese gerade aufkeimende vierte Generation des Commerce, in der die Kund*in wirklich die Kontrolle hat, wird angeführt von einer jüngeren und einflussreichen Generationen (Freeman, 2017). Diese ökonomische Welt wird laut dem Blogpost von Freeman (2017) bestimmt von schnelllebigen, mobilen, digital informierten und emotional engagierten Online-Erlebnissen – rund um die Uhr. Eine Welt, in der die C2C-Kommunikation den Einfluss der B2C-Kommunikation bei weitem übertrifft. Eine Welt, in der sowohl positive als auch negative Promotoren von Unternehmen die öffentliche Wahrnehmung einer Marke, eines Produkts oder einer Dienstleistung bestimmen (Freeman, 2017). Es verändert sich also nicht der Markt, sondern die

Kundschaft. Ganz ähnlich wie bei Industrie 4.0 – wo es weg von Masse-Standard-Produkten hin zu Losgröße 1 geht – erwartet auch die Kund*in 4.0 genau auf sie zugeschnittene Angebote (Weissman & Wegerer, 2018). Um den Kunden 4.0 verstehen zu können, sollte nochmal ein Blick auf die früheren Generationen geworfen werden, die Freeman (2017) wie folgt zusammenfasst:

- Kunde 1.0 (vor 1950): Kund*innen profitierten von der Produktions- oder Lieferfähigkeit des Unternehmens. Sie hatten wenig Auswahl und kauften Produkte, die verfügbar waren und die besten, die zu dieser Zeit erhalten werden konnten.
- Kunde 2.0 (1950-2000): Die Prinzipien des Kunden 1.0 gelten noch immer, sie werden jedoch unterstützt durch die Entwicklung von Konsum und Marketing. Die Konzepte der Marke tauchen erstmals auf und die Nachfrage der Kund*innen nach Produkten und Dienstleistungen wurden stark vom Marketing getrieben.
- Kunde 3.0 (2000-2015): Die Prinzipien des Kunden 1.0 und 2.0 werden fortgesetzt, durch die explosionsartigen Ausbreitung der Globalisierung und des Internets erweitert. Kundenbindung und Loyalität sind Schlüsselmerkmale, der Schwerpunkt des Wettbewerbs liegt auf der Bereitstellung des bestmöglichen Kundenerlebnisses.

Die Kund*in der vierten Generation (seit 2015) ist laut Freeman (2017) eine Person, deren Prinzipien sich (basierend auf den Erfahrungen der Kunden 1.0, 2.0 und 3.0) weiterentwickeln, sich jedoch jetzt durch eine große Anzahl an Auswahlmöglichkeiten, durch Online-Plattformen und der Entstehung eines neuen Kundentyps erschweren. Der gut informierte Kunde 4.0 ist mittlerweile ein wahrer Recherchekünstler, der dementsprechend viele Quellen befragt, bevor eine Kaufentscheidung getroffen wird. Eine gute Reputation ist dabei umso wichtiger (Weissman & Wegerer, 2018).

Der Kunde 4.0 ist ein „Netizen“ – eine Bürger*in des Internets. Als Digital Native ist es für ihn ein Leichtes, sich online mit anderen zu verbinden und Informationen

auszutauschen (Kotler, Kartajaya & Setiawan, 2016). Er ist basierend auf Forschungsergebnissen von Freeman (2017) einer, dessen persönliche Ziele über gutem Marketing stehen. Einer, der versucht andere Kund*innen zu beeinflussen und von ihnen beeinflusst zu werden. Einer, der die Erwartungen an den Anbieter hat, dass dieser die Art und Weise, wie er Produkte präsentiert und liefert, an die Ansprüche der Kund*innen anpasst. Wo die Customer Journey eine einzigartige, individuelle Reise ist, in der Unternehmen dazu beitragen, dass die Kund*innen ihre Ziele erreichen können. (Freeman, 2017)

Abbildung 7 stellt die verschiedenen Generationen der Kund*innen und die Änderungen der Einstellungen jener in – Bezug auf die Marktaktivität – dar (Nowacki, 2014).



Abbildung 7: Änderungen der Kunden-Einstellungen in Bezug auf die Marktaktivität über Jahrzehnte.

Quelle: eigene Darstellung, angelehnt an Nowacki, 2014

Laut Wereda & Woźniak (2019) lässt sich der Schluss ziehen, dass die Konsument*in der vierten Generation zu einer Einheit wird, da sie aufgrund von eigenen Vorteilen, moralischer Zufriedenheit und Engagement mit einem Unternehmen zusammenarbeiten möchte.

3.3 Die Customer Journey 4.0 im E-Commerce

Weissman & Wegerer (2018, S. 22) haben es bereits festgestellt: es verändern sich nicht die Märkte, sondern die Kund*innen. In der Zeit des allgegenwärtigen Handels verändern sich Einkaufsgewohnheiten nachhaltig aufgrund des ständig wachsenden Einflusses von E-Commerce und veränderten Kundenbedürfnissen bzw. Einkaufsverhalten von Kund*innen (Bailey, 2020, S. 6). Daher ist es wesentlicher denn je, den Fokus auf die Kund*in zu richten. Unternehmen erhalten im digitalen

Zeitalter das erste Mal die einheitliche Sicht auf ihre Kund*innen, da jegliche Kundenaktivität wie Käufe, Klicks, Retouren etc. auf jedem Kanal gesammelt, aggregiert und über eine zentrale Schnittstelle bereitgestellt werden kann (Bailey, 2020, S. 10). Durch die immer granularer werdenden Informationen zur Kanal- und Device-Nutzung wird auch diese Reise der Kund*in – die Customer Journey – immer komplexer (Heinemann, 2020, S. 13). Die Anzahl der möglichen Kontaktpunkte (Touchpoints) zwischen User*innen und einem Online Shop nimmt laut Heinemann (2020, S. 13) immer weiter zu, gängige Zielhandlungen beschränken sich nicht mehr nur auf Käufe, Bestellungen oder Anfragen. Außerdem ändert sich bei jeder individuellen Journey, welche Touchpoints von der Kund*in genutzt werden. Das hängt von vielen Faktoren ab, wie zum Beispiel dem Produktpreis, der Verfügbarkeit oder Zahlungsmöglichkeiten (Bailey, 2020, S. 11). Auch das Endgerät der potentiellen Kund*innen spielt basierend auf den Annahmen von Bailey (2020, S. 11) eine Rolle: die Customer Journey ist am Smartphone eine andere, als bei der Verwendung eines Laptops oder Tablets. Dazu kommt, dass sich der engagierte Kunde 4.0 Rat bei Freunden und Familie, über Bewertungen (Social Proof) und andere Vergleichsportale einholt, welche ebenfalls Touchpoints mit dem Unternehmen darstellen (Heinemann, 2020, S. 13).

Die Reise des Kunden vom ersten Gedanken daran, etwas zu kaufen, bis hin zum tatsächlichen Kaufabschluss und der Nutzung des Produkts kann wenige Minuten (bei sehr konkreten Kaufabsichten), mehrere Stunden (bei intensiveren Vergleichen) oder sogar über Monate hinweg (bei hochpreisigen Anschaffungen und Luxusartikeln) dauern (Weissman & Wegerer, 2018, S. 125). Die Customer Journey besagt, dass zwischen dem Kaufgedanken, dem Klick und dem Kaufabschluss eine Reihe an Stationen liegt, die eine Einflussnahme auf die Reise der Kund*in ausübt (Heinemann, 2020), sowohl analog als auch digital (Bailey, 2020, S. 11). Der neue Customer Journey Ansatz 4.0 muss also die Online- und Offline-Welt zusammenbringen, was eine qualifizierte Datenerhebung voraussetzt (Heinemann, 2020, S. 71). Typische Kontaktpunkte mit einem Online Shop könnten zum Beispiel Werbung im Radio oder TV, Printwerbemittel, Empfehlungen von Bekannten,

Newsletter, Bewertungen in Vergleichsportalen, Display- oder Retargeting-Marketing sein (Heinemann, 2020; Weissman & Wegerer, 2018). Aber auch jeder Kontakt mit dem Call-Center, mit einer Bestellung, mit einer Rechnung, mit einer Retoure etc. stellt einen Touchpoint mit dem Online Shop dar (Weissman & Wegerer, 2018, S. 125).

In der Literatur werden fünf bis sechs Phasen definiert, in denen potentielle Kund*innen Kontakt mit einer Marke, Unternehmen oder Online Shop haben können. Die fünf Phasen laut maximizer crm (2018) werden in Abbildung 8 dargestellt:



*Abbildung 8: Die neue Customer Journey 4.0 im E-Commerce.
Quelle: maximizer crm, 2018*

- **Awareness:** sich mit der Brand oder dem Online Shop vertraut machen
- **Consideration:** Unternehmen, Produkte oder Service-Angebote aktiv prüfen und recherchieren
- **Purchase:** Bewertung der Alternativen und Kaufentscheidung treffen
- **Retention:** Nutzung des Produkts und Erfahrungen mit Serviceangeboten und Support sammeln
- **Advocacy:** Zufriedenheit mit Kaufentscheidung entweder hoch oder niedrig; Meinung wird abgegeben – ob positiv oder negativ

Dabei spielen sich zwei der Phasen vor der Kaufentscheidung und drei Phasen mit und nach der Entscheidung der Kund*in ab (Heinemann, 2020).

Wie nun eine Customer Journey inklusive der verschiedenen On- und Offline-Touchpoints konkret in einem Online Shop aussehen kann, wird am Beispiel der Customer Journey von ottoversand.at veranschaulicht (Abbildung 9).

„Wir wollen unseren Kunden an jedem Touchpoint begeistern und mit exzellentem Service bedienen. Unser Anspruch ist die absolute Zufriedenheit unserer Kunden entlang der gesamten Kaufkette. Wir denken immer Mobile First.“ Mit diesem Statement beschreibt OTTO GmbH (2020) das Ziel und die Vision der eigenen Customer Journey.

Ist die erste Phase (Wahrnehmen & Bedürfnis wecken) noch geprägt von verschiedenen Marketingaktivitäten, Trigger oder aktiven Sortiments- oder Service-Vergleichen über die verschiedensten Kanäle hinweg, spielt sich die Phase 2 der OTTO Customer Journey (Stöbern & Informieren) direkt auf der Shop-Plattform ab. Die entsprechenden Touchpoints und deren Ausprägungen der Phase „Stöbern und Informieren“ werden exemplarisch in Tabelle 1 dargestellt.

Stöbern & Informieren		
Touchpoint	Ausprägung	Online/Offline
Startseite	Aktionen	Online
	Waren-Angebot	Online
	Inspiration	Online
	Serviceangebote	Online
Landingpages	Lifestyle	Online
	Werbliche Landingpages	Online
	Themen Landingpages (Markenshop Introseiten)	Online
	Sortimentsintroseiten	Online
Interne Suche	Sucheingabefeld	
	Suchergebnisse	Online
	Filter	Online

	Null-Treffer-Seiten	Online
	Produktempfehlungen, Serienartikel	Online
	Vorschlagssuche	Online
	Suche-Feedback-Formular	Online
Mehrbildansicht	Produktliste	Online
	Filter	Online
Merkzettel	Produkte/Ware	Online
Detailview	Aktionen	Online
	Artikelbild	Online
	Artikeldetailinformationen	Online
	Serviceangebote	Online
	Größen- und Farbauswahl	Online
	Kundenbewertung (Artikelbewertung)	Online
	Produktverfügbarkeit	Online
	In-den-Warenkorb-Button	Online
	AtB (Add to Basket) Layer	Online
	Produktempfehlungen	Online
Service & Hilfe	FAQ	Online
	Serviceseiten (Serviceangebote, Liefermöglichkeiten, Größentabelle, ...)	Online
	Paketshopfinder	Online
	Newsletteranmeldung	Online
	Kundenfachberatung (Telefon, Mail, Chat)	Online
	Kontaktformular	Online
	Unternehmensinformationen (Datenschutz, AGBs, ...)	Online
	Callback	Offline

***Tabelle 1: Touchpoints und Ausprägungen der Phase "Stöbern & Informieren".
Quelle: OTTO GmbH, 2020***

Im Rahmen dieser Arbeit liegt der Fokus auf dem Touchpoint der internen Suche mit den verschiedenen Möglichkeiten, nach Produkten zu filtern, die genau den Bedürfnissen und Anforderungen der User*in entsprechen.

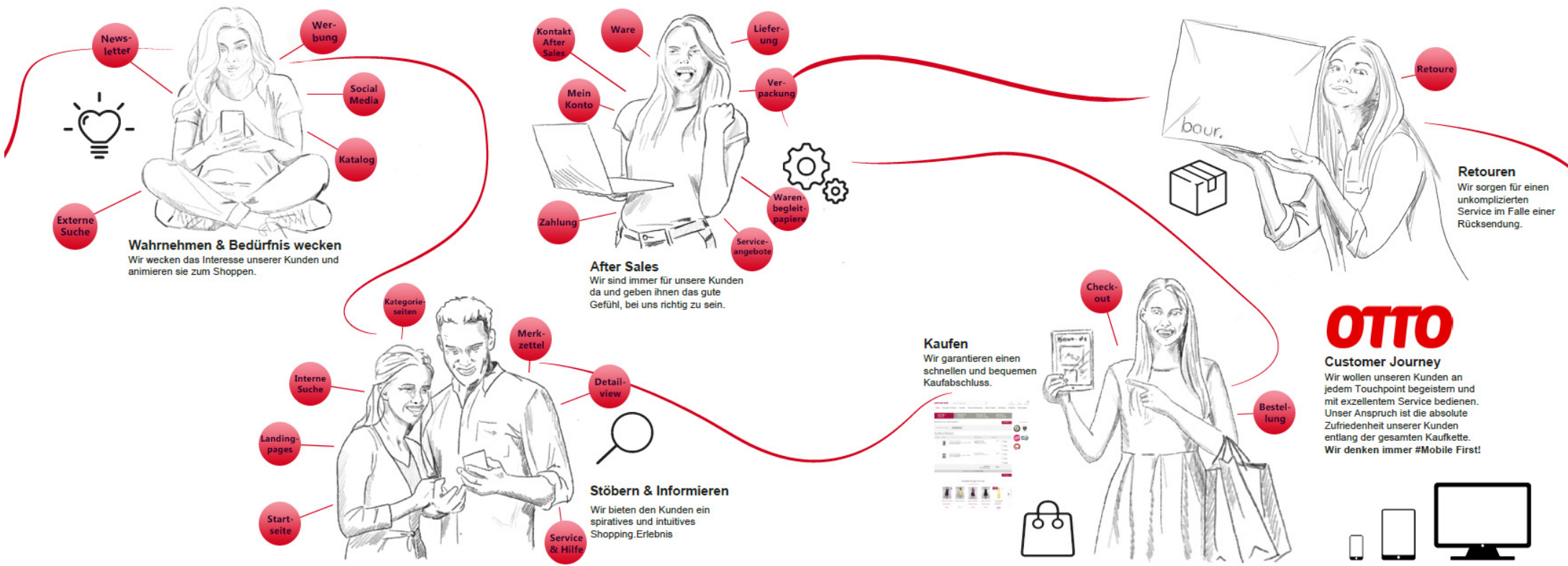


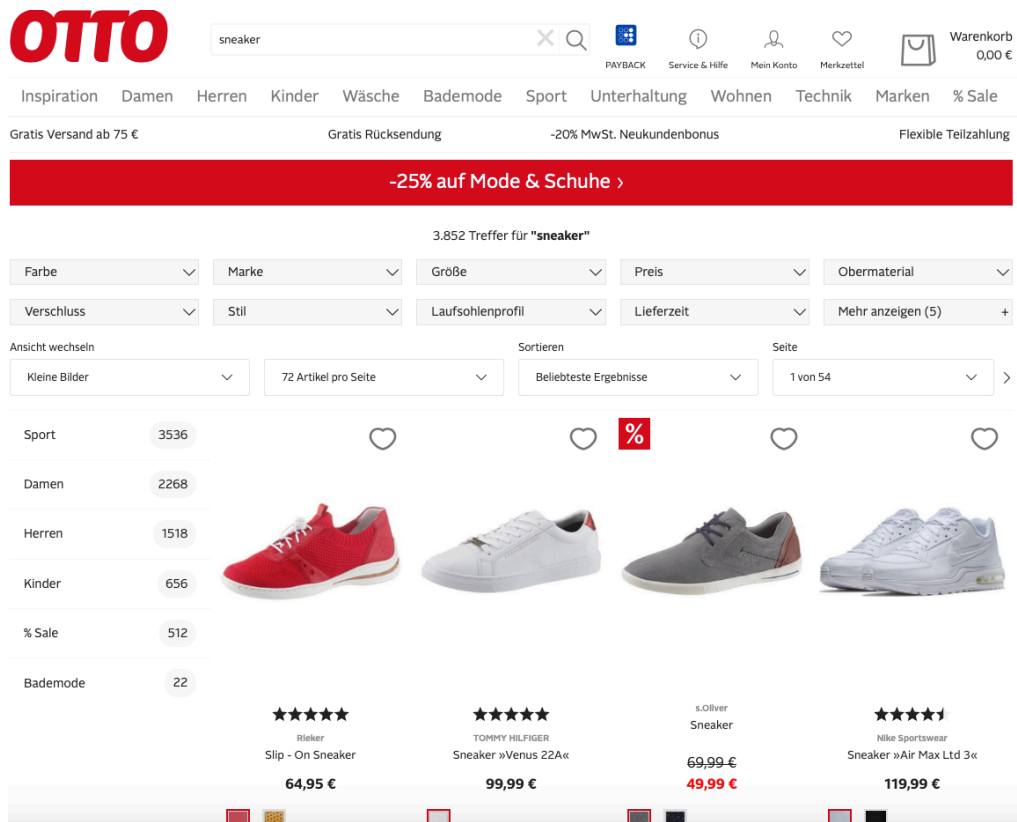
Abbildung 9: Customer Journey inklusive On- und Offline Touchpoints der OTTO GmbH.
Quelle: eigene Darstellung, angelehnt an OTTO GmbH, 2020

3.4 Onsite Search am Beispiel von OTTO Österreich

Einer der business-seitig relevantesten Touchpoints im Online Shop von OTTO Österreich ist sicherlich die Onsite Suche. So wurden im Jahr 2019 über die integrierte Suchfunktion mehr als 28 % Umsatz – gemessen am Gesamt-Umsatz – erwirtschaftet (OTTO GmbH, 2020a). Außerdem wurden 2019 laut OTTO GmbH (2020a) in über 21 % der Gesamtanzahl an Sessions Suchanfragen ausgeführt, jeder fünfte Shop-User benutzt basierend auf dieser Information also die Suchfunktion. Ausgehend von diesen Zahlen wird deutlich, wie wesentlich eine performante Onsite Search ist.

Nachfolgend wird erklärt, wie sich die Suche Frontend-seitig verhält, sprich was User*innen im Shop zu sehen bekommen, wenn eine Suche durchgeführt wird. Im OTTO Online Shop wird beim Ausführen einer Suchanfrage (z.B. über die Eingabe eines Suchbegriffs in den Suchschlitz im Header) eine Search Engine Result Page (SERP) aufgerufen (Abbildung 10). Auf dieser Suchergebnisseite werden verschiedene Elemente und Funktionalitäten dargestellt. Ganz oben auf der Seite ist der Header zu sehen, dieser zieht sich über alle Templates des Online Shops durch. So wird sichergestellt, dass die User*in die wichtigsten Seiten und Bereiche des Shops jederzeit ohne Umwege erreichen kann: die Startseite, den Service- und Hilfebereich, das persönliche Kundenkonto, den Merkmaltel und den Warenkorb. Aber auch weitere Suchanfragen können an jeder Stelle im Shop ausgeführt werden. Zudem ist die Hauptnavigation mit verschiedenen Untermenüpunkten ebenfalls im Header verankert und überall nutzbar.

Der eigentliche Inhalt der SERP beginnt ab der Anzeige des Suche-Recalls, sprich der Anzahl an gefundenen Artikel zu einem Suchbegriff. Im Beispiel der Abbildung 10 wurden 3.852 Treffer für den Suchbegriff „Sneaker“ gefunden (OTTO GmbH, 15.06.2020).



**Abbildung 10: SERP einer Suchanfrage mit Filtermöglichkeiten auf *ottoversand.at*.
Quelle: OTTO GmbH, 2020b**

Es folgen verschiedene Filtermöglichkeiten. Diese basieren auf den relevanten Produkt-Attributen des Recalls, wie Farben, Marken, Größen, Preis, Lieferbarkeiten, Produktbewertungen. Und je nach Suchanfrage bzw. Sortiment gibt es spezifische Filter wie Materialien, Absatzhöhe, Ärmellänge, Bildschirmdiagonalen, Möbel-Abmessungen, Ladevolumen Waschen, Touren, etc. Weitere Filter-Besonderheiten sind die Kategorie-Filter im linken Bereich sowie die Auswahl der Ansicht, der angezeigte Produkte, der Sortier-Logik und der Page.

Danach werden die zum Suchbegriff gefundenen Artikel angezeigt. Dabei sieht die User*in auf einen Blick die wichtigsten Produkt-Informationen wie die Marke, eine Kurzbeschreibung, den regulären Preis sowie eventuelle rabattierte Preise, eine Farb-Auswahl und wenn vorhanden Produktdatenblätter, Energieeffizienzklassen-Label und Bewertungssterne. Außerdem kann jedes Produkt von der SERP aus

direkt auf den Merktzettel zur besseren Vergleichbarkeit und zum Merken für später gelegt werden.

Als Suche-Backend hat der Online Shop von OTTO Österreich die Webshop Search Engine von *Magellan*-Explorer im Einsatz. Insbesondere die im Rahmen dieser Masterarbeit interessante Filter-Funktion wird über *Magellan* bestens bedient. Die Anzeige der Filter sowie auch die Sortierung kann im Back-Office gepflegt werden. Zudem ist es möglich, diese Filter automatisiert zu sortieren bzw. zu entscheiden, ob Filter angezeigt werden oder nicht. Die aktuell selbstlernende Sortier-Logik basiert auf der Nutzung von Filtern. Das ist eine grundvoraussetzende Funktion, um Filter beispielsweise auch personalisiert zu sortieren oder anzuzeigen. Zudem werden neben ein paar wenigen Filter, die immer angezeigt werden (z.B. Lieferbarkeit, Preis, Marke, Bewertungen, Farbe) nur die Filter eingeblendet, die eine Produktabdeckung von mindestens 35 % sicher stellen.

3.5 Die Herausforderungen und Probleme bei der Produktsuche

Basierend auf dem Business-Impact der Onsite Search bei OTTO Österreich (28 % des Gesamt-Umsatzes im Jahr 2019 durch die Suche generiert; in jeder 5. Session wurde die Suchfunktion verwendet), ist es umso wichtiger, dass die Kund*innen auf direktem und schnellem Weg die gesuchten Produkte finden. Blickt man hierbei auf die Bedürfnisse und Erwartungshaltungen von sowohl den Unternehmen, Online Shops oder Anbietern wie auch auf die der Such-Anwender*innen, wird klar, dass ein Zwiespalt entsteht (Riemer & Brüggemann, 2007). Denn einerseits soll die Suchfunktion des Online Shops eine möglichst breite, hohe und neutrale Abdeckung des Produktsortiments erzielen, um den vollen Umfang der Produktpalette darzustellen. Auf der anderen Seite sollen Suchmaschinen möglichst interessante und „die richtigen“ Suchergebnisse liefern und gut nutzbar präsentieren. Aus Sicht der Suche-Nutzenden stellt sich ein ganz ähnliches Problem: zum einen wollen die User*innen, dass keine potentiellen Ergebnisse abgeschnitten werden, zum anderen

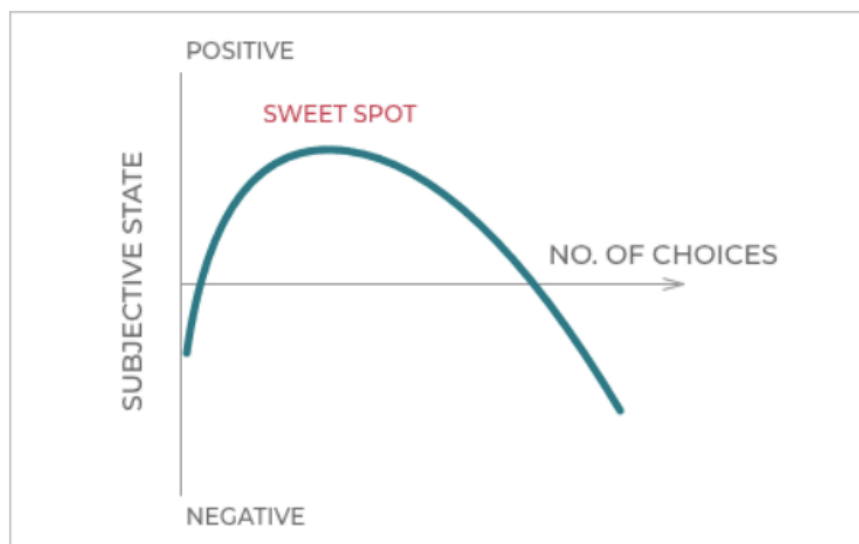
erwarten sie, dass sie mit möglichst genau den relevanten Ergebnissen versorgt werden, nach denen sie suchen (Riemer & Brüggemann, 2007).

Dieser Trade-Off verdeutlicht, dass im digitalen Zeitalter sowohl User*innen als auch die Online Shops wie OTTO oder Amazon erwarten, dass Artikel zielgerichtet und angepasst an die Bedürfnisse und Vorlieben der Kundschaften empfohlen werden, da insbesondere die Reizüberflutung und ein Information Overload eine Produktsuche deutlich erschweren können (Guo, Wu, Wang & Tan, 2016, S. 191). Denn im Beispiel des Suchergebnisses zum Suchbegriff „Sneaker“ (siehe Abbildung 10) wird dem User*in im OTTO Österreich Online Shop ein Recall von 3.852 Produkten geliefert. Eine wichtige Erkenntnis in Bezug auf einen auftretenden Information Overload beziehungsweise einer Vielzahl an potentiellen Produkten, die zur Kaufauswahl stehen, liefern die Forschungsergebnisse von Schwartz (2005). Er beschreibt in seinem Werk mit dem Effekt des „Paradox of Choice“, dass die Bereitstellung von mehreren Optionen oder einer großen Auswahl – insbesondere wenn sie von hoher Relevanz und die Errungenschaft des Produkts persönlich wichtig ist – zu einer schlechteren Wahl und geringeren Zufriedenheit führen (Schwartz, 2005).

Besonders für Online Shops, die ein großes und breites Produkt-Spektrum vertreiben (Long-Tail Strategie) ist das Paradox of Choice eine große Herausforderung. Denn ist ein großes Sortiment scheinbar ein Alleinstellungsmerkmal, kann es die Entscheidungsfindung bei der Kundenschaft deutlich erschweren und dazu führen, dass eine – im nachhinein gesehen – schlechte Entscheidung getroffen oder die Kaufentscheidung gar verhindert wird (Spreer, 2018, S. 132). Laut Schwartz (2005) hat dieser Effekt mehrere Ursachen: zum einen ist es so, dass je mehr Alternativen zur Auswahl stehen, desto höher der Aufwand für eine Bewertung bzw. Abwägung dieser ist. Das Gehirn ist allerdings dahingehend programmiert, mittels Heuristiken diese Aufwände zu reduzieren, wodurch ein Konflikt entsteht. Außerdem blähen viele Optionen die Erwartungen gegenüber der Auswahl auf, sodass die Zufriedenheit mit dem gewählten Artikel

oftmals hinter den Erwartungen bleibt (Schwartz, 2005). Den aus Neuromarketing-technischer Sicht fokussiert sich das Gehirn nicht auf das, was es hat, sondern auf das, was es nicht hat (Semeijn, 2018).

Die Existenz dieses Phänomens könnte wichtige Auswirkungen auf den Einsatz und die Optimierung von Suchmaschinen haben (Oulasvirta, Hukkinen & Schwartz, 2009, S. 516) – insbesondere auch auf Suchfunktionen in Online Shops. Das bedeutet nicht, dass das Beste darin besteht, überhaupt keine Auswahl zu bieten. Schwartz (2014) entdeckte im Paradox of Choice einen sogenannten „Sweet Spot“.



*Abbildung 11: Sweet Spot im Paradox of Choice.
Quelle: Semeijn, 2018*

Das ist jener Punkt, an dem die Anzahl der Auswahl-Möglichkeiten für das subjektive Wohlbefinden am effektivsten ist (Abbildung 11). Dieser Sweet Spot ermöglicht es den Suchenden, von der Vielfalt der Auswahl zu profitieren und nicht von ihr gelähmt zu werden (Schwartz, 2014).

Das Paradox of Choice hat essentielle Auswirkungen auf das Design von Suchergebnissen. Denn stellt sich nun die allgemeine Annahme als fragwürdig heraus, dass mehr Auswahl und eine große Anzahl an Suchergebnissen eine höhere

Wirksamkeit bei den Kund*innen erzielt, ist die Frage berechtigt, ob Suchmaschinen weniger „stupiden Reportern“ ähneln sollten, sondern eher persönlichen Assistenten, die Kundschaften zu den vernünftigsten und besten Optionen in einem Shop führen (Oulasvirta et al., 2009, S. 516).

Denn auch Vogt & Quelle-Korting stellten 2017 fest, dass im Besonderen Relevanz und Präzision zwei wesentliche Merkmale für ein Suchergebnis sind, welches eine hohe (subjektive) Qualität und dadurch auch die Erwartungskonformität der Nutzer*innen sicherstellen. Sie behaupten basierend auf ihren Forschungen, dass auf der ersten Seite eines Suchergebnisses ausschließlich für den Suchende*n relevante Produkte angezeigt werden sollen. Falls der Recall auch irrelevante Produkte liefert, sollten diese hinter den relevanten Artikeln einsortiert werden, wobei sich der Begriff „relevant“ auf die Erwartungskonformität des Produktes aus Nutzersicht in Bezug auf den Suchbegriff bezieht (Vogt & Quelle-Korting, 2017, S. 16).

