

Hochschule Rhein-Waal
Fakultät Kommunikation und Umwelt
Prof. Dr. Frank Zimmer
Prof. Dr.-Ing. Ido Iurgel
Abgabedatum: 06.Februar 2018

**Die (gesellschaftliche)
Digitalisierung in Deutschland
am Beispiel von Twitter-Nachrichten
—
Visualisierung in R und Shiny**

Bachelorarbeit
im Studiengang
Medien- und Kommunikationsinformatik
zur Erlangung des akademischen Grades

Bachelor of Science

vorgelegt von
Kristina-Susann Baudach
Matrikelnummer: 16882

Friedrich-Heinrich-Allee 35
47475 Kamp-Lintfort

Abstract

This Bachelor-Thesis is about the implementation of a web-application with R and Shiny to visualise the grade of (social) digitalisation in Germany and Germany's federal states with Twitter-Messages, named tweets, in an interactive, cartographic infographic.

R is used as a programming language in the environment RStudio. The R-Package "shiny" makes it possible to implement a web-application straight from R. The R-Package "leaflet" implements the map. The data of the tweets comes from the Twitter-Streaming-API and is used with the R-Package "rtweets".

The implementation includes the scripts "ui.R", for the User Interface, "server.R" and 22 more scripts, for the implementation of the function. Also, it includes a file named "style.css" for the style.

For presenting that it is possible to visualise different questions about the digitalisation of society with tweets, two examples are given. The first question is on which daytime the digitalisation is the strongest and the other question is on which day the digitalisation is the strongest. Therefore, there are four daily streams for an hour from 07.12.2017 to 13.12.2017, which show that the digitisation is the strongest in North Rhine-Westphalia and Berlin in the evening. While North Rhine-Westphalia has the strongest digitalisation on 10.12.2017 and Berlin's digitalisation is strongest on 08.12.2017.

After that the critic about this application is described. The strongest critic is the question about the representativity of tweets for the digitalisation. And there are many possibilities for additional given examinations.

Digitalisierung • Twitter • R • Shiny • Visualisierung

Inhaltsverzeichnis

Abkürzungsverzeichnis	VII
Abbildungsverzeichnis	VIII
Tabellenverzeichnis	XI
1 Einleitung	01
1.1 Relevanz des Themas	01
1.2 Zielsetzung	02
1.3 Vorgehensweise	03
2 Digitalisierung	05
2.1 Definition „Digitalisierung“	05
2.2 Stand der Digitalisierung in Deutschland	06
2.2.1 Die digitale Agenda der deutschen Bundesregierung	07
2.2.2 Studien der europäischen Union zur Digitalisierung in Deutschland	08
2.2.3 Deutsche Studien zum Stand der Digitalisierung	09
3 Soziale Medien im Digitalisierungsprozess	12
3.1 Definition „soziale Medien“	12
3.2 Soziale Medien als ein Indikator der Digitalisierung in Deutschland	13
4 Twitter als Teil sozialer Medien im Digitalisierungsprozess ..	15
4.1 Twitter als Teil sozialer Medien	15
4.2 Tweets als ein Indikator der Digitalisierung	15
5 Visualisierung von Daten	17
5.1 Definition „Informationsgrafik“ („Infografik“)	17
5.2 Interaktive, kartografische Infografik	17
5.3 Visualisierung von Daten aus sozialen Medien	18
6 Verwendete Werkzeuge	20
6.1 R und RStudio	20
6.2 Das Twitter-API	20
6.2.2 Twitter-Search-API versus Twitter-Streaming-API	21
6.2.3 Verwendung des Twitter-API im Rahmen dieser Arbeit	23
6.3 R-Package „rtweets“	24
6.4 R-Package „shiny“	25
6.4.1 Die Benutzeroberfläche	26
6.4.2 Die Server-Funktion	26
6.4.3 Die Veröffentlichung einer Shiny-App	27
6.5 R-Package „leaflet“	28

7 Visualisierung der Digitalisierung in den Bundesländern . . .	30
7.1 „ui.R“ und „style.css“ der Shiny-App	31
7.2 „server.R“ der Shiny-App	32
7.3 Weitere eingebundene R-Scripte in alphabetischer Reihenfolge	35
7.3.1 entscheidungKarte.R	35
7.3.2 ermittlungBunNr.R	35
7.3.3 erzeugeDateiname.R und erzeugeDateinameDownload.R	35
7.3.4 erzeugeDynText.R	35
7.3.5 fehlermeldung.R	36
7.3.6 focus.R	36
7.3.7 geoDaten.R	37
7.3.8 karteLeaflet.R	37
7.3.9 leseCSV.R	37
7.3.10 liegtIn.R	37
7.3.11 react.R	37
7.3.12 schreibeCSV.R	38
7.3.13 sortiereDaten.R	38
7.3.14 spaltennamenKorrigieren.R	39
7.3.15 startKarte.R	39
7.3.16 streamGeo.R	39
7.3.17 streamZeichnen.R	40
7.3.18 twitterAuthentifizierung.R	40
7.3.19 vergleichZeichnen.R	41
7.3.20 zaehleTweets.R	41
7.3.21 zeichnen.R	41
8 Beispiele für die Verwendung und Auswertung der App	42
8.1 Tageszeiten mit der stärksten Digitalisierung durch Tweets	42
8.2 Tag mit der stärksten Digitalisierung durch Tweets	44
9 Abschlussbetrachtung	49
9.1 Kritische Betrachtung	49
9.2 Ausblick für die Forschung	51
9.3 Fazit	52
Literaturverzeichnis	55
Anhang	61
Anlage A: Spezifikation der verwendeten Versionen der Werkzeuge	61
A.1 Technische Spezifikationen des verwendeten Rechners	61
A.2 R-Version	61
A.3 RStudio-Version	62
A.4 Rtweet-Version	62

