

BACHELORARBEIT

Google Trends als moderne Wahrsagekugel: Wie Google anhand von Suchbegriffen künftige Ereignisse vorhersehen kann

vorgelegt im September 2020
von **Selcuk Acikbas**

1. Prüfer: Prof. Dr. Hardy Gundlach
 2. Prüferin: Diana Jaber
-

**HOCHSCHULE FÜR ANGEWANDTE
WISSENSCHAFTEN HAMBURG**
Department Information
Studiengang Medien und Information

**HOCHSCHULE FÜR ANGEWANDTE
WISSENSCHAFTEN HAMBURG**
Hamburg University of Applied Sciences

Google Trends als moderne Wahrsagekugel: Wie Google anhand von Suchbegriffen künftige Ereignisse vorhersehen kann

Bachelorarbeit vorgelegt von
Selcuk Acikbas

Abstract

Die nachfolgende Arbeit beschäftigt sich mit den Möglichkeiten von Google Trends, künftige Ereignisse anhand von Suchbegriffen der Nutzer vorherzusagen. Google Trends ist ein separates Online-Werkzeug, welches dazu dient, das Suchinteresse der Google-Nutzer im Netz zu repräsentieren. Es wertet die kumulierten Suchbegriffe regional oder global aus und stellt sie in Diagrammen dar. Dadurch lassen sich mögliche Trends in der Zukunft abzeichnen. Dazugehörig ist das Tool Google Flu Trends, welches die Funktionen und Datenbestand von Google Trends verwendet, um Krankheitsherde zu erkennen und auf Basis der eingegebenen Suchbegriffe vorherzusagen.

Dazu wird die folgende Forschungsfrage gestellt: *Kann ein hohes Medieninteresse zu einer bestimmten Krankheit dazu führen, dass sich das Suchverhalten der Nutzer, in Bezug zu Krankheitsthemen, verändert.* Hierfür wurde eine Online-Umfrage entwickelt, die das Suchverhalten von Google-Nutzern untersuchen soll. Die Befragten bewerten ihr eigenes Suchverhalten, als sie zuletzt krank oder gesund waren mit Blick auf die letzten drei Monate. Die aufgestellten Hypothesen untersuchen, ob eine hohe bzw. geringe Medienberichterstattung zu einer vermehrten oder zu einer geringeren Suche nach Krankheitssymptomen im Internet führt.

Die Auswertung der Online-Umfrage zeigt, dass Google-Nutzer in den letzten drei Monaten häufiger nach krankheitsbedingten Themen recherchiert haben. Das bedeutet, dass eine verstärkte Medienpräsenz zu Krankheiten dazu führt, sich vermehrt nach Krankheitssymptomen zu informieren und nicht darauf hindeutet, dass die Personen tatsächlich erkrankt sind.

Auf dieser Grundlage ist es empfehlenswert, die Suchbegriffe differenzierter zu analysieren. Es muss zwischen einer reinen Informationsrecherche und einer krankheitsmotivierten Suche unterschieden werden, um zukünftige Falschprognosen zu vermeiden.

Schlagwörter

Google Trends, Google Flu Trends, Grippewellen, Krankheitswellen, mediale Berichterstattung, GrippeWeb, SORMAS, Covid-19

Abbildungsverzeichnis

Abbildung 1 Fünf Suchbegriffe im Vergleich	S. 3
Abbildung 2 Vergleich Akne und Neurodermitis	S. 5
Abbildung 3 Suchvolumen von Akne und Neurodermitis	S. 6
Abbildung 4 Suchbegriffe mit Anführungszeichen	S. 6
Abbildung 5 Suchbegriffe ohne Anführungszeichen	S. 7
Abbildung 6 Falschaussage von Google Flu Trends 2013 in den USA	S. 10
Abbildung 7 Gewinnspielteilnehmer vs. nicht Gewinnspielteilnehmer	S. 13
Abbildung 8 Wöchentliche Meldungen zwischen 2011-2017	S. 13
Abbildung 9 Ergebnisse der Woche 30/2020	S. 14
Abbildung 10 SORMAS Benutzeroberfläche	S. 15
Abbildung 11 Workflow SORMAS	S. 16
Abbildung 12 Altersangabe	S. 22
Abbildung 13 Geschlechterverteilung	S. 22
Abbildung 14 Angaben zum Wohnort	S. 23
Abbildung 15 Frage zum Berufsstand	S. 24
Abbildung 16 Welche Suchmaschinen werden genutzt	S. 24
Abbildung 17 Von welchen Krankheiten haben die Probanden schon gehört	S. 25
Abbildung 18 Von wo haben die Probanden von den Krankheiten erfahren	S. 26
Abbildung 19 Wann waren die Probanden zuletzt krank	S. 27
Abbildung 20 Suchmaschine-Nutzung bei Krankheit	S. 28
Abbildung 21 Alternative zur Nutzung von Google	S. 29
Abbildung 22 Informationsquellen	S. 30
Abbildung 23 Recherche, während einer Erkrankung	S. 32
Abbildung 24 Vergleich der Antworten	S. 32
Abbildung 25 Recherche in den letzten drei Monaten	S. 33

Inhaltsverzeichnis

Abbildungsverzeichnis	II
Abkürzungsverzeichnis	IV
1. Einleitung	1
2 Frage -und Problemstellung	2
3 Google Trends-Eine kurze Einführung	3
4 Fehlinterpretationen bei der Nutzung von Google Trends	4
4.1 <i>Nicht alle Suchanfragen werden berücksichtigt</i>	4
4.2 <i>Keine absoluten Zahlen</i>	5
4.3 <i>Eine Suchanfrage bei Google Trends ist anders definiert</i>	6
4.4 <i>Trend ist nicht gleich Trend!</i>	7
4.5 <i>Zwischenfazit: Möglichkeiten und Grenzen von Google Trends</i>	8
5 Ergänzende Programme zu Google Trends	8
5.1 <i>Google Correlate</i>	9
5.2 <i>Google Flu Trends</i>	9
5.3 <i>Zwischenfazit: Google Correlate und Google Flu Trends</i>	10
6 Weitere Modelle	11
6.1 <i>GrippeWeb</i>	11
6.2 <i>GrippeWeb: Daten und Auswertung</i>	12
6.3 <i>SORMAS</i>	15
6.4 <i>SORMAS: Daten und Auswertung</i>	15
6.5 <i>Zwischenfazit: GrippeWeb und SORMAS</i>	17
7 Forschungsfrage und Hypothesen	18
8 Methodisches Vorgehen: Online-Umfrage	19
8.1 <i>Zielgruppe</i>	19
8.2 <i>Fragenkatalog</i>	20
9 Auswertung der Online-Umfrage	21

9.1	<i>Ergebnisse und Analyse der Online-Umfrage: Demografische Daten</i>	21
9.2	<i>Ergebnisse und Analyse der Online-Umfrage: Suchverhalten im Internet</i>	24
10	Fazit und Ausblick	34
11	Literaturverzeichnis	38
12	Anhänge	40
12.1	<i>GrippeWeb Anmeldung</i>	40
12.2	<i>Fragenkatalog</i>	43
12.3	<i>SPSS-Tabellen</i>	46

Abkürzungsverzeichnis

AGI	Arbeitsgemeinschaft Influenza
ARE	Akute respiratorische Erkrankungen
BSE	Bovine spongiforme Enzephalopathie
COVID-19	Coronavirus disease 2019
CDC	Centers for Disease Control and Prevention
CSV	Comma-separated values
EHEC	Enterohämorrhagische Escherichia Coli
GT	Google Trends
GFT	Google Flu Trends
GC	Google Correlate
HZI	Helmholtz Centre for Infection Research
ILI	Influenza-like-illness
KW	Kalenderwoche
RKI	Robert Koch-Institut
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SEO	Search engine optimization
SORMAS	Surveillance, Outbreak Response Management and Analysis System

1. Einleitung

“We are not claiming that Google Trends data can help in predicting the future. Rather we are claiming that Google Trends may help in predicting the present.”¹

(Choi und Varien 2011: 1)

Dieses Zitat ist eines der ersten Sätze in den alljährlichen Papers (2009-2012) von Hyunyoung Choi und Hal Varien mit dem Titel „*Predicting the present with Google Trends*“. Varien, seit 2007 als Chefökonom bei Google Inc. tätig, hat zwei Lehrbücher veröffentlicht. Die Autoren beleuchten die Motivation und die Methodik mit der Google Trends arbeitet. Anhand diverser Beispiele werden die Ergebnisse analysiert und letztendlich mit der Aussage, dass Google Trends nicht versucht die Zukunft vorherzusagen, sondern vielmehr die Gegenwart bestimmen möchte, in Einklang gebracht.

Immer mehr Menschen nutzen das Internet, um nach den unterschiedlichsten Themen zu recherchieren. Oft werden einzelne Suchwörter eingegeben, um komplexere Themen zu beschreiben, die Google mit Suchergebnissen beantwortet. Google wird als Gesundheitsratgeber immer beliebter. Etwa 46 Prozent der befragten Deutschen geben an, regelmäßig im Internet nach Gesundheitsthemen zu recherchieren (vgl. Universität zu Köln 2020). Von Ernährungstipps bis hin zu krankheitsbedingten Symptomen wird die Suchmaschine genutzt, um sich zu informieren. Google Trends schöpft aus dieser enormen Masse an Suchanfragen, um gesellschaftliche Trends in visuell ansprechenden Diagrammen darzustellen. Aus diesem Programm erfolgte Google Flu Trends, das speziell Anfragen rund um Krankheitsthemen sammelt und auswertet. Sucheingaben zu Symptomen werden erfasst, die Aufschluss über eine anbahnende Grippewelle geben sollen. Dieses Konzept lässt sich regional sowie global einsetzen.

Diese Arbeit beschäftigt sich mit dem Prinzip von Google Flu Trends, Daten zu sammeln und Rückschlüsse auf dessen Informationsqualität zu ziehen. Zunächst wird Google Trends in seinen Funktionen beleuchtet, um zu verstehen, wie es Sucheingaben sammelt und auswertet. Anschließend werden Google Flu Trends sowie verwandte Systeme untersucht und verglichen. Im Methodenteil der Arbeit wird die Forschungsfrage sowie Hypothesen aufgestellt und mit Hilfe der Ergebnisse beantwortet. Die Forschungsfrage untersucht, ob ein hohes Medieninteresse zu einer bestimmten Krankheit dazu führen kann, dass sich das Suchverhalten von Nutzern, bezüglich Krankheitsthemen, verändert.

Die Ergebnisse sollen erste Hinweise auf das Suchverhalten von Google-Nutzern geben und wie es sich durch äußere Faktoren wie die Medienberichterstattung verändern kann.

¹ Übersetzung: Wir behaupten nicht, dass wir mit Hilfe der Daten von Google Trends die Zukunft vorhersagen können. Vielmehr behaupten wir, dass Google Trends dabei helfen kann die *Gegenwart zu bestimmen*. (Acikbas 2020)

2 Frage -und Problemstellung

In diesem Abschnitt wird die Problemstellung dieser Arbeit im kurzen Rahmen erläutert, um einen leichteren Einstieg für den Leser zu ermöglichen. Eine ausführlichere Erklärung folgt in Kapitel 7.

Google Trends ist ein Online-Werkzeug, mit dessen Hilfe das Suchinteresse von Nutzern wiedergespiegelt werden kann. Dabei lassen sich bis zu fünf Schlagwörter in einem zeitlichen Ablauf miteinander vergleichen und in einem anschaulichen Diagramm darstellen. Dadurch lässt sich die Popularität der einzelnen Begriffe nachvollziehen und eventuelle Trends in der Zukunft erkennen.

In dieser Arbeit soll die Informationsqualität von GT überprüft werden. Es wird geprüft, ob die Informationen, die GT zur Verfügung stellt, zu einem zuverlässigen und aussagekräftigen Ergebnis führen kann. Erste Versuche gab es bereits mit Google Flu Trends, mit der potenzielle Grippewellen vorhergesagt werden sollten. Mit Hilfe von Suchbegriffen der Google-Nutzer, die nach Krankheitssymptomen recherchieren, soll es GFT möglich machen Krankheitswellen Tage oder Wochen im Voraus zu erkennen, wie beispielsweise der CDC-Report aus den Vereinigten Staaten (vgl. Carneiro und Mylonakis: 2009). GFT wurde im Jahr 2015 eingestellt, da es mehrere Jahre in Folge zu falschen Prognosen zu kommenden Grippewellen traf. Die Prognosen waren teilweise mehr als doppelt so hoch angesetzt. Eine erhöhte Medienberichterstattung und häufige Präsenz von grippeähnlichen Erkrankungen in den Medien scheinen die Ursache für die inkorrekten Prognosen gewesen zu sein, da die Nutzer öfter nach Krankheitssymptomen recherchiert haben. Schlussfolgernd konnte gesagt werden, dass nicht alle Personen, die nach Symptomen im Internet gesucht haben, auch tatsächlich krank waren (vgl. Budras 2014; Cervellin et al. 2017).

Die vorliegende Arbeit soll überprüfen, ob ein verstärktes Medieninteresse zu einem verzerrten Suchverhalten bei den Probanden führen kann. Daher wurde eine Online-Umfrage entwickelt, mit denen die Teilnehmer ihr eigenes Suchverhalten in einem Krankheitsfall sowie im gesunden Zustand bewerten sollen. Die Ergebnisse sollen Aufschluss über das generelle Suchverhalten von Personen geben und wie eine Beeinflussung durch die Medien zu veränderten Rechercheaktivitäten führen kann. Die Forschungsfrage lautet daher, ob ein hohes Medieninteresse das Suchverhalten von Personen, bezüglich Krankheitsthemen, beeinflussen kann.

3 Google Trends-Eine kurze Einführung

In diesem Kapitel wird das grundlegende Gerüst von GT beleuchtet. Hierbei betrachten wir dessen Funktionen und Möglichkeiten zur Auswertung der eingegebenen Suchbegriffe. GT bietet diverse Funktionen, die seit seiner Einführung im Jahre 2004 dazu gekommen sind. So zum Beispiel *Google Flu Trends*, welches dabei helfen soll zukünftige Grippewellen zu erkennen. Dabei ist es wichtig auf vier Vorannahmen von GT einzugehen, da sie bei der Auswertung der Suchanfragen eine wichtige Rolle spielen können. Denn auf welcher Basis GT die Daten auswertet und letztendlich darstellt, ist für Laien nur schwer zu nachzuvollziehen (vgl. Alby 2016). Diese werden in einem separaten Kapitel nochmals detaillierter beleuchtet.

GT ist ein Online-Suchwerkzeug, das primär dazu dient, Suchanfragen, Themen oder Phrasen in einem bestimmten Zeitraum abzufragen. Diese Abfragen wiederum werden in Diagrammen präsentiert, die das öffentliche Interesse von Nutzern widerspiegeln und Prognosen erlauben sollen. Es lässt sich die Popularität der einzelnen Suchbegriffe in einem Zeitablauf analysieren und in Relation zu anderen Begriffen regional oder weltweit vergleichen. Maximal fünf Begriffe lassen sich gleichzeitig miteinander parallelisieren:

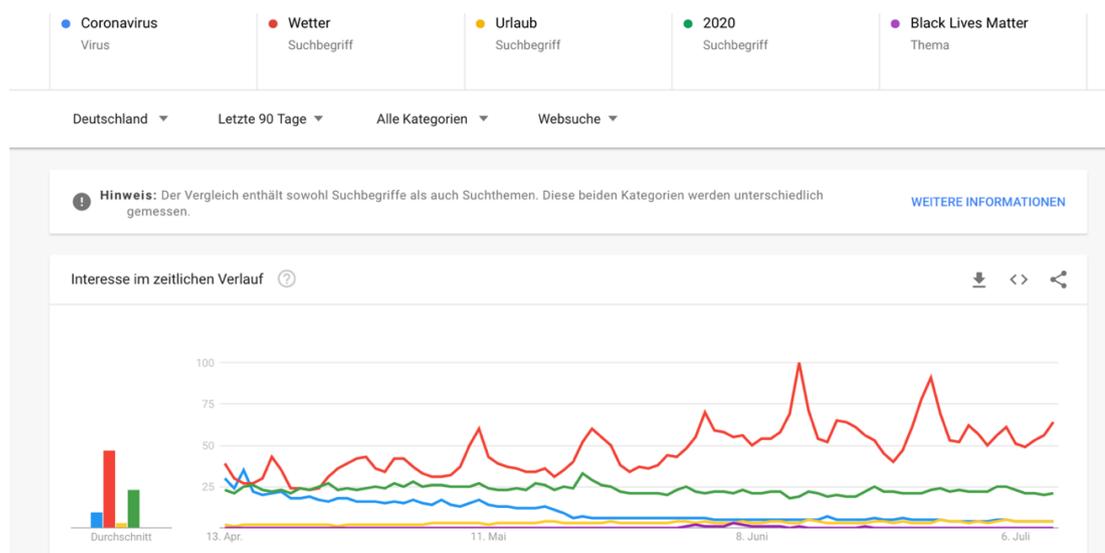


Abbildung 1: Fünf Suchbegriffe im Vergleich
(Quelle: <https://trends.google.de/trends/explore>)

Wie in der Abbildung zu sehen ist, wurden hier fünf Suchbegriffe miteinander verglichen und in einem zeitlichen Ablauf von 90 Tagen in einem Diagramm angezeigt. Der zeitliche Rahmen sowie die Region lassen sich manuell jederzeit ändern. Neben der zeitlichen Abgrenzung lassen sich die Suchbegriffe auf Kategorien eingrenzen oder speziell auf die Bildersuche, YouTube-Videos oder auf Google News einschränken. Bei einer zeitlichen Eingrenzung von drei Monaten oder kürzer werden die Daten auf täglicher Basis angezeigt. Bei einer Eingrenzung über drei Monate werden die Ergebnisse wöchentlich gestaffelt (vgl. Stephens-Davidowitz und Varien 2014: 9).

Die Ergebnisse können als CSV-Datei in Excel eingelesen und auch zur Kalkulation verwendet werden. Auch lassen sich im Bereich Top Charts täglich die 10 häufigsten Suchanfragen einer Region einsehen.

Wichtig hierbei ist die Unterscheidung zwischen den beiden Begriffen „Trends“ und „angesagten Suchanfragen“ genauer zu definieren. Google unterscheidet nämlich zwischen der absoluten Anzahl von Suchbegriffen und den tatsächlichen Trend. Der Trend muss nicht dem häufigsten Suchbegriff entsprechen und wird nur in zeitlicher Relation betrachtet. Ein Trend deutet zwar darauf hin, dass es zurzeit häufig gesucht wird, bedeutet aber nicht, dass es den höchsten Anteil an Suchanfragen ausmacht.

Beide Faktoren können auch in Korrelation zueinander in Beziehung gesetzt werden. Das bedeutet, dass die absolute Anzahl der Suchanfrage in Verbindung ihres zeitlichen Ablaufs dargestellt wird. Dadurch lässt sich die Veränderung der Popularität der Suchanfrage abbilden. Als Beispiel sei hier das Wort SEO genannt. Diese Suchanfrage wird weitaus öfter verwendet als sein ausgeschriebenes Pendant Suchmaschinenoptimierung. Man kann annehmen, dass dies auch in Zukunft so sein wird, da Nutzer es bevorzugen kurze und prägnante Begriffe als Suchanfrage zu verwenden (vgl. Rouse 2017).

4 Fehlinterpretationen bei der Nutzung von Google Trends

Bevor sich hier tiefergehend mit den Funktionen von GT auseinandergesetzt wird, sollte zuerst betrachtet werden, wie überhaupt die Daten für die Prognosen erhoben und verarbeitet werden. Dies soll vor allem dabei helfen, Missverständnisse des Online-Werkzeugs aufzuklären. Auch lassen sich erste Aussagen über die ausgeworfene Informationsqualität treffen. Sind die Daten überhaupt korrekt? Kann man sie ohne weiteres verwenden? Die folgenden Aussagen werden in den kommenden Kapiteln erklärt:

1. Nicht alle Suchanfragen werden berücksichtigt
2. Keine absoluten Zahlen
3. Eine Suchanfrage bei Google Trends ist anders definiert
4. Trend ist nicht gleich Trend!

4.1 Nicht alle Suchanfragen werden berücksichtigt

Eines der größten Missverständnisse, denen Nutzer bei der Verwendung von GT erlegen, ist die Tatsache, dass Google nicht alle Suchanfragen berücksichtigt. Es wird lediglich nur ein repräsentatives Sample angezeigt. Nicht jeder Suchbegriff hat einen ausreichend hohes Suchvolumen² und wird daher nicht berücksichtigt (vgl. Nuti et al. 2014).

² Anmerkung: Das Suchvolumen gibt die zu erwartende Anzahl an Suchanfragen zu einem bestimmten Keyword in einem bestimmten Zeitraum an. Daher sind diese Angaben nur Schätzungen und erheben keinen Anspruch auf absolute Richtigkeit und können starken Schwankungen ausgesetzt sein. Das Suchvolumen kann letztendlich dazu dienen, das Nutzerinteresse an einem bestimmten Produkt oder Dienstleistung zu ermitteln. Auf GT bezogen, eignen sich Keywordvergleiche als Methode, um grobe Tendenzen zu bestimmen. Wenn im deutschsprachigen Raum von Suchvolumina gesprochen wird, dann beziehen sich die Angaben zumeist auf Google (vgl. Wurm 2020).