3.6 Die Herausforderungen der Personalisierung hinsichtlich Datenschutz

Auch wenn es in den letzten Jahren einen Boom an personalisierten Angeboten gab, wird die Personalisierung öffentlich trotzdem immer wieder kritisch diskutiert (Lewandowski, 2018). Denn eine erfolgreiche Personalisierungsstrategie hängt nicht nur davon ab, ob Online Shop Betreiber Informationen über ihre Kund*innen entsprechend erfassen, analysieren und verarbeiten können. Personalisierung hängt auch maßgeblich von der Bereitschaft der User*innen ab, Daten über das Nutzungs- und Kaufverhalten im Online Shop preiszugeben, Informationen auszutauschen und Personalisierungsdienste zu nutzen (Chellappa & Sin, 2005, S. 181). Zwar schätzen die Kund*innen personalisierte und auf sie maßgeschneiderte Services, Angebote und Inhalte, jedoch vermuten sie, dass ihre privaten Informationen ohne ihr Wissen getrackt werden (Phelps, Novak & Ferrell, 2000; Sheehan & Hoy, 2000), was wiederum negative Gefühle in Bezug auf Personalisierung hervorruft (Andrade,

Kaltcheva & Weitz, 2002). Trotzdem ist die Personalisierung im E-Commerce eine wichtige Strategie, da sie einen positiven Beitrag zum Unternehmenserfolg leistet (Lee & Cranage, 2011, S.987).

Dieses Phänomen, dass User*innen zwar nicht wollen, dass ihre Daten für unternehmerische Zwecke oder Empfehlungen verwendet werden, im Alltag der Nutzen von datengetriebenen Services diese Risiken dann aber aufwiegt, wird als „Paradoxon der Personalisierung und Privatsphäre“ bezeichnet (Awad & Krishnan, 2006). Werden Bedenken zum Datenschutz von Online Shops angemessen berücksichtigt, ist es denkbar, dass die Wahrnehmung hinsichtlich der kritisch gesehenen Personalisierung durch die Kund*innen sogar positiv wird (Lee & Cranage, 2011, S. 988). Es kann also zum einen passieren, dass ein Mangel an Datenschutz die Bedenken der Kund*innen bezüglich dem Umgang mit den personenbezogenen und privaten Daten bestärkt. Zum anderen kann die Kombination aus hochgradig personalisierten Services mit ernst genommenen und aufgeklärten Datenschutz das Potential steigern und einen enormen Mehrwert für die Kund*innen schaffen (Lee & Cranage, 2011, S. 988).

Eine einfache und skalierbare Möglichkeit, Personalisierung in Online Shops einzusetzen und Benutzerprofile zu erstellen, bieten Big Data Technologien (Habegger et al., 2014). Dabei weisen verschiedene Analysetechniken basierend auf Big Data zwei große Stärken auf:

- 1) Big Data Methoden können neben strukturierten Daten auch unstrukturierte Daten verarbeiten. Diese werden von Benutzer*innen implizit generierte und nehmen aufgrund der vielen Informationen mit hoher Geschwindigkeit an Volumen zu (Habegger et al., 2014). Mit dem Aufkommen des technologischen Fortschritts im digitalen Zeitalter ist das Speichern und Übertragen derart großer Datenmengen keine Herausforderung mehr, zudem ist auch die Analyse solcher Datenmengen in Rekordzeit – aufgrund der Verfügbarkeit neuer Algorithmen und Methoden – kein Problem (Punagin & Arya, 2015).

2) Mittels Big Data Analysetechniken können Daten aus mehreren Quellen verarbeitet werden, wodurch Benutzerattribute aus verschiedenen Quellen zu einem einzigen Benutzerprofil zusammengefasst werden (Habegger et al., 2014).

Zwar ist genau diese umfangreiche und aggregierende Userprofil-Erstellung ein großer Mehrwert aus Unternehmenssicht, da zielgenauere und relevantere Dienste und Services bereitgestellt werden können (Habegger et al., 2014). Dank dieser Benutzerprofile ist es beispielsweise auch möglich, den Kund*innen personalisierte Suchergebnisseiten, Produktlisten oder andere Produkt- und Content-Empfehlungen anzubieten (Punagin & Arya, 2015). Auf der anderen Seite stellt die Generierung von Benutzerprofilen eine erhebliche Bedrohung für die Privatsphäre der Benutzer*innen dar, denn Webseiten, die nur bedingt zum Schutz der Daten neigen, können Personalisierungsinformationen missbrauchen, wenn die User*innen einer Verwendung nicht zugestimmt haben (Habegger et al., 2014).

Für eine datenschutzfreundliche Gestaltung von Benutzerprofilen und als Schutzmaßnahme für die Verarbeitung von personenbezogenen Daten sind die Pseudonymisierung und die Anonymisierung anerkannte Instrumente (Knopp, 2015, S. 527). Beide technischen Konzepte verfolgen das Prinzip der Datensparsamkeit. Dabei wird in § 3a BDSG festgehalten, dass die Erhebung, Verarbeitung und Nutzung von personenbezogenen Daten und die Auswahl sowie Gestaltung der Datenverarbeitungssysteme sich an dem Ziel auszurichten haben, so wenig personenbezogene Daten wie möglich zu erheben, zu verarbeiten oder zu nutzen. *„Insbesondere sind personenbezogene Daten zu anonymisieren und zu pseudonymisieren, soweit dies nach dem Verwendungszweck möglich ist und keinen im Verhältnis zu dem angestrebten Schutzzweck unverhältnismäßigen Aufwand erfordert“* (§ 3a BDSG). Anonymisierung und Pseudonymisierung unterscheiden sich darin, dass Daten, die unter einem Pseudonym gespeichert werden, oftmals noch einzelnen Personen zuzuordnen sind (Schaar, 2012, S. 369). Eine Pseudonymisierung kann vor allem durch eine Kodierung erreicht werden, wobei eine Zeichenfolge generiert wird, um damit eine Zuordnung zu den Daten zu

ermöglichen (Dorschel, 2015, S. 187). Diese Zuordnung sollte nur bei Bedarf durchgeführt werden und darf nur unter Einhaltung vorher definierter Rahmenbedingungen den Personenbezug wiederherstellen (Schaar, 2012, S. 369). Methoden zur Pseudonymisierung sind die Listenerstellung (Datensätze anhand einer Tabelle bestimmten Pseudonymen zuordnen) bzw. Berechnungsverfahren durch die Anwendung von kryptographischen Schlüsseln (Wegeler, 2019). Im Gegensatz dazu erfordert die Anonymisierung gemäß § 3 Abs. 6 BDSG eine derartige Veränderung der personenbezogenen Daten, dass keine Rückschlüsse mehr auf die bezogene Person möglich sind bzw. nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten & Arbeitskraft. Mögliche Varianten der Anonymisierung könnten eine Generalisierung, eine Löschung, eine Mikroaggregation oder eine Verfälschung sein (Winter, Battis & Halvani, 2019, S. 342).

3.7 Die Herausforderungen der Personalisierung hinsichtlich Filter-Bubbles

Maßgeschneiderte Informationen sollten die Fähigkeit einer Person stärken, Entscheidungen zum eigenen Vorteil zu treffen (Becker, 2019, S. 310). Die Kontrolle und das Wissen der Individuen über den Informationsfluss gehen dabei allerdings verloren. User*innen werden in eine Filter-Bubble versetzt, die auf Algorithmen und Unternehmensrichtlinien basiert, welche den Zielpersonen unbekannt sind (Becker, 2019, S. 310). Doch Nutzer*innen müssen sich im Klaren sein, dass die Daten, die im Rahmen der Personalisierung gesammelt werden, auch dazu beitragen, dass dort die Bildung von Filter-Bubbles beginnt (Fitzgerald, 2017). Eli Pariser (2011, S. 23) stellt in seinem Buch die Behauptung auf, dass Personalisierung ein Service für jedes Individuum sein sollte, doch die gesammelten Nutzerdaten sind vor allem auch anderen dienlich: Unternehmen, Online Shops, Werbetreibenden, ja sogar anderen User*innen, die sich ähnlich wie das Individuum verhalten. Denn Personalisierung wird häufig als das Anzeigen von relevanten Informationen verkauft und auch so verstanden (Fitzgerald, 2017). Doch oftmals werden „relevante Produkte und Services“ angezeigt, damit die Bedürfnisse des Unternehmens gestillt werden.

Zuerst werden die Ziele des Online Shops verfolgt, danach erst die Wünsche der Benutzer*innen befriedigt (Fitzgerald, 2017). Fitzgerald (2017) geht in seinem Beitrag sogar noch weiter und sagt: *„Relevanter Inhalt ist in einer Best-Case-Definition „Information, an der das Individuum am meisten interessiert ist“: Relevanter Inhalt ist aber möglicherweise genau das, was User*innen auf einer Webseite hält und zu einer Konvertierung bringt. Es wird aber wahrscheinlich keine große Überschneidung mit Informationen geben, die Vorurteile in Frage stellen, unsere Annahmen brechen oder die Weltansicht erweitern.“* Personalisierung verändert nicht nur die Wahrnehmung von Informationen, Produkten und Inhalten, sondern sie verändert auch die ökonomische Ausgangslage, die bestimmt, welche Inhalte überhaupt produziert werden (Pariser, 2011, S. 76).

Im Zeitalter der Digitalisierung ist die Menge an Informationen schier uneingeschränkt, nach denen Interessierte im Web suchen können (Volkens & Anderson, 2017, S. 81). Dabei kann, wie bereits vorangegangen erläutert, eine Reizüberflutung beziehungsweise ein Information Overload auftreten (Chen et al., 2009). Um den Effekt des Information Overloads zu entkräften, gibt es unterschiedliche Filtermechanismen. Laut Volkens & Anderson (2017, S. 81) liegt dies zum einen im Menschen selber. So neigen Informationssuchende im Unterbewusstsein dazu, sich eher bestätigenden Informationen zuzuwenden. Der Effekt wird „Confirmation Bias“ bezeichnet, dabei tendiert ein Individuum dazu, sich den Informationen, Services oder Artikel zuzuwenden, die eine vorausgewählte Alternative unterstützen (Schulz-Hard, Frey, Lüthgens & Moscovici, 2000). Dieser „Bestätigungsfehler“ bezeichnet die Neigung, Dinge als wahr anzusehen, die die eigene bestehende Sicht bestärken und vor allem das zu sehen, was man sehen möchte (Pariser, 2011, S. 93).

Zum anderen beruhen auch technische Filtersysteme auf dem Prinzip der maximalen Ähnlichkeit, wobei Filtermechanismen ihre Vorhersagen aus der Vergangenheit schließen und sie mit ähnlichen Verhaltensmustern anderer Nutzer*innen abgleichen (Volkens & Anderson, 2017). Diese Filtermechanismen

sind Algorithmen, anhand derer einzelne Individuen Segmenten zugeordnet werden und beispielsweise ihre Suchergebnisse den Vorlieben der Gruppe angepasst werden (Pariser, 2011, S. 42). So werden durch Klicksignale gespeiste Informationswolken User*innen eher mit Inhalten umgeben, die bestehende Ansichten bestärken, als diese zu hinterfragen (Pariser, 2011, S. 96). Abbildung 12 symbolisiert laut Pariser (2011) diese Aspekte als das Phänomen der Filter-Bubble. Dabei erklärt er in seinem TEDTalk: *„Die Filterblase ist ein ganz persönliches einzigartiges Informationsuniversum, in dem Individuen online leben. Was in der Filter-Bubble ist, hängt davon ab, wer dieses Individuum ist und was es tut. Aber es bestimmt nicht, was hineinkommt. Und noch wichtiger ist, dass man nicht sehen kann, was aussortiert wird.“* (Pariser, 2011a)

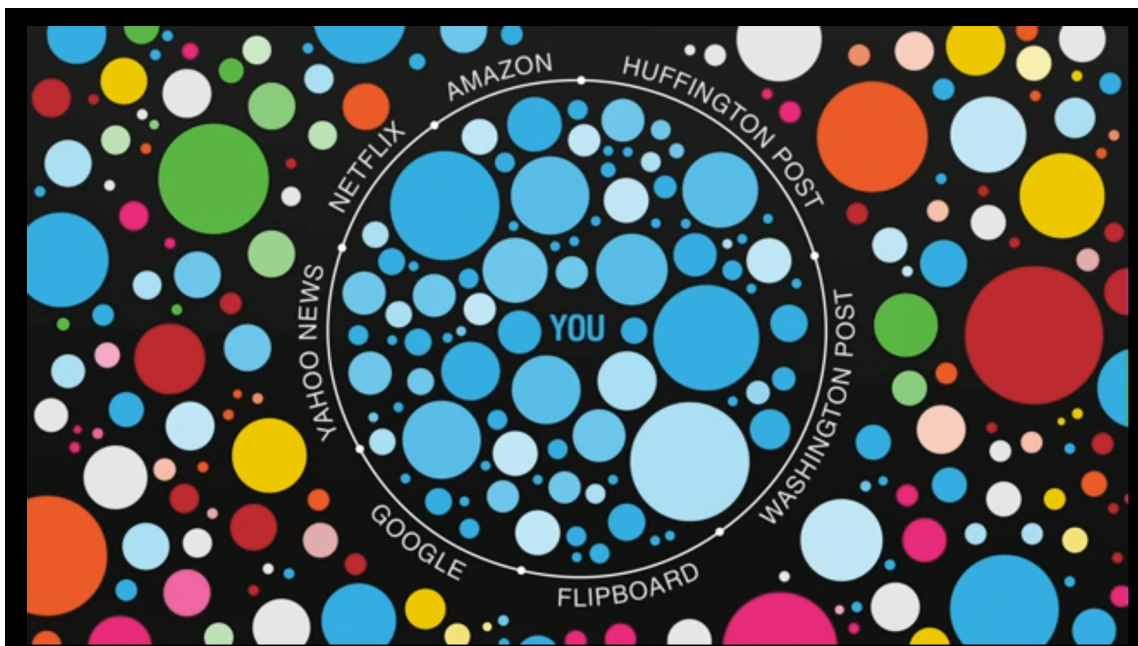


Abbildung 12: Die Filter-Bubble.
Quelle: Pariser, 2011

3.8 Die Chancen und Möglichkeiten der Personalisierung im Online Shop

Warum Personalisierung also trotzdem benötigt wird, ist jedenfalls eine berechtigte Frage, weshalb die Autorin im Rahmen dieser Masterarbeit die Notwendigkeit sieht, die Chancen und Vorteile der Personalisierung mit Fokus auf die Onsite Suche genauer zu beleuchten.

Unterschiedliche User*innen eines Online Shops haben beim Stellen einer Suchanfrage unterschiedliche Anforderungen und Bedürfnisse (Rastegari & Shamsuddin, 2010). Die Erwartungshaltung ist allerdings wohl bei allen Nutzenden die gleiche: die in der jeweiligen Situation der Nutzer*in notwendigen Informationen mit möglichst geringem Aufwand in der richtigen Qualität und Granularität zu finden (Riemer & Brüggemann, 2007, S. 116). Dabei stehen Suchmaschinen im Web, aber auch Suchfunktionen in Online Shops oder auf Webseiten vor der Herausforderung, aus riesigen Mengen an Dokumenten, Produktdaten und Informationen die für die Nutzer*in passenden Ergebnisse zu finden (Riemer & Brüggemann, 2007, S. 116). Wie bereits vorangehend ausführlich diskutiert, kann dieser Information Overload zu einer Überforderung der Individuen führen, es herrscht das Paradox of Choice (Chen et al., 2009; Schwartz, 2005). Darum wird insbesondere im E-Commerce mit Vorselektion, Verknappung, Bewertungen und mit Personalisierung gearbeitet (Volkens & Anderson, 2017, S. 82).

Setzt man das Prinzip der Personalisierung in Beziehung mit der Informationssuche, dann ist die Anordnung von Informationsobjekten so zu gestalten, dass sie an eine individuelle Nutzer*in angepasst ist (Lewandowski, 2018). Ein großer Mehrwert für die Nutzenden ist dabei, dass die Informationsbedürfnisse eines jeden Einzelnen besser befriedigt werden können, als das mit konventionellen Verfahren möglich ist, bei denen alle Nutzer*innen die gleichen Produktergebnisse erhalten (Lewandowski, 2018). Denn durch das Berücksichtigen der individuellen Vorlieben und Präferenzen sollen die

unterschiedlichen Wünsche und Ziele berücksichtigt werden, um aus Sicht des Suchenden relevantere und genauere Suchergebnisse zu erhalten (Jürgens et al., 2014, S. 106).

Um relevant zu sein, benötigen Algorithmen zur Personalisierung – egal ob zur Personalisierung der Onsite Suche, für Produktempfehlungen, personalisierten Content etc. – möglichst viele Daten (Pariser, 2011, S. 46). Genauer gesagt basieren die meisten Personalisierungstechniken im Kern auf Nutzerprofilen und der Erhebung persönlicher Daten der User*innen (Riemer & Brüggemann, 2007, S. 118), denn Personalisierung funktioniert dann am besten, wenn die Kund*in genügend Daten über sich preisgibt (Volgens & Anderson, 2017, S. 82). Dabei werden die Daten in Profilen gespeichert, welche von Personalisierungssysteme abgefragt, analysiert und verarbeitet werden können (Datta, Dutta, VanderMeer, Ramamritham & Navathe, 2001). Zur Gewinnung dieser Präferenzdaten der Online Shop Nutzenden können entweder explizite oder implizite Techniken und Methoden eingesetzt werden (Liu, Reiner, Frigessi & Scheel, 2019, S. 1).

Bei der expliziten Personalisierung legt die Nutzer*in ein Profil an, in dem Vorlieben, Wünsche und Interessen ausdrücklich von ihm selber benannt werden (Lewandowski, 2018, S. 5). Das kann beispielsweise durch die Bewertungen von Interessenskategorien oder das Ausfüllen von Fragebögen gelöst werden (Greco, Grego & Zumpano, 2004, S. 158). Die User*in ist sich dabei dessen bewusst, dass das angelegte Profil und die Interaktion mit Fragebögen oder das Kundgeben von Interessen die Suchergebnisse, Online Shop Inhalte, Produktempfehlen, etc. beeinflussen (Lewandowski, 2018, S. 5). Die explizite Profilerstellung bringt den Vorteil mit sich, dass die Individuen Präferenzen direkt angeben und die Kontrolle über die Personalisierung behalten können (Ansari & Mela, 2003, S. 7). Auf der anderen Seite stellt der explizite Personalisierungsansatz das Problem dar, dass Anwender oft nicht in der Lage sind, ihre Präferenzen richtig auszudrücken (Teevan, Dumais & Horvitz, 2005). Außerdem veralten einmal gepflegte Präferenzen im

Zeitablauf, wenn sich nicht wieder manuell bei Interessensänderungen bearbeitet werden (Ansari & Mela, 2003).

Im Gegensatz zu der expliziten Methode der Personalisierung basiert die implizite Technik auf der automatisierten Gewinnung und Analyse von Daten (Mertens, Stößlein & Zeller, 2004, S. 8). Meist sind sich die User*innen bei der impliziten Personalisierung nicht darüber bewusst, dass der Online Shop bzw. die Suchergebnisse personalisiert werden beziehungsweise die Gründe für die Anzeige bestimmter Ergebnisse meistens nicht transparent sind (Lewandowski, 2018, S. 5). Anhand verschiedener Faktoren bestimmt das Personalisierungssystem, woran die User*in interessiert sein könnte: vorheriger Interaktion mit dem Online Shop oder der Webseite wie Klicks, Browserverlauf, Suchanfragen, Käufen, Zeitaufwand für das Lesen von Informationen zu einem Produkt, etc. sowie IP-Adresse, Cookies, Browser-Informationen und anderen Faktoren (Bozdag, 2013, S. 213). Der große Vorteil impliziter Personalisierungstechniken ist für E-Commerce Retailer wie OTTO Österreich, dass im Online Shop zum einen umfangreiche Datenmengen über jedes einzelne Individuum gesammelt werden können und zum anderen Nutzer*innen nicht aktiv zur eigenen Profilerstellung beitragen müssen (Lewandowski, 2018, S. 6).

Zudem werden User-Profile automatisch aktualisiert, sobald Kund*innen mit dem Online Shop interagieren (Bozdag, 2013). Dadurch können auch Veränderungen der Nutzerinteressen und -präferenzen im zeitlichen Verlauf besser abgebildet werden, weil eine kontinuierliche Überprüfung die Erstellung von dynamischen Profilen unterstützt (Datta et al., 2001, S. 104). Die Generierung der Daten geschieht automatisch, wodurch Profile mit sehr geringen Kosten erstellt werden können (Wiedmann, Buxel & Walsh, 2002, S. 172).

Im Rahmen dieser Arbeit wird das implizite Sammeln von Klickdaten – insbesondere das Setzen von Produktfilter auf Suchergebnisseiten – beleuchtet. Clickstream-Daten können einfach erfasst und gesammelt werden, da diese

Information in großen Mengen vorkommt (Liu et al., 2019, S. 1). Eine Herausforderung dieser impliziten Methode ist jedoch, dass als Rückkopplung an das System oftmals nur „positives“ Feedback erfasst wird (Gauch, Speretta, Chandramouli & Micarelli, 2007, S. 64). Aus dieser Annahme lässt sich ableiten, dass, wenn von dem oder der Suchenden ein Produktfilter gesetzt wird, davon auszugehen ist, dass ein gewisses Interesse des Nutzens auf eine Filter-Eigenschaft bzw. -Wert gegeben ist (Gauch et al., 2007). Diese gesammelten Präferenzangaben können im Beispiel des Online Shops ottoversand.at durch die Möglichkeit des Löschs von Filtern verfälscht werden. Denn werden bereits gesetzte Filter wieder gelöscht, wird diese Aktion initial ebenfalls nur als Klick auf einen Button gewertet, der jedoch eine besondere Information enthält: und zwar das „Widerrufen“ einer Präferenz bzw. Interesse. Die Möglichkeit des Löschs von Filtern stellt also eine negative Feedback-Rückkopplung an das System dar, welche im Personalisierungsalgorithmus betrachtet werden sollte.

Die Personalisierung von Diensten, Produkten und Inhalten ist untrennbar mit der Sammlung und Verarbeitung von persönlichen und personenbezogenen Daten verbunden, die daraus erstellten Kundenprofile werden zu einer wirtschaftlich bedeutenden Ressource (Müller, Sackmann, Günther & Spiekermann, 2007, S. 1). Doch fühlt sich die Kundin durch Personalisierung überwacht, könnte genau das Gegenteil vom eigentlichen Ziel erreicht werden (Einfinger, 2003, S. 81). Die Internet-Kundschaft ist besonders was die Wahrung der Privatsphäre angeht sensibel, zudem gibt es eindeutige und klar offen gelegte Richtlinien für das Unternehmen, welche Daten über User*innen gesammelt und weiterverarbeitet werden dürfen (Einfinger, 2003, S. 81).

4. Durchführung der Datensammlung im Online Shop

4.1 Implementierung der Filter-Funktion im OTTO Online Shop

Die unterschiedlichen Filter und Filterwerte auf den Suchergebnisseiten werden basierend auf den Produktattributen der angezeigten Artikel erzeugt. Im Produktinformationssystem des Online Shops wird an den Produktkategorien gepflegt, welche Attribute auf Produktlisten als Filter dargestellt werden sollen. Das Suche-Backend kommuniziert mit diesem PIM und schaut sich die Kategorie-Nachrichten an. Es wird dann ein Abgleich gemacht, ob es zu den abgerufenen Attributen auch schon Filter gibt. Wenn nicht, werden die Filter inkl. Filter-Werte automatisch neu angelegt. Außerdem läuft einmal am Tag eine Bereinigung, die Filter auch wieder löscht, wenn das zugehörige Attribut in keiner einzigen Kategorie mehr als Filter gepflegt ist.

Die Anzeige eines Filters auf der Suchergebnisseite ist durch einen Algorithmus gesteuert. Dieser fragt folgende Bedingungen ab, um festzustellen, ob ein Filter angezeigt wird, oder nicht. Es werden also nur Filter angezeigt:

- die vom Suche-Backend auch bereits gebaut wurden, weil ein Attribut als Filter an einer Kategorie gepflegt war.
- die mehr als einen Filter-Wert (= Filter-Ausprägung) haben.
- die Filter-Werte von mehr als sechs verschiedene Artikeln haben.
- bei denen mindestens 35 % der gefundenen Artikel Filter-Ausprägungen zu den entsprechenden Filtern haben.

Wird eine Regel verletzt, wird der Filter standardmäßig nicht angezeigt. Diese Logik ist jedoch auch übersteuerbar. Im Suche-Backend können einzelne Filter manuell fest ein- oder ausgeblendet werden.

Die Sortierung der Filter untereinander wird über eine Heuristik im *Magellan* Suche-Backend festgelegt. Diese Heuristik ermöglicht bereits heute eine selbstlernende Sortierung, die für jeden Suchbegriff einzeln anhand von Klickdaten die Reihenfolge der Filter bestimmt. Oft benutzte Filter rutschen so weiter nach vorne, weniger häufig verwendete Filter werden nach hinten sortiert. Sollte es für einen Suchbegriff zu wenig Klickdaten geben, um eine gelernte Sortierung auszuspielen, kommt als Fallback eine globale Positionsangabe als Filter-Sortierung zum Einsatz.

Auch die Filter-Werte innerhalb eines Filters unterliegen einer Sortierung. Diese kann ebenfalls im *Magellan*-Backend eingestellt werden. Mögliche Sortier-Logiken sind hier eine alphabetische Sortierung oder nach Anzahl der gefilterten Artikel.

4.2 Erhebung und Aufzeichnung von Filternutzungs-Trackingdaten

Um die Klick-Performance der Filter auf Suchergebnisseiten messen zu können, wurde ein Webtracking implementiert. Dafür sind sogenannte *Landmarks* im Code verbaut, die gefeuert werden, sobald eine Aktion ausgelöst wird. Eine Aktion kann zum Beispiel das Setzen eines Marken-Filters, das Einstellen des Preisreglers im Preisfilter, oder das Löschen eines gesetzten Filters sein. Durch das Feuern der Tracking-Landmarks können einzelne Klicks aufgezeichnet werden. So ist sehr genau nachvollziehbar, zu welchem Suchbegriff eine User*in wann einen Filter gesetzt, verändert oder wieder gelöscht hat.

Ein Landmark besteht aus mehreren verschiedenen Landmark-Attributen, die wiederum verschiedene Attributs-Werte haben können. In Tabelle 2 werden die für diese Masterarbeit relevanten Tracking-Attribute des *suche*-Landmarks angeführt.

Landmark-Attribut	Beschreibung des Attributs	Mögliche Attribut-Werte
stern	Suchbegriff	z.B. „hose“
search-id	Jede Suchanfrage erhält eine eindeutige ID	z.B. „bjxehaxcylg20181024“
mkz	Major MKZ des Suchbegriffs	z.B. „131“; -1 bei Nulltreffer
sges	Filter nach Geschlecht	boy, male, female, girl, babymale, babyfemale, child, babykleidung
smark	Markenfilter	z.B. „Tommy Hilfiger“
spreis	Preisfilter	Nur reduzierte Artikel, Preisbereich (z.B. 15-788), - (wird gefeuert, wenn keine Option gewählt)
ssize	Gefilterte Größe	z.B. „38“
filter.filter_color	Farbfilter	z.B. „blau“
filter.s_filter_Availability	Lieferbarkeitsfilter	z.B. „sofort lieferbar“

Table 2: Attribute und Ausprägungen des suche-Landmarks.
Quelle: OTTO GmbH, 2020

Für die Datensammlung werden also bei vordefinierten Aktionen – wie zum Beispiel dem Setzen der oben angeführten Filter – Landmark-Attribute mit den entsprechenden Werten in Form von Informationen in eine Analytics-Datenbank des Anbieters *Exasol* weggeschrieben und aufgezeichnet. Um Erkenntnisse gewinnen und Entscheidungen ableiten zu können, werden die gesammelten Informationen in einfacher dargestelltes und aufbereitetes Wissen umgewandelt. Dabei werden die Mitarbeiter*innen von OTTO GmbH beispielsweise mit dem Tool *Tableau* unterstützt, dass die Daten aus der Exasol Analytics-Datenbank grafisch in einzelnen Views aufbereitet und greifbarer macht. Außerdem können User-Interaktionen anhand der Customer Journey basierend auf den getrackten Landmarks in einer OTTO-Eigenentwicklung, dem „*LMTTestTool*“, nachgestellt und nachvollzogen werden.

In den nachfolgenden Abbildungen wird eine Beispiel-Suchanfragen nach „ballerina“, in der der Größen- und der Farbfilter gesetzt werden, im Online Shop durchgeführt (Abbildung 13) und die getätigten Interaktionen im LMTTestTool nachvollzogen (Abbildung 14).