A.5 Leaflet-Version	63
A.6 Shiny-Version	64
Anlage B: R-Scripte	65
B.1 ui.R	65
B.2 server.R	68
B.3 entscheidungKarte.R	70
B.4 ermittlungBunNr.R	70
B.5 erzeugeDateiname.R	71
B.6 erzeugeDateinameDownload.R	71
B.7 erzeugeDynText.R	71
B.8 fehlermeldung.R	72
B.9 focus.R	72
B.10 geoDaten.R	72
B.11 karteLeaflet.R	73
B.12 leseCSV.R	73
B.13 liegtIn.R	73
B.14 react.R	73
B.15 schreibeCSV.R	73
B.16 sortiereDaten.R	74
B.17 spaltennamenKorrigieren.R	74
B.18 startKarte.R	74
B.19 streamGeo.R	74
B.20 streamZeichnen.R	75
B.21 twitterAuthentifizierung.R	75
B.22 vergleichZeichnen.R	75
B.23 zaehleTweets.R	76
B.24 zeichnen.R	76
B.25 style.css	76
B.26 Einwohnerzahlen.csv	77
Anlage C: Visualisierungen der Streams vom 07. bis 13.12.2017 in Berlin	78
C.1 Daten vom 07.12.2017	78
C.2 Daten vom 08.12.2017	79
C.3 Daten vom 09.12.2017	80
C.4 Daten vom 10.12.2017	81
C.5 Daten vom 11.12.2017	82
C.6 Daten vom 12.12.2017	83
C.7 Daten vom 13.12.2017	84
Anlage D: Visualisierungen der Streams vom 07. bis 13.12.2017 in NRW	85
D.1 Daten vom 07.12.2017	85
D.2 Daten vom 08.12.2017	86
D.3 Daten vom 09.12.2017	87
D.4 Daten vom 10.12.2017	88

D.5 Daten vom 11.12.2017	89
D.6 Daten vom 12.12.2017	90
D.7 Daten vom 13.12.2017	91
Anlage E: CSV-Datei vom Stream am 07.12. 2017 morgens	92
Selbstständigkeitserklärung	105

Abkürzungsverzeichnis

API	Application Programming Interface
App	Web-Anwendung
BMI	Bundesministerium des Innern
BMVI	Bundesministerium für Verkehr und digitale Infrastruktur
BMWI	Bundesministerium für Wirtschaft und Energie
BPA	Presse- und Informationsamt der Bundesregierung
CSS	Cascading Style Sheets
DESI	Digital Economy and Society Index
EDPR	Europe's Digital Progress Report
HTTP	Hypertext Transfer Protocol
IDE	integrated development environment
JSON	JavaScript Object Notation
IKT	Informations- und Kommunikationstechnik
IuK	Information und Kommunikation
NRW	Nordrhein-Westfalen
ÖFIT	Kompetenzzentrum Öffentliche IT
SSL	Secure Sockets Layer
Tweet	Twitter-Nachricht
UI	Benutzeroberfläche (von engl. User Interface)