Ab wann ein ausreichend hohes Suchvolumen existiert, ist nicht bekannt. Es sollte sich bewusst sein, dass die Daten die GT wiedergibt kein exaktes Bild von reellen Suchaktivitäten darstellt (vgl. Alby 2016).

4.2 Keine absoluten Zahlen

GT verwendet keine absoluten Zahlen bei der Darstellung der Diagramme. Das bedeutet, dass eine steigende Linie nicht darauf hindeutet, ob ein Suchbegriff besonders oft gesucht wird. Im Gegensatz bedeutet eine abfallende Linie nicht, dass der Suchbegriff weniger gesucht wird. Es deutet eher darauf hin, dass der prozentuale Anteil der Suche nach diesem Suchbegriff sich verändert hat. Die Daten werden lediglich in einer Skala von 0 bis 100 abgebildet (vgl. Stephens-Davidowitz und Varien 2014: 12). Als Beispiel wird hier der Vergleich zwischen zwei Suchbegriffen genommen:

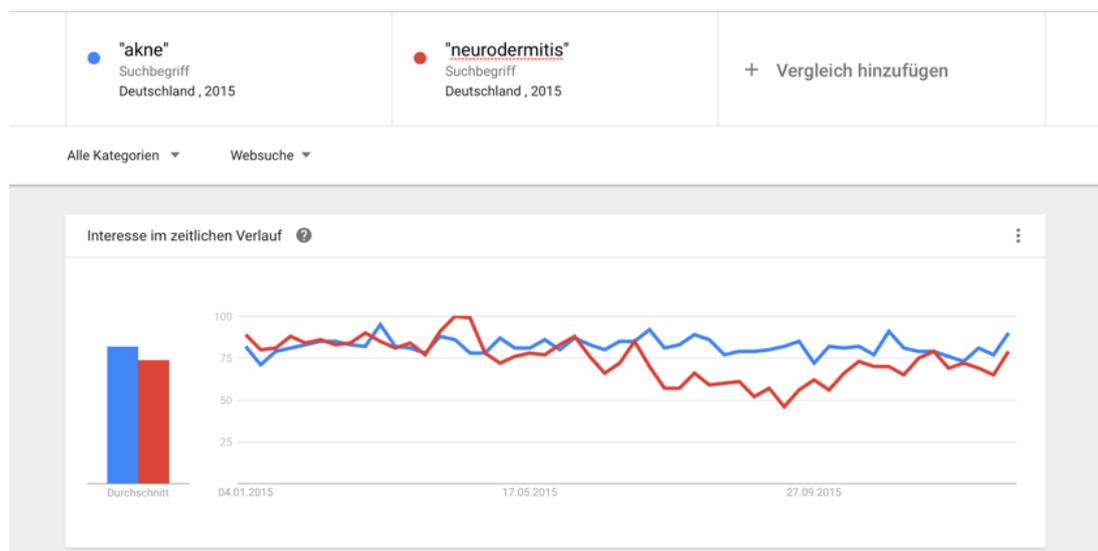


Abbildung 2: Vergleich Akne und Neurodermitis

(Quelle: <https://tom.alby.de/5-gruende-warum-du-google-trends-falsch-verstehst/>)

Beide Begriffe wurden in dem Zeitraum von 2015 betrachtet. Akne und Neurodermitis wechseln sich in ihrer Häufigkeit ab, allerdings scheint Akne zum Ende des Jahres eine höhere Popularität zu haben. Wenn der derselbe Zeitraum bei Google AdWords Keyword Planner³ betrachtet wird, fällt das erhöhte Suchvolumen von Neurodermitis besonders stark auf:

³ Anmerkung: Mit Google Ads Keyword Planner lassen sich Suchvolumina und mögliche Besucherzahlen zu bestimmten Keywords ermitteln. Die Suchvolumina werden in Zahlen ausgedrückt, die man bei Google Trends nicht einsehen kann.

Suchbegriffe	Durchschnittl. Suchanfragen pro Monat Jan. 2015 bis Dez. 2015	Wettbewerb	Vorgeschlagenes Gebot	Anteil an Anz.Im	Zu Plan hinzufügen
neurodermitis	74.000	Mittel	0,98 €		»
akne	18.100	Mittel	0,87 €		»

Abbildung 3: Suchvolumen von Akne und Neurodermitis

(Quelle: <https://tom.alby.de/5-gruende-warum-du-google-trends-falsch-verstehst/>)

Beide Begriffe haben ein unterschiedlich hohes Suchvolumen. Neurodermitis hat ein Suchvolumen von 74.000 und Akne von 18.100. Die Linien im Diagramm von GT dürften sich demnach nicht so nah sein. Deswegen ist es ratsam zusätzliche Tools zu verwenden, um kein falsches Suchinteresse abzubilden. Zwar wird Google Trend gerne für solche Statistiken verwendet, allerdings sollte man nicht vergessen, dass hier keine absoluten Zahlen als Basis genommen werden (vgl. Alby 2016).

4.3 Eine Suchanfrage bei Google Trends ist anders definiert

Wenn man ein Suchbegriff bei Google eingibt, so erscheinen üblicherweise relevante Suchergebnisse zu diesem eingegebenen Begriff. Das bedeutet, es werden nur Ergebnisse angezeigt, die dieses Suchwort beinhalten.

Bei GT ist eine Suchanfrage anders definiert. Wird ein Suchbegriff eingegeben, wird es zusätzlich in unterschiedliche Kategorien sortiert. Unter welchen Kriterien diese Einteilung erfolgt, ist unbekannt. Diesen Umstand kann umgangen werden, die Suchbegriffe in Anführungszeichen gesetzt werden. Den Unterschied macht das folgende Beispiel deutlich:

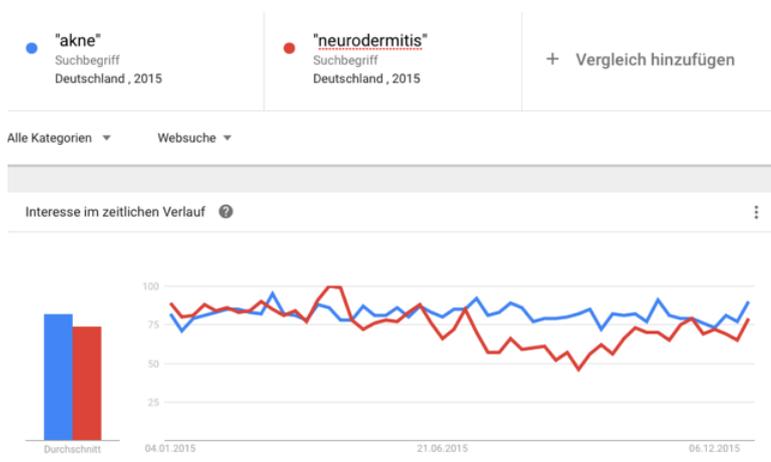


Abbildung 4: Suchbegriffe mit Anführungszeichen

(Quelle: <https://tom.alby.de/5-gruende-warum-du-google-trends-falsch-verstehst/>)

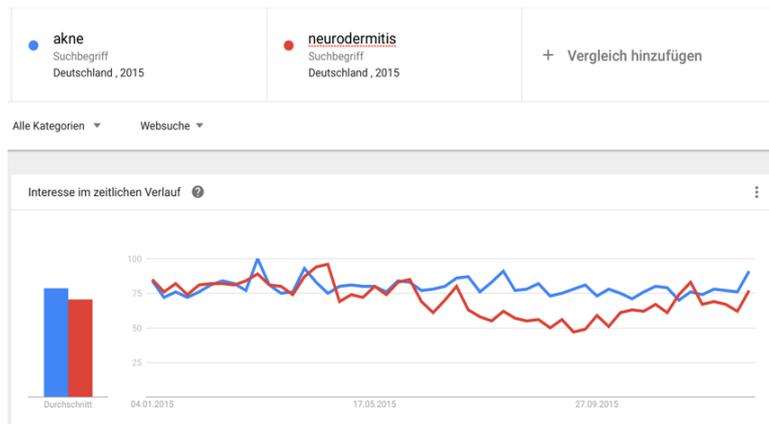


Abbildung 5: Suchbegriffe ohne Anführungszeichen

(Quelle: <https://tom.alby.de/5-gruende-warum-du-google-trends-falsch-verstehst/>)

Wie auf den Diagrammen zu sehen ist, ergibt sich ein Unterschied. Wie GT die Suchergebnisse anreichert, ist nicht bekannt. Aus diesem Diagramm ist nicht ersichtlich, welche Gruppen oder Kategorien zusätzlich zum Suchbegriff verwendet werden, um es letztendlich so abzubilden (vgl. Alby 2016).

4.4 Trend ist nicht gleich Trend!

Da GT keine absoluten Zahlen verwendet, kann es auch bei Trend Beobachtungen zu Fehlinterpretationen kommen. Wie zuvor erwähnt, verwendet GT eine Skala von 0-100, um das Suchinteresse darzustellen. Wenn ein Suchbegriff den Wert der 100 erreicht, wird von diesem Maximum aus errechnet wie hoch das Suchinteresse ist. Als Basis wird das Gesamtvolumen der Suchbegriffe verwendet, die man allerdings als Nutzer nicht kennt. Um ein verlässliches Ergebnis zu bekommen, sollte unbedingt mehrere Zeiträume betrachtet werden.

Wenn auf einem Diagramm zwei Begriffe miteinander verglichen werden, beispielsweise Trump und Wetter für das Jahr 2016 und einen Zeitraum von 12 Monaten aussucht, kann es passieren, dass der Begriff Trump eher ein geringes Interesse aufweist als das Wetter. Wird ein Zeitraum von den letzten Sieben Tagen herausgesucht, kann es dazu führen, dass das Suchinteresse für Trump nun deutlich höher ist. Ganz besonders zum amerikanischen Wahltag 2016. Diese Isolation von bestimmten Zeiträumen kann dazu verleiten, sich bestimmte Augenblicke rauszusuchen, die die eigene These unterstützen. Neben dem Abgleich von mehreren Zeiträumen, sollte ein zusätzlicher Begriff verwendet werden. Dieser Begriff fungiert als Bezugspunkt und wird im Verhältnis in der Popularität zu anderen Suchbegriffen gesehen. Dabei ist es unerheblich, ob der Begriff in einem Sinneszusammenhang mit den eigenen gesuchten Schlagwörtern steht. Darüber hinaus sollte es zu diesem Bezugspunkt, auch *Benchmark* genannt, auch genug Hintergrundwissen vorhanden sein, um eventuelle Ausreißer zu erkennen, mit der aktuellen Suchanfrage vergleichen und richtig interpretieren zu können (vgl. Alby 2016).

4.5 Zwischenfazit: Möglichkeiten und Grenzen von Google Trends

Google Trends scheint in seiner Nutzung und Auswertung wie das perfekte Online-Werkzeug, um das Suchinteresse von Nutzern ideal widerzuspiegeln. Das leicht zu bedienende Interface sowie die klaren ausgegebenen Daten vermitteln den Eindruck, dass GT ein einfaches und zuverlässiges Tool zur Bestimmung von Popularitäten ist.

Es stellt sich also die Frage, ob von einer hohen Informationsqualität gesprochen werden kann? Bedingt. Wie zuvor beschrieben, verwendet GT nicht alle Daten, die im Internet zur Verfügung stehen. Es werden keine absoluten Zahlen verwendet und auch nur ein kleiner Teil der Suchbegriffe wird repräsentiert. Der zeitliche Rahmen, in denen die Bekanntheit und Verwendung von Suchbegriffen betrachtet wird, ist ebenso wichtig wie einen konkreten Bezugspunkt zu den Ereignissen herzustellen. Eine steigende Kurve bedeutet nicht unbedingt, dass es aktuell ein erhöhtes Interesse gibt oder in Zukunft geben wird. Es funktioniert am effektivsten, wenn andere Tools zur Ergänzung genommen werden, die dabei helfen können eventuelle Diskrepanzen aufzudecken.

Allerdings ist die Fülle an Daten, die GT zur Verfügung stellt, beachtlich. Es lassen sich, bis auf die Stunde genau, Datensätze aus den letzten 16 Jahren anzeigen (bis 2004). Viele andere Tools, wie zum Beispiel das zuvor erwähnte Google Ads Keyword Planer, können nur Datensätze aus den letzten vier Jahren abrufen.

Es stellt sich nun die Frage, wofür man GT nun verwenden sollte. Alle zuvor genannten Punkte, die Google im Hintergrund ausführt, um die Daten auszuwerten, sollten gründlich bedacht werden. Zwar scheinen die Daten vollständig und mit der eigenen Wahrnehmung gleichwertig zu sein, jedoch lässt sich der Prozess der Verwertung der Daten durch GT nicht vollständig nachvollziehen.

GT hat im Laufe seiner Existenz unterschiedliche Funktionen dazubekommen. Eine der wichtigsten davon sind GFT und GC. GFT macht den Grundgedanken von GT zunutze, in dem es das Nutzerinteresse widerspiegelt und Prognosen erstellt und versucht zukünftige Grippewellen mithilfe von eingegebenen Suchbegriffen zu ermitteln. GC ist ein weiteres Tool zur Bestimmung von Korrelationen zwischen zwei oder mehreren Suchbegriffen. Beide Funktionen wurden bereits eingestellt. Dadurch lässt sich eine Aussage über deren Funktionalität und Effizienz machen.

5 Ergänzende Programme zu Google Trends

In den kommenden Abschnitten werden die zusätzlichen und ergänzenden Programme zu GT kurz vorgestellt. GT bietet viele Möglichkeiten. Für dieses Projekt wird allerdings der Fokus auf die Funktionen GFT und GC gelegt. Im Rahmen dieser Arbeit genießt GFT einen besonderen Fokus.

5.1 Google Correlate

GC wurde seit 2003 eingeführt. Etwa ein Jahr bevor GT gestartet ist. Die primäre Funktion ist die Analyse der Suchfrequenz von einzelnen Suchbegriffen in Relation zu anderen ähnlichen Suchbegriffen in einem festgelegten Zeitraum zu bestimmen. Ähnlich wie GT lassen sich hier auch zukünftige Trends in der Gesellschaft abbilden (vgl. Wagner 2016).

Allerdings muss hier eine Abgrenzung zu GT gemacht werden und wie es die Prognosen darstellt. Bei GT wird die Häufigkeit eines Suchwortes analysiert. Bei GC wird die Häufigkeit eines Terms in einem bestimmten Zeitraum in Bezug zu anderen Termen ausgewertet. Wenn man einen Begriff bei GC eingibt, werden ähnliche Begriffe im selben zeitlichen Ablauf gezeigt. Diese werden in einem sogenannten Korrelationskoeffizienten ausgedrückt. Dieser kann zwischen 1.0 und 0.0 liegen. Dabei bedeutet 1.0 eine Übereinstimmung und 0.0 bedeutet keine Übereinstimmung. Die Übereinstimmung misst die Ähnlichkeit von zwei oder mehreren Suchbegriffen.

Wenn man in der Suchleiste ein Begriff eingibt, dann erscheinen ähnliche Begriffe, die im selben Zeitraum ein ähnliches hohes Suchaufkommen vorweisen. Dabei müssen die Begriffe in keinem Zusammenhang stehen (vgl. Stephens-Davidowitz und Varien: 1).

Bedingt einer zu geringen Nachfrage, wurde GC im Dezember 2019 eingestellt.

5.2 Google Flu Trends

Google Flu Trends wurde als eigenständiges Projekt von Google Inc. im Jahre 2008 gestartet. Der Grundgedanke war, die Daten, die es durch die Sucheingaben von Nutzern in Verbindung von Krankheitssymptomen erhält, auszuwerten. Diese Datensätze sollen dazu beitragen, zukünftige Grippewellen bzw. Epidemien in Echtzeit vorherzusagen und so Krankenhäuser und Arztpraxen zu entlasten. Für Statistiken und Verläufe werden hier die Daten aus den US-amerikanischen Raum verwendet, da dort die Fülle an Daten am höchsten ist (vgl. Carneiro und Mylonakis: 2009; Nuti et al. 2014; Cervellin et al. 2017). Hierbei lassen sich auch die Möglichkeiten und vor allem die Grenzen von GFT verdeutlichen.

In seiner Anfangsphase erschien GFT als vielversprechendes und zuverlässiges Mittel, um bevorstehende Krankheitswellen aufzudecken. Im Jahr 2009 konnte somit eine Grippewelle in der USA, sieben bis zehn Tage vor den Überwachungssystemen der CDC vorhergesagt werden. Um das zu erreichen, wurden in einem Zeitraum von fünf Jahren die 50 Millionen häufigsten gesuchten Begriffe mit dem CDC-Report verglichen.

Es konnte eine hohe Korrelation zwischen den beiden Datensätzen festgestellt werden, die das Potenzial von GFT bestätigen konnte. In den Jahren 2004 bis 2008 wurde eine hohe Übereinstimmung zwischen dem Bericht des CDC und der Auswertung der Sucheingaben von Nutzern beobachtet werden. Vor allem in den Krankheiten des West-Nil- Fiebers, Respiratory-Syncytial-Virus und der Vogelgrippe (vgl. Carneiro und Mylonakis: 1157).

Die Grippesaison 2012/2013 in den Vereinigten Staaten zeigte das erste Mal die Schwächen von GFT auf. Die Prognosen, die dargestellt wurden, erwiesen sich als überschätzt. Die Grippe

hätte sich laut GFT Prognosen etwa doppelt so stark verbreiten müssen als vom CDC-Report angegeben:

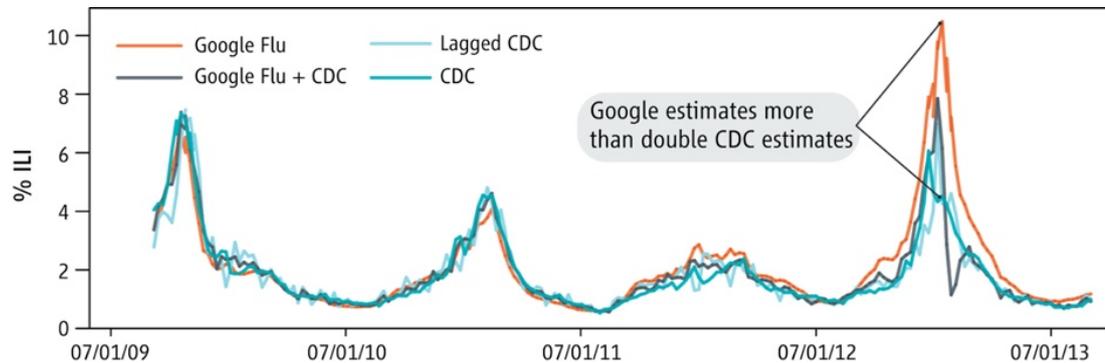


Abbildung 6: Falschaussage von Google Flu Trends 2013 in den USA
(Quelle: <https://science.sciencemag.org/content/343/6176/1203.full>)

Warum GFT auf diese Grippewelle so anders reagiert, lässt möglicherweise anhand eines neuen Features von Google erklären. Das Unternehmen führte das sogenannte „Autocomplete“ ein, das einzelne Buchstaben zu ganzen Wörtern vervollständigen konnte. Somit konnte Google den Nutzern Wörter vorschlagen, die sie bei der Suche nicht gewählt hätten. Wörter wie Fieber oder Grippe wurden womöglich öfter verwendet als vom Nutzer gewollt. Auch das allgemeine Medieninteresse kann dazu führen, dass Krankheitssymptome vermehrt im Netz gesucht werden. Wenn ein bestimmter Erreger häufiger in der Medienberichtserstattung erwähnt wird, führt dies in der Regel zu einem erhöhten Suchaufkommen der Nutzer (vgl. Budras 2014; Cervellin et al. 2017). GFT funktioniert am effektivsten, wenn es eine bestimmte Menge an Daten zur Verfügung steht. Deswegen sind Überwachungen von potenziellen Krankheitswellen in Gebieten mit einer guten Infrastruktur und Netzabdeckung besser geeignet, als Regionen ohne diese Eigenschaften.

Drei Jahre in Folge konnte GFT keine zuverlässige Vorhersage zu kommenden Grippewellen machen. Die Daten waren zum Teil doppelt so hoch angesetzt. Warum die Daten so ungenau sind, lässt sich nur schwer analysieren, da Google keinen Einblick in den Algorithmus erlaubt. GFT wurde 2015 eingestellt. Bis dahin war es in über 25 Ländern verfügbar. Auch wenn es auf nationaler Ebene keine akkuraten Ergebnisse liefern konnte, auf regionaler Basis kann es noch Vorteile bieten, vor allem um in Städten oder Bundesstaaten auf Epidemien zu reagieren (vgl. Lazer et al. 2014; Thoma 2014).

5.3 Zwischenfazit: Google Correlate und Google Flu Trends

Zusammengefasst lässt sich bestimmen, dass beide Werkzeuge dabei helfen können, auf Entwicklungen in der Gesellschaft zu reagieren. In ihren Grundzügen basieren GC und GFT auf GT, indem sie Suchbegriffe analysieren und dessen zeitliche Popularität auswerten.