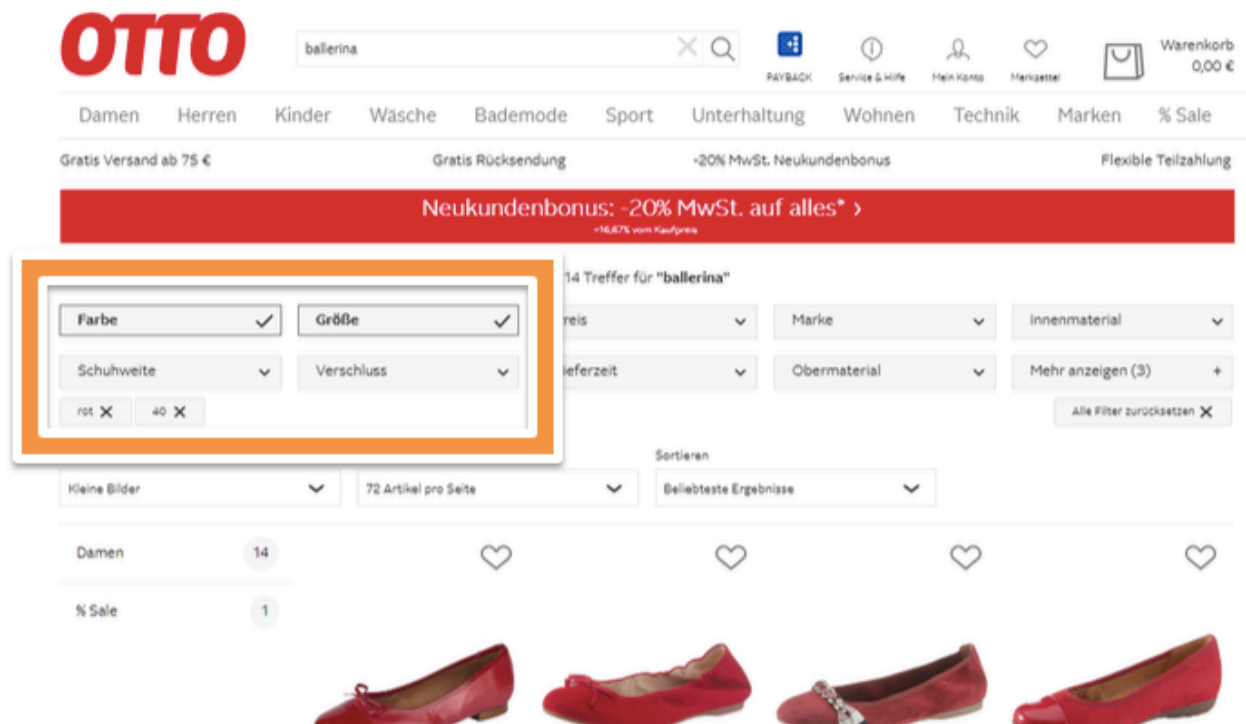


Abbildung 13: Suchanfrage im OTTO Online Shop mit gesetztem Farb- und Größenfilter.
Quelle: OTTO GmbH, 2020

Abbildung 14 zeigt das Landmark „suche“ und dessen Landmark-Attribute und Attribut-Werte, welches beim Ausführen einer Suchanfrage im Online Shop gefeuert wird. Es kann exakt das Verhalten der oben dargestellten Suchanfrage inkl. Filterung nachvollzogen werden.

Für die Erläuterung und Nachvollziehung der Filternutzung sind insbesondere die Attribute „mkz“, „stern“, „ssize“ und „filter.filter_color“ relevant. Dabei bezeichnet „mkz“ das sogenannte Marktkennzeichen, welches für die führende Sortimentsgruppe des Suchergebnisses steht (z.B. Garten, Möbel, Schuhe etc.). „stern“ beschreibt die Such-Query (z.B. „ballerina“), „ssize“ die gefilterte Größe (z.B.

„40“) und „filter.filter_color“ die gefilterte Farbe (z.B. „rot“). Die entsprechenden Filter-Attribute werden lediglich gefeuert, wenn sie im Shop-Frontend auch gesetzt werden.

```
seq 17 12/08/2020 16:28:01 lm suche cli 1017
categories = 1,0.0;2,0.0;3,0.0;4,100.0;5,0.0;6,0.0;7,0.0;8,0.0;9,0.0;10,0.0
filter.filter_color = rot
filtered = 2
FP_NITRO = 1.eacde02d-83a1-4237-8fbc-65dc9005d921.0.1393d3fb-77b6-443a-8811-
micro_caching = 0
mkz = 509
pagenu = 0
pc = Suchergebnis
pc2 = Suchergebnis
productIdsOnPage =
AKLBB473706576,AKLBB976468703,AKLBB942241966,AKLBB843536664,AKLBB738053030,AKLBB
sanz = 14
search-id = d89dbcdc-120200812
setLastShopName = Suche
setLastShopPath = Suche
setLastSourceDetailSequence = 1
slocale = de
spreis = -
ssize = 40
ssort = topseller
stern = ballerina
sview = intern
template_default = gallery_1
webservice_id = busca17_1597242481253_32
```

**Abbildung 14: OTTO Eigenentwicklung "LMTestTool" zum Nachvollziehen des Suchverhaltens.
Quelle: OTTO GmbH, 2020**

Grundsätzlich sind alle weiteren Filter auf den Produktlisten und Suchergebnisseiten ebenfalls vertrackt. Im Rahmen dieser Arbeit wurden allerdings nur jene Filter zur Betrachtung ausgewählt, die auch Business-seitig einen Impact in Bezug auf relevante Personalisierungs-Use Cases der Datenerhebung und -weiterverarbeitung haben können.

4.3 Beschreibung der Auswertung der Filternutzungs-Daten

Um Analysen über die allgemeine Filternutzung basierend auf Sortiments-Gruppen des OTTO Online Shops tätigen zu können, wurden gesammelte Tracking-Daten über den Zeitraum von 1. bis 31. Juli 2020 aggregiert. In die Datenerhebung

miteinanderbezogen sind dabei die Devices App, Mobile, Tablet und Desktop. Über eine SQL-Abfrage (siehe Quellcode 1 in Anhang A.1.) werden die entsprechenden Daten konsolidiert und an Excel zurück gegeben.

Diese initiale Abfrage liefert folgende Informationen, in Form einer Pivot-Tabelle dargestellt, zurück:

Major-Sortiment aus Suchanfrage	Anzahl der Suches gesamt	Suches		Suches m.		Suches m.		Suches m.		Suches m.		Suches m. "Für wen?"		Suches m. Verfüg- barkeitsfilter	
		m. Filter	... in %	Markenfilter	... in %	Preisfilter	.. In %	Größenfilter	...in %	Farbfilter	...in %	Filter	...in %	barkeitsfilter	...in %
Bad	34991	2840	8,12%	444	15,63%	569	20,04%	210	7,39%	1766	62,18%	7	0,25%	5	0,18%
Badebekleidung	49536	4765	9,62%	860	18,05%	417	8,75%	2783	58,41%	1452	30,47%	74	1,55%	13	0,27%
Büro/Foto/Telekom	53388	2037	3,82%	1073	52,68%	913	44,82%	0	0,00%	245	12,03%	0	0,00%	13	0,64%
Camping	2471	72	2,91%	38	52,78%	29	40,28%	0	0,00%	6	8,33%	0	0,00%	1	1,39%
DOB	170019	18499	10,88%	2073	11,21%	2214	11,97%	10828	58,53%	6455	34,89%	394	2,13%	39	0,21%
Garten	36848	1498	4,07%	427	28,50%	604	40,32%	0	0,00%	534	35,65%	0	0,00%	8	0,53%
Gesundheit	11642	384	3,30%	256	66,67%	103	26,82%	0	0,00%	33	8,59%	0	0,00%	0	0,00%
Groß-Elektro	43375	4412	10,17%	2016	45,69%	1880	42,61%	0	0,00%	918	20,81%	0	0,00%	10	0,23%
HAKA	24943	2836	11,37%	585	20,63%	433	15,27%	1413	49,82%	832	29,34%	159	5,61%	2	0,07%
Haushalt	33002	1106	3,35%	460	41,59%	212	19,17%	1	0,09%	473	42,77%	0	0,00%	3	0,27%
Haustext	35103	2728	7,77%	304	11,14%	942	34,53%	0	0,00%	1649	60,45%	0	0,00%	5	0,18%
Heimtext	40946	4461	10,89%	306	6,86%	724	16,23%	0	0,00%	3765	84,40%	0	0,00%	8	0,18%
Heimwerker	17969	809	4,50%	344	42,52%	170	21,01%	0	0,00%	318	39,31%	10	1,24%	4	0,49%
Klein-Elektro	32619	1568	4,81%	920	58,67%	539	34,38%	0	0,00%	185	11,80%	0	0,00%	2	0,13%
KOB	11562	820	7,09%	88	10,73%	109	13,29%	452	55,12%	109	13,29%	157	19,15%	0	0,00%
Lederwaren	16598	1711	10,31%	425	24,84%	587	34,31%	0	0,00%	740	43,25%	117	6,84%	6	0,35%
Leuchten	6551	447	6,82%	29	6,49%	195	43,62%	0	0,00%	260	58,17%	0	0,00%	4	0,89%
Möbel	167514	14149	8,45%	1120	7,92%	6478	45,78%	0	0,00%	8236	58,21%	45	0,32%	38	0,27%
Parfum/Kosmetik	3311	77	2,33%	62	80,52%	7	9,09%	0	0,00%	6	7,79%	3	3,90%	1	1,30%
Schuhe	31608	4594	14,53%	633	13,78%	266	5,79%	2659	57,88%	1820	39,62%	153	3,33%	12	0,26%
Spielwaren / Musik	13073	238	1,82%	94	39,50%	112	47,06%	0	0,00%	33	13,87%	3	1,26%	5	2,10%
Sport/Fahrrad	17201	586	3,41%	200	34,13%	319	54,44%	14	2,39%	41	7,00%	19	3,24%	7	1,19%
Sportbekleidung	39831	4950	12,43%	946	19,11%	615	12,42%	2856	57,70%	1142	23,07%	268	5,41%	15	0,30%
Sportschuhe	31436	3155	10,04%	727	23,04%	279	8,84%	1745	55,31%	982	31,13%	60	1,90%	5	0,16%
Uhren/Schmuck	24656	1655	6,71%	338	20,42%	637	38,49%	35	2,11%	741	44,77%	74	4,47%	6	0,36%
Unterhaltungselektro	24049	1074	4,47%	646	60,15%	382	35,57%	0	0,00%	114	10,61%	0	0,00%	3	0,28%
Wäsche	40942	4090	9,99%	641	15,67%	311	7,60%	2628	64,25%	1276	31,20%	44	1,08%	7	0,17%
Gesamtergebnis	1015184	85561	8,43%	16055	18,76%	20046	23,43%	25624	29,95%	34131	39,89%	1587	1,85%	222	0,26%

Tabelle 3: Ausgabe der Filternutzungsrate auf Sortiments-Gruppen Ebene in Form einer Pivot-Tabelle.

Tabelle 3 bildet eine Aufstellung der Anzahl an durchgeführten Suchanfragen im untersuchten Zeitraum und der Filternutzung im allgemeinen Überblick ab. So wurden im Juli 2020 im OTTO Online Shop mehr als eine Millionen Suchanfragen durchgeführt. Diese Suchanfragen wurden aufgesplittet auf die führenden Marktkennzeichen (= Sortimentsgruppen) innerhalb eines Suchergebnisses. Dadurch lässt sich als erster grober Überblick ableiten, welche Filter innerhalb welcher Sortimentsgruppe wie relevant sind. Auf eine detaillierte Interpretation und Beschreibung dieser Informationen wird im Kapitel 5 eingegangen.

Um weitere Erkenntnisse zu den einzelnen Filterwerten der untersuchten Filter zu erhalten, wurde die unter A.1. angeführte SQL-Abfrage erweitert (siehe A.2.). Diese

Erweiterung des SQL-Codes führt dazu, dass allesamt ausgewählten Filterwerte innerhalb eines Filters ebenfalls von der Datenbank zurückgegeben werden. Auch hier werden detailliertere Analysen und Interpretationen in Kapitel 5 durchgeführt und erläutert. Dazu wurden neben Aggregationsanalysen auf Filterwert-Ebene auch Clusteranalysen durchgeführt, um die statistische Signifikanz der Zusammenhänge zu ermitteln. Diese Clusteranalysen basieren auf dem von der Visualisierungs-Suite *Tableau* verwendeten k-Means-Algorithmus, welcher die Daten in k-Cluster unterteilt.

5. Ergebnisse und Interpretation der empirischen Forschung

In diesem Kapitel werden die gesammelten Daten analysiert, grafisch ausgewertet und interpretiert. Um basierend auf den Ergebnissen der empirischen Forschung auch relevante Ableitungen treffen zu können, geht die Autorin auf verschiedene Fragestellungen ein, die nachfolgend in Bezug auf die Filternutzung und das Verhalten der User*innen beantwortet werden.

5.1 Wie hoch ist die Filternutzungsrate allgemein?

Im Allgemeinen ist erkennbar, dass die durchschnittliche Filternutzungsrate der untersuchten Filter im OTTO Online Shop auf Suchergebnisseiten im Monat Juli 2020 bei 8,43 % liegt. Es wurde also in 8,43 % aller durchgeführten Suchanfragen auch mindestens einer der untersuchten Filter gesetzt (siehe Tabelle 3). Zu den untersuchten Filtern zählen der Farb-, Geschlechts-, Größen-, Lieferzeit-, Preis-, und Marken-Filter.

5.2 Sind Sortimente erkennbar, bei denen Filter relevanter bzw. weniger relevant sind?

In Suchanfragen rund um die Sortimente Schuhe, Sportbekleidung, HAKA (Herren- & Knabenmode), DOB (Damenoberbekleidung), KOB (Kinderoberbekleidung), Lederwaren, Sportschuhe, Wäsche und Badebekleidung haben zwischen 10-14 % der User*innen bei ihrer Suchanfrage auch mindestens einen Filter zur besseren Navigation gesetzt (siehe auch Tabelle 3).

Einen detaillierteren Blick auf die Relevanz einzelner Filter innerhalb verschiedener Sortimente veranschaulicht eine Farbskala – angewendet auf die Filternutzungsraten (Tabelle 4). Dabei lassen sich folgende Aussagen in Bezug auf die Sortimente und der Relevanz von verschiedenen Filtern ableiten.

Insbesondere in den Fashion-Sortimenten haben die untersuchten Filter eine hohe Relevanz.

Major-Sortiment aus Suchanfrage	Anzahl der Suchanfragen gesamt	Markenfilter in %	Preisfilter in %	Größenfilter in %	Farbfilter in %	"Für wen?-Filter in %	Verfügbarkeitsf. In %
Bad	34991	15,63%	20,04%	7,39%	62,18%	0,25%	0,18%
Badebekleidung	49536	18,05%	8,75%	58,41%	30,47%	1,55%	0,27%
Büro/Foto/Telekom	53388	52,68%	44,82%	0,00%	12,03%	0,00%	0,64%
Camping	2471	52,78%	40,28%	0,00%	8,33%	0,00%	1,39%
DOB	170019	11,21%	11,97%	58,53%	34,89%	2,13%	0,21%
Garten	36848	28,50%	40,32%	0,00%	35,65%	0,00%	0,53%
Gesundheit	11642	66,67%	26,82%	0,00%	8,59%	0,00%	0,00%
Groß-Elektro	43375	45,69%	42,61%	0,00%	20,81%	0,00%	0,23%
HAKA	24943	20,63%	15,27%	49,82%	29,34%	5,61%	0,07%
Haushalt	33002	41,59%	19,17%	0,09%	42,77%	0,00%	0,27%
Haustex	35103	11,14%	34,53%	0,00%	60,45%	0,00%	0,18%
Heimtex	40946	6,86%	16,23%	0,00%	84,40%	0,00%	0,18%
Heimwerker	17969	42,52%	21,01%	0,00%	39,31%	1,24%	0,49%
Klein-Elektro	32619	58,67%	34,38%	0,00%	11,80%	0,00%	0,13%
KOB	11562	10,73%	13,29%	55,12%	13,29%	19,15%	0,00%
Lederwaren	16598	24,84%	34,31%	0,00%	43,25%	6,84%	0,35%
Leuchten	6551	6,49%	43,62%	0,00%	58,17%	0,00%	0,89%
Möbel	167514	7,92%	45,78%	0,00%	58,21%	0,32%	0,27%
Parfum/Kosmetik	3311	80,52%	9,09%	0,00%	7,79%	3,90%	1,30%
Schuhe	31608	13,78%	5,79%	57,88%	39,62%	3,33%	0,26%
Spielwaren / Musik	13073	39,50%	47,06%	0,00%	13,87%	1,26%	2,10%
Sport/Fahrrad	17201	34,13%	54,44%	2,39%	7,00%	3,24%	1,19%
Sportbekleidung	39831	19,11%	12,42%	57,70%	23,07%	5,41%	0,30%
Sportschuhe	31436	23,04%	8,84%	55,31%	31,13%	1,90%	0,16%
Uhren/Schmuck	24656	20,42%	38,49%	2,11%	44,77%	4,47%	0,36%
Unterhaltungselektronik	24049	60,15%	35,57%	0,00%	10,61%	0,00%	0,28%
Wäsche	40942	15,67%	7,60%	64,25%	31,20%	1,08%	0,17%
Gesamtergebnis	1015184	18,76%	23,43%	29,95%	39,89%	1,85%	0,26%

Tabelle 4: Farbskala der Filternutzungsraten, angewendet auf die einzelnen Sortimente und beleuchteten Filter.

Im Durchschnitt sind Filter im Sortiment Parfum / Kosmetik eher unwichtiger, hier sticht allerdings der Markenfilter mit einer besonders hohen Nutzungsrate heraus. Auch in Suchanfragen rund um Heimtex-Artikel sind Filter eher irrelevant, eine Ausnahme stellt hier der Farbfilter mit einer Nutzungsrate von 84,4 % dar.

Des Weiteren ist auffällig, dass in allen Sortimenten der Geschlechtsfilter sowie der Lieferverfügbarkeitsfilter irrelevant sind. Auch der Größenfilter ist außerhalb der Fashion-Sortimente für die User*innen weniger wichtig. Auf eine detailliertere Interpretation wird unter Abschnitt 5.4 eingegangen.

5.3 Sind Suchbegriffe erkennbar, bei denen Filter relevanter bzw. weniger relevant sind?

Zur Beantwortung dieser Frage wird mittels der Visualisierungs-Suite *Tableau* eine Clusteranalyse durchgeführt. Dabei werden zum einen die Anzahl der Sessions mit Suchanfrage und zum anderen die insgesamt gesetzten Filter pro Suchbegriff aggregiert. Untersucht werden pro Suchbegriff die Summen der sechs untersuchten Filter (Farbe, Größe, Marke, Preis, Geschlecht und Lieferzeit) und in Beziehung mit der Anzahl der Sessions und der Anzahl gesamt gesetzten Filter in einem Streudiagramm dargestellt. Unterteilt man dieses Streudiagramm mit einem k-Means-Algorithmus in Cluster, ergibt sich folgendes Diagramm in Abbildung 15.

Clusteranalyse Suchbegriffe

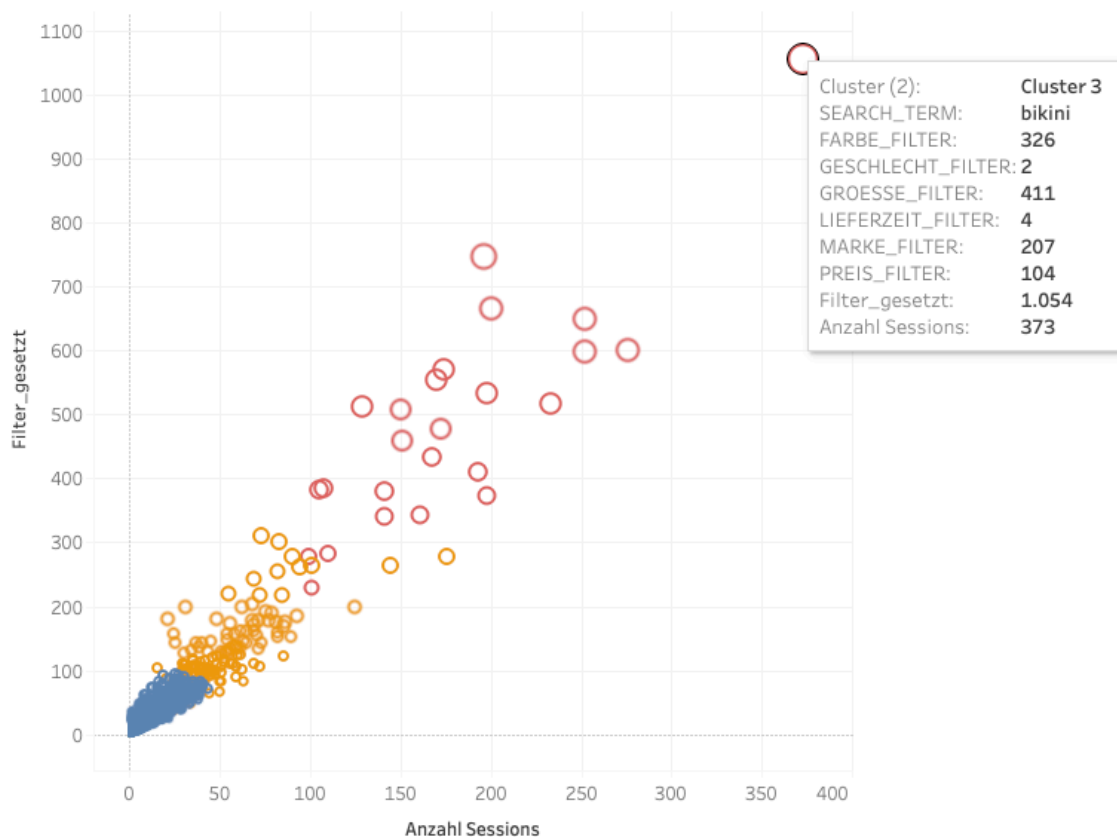


Abbildung 15: Ermittelte Cluster der Suchbegriffe basierend auf den gesetzten Filtern pro Session.

Der k-Means-Algorithmus ermittelt dabei 3 verschiedene Cluster. Die Größe der einzelnen Kreise stellt dar, wie oft für einen Suchbegriff insgesamt Filter gesetzt wurden – je größer der Kreis, desto mehr Filterwerte wurden von den Nutzer*innen im Juli 2020 gewählt.

Der rote Cluster mit 25 Einheiten beinhaltet die Top Suchbegriffe nach Sessions, also jene die am häufigsten gesucht und bei denen gleichzeitig am häufigsten Filter gesetzt wurden. Diese können in Bezug auf die Filternutzung als die relevantesten Suchbegriffe identifiziert werden. Abbildung 15 betrachtet auch einen Suchbegriff beispielhaft im Detail. Dabei handelt es sich um die Search-Query „bikini“, nach der im Beobachtungszeitraum insgesamt in 373 Sessions gesucht und insgesamt 1.054 Mal ein Filterwert in den untersuchten Filtern von User*innen ausgewählt wurde. Außerdem ist ablesbar, dass der Größen- und der Farbfilter die relevantesten Filter für diesen Suchbegriff sind. Diese sollten in der Filter-Personalisierungsstrategie des Suchbegriffs „bikini“ besonders beachtet werden. Tabelle 5 listet die Suchbegriffe aus dem roten Cluster auf.

Cluster (2)	SEARCH_TERM	Anzahl Sessions ▼	Filter_gesetzt	FARBE...	GESCHLECHT...	GROESSE...	LIEFERZEIT...	MARKE...	PREIS...
Cluster 3	bikini	373	1.054	326	2	411	4	207	104
Cluster 3	waschmaschine	276	600	37	0	0	1	284	278
Cluster 3	bettwäsche	252	599	326	0	0	0	54	219
Cluster 3	dirndl	252	650	97	1	392	0	30	130
Cluster 3	T-Shirts	233	517	64	167	59	0	215	12
Cluster 3	große großen damenhosen	200	666	138	0	439	0	55	34
Cluster 3	couchtisch	198	533	328	0	0	0	8	197
Cluster 3	Dirndl	198	373	32	0	77	0	237	27
Cluster 3	teppich	196	746	559	0	0	1	29	157
Cluster 3	kühlschrank	193	409	98	0	0	1	85	225
Cluster 2	birkenstock	176	278	71	0	198	0	0	9
Cluster 3	kleider	174	571	133	0	286	2	35	115
Cluster 3	schuhe	172	477	103	4	199	2	131	38
Cluster 3	boxspringbett	170	553	264	0	0	2	16	271
Cluster 3	geschirrspüler	167	433	83	0	0	1	174	175
Cluster 3	Teppiche	161	342	300	0	0	0	38	4
Cluster 3	bh	151	458	141	0	258	2	31	26
Cluster 3	kommode	150	507	269	0	0	2	44	192
Cluster 2	skechers	144	264	71	1	161	1	27	3
Cluster 3	home affaire	141	379	38	0	0	0	333	8
Cluster 3	lascana	141	340	88	0	201	0	25	26
Cluster 3	ecksofa	129	513	193	0	0	0	37	283
Cluster 2	comma	125	200	22	0	14	0	161	3
Cluster 3	schreibtisch	110	283	145	0	0	0	9	129
Cluster 3	kleiderschrank	107	384	218	0	0	0	21	145
Cluster 3	sofa	105	382	122	0	0	1	31	228
Cluster 3	laptop	101	229	8	1	0	1	70	149
Cluster 2	sideboard	101	265	145	0	0	0	39	81
Cluster 3	wohnlandschaft	99	277	98	0	0	1	13	165

Tabelle 5: Die relevantesten Suchbegriffe in Bezug auf die Filternutzung auf ottoversand.at.

Im blauen Cluster in Abbildung 15 sind jene Suchbegriffe zusammengefasst, die nur eine geringe Filternutzung und selten in User-Sessions gesucht wurden. Er ist zugleich mit 13.500 Suchbegriffen der größte Cluster. Der orange Cluster bildet die Mitte dieser beiden Faktoren und umfasst 269 untersuchte Suchbegriffe.

Tabelle 6 beschreibt das Streudiagramm in Abbildung 15 mit deskriptiven Daten. Durch die Varianzanalyse (ANOVA), welche automatisch von *Tableau* durchgeführt wird, konnte gezeigt werden, dass es einen statistisch signifikanten Unterschied ($p = 0.000$) im Zusammenhang der Suchbegriffe zwischen allen untersuchten Filtern gibt.

Analyse der Varianz:

Variable	F-Statistik	p-Wert	Modell		Fehler	
			Quadratsumme	DF	Quadratsumme	DF
Filter_gesetzt	4792.0	0.0	6.478	2	9.322	13791
Anzahl Sessions	4548.0	0.0	7.074	2	10.73	13791
Summe von PREIS_FILTER	3651.0	0.0	6.185	2	11.68	13791
Summe von LIEFERZEIT_FILTER	3252.0	0.0	6.629	2	14.06	13791
Summe von FARBE_FILTER	3187.0	0.0	2.655	2	5.746	13791
Summe von MARKE_FILTER	2271.0	0.0	2.142	2	6.505	13791
Summe von GROESSE_FILTER	1491.0	0.0	1.538	2	7.112	13791
Summe von GESCHLECHT_FILTER	179.6	0.0	0.0462	2	1.774	13791

Tabelle 6: Analyse der Varianz des Clusterings der Suchbegriffe.

5.4 Welche Erkenntnisse lassen sich aus dem Filternutzungsverhalten der User*innen ableiten?

In diesem Abschnitt werden die sechs untersuchten Filter Markenfilter, Preisfilter, Größenfilter, Farbfilter, Geschlechtsfilter und Lieferzeitenfilter genauer beleuchtet und mit unterschiedlichen Faktoren und Blickweisen in Beziehung gesetzt. Dabei werden folgende Fragen je Filter beantwortet und mit Interpretationen zu den Erkenntnissen und Analysen untermauert:

- Wie häufig werden die verschiedenen Filter innerhalb der Suchbegriffe bzw. Sortimente auf Suchergebnisseiten gesetzt?
- Welche Filterwerte innerhalb eines Filters werden am häufigsten ausgewählt?
- Wie häufig kommen Mehrfach-Auswahlen von Filterwerten vor?
- Sind verschiedene Segmente oder Cluster in Bezug auf die Nutzung von Filter und Filterwerten auf Suchergebnisseiten erkennbar?

5.4.1 Markenfilter

Tabelle 7 betrachtet den Markenfilter genauer. Dieser macht in absoluten Zahlen einen Anteil von 16.055 Searches (= 18,76 %) der Gesamtanzahl an Suchanfragen mit gesetztem Filter aus.

Major-Sortiment aus Suchanfrage	Anzahl der Searches gesamt	Searches m. Filter	Searches m. Markenfilter	... in%
Bad	34991	2840	444	15,63%
Badebekleidung	49536	4765	860	18,05%
Büro/Foto/Telekom	53388	2037	1073	52,68%
Camping	2471	72	38	52,78%
DOB	170019	18499	2073	11,21%
Garten	36848	1498	427	28,50%
Gesundheit	11642	384	256	66,67%
Groß-Elektro	43375	4412	2016	45,69%
HAKA	24943	2836	585	20,63%
Haushalt	33002	1106	460	41,59%
Haustex	35103	2728	304	11,14%
Heimtex	40946	4461	306	6,86%
Heimwerker	17969	809	344	42,52%
Klein-Elektro	32619	1568	920	58,67%
KOB	11562	820	88	10,73%
Lederwaren	16598	1711	425	24,84%
Leuchten	6551	447	29	6,49%
Möbel	167514	14149	1120	7,92%
Parfum/Kosmetik	3311	77	62	80,52%
Schuhe	31608	4594	633	13,78%
Spielwaren / Musik	13073	238	94	39,50%
Sport/Fahrrad	17201	586	200	34,13%
Sportbekleidung	39831	4950	946	19,11%
Sportschuhe	31436	3155	727	23,04%
Uhren/Schmuck	24656	1655	338	20,42%
Unterhaltungselektronik	24049	1074	646	60,15%
Wäsche	40942	4090	641	15,67%
Gesamtergebnis	1015184	85561	16055	18,76%

Tabelle 7: Nutzungsrate des Markenfilters über alle führenden Sortimente innerhalb von Suchanfragen.

Dabei ist erkennbar, dass der Markenfilter insbesondere bei Suchanfragen rund um Parfum/Kosmetik mit mehr als 80 % Filternutzungsrate äußerst wichtig für die User*innen ist. Auch bei Suchen nach Produkten in den Segmenten Gesundheit, Unterhaltungselektronik, Klein-Elektro, Camping und Büro/Foto/Telekom wird der Markenfilter bei mehr als 50 % der Suchanfragen gesetzt. Hingegen ist im Bereich Living, Heimtex und Möbel eine Filterung basierend auf verschiedenen Marken relativ gesehen weniger relevant – hier liegt die Rate unter 10 %.