Abbildungsverzeichnis

Abbildung 1.1:	Übersicht über die Struktur dieser Arbeit.....	03
Abbildung 2.1:	Digitalisierung als globaler Megatrend	05
Abbildung 2.2:	DESI 2017 - relative Leistung nach Bereichen	08
Abbildung 2.3:	Deutschland-Index der Digitalisierung 2017	10
Abbildung 3.1:	Soziale Medien in Deutschland	13
Abbildung 5.1:	Beispiel einer interaktive, kartografische Infografik in dieser Arbeit	18
Abbildung 6.1:	Aufbau einer Shiny-App - Die „app.R“	27
Abbildung 7.1:	Ordnerstruktur der Shiny-App	30
Abbildung 7.2:	Übersicht über den strukturellen Aufbau der Shiny-App	30
Abbildung 8.1:	Tweets in Berlin morgens (blau) und mittags (grün).....	43
Abbildung 8.2:	Tweets in Berlin nachmittags (rot) und abends (gelb)	43
Abbildung 8.3:	Tweets in NRW morgens (blau) und mittags (grün)	43
Abbildung 8.4:	Tweets in NRW nachmittags (rot) und abends (gelb).....	43
Abbildung 8.5:	Tweets in Berlin mittags (grün) und abends (gelb)	44
Abbildung 8.6:	Tweets in NRW mittags (grün) und abends (gelb).....	44
Abbildung 8.7:	Tweets in Berlin am 07.12.2017 (blau) und am 08.12.2017 (grün).....	45
Abbildung 8.8:	Tweets in NRW am 07.12.2017 (blau) und am 08.12.2017 (grün).....	45
Abbildung 8.9:	Tweets in Berlin am 09.12.2017 (rot) und am 10.12.2017 (gelb).....	46
Abbildung 8.10:	Tweets in NRW am 09.12.2017 (rot) und am 10.12.2017 (gelb)	46
Abbildung 8.11:	Tweets in Berlin am 11.12.2017 (lila) und am 12.12.2017 (pink)	46
Abbildung 8.12:	Tweets in NRW am 11.12.2017 (lila) und am 12.12.2017 (pink).....	46
Abbildung 8.13:	Tweets in Berlin am 13.12.2017 (orange) und am 07.12.2017 (blau)	47
Abbildung 8.14:	Tweets in NRW am 13.12.2017 (orange) und am 07.12.2017 (blau)	47
Abbildung 8.15:	Tweets in Berlin am 08.12.2017 (grün) und am 10.12.2017 (gelb).....	47
Abbildung 8.16:	Tweets in NRW am 08.12.2017 (grün) und am 10.12.2017 (gelb).....	47
Abbildung 8.17:	Tweets in Berlin am 12.12.2017 (pink) und am 13.12.2017 (orange)	48
Abbildung 8.18:	Tweets in NRW am 12.12.2017 (pink) und am 13.12.2017 (orange)	48
Abbildung 8.19:	Tweets in Berlin am 08.12.2017 (grün) und am 12.12.2017 (pink).....	48
Abbildung 8.20:	Tweets in NRW am 10.12.2017 (gelb) und am 13.12.2017 (orange)	48
Abbildung 9.1:	1. Vergleich: 09.12.2017 (blau) und 10.12.2017 (gelb)	49
Abbildung 9.2:	2. Vergleich: 09.12.2017 (gelb) und 10.12.2017 (blau)	49
Abbildung 9.3:	3. Vergleich: 10.12.2017 (blau) und 09.12.2017 (gelb)	50
Abbildung 9.4:	4. Vergleich: 10.12.2017 (gelb) und 09.12.2017 (blau)	50
Abbildung A.1:	Die technischen Spezifikationen des verwendeten Rechners	61
Abbildung A.2:	Die verwendete Version von R	61
Abbildung A.3:	Die verwendete Version von RStudio	62
Abbildung A.4:	Die verwendete Version von Rtweet	62
Abbildung A.5:	Die verwendete Version von Leaflet.....	63
Abbildung A.6:	Die verwendete Version von Shiny.....	64
Abbildung C.1:	Tweets in Berlin morgens am 07.12.2017	78
Abbildung C.2:	Tweets in Berlin mittags am 07.12.2017	78
Abbildung C.3:	Tweets in Berlin nachmittags am 07.12.2017	78
Abbildung C.4:	Tweets in Berlin abends am 07.12.2017	78

Abbildung C.5:	Tweets in Berlin morgens am 08.12.2017.....	79
Abbildung C.6:	Tweets in Berlin mittags am 08.12.2017	79
Abbildung C.7:	Tweets in Berlin nachmittags am 08.12.2017	79
Abbildung C.8:	Tweets in Berlin abends am 08.12.2017	79
Abbildung C.9:	Tweets in Berlin morgens am 09.12.2017.....	80
Abbildung C.10:	Tweets in Berlin mittags am 09.12.2017	80
Abbildung C.11:	Tweets in Berlin nachmittags am 09.12.2017	80
Abbildung C.12:	Tweets in Berlin abends am 09.12.2017	80
Abbildung C.13:	Tweets in Berlin morgens am 10.12.2017.....	81
Abbildung C.14:	Tweets in Berlin mittags am 10.12.2017	81
Abbildung C.15:	Tweets in Berlin nachmittags am 10.12.2017	81
Abbildung C.16:	Tweets in Berlin abends am 10.12.2017.....	81
Abbildung C.17:	Tweets in Berlin morgens am 11.12.2017	82
Abbildung C.18:	Tweets in Berlin mittags am 11.12.2017.....	82
Abbildung C.19:	Tweets in Berlin nachmittags am 11.12.2017	82
Abbildung C.20:	Tweets in Berlin abends am 11.12.2017	82
Abbildung C.21:	Tweets in Berlin morgens am 12.12.2017.....	83
Abbildung C.22:	Tweets in Berlin mittags am 12.12.2017	83
Abbildung C.23:	Tweets in Berlin nachmittags am 12.12.2017	83
Abbildung C.24:	Tweets in Berlin abends am 12.12.2017	83
Abbildung C.25:	Tweets in Berlin morgens am 13.12.2017.....	84
Abbildung C.26:	Tweets in Berlin mittags am 13.12.2017	84
Abbildung C.27:	Tweets in Berlin nachmittags am 13.12.2017	84
Abbildung C.28:	Tweets in Berlin abends am 13.12.2017	84
Abbildung D.1:	Tweets in NRW morgens am 07.12.2017	85
Abbildung D.2:	Tweets in NRW mittags am 07.12.2017.....	85
Abbildung D.3:	Tweets in NRW nachmittags am 07.12.2017.....	85
Abbildung D.4:	Tweets in NRW abends am 07.12.2017	85
Abbildung D.5:	Tweets in NRW morgens am 08.12.2017	86
Abbildung D.6:	Tweets in NRW mittags am 08.12.2017.....	86
Abbildung D.7:	Tweets in NRW nachmittags am 08.12.2017.....	86
Abbildung D.8:	Tweets in NRW abends am 08.12.2017	86
Abbildung D.9:	Tweets in NRW morgens am 09.12.2017	87
Abbildung D.10:	Tweets in NRW mittags am 09.12.2017.....	87
Abbildung D.11:	Tweets in NRW nachmittags am 09.12.2017.....	87
Abbildung D.12:	Tweets in NRW abends am 09.12.2017	87
Abbildung D.13:	Tweets in NRW morgens am 10.12.2017	88
Abbildung D.14:	Tweets in NRW mittags am 10.12.2017.....	88
Abbildung D.15:	Tweets in NRW nachmittags am 10.12.2017.....	88
Abbildung D.16:	Tweets in NRW abends am 10.12.2017	88
Abbildung D.17:	Tweets in NRW morgens am 11.12.2017	89
Abbildung D.18:	Tweets in NRW mittags am 11.12.2017	89
Abbildung D.19:	Tweets in NRW nachmittags am 11.12.2017.....	89
Abbildung D.20:	Tweets in NRW abends am 11.12.2017.....	89