GC zeigt, anders als GT, Ähnlichkeiten zwischen Suchbegriffen auf. Also wie hoch die Ähnlichkeit zwischen zwei oder mehreren Begriffen in einem ausgewählten Zeitraum ist. GFT hingegen nutzt die Popularität von Google, um bevorstehende Grippewellen aufzudecken. Die Menge an Suchanfragen über medizinische Themen lassen GFT erste Aussagen über möglicherweise künftige Epidemien treffen. Allerdings lassen sich keine Aussagen darüber

treffen, ob die Nutzer, die nach medizinisch relevanten Themen recherchieren, auch selbst erkrankt sind oder nicht. Dadurch können erhebliche Diskrepanzen zwischen der Aussage von GFT und beispielsweise dem CDC-Report entstehen (siehe Grafik oben). Wie zuvor erwähnt, können neue Funktionen wie das *Autocomplete* dazu beitragen, das Nutzerinteresse falsch wiederzugeben. Auch das verstärkte Medieninteresse für bestimmte Krankheiten, kann zu einem erhöhten Aufkommen an Suchanfragen zu einem speziellen Krankheitsbild resultieren (vgl. Budras 2014).

In den folgenden Abschnitten wird sich mit zwei weiteren Modellen zur Aufdeckung von potenziellen Krankheitswellen beschäftigt: GrippeWeb vom Robert-Koch-Institut und SORMAS.

6 Weitere Modelle

Dieser Abschnitt stellt zwei weitere verwandte Modelle zu GFT vor. Beide Systeme sind auf die Mitarbeit der Bevölkerung angewiesen, um Krankheitsherde zu erkennen und medizinisch darauf zu reagieren. Während GrippeWeb die Mitarbeit von Privatpersonen erfordert, nutzt SORMAS die Daten, die es von Praxen, Kliniken oder Krankenhäuser sammelt, um Krankheitsherde zu erkennen und frühzeitig darauf zu reagieren. Beide Modelle nutzen die Idee, durch die Partizipation von Nutzern ein übersichtlicheres Bild über die derzeitige medizinische Lage zu erhalten. In Gegensatz zu GT wird hier nicht mit einzelnen Suchbegriffen gearbeitet, sondern mit klar definierten Abfragen zum Krankheitszustand. Die Nutzer beantworten Fragen zum Wohlbefinden und Symptomen. Dadurch ergeben sich präzisere Ergebnisse, die genauere Rückschlüsse auf anbahnende Krankheitswellen zulassen.

6.1 GrippeWeb

GrippeWeb ist ein Projekt des RKI und ging 2011 als Web-Portal online. Es dient primär dazu die Aktivität von Atemwegserkrankungen zu überwachen und aufzuzeichnen. GrippeWeb ist nicht das einzige System seiner Art. Im europäischen Raum gibt es diverse Überwachungssysteme, die sich mit dem Auftreten von ARE⁴ beschäftigt (vgl. Epidemiologisches Bulletin 2017). Dabei verwendet es die Informationen direkt aus der Bevölkerung. Privatpersonen mit einer ARE können sich im Portal anmelden und Symptome beschreiben und eintragen. In Wochenberichten wird der Anteil der Gesamtbevölkerung, die an akuten ARE erkrankt sind, gemessen. Die gespeicherten Informationen werden mit den

⁴ Anmerkung: „Akute Atemwegserkrankung (ARE):

Unter akuten Atemwegserkrankungen wird eine Vielzahl von Erkrankungen der oberen und unteren Atemwege zusammengefasst, hierzu gehören vor allem Erkältungen (grippale Infekte) und die schwerer verlaufenden grippeähnliche Erkrankungen (ILI). Die meisten akuten Atemwegserkrankungen sind den Erkältungen (grippalen Infekte) zuzuordnen. Sie beginnen meist schleichend, häufig steht ein Schnupfen im Vordergrund, aber auch Fieber, Husten oder Halsschmerzen können vorkommen“ (Robert Koch-Institut 2020).

Ergebnissen der Arbeitsgemeinschaft Influenza ergänzt. Auf der Ebene der primärärztlichen Versorgung, überwacht die AGI die Aktivität von akuten ARE. Seit 2014 gibt es über 8000 registrierte Teilnehmer (vgl. Robert Koch Institut 2020).

Auf der Startseite von GrippeWeb wird für jede Kalenderwoche eine Zusammenfassung der eingegangenen Daten präsentiert. Es werden die Daten aus der Vorwoche mit der aktuellen Woche verglichen und steigende oder fallende Raten von ARE beschrieben:

Die für die Bevölkerung in Deutschland geschätzten Raten der 29. KW beruhen auf den Angaben von 6.183 GrippeWeb-Teilnehmern und -Teilnehmerinnen, von diesen meldeten 217 eine ARE, 37 meldeten eine ILI (Datenstand: Di, 21.07.2020). Durch Nachmeldungen können sich aber noch Änderungen ergeben. Zum Beispiel beruhte der Bericht der Vorwoche (28. KW 2020) auf den Meldungen von 6.079 Teilnehmern. Durch Nachmeldungen liegen inzwischen 7.598 Meldungen für die 28. KW 2020 vor. (Robert Koch-Institut: 2020)

Anschließend werden die Ergebnisse mit den Daten der AGI verglichen und ergänzt:

Die Arbeitsgemeinschaft Influenza (AGI), die Meldungen von Haus- und Kinderarztpraxen auswertet, berichtet, dass in der 29. KW 2020 im Vergleich zur Vorwoche insgesamt etwas weniger Arztbesuche wegen ARE registriert wurden. (Robert Koch-Institut: 2020)

6.2 GrippeWeb: Daten und Auswertung

Ab einem Alter von 14 Jahren dürfen sich Personen auf diesem System registrieren. Die Nutzer werden gebeten Angaben zum Alter, Wohnort und eventuelle Vorerkrankungen zu machen. Für die Vorerkrankungen wird ein Aufnahmebogen bereitgestellt, der ausgefüllt werden muss (die Bilder zu den einzelnen Anmeldeschritten werden in den Anhängen mit Screenshots nochmals erläutert). Zudem ist es erlaubt, dass Eltern Angaben über ihre Kinder machen können, die noch unter 14 Jahren sind.

Ab dem Zeitpunkt der Anmeldung erhalten alle Teilnehmer jeden Montagmorgen eine Mail mit der Bitte, anzugeben, ob eine ARE bei ihnen aufgetreten sind oder nicht. Die Ergebnisse werden zeitnah veröffentlicht und für die Nutzer in einem passwortgeschützten Bereich aufbereitet. Dort können sie sich, ähnlich eines Tagebucheintrags, ihre eigenen Verläufe rückwirkend seit der Anmeldung betrachten (vgl. Epidemiologisches Bulletin 2017).

Um verlässliche Ergebnisse zu bekommen, ist eine regelmäßige Teilnahme der registrierten Nutzer notwendig (vgl. Friesema: 6353). Hierfür wurde bei GrippeWeb ein Gewinnspiel eingeführt. Für jede Frage, die man in der wöchentlichen Email beantwortet, gibt es Punkte, die die Chancen für die Preise steigert. Die Preise können verschiedene Laptops und Espressomaschinen sein. Eine mögliche Verzerrung der Ergebnisse durch das Gewinnspiel lassen sich ausschließen. In allen Altersgruppen lässt sich kein signifikanter Unterschied feststellen zwischen den Gewinnspielteilnehmern und denen, die nicht am Gewinnspiel teilgenommen haben (vgl. Bayer et al. 2014: 4).

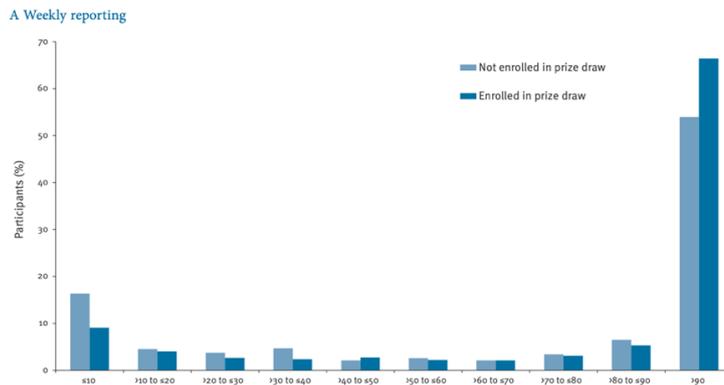


Abbildung 7: Gewinnspielteilnehmer vs. nicht Gewinnspielteilnehmer
 (Quelle: Internet-based syndromic monitoring of acute respiratory illness in the general population of Germany, weeks 35/2011 to 34/2012)

Die Gewinne werden einmal im Quartal ausgelost. Ab September jeden Jahres wird jede Woche zusätzlich abgefragt, ob der Teilnehmer sich gegen Influenza impfen lassen hat. Der Befragte kann angeben, im welchen Monat die Impfung erfolgte. Diese Frage wird bis März des kommenden Jahres wöchentlich gestellt, bis der Befragte dies bejaht (vgl. Epidemiologisches Bulletin 2017).

Die erstmalige Auswertung der Ergebnisse erfolgte in der 22 Kalenderwoche. GrippeWeb ging im März 2011 online. Um verlässliche Aussagen zu treffen, musste eine ausreichend hohe Teilnehmerzahl vorhanden sein. Zum Zeitpunkt der Auswertung gab es bis zu 800 wöchentliche Meldungen. Die hier betrachteten Werte beziehen sich auf einen Zeitraum von 2011 bis 2017:

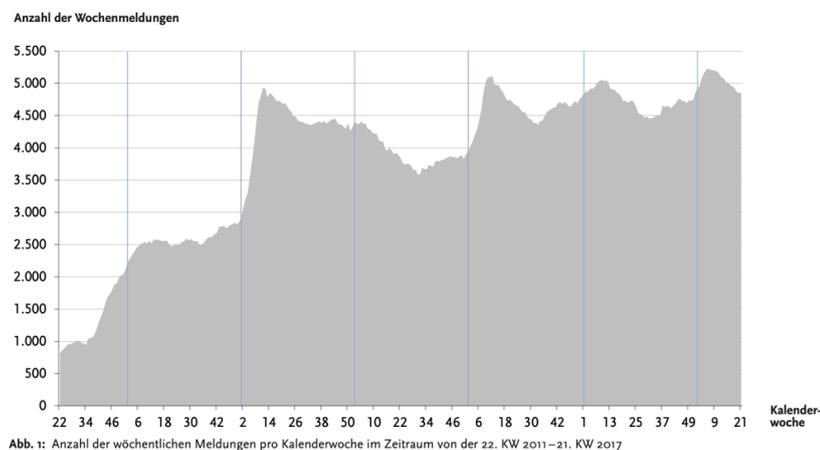


Abbildung 8: Wöchentliche Meldungen zwischen 2011-2017
 (Quelle: Epidemiologisches Bulletin 2017)

Die obige Grafik zeigt die wöchentlichen Meldungen, die bei GrippeWeb seit 2011 bis 2017 eingegangen sind. Dabei werden die Kalenderwochen in 12 Wochen Schritten dargestellt. Zwischen 2011 und 2017 gingen 1,2 Millionen Meldungen bei GrippeWeb ein, bei einer Teilnehmerzahl von 12.286. Die anfänglichen wöchentlichen Meldungen in der KW 22 2011 betragen etwa 800, in der 21 KW 2017 betragen sie 4.800 Meldungen. Zwischen diesen beiden Zeitpunkten haben 53 Prozent aller Teilnehmer in jeder Woche die Fragen zum Krankheitszustand beantwortet. Das Durchschnittsalter der Befragten beläuft sich auf 42 Jahre mit einer Altersspannweite von 0- 92 Jahren. Etwa 44 Prozent der Befragten waren männlich, 56 Prozent waren weiblich (vgl. Epidemiologisches Bulletin 2017).

Seit 2011 hat GrippeWeb kontinuierlich an wöchentlichen Meldungen und Teilnehmern dazugewonnen. Neben der erhöhten Anzahl der Teilnehmer spielt auch die Zuverlässigkeit der Befragten eine wichtige Rolle. Wie bereits beschrieben, haben in dem Zeitraum von 2011 bis 2017 53 Prozent der Befragten das System genutzt und auch die wöchentlichen Fragen beantwortet. Grundsätzlich werden die Altersgruppen in der Bevölkerung gut repräsentiert. Allerdings bilden Personen ab 60 Jahren sowie junge Erwachsene bis 25 Jahren eine Ausnahme. Beide Altersgruppen sind unterrepräsentiert, was daran liegen könnte, dass Personen ab 60 Jahren noch einen geringeren Bezug zum Internet haben (vgl. Epidemiologisches Bulletin 2017). Jugendliche bis junge Erwachsene müssen sich ab 14 Jahren selbst anmelden. Die Altersgruppe von unter 14 Jahren ist jedoch sehr gut vertreten, da die Eltern die wöchentlichen Angaben übernehmen.

GrippeWeb kann zukünftig dabei helfen, Pandemien zu erkennen und so Impfstoffe rechtzeitig zur Verfügung zu stellen. Auch wäre eine technische Weiterentwicklung für die Bekämpfung von Influenzawellen überaus vorteilhaft. Eine App könnte zeitnah die eingegebenen Daten der Nutzer auswerten und so das Email Verfahren ersetzen (vgl. Epidemiologisches Bulletin 2017).

Die abschließende Grafik zeigt die zusammengefassten Meldungen in einem Diagramm für die KW 30. Für jede Kalenderwoche werden die Ergebnisse angezeigt und mit vorherigen Kalenderwochen verglichen. Darüber hinaus wird die Grippewellensaison hervorgehoben, um sie mit den Meldungen der Teilnehmer zu betrachten:

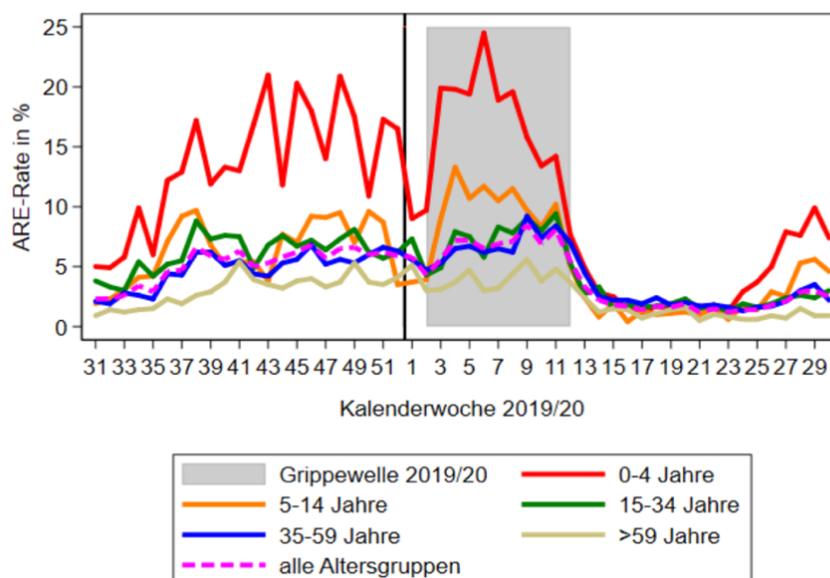


Abbildung 9: Ergebnisse der Woche 30/2020
(Quelle: grippeweb.rki.de)

Die Grippewelle wird mit einem grauen Feld hervorgehoben und die Spannweite der Grippesaison lässt sich aus den Kalenderwochen ablesen, hier von der ersten Kalenderwoche bis zur elften. Die unterschiedlichen Altersgruppen werden ebenfalls in unterschiedlichen Farben dargestellt. Die ARE-Rate in Prozent gibt den Anstieg von akuten Atemwegserkrankungen in der Bevölkerung an. Demnach ist die Zunahme von ARE bei der Altersgruppe null bis vier Jahren um etwa 25 Prozent gestiegen.

6.3 SORMAS

Das *Surveillance Outbreak Response Management and Analysis System*⁵ oder kurz SORMAS, wurde 2014 am Helmholtz-Zentrum für Infektionsforschung in Zusammenarbeit mit westafrikanischen Gesundheitsämtern entwickelt. Als Reaktion auf den Ebola-Ausbruch 2014 in Westafrika wurde eine App entwickelt, die als Frühwarnsystem für zukünftige Ebola Ausbrüche fungieren soll. Die App nutzt die Cloud Technologie, um auch in Regionen mit schwacher Internet Infrastruktur zuverlässig mit Daten versorgt zu werden. (vgl. Helmholtz-Zentrum für Infektionsforschung 2020).

Hierbei geht SORMAS einen technischen Schritt weiter als vom RKI entwickelten GrippeWeb. Anders als GrippeWeb wird SORMAS nicht von Privatpersonen verwendet. Vielmehr vernetzt es alle relevanten Schnittstellen bei Seuchenausbrüchen wie zum Beispiel Praxen, Kliniken oder Krankenhäuser. Ziel ist es, Infizierte schnell zu identifizieren und Gesundheitseinrichtungen zu informieren (vgl. vitagroup AG 2020).

6.4 SORMAS: Daten und Auswertung

Privatpersonen haben in der Regel keinen Zugriff auf SORMAS. Die App funktioniert speziell zwischen allen relevanten Gesundheitsstellen und Personen. Das bedeutet, dass eine Ärztin oder Arzt die App auf ihr mobiles Gerät herunterladen und die Symptome des Betroffenen eingeben kann, die dann an Krankenhäuser, Praxen, Behörden und Kliniken zum Abruf bereitstehen. Dies kann sogar offline geschehen. Wenn zurzeit keine Internetverbindung vorhanden ist, werden bei nächster Gelegenheit mit einer Internetverbindung die Daten automatisch in die Datenbank übermittelt (vgl. vitagroup AG 2020). SORMAS kann auch auf stationären Computern verwendet werden, wie auch auf mobilen Geräten. Die untenstehende Abbildung zeigt die Benutzeroberfläche der mobilen Version:



Abbildung 10: SORMAS Benutzeroberfläche

(Quelle: https://www.eurosurveillance.org/content/10.2807/1560-7917.ES2015.20.12.21071#html_fulltext)

Das Interface ist benutzerorientiert und leicht zu navigieren. Sie Symptome werden einzeln abgefragt und übersichtlich dargestellt (Hier: *Muscle Pain, Headache, Pharyngitis, Vomiting, Diarrhoea*). Die Daten werden anschließend in die Datenbank eingespeist und stehen künftig

⁵ Anmerkung: Digitales Melde -und Managementsystem

den behandelnden Akteuren zur Verfügung. Der Patient kann anschließend im nächsten Krankenhaus, welches auch SORMAS nutzt, weiter behandelt werden (vgl. Fähnrich et al. 2015).

Die Nutzung von SORMAS lässt sich in drei Schritten beschreiben. Der erste Schritt ist die mögliche Erkennung von Krankheitswellen und das Sammeln von Daten. Die Daten gelangen über verschiedene Kanäle in das System. Diese können das Bürgertelefon (*Citizen Hotline*), Krankenhäuser, Mediziner oder generelle Verdachtsfälle sein, die von Informanten geäußert werden können.

Der zweite Schritt ist die Überprüfung und Bestätigung der gesammelten Daten. Überwachungs-Officer (*Surveillance officers*) und Überwachungs-Supervisoren (*Surveillance supervisor*) analysieren die Daten und können weitere Personen mit der Kontrolle beauftragen. Der dritte Schritt bezeichnet die Eindämmung der Krankheit. Der Infizierte wird identifiziert und behandelt. Falls weitere Maßnahmen zur Eindämmung vonnöten sind, werden sie eingeleitet (vgl. vitagroup AG 2020). Abschließend eine Abbildung, die den Workflow von SORMAS visualisiert:

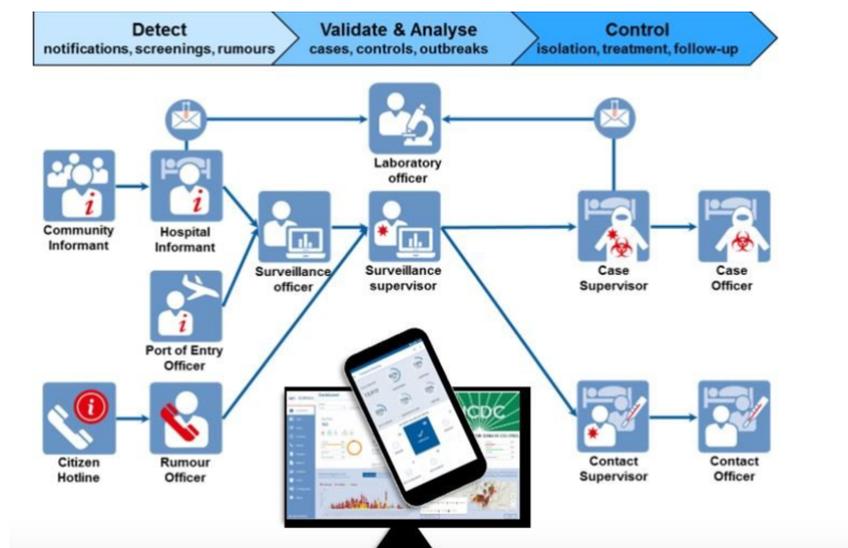


Abbildung 11: Workflow SORMAS

(Quelle: https://www.vitagroup.ag/de_DE/News-Insights/Sormas-Epidemie-Management-Programm)

Hier werden, neben den zuvor genannten zuständigen Officers, auch weitere Personen genannt, die eine wichtige Funktion im SORMAS-System innehaben. Nicht nur durch das Bürgertelefon gelangen Informationen an die Überwachungs-Officers, sondern auch durch sogenannte Krankenhaus Informanten (*Hospital Informant*). Die Daten der Informanten werden analysiert und an die Sachverständigen weitergeleitet, die die Eindämmung der Krankheit vorbereiten (*Case Supervisor, Case Officer, Contact Supervisor, Contact Officer*). Die Daten, die durch die Informanten und durch die Sachverständigen vor Ort gesammelt wurden, werden an das zuständige Labor übermittelt (*Laboratory Officer*). Dort werden die Daten und die Proben auf eine etwaige Infektion überprüft.