Eine interessante Feststellung in Bezug auf die Relevanz und die Klick-Performance des Markenfilters ist, dass dieser in absoluten Zahlen gesehen innerhalb der verschiedenen Sortimente eine andere Wichtigkeit hat, als in der relativen Dimension. Denn absolut sind die Sortimente „Damenoberbekleidung“ und „Groß-Elektro“ jene Sortimente, in denen der Markenfilter am häufigsten von User*innen genutzt wird (siehe Abbildung 16).

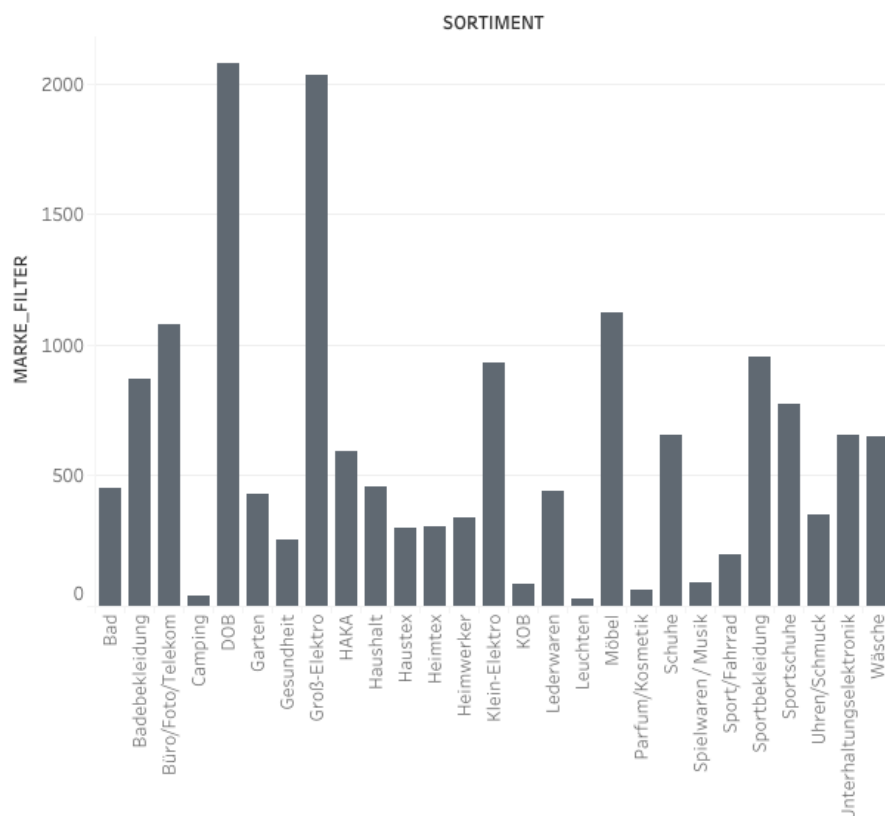


Abbildung 16: Absolute Anteile des Markenfilters innerhalb der Sortimente.

Stellt man die Filternutzung in Prozent dar, setzten nur 11,21 % der Nutzenden einen Marken-Filterwert im DOB-Segment. Im Gegensatz dazu nutzten 45,69 % bei Such-Queries rund um das Groß-Elektro-Sortiment den Markenfilter (Tabelle 7). Diese Verschiebung der absoluten Anteile im Vergleich zu den relativen ist dadurch erklärbar, dass wie im Beispiel des Damenfashion-Sortiments eine sehr hohe Anzahl an Suchanfragen innerhalb des DOB-Sortiments ausgeführt wird (insgesamt mehr als 170.000 Suchanfragen), jedoch bei nur knapp 11 % an Fashion-Suchbegriffen überhaupt irgendeiner der untersuchten Filter gesetzt wird. So sind prozentual gesehen 2000 gesetzte Markenfilter ein relativ geringer Anteil im Vergleich zu 2000 gesetzten Markenfilter im Groß-Elektro-Sortiment, wo bei nur 4400 Suchanfragen mindestens ein Filter von den User*innen gesetzt wurde.

Diese differenzierte Betrachtungsweise zwischen den absoluten und relativen Anteilen des Markenfilters innerhalb verschiedener Sortimente sollte auch bei weiteren Implikationen in Bezug auf mögliche Personalisierungs-Strategien des Markenfilters beachtet werden.

Für detailliertere Ableitungen in Bezug auf die Relevanz einzelner Filterwerte werden in Tabelle 8 die 30 am häufigsten ausgewählten Values des Markenfilters aufgelistet. Dabei ist erkennbar, dass insbesondere die Möbel-Marke „Home Affaire“ eine sehr relevante Marke ist. Auch verschiedene Technik-Marken wie Samsung, Miele, BOSCH, SIEMENS und AEG haben einen großen Anteil an den am häufigsten gesetzten Filterwerten. Im Fashion-Sortiment gehören LASCANA, Comma oder TOMMY HILFIGER zu den Top-Marken.

	Filterwert des Markenfilters	Häufigkeit des Filterwerts
1	⊕ Home affaire	762
2	⊕ Samsung	727
3	⊕ LASCANA	582
4	⊕ Miele	552
5	⊕ BOSCH	417
6	⊕ s.Oliver	398
7	⊕ SIEMENS	356
8	⊕ Comma	349
9	⊕ AEG	262
10	⊕ TOMMY HILFIGER	249
11	⊕ Apple	227
12	⊕ my home	227
13	⊕ WMF	210
14	⊕ LG	208
15	⊕ MarJo	198
17	⊕ Einhell	183
18	⊕ HP	183
19	⊕ Skechers	159
21	⊕ adidas Performance	158
22	⊕ Bench.	149
23	⊕ Nike	142
24	⊕ Buffalo	133
25	⊕ adidas Originals	131
26	⊕ Hanseatic	130
27	⊕ Jack Wolfskin	129
28	⊕ Tom Tailor	119
29	⊕ s.Oliver Beachwear	118
30	⊕ GORENJE	115

Tabelle 8: Top 30 Filterwerte des Markenfilters.

Tabelle 9 zeigt die häufigsten Marken-Filterwerte und die dazugehörigen Suchbegriffe, mit der Einschränkung, dass bei einem Suchbegriff mindestens 30x der entsprechende Filterwert gesetzt wurde. Diese Einschränkung wird vorgenommen, um die relevantesten Sortimente bzw. Suchbegriffe herauszuheben. Dabei lässt sich erkennen, dass einige Filter-Werte der Top 30 aus Tabelle 8 rausfallen bzw. sich die Ränge verschieben.

Zeilenbeschriftungen	IT	Summe von MARKE_FILTER
[-] Home affaire		500
home affaire		421
Sideboard		40
kommode		39
[-] Comma		283
Comma		249
Jumpsuits		34
[-] Miele		253
staubsauger		119
Waschmaschinen		96
Einbau-Geschirrspüler		38
[-] my home		173
my home		173
[-] Samsung		163
Smart-TV		62
kühlschränke		37
tablets		33
fernseher		31
[-] MarJo		145
dirndl		145
[-] WMF		132
Topfsets		132
[-] Einhell		109
akku-rasenmäher		109
[-] s.Oliver		109
s.oliver		109
[-] Smeg		73
kühlschränke		73
[-] EUFAB		66
fahrradträger		66
[-] Andreas Gabalier Kollektion		59
dirndl		59
[-] Step by Step		59
Schultaschen		59
[-] Smoby		56
Spielhaus		56
[-] De'Longhi		50
kaffeevollautomat		50
[+] Vans		48
[+] Hammerschmid		45
[+] Apple		43
[+] Garmin		43
[+] MERXX		42
[+] Skechers		42
[+] Sportstech		41
[+] LASCANA		41
[+] Bench.		40
[+] s.Oliver Beachwear, Buffalo, Venice Beach, Sunseeker, KangaROOS, Bench., Sunflair, JETTE, Bruno Banani, Chiemsee, Olympia, PETITE FLEUR, Homeboy		40
[+] Asics		40
[+] BOSCH		39
[+] SIEMENS		37
[+] Tefal		35
[+] HP		32
[+] KitchenAid		31
[+] AEG		31
[+] Viita		30

Tabelle 9: Filterwerte des Markenfilters bei Suchbegriffen mit mindestens 30 Mal gesetztem Filterwert.

Der nachfolgende Ausschnitt der Mehrfach-Auswahlen von Marken-Filterwerten in Tabelle 10 lässt erkennen, dass Sportmarken wie Adidas Performance & Adidas Original, Nike Sportswear & Nike, Lascana & Lascana Active häufig miteinander gesetzt werden und dabei eine gewisse Markentreue herrscht. Es werden also nur selten beispielsweise Nike- und Adidas-Artikel gemeinsam gefiltert. Bei den

Technik-Marken wird häufig gemeinsam nach Siemens- & Bosch-Produkten oder nach Samsung- & LG-Artikel gefiltert.

Zeilenbeschriftungen	Summe von MARKE_FILTER
Gesamtergebnis	6272
adidas Performance,adidas Originals	106
adidas Originals,adidas Performance	62
SIEMENS,BOSCH	55
TOMMY HILFIGER,TOMMY JEANS	45
LASCANA,LASCANA ACTIVE	44
BOSCH,SIEMENS	42
s.Oliver Beachwear,Buffalo,Venice Beach,Sunseeker,KangaROOS,Bench.,Sunflair,JETTE,Bruno Banani	40
Samsung,LG	32
adidas Performance,adidas Originals,adidas	29
Sheego,sheego by Joe Browns,sheego by Miyabi Kawai	26
Rieker,Gabor,Gabor Comfort	25
Nike,Nike Sportswear,Nike SB	24
Salonloewe,wash+dry by Kleen-Tex	22
Esprit,edc by Esprit,Esprit Collection	21
FREDSBRUDER,TOMMY HILFIGER,Bruno Banani, Tom Tailor,Liebeskind Berlin,Marc O'Polo,Joop!,Desig	21
Nike Sportswear,Nike	21
adidas Originals,adidas	19
Nike Sportswear,Nike,Nike SB	19
s.Oliver,s.Oliver BLACK LABEL	19
Home affaire,Home affaire Collection	18
adidas Performance,Nike,Puma,Nike Sportswear,Reebok,adidas Originals,Fila,Kappa,Bruno Banani,	17
BOSCH,Samsung	17
Nike,Nike Sportswear	17
Privileg,Privileg Family Edition	17
Home affaire,heine home	16
Only,ONLY CARMAKOMA	16
Vadobag,Lässig,Kidzroom,Fjällräven,Jack Wolfskin,Trespass,Liebeskind Berlin	16
adidas Performance,adidas TERREX	15
BOSCH,BEKO	15
Sheego,sheego by Joe Browns,sheego by Miyabi Kawai,Zizzi	15
Tom Tailor, Tom Tailor Denim	15
AEG,BOSCH	14
Country Line ,Love Nature	14
FREDSBRUDER,bugatti,Liebeskind Berlin,WouWou,Landleder,BODENSCHATZ,Bruno Banani,Marc O'P	14
LASCANA,Buffalo,Melrose	14
LASCANA,Buffalo,Sunseeker,Venice Beach,Sunflair	14
Levi's®,Nike,Puma,Bench.,adidas Originals,Fila,Champion	14

Tabelle 10: Mehrfach-Auswahlen von Filterwerten des Markenfilters.

Abbildung 17 stellt eine Baumkarte dar, die drei Cluster der gesetzten Marken-Filterwerte abbildet. Die Cluster in roter, oranger und blauer Farbe beschreiben jeweils die Häufigkeit der gesetzten Filterwerte basierend auf den von User*innen eingegebenen Suchbegriffen. Der rote Cluster bildet die beiden häufigsten

Filterwerte innerhalb des Markenfilters ab, der orange Cluster erklärt Filterwerte für Suchbegriffe, die verhältnismäßig oft gesetzt wurden und bei dem blauen Cluster wurden die Filter-Values seltener ausgewählt bzw. zum Teil sogar nur einmal pro Suchbegriff gesetzt.

Markencluster

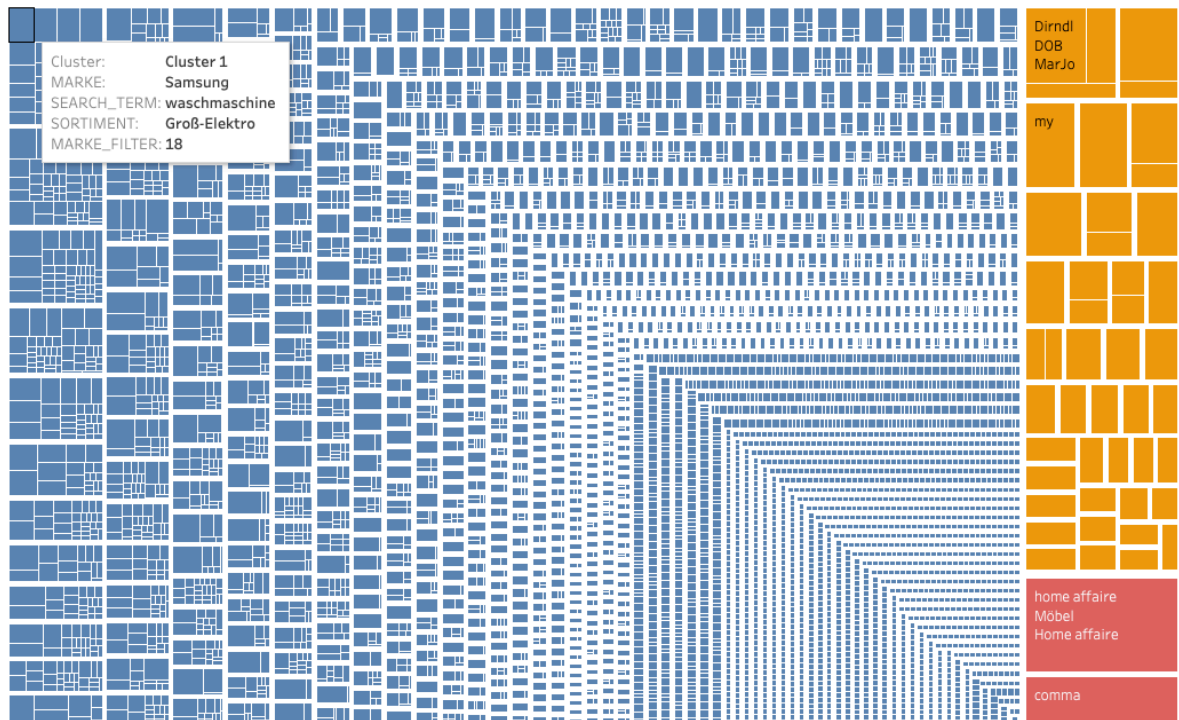


Abbildung 17: Baumkarte mit drei identifizierten Markenfilter-Clustern.

Die geclusterte Baumkarte kann nun also so gelesen und interpretiert werden, dass einzelne Segmente (getrennt durch eine dickere, weiße Linie) die unterschiedlichen Suchbegriffe darstellen. Im linken oberen Eck im blauen Cluster wird zum Beispiel die Einheit bzw. der Suchbegriff „waschmaschine“ dargestellt, wobei der Markenfilter insgesamt 18 Mal bei diesem Suchbegriff auf „Samsung“ gesetzt wurde. Direkt rechts daneben schließt die Kachel mit dem Filterwert „Privileg“ für den Suchbegriff „waschmaschine“ an, der insgesamt 17 Mal gesetzt wurde. Anhand der Größe dieser einzelnen markierten Kachel lässt sich auch erkennen, dass für den Suchbegriff „waschmaschine“ Samsung die am häufigsten gewählte Marke im Juli 2020 ist.

Je kleiner die Segmente und die einzelnen Kacheln innerhalb der Segmente werden, desto weniger unterschiedliche Marken-Filterwerte wurden für einen Suchbegriff gesetzt bzw. desto seltener wurden diese Filterwerte von User*innen ausgewählt. Anhand dieser Informationen ist einfach ablesbar, bei welchen Sortimenten bzw. Suchbegriffen eine segmentierte Gestaltung/Positionierung/Sortierung des Farbfilters Sinn machen kann.

Gibt eine Kund*in beispielsweise den Suchbegriff „dirndl“ ein (rechts oben im orangen Cluster), ist es relativ wahrscheinlich, dass auch eine Filterung nach der Marke „MarJo“ getätigt wird. In diesem Suchbegriffs-Segment sind auch noch die Marken „Andreas Gabalier Kollektion“ und „Hammerschmid“ sehr relevante Filter-Values, die bei einer Suche nach dem Begriff „dirndl“ im OTTO Online Shop beispielsweise personalisiert beworben oder besonders hervorgehoben werden können.

Blickt man auf die ANOVA der Clusteranalyse in Abbildung 17, kann gesagt werden, dass sich die Anzahl der gesetzten Marken-Filterwerte signifikant zwischen den identifizierten Clustern unterscheiden (siehe Tabelle 11).

Analyse der Varianz:

Variable	F-Statistik	p-Wert	Modell		Fehler	
			Quadratsumme	DF	Quadratsumme	DF
Summe von MARKE_FILTER	6500.0	0.0	1.818	2	2.759	19727

Tabelle 11: Varianzanalyse der Marken-Values-Custer.

5.4.2 Preisfilter

Die Nutzungsrate des Preisfilter auf Suchergebnisseiten wird in Tabelle 12 detaillierter betrachtet. Dabei lässt sich ableiten, dass auch der Preisfilter ein ähnliches Bild wie der Markenfilter darstellt, wenn man die relative Filternutzungsrate und die absoluten Klickzahlen von Filter-Werten in Beziehung zu einander setzt.

So ist im Sport- & Fahrrad-Sortiment der Preisfilter prozentuell im Vergleich zu den anderen Sortimenten am relevantesten. Denn in 54,44 % aller Suchanfragen wird in diesem Sortiment der Preisfilter gesetzt, absolut gesehen sind das 319 gefilterte Preis-Ranges bei Suchbegriffe rund um Sport & Freizeit. Auch bei Spielwaren & Musik, im Möbel- und Leuchten-Bereich, bei Suchanfragen rund um Groß-Elektro, Büro/Foto/Telekom, Garten und Camping ist der Preisfilter mit einer Filternutzungsrate von mehr als 40 % ein wichtiger Filter.

Major-Sortiment aus Suchanfrage	Anzahl der Searches gesamt	Searches m. Filter	Searches m. Preisfilter	.. In %
Bad	34991	2840	569	20,04%
Badebekleidung	49536	4765	417	8,75%
Büro/Foto/Telekom	53388	2037	913	44,82%
Camping	2471	72	29	40,28%
DOB	170019	18499	2214	11,97%
Garten	36848	1498	604	40,32%
Gesundheit	11642	384	103	26,82%
Groß-Elektro	43375	4412	1880	42,61%
HAKA	24943	2836	433	15,27%
Haushalt	33002	1106	212	19,17%
Haustex	35103	2728	942	34,53%
Heimtex	40946	4461	724	16,23%
Heimwerker	17969	809	170	21,01%
Klein-Elektro	32619	1568	539	34,38%
KOB	11562	820	109	13,29%
Lederwaren	16598	1711	587	34,31%
Leuchten	6551	447	195	43,62%
Möbel	167514	14149	6478	45,78%
Parfum/Kosmetik	3311	77	7	9,09%
Schuhe	31608	4594	266	5,79%
Spielwaren / Musik	13073	238	112	47,06%
Sport/Fahrrad	17201	586	319	54,44%
Sportbekleidung	39831	4950	615	12,42%
Sportschuhe	31436	3155	279	8,84%
Uhren/Schmuck	24656	1655	637	38,49%
Unterhaltungselektronik	24049	1074	382	35,57%
Wäsche	40942	4090	311	7,60%
Gesamtergebnis	1015184	85561	20046	23,43%

Tabelle 12: Nutzungsrate des Preisfilters über alle führenden Sortimente innerhalb von Suchanfragen.

Vermeintlich weniger relevant ist der Preisfilter in den Sortimenten DOB, Sportschuhe, Badebekleidung, Wäsche und Schuhe mit einer Nutzungsrate von kleiner 10 %. Blickt man aber beispielsweise im DOB-Sortiment wiederum auf die Nutzungszahlen, zeigt sich, dass mit über 2200 gesetzten Preis-Filterwerten die absolute Nutzung der Preisfilterung im Damenoberbekleidungs-Sortiment am zweithöchsten über alle Sortimente hinweg ist (siehe Abbildung 18).

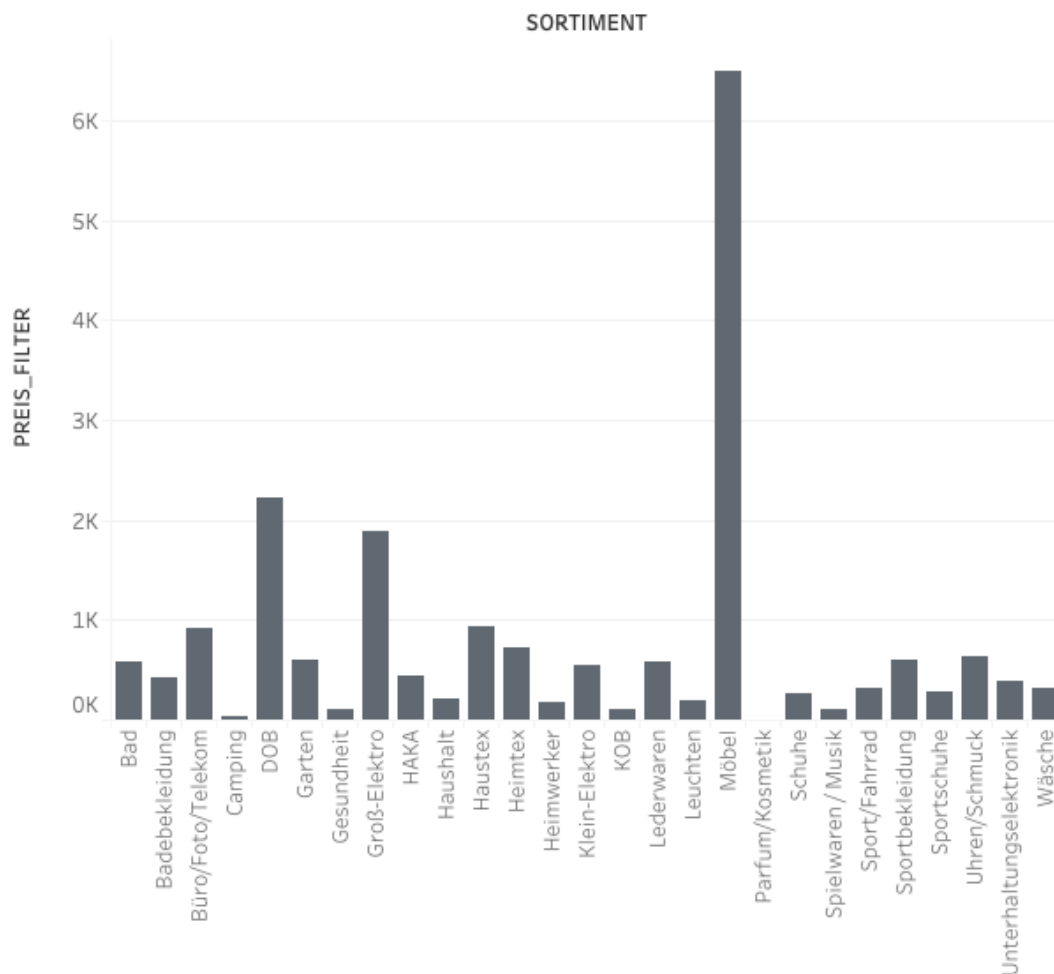


Abbildung 18: Absolute Nutzung des Preisfilters über alle Sortimente.

In Tabelle 13 werden die häufigsten Filterwerte des Preisfilters angeführt. Dabei lässt sich ableiten, dass der Preisfilter mit 3242 Mal zum allergrößten Teil von User*innen verwendet wird, um nach SALE-Artikeln ("nur reduzierte Artikel") zu

filtern. Außerdem werden verschiedenste Preis-Ranges durch einen im Shop implementierten Schiebe-Regler angegeben. Auch dabei ist erkennbar, dass der Großteil der gesetzten Produktpreis-Bereiche eher im LOW-Price-Segment (bis max. 200 €) ist und daher eine gewisse Preissensitivität der OTTO-Kund*innen angenommen werden kann.

	Filterwerte des Preisfilters	Häufigkeit der Filterwerte
1	⊕ Nur reduzierte Artikel	3242
2	⊕ reduziert	179
3	⊕ 9-50	149
4	⊕ 19-200	121
5	⊕ 19-50	108
6	⊕ 79-500	97
7	⊕ 9-40	93
8	⊕ 12-50	85
9	⊕ 9-30	85
10	⊕ 14-50	78
11	⊕ 14-40	70
12	⊕ 19-30	64
13	⊕ 8-25	61
14	⊕ 24-790	61
15	⊕ 289-1000	60
17	⊕ 29-100	56
18	⊕ 9-35	54
19	⊕ 289-600	54
21	⊕ 11-40	53
22	⊕ 14-60	52
23	⊕ 12-30	51
24	⊕ 9-25	50
25	⊕ 19-300	50
26	⊕ 79-600	50
27	⊕ 9-20	49
28	⊕ 15-1000	47
29	⊕ 24-375	47
30	⊕ 8-50	46

Tabelle 13: Top 30 Filterwerte des Preisfilters.

Blickt man auf die führenden Sortimente einer Suchanfrage, ist zu sehen, dass insbesondere im Möbel-, Groß-Elektro-, Damenoberbekleidung- (DOB) und im Haustex-Sortiment nach reduzierten Produkten und Sale-Angeboten gefiltert wird. Neben Suchanfragen in das DOB- bzw. Möbel-Sortiment sind die Kund*innen auch im Bademodebereich preissensitiver. Tabelle 14 listet dabei die häufigsten fünf

Preisfilter-Werte in den Sortimenten auf, wo mindestens 30x entweder „reduziert“ und/oder eine bestimmte Preis-Range ausgewählt wurde.

Filterwerte des Preisfilter: ▾	Häufigkeit der Filterwerte ▾
Nur reduzierte Artikel	3187
Möbel	689
Groß-Elektro	413
DOB	411
Haustex	244
Büro/Foto/Telekom	149
Sportbekleidung	147
Badebekleidung	121
Bad	111
HAKA	111
Unterhaltungselektronik	103
Heimtex	99
Sportschuhe	86
Klein-Elektro	84
Garten	68
Wäsche	68
Uhren/Schmuck	64
Sport/Fahrrad	59
Lederwaren	52
Schuhe	44
Heimwerker	34
Haushalt	30
9-50	119
DOB	78
Badebekleidung	41
19-200	89
Möbel	89
19-50	36
DOB	36
79-500	91
Möbel	91
9-40	32
Badebekleidung	32

Tabelle 14: Filterwerte des Preisfilters bei Sortimente mit mindestens 30 Mal gesetztem Filterwert.

Abbildung 19 veranschaulicht die in Tabelle 14 erkennbare Relevanz der „Sale-Affinität“ sehr deutlich in Form einer verclusterten Baumkarte. So konnten in der Visualisierungs-Software *Tableau* vier Cluster in Bezug auf den Preisfilter und die Sortimente identifiziert werden, wobei der türkise Cluster alleinig das Sortiment

Möbel mit dem größten Anteil an Sale-Filterung darstellt – das kann auch aus Tabelle 14 nachvollzogen werden. Der orange Cluster beinhalten die Sortimente DOB und Groß-Elektro mit dem Filterwert „nur reduzierte Artikel anzeigen“. Der rote Cluster stellt neben dem überwiegenden Segment „Sale-Affinität“ noch die beiden kleinen Preis-Segmente „9-50“ und „19-200“, auch diese Segmentierung der Preis-Filtervalues kann in der vorherigen Tabelle 14 nachvollzogen werden. Der große blaue Cluster umfasst alle weiteren Preis-Ranges als Segmente.

Preiscluster

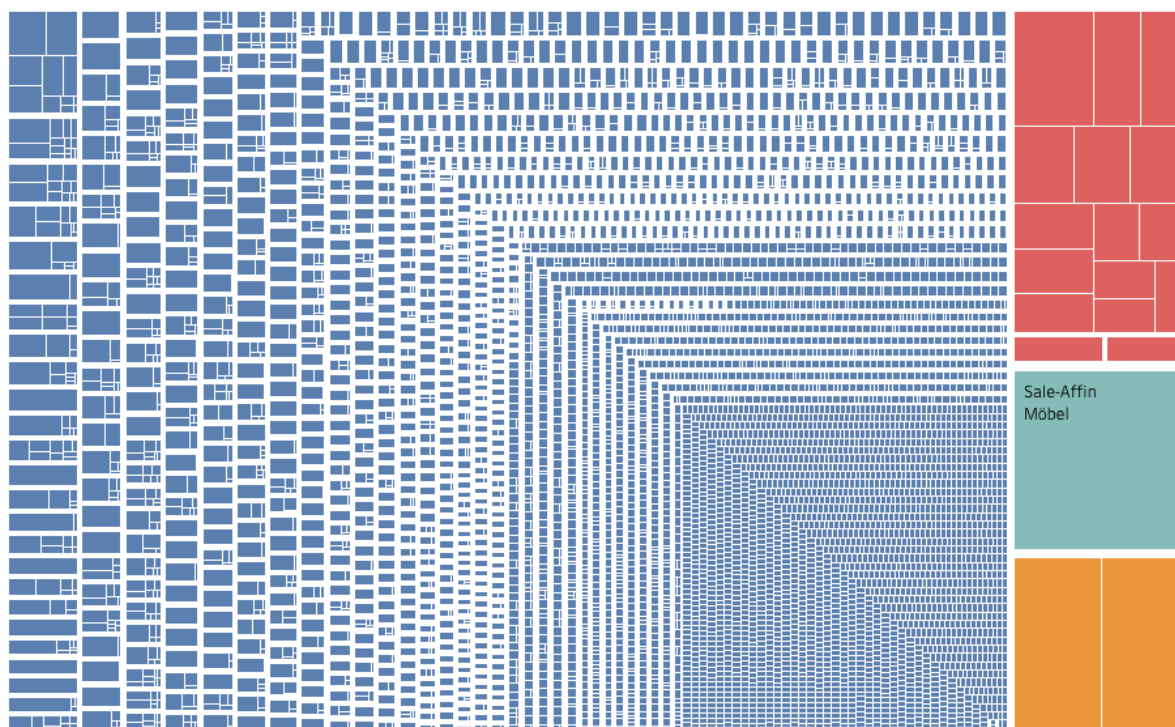


Abbildung 19: Clustering der Filterwerte des Preis-Filters innerhalb Sortimente.