Abbildung D.21:	Tweets in NRW morgens am 12.12.2017	90
Abbildung D.22:	Tweets in NRW mittags am 12.12.2017	90
Abbildung D.23:	Tweets in NRW nachmittags am 12.12.2017	90
Abbildung D.24:	Tweets in NRW abends am 12.12.2017	90
Abbildung D.25:	Tweets in NRW morgens am 13.12.2017	91
Abbildung D.26:	Tweets in NRW mittags am 13.12.2017	91
Abbildung D.27:	Tweets in NRW nachmittags am 13.12.2017	91
Abbildung D.28:	Tweets in NRW mittags am 13.12.2017	91

Tabellenverzeichnis

Tabelle 6.1: Twitter-Search-API versus Twitter-Streaming-API	22
Tabelle 8.1: Anzahl der Tweets pro Tageszeit vom 07.12.2017 bis 13.12.2017	42
Tabelle 8.2: Anzahl der Tweets pro Tag vom 07.12.2017 bis 13.12.2017	45

1 Einleitung

1.1 Relevanz des Themas

Der Begriff „Digitalisierung“ wird verwendet, um die Umstellung vom Analogen ins Digitale zu beschreiben. In dieser Bachelorarbeit wird die gesellschaftliche Digitalisierung als Umstellung gesellschaftlicher Bereiche vom Analogen ins Digitale betrachtet⁰¹.

„Deutschland und die Digitalisierung: Der digitale Nachzügler“ (Gillmann, 2017) und „Digitalisierung: Deutschland im Internet nur Mittelmaß“ (Zeit Online, 2015) sind Beispiele für Titel von Zeitungsartikeln, die das große Interesse der Öffentlichkeit am Themenfeld der Digitalisierung in Deutschland belegen. Selbst die Bundesregierung verpflichtet sich mit ihrer „Digitale[n] Agenda 2014 – 2017“ (BMW, BMI, & BMVI, 2014, S. Titel) zur Umsetzung von Zielen, die den Digitalisierungsprozess in Deutschland fördern.

Es gibt bereits viele jährliche Studien zum Stand der Digitalisierung, wie beispielsweise den „Deutschland-Index der Digitalisierung“ (Opiela, et al., 2017a, S. Titel). Diese Studien betrachten verschiedene Bereiche der Digitalisierung. Kaum eine der Studien untersucht jedoch einen so kleinen Unterbereich, wie die Verwendung eines einzelnen sozialen Mediums, wie beispielsweise Twitter. An welchen geografischen Orten in Deutschland die Digitalisierung stattfindet, wird ebenfalls von keiner dieser Studien untersucht. Daher ist dies ein interessanter Ansatzpunkt für die wissenschaftliche Forschung.

Darüber hinaus werden Daten von Twitter bereits in vielen wissenschaftlichen Forschungen angewandt und publiziert (vgl. Rahlf & Weller, 2014, S. 142). So gibt es beispielsweise die Webanwendung „Tweetping“ (SAS Lightstream, 2017) der Firma SAS Lightstream, die weltweit Tweets in Echtzeit auf einer Karte visualisiert, oder auch die mit R und Shiny erstellte Anwendung „Tweet Analyzer“ (Adoptitude Analytics, 2017), die alle Emotionen aus den Tweets zu einem Twitter-Nutzer bewertet und darstellt. Auch Titel wie „Twitter Data Analytics“ (Kumar, Morstatter & Liu, 2014, S. Titel), oder „Twitter-Archive und die Herausforderungen von »Big Social Data« für die Medien- und Kommunikationswissenschaft“ (Burgess & Bruns, 2014, S. 191) lassen erkennen, dass Daten von Twitter in der Forschung verwendet werden. Bisher gibt es noch keine Visualisierung der geografischen Daten von Twitter, mit Bezug zur Digitalisierung, die ausschließlich aus Deutschland stammen. Geografische Twitter-Daten aus Deutschland zu nutzen, um zu zeigen, wo Digitalisierung in Deutschland stattfindet, ist demnach ein Punkt, den es noch zu untersuchen gilt.