6.5 Zwischenfazit: GrippeWeb und SORMAS

Im abschließenden Fazit werden die beiden Seuchenbekämpfungstools nochmals betrachtet und die jeweiligen Unterschiede genauer benannt.

GrippeWeb und SORMAS haben beide das Ziel, Krankheiten in bestimmten Regionen aufzudecken und die Eindämmung eben solcher zu ermöglichen. Durch die Mithilfe von Bürgerinnen und Bürger, Ärztinnen und Ärzten, Praxen, Kliniken und Krankenhäuser, soll ein engmaschiges Netz aus allen relevanten Beteiligten geschaffen werden, die dabei helfen können, Infizierte schnell und unkompliziert zu identifizieren. Dabei ist das Prinzip bei beiden Systemen ähnlich: Symptome werden lokalisiert, benannt und in ein eigens dafür eingerichtetes Online-Portal eingetragen.

Der signifikanteste Unterschied besteht hier, wer die Daten sammelt und ins System eingibt. Während bei GrippeWeb jede Privatperson sich im Web-Portal anmelden kann und Auskunft über seine derzeitige Gesundheitslage im Bereich der ARE treffen kann, können bei SORMAS nur behandelnde Ärztinnen und Ärzte bzw. zuständige Gesundheitsmitarbeiter das Programm nutzen. Auch ist in der Identifizierung der Krankheiten ein Unterschied. GrippeWeb beobachtet die Anzahl der akuten Atemwegserkrankungen in der Bevölkerung. SORMAS wurde, in seiner ursprünglichen Intention, als Maßnahme gegen künftige Ebola Ausbrüche entwickelt.

Hier zeigen sich schon erste Stärken von SORMAS gegenüber GrippeWeb. Während GrippeWeb den derzeitigen Krankheitszustand nur an stationären Computern abfragen kann, bietet SORMAS ein benutzerorientiertes Interface, welches als App auf jedes Smartphone verwendet werden kann. Darüber hinaus ist SORMAS ein adaptierbares System. Es wurde zwar als Frühwarnsystem für Ebola entwickelt, lässt sich jedoch auch auf andere Krankheitsbilder anwenden. 2017 konnte erfolgreich in weiten Teilen Nigerias Krankheitsherde der Affenpocken aufgedeckt werden. Durch die Umstellung auf die Symptomatik von Affenpocken, konnte SORMAS in wenigen Tagen die betroffenen Gebiete lokalisieren. SORMAS konnte durch ihr adaptives Design mittlerweile 37 Infektionskrankheiten erkennen (vgl. vitagroup AG 2020). Jüngst wurde auch die COVID-19 Pandemie in das System integriert (vgl. Krause et al. 2020). Zukünftig soll auch eine App für Privatpersonen erfolgen, die mit bestimmten Krankheiten in Kontakt kamen. Diese Entwicklung erfolgte als Reaktion auf die zuvor genannte Pandemie. Allerdings könnte ein datenrechtliches Problem zwischen den Patientendaten und Mobilfunkdaten entstehen, dass eine Entwicklung einer App entgegensteuern könnte. Aktuell arbeiten 39 Gesundheitsämter mit SORMAS (vgl. Norddeutscher Rundfunk 2020).

Die hier vorgestellten Seuchenbekämpfungstools wurden als Frühwarnsysteme entwickelt, um Krankheitswellen zuverlässig und präzise zu bestimmen. Ein ähnliches Prinzip verfolgt Google Trends und speziell mit Google Flu Trends. Bei Google wird die Masse an Sucheingaben verwendet, um eine möglichst präzise Vorhersage von bestimmten gesellschaftlichen oder Krankheitsausbrüchen zu bestimmen. Eine erhöhte Ballung von Suchbegriffen zu einem speziellen Symptom, kann Rückschlüsse auf eine künftige Grippewelle erlauben.

In den nächsten Kapiteln werden die Forschungsfragen sowie die angewandte Methodik dieser Bachelor-Arbeit vorgestellt und erläutert.

7 Forschungsfrage und Hypothesen

Der Titel dieser Arbeit soll überprüfen, ob und wie GT durch die Auswertung der Sucheingaben eine Datenbasis schaffen kann, mit der künftige Ereignisse, wie zum Beispiel Krankheitswellen, vorhergesehen werden können.

In den Abschnitten 3 und 4 wurde auf die Funktionen und Möglichkeiten von GT eingegangen. GT zieht seine Basis für dessen Prognosen aus den kumulierten Eingaben aller Suchbegriffe und stellt sie als Diagramme dar. Steigende und fallende Kurven, sollen das erhöhte beziehungsweise verringerte Suchinteresse von Google Nutzern repräsentieren. Zusätzliche Systeme nutzen die Datenlage von Google und GT als Basis, um bestimmte Ereignisse besser vorherzusehen und darauf zu reagieren. GFT war eines dieser neuen Systeme, das 2008 erstmals online ging. Es sollte Suchbegriffe im Bereich der Gesundheit und Krankheit analysieren und etwaige regionale Ballungen erkennen, um Vorhersagen zu möglichen Ausbrüchen von Grippewellen treffen. Obwohl GFT in seinen ersten Jahren zuverlässige Aussagen zu Grippewellen treffen konnte (vgl. Carneiro und Mylonakis: 1157), wurde der Dienst 2015 eingestellt, da wiederholt falsche Aussagen zu Krankheitswellen getroffen wurden (vgl. Lazer et al. 2014; Thoma 2014).

Im Abschnitt 6 wurden zwei weitere Systeme vorgestellt, die eine Frühwarnfunktion innehaben. GrippeWeb fragt auf wöchentlicher Basis die Teilnehmer nach akuten Atemwegserkrankungen ab und SORMAS dient als Überwachung von Ebola Ausbrüchen. Während GrippeWeb die Partizipation von Privatpersonen erfordert, funktioniert SORMAS als Schnittstelle für Gesundheitsbehörden, Ärztinnen und Ärzte, Praxen und Kliniken. Krankheitsherde sollen schnell und zuverlässig identifiziert werden, um die bestmögliche Behandlung zu ermöglichen.

Wie bereits im Abschnitt 5.2 erwähnt, kann auch die Medienberichtserstattung zu einem hohen Suchaufkommen von Krankheitssymptomen führen (vgl. Budras 2014; Cervellin et al. 2017). Es soll untersucht werden, ob Personen Google nutzen, die selbst an einer Krankheit leiden oder ob sie durch das verstärkte Medieninteresse, wie aktuell im Falle COVID-19-Pandemie, nach Symptomen die Suchmaschine zu Rate ziehen. Aus dieser Forschungsfrage lassen sich wie folgt zwei Hypothesen aufstellen:

- **H1:** Je höher das Medieninteresse zu einer Erkrankung ist, desto häufiger wird im Internet nach Krankheitssymptomen recherchiert.
- **H2:** Je geringer das Medieninteresse zu einer Erkrankung ist, desto weniger wird im Internet nach Krankheitssymptomen recherchiert.

Die Hypothesen sollen primär die Tatsache überprüfen, ob eine Google-Suche zu einer Krankheit aus eigenmotivierten Zwecken vollzogen wird, weil man selber krank ist, oder, ob es äußerlichen Einflüssen unterworfen ist, wie zum Beispiel durch ein verstärktes

Medieninteresse. Durch das Abfragen des Suchverhaltens soll eine klare Trennung zwischen innerer Motivation und äußerlichen Einflüssen möglich gemacht werden. Im folgenden Abschnitt wird das methodische Vorgehen zur Belegung oder Wiederlegung der Hypothesen vorgestellt.

8 Methodisches Vorgehen: Online-Umfrage

Durch die Online-Umfrage soll das Suchverhalten von Personen untersucht werden, die das Internet nutzen, speziell durch Google, um sich über Erkrankungen und Krankheitssymptomen zu informieren. Es soll überprüft werden, ob eine hohe Medienpräsenz zu einer bestimmten Krankheit auch zu erhöhten Suchaktivitäten bei den Nutzern zu krankheitsbedingten Themen führen kann. Dabei sollten alle Probanden Google als primäre Suchmaschine nutzen, da die grundsätzliche Idee, Schlagwörter zu Krankheitsthemen zu analysieren aus GT und GFT ableiten lässt. Ein weiterer wichtiger Aspekt sind die demografischen Daten, mit der die Probanden in verschiedenen Kategorien betrachtet und untersucht werden können.

Die Umfrage wurde ausschließlich online durchgeführt und auf unterschiedliche Kanäle verbreitet, um an die notwendigen Probanden heranzutreten. Die dafür genutzten Kanäle waren Social Media Kanäle wie Facebook, Instagram und WhatsApp. Facebook-Seiten wie *Umfragen für Studienarbeiten* haben sich dafür geeignet zusätzliche Probanden zu finden. Auf die angesteuerte Zielgruppe wird im Abschnitt 9.1 präziser eingegangen. Die Umfrage sollte für alle Probanden eigenständig und einfach durchführbar sein. Damit die Teilnehmer nicht vorzeitig das Interesse verlieren, wurde die Dauer des Fragebogens auf 3-5 Minuten ausgelegt sein und maximal aus 13 Fragen bestehen. Die Fragen waren kurz und prägnant gestellt, um Probleme bei der Beantwortung und anschließender Auswertung vorzubeugen. Die Probanden bewerten in mehreren Schritten ihr eigenes Suchverhalten im Internet mit aussagekräftigen vorgegeben Antworten. Eine Frage beinhaltet die Eingabe von Text, die allerdings auch freiwillig ist, damit die Teilnehmer nicht vorzeitig abbrechen. Näheres wird im Abschnitt 9.2 beschrieben. Die Fragen sollen dabei helfen das Suchverhalten besser zu differenzieren in Hinblick auf die Recherche von Krankheitsthemen.

8.1 Zielgruppe

Die Online-Umfrage richtet sich an alle Personen, die Google als primäre Suchmaschine nutzen, um sich über Krankheitssymptome und verschiedenen Erkrankungen zu informieren. Daher ist die Auswahl der potenziellen Teilnehmer sehr hoch, da Google eine Monopolstellung im Bereich der genutzten Suchmaschinen in Deutschland innehat. Dabei wird hier keine Eingrenzung beim Alter vorgenommen, da die Nutzung von Google auf alle Altersgruppen anwendbar ist. Daher sollte die Durchmischung der Probanden gleichmäßig sein, wobei die Anzahl an Studenten voraussichtlich am höchsten sein wird, weil die Akquise hauptsächlich durch den Universitätsverteiler erfolgt ist.

Um sicherzustellen, auch die richtige Zielgruppe angesprochen zu haben, werden Fragen zur Suchmaschinen-Nutzung gestellt. Die Teilnehmer sollen angeben, welche Suchmaschine primär zur Recherche von Krankheitsthemen genutzt wird. Hierbei durften die Teilnehmer mehrere Suchmaschinen angeben. Google wurde zu 98,5 Prozent genannt, darunter noch Yahoo, Bing und DuckDuckGo.

Die Zielgruppe wird auf unterschiedliche Kanäle angesprochen. Wie zuvor erwähnt, wurden potenzielle Probanden auf diversen Facebook-Seiten angesprochen, Instagram, mit der Bitte den folgenden Link zur Umfrage zu verwenden und persönliche Kontakte gebeten daran teilzunehmen. Dadurch konnte eine weite Streuung der Teilnehmer erreicht werden. Durch die Maßnahmen konnten 137 Probanden für die Online-Umfrage generiert werden.

8.2 Fragenkatalog

Der Fragebogen beinhaltet 13 Fragen, wobei vier davon demografische Daten abfragen. Die restlichen acht Fragen beschäftigen sich mit dem Suchverhalten der Probanden.

Die Online-Umfrage wurde zunächst mit einem kurzen Text eingeleitet, um den Sinn und Zweck der Befragung zu erläutern. Es wurde auf eine längere Erklärung verzichtet, da zu lange Texte das Interesse der Probanden verringern konnte. Der Teilnehmer soll kurz und knapp informiert werden, ohne auf wichtige Informationen verzichten zu müssen. Da es hier um das persönliche Suchverhalten von individuellen Nutzern geht, wird den Probanden versichert, dass alle erhobenen Daten anonym sind und zudem keine Rückschlüsse auf einzelne Personen zugelassen werden können. Um zusätzlich die Teilnahmebereitschaft zu erhöhen, wurde die voraussichtliche Dauer der Umfrage auf 3-5 Minuten geschätzt.

Der Fragenkatalog war optisch sowie inhaltlich klar gestaltet, damit die Probanden nicht abgelenkt oder verwirrt werden. Die Fragen waren verständlich formuliert, damit die Probanden die Umfrage eigenständig und ohne Unterstützung ausfüllen können. Da das Thema eine breite Altersspanne anspricht, waren die Fragen auch für Schüler ab 14 Jahren leicht verständlich. Fragen, die sich um das Suchverhalten rund um Krankheitsthemen beschäftigen, wurden als Pflichtfragen festgelegt, damit die zuvor aufgestellten Hypothesen auch belegt oder widerlegt werden können. Der Fragebogen enthält eine offene Frage, die bestimmte Internetseiten, TV-Sendungen, Zeitungen oder Zeitschriften zur Recherche von Gesundheitsthemen abfragt. Allerdings ist diese nicht zwingend zu beantworten und kann übersprungen werden. Dadurch soll die Teilnahmebereitschaft sowie Motivation nicht beeinflusst werden. Zudem sind Fragen zum Berufsstand und Aufenthaltsort keine Pflichtangaben. Das soll das Vertrauen der Probanden in die Anonymität gestärkt werden.

Der Fragebogen wird mit einem kurzen Einleitungstext vorgestellt. Danach folgt die Abfrage nach dem Alter und Geschlecht. Die anschließenden drei Fragen sollen ermitteln, welche Suchmaschinen die Probanden nutzen, ob sie schon von bestimmten Krankheiten in den Medien gehört haben und wann sie zuletzt krank waren. Bei den nächsten vier Fragen sollen die Nutzer ihr eigenes Suchverhalten bewerten und angeben, ob sie im Krankheitsfall oder wenn sie gesund sind die Suchmaschine nutzen, um sich über Erkrankungen zu informieren. Hier wird auch abgefragt, ob die Probanden in den letzten drei Monaten verstärkt nach

Krankheitssymptomen geforscht haben. Diese Frage ist essenziell zur Untersuchung der aufgestellten Hypothesen. Abschließend werden zusätzliche demografische Daten zum Berufsstand und Wohnort abgefragt. Nach dem Beenden der Umfrage folgt eine erweiterte Erklärung zur Umfrage und beleuchtet die Bedeutung der gesammelten Daten und wie sie in der Forschung helfen können.

9 Auswertung der Online-Umfrage

Nach dem Abschluss der Online-Umfrage müssen die Ergebnisse ausgewertet werden. Da die Zielgruppe sehr weit gefasst werden kann, muss keine Überprüfung der Probanden zur Eignung dieser Umfrage vorgenommen werden. Die Evaluierung der Ergebnisse erfolgt hauptsächlich über die einfache Häufigkeitsverteilung und werden in Diagrammen erläutert. Insgesamt haben 141 Probanden an der Umfrage teilgenommen. Wobei letztlich 123 Teilnehmer die Umfrage auch tatsächlich beendeten. Das bedeutet, es gibt ein Defizit an 18 Personen, die die Teilnahme vorzeitig abbrachen. Davon haben zwölf Befragte die Teilnahme schon bei der dritten Frage abgebrochen, sechs Probanden haben nach der siebten Frage aufgehört. Daher nimmt die Anzahl der Teilnehmer zu Beginn bis zum Ende der Umfrage ab. Die Teilnehmeranzahl ist bei der jeweiligen Frage nochmals angegeben.

9.1 Ergebnisse und Analyse der Online-Umfrage: Demografische Daten

Nach dem Einleitungstext folgen die Fragen zum Alter und Geschlecht. Diese Fragen sind schnell zu beantworten und sollen einen schnellen und unkomplizierten Einstieg in die Umfrage ermöglichen. Alle Angaben zu demografischen Daten sind freiwillig und können übersprungen werden. Damit soll der Eindruck genommen werden, dass die Umfrage nur dem Zweck dient persönliche Daten zu sammeln. Da die Fragen zu demografischen Daten aufgeteilt sind, zwei Fragen zu Beginn und zwei Fragen zum Ende, werden die Ergebnisse im Zuge der Übersicht hier zuerst analysiert. Das bedeutet, dass die Fragen eins, zwei und 13 hier vorab behandelt werden. Frage zwölf wird gesondert analysiert, da das Suchverhalten nach Krankheitssymptomen regional betrachtet wird. Die Anzahl der Teilnehmer hat sich bis zum Ende des Fragebogens verringert. Es begann mit 137 Probanden und 123 Befragte haben die Umfrage beendet:

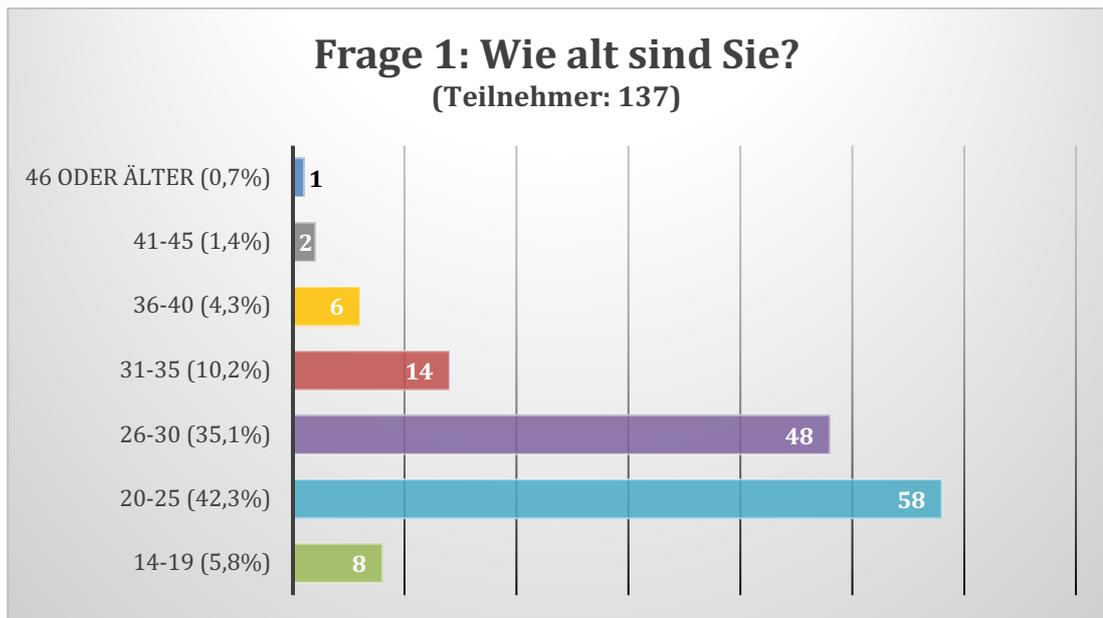


Abbildung 12: Altersangabe (Online-Umfrage 2020)

Die Altersspannweite aller Befragten beläuft sich auf 40 Jahren. Somit sind alle Altersgruppen zwischen 14 und 54 Jahren vertreten. Dabei macht die Altersgruppe von 20 bis 30 Jahren 77,4 Prozent der Befragten aus. Der hohe Anteil lässt sich durch die Teilnahme der Studierenden erklären, die in der Frage 13 nochmals genauer beschrieben werden. Der Altersdurchschnitt beläuft sich demnach auf 26,4 Jahren. Erwartungsgemäß sind die Altersgruppen ab 31 Jahren am geringsten vertreten. Dennoch ist das Ergebnis breit gefächert und erlaubt auch differenziertere Aussagen zum Suchverhalten. Daher soll in den kommenden Fragen auch geprüft werden, ob das Alter einen beeinflussenden Faktor auf das Suchverhalten im Internet ausübt. Da ein Hauptteil der Teilnehmer unter 30 Jahren ist, scheint es sinnvoll ab dieser Altersgrenze das Suchverhalten bei älteren Personen zu analysieren. Hier lässt sich eventuell ein Zusammenhang zwischen dem Alter und der generellen Nutzung von Google feststellen. Beispielsweise lässt sich bei Frage vier leichte Tendenzen im Altersdurchschnitt beobachten. Näheres dazu im Abschnitt zu Frage 4.