Um der Vollständigkeit halber auch Erkenntnisse in höheren Preissegmente (bis 1000 €) zu erhalten, werden in Tabelle 15 Sortimente angeführt, bei denen davon ausgegangen werden kann, dass Kund*innen dazu eher bereit sind, mehr Geld für Produkte aus diesen Produktsegmenten auszugeben. Dabei lässt sich sagen, dass

Produkte mit höheren Preis-Ranges beinahe ausschließlich im Möbel-Sortiment gefiltert werden.

Filterwerte des Preisfilters	Häufigkeit der Filterwerte
79-500	91
Möbel	91
24-790	61
Möbel	61
15-1000	47
Möbel	47
19-300	45
Groß-Elektro	21
Möbel	24

Tabelle 15: Sortimente des Preisfilters bis 1000 €, mit der Einschränkung, dass eine Preisrange mind. 20x in einem Sortiment gesetzt wurde.

Da der Preisfilter im OTTO Online Shop in Form eines Schiebereglers implementiert ist, mit dem entweder Preisbereiche eingestellt werden können oder mittels Checkbox nach reduzierte Artikel gefiltert werden kann, sind hier keine Mehrfachauswahlen von Filterwerten wie bei anderen Filtern möglich, mit der Ausnahme dass gleichzeitig zu einer eingestellten Preis-Range auch der Sale-Filter gesetzt werden kann. Tabelle 16 zeigt einen Ausschnitt, wie häufig das vorkommt. Insgesamt wurden 1241 Mal im Juli 2020 nach reduzierte Artikel gemeinsam mit einer manuell eingegebenen Preisrange gefiltert.

Value	Summe von PREIS_FILTEI
Gesamtergebnis	1241
299-700 Nur reduzierte Artikel	20
12-50 Nur reduzierte Artikel	19
394-925 Nur reduzierte Artikel	15
79-500 Nur reduzierte Artikel	15
199-805 Nur reduzierte Artikel	14
399-790 Nur reduzierte Artikel	12
79-470 Nur reduzierte Artikel	12
12-48 Nur reduzierte Artikel	11
2-66 Nur reduzierte Artikel	11
19-401 Nur reduzierte Artikel	10
6-20 Nur reduzierte Artikel	10

Tabelle 16: Mehrfach-Auswahl von Preisranges und der Sale-Checkbox

Betrachtet man eine Segmentierung auf Suchbegriffs-Ebene ergeben sich im Nutzungsverhalten bei der Preisfilterung vier Cluster (Abbildung 20). Auch hier leiten die Cluster wieder eine Gruppierung nach der Häufigkeit an gesetzten Filtern ab und stellen durch die einzelnen Segmente verschiedene Suchbegriffe innerhalb der Cluster bzw. die Kacheln innerhalb der Segmente die verschiedenen Filterwerte/Preis-Ranges dar.

So beschreibt beispielsweise der türkise Cluster rechts unten jenen Cluster, welcher die Suchbegriffe mit den häufigsten Preisfilter-Values enthält. Suchten Kund*innen nach „boxspringbett“ wurde im Juli 2020 44 Mal die Preisrange 289-600 über den im Shop implementierten Preisfilter manuell eingestellt. Bei diesem Suchbegriff wurde zudem 31 Mal eine Preisrange von 289 € bis 1000 € von User*innen angegeben und 27 Mal der Hacken für Sale-Artikel gesetzt.

Preiscluster

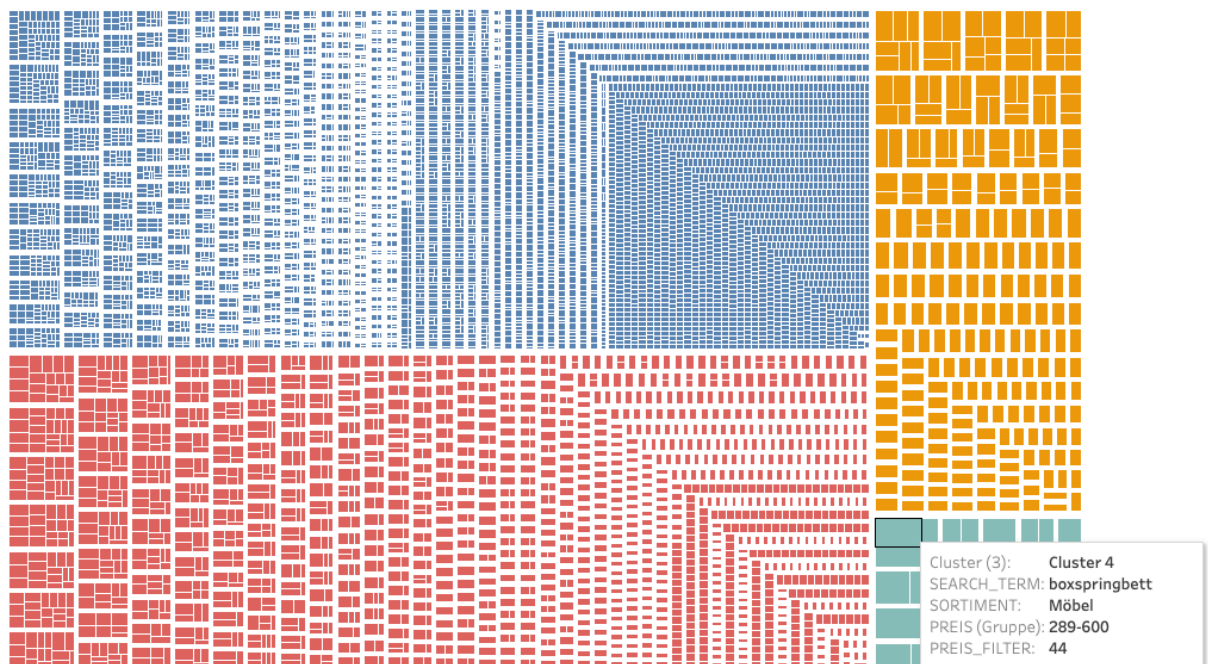


Abbildung 20: Baumkarte mit vier identifizierten Preisfilter-Clustern.

Die Varianzanalyse der identifizierten Preisfilter-Cluster in Abbildung 20 beschreibt, dass sich die Summe der gesetzten Preis-Filterwerte statistisch signifikant ($p=0,000$) zwischen den Clustern unterscheiden (siehe Tabelle 17).

Analyse der Varianz:

Variable	F-Statistik	p-Wert	Modell		Fehler	
			Quadratsumme	DF	Quadratsumme	DF
Summe von PREIS_FILTER	5997.0	0.0	17.57	3	21.16	21662

Tabelle 17: Varianzanalyse der identifizierten Preisfilter-Cluster

Es empfiehlt sich grundsätzlich, diese Clusterung auch für Handlungsempfehlungen in Bezug auf die Personalisierung des Preisfilters innerhalb verschiedener Suchbegriffe zu berücksichtigen (siehe Kapitel 6).

5.4.3 Größenfilter

Der untersuchte Größenfilter (siehe Tabelle 18) ist lediglich in den Fashion-Sortimenten relevant. Das liegt daran, dass der implementierte Größenfilter zum allergrößten Teil Kleider- und Schuhgrößen abbildet. Andere Größenangaben, wie z.B. Möbelabmessungen, Fernseher- oder Laptop-Diagonalen, Größenangaben in den Klein- und Groß-Elektro-Sortimenten etc. haben eigene Größen- bzw. Abmessungsfilter.

Major-Sortiment aus Suchanfrage	Anzahl der Searches gesamt	Searches m. Filter	Searches m. Größenfilter	...in %
Bad	34991	2840	210	7,39%
Badebekleidung	49536	4765	2783	58,41%
Büro/Foto/Telekom	53388	2037	0	0,00%
Camping	2471	72	0	0,00%
DOB	170019	18499	10828	58,53%
Garten	36848	1498	0	0,00%
Gesundheit	11642	384	0	0,00%
Groß-Elektro	43375	4412	0	0,00%
HAKA	24943	2836	1413	49,82%
Haushalt	33002	1106	1	0,09%
Haustex	35103	2728	0	0,00%
Heimtex	40946	4461	0	0,00%
Heimwerker	17969	809	0	0,00%
Klein-Elektro	32619	1568	0	0,00%
KOB	11562	820	452	55,12%
Lederwaren	16598	1711	0	0,00%
Leuchten	6551	447	0	0,00%
Möbel	167514	14149	0	0,00%
Parfum/Kosmetik	3311	77	0	0,00%
Schuhe	31608	4594	2659	57,88%
Spielwaren / Musik	13073	238	0	0,00%
Sport/Fahrrad	17201	586	14	2,39%
Sportbekleidung	39831	4950	2856	57,70%
Sportschuhe	31436	3155	1745	55,31%
Uhren/Schmuck	24656	1655	35	2,11%
Unterhaltungselektronik	24049	1074	0	0,00%
Wäsche	40942	4090	2628	64,25%
Gesamtergebnis	1015184	85561	25624	29,95%

Tabelle 18: Nutzungsrate des Größenfilters über alle führenden Sortimente innerhalb von Suchanfragen.

Blickt man auf die Filterwerte des Größenfilters ist anzumerken, dass aus der Datenanalyse zum Teil nur schwer unterscheidbar ist, ob es sich um einen Filterwert einer Kleidungsgröße oder einer Schuhgröße handelt (insbesondere bei den Filterwerten bzw. Größen „38“, „40“ und „42“ ist dieses Problem festzustellen). Dafür ist eine detailliertere Analyse basierend auf Sortiments-Ebene notwendig (siehe Tabelle 20).

	Filterwerte des Größenfilter	Häufigkeit der Filterwerte
1	⊕ 38	2265
2	⊕ 40	1780
3	⊕ 48	1727
4	⊕ 42	1714
5	⊕ 36	1592
6	⊕ 44	1469
7	⊕ 46	1354
8	⊕ 50	1047
9	⊕ 34	915
10	⊕ 52	795
11	⊕ 39	697
12	⊕ 54	586
13	⊕ 46,48	546
14	⊕ 48,50	429
15	⊕ L	426
17	⊕ 43	401
18	⊕ 44,46	400
19	⊕ 36,34	399
21	⊕ M	339
22	⊕ 37	326
23	⊕ 38,M	319
24	⊕ XXL	305
25	⊕ 85	305
26	⊕ 36,S	293
27	⊕ 38,36	278
28	⊕ 32	275
29	⊕ 40,38	271
30	⊕ XXXL	264

Tabelle 19: Top 30 Filterwerte des Größenfilters.

Tabelle 19 stellt die 30 häufigsten Werte des Größenfilter im Juli 2020 auf ottoversand.at dar. Dabei ist erkennbar, dass insbesondere bei den Damen-Großen-Größen die entsprechenden Größen-Filter verhältnismäßig häufig gesetzt werden. Dazu gehören die Filterwerte „44“, „46“, „48“, „50“, „52“, „54“.

Außerdem lässt sich durch Tabelle 19 auch ableiten, dass ein entsprechend großer Anteil an User*innen eine Zwischengröße trägt. Zum Beispiel durch 546 gleichzeitige Filtersetzungen von Größe „46,48“ oder die gemeinsame Filterauswahl von „38,M“, welche in Summe 319 Mal getätigt wurde.

Tabelle 20 liefert weitere Insights über den Größenfilter auf Sortimentsebene. Dabei werden die Filterwerte mit den größten Anteilen jeweils herausgehoben. So ist ersichtlich, dass der häufigste Größenfilter-Wert eine Plus-Size Größe ist. Im Juli 2020

wurde 1160 Mal der Größen-Filter „48“ gesetzt. Grundsätzlich lässt die Anzahl an gesetzten DOB-Größenfiltern vermuten, dass die User*innen eher nach größer ausfallender Mode suchen bzw. filtern.

Filterwerte des Größenfilters	Summe von GROESSE_FILTER
DOB	5790
48	1160
38	950
36	784
44	783
46	780
42	669
50	664
Badebekleidung	1933
40	378
42	376
38	359
36	315
44	278
46	227
Sportbekleidung	1018
48	180
36	170
50	146
L	145
38	134
40	126
46	117
Schuhe	2239
38	500
39	450
40	403
42	260
37	235
41	222
43	169
HAKA	706
XL	173
L	148
XXXL	144
M	135
XXL	106

Tabelle 20: Filterwerte des Größenfilters auf Sortimentsebene.

Auch im Sportmode-Segment ist auffällig, dass ein größerer Anteil der gesetzten Werte eine Filterung nach Großen Größen darstellt. Im Badebekleidungs-Sortiment

sieht das Bild etwas anders aus, hier sind die am häufigsten vorkommenden Filterwerte zwischen Kleidergröße 36 und 42, was auf eher schlankere Frauen schließen lässt. Bezieht man die einzelnen Größen-Values auf das Schuh-Sortiment, wird in Tabelle 20 ersichtlich, dass ein großer Anteil der bereits genannten Größen „38“, „40“ und „42“ unter anderem auf Suchbegriffe zurückzuführen ist, die als Haupt-Sortimentsgruppe Schuhe umfassen. Eine weitere interessante Beobachtung liefert Tabelle 20 im Herrenmode-Segment (HAKA). So filtern User*innen eher nach den amerikanischen Größenmaßen (M, L, XL, XXL etc.). Dies kann im DOB Sortiment eher nicht bzw. wenn, dann in Filter-Kombination mit Konfektionsgrößen (z.B. „38,M“ – siehe Tabelle 21) festgestellt werden.

Value	Summe von GROESSE_FILTER
Gesamtergebnis	15263
46,48	546
48,50	429
44,46	400
36,34	399
38,M	319
36,S	293
38,36	278
40,38	271
40,42	230
34,32	218
42,44	218
50,52	203
52,54	196
36,38	165
34,XS	163
42,40	148
38,40	135
42,L	111
56,58	109
44,42	107
40,M	103
44,XL	81
46,XL	81
48,XXL	76
40,L	74
48,XL	71

Tabelle 21: Mehrfach-Auswahlen von Filterwerten des Markenfilters.

Tabelle 21 streicht nochmal heraus, dass User*innen oftmals zwischen zwei Größen tendieren. Denn insgesamt über 15.000 Mal wurden mehrere Größen-Werte gleichzeitig gesetzt. Am häufigsten ist das im Übergrößen-Sortiment erkennbar, da die Top-3-Mehrfachauswahlen zwischen Größe 44 und 50 gefiltert wurden.

Abbildung 21 veranschaulicht die Suchbegriffe basierend auf den gesetzten Größen-Values innerhalb des Größenfilters. Dabei ergeben sich fünf Cluster, wobei der blaue und der orange Cluster für Ableitungen hin zu Personalisierungs-Strategien des Größenfilters vernachlässigt werden können. Denn innerhalb dieser Cluster wird ersichtlich, dass eine Vielzahl an Suchbegriffen sehr selten bis hin zu nur einmal einen Größen-Filterwert bedient (zum Beispiel bei selten genutzten Größen-Filterkombinationen).

Größencluster

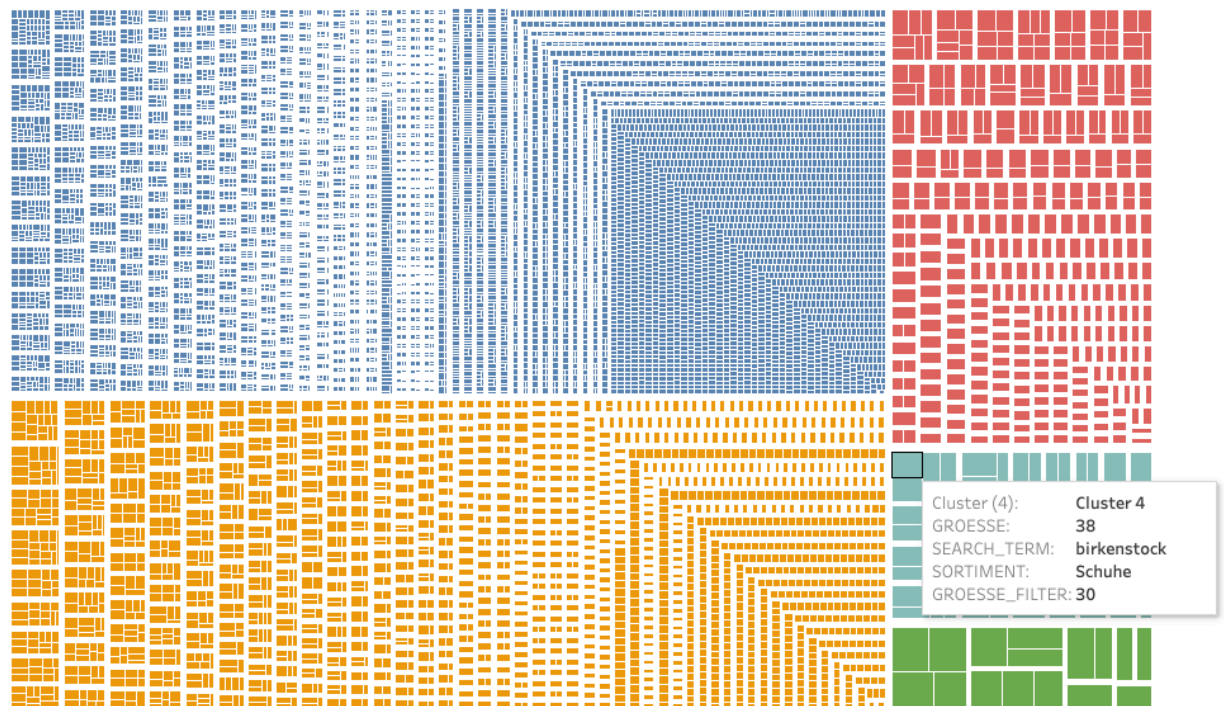


Abbildung 21: Baumkarte mit fünf identifizierten Größenfilter-Clustern.

Business-seitig interessanter sind hierbei der grüne, der türkise und der rote Cluster. Innerhalb dieser werden Suchbegriffe abgedeckt, die auch eine höhere

Anzahl an gesetzten Größen-Filterwerten vorweisen. Die zurückgelieferten Produkte des Suchbegriff „birkenstock“ wurden im Juli 2020 beispielsweise 30 Mal auf die Schuhgröße 38 eingeschränkt.

Die Analyse der Varianz beschreibt, dass es einen statistisch signifikanten Unterschied im Zusammenhang der Anzahl an gesetzter Größenfilter-Werte zwischen den verschiedenen Clustern auf Basis der Suchbegriffe gibt (siehe Tabelle 22).

Analyse der Varianz:

Variable	F-Statistik	p-Wert	Modell		Fehler	
			Quadratsumme	DF	Quadratsumme	DF
Summe von GROESSE_FILTER	5004.0	0.0	30.91	4	35.25	22826

Tabelle 22: Varianzanalyse der erkannten Größenfilter-Cluster

Durch diese Clusteranalyse kann abgeleitet werden, bei welchen Suchbegriffen bzw. Sortimenten welche Größenfilter am häufigsten gewählt werden und dadurch personalisiert prominenter dargestellt werden können. Konkrete Handlungsempfehlungen rund um die Informationen basierend auf dem Größenfilter sind in Kapitel 6 zusammen gefasst.

5.4.4 Farbfilter

Eine Filterung nach Farben ist mit einer Filternutzungsrate von 84,4 % bei Suchanfragen im Heimtex-Sortiment auffällig relevant (siehe Tabelle 23). Dieses enthält beispielsweise Suchbegriffe rund um Vorhänge, Bettwäschen, Teppiche etc. Auch Sortimente wie Bad, Haustex, Möbel und Leuchten sind mit einer Nutzungsrate von größer 50 % noch auffällig beliebt für eine Filterung anhand der verfügbaren Farben. Deutlich weniger nach Farben gefiltert wird in Sortimenten wie Sport/Fahrrad, Parfum/Kosmetik, Camping und Gesundheit, dort werden nicht mal bei 10 % aller Suchanfragen in die entsprechenden Sortimente Farbfilter gesetzt.

Major-Sortiment aus Suchanfrage	Anzahl der Searches gesamt	Searches m. Filter	Searches m. Farbfiler	...in%
Bad	34991	2840	1766	62,18%
Badebekleidung	49536	4765	1452	30,47%
Büro/Foto/Telekom	53388	2037	245	12,03%
Camping	2471	72	6	8,33%
DOB	170019	18499	6455	34,89%
Garten	36848	1498	534	35,65%
Gesundheit	11642	384	33	8,59%
Groß-Elektro	43375	4412	918	20,81%
HAKA	24943	2836	832	29,34%
Haushalt	33002	1106	473	42,77%
Haustex	35103	2728	1649	60,45%
Heimtex	40946	4461	3765	84,40%
Heimwerker	17969	809	318	39,31%
Klein-Elektro	32619	1568	185	11,80%
KOB	11562	820	109	13,29%
Lederwaren	16598	1711	740	43,25%
Leuchten	6551	447	260	58,17%
Möbel	167514	14149	8236	58,21%
Parfum/Kosmetik	3311	77	6	7,79%
Schuhe	31608	4594	1820	39,62%
Spielwaren / Musik	13073	238	33	13,87%
Sport/Fahrrad	17201	586	41	7,00%
Sportbekleidung	39831	4950	1142	23,07%
Sportschuhe	31436	3155	982	31,13%
Uhren/Schmuck	24656	1655	741	44,77%
Unterhaltungselektronik	24049	1074	114	10,61%
Wäsche	40942	4090	1276	31,20%
Gesamtergebnis	1015184	85561	34131	39,89%

Tabelle 23: Nutzungsrate des Farbfilters über alle führenden Sortimente innerhalb von Suchanfragen.

Tabelle 24 listet die 30 häufigsten Farb-Filterwerte im OTTO Online Shop auf, die im untersuchten Zeitraum von 1. bis 31. Juli 2020 gesetzt wurden. Dabei ist erkennbar, dass zum einen die Basic-Farben „weiß“ und „schwarz“ bei den Kund*innen sehr beliebt sind. Aber auch weitere Grundfarben wie blau, grau, braun, rot, beige, grün etc. werden häufig im Farb-Filter ausgewählt und für eine feinere Filterung des Suchergebnisses basierend auf Farben herangezogen.

	Filterwerte des Farbfilter	Häufigkeit der Filterwerte
1	⊕ weiß	9658
2	⊕ schwarz	7055
3	⊕ blau	3708
4	⊕ grau	3258
5	⊕ braun	2312
6	⊕ rot	1884
7	⊕ beige	1768
8	⊕ grün	1644
9	⊕ silberfarben	1061
10	⊕ rosa	1058
11	⊕ bunt	766
12	⊕ gelb	731
13	⊕ goldfarben	549
14	⊕ weiß,beige	453
15	⊕ beige,braun	424
17	⊕ grau,silberfarben	355
18	⊕ weiß,grau	339
19	⊕ grau,schwarz	328
21	⊕ weiß,schwarz	270
22	⊕ braun,beige	249
23	⊕ schwarz,grau	246
24	⊕ schwarz,blau	223
25	⊕ schwarz,weiß	209
26	⊕ blau,grün	200
27	⊕ beige,weiß	197
28	⊕ grau,beige	196
29	⊕ braun,schwarz	187
30	⊕ blau,bunt	164

Tabelle 24: Top 30 Filterwerte des Farbfilters.

Auch bei der Analyse des Farbfilters macht es Sinn – aufgrund des Long-Tail-Sortiments im OTTO Online Shop – einen detaillierteren Blick der Farb-Relevanzen und -Filternutzung auf die Ebene der Sortimente bzw. Suchbegriffe zu legen. Schränkt man die vorhandenen Daten wieder auf die Suchbegriffe ein, bei denen ein Farbfilter mindestens 30 Mal gesetzt wurde, ergibt sich das Bild in Tabelle 25.

Filterwerte des Farbfilters	Häufigkeit der Filterwerte
weiß	1568
⊕ Bad	87
⊕ DOB	79
⊕ Haushalt	30
⊕ Heintex	291
⊕ Möbel	1051
⊕ Sportschuhe	30
grau	382
⊕ Bad	34
⊕ Haustex	32
⊕ Heintex	221
⊕ Möbel	95
schwarz	229
⊕ Badebekleidung	68
⊕ Möbel	122
braun	208
⊕ Heintex	124
⊕ Möbel	84
grün	183
⊕ Badebekleidung	31
⊕ Haustex	94
⊕ Heintex	58
rosa	150
⊕ Haustex	31
⊕ Heintex	80
⊕ Möbel	39
rot	128
⊕ Heintex	83
⊕ Schuhe	45
beige	112
⊕ Heintex	82
⊕ Schuhe	30

Tabelle 25: Filterwerte des Farbfilters mit der Einschränkung, dass ein Farb-Filterwert mind. 30x bei Suchanfragen in ein bestimmtes Sortiment gesetzt wurde.

Es ist erkennbar, dass die unterschiedlichen Farbwerte hauptsächlich in den Sortimenten Möbel, DOB, Badebekleidung und Schuhe gesetzt werden. Hier sind Suchbegriffe wie Couchtisch, Bettwäsche, Bikini, Teppich, Vorhänge die meistgesuchten Queries, bei denen User*innen nach entsprechenden Farben filtern.

Ähnlich wie bei den Größen-Filterwerten ist erkennbar, dass sehr häufig mehrere (zum Teil sehr ähnliche) Farbtöne im Filter gesetzt wurden. Über 15.000 Mal filterten im Juli 2020 User*innen nach zwei Farben oder mehr auf den Suchergebnisseiten. Dabei sind die beliebtesten Mehrfach-Auswahlen bei den

Farben Weiß & Beige, Beige & Braun, Grau & Silberfarben, Weiß & Grau oder Grau & Schwarz (siehe Tabelle 26).

Zeilenbeschriftungen	Summe von FARBE_FIL
Gesamtergebnis	15128
weiß,beige	453
beige,braun	424
grau,silberfarben	355
weiß,grau	339
grau,schwarz	328
weiß,schwarz	270
braun,beige	249
schwarz,grau	246
schwarz,blau	223
schwarz,weiß	209
blau,grün	200

Tabelle 26: Mehrfach-Auswahlen von Filterwerten des Farbfilters.

Eine durchgeführte Clusteranalyse veranschaulicht die von User*innen eingegebenen Suchbegriffe, aufgeteilt in fünf erkannte Cluster nach Häufigkeit der gesetzten Farb-Filterwerte (siehe Abbildung 22).

Farbcluster

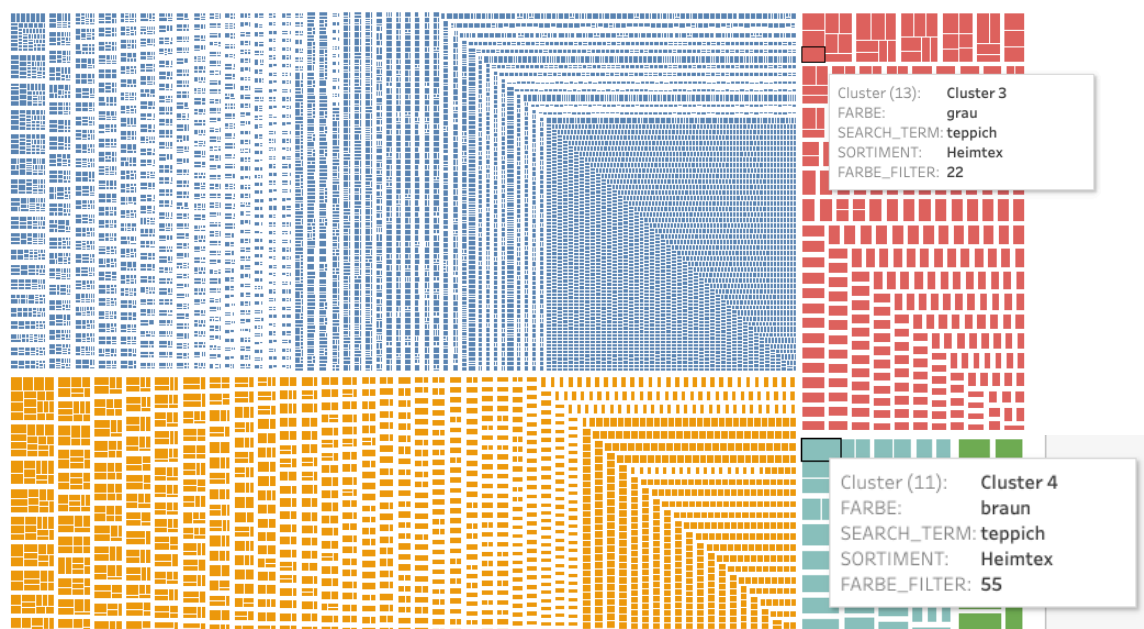


Abbildung 22: Baumkarte mit fünf identifizierten Farbfilter-Clustern.

Dabei sind der blaue und der orange Cluster wieder diejenigen mit den meisten Suchbegriffen. Sie beinhalten aber auch jene Suchbegriffe, welche nur wenige Farbfilter bzw. Farbfilter-Kombinationen nachweisen und daher für eine Personalisierungs-Strategie nur bedingt relevant sind. Der rote, türkise und grüne Cluster beschreiben die Suchbegriffe, die eine größere Anzahl an gesetzten Filterwerten vorweisen.