Über dies wird in dieser Arbeit die Programmiersprache R verwendet. Diese ist ursprünglich für Grafiken und Statistiken entwickelt worden (vgl. The R Foundation, 2017). Die Nutzung des R-Packages „Shiny“, für die Umsetzung einer derartigen Visualisierung, ermöglicht es die Funktionen von R als Webanwendung bereitzustellen (vgl. RStudio, Inc., 2017m). Dadurch kann die erzeugte Anwendung auch ohne Kenntnisse in R für weitere Forschungen verwendet werden. So bereichert sie die Wissenschaft um ein Werkzeug zur Visualisierung der Digitalisierung anhand von Twitter-Nachrichten in Deutschland. Dieses kann dann für unterschiedlichste Fragestellungen in der Forschung zum The-

⁰¹ Siehe dazu Kapitel 2.2.2 und Kapitel 2.2.3.

menfeld der Digitalisierung in Deutschland hinzugezogen werden. Beispielsweise können so Fragestellungen, wie „Zu welcher Tageszeit gibt es die meisten Tweets?“, also „Wann ist die Digitalisierung täglich am stärksten?“, oder „An welchem Wochentag gibt es die meisten Tweets und ist somit die Digitalisierung am stärksten?“, beantwortet werden.

1.2 Zielsetzung

Ziel der Bachelorarbeit ist es, eine Web-Anwendung (App) mittels R und Shiny zu schaffen, die es Nutzern ermöglicht, den Ort der (gesellschaftlichen) Digitalisierung zu einem von Ihnen bestimmten Zeitraum am Beispiel von Twitter-Nachrichten in Deutschland beziehungsweise in den einzelnen Bundesländern zu visualisieren und mit anderen selbst gewählten Zeiträumen visuell zu vergleichen.

Überdies wird außerdem geklärt, was die Digitalisierung ist, wie der Stand der Digitalisierung aktuell in Deutschland ist und wo Twitter im Rahmen der Digitalisierung zu verorten ist. Die Visualisierungsform „geografische Informationsgrafik“, welche die gewählt Darstellungsform für die Verortung der Twitter-Nachrichten in Deutschland ist, wird ebenfalls vorgestellt. Auch wird ein Überblick über R und Shiny und deren Möglichkeiten, im Allgemeinen und speziell bezogen auf das Thema dieser Arbeit, gegeben.

Es ist jedoch kein Ziel dieser Arbeit, den Grad der Digitalisierung in Deutschland oder die prozentuale Verteilung der Twitter-Nachrichten in Abhängigkeit zur Einwohnerzahl der Bundesländer oder ähnliches zu visualisieren. Auch sollen alle Twitter-Nachrichten, die innerhalb Deutschlands veröffentlicht werden, berücksichtigt werden und nicht nach bestimmten Inhalten gefiltert werden.