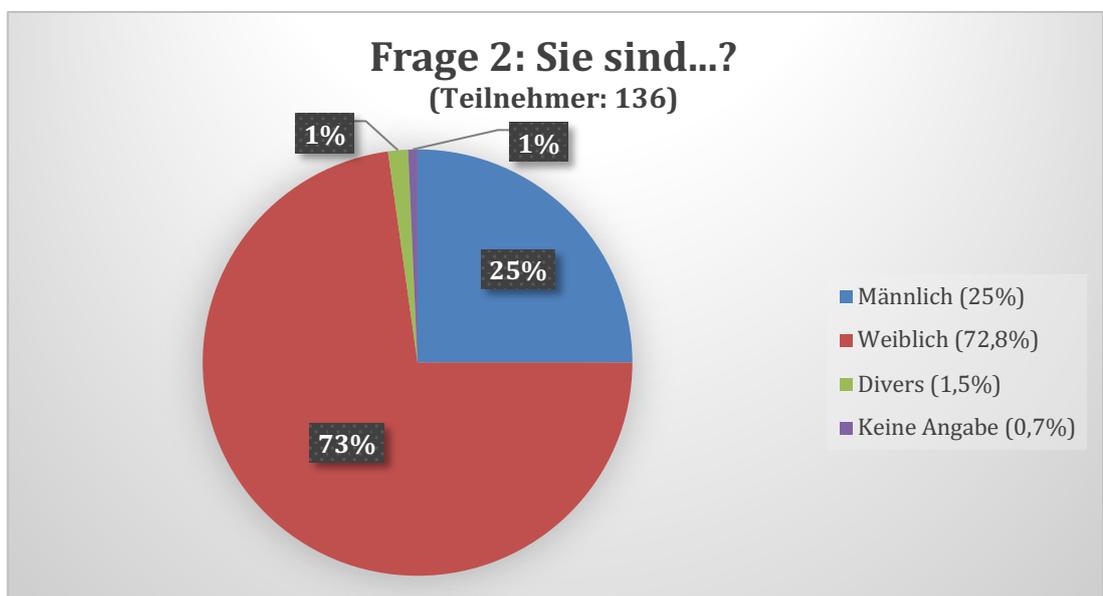


Abbildung 13: Geschlechterverteilung (Online-Umfrage 2020)

Bei der zweiten Frage konnten die Probanden ihr Geschlecht angeben. Hier fällt besonders die unausgewogene Geschlechterverteilung auf. Etwa 72,8 Prozent der Befragten sind weiblich, 25 Prozent sind männlich. Eine kleine Gruppe hat mit 1,5 Prozent divers angegeben und 0,7 Prozent hat keine Angabe gemacht. Daher kann auch hier nicht von einem repräsentativen Ergebnis gesprochen werden, da die Aufteilung der Geschlechter nicht mit der Aufteilung der deutschen Bevölkerung übereinstimmt, die 50,7 Prozent weiblich und 49,3 Prozent männlich sind (vgl. Bundeszentrale für politische Bildung 2018).

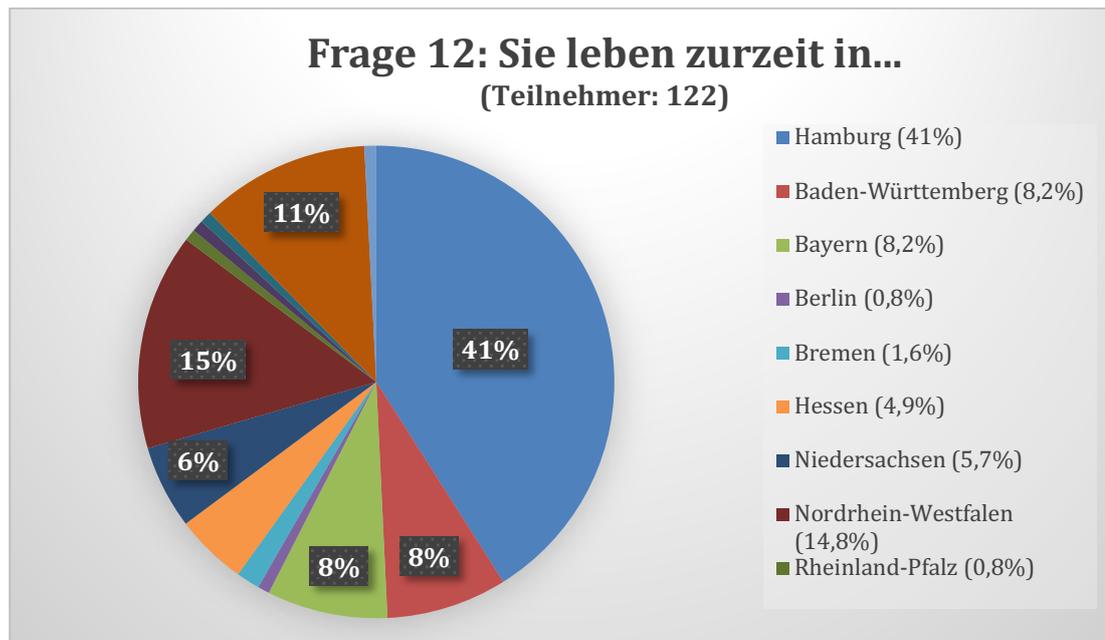


Abbildung 14: Angaben zum Wohnort (Online-Umfrage 2020)

Die Abfrage zum Wohnort erfolgt als vorletzte Frage. Aus 13 Bundesländern haben Probanden an der Online-Umfrage teilgenommen. Auffällig ist der Anteil der Befragten aus Hamburg. Da die Umfrage an Familie, Freunde, Kommilitonen und Arbeitskollegen versendet wurde, ist der hohe Anteil zu erklären. Die Abfrage soll aufzeigen, wie sich die Probanden in Deutschland verteilen. Dies ist insofern wichtig, da GFT in der Auswertung regionale Daten verwendet, um potenzielle Krankheitsherde zu lokalisieren. Natürlich sind die Daten dieser Online-Umfrage keineswegs repräsentativ, aufgrund der zu kleinen Datenmenge. Daher lassen sich mit dieser kleinen Menge an Daten keine zuverlässigen Tendenzen erkennen, ob eine vermehrte Suche nach Krankheitssymptomen in einem bestimmten Bundesland auf eine kommende Krankheitswelle hindeutet.

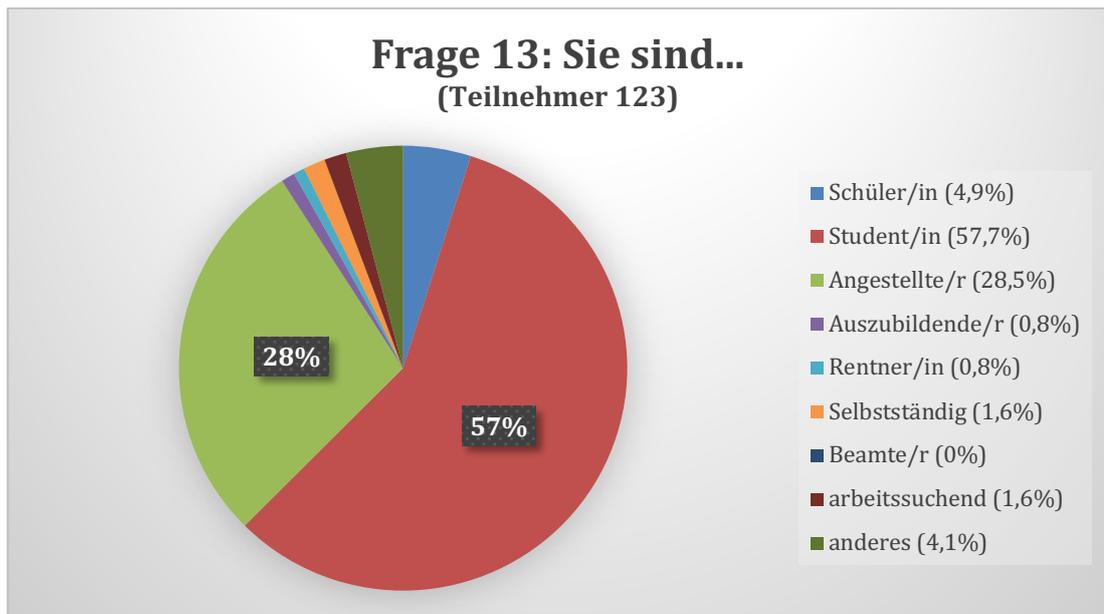


Abbildung 15: Frage zum Berufsstand (Online-Umfrage 2020)

Bei der Frage zum Berufsstand sind fast alle Berufsgruppen durch die Probanden abgebildet. Keiner der Befragten die Option Beamte/r gewählt. Allerdings überwiegt der Anteil der Studenten. Daher kann hier nicht von einem repräsentativen Ergebnis für alle Berufsgruppen gesprochen werden. Der zweithöchste Anteil an Teilnehmern verteilt sich auf die Angestellten, die 28,5 Prozent der Menge ausmachen. Auch finden sich mit 4,9 Prozent Schüler, mit jeweils 0,8 Prozent Rentner und Auszubildende sowie jeweils mit 1,6 Prozent Selbstständige und Arbeitssuchende. Die restlichen Antworten verteilen sich auf die Option anderes mit 4,1 Prozent. Die Verteilung ist wenig überraschend, da die Verbreitung der Online-Umfrage hauptsächlich über Universitätsinterne E-Mail-Verteiler und Kommilitonen erfolgte. Die Berufsgruppe der Angestellten wurde durch Arbeitskollegen akquiriert.

9.2 Ergebnisse und Analyse der Online-Umfrage: Suchverhalten im Internet

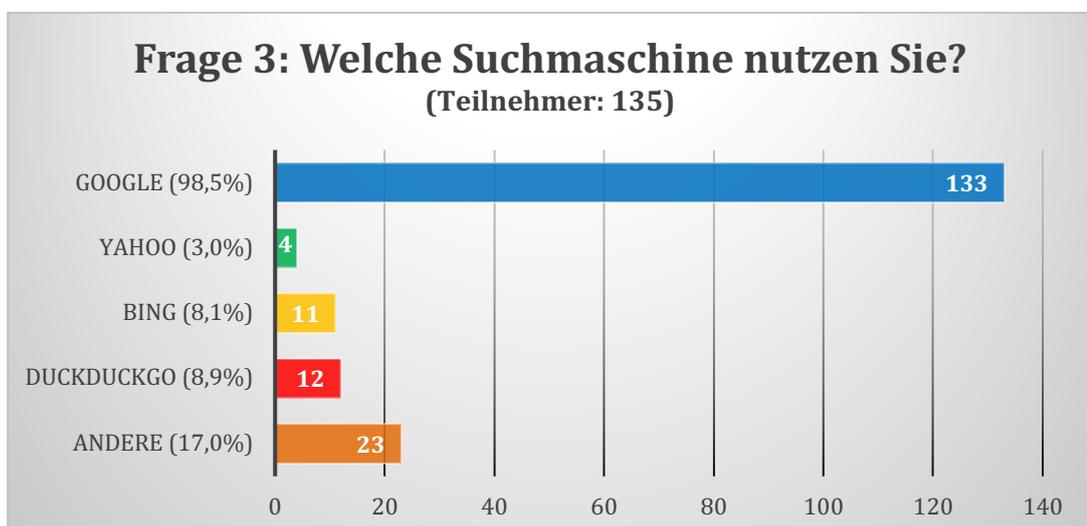


Abbildung 16: Welche Suchmaschinen werden genutzt (Online-Umfrage 2020)

Wie zu sehen, ergibt sich eine deutliche Präferenz, welche Suchmaschine am meisten genutzt wird. Etwa 98,5 Prozent haben auf die Frage, welche Suchmaschine sie nutzen, Google gewählt. Am wenigsten, etwa 3 Prozent, haben Yahoo angegeben. Bing und DuckDuckGo wurden mit 8,1 Prozent und 8,9 Prozent beantwortet. 17 Prozent der Befragten haben die Option Andere ausgewählt. Das lässt darauf schließen, dass die Teilnehmer entweder keines der oben genannten Suchmaschinen nutzen oder zusätzlich zu den genannten noch Weitere verwenden. Zudem deckt sich der Anteil an Google-Nutzern mit offiziellen Statistiken zur Nutzung von Suchmaschinen. 2018 betrug der Suchmaschinenmarktanteil von Google in Deutschland 95,5 Prozent, gefolgt von Bing mit 2,3 Prozent, Yahoo mit 0,9 Prozent und DuckDuckGo mit 0,5 Prozent (vgl. Wollenberg 2019). Daher kann man das Diagramm als repräsentativ betrachten. Zählt man die Fälle zusammen ergibt sich ein prozentualer Wert von 135,6. Das bedeutet, dass durchschnittlich 1,3 Suchmaschinen genannt wurden. Die Möglichkeit mehrere Antworten zu nennen wurde demnach nicht ausgenutzt. Daraus lässt sich eindeutig schließen, dass ein Großteil der Probanden nur eine Suchmaschine nutzen.

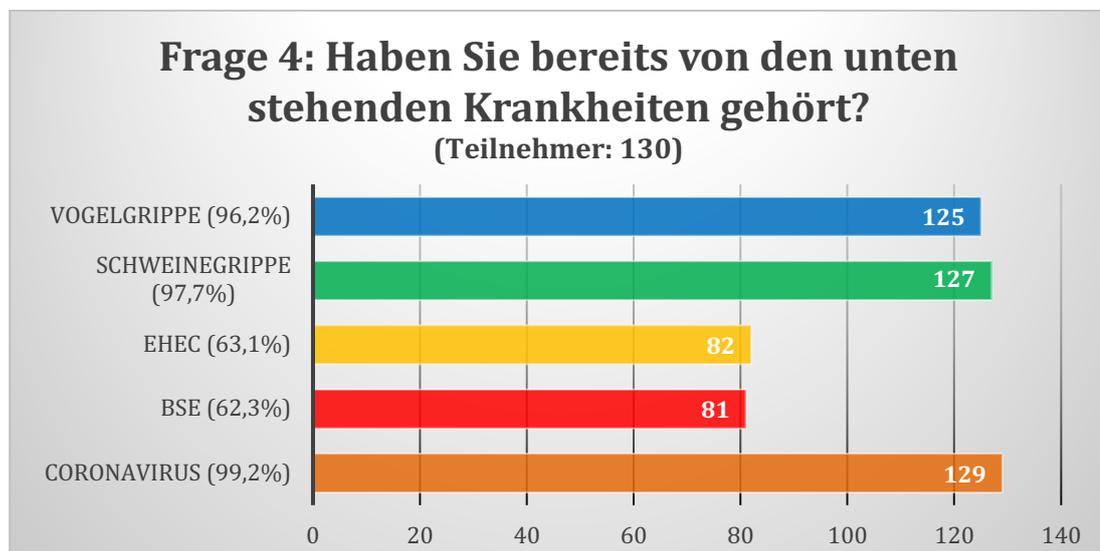


Abbildung 17: Von welchen Krankheiten haben die Probanden schon gehört (Online-Umfrage 2020)

Die zweite Frage beschäftigt sich mit den unterschiedlichen Krankheiten, die zwischen 2000 und 2019 ein erhöhtes Medieninteresse aufwiesen. Hier soll eine einfache Abfrage das Wissen der Probanden überprüfen. Da diese Krankheiten öfter in den Medien repräsentiert wurden, soll getestet werden, ob sie grundsätzlich die aufgezählten Erkrankungen kennen. Erst in der anschließenden Frage wird geklärt aus welchen Kanälen sie von den Krankheiten erfahren haben. Auch hier durften die Probanden mehrere Antworten auswählen.

Das Diagramm zeigt deutliche Tendenzen. Das Coronavirus wurde zu 99,2 Prozent von nahezu allen Teilnehmern genannt. Die Schweinegrippe kommt auf 97,7 Prozent. Die Vogelgrippe wurde zu 96,2 Prozent ausgewählt. EHEC und BSE fallen deutlich ab. Beide Krankheiten wurden zu 63,1 Prozent und 62,3 Prozent genannt. Dies lässt sich wahrscheinlich durch die zeitliche Distanz der Ausbrüche erklären. Die ersten BSE-Fälle wurden im Jahr 2000 bestätigt. EHEC Ausbrüche folgten 2011. Wenn man den Altersdurchschnitt von 26,4 Jahren der Teilnehmer betrachtet, kann daraus abgeleitet werden, dass die meisten der Befragten noch zu jung waren, um von diesen Krankheiten gehört zu haben. Der Altersdurchschnitt bei den Personen, die EHEC und BSE gewählt haben beträgt 27,3 und 28,9 Jahre. Das ist etwas höher

als der Altersdurchschnitt aller Befragten, dennoch lässt sich ein Zusammenhang zwischen dem Alter und der Zeit, in denen die Krankheiten erstmals ausbrachen, feststellen. Bei einer höheren Teilnehmerzahl ließe sich ein verlässlicheres Ergebnis abbilden. Berichterstattungen zu der Schweinegrippe und der Vogelgrippe gibt es kontinuierlich seit 2009 beziehungsweise 2014. Dass fast alle Probanden das Coronavirus angegeben haben ist hier nachvollziehbar. Die kumulierten prozentualen Anteile ergeben 418,5. Das bedeutet, dass im Durchschnitt 4,1 Antworten gegeben wurden. Die Möglichkeit mehrere Antworten zu nennen wurde demnach sehr häufig genutzt.

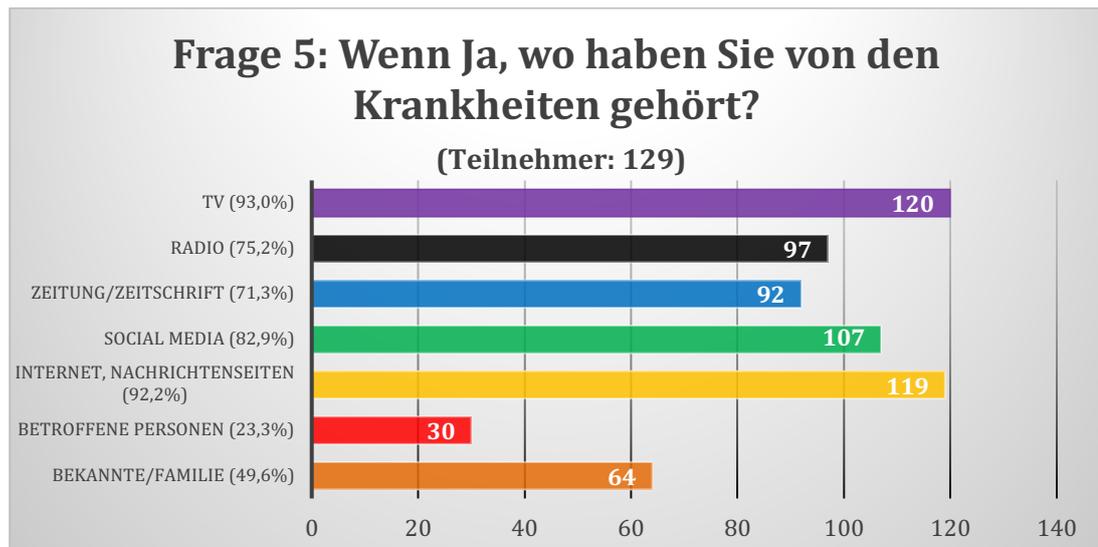


Abbildung 18: Von wo haben die Probanden von den Krankheiten erfahren (Online-Umfrage 2020)

Die dritte Frage ergänzt die vorherige Abfrage zu den Erkrankungen. Hier soll geprüft werden, aus welchen Kanälen die Informationen zu den bereits genannten Erkrankungen stammen. Dabei durften die Befragten mehrere Antworten geben, da die Beschränkung auf ein Medium nicht sinnvoll wäre. Diese Frage dient primär dazu das breite Spektrum des Medieninteresses abzubilden und welche Kanäle bevorzugt genutzt wurden.

Neben den klassischen Medien, wie TV, Radio, Zeitung, Zeitschrift, Internet und Social Media, wurden auch andere Informationskanäle als Antwortmöglichkeiten gegeben. Die Probanden unter den zuvor genannten Antworten, auch die Möglichkeit betroffene Personen und Bekannte und Familie auszuwählen. Diese zwei Angaben sollen als Kontrollantworten dienen, um herauszufinden, woher hauptsächlich das Wissen über die Erkrankungen folgen. Wie in der Abbildung zu sehen, zeigt das Medium TV mit 93 Prozent den höchsten Anteil der Antworten aus. Gefolgt von Internet mit 92,2 Prozent. Informationswege über Social Media wurden zu 82,9 Prozent angegeben. Hier sollte ein klarer Unterschied zwischen Internet und Social Media definiert werden. Social Media beinhaltet Inhalte Facebook, Instagram, Twitter oder ähnliche soziale Medien. Internet bezeichnet bestimmte Nachrichtenseiten wie ZEIT Online, Welt, Focus, BILD und mehr. Radio wurde zu 75,2 Prozent der Teilnehmer genannt und Zeitung und Zeitschriften zu 71,3 Prozent. Die letzten zwei Antwortmöglichkeiten fallen im Vergleich stark ab. Betroffene Personen machen einen Anteil von 23,3 Prozent aus, während Bekannte und Familie mit 94,6 Prozent genannt wurden. Die Auswertung dieser Frage zeigt deutlich, dass Informationen zu Erkrankungen durch den TV und online Nachrichtenseiten eingeholt werden. Informationen durch Betroffene Personen oder Bekannte

und Familie werden deutlich weniger in Anspruch genommen. Die kumulierten prozentualen Anteile ergeben 487,6. Das bedeutet, dass jeder Teilnehmer durchschnittlich 4,8 Informationsquellen genannt haben.