Dabei kann die Baumkarte wieder so interpretiert werden, dass beispielsweise das linke obere Segment im roten Cluster den Suchbegriff „teppich“ darstellt. Die markierte Kachel legt im Detail dar, dass bei diesem Suchbegriff insgesamt 22 Mal der Farbwert „grau“ gesetzt wurde. Je mehr Kacheln ein Segment bzw. Suchbegriff enthält, desto mehr verschiedene Farbwerte wurden einzeln oder in Kombination gesetzt und desto relevanter ist der Farbfilter auch für diesen Suchbegriff. Wichtig anzumerken ist, dass ein Suchbegriff auch in verschiedenen Clustern vorkommen kann, da die Suchbegriffe aufgrund der Summe an gesetzten Farbfilter-Werten geclustert werden – in welche(n) Cluster ein Suchbegriff reinfällt, hängt demnach davon ab, wie hoch die Anzahl der gesetzten Farbfilter-Werte ist. In Abbildung 22 ist erkennbar, dass der Suchbegriff „teppich“ auch als Segment im türkisen Cluster wiederzufinden ist. Denn beispielsweise der Filterwert „braun“ wurde in Bezug auf den entsprechenden Suchbegriff mit einer Gesamtanzahl von 55 mehr als doppelt so häufig gesetzt, wie der Farbwert „grau“ im roten Cluster“. Es lässt sich also ableiten, dass der braune Farb-Filterwert für die Suchanfrage „teppich“ deutlich relevanter ist, als „grau“.

Analyse der Varianz:

Variable	F-Statistik	p-Wert	Modell		Fehler	
			Quadratsumme	DF	Quadratsumme	DF
Summe von FARBE_FILTER	5155.0	0.0	14.34	4	16.38	23556

Tabelle 27: Varianzanalyse der identifizierten Farbfilter-Cluster.

Bei Durchführung der Analyse der Varianz ist erkennbar, dass mit $p=0,000$ ein signifikanter Unterschied im Zusammenhang der Anzahl an Farbfilter-Werte je Suchbegriff zwischen den Clustern gegeben ist (siehe Tabelle 27).

5.4.5 Geschlechtsfilter & Verfügbarkeitsfilter

Tabelle 28 betrachtet den Filter nach Geschlecht (z.B. Damen, Herren, Jungen, Mädchen etc.) und den Verfügbarkeitsfilter (z.B. sofort lieferbar, innerhalb 24 Stunden, 2-4 Werktage, 3 Wochen etc.) in einer gemeinsamen Aufstellung. Auffällig ist, dass im Vergleich zu den vorangegangenen Filtern diese beiden über alle Sortimente hinweg eine sehr geringe Filternutzungsrate auf entsprechenden SERPs haben.

Major-Sortiment aus Suchanfrage	Anzahl der Searches gesamt	Searches m. Filter	Searches m. "Für wen?" Filterin%	Searches m. Verfügbarkeitsfilterin %
Bad	34991	2840	7	0,25%	5	0,18%
Badebekleidung	49536	4765	74	1,55%	13	0,27%
Büro/Foto/Telekom	53388	2037	0	0,00%	13	0,64%
Camping	2471	72	0	0,00%	1	1,39%
DOB	170019	18499	394	2,13%	39	0,21%
Garten	36848	1498	0	0,00%	8	0,53%
Gesundheit	11642	384	0	0,00%	0	0,00%
Groß-Elektro	43375	4412	0	0,00%	10	0,23%
HAKA	24943	2836	159	5,61%	2	0,07%
Haushalt	33002	1106	0	0,00%	3	0,27%
Haustex	35103	2728	0	0,00%	5	0,18%
Heimtex	40946	4461	0	0,00%	8	0,18%
Heimwerker	17969	809	10	1,24%	4	0,49%
Klein-Elektro	32619	1568	0	0,00%	2	0,13%
KOB	11562	820	157	19,15%	0	0,00%
Lederwaren	16598	1711	117	6,84%	6	0,35%
Leuchten	6551	447	0	0,00%	4	0,89%
Möbel	167514	14149	45	0,32%	38	0,27%
Parfum/Kosmetik	3311	77	3	3,90%	1	1,30%
Schuhe	31608	4594	153	3,33%	12	0,26%
Spielwaren / Musik	13073	238	3	1,26%	5	2,10%
Sport/Fahrrad	17201	586	19	3,24%	7	1,19%
Sportbekleidung	39831	4950	268	5,41%	15	0,30%
Sportschuhe	31436	3155	60	1,90%	5	0,16%
Uhren/Schmuck	24656	1655	74	4,47%	6	0,36%
Unterhaltungselektronik	24049	1074	0	0,00%	3	0,28%
Wäsche	40942	4090	44	1,08%	7	0,17%
Gesamtergebnis	1015184	85561	1587	1,85%	222	0,26%

Tabelle 28: Nutzungsraten des Geschlechts- und Verfügbarkeits-Filter über alle führenden Sortimente innerhalb von Suchanfragen.

Wenn der Geschlechts-Filter gesetzt wird, ist er für die User*innen mit einer Nutzungsrate von 19,15 % im Kinderoberbekleidungssortiment verhältnismäßig

wichtig. Das liegt vor allem daran, dass im Sortiment KOB sowohl Jungen- als auch Mädchenkleidung in einem Sortimentsstrang zusammengefasst ist und so eine leichtere Navigation ermöglicht wird.

Unterbei stellt Tabelle 29 die gesetzten Filterwerte des Geschlechtsfilters dar. Wenn der Filter verwendet wird, ist der häufigste Filterwert, der gesetzt wird „female“. Nach Artikeln für Männer wurde im Juli 2020 461 gefiltert („male“). Auch bei Suchanfragen, die Artikel des Kindersortiments enthalten, wird der Filter mit unterschiedlichen Filterwert-Kombinationen genutzt.

	Filterwerte des Geschlechts-Filte	Häufigkeit der Filterwerte
1	⊕ female	1354
2	⊕ male	461
3	⊕ child	193
4	⊕ girl	173
5	⊕ boy	137
6	⊕ baby	38
7	⊕ child,boy	23
8	⊕ child,girl	17
9	⊕ girl,baby	2
10	⊕ boy,girl	2
11	⊕ child,baby	1
12	⊕ child,boy,girl	1
13	⊕ boy,baby	1
	Gesamtergebnis	2403

Tabelle 29: Top 12 Filterwerte des Geschlechts-Filters („Für Wen?“).

Blickt man auf die Suchbegriffe, bei denen ein Filterwert mindestens 20 Mal von Nutzer*innen ausgewählt wurde, ist anhand Tabelle 30 ersichtlich, dass bei geschlechts-neutralen Suchbegriffen wie „T-Shirts“, „Zehentrenner“, „Geldbörsen“, „Fahrradtrikots“ oder „Lederhosen“ der Geschlechts-Filter verwendet wird. Im Kindersortiment wurde ein Filterwert lediglich bei den Suchbegriffen „noppies“ bzw. „kinderbett“ mindestens 20 Mal gesetzt, andere Suchbegriffe sind für den Geschlechtsfilter demzufolge weniger relevant (siehe Tabelle 30).

Zeilenbeschriftungen	Summe von GESCHLECHT_FILTER
female	648
blusen	35
bügel-bikinis	20
Fahrradtrikots	56
Funktionsshirts	21
Geldbörsen	52
Hoodies	28
Jeansjacken	22
Jogginganzüge	23
Jumpsuits	57
Regenjacken	28
Sandaletten	28
Shorts	35
Strohhüte	40
T-Shirts	86
Umhängetaschen	25
wanderschuhe	34
zehentrenner	58
male	136
Lederhosen	40
sandalen	20
Trachtenjacken	26
T-Shirts	50
boy	45
noppies	45
child	26
kinderbett	26

Tabelle 30: Filterwerte des Geschlechtsfilters mit der Einschränkung, dass ein Filterwert mind. 20x bei Suchanfragen in ein bestimmtes Sortiment gesetzt wurde.

Abbildung 23 beschreibt eine Clusteranalyse der Geschlechts-Filterwerte mit einer Segmentierung der Suchbegriffe. Die mittels k-Means-Algorithmus identifizierten vier Cluster werden – wie auch schon bei den vorangegangenen Untersuchungen der Filter – in Form einer Baumkarte über die Visualisierungs-Software *Tableau* dargestellt. Der türkise Cluster beinhaltet dabei jene Filterwerte, die eine hohe Anzahl an Geschlechtsfilter-Values je segmentierten Suchbegriff aufweisen. So wurde beispielsweise im Zeitraum von 01. bis 31.07.2020 beim Suchbegriff „T-Shirts“ 62 Mal der „female“-Filterwert gesetzt. Unter der „female“-Kachel ist ersichtlich, dass bei diesem Suchbegriff auch der Filterwert für männliche Artikel („male“)

häufig ausgewählt wurde. Auch in anderen Clustern ist der Suchbegriff enthalten, dies liegt daran, dass dabei andere führende Sortimente ausgegeben wurden (z.B. das Sortiment Sportbekleidung im roten Cluster links oben).

Geschlechtsfilter

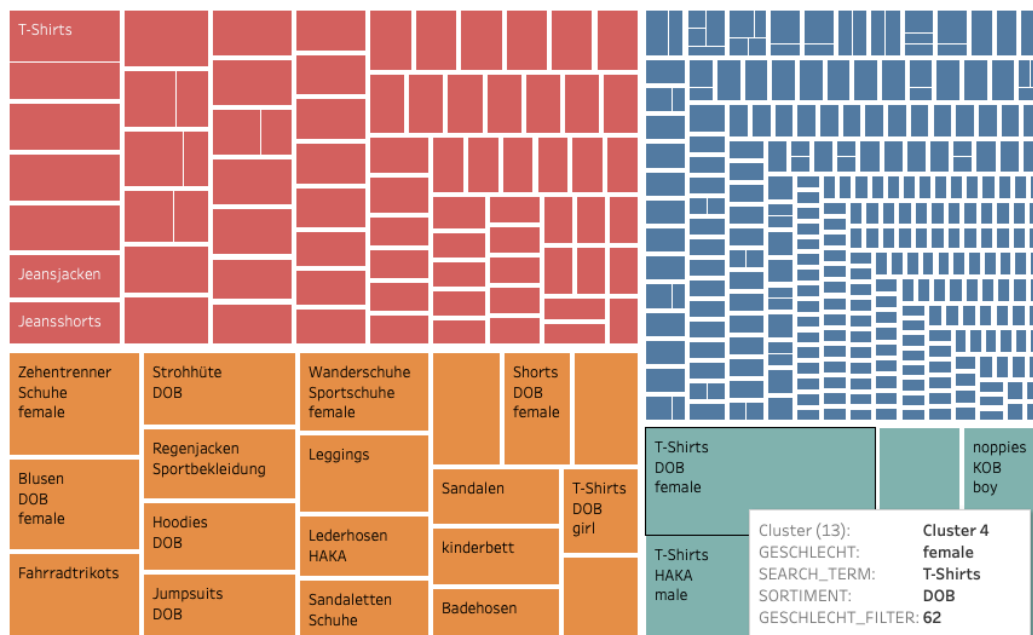


Abbildung 23: Clusteranalyse der Geschlechts-Filterwerte in Form einer Baumkarte.

Der orange Cluster beinhaltet jene Suchbegriffe, die ebenfalls noch eine entsprechend hohe Anhäufung an gesetzten Geschlechtsfilter-Werte aufweisen. Der rote und blaue Cluster sind aufgrund der sowieso schon geringen Relevanz des Geschlechtsfilters für den OTTO Online Shop eher zu vernachlässigen und müssen nicht zwingend in die Personalisierungs-Strategie der Filter miteinbezogen werden.

Tabelle 31 stellt die ANOVA für die Cluster des Geschlechtsfilters dar und beschreibt hochsignifikante Unterschiede ($p=0,000$) im Zusammenhang der Anzahl an Geschlechts-Filterwerte zwischen den jeweiligen Clustern.

Analyse der Varianz:

Variable	F-Statistik	p-Wert	Modell		Fehler	
			Quadratsumme	DF	Quadratsumme	DF
Summe von GESCHLECHT_FILTER	4738.0	0.0	6.618	3	7.189	15440

Tabelle 31: Varianzanalyse der identifizierten Geschlechtsfilter-Cluster.

In Hinblick auf den Verfügbarkeitsfilter lässt sich ableiten, dass dieser Filter bei Suchanfragen im Spielwaren/Musik-Sortiment mit einer Nutzungsrate von 2,10 % relativ gesehen am häufigsten verwendet und nach einer gewünschten Lieferverfügbarkeit gefiltert wird. Absolut sind das aber nur fünf ausgewählte Verfügbarkeits-Filterwerte (siehe Tabelle 28).

Blickt man auf die einzelnen Filterwerte des Lieferzeitenfilters zeigt Tabelle 32, dass bei einer allgemein geringen Nutzung des Filters den größten Anteil der „in 2-4 Werktagen bei dir“-Filterwert ausmacht. Daraus lässt sich ableiten, dass User*innen, die diesen Filter-Value auswählen, Wert auf eine schnelle Lieferung legen. Aus der gleichen Tabelle lässt sich auch erkennen, dass in absoluten Zahlen nur sehr selten mehrere Verfügbarkeits-Filterwerte gleichzeitig gesetzt werden.

Filterwerte des Verfügbarkeitsfilters	Häufigkeit der Filterwerte
1 ⊕ vorrätig - in 2-4 Werktagen bei dir	170
2 ⊕ vorrätig - in 2-4 Werktagen bei dir,vorrätig - in 5-6 Werktagen bei dir	20
3 ⊕ vorrätig - in 5-6 Werktagen bei dir	18
4 ⊕ vorrätig - in 4-5 Werktagen bei dir	15
5 ⊕ vorrätig - in 24 Stunden bei dir	15
6 ⊕ vorrätig - in 2-4 Werktagen bei dir,vorrätig - in 5-6 Werktagen bei dir,Lieferbar in 1 Woche	9
7 ⊕ vorrätig - in 2-4 Werktagen bei dir,Lieferbar in 1 Woche	6
8 ⊕ Lieferbar in 2 Wochen	5
9 ⊕ vorrätig - in 24 Stunden bei dir,vorrätig - in 2-4 Werktagen bei dir	5
10 ⊕ vorrätig - in 3-4 Werktagen bei dir	5
11 ⊕ vorrätig - in 2-4 Werktagen bei dir + weitere	5
12 ⊕ vorrätig - in 5-6 Werktagen bei dir,Lieferbar in 1 Woche	4
13 ⊕ Lieferbar in 3 Wochen	4
14 ⊕ vorrätig - in 2-4 Werktagen bei dir,vorrätig - in 3-4 Werktagen bei dir,vorrätig - in 4-5 Werktagen bei dir,vorrätig - in 5-6 Werktagen bei dir,Lieferbar in 1 Woche	4
15 ⊕ vorrätig - in 2-4 Werktagen bei dir,vorrätig - in 3-5 Werktagen bei dir	3
17 ⊕ vorrätig - in 2-4 Werktagen bei dir,vorrätig - in 3-5 Werktagen bei dir,vorrätig - in 5-6 Werktagen bei dir	3
18 ⊕ vorrätig - in 2-4 Werktagen bei dir,vorrätig - in 3-4 Werktagen bei dir	3
19 ⊕ Lieferzeit länger als 4 Wochen	3
21 ⊕ vorrätig - in 4-5 Werktagen bei dir,Lieferbar in 1 Woche,Lieferbar in 2 Wochen	2
22 ⊕ vorrätig - in 24 Stunden bei dir + weitere	2
23 ⊕ vorrätig - in 4-5 Werktagen bei dir,vorrätig - in 5-6 Werktagen bei dir,Lieferbar in 1 Woche	1
24 ⊕ vorrätig - in 2-4 Werktagen bei dir,vorrätig - in 3-4 Werktagen bei dir,vorrätig - in 4-5 Werktagen bei dir,Lieferbar in 2 Wochen,Lieferbar in 4 Wochen	1
25 ⊕ vorrätig - in 5-6 Werktagen bei dir,Lieferbar in 2 Wochen	1
26 ⊕ vorrätig - in 2-4 Werktagen bei dir,vorrätig - in 3-4 Werktagen bei dir,vorrätig - in 5-6 Werktagen bei dir	1
27 ⊕ vorrätig - in 4-5 Werktagen bei dir,Lieferbar in 2 Wochen,Lieferbar in 3 Wochen,Lieferbar in 4 Wochen	1
28 ⊕ vorrätig - in 2-4 Werktagen bei dir,Lieferbar in 2 Wochen	1
29 ⊕ vorrätig - in 24 Stunden bei dir,vorrätig - in 3-4 Werktagen bei dir,vorrätig - in 4-5 Werktagen bei dir,Lieferbar in 1 Woche	1
30 ⊕ vorrätig - in 2-4 Werktagen bei dir,vorrätig - in 5-6 Werktagen bei dir,Lieferbar in 2 Wochen,Lieferbar in 3 Wochen,Lieferbar in 4 Wochen	1

Tabelle 32: Top-30-Filterwerte des Lieferzeitenfilters & Häufigkeit von Mehrfach-Auswahlen.

In einer tiefergehenden Sortiments-Analyse (siehe Tabelle 33) wird ersichtlich, dass der Top-1-Filter vorrangig im Fashion-Segment (DOB, Badebekleidung, Schuhe, Sportbekleidung, Wäsche) gesetzt wird.

Aufgrund der business-seitigen Relevanz der 24-Stunden-Lieferung, die OTTO Österreich anbietet, wird auch der „in 24 Stunden bei dir“-Filterwert beleuchtet. Dabei lässt sich anhand Tabelle 33 erkennen, dass mit 15 Einzel-Auswahlen des Filterwertes lediglich sehr wenige User*innen aktiv nach diesem Angebot im Juli gefiltert haben. Gesetzt wurde er im Büro/Foto/Telekom-, Klein-Elektro-, Unterhaltungselektronik-, Groß-Elektro- und Heimwerker-Sortiment.

Filterwerte des Verfügbarkeitsfilters	Häufigkeit der Filterwerte
vorrätig - in 2-4 Werktagen bei dir	110
DOB	45
Badebekleidung	14
Schuhe	14
Sportbekleidung	14
Möbel	13
Wäsche	10
vorrätig - in 24 Stunden bei dir	15
Büro/Foto/Telekom	7
drucker	2
Gaming Monitor	1
gaming pc	1
laptop	1
Monitor	1
pc	1
Klein-Elektro	3
heissluftfriteuse	1
mikrowelle	1
staubsauger sale	1
Unterhaltungselektronik	3
fernseher	2
jbl bluetooth	1
Groß-Elektro	1
Kühlschränke mit G	1
Heimwerker	1
klimageräte	1

Tabelle 33: Filterwerte "in 2-4 Werktagen bei dir" und „in 24 Stunden bei dir“ inkl. Sortimente bzw. Suchbegriffe, innerhalb deren die Filterwerte am häufigsten gesetzt werden.

Auch für den Lieferzeitenfilter wurde eine Clusteranalyse durchgeführt. Anhand Abbildung 24 ist erkennbar, dass die wenigen gesetzten Verfügbarkeitswerte im Filter nur in zwei Cluster unterteilbar sind. Dabei beschreibt der rote Cluster jene Suchbegriffe, bei denen ein Lieferaussagewert mindestens 2 Mal im betrachteten Zeitraum von User*innen gesetzt wurde. Der orange Cluster stellt all jene Suchbegriffe dar, wo Filterwerte nur einmal gewählt wurden.

Lieferverfügbarkeitcluster

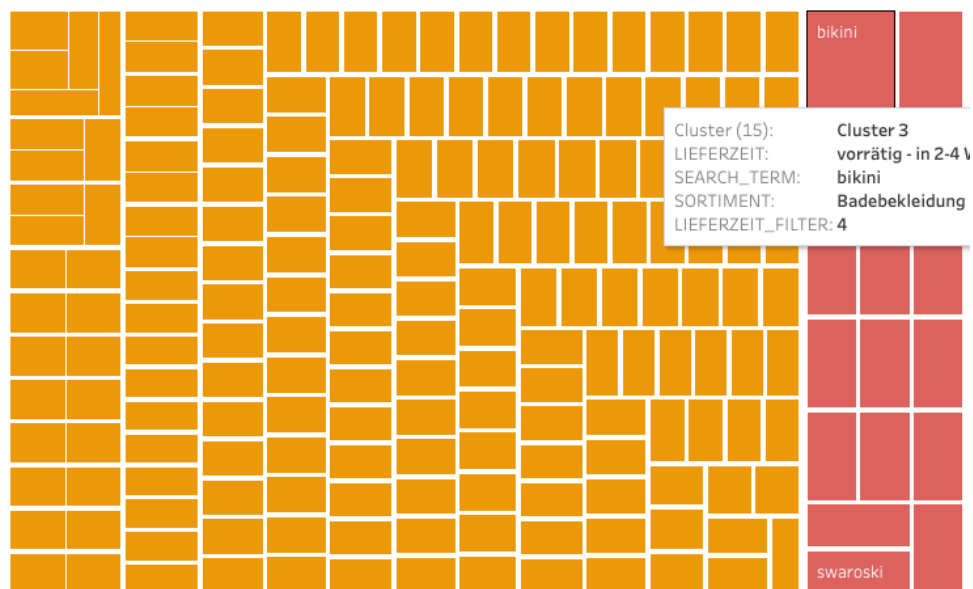


Abbildung 24: Clusteranalyse der Lieferzeiten-Filterwerte in Form einer Baumkarte.

Wie schon zuvor angemerkt ist der Verfügbarkeitsfilter – ähnlich wie der Geschlechtsfilter – business-seitig aufgrund fehlender Nutzung seitens der Kund*innen kaum relevant und kann für weitere Handlungsempfehlungen vernachlässigt werden.

Auch für die oben dargestellte Clusteranalyse in Abbildung 24 wird eine Varianzanalyse durchgeführt, um eine Aussage hin zur Signifikanz geben zu können. Dabei lässt sich als Beschreibung der deskriptiven Daten in Tabelle 34

ableiten, dass sich die Summe der gesetzten Lieferzeiten-Filterwerte signifikant zwischen den beiden identifizierten Clustern unterscheiden ($p=0,000$).

Analyse der Varianz:

Variable	F-Statistik	p-Wert	Modell		Fehler	
			Quadratsumme	DF	Quadratsumme	DF
Summe von LIEFERZEIT_FILTER	7555.0	0.0	16.67	2	16.99	15396

Tabelle 34: Varianzanalyse der identifizierten Lieferzeitenfilter-Cluster.

5.5 Gibt es User*innen, die Suchergebnisse häufig mit Filter einschränken und in mehreren Sessions Filter setzen?

Zur Beantwortung dieser Frage wird eine weitere Clusteranalyse durchgeführt und in Form eines Streudiagramms dargestellt. Dabei werden – ähnlich wie bei der Clusteranalyse mit der untersuchten Ebene der Suchbegriffe in Abbildung 15 – sowohl die Anzahl der Sessions mit Suchanfrage als auch die insgesamt gesetzten Filter aggregiert; dieses Mal aber auf Ebene der verhashten User-IDs. Mittels k-Means-Algorithmus wird das Streudiagramm in zwei Cluster geteilt (siehe Abbildung 25).

Der blaue Cluster beinhaltet dabei jene User, die im untersuchten Zeitraum nur wenige Sessions mit einer Suchanfrage durchgeführt haben und dabei auch nur wenige unterschiedliche Filter genutzt haben. Anhand der User in diesem Cluster können also nur bedingt Interpretationen zur oben gestellten Frage getroffen werden.

Business-seitig relevanter ist dabei jener Cluster, der in mehreren Sessions mit Suchanfrage auch mehrere Filter gesetzt hat. Dieser orange Cluster beinhaltet mehr als 1500 individuelle User*innen, die über das Tracking-System wiedererkannt werden konnten.

Clusteranalyse User

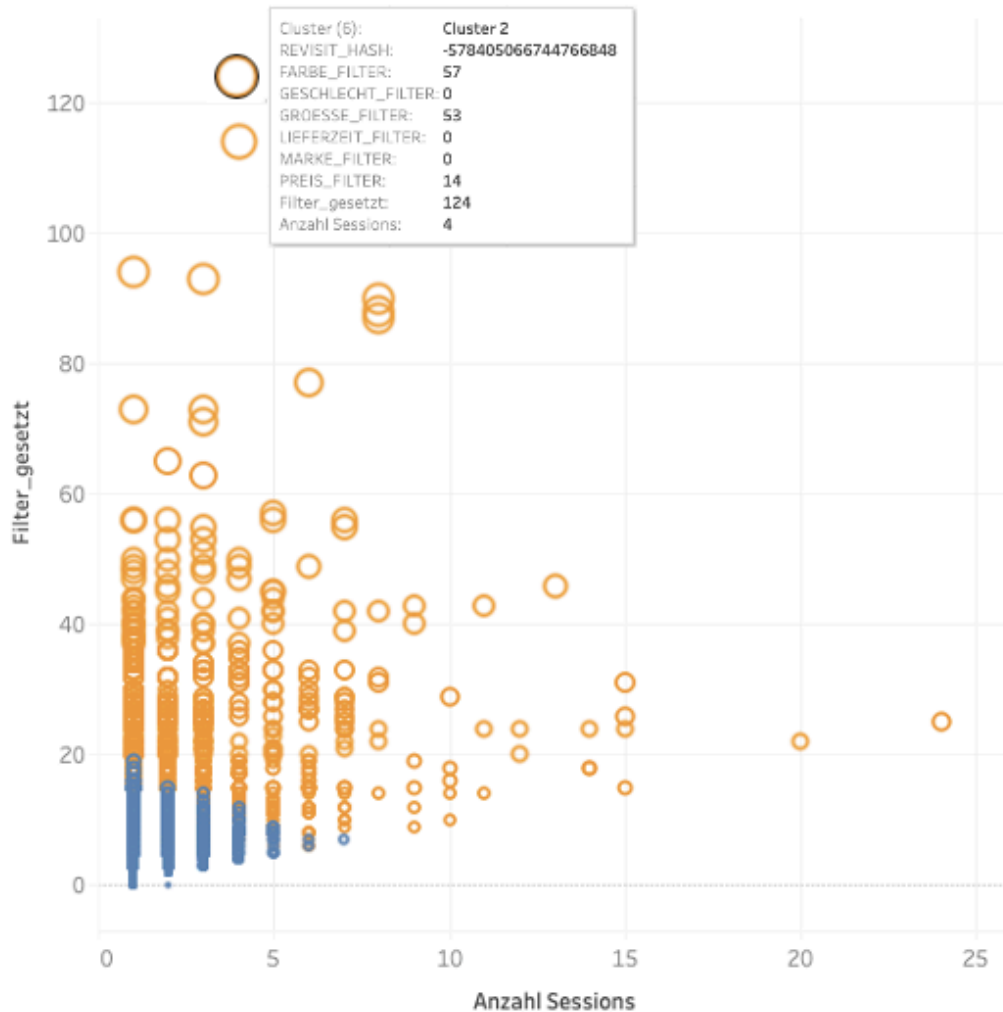


Abbildung 25: Ermittelte Cluster der User*innen basierend auf den gesetzten Filtern pro Session mit Suchanfrage.

So setzte beispielsweise die markierte User*in (Abbildung 25) in 4 Sessions mit Suchanfragen insgesamt 57 Mal den Farbfilter mit entsprechenden Filterwerten, 53 Mal den Größenfilter und 14 Mal einen Preisfilter. Diese Information über einzelne User kann sehr interessant sein, ist aber auch mit sehr großer Vorsicht hinsichtlich der neuen EU-DSGVO zu walten.

Die durchgeführte Analyse der Varianz in Tabelle 35 zeigt, dass es einen signifikanten Unterschied im Zusammenhang der User zwischen den gesetzten

Filtern Farbe, Größe, Preis und Marke gibt ($p=0,000$). Es gibt keinen statistisch signifikanten Unterschied im Zusammenhang der User zwischen den gesetzten Geschlechtsfiltern und den Lieferzeitenfiltern.

Analyse der Varianz:

Variable	F-Statistik	p-Wert	Modell		Fehler	
			Quadratsumme	DF	Quadratsumme	DF
Filter_gesetzt	1.698e+04	0.0	21.65	1	41.79	32764
Summe von FARBE_FILTER	1.239e+04	0.0	24.47	1	64.72	32764
Anzahl Sessions	4991.0	0.0	3.542	1	23.25	32764
Summe von GROESSE_FILTER	3691.0	0.0	3.278	1	29.09	32764
Summe von PREIS_FILTER	3610.0	0.0	4.117	1	37.37	32764
Summe von MARKE_FILTER	718.8	0.0	0.657	1	29.95	32764
Summe von GESCHLECHT_FILTER	19.17	1.202e-05	0.006483	1	11.08	32764
Summe von LIEFERZEIT_FILTER	0.2583	0.6113	6.3e-05	1	7.992	32764

Tabelle 35: Analyse der Varianz des Clusterings der User.

6. Diskussion

6.1 Konkrete Handlungsempfehlungen & Implikationen

Zur Beantwortung der zu Grunde liegenden Forschungsfrage dieser Arbeit

Welche Personalisierungs-Maßnahmen lassen sich durch die Filter-Nutzung auf Suchergebnisseiten im OTTO Online Shop ableiten?

werden basierend auf den Ergebnissen der Auswertung, Analyse und Interpretation der gesammelte Web-Tracking-Daten Handlungsempfehlungen und Implikationen für den OTTO Online Shop abgeleitet.

Als einleitende und übergreifende Empfehlung wird abgegeben, alle nachfolgend erläuterten Themen auch aus fachlicher Sicht zu beleuchten und den Business-Impact für jede einzelne Implikation zu beziffern – sowohl in Bezug auf den Umsetzungs-Aufwand als auch auf den Nutzen, den verschiedene Use Cases und (Teil-)Projekte mit sich bringen können.