1.3 Vorgehensweise

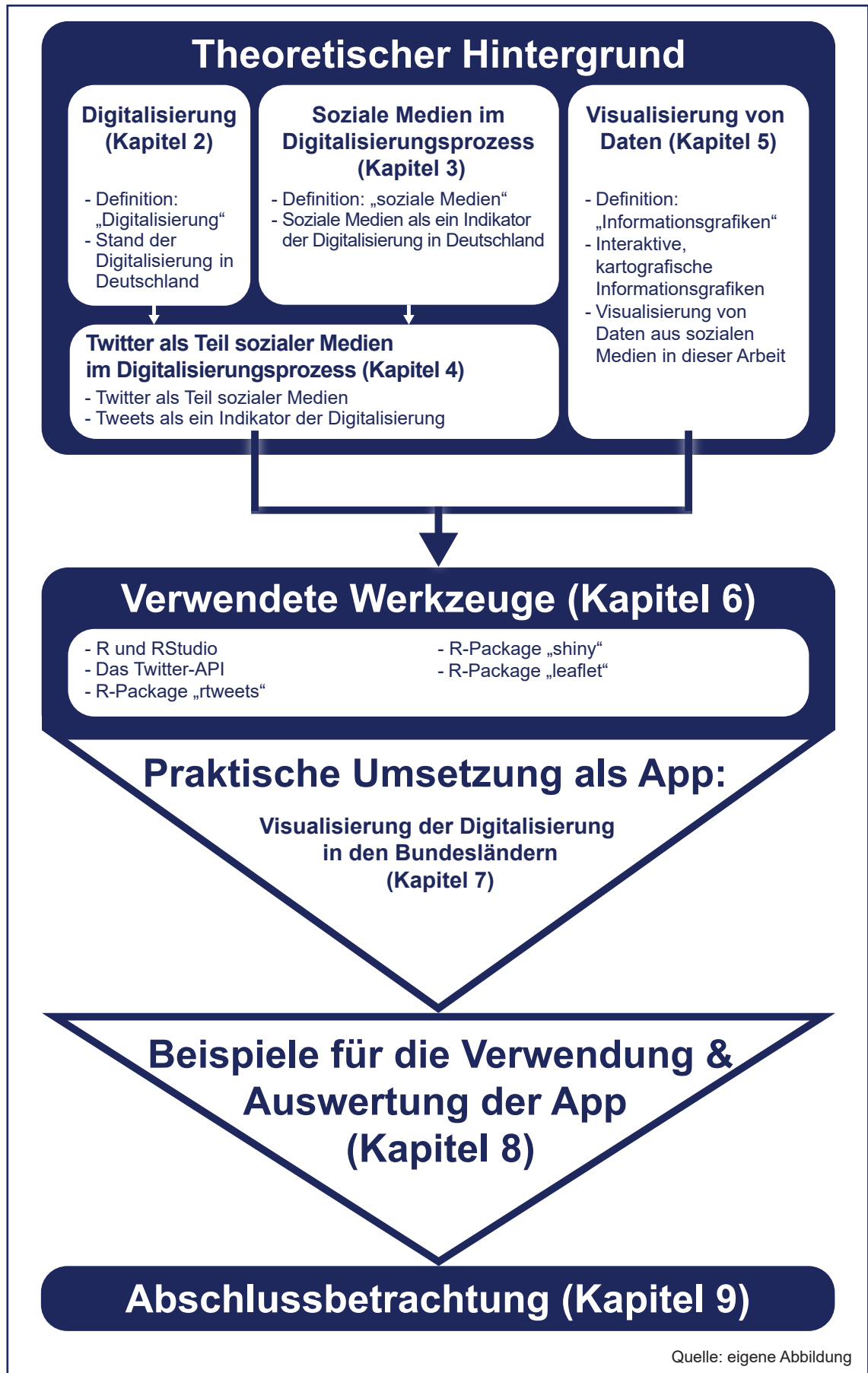


Abbildung 1.1: Übersicht über die Struktur dieser Arbeit

Zur Umsetzung der Zielsetzung dieser Arbeit wird, wie in Abbildung 1.1 zu sehen, zunächst der theoretische Hintergrund⁰² erläutert. Dann werden die zur Visualisierung verwendeten Werkzeuge⁰³ vorgestellt und die praktische Umsetzung der App⁰⁴ erklärt. Anschließend wird die Verwendung und Auswertung der App an Hand zweier Beispiele⁰⁵ dargestellt. Zuletzt erfolgt dann eine Abschlussbetrachtung⁰⁶.

Der theoretische Hintergrund besteht aus vier Teilen. Der erste Teil ist die Digitalisierung und beginnt mit einer Definition des Begriffs der Digitalisierung. Danach wird der Stand der Digitalisierung in Deutschland anhand von Studien, sowie der „Digitale[n] Agenda 2014 - 2017“ (BMW, BMI, & BMVI, 2014, S. Titel) der deutschen Bundesregierung, präsentiert.

Der zweite Teil des theoretischen Hintergrundes behandelt die sozialen Medien im Digitalisierungsprozess, in dem erst der Begriff der sozialen Medien definiert und anschließend aufgezeigt wird, dass diese ein Indikator für die Digitalisierung in Deutschland sind.

Diese beiden Teile führen dann zum dritten Teil, der Twitter als Teil sozialer Medien im Digitalisierungsprozess verortet. Dazu wird Twitter in den Bereich soziale Medien eingeordnet und Twitter-Nachrichten als ein Indikator für die Digitalisierung in Deutschland betrachtet.

Außerdem beschreibt der vierte Teil des theoretischen Hintergrundes die Visualisierung von Daten. Dazu erfolgt eine Definition des Begriffs „Informationsgrafiken“, interaktive, kartografische Informationsgrafiken werden konkretisiert und die Visualisierung von Daten aus sozialen Medien wird angesprochen.

Die für die Visualisierung verwendeten Werkzeuge, wie die Programmiersprache R und die Entwicklungsumgebung RStudio, werden in Kapitel 6 beschrieben. Außerdem gehören das Application Programming Interface (API) von Twitter und die R-Packages „rtweets“, „shiny“ und „leaflet“ ebenfalls zu den verwendeten Werkzeugen, während die Skripte der App in Kapitel 7 dokumentiert werden.

Danach wird die Verwendung und Auswertung der App an zwei beispielhaften Fragestellungen präsentiert. Dabei ist eine Fragestellung, zu welcher Tageszeit die Digitalisierung am stärksten ist, also wann und wo die meisten Tweets veröffentlicht worden sind. Die andere lautet, an welchem Tag innerhalb eines Zeitraumes die Digitalisierung am stärksten ist.

In der Abschlussbetrachtung wird die Arbeit kritisch betrachtet. Außerdem wird ein Ausblick für weitere Forschungen und ein Fazit gegeben.

02 Siehe dazu Kapitel 2 bis 5.

03 Siehe dazu Kapitel 6.

04 Siehe dazu Kapitel 7.

05 Siehe dazu Kapitel 8.

06 Siehe dazu Kapitel 9.

2 Digitalisierung

2.1 Definition „Digitalisierung“

Digitalisierung ist ein Begriff, der verschiedene Bedeutungen besitzen kann (vgl. Bendel, 2015). Es kann zum Beispiel das Umwandeln eines analogen in ein digitales Signal gemeint sein, oder auch das digitale Darstellen von Daten und Informationen (vgl. Dudenredaktion, 2015a).