Abbildung 19: Wann waren die Probanden zuletzt krank (Online-Umfrage 2020)

Diese Frage dient als Vorbereitung zur nächsten Frage und soll die Probanden dabei helfen sich an eine Situation, in denen sie krank waren, zu erinnern. Die kommenden Fragen untersuchen das Suchverhalten der Teilnehmer in Krankheitssituationen.

Die Probanden konnten in ein freies Textfeld das Jahr eingeben, in denen Sie zuletzt krank waren. Den größten Anteil macht das Jahr 2019. Dort haben 43 Prozent aller Teilnehmer angegeben, dass sie zuletzt krank waren. Im Jahr 2020 haben 39 Prozent der Befragten angegeben zuletzt krank gewesen zu sein. 2018 waren es noch 15 Prozent, 2017, 2016 und 2010 waren es jeweils nur 1 Prozent der Teilnehmer. Der Zweck dieser Frage ist es, herauszufinden wann die Probanden zuletzt krank waren, damit sie ihr Suchverhalten im Internet noch bewerten können. In diesem Fall reicht eine simple Angabe des Jahres aus, um die Antworten auszuwerten. Ein genaues Datum der Erkrankung zu nennen, würde eher dazu führen, dass die Probanden den Fragebogen abbrechen, da nicht jeder Teilnehmer den exakten Zeitraum einer Grippe, Erkältung oder Fieber benennen kann.

Frage 7: Wenn ich krank bin nutze ich die Suchmaschine Google, um nach Symptomen und Heilungsmethoden zu recherchieren.
(Teilnehmer: 128)

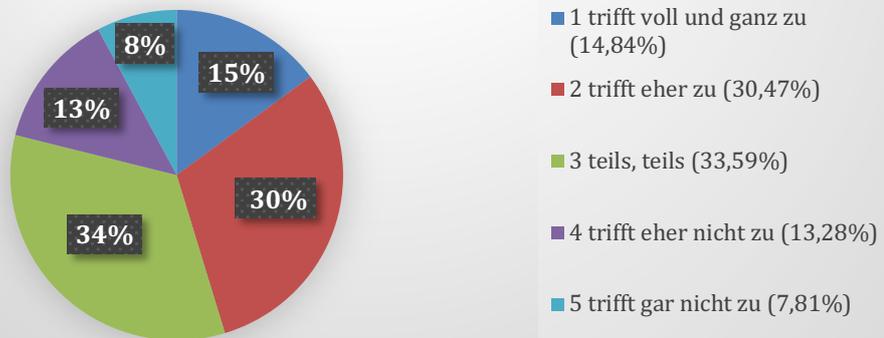


Abbildung 20: Suchmaschinen-Nutzung bei Krankheit (Online-Umfrage 2020)

Bei dieser Frage sollen die Probanden ihr eigenes Suchverhalten während einer Erkrankung bewerten. Dabei konnten sie zwischen fünf Antworten auswählen. Wichtig für die Ergebnisse ist die Tatsache, dass die Probanden ihr Suchverhalten bewerten müssen, als sie zuletzt krank waren. Deswegen war die vorherige Frage bedeutend, da sich die Teilnehmer an eine Situation zurückerinnern und ihr Verhalten rekapitulieren sollen. Bei 14,8 Prozent der Befragten „trifft es voll und ganz zu“, dass sie in einem Krankheitsfall Google zur Hilfe nehmen. Etwa 30,4 Prozent haben „trifft eher zu“ angegeben. 33,5 Prozent haben mit „teils, teils“ geantwortet. Mit 13,2 Prozent wurde auf „trifft eher nicht zu“ geantwortet. Am wenigsten wurde die Antwort „trifft gar nicht zu“, mit 7,8 Prozent, gewählt. Auf den ersten Blick ist sofort zu erkennen, wie sich die Antworten verteilen. Der größte Teil der Angaben verteilt sich auf die Antworten „teils, teils“ und „trifft eher zu“, mit 64 Prozent. Die restlichen Ergebnisse machen 36 Prozent der gesamten Antworten aus. Der Durchschnitt der Antworten beläuft sich auf 2,69. Also bewegen sich die Antworten der Befragten zwischen „trifft eher zu“ und „teils, teils“. Die Standardabweichung beträgt 1,12 rund um den Mittelwert. Daher kann eine Aussage über die Ähnlichkeiten der gegebenen Antworten der Teilnehmer machen. Weit über die Hälfte der Probanden haben übereinstimmende Angaben zum Suchverhalten angegeben. Die darauffolgende Frage wird nur von den Probanden vorgeschlagen, die zwischen „teils, teils“ und „trifft gar nicht zu“ entschieden haben. Infolgedessen beläuft sich der Anteil der Befragten auf nur 70 Personen.

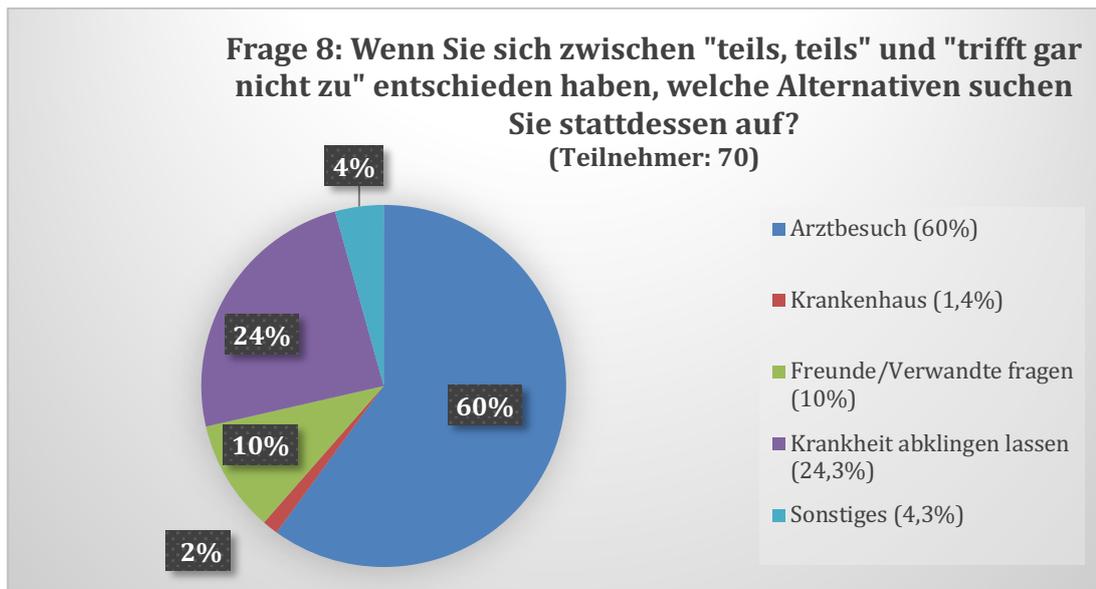


Abbildung 21: Alternativen zur Nutzung von Google (Online-Umfrage 2020)

Da zuvor die Personen betrachtet wurden, die im Falle einer Erkrankung die Suchmaschine Google nutzen, um sich über diverse Symptome und Heilungsmethoden zu informieren, werden hier die Teilnehmer betrachtet, die nicht auf eine Internetrecherche zurückgreifen. Die Spannweite der Antworten, die ausgewählt wurden, um auf diese Frage weiterzuleiten, umfasst alle Angaben zwischen den Möglichkeiten „teils, teils“ und „trifft gar nicht zu“. Die Anzahl der Personen, die auf diese Frage weitergeleitet wurden, ist geringer als die Gesamtanzahl der Probanden. Die Antwortmöglichkeit „teils, teils“ wurde bewusst in dieses Spektrum aufgenommen, da die Intention Google zu nutzen auch in beide Richtungen interpretiert werden kann. Personen können Google verwenden bei einem Krankheitsfall oder es nicht verwenden.

Die deutliche Mehrheit von 60 Prozent würde einen Arztbesuch als Alternative in Erwägung ziehen, wenn sie keine Suchmaschine zur Recherche gebrauchen. Weitere 24 Prozent der Befragten würden die Krankheit einfach ausklingen lassen und keine Hilfe in Anspruch nehmen. In weiteren Studien müsste natürlich betrachtet werden, ob die Schwere der Krankheit eine Auswirkung auf das Verhalten hat. Etwa 10 Prozent der Teilnehmer würden Freunde oder Verwandte fragen. 1,4 Prozent würden ein Krankenhaus aufsuchen. Die restlichen 4,3 Prozent wurden als Sonstige angegeben. Da es unzählige Antwortmöglichkeiten als Alternativen aufgezählt werden konnten, wurde die Möglichkeit, keine fest definierte Antwort angeben zu müssen, eingebaut. Da nur ein kleiner Teil diese Antwortmöglichkeit genutzt hat, kann hier von einer passenden Auswahl an Antwortmöglichkeiten gesprochen werden.

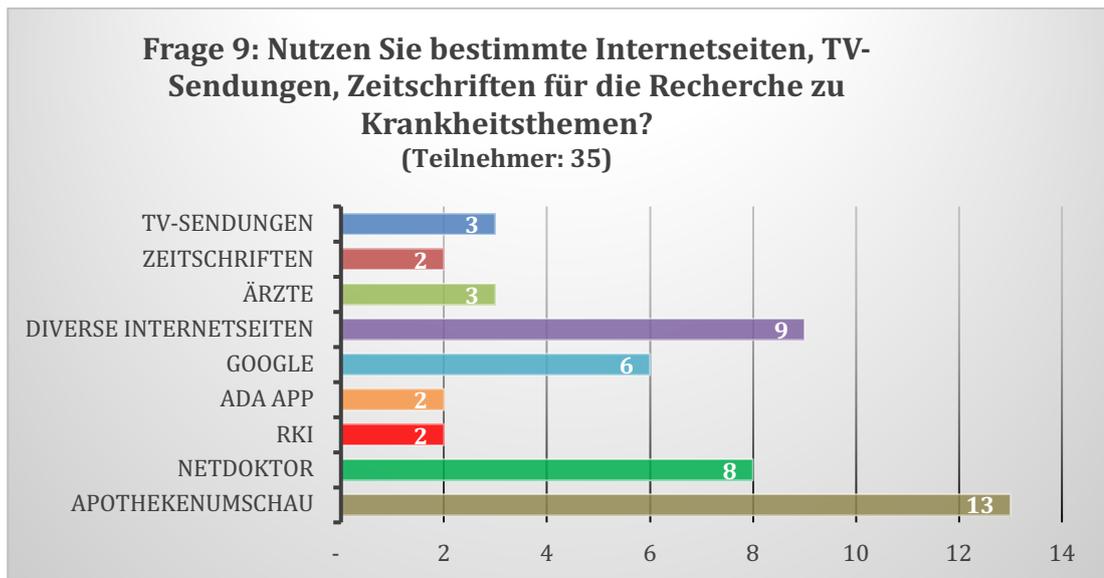


Abbildung 22: Informationsquellen (Online-Umfrage 2020)

Die folgende Frage beschäftigt sich mit den Informationsquellen, die die Probanden nutzen, um sich über verschiedene Krankheitsthemen zu informieren. Hier sollen die Befragten angeben, ob sie auf bestimmte Internetseiten, TV-Sendungen oder Zeitschriften zur Recherche zurückgreifen. Die Teilnehmer können in ein freies Textfeld ihre Antworten eintragen. Die Frage wurde bewusst nicht als Pflichtfrage integriert und mit einem entsprechenden Hinweis markiert. Damit sollte einem vorzeitigen Abbruch des Fragebogens entgegenwirken. Insgesamt haben 35 Teilnehmer die Frage beantwortet. Es gab 48 Antworten, da jeder Proband mehrere Quellen angeben durfte. In der obigen Grafik werden die Häufigkeiten der verschiedenen Antworten dargestellt. Im kommenden Text werden die Antworten nochmals differenzierter beschrieben.

Der Großteil der Antworten beziehen sich auf Internetquellen. Am häufigsten wurde die Internetseite der Apothekenumschau genannt.⁶ Netdoktor⁷ folgt mit acht Nennungen. Google wurde als Quelle ebenfalls angegeben. Diese Antworten beziehen sich auf die Suchergebnisse, die die Suchmaschine auswirft. Es wurden hier keine spezifischen Internetseiten im Zusammenhang genannt. Das RKI wurde als Informationsquelle von zwei Personen angegeben. Eine spezielle Gesundheits-App, die Ada App⁸, wurde ebenfalls von zwei Befragten erwähnt.

Um die Übersicht zu wahren, wurden in der obigen Grafik die Antworten zu Diverse Internetseiten, Ärzte, Zeitschriften und TV-Sendungen zusammengefasst, da es zu jedem Punkt unterschiedliche Antworten existieren. Diese werden im Folgenden nochmals im Detail aufgezählt:

⁶ Anmerkung: Die Internetseite ermöglicht die Suche nach Erkrankungen und Symptomen. Ist die Suche erfolgreich, wird das Krankheitsbild und die dazugehörigen Symptome beschrieben.

⁷ Anmerkung: Netdoktor bietet Fachwissen über Volkskrankheiten. Artikel werden verständlich von Experten für Laien verfasst.

⁸ Anmerkung: Es ist möglich, Symptome in die App einzugeben. Die Daten werden mit Tausenden anderen Personen verglichen, um die mögliche Ursache der Symptome zu ermitteln.

Diverse Internetseiten

- Wikipedia
- Gesundheitsamt
- Krankenkassenseiten
- Ärzteblatt
- Gesundheit.de
- Diverse medizinische Internetseiten
- Zentrum der Gesundheit
- Qmul.ac.uk⁹
- Internetseiten für natürliche Heilmittel

Ärzte

- Arzt
- HNO-Ärzte
- Frauenärzte

Zeitschriften

- BILD der Frau
- Psychologie Heute

TV-Sendungen

- RTL
- SAT 1
- Tagesschau

Die obigen Antworten wurden in Stichwörtern von den Probanden angegeben. Auf die Frage hin, ob die Teilnehmer bestimmte Internetseiten, TV-Sendungen oder Zeitschriften zur Recherche von Krankheitsthemen verwenden, gab es neben spezifischen Angaben, auch sehr allgemein gehaltene Antworten. Als klar definierte TV-Sendung wurde die Tagesschau genannt. Bei den Zeitschriften sind BILD der Frau sowie Psychologie Heute als Quellen angegeben. Es wurden drei Internetseiten erwähnt, Gesundheit.de, Qmul.ac.uk und das Ärzteblatt. Als Alternative wurden auch Ärzte, HNO-Ärzte und Frauenärzte geantwortet. Ansonsten wurden nicht näher beschriebene Internetseiten als Antwort eingetragen. Auffällig ist der große Anteil an Internetquellen. Von 48 Antworten waren 40 davon Quellen von diversen Internetseiten.

⁹ Anmerkung: Queen Mary-University of London.

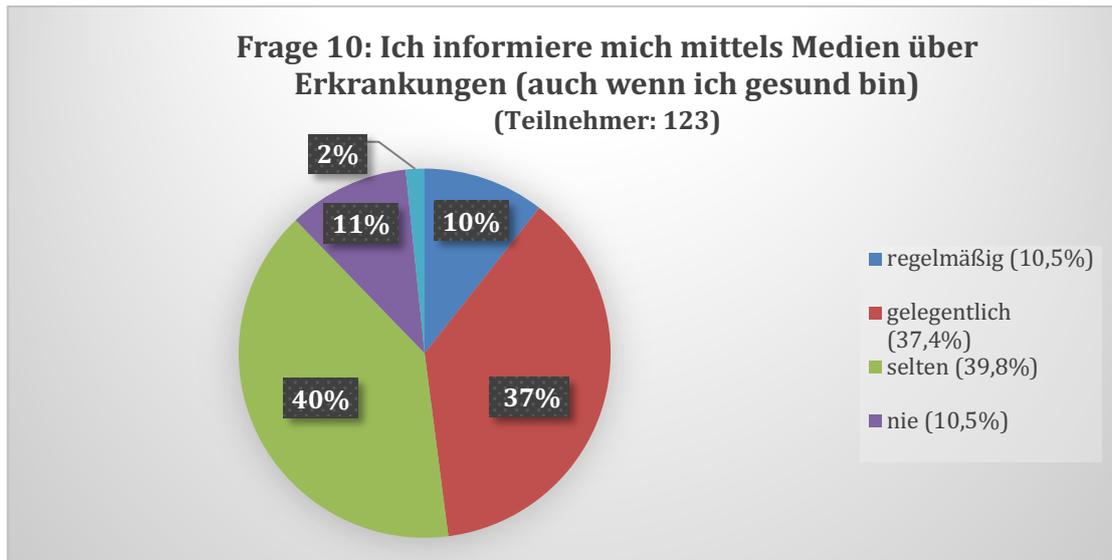


Abbildung 23: Recherche, während einer Erkrankung (Online-Umfrage 2020)

Bei dieser Frage sollen die Befragten ihr Suchverhalten bewerten, wenn sie gesund sind. Es soll untersucht werden, ob das Suchverhalten im Internet mit einer Erkrankung oder ohne eine Erkrankung verändert. Die Frage steht demnach im Gegensatz zu der Frage 7 und dient als Basis, um die zuvor aufgestellten Hypothesen zu belegen oder zu widerlegen.

Die Probanden konnten zwischen fünf Antwortmöglichkeiten auswählen, wobei die Möglichkeit zur Enthaltung gegeben wurde. Auffällig ist die Verteilung der einzelnen Angaben. Die beiden Extreme, „regelmäßig“ und „nie“, wurden mit jeweils 10,5 Prozent von den Teilnehmern gewählt. Nahezu identisch sind die Angaben bei „gelegentlich“ und „selten“. Beide haben eine ähnliche prozentuale Verteilung, 37,4 und 39,8 Prozent. Nur ein kleiner Teil von 1,6 Prozent haben sich der Antwort enthalten. Die Anzahl an Personen, die sich regelmäßig mittels Medien über Erkrankungen informieren ist in dieser Befragung genauso hoch, wie die Teilnehmer, die sich nie über Erkrankungen informieren. Die anschließende Abbildung zeigt den Vergleich zwischen den Fragen 7 und 10:

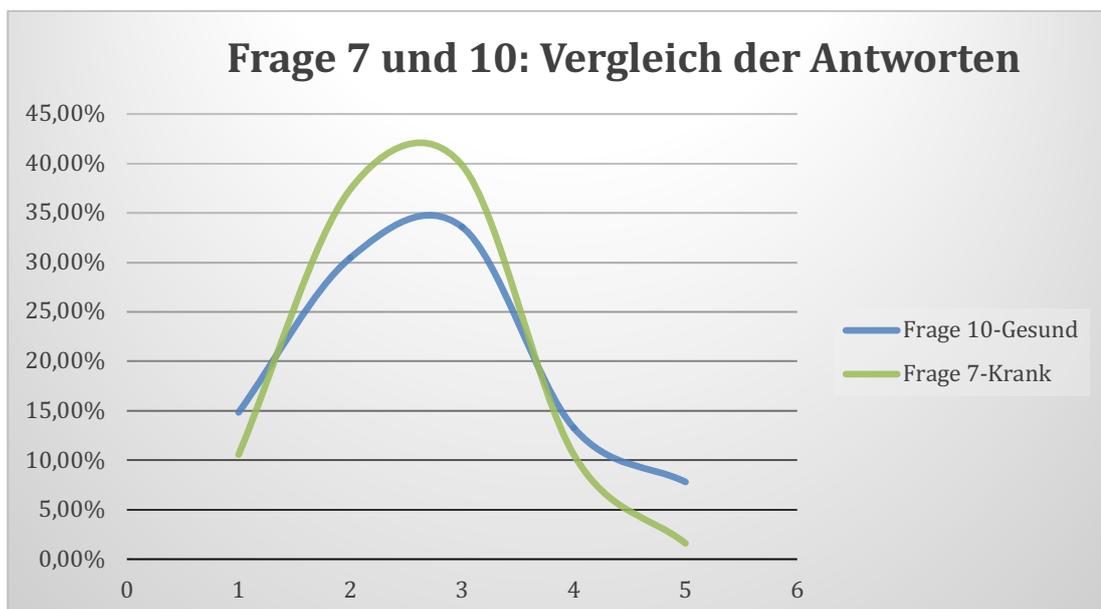


Abbildung 24: Vergleich der Antworten (Online-Umfrage 2020)

Deutlich zu erkennen ist der Zuwachs an Google Suchaktivitäten, während einer Erkrankung. Zwar ist der Unterschied nicht sehr groß, dennoch lässt sich eine Tendenz bezüglich des eigenen Gesundheitszustandes erkennen. Grundsätzlich ist mit dieser Datenmenge kein repräsentatives Ergebnis möglich. Die kommende Frage geht nochmals genauer auf das Suchverhalten in den letzten drei Monaten ein:

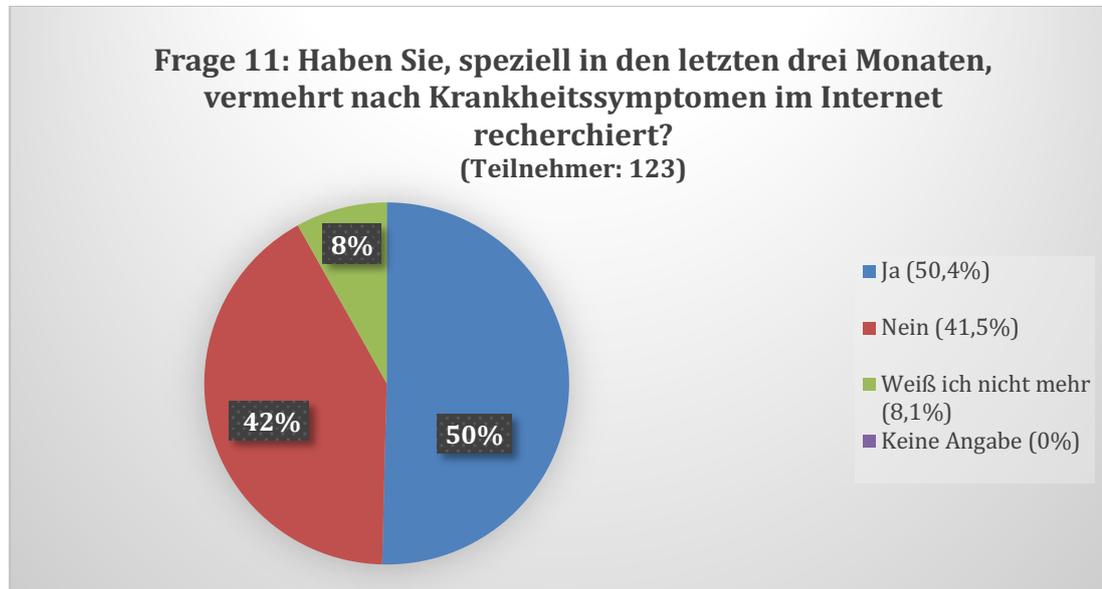


Abbildung 25: Recherche in den letzten drei Monaten (Online-Umfrage 2020)

Während die Fragen 7 und 10 das Suchverhalten in zwei unterschiedlichen Situationen untersucht, beschäftigt sich diese Frage mit dem Suchverhalten der letzten drei Monate. Dies ist besonders wichtig, da in den letzten Monaten bis heute ein hohes Medieninteresse zu COVID-19 herrscht. Die obige Abbildung fragt nach einem spezifischen Zeitraum, in denen die Probanden vermehrt nach Krankheitssymptomen recherchiert haben. Hier wurde kein klar definierter Grund genannt, warum es in den letzten drei Monaten zu einer erhöhten Suchaktivität kommen sollte, da es durch die aktuelle Lage der COVID-19-Pandemie offensichtlich ist. Die Befragung startete am zwölften August und umfasst rückblickend den Zeitraum bis zum Monat Mai. Es sollte vorab keine Beeinflussung durch bestimmte Schlagwörter stattfinden. Da ein signifikanter Anteil der Befragten einen Unterschied in ihrem Suchverhalten im Internet angegeben hat, kann man davon ausgehen, dass die Probanden die Frage richtig interpretiert haben.