In einem ersten Schritt ist es zu empfehlen, ein klares Ziel festzulegen, was mit der Personalisierung von Filtern bzw. mit den gesammelten Daten aus der Filternutzung im OTTO Online Shop erreicht werden soll. Die Markenverantwortlichen können sich beim Identifizieren entsprechender Ziele an die SMART-Methode halten. Dabei sollten die Ziele spezifisch, messbar, akzeptiert, realistisch und terminiert sein. So ist sichergestellt, dass erreichbare Ziele aufgestellt werden und im Monitoring zu jedem Zeitpunkt klar ist, wie der aktuelle Stand der Ziel-Umsetzung bzw. -Erreichung aussieht.

Sind die Ziele definiert, kann die auf den Produktfilter basierende Personalisierungsstrategie erarbeitet und mittels Priorisierung der umzusetzenden Themen geplant werden. Durch die so entstehende Roadmap können die einzelnen Use Cases iterativ entwickelt und im Online Shop implementiert werden. Dies hat

den Vorteil, dass Feature-Quick-Wins schnell Mehrwert auf Kundenseite generieren können, einfach und in kurzer Zeit Testings aufgesetzt und Insights zum Erfolg verschiedener Maßnahmen getroffen werden können sowie Verbesserungen auf kurzem Wege vorgenommen werden können. Ein Ansatz des iterativen Vorgehens beim Aufbau einer entsprechenden Personalisierungsstrategie könnte sein, sich vorrangig auf die relevantesten Suchbegriffe zu fokussieren. Das sind für ottoversand.at im Juli 2020 jene Queries, die in Tabelle 5 angeführt werden und den roten Cluster in Abbildung 15 beschreiben. Da die Ergebnisse der empirischen Forschung auch gezeigt haben, dass der Verfügbarkeitsfilter sowie der Geschlechtsfilter im Vergleich zu den anderen vier untersuchten Filtern (Größe, Farbe, Marke, Preis) nur wenig Nutzungsrelevanz haben, sollten diese Filter im MVP-Ansatz (Minimum Viable Product) vernachlässigt und entsprechende Use Cases erst zu einem späteren Zeitpunkt erarbeitet und umgesetzt werden.

Zu empfehlen ist ein Mix aus impliziter Datensammlung – durch beispielsweise der Speicherung des Klickverhaltens von User*innen – und einer expliziten Abfrage nach Lieblingsmarken, bevorzugter Größen etc., wo Kunden entsprechende Präferenzen selber angeben und auch wieder ändern können. Beim Generieren, Sammeln und Speichern der Daten hat ein Schutz und hochsensibler Umgang der personenbezogenen Informationen der OTTO User*innen oberste Priorität. Für die Sammlung und Weiterverarbeitung der Clickstream-Daten könnte beispielsweise angedacht werden, dass die Klickdaten verhaltensbasiert weiterverarbeitet werden. So kann beispielsweise ein Modell mit historischen Daten der Filternutzung trainiert werden, welches basierend auf dem In-Session-Verhalten der Individuen vorhersagt, wie Filter und Filterwerte sortiert werden sollten, um die höchste Relevanz für Nutzer*innen zu bieten. Der Vorteil ist, dass diese Informationen auch verwendet werden können, wenn aufgrund eines Opt-Outs der User*in durch den Cookie-Banners keiner Datenspeicherung zugestimmt wird. So können relevante Informationen in Bezug auf die Filternutzung je Suchbegriff trotzdem im Laufe einer Session ausgespielt werden.

Eine explizite Sammlung von Präferenzen basierend auf den untersuchten Filtern kann dann vollzogen werden, wenn User*innen der Speicherung von Daten zustimmen. Dabei ist zu empfehlen, Abfrage-Masken für die relevantesten Filter zu gestalten und implementieren. Eine große Herausforderung stellt hier die Vielfältigkeit des Sortiments und die Long-Tail-Produktstrategie von ottoversand.at. Wie in Kapitel 5 festgestellt, sind für die Kund*innen pro Sortiment andere Filter mehr oder weniger relevant. So wird beispielsweise im Fashion-Sortiment der Größenfilter am häufigsten genutzt. Der Markenfilter hingegen ist insbesondere bei Suchanfragen im Living- und Technik-Sortiment für die User*innen besonders relevant. Um erste und schnelle Erfahrungswerte in Bezug auf die explizite Datenspeicherung zu sammeln, kann daher empfohlen werden, sich auf die zwei meistgenutzten Filter (Farbe und Größe) zu fokussieren und bei den relevantesten Suchbegriffe (siehe Tabelle 5) je nach Sortiment eine Eingabemaske auszuspielen. Beispiele von Online Shops, wo eine entsprechende Funktionalität bereits erfolgreich im Einsatz ist, sind AboutYou und Zalando. Bei diesen Shops haben Kund*innen die Möglichkeit, ihre verschiedenen Präferenzen, Lieblingsfarben und Marken oder Kleidungs- und Schuhgrößen zu speichern. Basierend auf diesen Angaben werden dann entsprechende Produkte auf Suchergebnisseiten oder Produktlisten vorgefiltert oder vorgereiht. Am Beispiel von AboutYou (Abbildung 26) sieht man, wie eine entsprechende Eingabemaske aussehen kann. Diese Daten können dann bei der Sortierung von Produktlisten oder bei einer Hervorhebung von entsprechenden Produkten verarbeitet werden. Auch in anderen Kanälen wie E-Mail, Affiliate, App und App-Push oder auch in Print-Werbemittel etc. könnten diese Userdaten weiterverwendet werden, wenn ein entsprechendes Opt-In bzw. Permission vorliegt. Eine andere Möglichkeit der expliziten Datensammlung kann sein, dass User*innen ihren Lieblingsmarken folgen können. Das könnte beispielsweise so umgesetzt werden, dass sie auf der Marken-Übersichtsseite im OTTO Online Shop via Herz (ähnlich wie das „Merkzettel“-Herzchen) die bevorzugten Marken speichern. Diese Präferenz sollte

auch wieder änderbar/rückgängig gemacht werden können (z.B. direkt über das „entliken“ der Marke oder im Mein Konto-Bereich).

Dieser Mix aus explizit gesammelten Daten bei bekannten User*innen und der Datenanreicherung durch die implizite Sammlung von Klick-Informationen basierend auf dem Nutzungsverhalten kann eine sehr gute Grundlage für eine Filterungs-Personalisierungs-Strategie darstellen.

ABOUT EDITH

Deine Einstellungen

Lieblingsmarken

Deine Größen

Benachrichtigungen

Zu Deinen Bestellungen

Welche Größen trägst du?

Konfektionsgröße

26	28	30	32	34	36	38	40	42
44	46	48	50	52	54	56	58	60
62	64	66	68	70	72	74	76	80
84	86							

EU-Größe

XXXS	XXS	XS	S	M	L	XL	4XS	10XL
XXL	XXXL	4XL	5XL	6XL	7XL	8XL	9XL	

Jeans-Weite

21	22	23	24	25	26	27	28	29
30	31	32	33	34	35	36	37	38

Abbildung 26: Explizite Eingabe von Größenangaben einer User*in.
Quelle: aboutyou.at, 2020

Die Filterung nach Farben ist für die OTTO-Kund*innen am relevantesten. Um eine optimale Bedienung der Filterung und eine schnelle Nutzung zu ermöglichen, sind

verschiedene Ideen denkbar. Zum einen können je Sortiment die relevantesten Farbwerte innerhalb des Filters hervorgehoben und beispielsweise als „beliebteste Farben für deinen Suchbegriff“ vorgereicht und gekennzeichnet werden. Eine andere Möglichkeit, die User*innen ohne Umwege zu den relevanten Artikeln zu navigieren kann sein, die relevanteste Farbe pro User aufgrund trainierter Modelle und auch explizit gesammelter Daten zu vorhersagen und die verschiedenen Ausführungen der Artikel auf der Produktliste anzuzeigen. Ist beispielsweise eine Couch in grau, beige, schwarz und blau verfügbar und eine Prädiktion liegt vor, dass ein Individuum aufgrund vorrangigener Filternutzung oder manuell hinterlegten Farb-Affinitäten auf der Suche nach grauen Couchen ist, kann diese Information bei der Darstellung des Artikel-Repräsentanten schon verarbeitet und jeweils der passende Repräsentant angezeigt werden.

Eine weitere Möglichkeit kann sein, die relevantesten Farbwerte nicht innerhalb des Farbfilters optisch hervorzuheben, sondern beispielsweise die Top-3-Farben des relevantesten Filters – als eigene Filter herausgestellt – in Form einer „1-Click-Filterung“ bei entsprechenden Suchbegriffen zu ermöglichen. Dies hat den Vorteil, dass dynamisch und datengetrieben je Suchbegriff die wichtigsten Filter schnell und unkompliziert genutzt werden können. Zum Abschluss kann auch angedacht werden, in der Suche ein Feature zu verbauen, dass anhand der Such-Query die relevantesten Farben erkennt und so direkt den entsprechenden Farb-Filterwert setzt. Sucht ein Kunde beispielsweise nach „Ballerina rot“, erkennt die *Magellan*-Suche den Suchstring Ballerina und setzt automatisch den Farb-Filterwert „rot“, um ein optimales Suchergebnis zu liefern.

Obwohl der Größenfilter nur im Fashion-Sortiment relevant ist, ist es nach den Erkenntnissen der empirischen Untersuchungen in Kapitel 5 der zweit-wichtigste Filter im OTTO Online Shop. Die Schwierigkeit stellt auf alle Fälle die Vielzahl an verschiedenen Größenläufen dar, die sich auch kaum bis gar nicht matchen oder verheiraten lassen. Daher sollten im ersten Schritt einfache Use Cases identifiziert werden. Ein Quick-Win könnte der Fokus auf Suchbegriffe sein, bei denen

Filterungen nach Große Größen eine Relevanz bei Kund*innen ergeben haben. Oder ähnlich wie beim Farbfilter direkt über die *Magellan*-Suchfunktion anhand der Such-Query jene Begriffe ermitteln, wo etwa „Große Größe“, „Plus-Size“, „Übergröße“ etc. bereits Teil der Suchanfrage ist. Auf den entsprechenden Suchergebnisseiten können dann z.B. die erkannten Größenfilter bereits voreingestellt sein. Eine weitere Möglichkeit stellt der Use Case dar, dass bei Suchanfragen rund um Große Größen ein Teaser in Banner- oder Produktkachel-Form auf die entsprechende Produktfilterung verlinkt und so der Filter in inspirativer Weise dargestellt wird.

Außerdem lohnt es sich, wie auch bereits anhand Abbildung 26 gezeigt, eine Eingabe-Maske für die bevorzugten Kleidungs- und Schuhgrößen anzubieten. So können User*innen selber dazu beitragen, dass der Online Shop lernt, welche Artikel für das Individuum die relevantesten sind und präferiert ausgespielt werden sollten. Diese gesammelten Informationen lassen sich wiederum mit den Web-Tracking Informationen verknüpfen. Auch an anderer Stelle ist diese Information sehr gut in eine Personalisierungsstrategie integrierbar. So können beispielsweise in Produkt-Recommendations die entsprechenden Style-Varianten mit der angegebenen Größe dargestellt werden oder im E-Mail-Marketing perfekt zugeschnittene Newsletter-Trigger und -Kampagnen versendet werden.

In Bezug auf den Preisfilter ist insbesondere die Sale-Affinität der Kunden herauszustreichen. So kann der Sale-Filter bei den entsprechenden Suchanfragen noch prominenter als eigener Filter dargestellt werden und nicht nur als Checkbox innerhalb des Preisfilters. Außerdem sollte bei Sale-affinen Kunden die Strategie durchdacht werden, dass anstatt Gutschein- und VF-Kampagnen verstärkt Sale-Kampagnen geplant werden, da so Einsparungen bei den Marketingkosten erreichbar sind. Eine entsprechende Segmentierung der Kunden kann über verschiedene Tools und Dienstleister im OTTO Online Shop umgesetzt werden. So sind beispielsweise *Dynamic Yield* und *prudsys* im Einsatz, mit denen nicht nur Segmentierung sondern 1:1 Individualisierung vorgenommen werden kann.

Blickt man auf den Markenfilter können hier zum einen bei entsprechendem Opt-In Kurz- und Langzeit-Affinitäten der einzelnen User berechnet und gespeichert werden. Zum anderen kann bei fehlendem Cookie-Banner-Opt-In im Shop aufgrund des Nutzungs- und Filterungsverhalten der Individuen verhaltensbasierte Personalisierung betrieben werden. Durch der Filterung auf Suchergebnisseiten sind sortimentsbasierte Marken-Cluster, häufig zusammen gefilterte Marken und schlussendlich auch zusammen gekaufte Cluster bildbar. Mittels dieser Cluster können auch ähnliche Marken in Form von inspirativen Content-Elementen im Shop beworben werden – beispielsweise durch Banner oder einer Markenleiste. Eine passende Headline für diese entsprechenden personalisierten Content-Elemente könne „Basierend auf deinen Lieblingsmarken wählten andere Kunden auch...“ sein. In Bezug auf den Markenfilter lassen sich die Informationen wie auch schon beim Größenfilter bei entsprechendem Opt-In wieder explizit über eine Eingabemaske von den User*innen abfragen und wegspeichern; sie sollten natürlich auch jederzeit änderbar sein.

Ein wichtiger Punkt in Bezug auf die vorgeschlagenen Implementierungen ist, diese auch unbedingt direkt am Kunden in Form von A/B-Tests live abzutesten und zu bewerten, ob die aufgestellten Use-Case-Annahmen auch zutreffen. So könnte zum Beispiel getestet werden, ob die entsprechenden SERPs besser performen, wenn eine transparente Kommunikation stattfindet, dass Suchergebnisseiten basierend auf gesammelten Daten über Interessen und Präferenzen der User*innen vorgefiltert werden. Im Gegensatz zu Produktlisten, bei denen die Produkt-Auswahl ohne einer Kommunikation an die Kund*innen gefiltert und die Produkte entsprechend eingeschränkt werden. Mögliche Kennzahlen, die darüber eine Aussagekraft geben, sind beispielsweise die Ausstiegsrate, die CTR auf eine Detailview oder die Conversion Rate.

6.2 Zusammenfassung und Fazit

Das Zeitalter der Digitalisierung hat dazu geführt, dass ein neues Paradigma des Handelns eingeleitet worden ist. Der E-Commerce hat sich zum U-Commerce – dem Ubiquitous Commerce – weiterentwickelt und führt dazu, dass Kund*innen allgegenwärtig Handeln können. Es findet ein ständiger Austausch, ein nahtloser Strom bestehend aus Kommunikation, Content und Dienstleistungen zwischen Lieferant*innen, Mitarbeiter*innen, Kund*innen und Produkten statt.

Diese ultimative Form des Handels führt vor allem dazu, dass nicht der Markt sich verändert, sondern die Kund*innen. Die 4. Generation des Commerce wird angeführt von jüngeren und einflussreichen Individuen, welche sich mobile, digital informierte, schnelllebige und emotional engagierte Online-Erlebnisse rund um die Uhr erwarten (Freeman, 2017). Mit Fokus auf den Touchpoint der Onsite Search, streicht diese Masterarbeit heraus, dass es sich Online-Retailer zur Aufgabe machen müssen, basierend auf den Bedürfnissen, Präferenzen und Affinitäten der User*innen anhand von Personalisierung, die relevantesten Produkte und Inhalte auf Suchergebnisseiten anzubieten. Personalisierung sollte insbesondere dabei helfen, eine Reizüberflutung bei den Kund*innen zu vermeiden und auch dem Prinzip des Paradox of Choice – also dem Auswahl-Dilemma von Schwartz (2005) – entgegenzuwirken. Einen großen Einfluss auf eine gezielte und datengetriebene Präsentation der relevantesten Artikel und Services bringen die Filter auf den SERPs mit sich. Schafft es OTTO Österreich, die angeführten Handlungsempfehlungen iterativ zu konzeptionieren, umzusetzen und livezustellen, ist sich die Autorin sicher, dass dadurch nicht nur auf Seiten des Online Shops Mehrwert in Form von einer Hebelung der Conversion erzielt werden kann. Sondern dass auch der Mehrwert für die Kund*innen deutlich gesteigert und die Customer Experience verbessert werden kann.

Zur Beantwortung der zu Grunde liegenden Forschungsfrage wurden in Abschnitt 6 wichtige Erkenntnisse dargestellt und näher erläutert. Diese umfassen

beispielsweise, dass es große Unterschiede innerhalb der gesuchten Sortimente gibt, welche Filter wie relevant sind. Daher macht es Sinn, in einem Proof of Concept (PoC) erstmals die häufigsten 25-30 Suchbegriffe mit der höchsten Filternutzung genauer zu analysieren und erste Use Cases bzw. Algorithmen zu definieren. Außerdem wird empfohlen, nicht nur eine implizite Datensammlung zu praktizieren, sondern die User*innen auch explizit dazu aufzufordern, mittels Eingabemasken Informationen zu präferierten Größen, Marken und Farben zu geben. Dieses Prinzip der Datensammlung kann auch auf die Kommunikation der Marketing-Strategie umgemünzt werden. So sollte beispielsweise getestet werden, ob ein größerer Mehrwert hinsichtlich erzeugtes Vertrauen und Transparenz geschafft werden kann, wenn die Personalisierungsmaßnahmen explizit kommuniziert werden. Beispielsweise mit „Um dir die beste Auswahl an Artikel deiner Lieblingsmarken bieten zu können, haben wir dein Suchergebnis vorgefiltert. Möchtest du lieber alle Artikel zu deiner Suchanfrage sehen, lösche einfach den Markenfilter.“ kann OTTO transparent und verständlich kommunizieren, dass die angezeigten Artikel personalisiert ausgespielt werden. Dadurch kann der Online-Retailer auch dem „Paradoxon der Personalisierung und Privatsphäre“ von entgegengewirkt. Denn auch wenn User*innen in Umfragen und Marktforschungen immer wieder angeben, dass sie skeptisch sind, wenn ihre persönlichen Daten für unternehmerische Zwecke verwendet werden, wiegt im Alltag der Nutzen von datengetriebenen Services diese Bedenken auf.

Um den Herausforderungen hinsichtlich Datenschutz und möglichen Filter-Bubbles Herr zu werden, muss unbedingt auf den Schutz der personenbezogenen Daten und eine datenschutzfreundliche Gestaltung der Benutzerprofile geachtet werden. Dabei können die technischen Konzepte der Anonymisierung oder Pseudonymisierung angewendet werden, welche das Prinzip der Datensparsamkeit verfolgen.

6.3 Limitationen der Arbeit und Ausblick

Im Zuge dieser Masterarbeit wurde ein sehr eng gestecktes und spitzes Themengebiet beleuchtet. Natürlich konnten in diesem Rahmen einige Gesichtspunkte nicht genauer betrachtet werden. Die nachfolgenden Punkte und Themen zeigen auf, in welche Richtung weiterführenden Forschungen gehen können.

Zum einen beziehen sich die Forschungsergebnisse auf Suchergebnisseiten, Kategorieseiten, welche über die Navigation erreichbar sind, sind im Forschungsspektrum nicht beinhaltet. Dies hat den Grund, dass der Umfang der Forschung den Rahmen ansonsten gesprengt hätte. Die Intension der User*innen ist beim Navigieren auf Kategorie-Produktlisten meist viel stöbernder, die Kund*innen lassen sich auf der Customer Journey gerne inspirieren. Je nachdem, auf welcher (Unter-)Menüebene sich User*innen befinden, wird auf den Kategorieseiten oftmals noch ein viel breiteres Feld mit einer Vielzahl an Produkten angesteuert. Daher ist die Hypothese, dass die Nutzung der Filter auf Produktlisten viel relevanter ist, als auf Suchergebnisseiten. Denn beim Ausführen einer Suchanfrage werden deutlich häufiger sehr konkrete Suchbegriffe eingegeben, die die Produktauswahl sowieso schon einschränken. Mit dieser Annahme könnte es also sehr interessant für die Online Shop-Verantwortlichen sein, wie hoch die Filternutzungsrate auf Produktlisten ist und ob sich ähnliche Erkenntnisse wie auf Suchergebnisseiten ableiten lassen.

Zudem wurden nur sechs ausgewählte Filter im Rahmen dieser Arbeit beleuchtet, welche basierend auf den Produktattributen auch innerhalb fast allen Sortimenten vorkommen. Auf SERPs im OTTO Online Shop sind aber noch viele weitere Filter implementiert, wie beispielsweise Laptop- oder Fernsehdiagonalen, Schaft- oder Absatzhöhe bei Schuhsortimenten, Kleider- oder Ärmellängefilter, Abmessungen im Möbelsortiment, Spezialfilter bei Haushaltselektro-SERPs etc. Diese könnten in

Bezug auf ihre Relevanz ebenfalls noch in weiteren Forschungen analysiert, bewertet und interpretiert werden.

Eine interessante Untersuchung, die in weiterführenden Forschungen angestrebt werden kann, ist die Performance der Filter in Hinblick auf die Click-Through-Rate auf Artikeldetailseiten und die Conversion Rate bzw. Kaufabschlüssen. Hier könnte beispielsweise bewertet werden, ob SERPs mit einer Filterung besser konvertieren, als SERPs ohne Filternutzung. Die Faktoren CTR und generierter Online-Umsatz je Filter sind in den Handlungsempfehlungen und Implikationen grundsätzlich noch nicht mitbeachtet, das sollte auf alle Fälle aufgrund gesammelter Informationen und Daten weiter verfolgt werden. Im Zuge dieser Masterarbeit war dies allerdings nicht möglich, da diese Daten in den Datensätzen nicht vorhanden waren.

Außerdem werden auch Filter, die von User*innen wieder gelöscht werden, noch nicht betrachtet. Hier werden also nur „positive“ Klicks bewertet, sprich das Setzen von entsprechenden Filterwerten. In weiteren Forschungen kann analysiert werden, wie sich die Performance von Suchergebnisseiten ändert, wenn beispielsweise ein Algorithmus implementiert wird, der gelöschte Filterwerte „abstrafft“ und so ein noch genaueres Bild über die Relevanz einzelner Filter und Filterwerte liefert.

Nachfolgend werden Fragestellungen aufgelistet, die im Zuge weitere Analysen und Forschungen beleuchtet werden können:

- Performen Suchanfragen mit gesetzten Filtern besser, als Suchanfragen ohne gesetzten Filtern? (basierend auf Conversion Rate, durchschnittlichen WK-Wert, o.ä.)
- Welche Filter werden häufig wieder gelöscht?
- Welche Erkenntnisse lassen sich in Hinblick auf die Suchanfrage-Performance bei gelöschten Filtern ableiten?
- Wie verhalten sich User*innen in ihrer Customer Journey, die bestimmte Filter setzen?

- Mit welchen Filter(-kombinationen) werden (überdurchschnittlich) hohe Warenkörbe generiert und abgeschickt?
- Gibt es Unterschiede bei der Filternutzung zwischen den verschiedenen Devices App, Desktop, Mobile und Tablet?

Basierend auf Erkenntnissen und eventuellen Forschungsergebnissen der genannten möglichen Next Steps kann in einer weiteren Phase die Personalisierungsstrategie der Filternutzung ausgebaut und um neue relevante Use Cases erweitert werden.

Da die durchgeführte Literaturrecherche und die Forschungslücke zeigen, dass in dem Bereich der Filternutzung im Online Shop und daraus resultierende Personalisierungsmaßnahmen kaum relevante Literatur und Forschungsergebnisse vorliegen, kann davon ausgegangen werden, dass weiterführende Forschungen dazu beitragen, Innovationen im E-Commerce voranzutreiben und Mehrwert sowohl auf Customer-Seite als auch auf Unternehmens-Seite generieren zu können.

Literaturverzeichnis

Aboutyou.at. (2020). <https://www.aboutyou.at/>

Adam, N. R., & Yesha, Y. (Hrsg.). (1996). *Electronic Commerce: Current Research Issues and Applications*. Springer-Verlag.

Alba, J., Lynch, J., Weitz, B., Janiszewski, C., Lutz, R., Sawyer, A., & Wood, S. (1997). Interactive Home Shopping: Consumer, Retailer, and Manufacturer Incentives to Participate in Electronic Marketplaces. *Journal of Marketing*, 61(3), 38.

Amazon.de. (2020). <https://www.amazon.de/>

Andrade, E. B., Kaltcheva, V., & Weitz, B. (2002). Self-Disclosure on the Web: The Impact of Privacy Policy, Reward, and Company Reputation. *NA - Advances in Consumer Research*, 29, 350–353.

Ansari, A., & Mela, C. F. (2003). E-Customization. *Journal of Marketing Research*, 40(2).

Anshari, M., Almunawar, M. N., Lim, S. A., & Al-Mudimigh, A. (2018). Customer relationship management and big data enabled: Personalization & customization of services. *Applied Computing and Informatics*.

Arora, N., Dreze, X., Ghose, A., Hess, J. D., Iyengar, R., Jing, B., Joshi, Y., Kumar, V., Lurie, N., Neslin, S., Sajeesh, S., Su, M., Syam, N., Thomas, J., & Zhang, Z. J. (2008). Putting one-to-one marketing to work: Personalization, customization, and choice. *Marketing Letters*, 19(3–4), 305–321.

Arthur, D., & Vassilvitskii, S. (2007). K-means++: The advantages of careful seeding. *In Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms*.

Awad, & Krishnan. (2006). The Personalization Privacy Paradox: An Empirical Evaluation of Information Transparency and the Willingness to Be Profiled Online for Personalization. *MIS Quarterly*, 30(1), 13.

Bailey, S. (2020). E-Commerce und stationärer Handel – Omnichannel Commerce für die perfekte Customer Journey. *Wirtschaftsinformatik & Management*, 12(1), 6–13.

Bauer, F., & Koth, H. (2014). *Der unvernünftige Kunde: Mit Behavioural Economics irrationale Entscheidungen verstehen und beeinflussen*. Redline Wirtschaft.

- Beasley, M. (2013). *Practical Web Analytics for User Experience: How Analytics Can Help You Understand Your Users*. Morgan Kaufmann.
- Becker, M. (2019). Privacy in the digital age: Comparing and contrasting individual versus social approaches towards privacy. *Ethics and Information Technology*, 21(4), 307–317.
- Bozdog, E. (2013). Bias in algorithmic filtering and personalization. *Ethics and Information Technology*, 15(3), 209–227.
- Bundesverband Digitale Wirtschaft (BVDW) e. V., aberratio G., Hamburg. (2012, September 14). *Customer Journey—Definitionen und Ausprägungen, DMEXCO Seminarfolien*.
<https://www.bvdw.org/themen/publikationen/detail/artikel/customer-journey-definitionen-und-auspraegungen-dmexco-seminarfolien/>
- Chellappa, R. K., & Sin, R. G. (2005). Personalization versus privacy: An empirical examination of the online consumer's dilemma. *Information Technology & Management. Information & Management*, 6, 181–202.
- Chen, Y.-C., Shang, R.-A., & Kao, C.-Y. (2009). The effects of information overload on consumers' subjective state towards buying decision in the internet shopping environment. *Electronic Commerce Research and Applications*, 8(1), 48–58.
- Datta, A., Dutta, K., VanderMeer, D., Ramamritham, K., & Navathe, S. B. (2001). An architecture to support scalable online personalization on the Web. *The VLDB Journal*, 10(1), 104–117.
- De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of Big Data based on its essential features. *Library Review*, 65, 122–135.
- de Ville, B. (2001). *Microsoft Data Mining*. Elsevier.
- Dorschel, J. (2015). *Praxishandbuch Big Data: Wirtschaft – Recht – Technik*. Springer-Verlag.
- Einfinger, M. (2003). *Web Content Management: Effizientes Management von Websites durch Content Management Systeme, mit einer empirischen Studie in Österreich und Kanada*. diplom.de.
- Esch, F.-R., & Knörle, C. (2015). Omni-Channel-Strategien durch Customer-Touchpoint-Management erfolgreich realisieren. In *Digitalisierung im Vertrieb: Strategien zum Einsatz neuer Technologien in Vertriebsorganisationen* (S. 123–137). Springer-Verlag.

- Fan, S., Lau, R. Y. K., & Zhao, J. L. (2015). Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix. *Big Data Research*, 2(1), 28–32.
- Fawcett, T., & Provost, F. (2013). *Data Science for Business*. O'Reilly Media.
- Fayyad, U., & Uthurusamy, R. (2002). Evolving data into mining solutions for insights. *Communications of the ACM*, 45(8), 28–31.
- Fitzgerald, B. (2017, Oktober 12). *Filter Bubbles and Privacy, and the Myth of the Privacy Setting*. FunnyMonkey. <https://funnymonkey.com/2017/filter-bubbles-and-privacy-and-the-myth-of-the-privacy>
- Franco, L. K., Rosa, J. H., Barbosa, J. L. V., Costa, C. A., & Yamin, A. C. (2011). MUCS: A model for ubiquitous commerce support. *Electronic Commerce Research and Applications*, 10(2), 237–246.
- Freeman, R. W. (2017, Mai 31). *Understanding Customer 4.0 – the Customer led revolution*. <https://wcomc.org/2017May31-2>
- Gauch, S., Speretta, M., Chandramouli, A., & Micarelli, A. (2007). User Profiles for Personalized Information Access. In P. Brusilovski, A. Kobsa, & W. Nejdl (Hrsg.), *The Adaptive Web* (S. 54–89). Springer-Verlag.
- Grabmeier, J., & Rudolph, A. (2002). Techniques of Cluster Algorithms in Data Mining. *Data Mining and Knowledge Discovery*, 6(4), 303–360.
- Graf, A., & Schneider, H. (2016). *The E-Commerce Book: About a channel that became an industry*. dfv Mediengruppe Fachbuch.
- Greco, A. (2012). *The Adobe SiteCatalyst Handbook: An Insider's Guide*. Pearson Education.
- Greco, G., Greco, S., & Zumpano, E. (2004). Collaborative Filtering Supporting Web Site Navigation. *AI Communications*, 17, 155–166.
- Grimmer, U., & Mucha, H.-J. (2013). Datensegmentierung mittels Clusteranalyse. In G. Nakhaeizadeh (Hrsg.), *Data Mining: Theoretische Aspekte und Anwendungen* (S. 109–141). Springer-Verlag.
- Groissberger, T., & Riedl, R. (2018). Interaktive Informationsmanagement-Tools in Online Shops: Studienergebnisse und Gestaltungsempfehlungen. In *Digital Customer Experience: Mit digitalen Diensten Kunden gewinnen und halten*. Springer-Verlag.