Darüber hinaus vermischen sich, wie in Abbildung 2.1 zu sehen, „aktuelle IT-Megatrends wie Industrie 4.0, Big Data, Cloud-Computing und Social Web (...) zum globalen Megatrend Digitalisierung“ (Gadatsch & Mangiapane, 2017, S. 1). Dies sind daher alles Begriffe, die mit dem Begriff „Digitalisierung“ in Verbindung stehen.

Mit „Industrie 4.0“ ist die vierte industrielle Revolution gemeint. Sie bezeichnet die schnelle Veränderung der Gesellschaft, „ausgelöst durch die Automation und Digitalisierung, welche sich wechselseitig noch verstärken“ (Wittpahl, 2017, S. 5). Es entsteht eine „Vernetzung von Menschen, Computern und Werkstücken“ (Gadatsch & Mangiapane, 2017, S. 1).

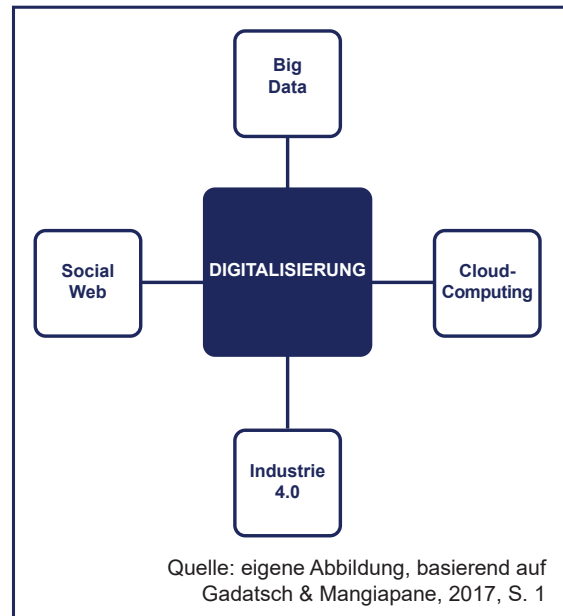


Abbildung 2.1: Digitalisierung als globaler Megatrend

Im 20. Jahrhundert wird der Fokus in der vierten industriellen Revolution auf „Automatisierung und Optimierung“ (Bendel, 2015) durch Computersysteme und -netze gesetzt, während im 21. Jahrhundert bereits begonnen wird „disruptive Technologien und innovative Geschäftsmodelle sowie Autonomisierung, Flexibilisierung und Individualisierung in der Digitalisierung“ (Bendel, 2015) anzuvizieren. Demnach ist die Digitalisierung ein Prozess, der die vierte Revolution durch den Einsatz von Computern in allen Lebensbereichen vorantreibt.

Die Digitalisierung betrifft viele Lebensbereiche. Manche „Branchen, wie die Schreibmaschinenindustrie, [sind] vollständig ausgelöscht und andere massiv umstrukturiert“ (Wittpahl, 2017, S. 5) worden. Die öffentliche Verwaltung betreibt die Digitalisierung unter dem Begriff eGovernment und Dienstleistungen werden heutzutage oft auch digital angeboten (vgl. Wittpahl, 2017, S. 5).

Auch die digitale Vernetzung, die einen schnellen Informationsaustausch bietet, ist ein wichtiger Teil der Digitalisierung. Heute ist „die Teilnahme am digitalen Netzwerk auch für eine breite Bevölkerung ohne IT-Kenntnisse möglich“ (Kollmann & Schmidt, 2016, S. 3). So kann jeder, der am digitalen Netzwerk, dem Internet, angeschlossen ist, Informationen austauschen, miteinander kommunizieren oder beispielsweise Transaktionen ausführen (vgl. Kollmann & Schmidt, 2016, S. 3).

Das digitale Netz ermöglicht unter anderem die Nutzung einer „netz-basierte[n] aktive[n] N:M Kommunikation“ (Gadatsch & Mangiapane, 2017, S. 1), das „Social Web“, beziehungsweise die „Sozialen Medien“⁰⁷ (vgl. Schmidt, 2017, S. 16).

Auch ermöglicht die digitale Vernetzung das sogenannte „Cloud-Computing“, welches die „flexible netzbasierte Informationsverarbeitung“ (Gadatsch & Mangiapane, 2017, S. 1) bezeichnet. Hierbei werden die Daten nicht auf der lokalen Festplatte „sondern in einem entfernten Rechenzentrum gespeichert“ (BPA, 2015, S. 8). Dies ermöglicht einen Zugriff von jedem beliebigen Ort aus (vgl. BPA, 2015, S.8).

Im Rahmen der Digitalisierung und der digitalen Vernetzung fallen viele Daten an. Für die „nahezu Echtzeit-Verarbeitung sehr großer (...) Datenmengen“ (Gadatsch & Mangiapane, 2017, S. 1) ist „Big Data“ ein Schlagwort in der Digitalisierung.