Etwa 50 Prozent aller Teilnehmer haben angegeben, dass sie in den letzten drei Monaten öfter nach Krankheitssymptomen im Internet recherchiert haben als gewöhnlich. 41,5 Prozent der Befragten haben keine Veränderung in ihrem Suchverhalten feststellen können. 8,1 Prozent der Teilnehmer konnten keine Einschätzung über ihr Suchverhalten geben. Positiv ist es, dass es keine Enthaltungen zu dieser Frage gab, somit kann der Datensatz als vollständig betrachtet werden.

Die Ergebnisse aus Frage 10 werden mit diesem Ergebnis in Zusammenhang gebracht. Es soll überprüft werden, ob die Antworten in Korrelation zueinander stehen. Haben Probanden, die angegeben haben, dass sie bei Gesundheit selten bis nie nach Krankheitssymptomen im Internet recherchieren, auch in den letzten drei Monaten eher weniger Google zur Recherche genutzt. Die Antwortmöglichkeiten aus Frage 10, „selten“ und „nie“, werden zu einer Antwort

zusammengefasst. Insgesamt haben 62 Probanden diese Angaben gemacht. Auf die Frage hin, ob die Teilnehmer speziell in den letzten drei Monaten, eine vermehrte Recherche nach Krankheitssymptomen im Internet getätigt haben, ergibt das folgende Ergebnis:

- Ja: 27 Probanden, 43,5 Prozent
- Nein: 31 Probanden, 50 Prozent
- Weiß ich nicht mehr: 5 Probanden, 8,06 Prozent

Die prozentuale Verteilung ist ähnlich zu der obigen Abbildung. Zwar haben mehr Probanden angegeben, dass sie in den letzten drei Monaten keine erhöhten Rechercheaktivitäten vorgenommen haben, allerdings haben 43 Prozent der Teilnehmer eine vermehrte Rechercheaktivität bei sich feststellen können. Das ist ein erfreuliches Ergebnis, da es aufzeigt, dass sich das Suchverhalten im Internet verändert hat und auch verändern kann, wenn äußere Faktoren, wie die hohe Medienpräsenz von bestimmten Krankheiten, das Verhalten zu krankheitsbedingten Themen beeinflussen kann. Die Ergebnisse werden im kommenden Kapitel mit der Forschungsfrage und Hypothesen in einem Fazit zusammengefasst.

10 Fazit und Ausblick

Zu Beginn der Arbeit wurde die Problemstellung der Arbeit herausgearbeitet. Es wurde beschrieben, wie Google Trends durch das Sammeln von eingegebenen Suchwörtern Prognosen erstellen kann, die auf zukünftige gesellschaftliche Interessen hindeuten können. Die theoretische Grundlage der Arbeit dient dazu, Google Trends zu verstehen und wie es seine Daten sammelt und auswertet. Darüber hinaus wurden zwei zusätzliche Systeme betrachtet, welche die Datengrundlage von Google nutzen, um auf verschiedenen Gebieten Prognosen zu erstellen. Google Flu Trends nutzt die gesammelten Suchbegriffe, die zu diversen krankheitsbedingten Themen eingegeben wurden, um auf zukünftige Grippewellen hinzuweisen. Google Correlate hingegen wurde im Marketing verwendet, um zu prüfen, ob bestimmte Suchwörter im Vergleich zueinander über einen festgelegten Zeitraum noch an Popularität gewinnen können. Beide Systeme wurden bereits eingestellt aus den zuvor genannten Gründen. Des Weiteren wurde auf zwei weitere Systeme eingegangen, die eine Frühwarnfunktion innehaben bezüglich grippeähnlicher Krankheiten. Das vom Robert Koch-Institut entwickelte GrippeWeb überwacht akute Atemwegserkrankungen in der Bevölkerung und nutzt die Mitarbeit von Privatpersonen, um sich ein Bild über die derzeitige Lage zu machen. SORMAS ist für Vertreter von Gesundheitseinrichtungen entwickeltes Programm, das ursprünglich Ebola-Ausbrüche in Westafrika überwachen und eindämmen sollte. Das adaptive System wird heute für die COVID-19 Pandemie auch in Deutschland eingesetzt. Die beiden Systeme nutzen die Idee von Google Trends die Eingaben von Personen zu verwenden, um auf krankheitsbedingte Entwicklungen besser reagieren zu können.

Mit dieser theoretischen Grundlage wurde die Basis für die Forschungsfrage sowie die Hypothesen geschaffen.

Die Fragestellung dieser Arbeit beschäftigt sich mit der Fähigkeit von GT, gesellschaftliche Trends durch die Eingabe von Suchbegriffen der Google-Nutzer zu erkennen und zuverlässige Prognosen zu erstellen. Nicht nur gesellschaftliche Trends sollen erkannt werden, auch das intervenieren in künftigen Krankheitswellen sollen möglich gemacht werden. GFT verwendet ebenfalls die Suchwörter der Nutzer als Basis, um anbahnende Krankheitswellen vorherzusagen. Wie zuvor erwähnt, wurde GFT, aufgrund von falschen Prognosen zu Grippewellen in mehreren aufeinanderfolgenden Jahren, eingestellt. Aktuelle Studien gehen von einer erhöhten Berichterstattung in bestimmten Zeiträumen zu grippeähnlichen Krankheiten als Ursache für die Falschprognosen aus (vgl. Budras 2014; Cervellin et al. 2017). Das führte dazu, dass die Prognosen über die Anzahl an Personen, die an einer Grippe erkrankt sind, doppelt so hoch eingeschätzt wurden. Das hohe Medienaufkommen führte zu einer vermehrten Aktivität im Internet nach Krankheitssymptomen zu recherchieren. Die häufige Präsenz von bestimmten Krankheitsbildern im Fernsehen oder Zeitungen konnte demnach die Informationsqualität von GT respektive GFT beeinflussen.

Die Hypothesen sollen den Faktor der Medienberichterstattung auf das Suchverhalten untersuchen. Die Online-Umfrage hatte zum Ziel, das Suchverhalten von Personen zu untersuchen, die nach Krankheitssymptomen oder Heilungsmethoden im Internet recherchieren. Dabei war es wichtig zu wissen, ob sie das in einem Krankheitsfall tun oder gesund sind. Die daraus resultierenden Ergebnisse werden in diesem Abschnitt nochmals zusammengefasst und die Hypothesen belegt oder widerlegt.

Annähernd alle Probanden nutzen Google als primäre Suchmaschine für diverse Recherchen im Internet. Auf die Frage hin, ob die Teilnehmer von den aufgeführten Krankheiten gehört haben, Vogelgrippe, Schweinegrippe, EHEC, BSE, Coronavirus, waren die Antworten für beide Grippearten und dem Coronavirus ähnlich hoch. EHEC und BSE fielen dabei stark ab, was wohl durch die geringe Medienpräsenz erklären lässt. 45 Prozent aller Probanden haben angegeben, dass sie, im Falle einer Krankheit, Google nutzen, um sich über Symptome und Heilungsmethoden zu informieren. Etwa 48 Prozent der Befragten nutzen Google auch wenn sie gesund sind. Die Probanden wurden gefragt, in welchem Jahr sie zuletzt krank waren. 82 Prozent haben das Jahr 2019 und 2020 angegeben, also eine gute Voraussetzung, um sich an das eigene Suchverhalten zu erinnern. Anschließend sollten die Probanden bestimmte Quellen nennen, die sie als Informationsquelle für krankheitsbedingte Themen verwenden.

Die letzte Frage beschäftigt sich speziell mit dem Suchverhalten in den letzten drei Monaten. Wie zuvor erläutert wurde kein klar definierter Grund genannt, warum die letzten drei Monate als Zeitraum ausgewählt wurde. Dies schien in Zusammenhang mit krankheitsbedingten Themen derzeit in den Medien offensichtlich. 50 Prozent der Probanden haben angegeben, in den letzten drei Monaten vermehrt nach Krankheitssymptomen im Internet recherchiert zu haben. Ungefähr 42 Prozent haben kein verändertes Suchverhalten bei sich feststellen können. Um ein valides Ergebnis zu erhalten, wurden Frage 10 und 11 gemeinsam betrachtet. Hierbei wurden die Probanden bedacht, die bei Gesundheit selten bis nie nach Symptomen recherchierten. Daraus ergab sich ein prozentualer Anteil von 43 Prozent, also 27 von 62 Teilnehmern. Daher kann von einem signifikanten Anstieg gesprochen werden, der als Beleg für Hypothese 1 betrachtet werden kann. Da der Anteil an Probanden bei fast 50 Prozent liegt,

kann innerhalb dieser Online-Umfrage von einem repräsentativen Ergebnis gesprochen werden.

Hypothese 2 kann im Zuge dieser Arbeit weder belegt noch widerlegt werden. Um ein Beweis der Hypothese zu erbringen, hätten zusätzliche Fragen den Probanden gestellt werden müssen. Die Hypothese untersucht, ob geringes Medieninteresse zu einer bestimmten Krankheit auch zu weniger Suchaktivitäten zu krankheitsbedingten Themen führen würde. Frage Vier hätte sich als Ansatzpunkt für diese Hypothese als geeignet erweisen können. Die Krankheiten EHEC und BSE wurden am wenigsten von den Befragten gewählt. Daher wäre eine Frage, ob die Teilnehmer in letzter Zeit nach zugehörigen Symptomen zu den genannten Krankheiten im Internet recherchiert haben sinnvoll, um feststellen zu können, wie viele nach spezifischen Krankheitssymptomen zu den oben genannten Krankheiten recherchiert haben. Da beide Krankheiten seit dem Jahr 2000 und 2010 nicht mehr so stark in den Medien vertreten sind, könnte eine Abfrage diesbezüglich die H2 unterstützen.

Die Ergebnisse der Online-Umfrage haben aufgezeigt, dass bestimmte Faktoren das Rechercheverhalten von Nutzern im Internet beeinflussen kann. In Krankheitsfällen wird die Suchmaschine Google verwendet, um nach Symptomen oder Heilungsmethoden zu suchen. Der Weg zum Arzt wird weniger als erste Wahl gesehen, sondern eher als letzte Möglichkeit. Darüber hinaus spielt die Medienberichterstattung eine wichtige Rolle, wie Google-Nutzer nach krankheitsbedingten Themen recherchieren. Die Schwierigkeit besteht darin, herauszufinden, ob Menschen tatsächlich krank waren zur Zeit der Recherche oder ob das hohe Medieninteresse dazu führt das Suchverhalten zu verändern. Der hohe Anstieg von krankheitsrelevanten Suchbegriffen führte bei GFT dazu überschätzte Prognosen zu erstellen. Obwohl GFT nicht mehr im Betrieb ist, so gibt es immer noch Versuche das Programm wiederzubeleben. Neueste Studien fanden heraus, dass die Art und Weise wie Suchbegriffe eingegeben werden eine wichtige Rolle spielen könnten, um Krankheitsherde zu lokalisieren. Erste Versuche wurden mit Symptomen bezüglich des COVID-19 Virus getestet. Suchbegriffe können auch kurze Sätze umfassen. Die Suchanfrage „Ich kann nicht riechen“ deutet eher darauf hin, dass eine Person tatsächlich an diesem Symptom leidet, als wenn die Suchanfrage nur „Geruchsverlust“ heißen würde (vgl. Stephens-Davidowitz 2020). Zukünftige Studien über Suchbegriffe müssten klarer definieren können, ob Personen tatsächlich erkrankt sind, um falsche Prognosen in Zukunft entgegenzuwirken.

GT hat mit seinem Pool an Suchbegriffen die Möglichkeit das Interesse von seinen Nutzern darzustellen. Es kann erkennen, ob ein Thema in seiner Popularität steigt oder fällt. Dadurch lassen sich eventuelle Trends in Zukunft ablesen. Diese Möglichkeit erlaubt es Prognosen in bestimmten Themenfeldern zu treffen. GFT hat sich auf die Vorhersage von grippeähnlichen Krankheiten in der Bevölkerung fokussiert. Es analysiert alle krankheitsrelevanten Suchbegriffe und erstellt Prognosen zum Ausbruch einer Grippewelle. Ärztinnen und Ärzte, Praxen und Krankenhäuser können diese Prognosen nutzen, um sich auf Krankheitswellen besser vorzubereiten. Ähnliche Systeme wie GrippeWeb und SORMAS greifen die Idee von GT bzw. GFT auf und gehen einen Schritt weiter. Wie bei GT werden Eingaben von Privatpersonen verwendet, allerdings sind es keine Suchbegriffe, sondern Antworten auf

spezifisch gestellte Fragen zum Krankheitszustand. Das erlaubt ein deutlich präziseres Bild der Lage, allerdings ist der Datenbestand deutlich kleiner als bei GT.

Diese Arbeit hat sich das Prinzip von Google Trends als Grundlage genommen, um ähnliche Systeme und deren Funktionsweisen zu untersuchen und zu vergleichen. Die Online-Umfrage untersuchte, wie die Sucheingaben von Personen den Datenbestand von GT bzw. GFT verändern kann und so falsche Prognosen entstehen können. Daher muss nicht nur die Medienberichterstattung beachtet werden, sondern auch wie die Suchbegriffe formuliert werden. Google Trends als moderne „Wahrsagekugel“ klingt vielversprechend, dennoch braucht es noch weitere Studien, um diese Wahrsagekugel besser zu lesen.

11 Literaturverzeichnis

- Alby, Tom (2016): 5 Gründe warum du Google Trends falsch verstehst, [online]
<https://tom.alby.de/5-gruende-warum-du-google-trends-falsch-verstehst>
[Abruf: 14.07.2020].
- Bayer, Christophe, et al. Internet-based syndromic monitoring of acute respiratory illness in the general population of Germany, weeks 35/2011 to 34/2012. *Eurosurveillance*, 2014, 19. Jg., Nr. 4, S. 20684.
- Budras, Corinna (2014): Google weiß wo die Grippe lauert, [online]
<https://www.faz.net/aktuell/wirtschaft/netzwirtschaft/google-flu-trends-big-data-kann-helfen-uns-gegen-krankheiten-zu-wappnen-13268389-p3.html>
[Abruf: 20.07.2020].
- Bundeszentrale für politische Bildung (2020): Bevölkerung nach Altersgruppen und Geschlecht, [online]
<https://www.bpb.de/nachschlagen/zahlen-und-fakten/soziale-situation-in-deutschland/61538/altersgruppen> [Abruf: 08.09.2020].
- Carneiro, Herman Anthony; Mylonakis, Eleftherios. Google trends: a web-based tool for real-time surveillance of disease outbreaks. *Clinical infectious diseases*, 2009, 49. Jg., Nr. 10, S. 1557-1564.
- Cervellin, Gianfranco; Comelli I, Ivan; Lippi, Giuseppe. Is Google Trends a reliable tool for digital epidemiology? Insights from different clinical settings. *Journal of epidemiology and global health*, 2017, 7. Jg., Nr. 3, S. 185-189.
- Choi, Hyunyoung; Varian, Hal. *Predicting the Present with Google Trends*. 2011.
- Fähnrich, C., et al. Surveillance and Outbreak Response Management System (SORMAS) to support the control of the Ebola virus disease outbreak in West Africa. *Eurosurveillance*, 2015, 20. Jg., Nr. 12, S. 21071.
- Friesema, I. H. M., et al. Internet-based monitoring of influenza-like illness in the general population: experience of five influenza seasons in The Netherlands. *Vaccine*, 2009, 27. Jg., Nr. 45, S. 6353-6357.
- Helmholtz-Zentrum für Infektionsforschung (2020): *SORMAS*, [online]
<https://www.helmholtz-hzi.de> [Abruf: 07.08.2020].
- Kang, Min, et al. *Using google trends for influenza surveillance in South China. PloS one*, 2013, 8. Jg., Nr. 1, S. e55205.
- Krause, G., et al. (2020): Mobile Seuchenüberwachung: Dem Weg der Viren auf der Spur, [online]
<https://www.aerzteblatt.de/archiv/212882/Mobile-Seuchenueberwachung-Dem-Weg-der-Viren-auf-der-Spur> [Abruf: 08.08.2020].
- Lazer, David, et al. *The parable of Google Flu: traps in big data analysis. Science*, 2014, 343. Jg., Nr. 6176, S. 1203-1205.

- Norddeutscher Rundfunk (2020): Coronavirus: Bessere Nachverfolgung durch SORMAS?, [online] https://www.ndr.de/nachrichten/niedersachsen/hannover_weser-leinegebiet/Coronavirus-Bessere-Nachverfolgung-durch-SORMAS,sormas102.html [Abruf: 03.09.2020].
- Nuti, Sudhakar V., et al. *The use of google trends in health care research: a systematic review. PloS one*, 2014, 9. Jg., Nr. 10, S. e109583.
- Robert Koch-Institut (2020): GrippeWeb, [online] <https://grippeweb.rki.de/Default.aspx> [Abruf: 02.08.2020]
- Robert Koch-Institut (2020): GrippeWeb, [online] <https://grippeweb.rki.de> [Abruf: 29.07.2020].
- Rouse, Margaret (2017): Google Trends, [online] <https://www.computerweekly.com/de/definition/Google-Trends> [Abruf: 13.07.2020].
- Stephens-Davidowitz, Seth (2020): Google Searches Can Help Us Find Emerging Covid-19 Outbreaks, [online] <https://www.nytimes.com/2020/04/05/opinion/coronavirus-google-searches.html> [Abruf: 11.08.2020].
- Stephens-Davidowitz, Seth; Varian, Hal. A hands-on guide to Google data. *Further details on the construction can be found on the Google Trends page*, 2014.
- Thoma, Jörg (2014): Überschätzte Grippewelle durch Google Flu Trends, [online] <https://www.golem.de/news/algorithmen-ueberschaetzte-grippewelle-durch-google-flu-trends-1403-105165.html> [Abruf: 21.07.2020].
- Universität zu Köln (2020): Dr. Google: Online nach Krankheitssymptomen zu suchen, wirkt sich negativ auf Psyche aus, [online] <https://portal.uni-koeln.de/universitaet/aktuell/presseinformationen/detail/dr-google-online-nach-krankheitssymptomen-zu-suchen-wirkt-sich-negativ-auf-psyche-aus> [Abruf: 09.09.2020].
- Wagner, Gidon (2016): Google Correlate: Der kleine, unbekannte Bruder von Google Trends, [online] <https://blog.hubspot.de/marketing/marketing-mit-google-correlate> [Abruf: 27.07.2020].
- Wollenberg, Saskia (2020): Suchmaschinen in Deutschland, [online] <https://www.luna-park.de/blog/26276-suchmaschinenmarktanteile-deutschland/> [Abruf:24.08.2020].
- Wurm, Marian (2020): Suchvolumen, [online] <https://www.loewenstark.com/wissen/suchvolumen> [Abruf: 13.07.2020].
- Vitagroup.ag (2020): SORMAS bekämpft Affenpocken-Epidemie in Nigeria, [online] https://www.vitagroup.ag/de_DE/News-Insights/Sormas-Epidemie-Management-Programm [Abruf: 07.08.2020].