- Gross, R. (2014). A Theoretical Consumer Decision Making Model: The Influence of Interactivity and Information Overload on Consumers Intent to Purchase Online. *International Journal of Business Management and Economic Research*, 5(4), 64–70.
- Guo, W., Wu, S., Wang, L., & Tan, T. (2016). Personalized ranking with pairwise Factorization Machines. *Neurocomputing*, 214, 191–200.
- Habegger, B., Hasan, O., Brunie, L., Bennani, N., Kosch, H., & Damiani, E. (2014). Personalization vs. Privacy in Big Data Analysis. *International Journal of Big Data*, 25–35.
- Hassler, M. (2011). *Web Analytics: Metriken auswerten, Besucherverhalten verstehen, Website optimieren* (3., überarbeitete Auflage 2012). mitp-Verlag.
- Heinemann, G. (2019). Geschäftsmodell des Online-Handels. In G. Heinemann, *Der neue Online-Handel* (S. 41–138). Springer Fachmedien Wiesbaden.
- Heinemann, G. (2020). *Der neue Online-Handel: Geschäftsmodelle, Geschäftssysteme und Benchmarks im E-Commerce*. Springer-Verlag.
- Henrici, M. (2012, April 24). *Verhaltensmuster „Paradox-of-Choice“: Große Auswahl, kleine Conversion?* konversionsKRAFT. <https://www.konversionskraft.de/analysen/grosse-auswahl-geringe-conversion-paradox-of-choice-in-der-praxis.html>
- Henry, W. A. (1980). The Effect of Information-Processing Ability on Processing Accuracy. *Journal of Consumer Research*, 7(1), 42–48.
- Holland, H., & Flocke, L. (2014). Customer-Journey-Analyse: Ein neuer Ansatz zur Optimierung des (Online-)Marketing-Mix. In *Digitales Dialogmarketing: Grundlagen, Strategien, Instrumente* (S. 826–853). Springer-Verlag.
- Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, 79(6), 995–1006.
- Jacoby, J., Speller, D. E., & Berning, C. K. (1974). Brand Choice Behavior as a Function of Information Load: Replication and Extension. *Journal of Consumer Research*, 1(1), 33–42.
- Junglas, I., & Watson, R. (2003). U-Commerce: A Conceptual Extension of E-Commerce and M-Commerce. *ICIS 2003 Proceedings*, 55.

- Jürgens, P., Stark, B., & Magin, M. (2014). Gefangen in der Filter Bubble? Search Engine Bias und Personalisierungsprozesse bei Suchmaschinen. In B. Stark, D. Dörr, & Aufenanger (Hrsg.), *Die Googleisierung der Informationssuche: Suchmaschinen zwischen Nutzung und Regulierung* (S. 98–135). De Gruyter.
- Kalyanaraman, S., & Sundar, S. S. (2006). The Psychological Appeal of Personalized Content in Web Portals: Does Customization Affect Attitudes and Behavior? *Journal of Communication*, 56(1), 110–132.
- Kawohl, J. (2017). *Digitale Customer Journey—Eine Untersuchung der DAX-Unternehmen*. Hochschule für Technik und Wirtschaft Berlin.
- Klug, K., & Strang, C. (2018). *Das Phänomen der Filter Bubble: Eine explorative Analyse der Wahrnehmung und der Akzeptanz personalisierter Informationen in Sozialen Medien am Beispiel des Facebook-Newsfeed*. 47.
- Knopp, M. (2015). Pseudonym–Grauzone zwischen Anonymisierung und Personenbezug. *Datenschutz und Datensicherheit - DuD*, 39(8), 527–530.
- Kotler, P., Armstrong, G., Wong, V., & Saunders, J. (2010). *Grundlagen des Marketing* (5. Aufl.). Pearson Studium.
- Kotler, P., & Bliemel, F. (2001). *Marketing-Management: Analyse, Planung und Verwirklichung*. Schäffer-Poeschel.
- Kotler, P., Kartajaya, H., & Setiawan, I. (2016). *Marketing 4.0: Moving from Traditional to Digital*. John Wiley & Sons.
- Krumm, J. (2018). *Ubiquitous Computing Fundamentals*. CRC Press.
- Kumar, S., Joshi, P., & Saquib, Z. (2015). *Ubiquitous Commerce: The New World of Technologies*. 1(2), 50–55.
- Kwon, K., & Kim, C. (2012). How to design personalization in a context of customer retention: Who personalizes what and to what extent? *Electronic Commerce Research and Applications*, 11(2), 101–116.
- Lee, C. H., & Cranage, D. A. (2011). Personalisation–privacy paradox: The effects of personalisation and privacy assurance on customer responses to travel Web sites. *Tourism Management*, 32, 987–994.
- Lewandowski, D. (2018). Personalisierung und Kontextualisierung. In F. Schade & U. Georgy (Hrsg.), *Praxishandbuch Informationsmarketing* (S. 334–342). De Gruyter.

- Liu, Q., Reiner, A. H., Frigessi, A., & Scheel, I. (2019). Diverse personalized recommendations with uncertainty from implicit preference data with the Bayesian Mallows model. *Knowledge-Based Systems*, 186, 27.
- Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2), 129–137.
- MacQueen, J. (1967). *Some methods for classification and analysis of multivariate observations*. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics. <https://projecteuclid.org/euclid.bsmsp/1200512992>
- maximizer crm. (2018, September 13). What Is The New Customer Journey? (And How to Master It). *Maximizer CRM Software*. <https://www.maximizer.com/what-is-the-new-customer-journey/>
- Mayer, S. (2017, April 27). „Paradox of Choice“ – Erleichtere Deinen Kunden die Wahl! *konversionsKRAFT*. <https://www.konversionskraft.de/ueberzeugungskraft/ueberzeugungskraft-video-paradox-of-choice.html>
- McGuire, W. J. (1976). Some Internal Psychological Factors Influencing Consumer Choice. *Journal of Consumer Research*, 2(4), 302–319. JSTOR.
- Meier, A., & Zumstein, D. (2012). *Web Analytics & Web Controlling: Webbasierte Business Intelligence zur Erfolgssicherung*. dpunkt.verlag.
- Menthe, T., & Sieg, M. (2017). *Kundennutzen – Schlüssel zum Verkaufserfolg: Wie Sie Mehrwert bieten, Preise leichter durchsetzen und Profitabilität sichern*. Springer-Verlag.
- Mertens, P., Stößlein, M., & Zeller, Th. (2004). *Personalisierung und Benutzermodellierung in der betrieblichen Informationsverarbeitung Stand und Entwicklungsmöglichkeiten*. Universität Erlangen-Nürnberg. <https://docplayer.org/19844573-Personalisierung-und-benutzermodellierung-in-der-betrieblichen-informationsverarbeitung-stand-und-entwicklungsmoeglichkeiten.html>
- Müller, G., Sackmann, S., Günther, O., & Spiekermann, S. (2007). Personalisierung und informationelle Selbstbestimmung Ein unauflösbarer Widerspruch? *WIRTSCHAFTSINFORMATIK*, 49(1), 1–2.
- Murthi, B., & Sarkar, S. (2003). The Role of the Management Sciences in Research on Personalization. *Management Science*, 49, 1344–1362.

- Nguyen, T. T., Hui, P.-M., Harper, F. M., Terveen, L., & Konstan, J. A. (2014). Exploring the filter bubble: The effect of using recommender systems on content diversity. *Proceedings of the 23rd International Conference on World Wide Web - WWW '14*, 677–686.
- Nowacki, F. (2014). Marketing 4.0—A new concept in the face of modern consumer changes. *Marketing i Rynek, 06*, 11–19.
- Official Google Blog. (2009, Dezember 4). Personalized Search for everyone. *Official Google Blog*. <https://googleblog.blogspot.com/2009/12/personalized-search-for-everyone.html>
- Orenga-Roglá, S., & Chalmeta, R. (2016). Social customer relationship management: Taking advantage of Web 2.0 and Big Data technologies. *SpringerPlus, 5*(1).
- OTTO GmbH. (2020a). *Ottoversand.at*. <https://www.ottoversand.at/>
- OTTO GmbH. (2020b, Juni 15). *Sneaker bei OTTO | Sneakers online shoppen*. <https://www.ottoversand.at/s/sneaker/>
- Oulasvirta, A., Hukkinen, J. P., & Schwartz, B. (2009). When more is less: The paradox of choice in search engine use. *Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '09*, 516.
- Pariser, E. (2011a, März). *Eli Pariser: Vorsicht vor „Filter-Blasen“ im Internet*. https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles?language=de
- Pariser, E. (2011b). *The Filter Bubble: What The Internet Is Hiding From You*. Penguin UK.
- Peppers, D., & Rogers, M. (1997). *Enterprise One to One: Tools for Competing in the Interactive Age A currency book One to One Series*. Crown Business.
- Phelps, J., Nowak, G., & Ferrell, E. (2000). Privacy Concerns and Consumer Willingness to Provide Personal Information. *Journal of Public Policy & Marketing, 19*(1), 27–41.
- Pierson, L. (2015). *Data Science For Dummies*. John Wiley & Sons.
- Piller, F. T. (2009, April 2). *Interview: Bruce Kasanoff of NowPossible.com on „Personal=Smarter“* [Blog]. Mass Customization & Open Innovation News. <https://mass->

customization.blogs.com/mass_customization_open_i/2009/04/interview-bruce-kasanoff-of-nowpossiblecom-on-personalsmarter.html

- Punagin, S., & Arya, A. (2015). Privacy in the age of Pervasive Internet and Big Data Analytics – Challenges and Opportunities. *International Journal of Modern Education and Computer Science*, 7, 36–47.
- Rastegari, H., & Shamsuddin, S. M. (2010). Web Search Personalization Based on Browsing History by Artificial Immune System. *International Journal of Advances in Soft Computing and Its Applications*, 2(3), 282–301.
- Riemer, K., & Brüggemann, F. (2007). Personalisierung der Internetsuche: Lösungstechniken und Marktüberblick. *WIRTSCHAFTSINFORMATIK*, 49(2), 116–126.
- Schaar, P. (2012). Systemdatenschutz – Datenschutz durch Technik oder warum wir eine Datenschutz-technologie brauchen. In J.-H. Schmidt & T. Weichert (Hrsg.), *Datenschutz: Grundlagen, Entwicklungen und Kontroversen*. Bpb, Bundeszentrale für politische Bildung.
- Schmecken, G. M. (2008). *Erfolgreiche Strategien für E-Commerce: Integrierte Kosten- und Leistungsführerschaft als Orientierungsmuster*. Springer-Verlag.
- Schneider, H., & Graf, A. (2017). *Das E-Commerce Buch: Marktanalysen - Geschäftsmodelle - Strategien*. dfv Mediengruppe Fachbuch.
- Scholz, J. (2019, Mai 31). *On-Site Search & SEO: Everything You Need to Know*. Search Engine Journal. <https://www.searchenginejournal.com/on-site-search-seo/309189/>
- Schulz-Hardt, S., Frey, D., Lüthgens, C., & Moscovici, S. (2000). Biased information search in group decision making. *Journal of Personality and Social Psychology*, 78(4), 655–669.
- Schwartz, B. (2005). *The Paradox of Choice: Why More Is Less* (Reprint). HarperCollins.
- Schwartz, B. (2014, Jänner 29). *Is the famous „paradox of choice“ a myth?* PBS NewsHour. <https://www.pbs.org/newshour/economy/is-the-famous-paradox-of-choic>
- Semeijn, E. (2018, Jänner 18). Paradox of choice: Why showing less to your customers is more! *Neurofied*. <https://neurofied.com/paradox-of-choice-why-less-more/>

- Sheehan, K. B., & Hoy, M. G. (2000). Dimensions of Privacy Concern among Online Consumers. *Journal of Public Policy & Marketing*, 19(1), 62–73.
- Spreer, P. (2018). *PsyConversion: 101 Behavior Patterns für eine bessere User Experience und höhere Conversion-Rate im E-Commerce* (1. Aufl. 2018). Springer Gabler.
- Statista Research Department. (2015, Mai 20). *Österreich—Einzelkriterien der Benutzerfreundlichkeit bei Online-Shops 2015*. Statista. <https://de.statista.com/statistik/daten/studie/305216/umfrage/einzelkriterien-der-benutzerfreundlichkeit-bei-online-shops-in-oesterreich/>
- Storey, M. (2011, Juni 13). Personalisation Is Changing Search Engine Optimisation. *Michael Storey - founder of Whype*. <http://whype.blogspot.com/2011/06/personalisation-is-changing-search.html>
- Stüber, E. (2013). *Personalisierung im Internethandel: Die Akzeptanz von Kaufempfehlungen in der Bekleidungsbranche*. Springer-Verlag.
- Stüber, E., & Hudetz, K. (2017). *Praxis der Personalisierung im Handel: Mit zeitgemäßen E-Commerce-Konzepten Umsatz und Kundenwert steigern*. Springer-Verlag.
- Teevan, J., Dumais, S. T., & Horvitz, E. (2005). Personalizing search via automated analysis of interests and activities. *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*, 449–456.
- Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, 79(4), 281–299.
- Villars, R. L., Eastwood, M., & Olofson, C. W. (2011). *Big Data: What It Is and Why You Should Care* (Analyze the Future, S. 14) [White Paper]. IDC.
- Vogt, A., & Quelle-Korting, T. (2017). Optimierung der Onsite-Suche am Beispiel von otto.de. *Datenbank-Spektrum*, 17(1), 15–20.
- Volgens, B., & Anderson, K. (2017). *Digital human: Der Mensch im Mittelpunkt der Digitalisierung*. Campus Verlag.
- Watson, R. T., Pitt, L. F., Berthon, P., & Zinkhan, G. M. (2002). U-Commerce: Expanding the Universe of Marketing. *Journal of the Academy of Marketing Science*, 30(4), 333–347.

- Watson, Richard T. (2000, Oktober 31). *U-Commerce: The Ultimate*. Ubiquity: An ACM IT Magazine and Forum. <https://ubiquity.acm.org/article.cfm?id=353882>
- Wegeler, C. (2019, August 27). Pseudonymisierung von personenbezogenen Daten. *activeMind AG*. <https://www.activemind.de/magazin/pseudonymisierung-daten/>
- Weissman, A., & Wegerer, S. (2018). *Digitaler Wandel in Familienunternehmen: Das Handbuch*. Campus Verlag.
- Wereda, W., & Woźniak, J. (2019). Building Relationships with Customer 4.0 in the Era of Marketing 4.0: The Case Study of Innovative Enterprises in Poland. *Social Sciences*, 8(6), 1–27.
- Wiedenbeck, M., & Züll, C. (2010). Clusteranalyse. In C. Wolf & H. Best (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (S. 525–552). VS Verlag für Sozialwissenschaften.
- Wiedmann, K.-P., Buxel, H., & Walsh, G. (2002). Customer profiling in e-commerce: Methodological aspects and challenges. *Journal of Database Marketing & Customer Strategy Management*, 9(2), 170–184.
- Winter, C., Battis, V., & Halvani, O. (2019). Herausforderungen für die Anonymisierung von Daten. In M. Lange & G. Stumme (Hrsg.), *INFORMATIK 2019: 50 Jahre Gesellschaft für Informatik – Informatik für Gesellschaft*. Gesellschaft für Informatik e.V.
- Wright, S. J. (2019). *Digitizing the Customer Journey: Using the Latest Digital Technologies to Support Growth, Efficiency and Delight Customers Throughout the Customer's Touchpoints*. Bluetrees GmbH.
- Zumstein, D., & Gächter, I. (2016). Digital Analytics – Strategien im digitalen Geschäft umsetzen und mit KPIs überprüfen. *HMD Praxis der Wirtschaftsinformatik*, 53(3), 371–388.
- Zwass, V. (1996). Electronic Commerce: Structures and Issues. *International Journal of Electronic Commerce*, 1(1), 3–23.

Abbildungsverzeichnis

Abbildung 1: Vergleich der modernen und der klassischen Customer Journey.....	7
Abbildung 2: Dynamische Änderung des Suchfelds je nach Browser-Breite.....	9
Abbildung 3: Such- & Filterfunktion bei ottoversand.at am Beispiel des Suchworts "Ballerina".....	9
Abbildung 4: Ein k-Means-Clustering mit 90 Punkten in einer Ebene und k=3 Zentroiden. Oben wird der anfängliche Satz an Punkten dargestellt, unten sind die Bewegungspfade der Zentroiden zu sehen.....	14
Abbildung 5: Das Konfitüren-Experiment von Iyengar & Lepper.	17
Abbildung 6: Auswirkungen der Filter-Bubble anhand des Suchbegriffs „BP“ im Jahr 2010.	19
Abbildung 7: Änderungen der Kunden-Einstellungen in Bezug auf die Marktaktivität über Jahrzehnte.	24
Abbildung 8: Die neue Customer Journey 4.0 im E-Commerce	26
Abbildung 9: Customer Journey inklusive On- und Offline Touchpoints der OTTO GmbH.	29
Abbildung 10: SERP einer Suchanfrage mit Filtermöglichkeiten auf ottoversand.at	31
Abbildung 11: Sweet Spot im Paradox of Choice.....	34
Abbildung 12: Die Filter-Bubble.....	40
Abbildung 13: Suchanfrage im OTTO Online Shop mit gesetztem Farb- und Größenfilter.	48
Abbildung 14: OTTO Eigenentwicklung "LMTesTool" zum Nachvollziehen des Suchverhaltens	49
Abbildung 15: Ermittelte Cluster der Suchbegriffe basierend auf den gesetzten Filtern pro Session.	54
Abbildung 16: Absolute Anteile des Markenfilters innerhalb der Sortimente.	58

Abbildung 17: Baumkarte mit drei identifizierten Markenfilter-Clustern.....	63
Abbildung 18: Absolute Nutzung des Preisfilters über alle Sortimente.	66
Abbildung 19: Clustering der Filterwerte des Preis-Filters innerhalb Sortimente.	69
Abbildung 20: Baumkarte mit vier identifizierten Preisfilter-Clustern.	71
Abbildung 21: Baumkarte mit fünf identifizierten Größenfilter-Clustern.	77
Abbildung 22: Baumkarte mit fünf identifizierten Farbfilter-Clustern.	82
Abbildung 23: Clusteranalyse der Geschlechts-Filterwerte in Form einer Baumkarte.	87
Abbildung 24: Clusteranalyse der Lieferzeiten-Filterwerte in Form einer Baumkarte.	90
Abbildung 25: Ermittelte Cluster der User*innen basierend auf den gesetzten Filtern pro Session mit Suchanfrage.....	92
Abbildung 26: Explizite Eingabe von Größenangaben einer User*in.	97

Tabellenverzeichnis

Tabelle 1: Touchpoints und Ausprägungen der Phase "Stöbern & Informieren". ...	28
Tabelle 2: Attribute und Ausprägungen des suche-Landmarks.	47
Tabelle 3: Ausgabe der Filternutzungsrate auf Sortiments-Gruppen Ebene in Form einer Pivot-Tabelle.	50
Tabelle 4: Farbskala der Filternutzungsraten, angewendet auf die einzelnen Sortimente und beleuchteten Filter.	53
Tabelle 5: Die relevantesten Suchbegriffe in Bezug auf die Filternutzung auf ottoversand.at.....	55
Tabelle 6: Analyse der Varianz des Clusterings der Suchbegriffe.	56
Tabelle 7: Nutzungsrate des Markenfilters über alle führenden Sortimente innerhalb von Suchanfragen.....	57
Tabelle 8: Top 30 Filterwerte des Markenfilters.....	60
Tabelle 9: Filterwerte des Markenfilters bei Suchbegriffen mit mindestens 30 Mal gesetztem Filterwert.	61
Tabelle 10: Mehrfach-Auswahlen von Filterwerten des Markenfilters.	62
Tabelle 11: Varianzanalyse der Marken-Values-Custer.	64
Tabelle 12: Nutzungsrate des Preisfilters über alle führenden Sortimente innerhalb von Suchanfragen.....	65
Tabelle 13: Top 30 Filterwerte des Preisfilters.....	67
Tabelle 14: Filterwerte des Preisfilters bei Sortimente mit mindestens 30 Mal gesetztem Filterwert.	68
Tabelle 15: Sortimente des Preisfilters bis 1000 €, mit der Einschränkung, dass eine Preisrange mind. 20x in einem Sortiment gesetzt wurde.....	70
Tabelle 16: Mehrfach-Auswahl von Preisranges und der Sale-Checkbox.....	70
Tabelle 17: Varianzanalyse der identifizierten Preisfilter-Cluster	72

Tabelle 18: Nutzungsrate des Größenfilters über alle führenden Sortimente innerhalb von Suchanfragen.....	73
Tabelle 19: Top 30 Filterwerte des Größenfilters.....	74
Tabelle 20: Filterwerte des Größenfilters auf Sortimentsebene.....	75
Tabelle 21: Mehrfach-Auswahlen von Filterwerten des Markenfilters.	76
Tabelle 22: Varianzanalyse der erkannten Größenfilter-Cluster.....	78
Tabelle 23: Nutzungsrate des Farbfilters über alle führenden Sortimente innerhalb von Suchanfragen.....	79
Tabelle 24: Top 30 Filterwerte des Farbfilters.	80
Tabelle 25: Filterwerte des Farbfilters mit der Einschränkung, dass ein Farb-Filterwert mind. 30x bei Suchanfragen in ein bestimmtes Sortiment gesetzt wurde.	81
Tabelle 26: Mehrfach-Auswahlen von Filterwerten des Farbfilters.	82
Tabelle 27: Varianzanalyse der identifizierten Farbfilter-Cluster.....	83
Tabelle 28: Nutzungsraten der Für-Wen- und Verfügbarkeits-Filter über alle führenden Sortimente innerhalb von Suchanfragen.....	84
Tabelle 29: Top 12 Filterwerte des Geschlechts-Filters („Für Wen?“).	85
Tabelle 30: Filterwerte des Geschlechtsfilters mit der Einschränkung, dass ein Filterwert mind. 20x bei Suchanfragen in ein bestimmtes Sortiment gesetzt wurde.	86
Tabelle 31: Varianzanalyse der identifizierten Geschlechtsfilter-Cluster.....	87
Tabelle 32: Top-30-Filterwerte des Lieferzeitenfilters & Häufigkeit von Mehrfach-Auswahlen.....	88
Tabelle 33: Filterwerte "in 2-4 Werktagen bei dir“ und „in 24 Stunden bei dir“ inkl. Sortimente bzw. Suchbegriffe, innerhalb deren die Filterwerte am häufigsten gesetzt werden.....	89
Tabelle 34: Varianzanalyse der identifizierten Lieferzeitenfilter-Cluster.....	91
Tabelle 35: Analyse der Varianz des Clusterings der User.	93

Anhang

A. Quellcode

A.1. SQL-Abfrage zur Aggregation der Filternutzung

```
1  with sessions as (  
2      select distinct  
3          a.val_session_id  
4          , a.session_end_date  
5          , a.revisit_hash  
6          , year(a.session_end_date)* 100 + month(a.session_end_date) as  
          monat  
7          , b.client_name_parent  
8          , b.client_name  
9      from  
10         argus_unito_raw.sessions_header a  
11     join  
12         argus_unito_raw.clients b  
13     on  
14         a.val_client_id = b.val_client_id  
15     where  
16         a.session_type not in (5, 6, 7) and a.val_type = 1 and  
          b.client_name_parent = 'ottoversand.at' and session_end_date  
          between '2020-07-01' and '2020-07-31'  
17 ),  
18 suche_landmarks as (  
19     select  
20         a.val_session_id  
21         , a.val_lm_sequenceno  
22         , b.search_term  
23         , b.search_id  
24         , b.mkz  
25         , b.geschlecht  
26         , b.marke  
27         , b.preis  
28         , b.groesse  
29         , b.sterne  
30         , b.farbe  
31         , b.lieferzeit  
32     from (  
33         select  
34             id  
35             , val_session_id  
36             , val_lm_sequenceno  
37         from  
38             argus_unito_raw.landmarks_header  
39         where  
40             lm_name = 'suche' and lm_date between '2020-07-01'  
              and '2020-07-31'  
41     ) a  
42     join (  
43         select  
44             id_landmarks_header  
45             , group_concat(case  
46                 when lm_attribute_name = 'stern' then  
47                     lm_attribute_value  
48                 else null  
                end) as search_term
```

```

49         , group_concat(case
50             when lm_attribute_name = 'search-id' then
51                 lm_attribute_value
52             else null
53         end) as search_id
54         , group_concat(case
55             when lm_attribute_name = 'mkz' then
56                 lm_attribute_value
57             else null
58         end) as mkz
59         , group_concat(case
60             when lm_attribute_name = 'sges' then
61                 lm_attribute_value
62             else null
63         end) as geschlecht
64         , group_concat(case
65             when lm_attribute_name = 'smark' then
66                 lm_attribute_value
67             else null
68         end) as marke
69         , group_concat(case
70             when lm_attribute_name = 'spreis' then
71                 lm_attribute_value
72             else null
73         end) as preis
74         , group_concat(case
75             when lm_attribute_name = 'ssize' then
76                 lm_attribute_value
77             else null
78         end) as groesse
79         , group_concat(case
80             when lm_attribute_name =
81                 'filter.filterReviewStars' then
82                 lm_attribute_value
83             else null
84         end) as sterne
85         , group_concat(case
86             when lm_attribute_name = 'filter.filter_color'
87             then lm_attribute_value
88             else null
89         end) as farbe
90         , group_concat(case
91             when lm_attribute_name =
92                 'filter.s_filter_Availability' then
93                 lm_attribute_value
94             else null
95         end) as lieferzeit
96     from
97         argus_unito_raw.landmarks_kv
98     where
99         lm_attribute_name in ('stern', 'search-id', 'sges',
100         'smark', 'spreis', 'ssize', 'filter.filterReviewStars',
101         'filter.filter_color', 'filter.s_filter_Availability', 'mkz') and
102         lm_attribute_value <> '-'
103     group by
104         1
105     ) b
106 on
107     a.id = b.id_landmarks_header
108 ),

```



```

95  gesamt as (
96      select
97          row_number() over (order by a.val_session_id, b.search_id,
98          b.val_lm_sequenceno) as rn
99          , a.client_name_parent
100         , a.client_name
101         , a.monat
102         , a.val_session_id
103         , a.revisit_hash
104         , b.val_lm_sequenceno
105         , b.search_term
106         , b.search_id
107         , b.mkz
108         , c.product_range_1 as textil_hardware
109         , c.product_range_2 as sortiment
110         , b.geschlecht
111         , b.marke
112         , b.preis
113         , b.groesse
114         , b.sterne
115         , b.farbe
116         , b.lieferzeit
117     from
118         sessions a
119     join
120         suche_landmarks b
121     on
122         a.val_session_id = b.val_session_id
123     left join
124         argus_unito_raw.mkz c
125     on
126         b.mkz = c.val_mkz and a.session_end_date between
127         c.val_date_from and c.val_date_to
128 ),
129 gesamt2 as (
130     select
131         client_name_parent
132         , client_name
133         , monat
134         , textil_hardware
135         , sortiment
136         , val_session_id
137         , search_id
138         , search_term
139         , revisit_hash
140         , max(geschlecht) as geschlecht
141         , max(marke) as marke
142         , max(preis) as preis
143         , max(groesse) as groesse
144         , max(sterne) as sterne
145         , max(farbe) as farbe
146         , max(lieferzeit) as lieferzeit
147     from
148         gesamt
149     group by
150         1, 2, 3, 4, 5, 6, 7, 8, 9
151 )
152 select
153     search_id
154     , search_term

```

```

153     , val_session_id
154     , sortiment
155     , textil_hardware
156     , revisit_hash
157     , count(distinct val_session_id) as sessions
158     , count(distinct search_id) as searches
159     , sum(case
160         when coalesce(geschlecht, marke, preis, groesse, sterne,
161             farbe, lieferzeit) is not null then 1
162             else 0
163         end) as filter_gesetzt
164     , sum(case
165         when geschlecht is not null then 1
166             else 0
167         end) as geschlecht_filter
168     , sum(case
169         when marke is not null then 1
170             else 0
171         end) as marke_filter
172     , sum(case
173         when preis is not null then 1
174             else 0
175         end) as preis_filter
176     , sum(case
177         when groesse is not null then 1
178             else 0
179         end) as groesse_filter
180     , sum(case
181         when sterne is not null then 1
182             else 0
183         end) as sterne_filter
184     , sum(case
185         when farbe is not null then 1
186             else 0
187         end) as farbe_filter
188     , sum(case
189         when lieferzeit is not null then 1
190             else 0
191         end) as lieferzeit_filter
192 from
193     gesamt2
194 group by
195     1, 2, 3, 4, 5, 6;

```

Quellcode 1: SQL-Abfrage zur Aggregation der Filternutzung auf Sortiments-Gruppen-Ebene

A.2. Ergänzung der einzelnen Filterwerte der untersuchten Filter

```
151     search_id
152     , search_term
153     , val_session_id
154     , sortiment
155     , textil_hardware
156     , revisit_hash
157     , geschlecht
158     , marke
159     , preis
160     , groesse
161     , sterne
162     , farbe
163     , lieferzeit
164     , count(distinct val_session_id) as sessions
165     , count(distinct search_id) as searches
166     , sum(case
167         when coalesce(geschlecht, marke, preis, groesse, sterne,
168                     farbe, lieferzeit) is not null then 1
169         else 0
170     )
```

Quellcode 2: Ergänzung der einzelnen Filterwerte der untersuchten Filter „geschlecht“, „marke“, „preis“, „groesse“, „sterne“, „farbe“ und „lieferzeit“.