12 Anhänge

12.1 GrippeWeb Anmeldung

Anbei die Anmeldung bei GrippeWeb über die offizielle Seite des Robert Koch-Instituts.

Wenn Sie wissen möchten, wie eine Teilnahme funktioniert, finden Sie nähere Informationen unter dem Menüpunkt „[Wie funktioniert GrippeWeb?](#)“. Registrieren können Sie sich unter dem Menüpunkt „[Mitmachen bei GrippeWeb](#)“ oder mit einem Klick auf das folgende Emblem „Jetzt mitmachen!“:



Unterstützen Sie **GrippeWeb** und

- nehmen Sie teil an einem neuartigen Projekt des Robert Koch-Instituts
- leisten Sie einen wichtigen Beitrag zur Erfassung akuter Atemwegserkrankungen in Deutschland
- erhalten Sie die **neuesten Informationen** zu akuten Atemwegserkrankungen, grippeähnlichen Erkrankungen und der „echten“ Grippe
- erleben Sie **Epidemiologie in Echtzeit** und zum Anfassen
- und – gewinnen Sie mit ein bisschen Glück einen der **wertvollen Preise**, als Dankeschön für Ihre Teilnahme

GrippeWeb - Kontoerstellung

E-Mail:

Passwort:

Passwort (Wiederholung):

Bitte übertragen Sie die links angezeigte fünfstellige Zahlen- und Buchstabenkombination in das Eingabefeld rechts.



Hiermit erkläre ich mich einverstanden, freiwillig bei GrippeWeb teilzunehmen. Meine Angaben werden ausschließlich vom Robert Koch-Institut gespeichert und nicht an Dritte weitergegeben. Weitere Informationen zum Datenschutz finden Sie [hier](#).

Sie sind hier: [Startseite](#) [Ihre Kontoangaben](#) [Ihre Basisdaten](#)

Liste der registrierten Teilnehmer im Haushalt

Als erstes müssen Sie Ihre eigenen Basisdaten anlegen.

Neuen Teilnehmer hinzufügen

Spitzname ?

Geburtsjahr

Geschlecht männlich weiblich

Körpergröße in cm

Körpergewicht in kg

Übliche tägliche Beschäftigung

Meist benutztes Verkehrsmittel:::

Keines der genannten Verkehrsmittel, sondern ?

Chronische Vorerkrankungen

Herzerkrankung ? Ja Nein Weiß nicht

chronische Bronchitis ? Ja Nein Weiß nicht

Asthma ? Ja Nein Weiß nicht

Diabetes ? Ja Nein Weiß nicht

andere Ja Nein Weiß nicht

Raucher Ja Nein

Teilnahme am Gewinnspiel

Ich bin mindestens 18 Jahre alt

Nur für mindestens 18-jährige: Ich nehme am [Gewinnspiel](#) teil

Hinweis: Diese Daten können Sie jederzeit bearbeiten!

Liste der registrierten Teilnehmer im Haushalt

Spitzname	Geburtsjahr	Geschlecht	Größe in cm	Gewicht in kg	Beschäftigung	Verkehrsmittel	Verkehrsmittel sonst.	Herzerkrankung	chr. Bronchitis	Asthma	Diabetes	andere	Raucher
Selcuk	1989	männlich	170	80	Ausbildung	OEP		nein	nein	nein	nein	nein	nein

Kind hinzufügen

Sie können nun weitere Personen aus Ihrem Haushalt unter 14 Jahren hinzufügen (s. Knopf „Kind hinzufügen“). Personen, die mindestens 14 Jahre alt sind, müssen sich selbst registrieren. Wenn Sie keinen weiteren Teilnehmer hinzufügen möchten, klicken Sie bitte auf den Knopf „Weiter“ zu den Haushaltsdaten.

Angaben zum Haushalt

Landkreis

Wohnen Sie in einem ... ?

"Spitzname" für Haushalt (Wohnheim, Gemeinschaftseinrichtung) ?

Anzahl der Personen im Haushalt unter 14 Jahren ?

Anzahl der Personen im Haushalt einschließlich Ihnen, die mindestens 14 Jahre alt sind ?

Wie sind Sie auf uns aufmerksam geworden?

Hinweis: Diese Daten können Sie jederzeit bearbeiten!

Teilnehmer Selcuk

Woche 30 vom 20.07.2020 bis 26.07.2020 Wöchentliche Fragen beantworten

Hatten Sie in der Woche vom Montag, 20.07.2020 - Sonntag, 26.07.2020 eine NEU aufgetretene Atemwegserkrankung (zum Beispiel neu aufgetretenen Husten, Schnupfen, Halsschmerzen, mit oder ohne Fieber)?*

Ja

Nein

Woche 29 vom 13.07.2020 bis 19.07.2020 Wöchentliche Fragen beantworten

Woche 28 vom 06.07.2020 bis 12.07.2020 Wöchentliche Fragen beantworten

Woche 27 vom 29.06.2020 bis 05.07.2020 Wöchentliche Fragen beantworten

Teilnehmer Selcuk

Woche 30 vom 20.07.2020 bis 26.07.2020 Wöchentliche Fragen beantworten 

Hatten Sie in der Woche vom Montag, 20.07.2020 - Sonntag, 26.07.2020 eine NEU aufgetretene Atemwegserkrankung (zum Beispiel neu aufgetretenen Husten, Schnupfen, Halsschmerzen, mit oder ohne Fieber)?*

Ja
 Nein

Bitte geben Sie an, ob Sie folgende Beschwerden hatten

Husten
 Ja
 Nein

Halsschmerzen
 Ja
 Nein

Fieber
 Ja
 Nein

Schnupfen
 Ja
 Nein

Woche 29 vom 13.07.2020 bis 19.07.2020 Wöchentliche Fragen beantworten 

Woche 28 vom 06.07.2020 bis 12.07.2020 Wöchentliche Fragen beantworten 

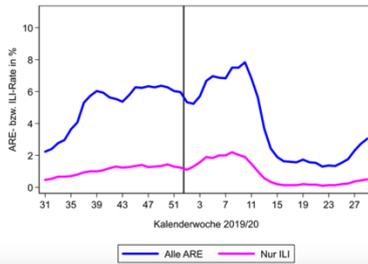
Woche 27 vom 29.06.2020 bis 05.07.2020 Wöchentliche Fragen beantworten 

Ergebnisse der Woche 29/2020

Auf der Startseite wird beschrieben, dass die für die Bevölkerung in Deutschland geschätzte Rate von neu aufgetretenen, akuten Atemwegserkrankungen (ARE) in der 29. Kalenderwoche (KW) (13.07. – 19.07.2020) stabil geblieben ist. Die Rate der grippeähnlichen Erkrankungen (GLE, definiert als ARE mit Fieber) ist im Vergleich zur Vorwoche ebenfalls stabil geblieben.

[Hier geht's zur Startseite](#)

Abbildung 1 zeigt die Gesamt-ARE- und Gesamt-ILI-Rate als geglättete Raten, d. h. sie sind jeweils über drei Wochen gemittelt. Die Grippezeit (2. – 12. KW 2020) stellt sich deutlich sowohl in der Gesamt-ARE- als auch in der Gesamt-ILI-Rate dar. Nach einem starken und abrupten Rückgang der Gesamt-ARE-Rate von der 11. KW bis zur 15. KW 2020 steigt sie, nach einer relativ stabilen Phase, seit etwa der 25. KW 2020 wieder an. Die Gesamt-ILI-Rate zeigt einen ähnlich markanten Verlauf, bleibt jedoch auch nach der 25. KW 2020 stabil. Die wegen der COVID-19-Pandemie geschlossenen Kitas und Schulen und die von der Bundesregierung beschlossenen Kontaktbeschränkungen scheinen zu einer deutlichen Reduzierung der ARE- und ILI-Aktivität vor allem in den jüngeren Altersgruppen beigetragen zu haben (Start der Maßnahmen mit Beginn der 11. KW 2020). Seit der 25. KW und 26. KW sind einzelne Bundesländer bei den Kitas wieder in den vollständigen Regelbetrieb übergegangen.



12.2 Fragenkatalog

Anbei der Fragenkatalog mit Einleitungstext

Liebe Teilnehmende,

Sie sind herzlichst dazu eingeladen bei meinem Fragebogen teilzunehmen. Im Rahmen meiner Bachelorarbeit an der HAW, untersuche ich das Suchverhalten von Google-Nutzern in Bezug auf Gesundheitsthemen.

An der Umfrage können alle teilnehmen, die Google als primäre Suchmaschine verwenden. Die Umfrage dauert maximal 3-5 Minuten.

Hinweis: Es gibt keine richtigen oder falschen Antworten.

Antworten Sie bitte so ehrlich wie möglich. Alle Antworten werden Anonym behandelt und erlauben keine Rückschlüsse auf individuelle Personen.

Vielen Dank für Ihre Unterstützung!

Bei Fragen: selcuk.acikbas@haw-hamburg.de

Selcuk Acikbas

Frage 1

Wie alt sind Sie?

Frage 2

Sie sind...?

- Männlich
- Weiblich
- Divers

Frage 3

Welche Suchmaschinen nutzen Sie? (Mehrfachantworten möglich)

- Google
- Yahoo
- Bing
- DuckDuckGo
- Andere

Frage 4

Haben Sie bereits von den unten stehenden Krankheiten gehört? (Mehrfachantworten möglich)

- Vogelgrippe
- Schweinegrippe

- EHEC
- BSE
- Coronavirus

Frage 5

Wenn Ja, wo haben Sie von den Krankheiten gehört? (Mehrfachantworten möglich)

- TV
- Radio
- Zeitung/Zeitschrift
- Social Media
- Internet, Nachrichtenseiten
- Betroffene Personen
- Bekannte/Familie

Frage 6

In welchem Jahr waren Sie zuletzt krank?

Frage 7

Wenn ich krank bin nutze ich die Suchmaschine Google, um nach Symptomen und Heilungsmethoden zu recherchieren.

- 1-trifft voll und ganz zu
- 2-trifft eher zu
- 3-teils, teils
- 4-trifft eher nicht zu
- 5-trifft gar nicht zu

Frage 8

Wenn Sie zwischen „teils, teils“ und „trifft gar nicht zu“ entschieden haben, welche Alternativen suchen Sie stattdessen auf?

- Arztbesuch
- Krankenhaus
- Freunde/Verwandte fragen
- Krankheit abklingen lassen
- Sonstiges

Frage 9

Nutzen Sie bestimmte Internetseiten, TV-Sendungen, Zeitschriften für die Recherche zu Krankheitsthemen?

Frage 10

Ich informiere mich mittels Medien über Erkrankungen (auch wenn ich gesund bin).

- 1-regelmäßig
- 2-gelegentlich
- 3-selten
- 4-nie
- 5-keine Angabe

Frage 11

Haben Sie, speziell in den letzten drei Monaten, vermehrt nach Krankheitssymptomen im Internet recherchiert?

- Ja
- Nein
- Weiß ich nicht mehr
- Keine Angabe

Frage 12

Sie leben zurzeit in...

Frage 13

Sie sind...

- Schüler/in
- Student/in
- Angestellte/r
- Auszubildende/r
- Rentner/in
- Selbständig
- Beamte/r
- Arbeitssuchend
- anderes

Geschafft!

Vielen Dank für die Teilnahme an meinem Fragebogen. Zweck der Umfrage war es, herauszufinden, ob Google-Nutzer nach Gesundheitsthemen im Netz recherchieren, auch wenn sie zurzeit gesund sind. Es soll vor allem untersucht werden, ob ein hohes Medieninteresse das Suchverhalten der Nutzer beeinflussen kann wie z.B. derzeit bei der Covid-19 Pandemie. Durch zukünftige Forschungen und Analysen der Suchanfragen, lassen sich eventuelle anbahnende Krankheitswellen eingrenzen.

Bitte klicken Sie auf "Fertig", um die Umfrage zu schließen.

12.3 SPSS-Tabellen

Frage 1

Statistiken

Alter

N	Gültig	137
	Fehlend	4
Mittelwert		26,43
Std.-Abweichung		5,888
Varianz		34,673
Spannweite		40
Minimum		14
Maximum		54

Alter

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	14	1	,7	,7
	15	1	,7	1,5
	16	1	,7	2,2
	17	2	1,4	3,6
	18	1	,7	4,4
	19	2	1,4	5,8
	20	7	5,0	10,9
	21	11	7,8	19,0
	22	9	6,4	25,5
	23	7	5,0	30,7
	24	10	7,1	38,0
	25	14	9,9	48,2
	26	9	6,4	54,7
	27	13	9,2	64,2
	28	9	6,4	70,8
	29	8	5,7	76,6
	30	9	6,4	83,2
	31	7	5,0	88,3
	32	1	,7	89,1
	33	1	,7	89,8
34	2	1,4	91,2	
35	3	2,1	93,4	
36	1	,7	94,2	
37	1	,7	94,9	
38	1	,7	95,6	
39	2	1,4	97,1	
40	1	,7	97,8	
44	1	,7	98,5	
45	1	,7	99,3	
54	1	,7	100,0	
Gesamt	137	97,2	100,0	
Fehlend	System	4	2,8	
Gesamt		141	100,0	

Frage 2

Geschlecht

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	3	2,1	2,1	2,1
2. Sie sind...?	1	,7	,7	2,8
Divers	2	1,4	1,4	4,3
insgesamt	1	,7	,7	5,0
Keine Angabe	1	,7	,7	5,7
Männlich	34	24,1	24,1	29,8
Weiblich	99	70,2	70,2	100,0
Gesamt	141	100,0	100,0	

Frage 3

Häufigkeiten von \$FrageSuchmaschine

		Antworten		Prozent der Fälle
		N	Prozent	
Frage_Suchmaschine ^a	Google	133	72,7%	98,5%
	Yahoo	4	2,2%	3,0%
	Bing	11	6,0%	8,1%
	DuckDuckGo	12	6,6%	8,9%
	Andere	23	12,6%	17,0%
Gesamt		183	100,0%	135,6%

a. Dichotomie-Gruppe tabellarisch dargestellt bei Wert 1.

Frage 4

Häufigkeiten von \$Krankheiten

		Antworten		Prozent der Fälle
		N	Prozent	
Krankheiten ^a	Vogelgrippe	125	23,0%	96,2%
	Schweinegrippe	127	23,3%	97,7%
	EHEC	82	15,1%	63,1%
	BSE	81	14,9%	62,3%
	Coronavirus	129	23,7%	99,2%
Gesamt		544	100,0%	418,5%

a. Dichotomie-Gruppe tabellarisch dargestellt bei Wert 1.

Zusammenfassung der Fallverarbeitung

	Fälle					
	Gültig		Fehlend		Gesamt	
	N	Prozent	N	Prozent	N	Prozent
Alter * Vogelgrippe	125	88,7%	16	11,3%	141	100,0%
Alter * Schweinegrippe	127	90,1%	14	9,9%	141	100,0%
Alter * EHEC	82	58,2%	59	41,8%	141	100,0%
Alter * BSE	81	57,4%	60	42,6%	141	100,0%
Alter * Coronavirus	129	91,5%	12	8,5%	141	100,0%

Verhältnistatistik für Alter / Vogelgrippe

Mittelwert	Preisgebundene Differenz	Streuungskoeffizient	Variationskoeffizient Zentrierter Median
26,688	1,000	,159	22,6%

Verhältnistatistik für Alter / Schweinegrippe

Mittelwert	Preisgebundene Differenz	Streuungskoeffizient	Variationskoeffizient Zentrierter Median
26,583	1,000	,161	22,6%

Verhältnistatistik für Alter / EHEC

Mittelwert	Preisgebundene Differenz	Streuungskoeffizient	Variationskoeffizient Zentrierter Median
27,293	1,000	,153	21,7%

Verhältnistatistik für Alter / BSE

Mittelwert	Preisgebundene Differenz	Streuungskoeffizient	Variationskoeffizient Zentrierter Median
28,963	1,000	,138	20,9%

Verhältnistatistik für Alter / Coronavirus

Mittelwert	Preisgebundene Differenz	Streuungskoeffizient	Variationskoeffizient Zentrierter Median
26,558	1,000	,159	22,6%

Frage 5

Häufigkeiten von \$VonWoGehört

		Antworten		Prozent der Fälle
		N	Prozent	
Von_Wo_Gehört ^a	TV	120	19,1%	93,0%
	Radio	97	15,4%	75,2%
	ZeitungZeitschrift	92	14,6%	71,3%
	SozialeMedien	107	17,0%	82,9%
	Internet	119	18,9%	92,2%
	BetroffenePersonen	30	4,8%	23,3%
	BekannteFamilie	64	10,2%	49,6%
Gesamt		629	100,0%	487,6%

a. Dichotomie-Gruppe tabellarisch dargestellt bei Wert 1.

Frage 6

ZuletztKrank

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	2010	1	,7	,8	,8
	2016	1	,7	,8	1,6
	2017	2	1,4	1,6	3,1
	2018	19	13,5	14,7	17,8
	2019	55	39,0	42,6	60,5
	2020	51	36,2	39,5	100,0
	Gesamt	129	91,5	100,0	
Fehlend	System	12	8,5		
Gesamt		141	100,0		

Frage 7

Statistiken

		KrankGoogle	Symptome
N	Gültig	0	128
	Fehlend	141	13
Mittelwert			2,69
Std.-Abweichung			1,121
Varianz			1,256
Spannweite			4

Symptome

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	1	19	13,5	14,8	14,8
	2	39	27,7	30,5	45,3
	3	43	30,5	33,6	78,9
	4	17	12,1	13,3	92,2
	5	10	7,1	7,8	100,0
	Gesamt	128	90,8	100,0	
Fehlend	System	13	9,2		
Gesamt		141	100,0		

Frage 8

WennNichtGoogle

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig		70	49,6	49,6	49,6
	8. Wenn Sie zwischen 'teils, teils' und 'trifft gar nicht zu' gewählt haben, welche Alternativen suchen Sie stattdessen auf?	1	,7	,7	50,4
	Arztbesuch	42	29,8	29,8	80,1
	Freunde/Verwandte Fragen	7	5,0	5,0	85,1
	Ich lasse die Krankheit abklingen	17	12,1	12,1	97,2
	Krankenhaus	1	,7	,7	97,9
	Sonstiges	3	2,1	2,1	100,0
	Gesamt	141	100,0	100,0	

Frage 10

Statistiken

		GesundGoogl e	Erkrankunge n
N	Gültig	0	123
	Fehlend	141	18
Mittelwert			2,55
Std.-Abweichung			,880
Varianz			,774
Spannweite			4

Erkrankungen

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	1	13	9,2	10,6	10,6
	2	46	32,6	37,4	48,0
	3	49	34,8	39,8	87,8
	4	13	9,2	10,6	98,4
	5	2	1,4	1,6	100,0
	Gesamt	123	87,2	100,0	
Fehlend	System	18	12,8		
Gesamt		141	100,0		

Frage 11

DreiMonateRecherche

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig		17	12,1	12,1	12,1
	11. Haben Sie, speziell in den letzten drei Monaten, häufiger nach Krankheitssymptomen im Internet recherchiert?	1	,7	,7	12,8
	Ja	62	44,0	44,0	56,7
	Nein	51	36,2	36,2	92,9
	Weiß ich nicht mehr	10	7,1	7,1	100,0
	Gesamt	141	100,0	100,0	

Frage 12

Wohnort

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig		19	13,5	13,5	13,5
	Baden-Württemberg	10	7,1	7,1	20,6
	Bayern	10	7,1	7,1	27,7
	Berlin	1	,7	,7	28,4
	Bremen	2	1,4	1,4	29,8
	Hamburg	50	35,5	35,5	65,2
	Hessen	6	4,3	4,3	69,5
	Niedersachsen	7	5,0	5,0	74,5
	Nordrhein-Westfalen	18	12,8	12,8	87,2
	Rheinland-Pfalz	1	,7	,7	87,9
	Saarland	1	,7	,7	88,7
	Sachsen	1	,7	,7	89,4
	Schleswig-Holstein	14	9,9	9,9	99,3
	Thüringen	1	,7	,7	100,0
	Gesamt	141	100,0	100,0	

Frage 13

Berufsstand

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	17	12,1	12,1	12,1
13. Sie sind...	1	,7	,7	12,8
anderes	5	3,5	3,5	16,3
Angestellte/r	35	24,8	24,8	41,1
arbeitssuchend	2	1,4	1,4	42,6
Auszubildende/r	1	,7	,7	43,3
Rentner/in	1	,7	,7	44,0
Schüler/in	6	4,3	4,3	48,2
Selbstständig	2	1,4	1,4	49,6
Student/in	71	50,4	50,4	100,0
Gesamt	141	100,0	100,0	

Ich versichere, die vorliegende Arbeit selbstständig ohne fremde Hilfe verfasst und keine anderen Quellen und Hilfsmittel als die angegebenen benutzt zu haben. Die aus anderen Werken wörtlich entnommenen Stellen oder dem Sinn nach entlehnten Passagen sind durch Quellenangabe kenntlich gemacht.