Zusammenfassung:

Der Begriff „Digitalisierung“ umfasst viele Bereiche der Gesellschaft und beschreibt den Prozess der Umstellung dieser Bereiche vom Analogen ins Digitale. Aus diesem Grund wird die Digitalisierung in den gesellschaftlichen Bereichen in dieser Arbeit auch als gesellschaftliche Digitalisierung bezeichnet.

2.2 Stand der Digitalisierung in Deutschland

Der Digitalisierungsprozess findet auch in Deutschland statt, wie einige Studien, sowie die Bemühungen der deutschen Bundesregierung zur Umsetzung der „Digitale[n] Agenda 2014 - 2017“ (BMWi, BMI, & BMVI, 2014, S. Titel), zeigen.

Studien zur Digitalisierung in Deutschland werden sowohl auf der europäischen Ebene, als auch auf nationaler Ebene durchgeführt. Die Europäische Union gibt beispielsweise jährlich den „Bericht über den Stand der Digitalisierung in Europa (Europe’s Digital Progress Report, EDPR)“⁰⁸ (European Union, 2017a, S. 1) und die „Ermittlung des Index für die digitale Wirtschaft und Gesellschaft DESI (Digital Economy and Society Index)“⁰⁹ (European Union, 2017a, S. 1) bekannt. In Deutschland gibt es dagegen jährlich die „ARD/ZDF-Onlinestudie“¹⁰ (ARD/ZDF-Medienkommission, 2017, S. 2) von ARD und ZDF, sowie den „Deutschland-Index der Digitalisierung“¹¹ (Opiela, et al., 2017a, S. Titel), vom Kompetenzzentrum Öffentliche IT (ÖFIT) und der „D21-Digital-Index“¹² (Initiative D21 e. V., 2016, S. Titel) der Initiative D21 e. V., die ebenfalls jährlich herausgegeben werden.

Darüber hinaus existieren noch weitere Studien, wie beispielsweise der Report „Digital in 2017“ (Himmelberg, 2017) der Agentur „We Are Social“ oder die Studie „Das digitale Wirtschaftswunder - Wunsch oder Wirklichkeit?“ (Windhagen, et al., 2017, S. Titel) des „McKinsey Global Institute“. Diese werden in dieser Arbeit jedoch nicht näher betrachtet werden. Ebenfalls sollen auch die Bemühungen der Landesregierungen, wie die der Landesregierung Nordrhein-Westfalen (2015), hier nur erwähnt werden.

⁰⁷ Siehe dazu auch Kapitel 3.

⁰⁸ Zuletzt herausgegeben für das Jahr 2017.

⁰⁹ Zuletzt herausgegeben für das Jahr 2017.

¹⁰ Zuletzt herausgegeben für das Jahr 2017.

¹¹ Zuletzt herausgegeben für das Jahr 2017.

¹² Zuletzt herausgegeben für das Jahr 2016.

2.2.1 Die digitale Agenda der deutschen Bundesregierung

Die deutsche Bundesregierung möchte mit der „Digitale[n] Agenda 2014 - 2017“ (BMW, BMI, & BMVI, 2014, S. Titel) Ziele definieren, um die „Chancen der Digitalisierung [zu] nutzen“ (BMW, BMI, & BMVI, 2014, S. 2). Dazu werden „sieben Handlungsbereiche“ (BPA, 2015, S. 1) definiert.

Im ersten Handlungsbereich werden Ziele für den Ausbau der digitalen Infrastruktur, sowohl durch den Breitbandausbau, als auch durch die Förderung der mobilen Nutzung, beschrieben (vgl. BMW, BMI, & BMVI, 2014, S. 8ff).

Ziele für die Unterstützung und das Vorantreiben von „digitale[r] Wirtschaft und digitale[m] Arbeiten“ (BMW, BMI, & BMVI, 2014, S. 12) werden im zweiten Handlungsbereich dargelegt. Dagegen ist der dritte Handlungsbereich als „Innovativer Staat“ (BMW, BMI, & BMVI, 2014, S. 18) betitelt. Hierbei handelt es sich um die Digitalisierung der Verwaltung und ihrer Prozesse, so dass beispielsweise „Behördengänge (...) ganz bequem am heimischen Computer“ (BPA, 2015, S. 12) erledigt werden können.

„Digitale Lebenswelten in der Gesellschaft gestalten“ (BMW, BMI, & BMVI, 2014, S. 22) ist der vierte Handlungsbereich. In diesem Bereich wird die Förderung der Teilnahme an digitalen Angeboten des Alltags aller gesellschaftlichen Gruppen thematisiert (vgl. BMW, BMI, & BMVI, 2014, S. 23). Dieser Bereich berücksichtigt die Barrierefreiheit, zum Beispiel im Internet, um unter anderem auch Menschen mit Behinderung die Teilhabe zu ermöglichen (BPA, 2015, S. 14f.)

Der Bereich „Bildung, Forschung, Wissenschaft, Kultur und Medien“ (BMW, BMI, & BMVI, 2014, S. 26) wird als Handlungsbereich fünf von der Bundesregierung im Rahmen der Digitalisierung festgelegt. Hierbei soll der digitale Wandel in der Bildung, der Wissenschaft und der Forschung voran getrieben werden (vgl. BMW, BMI, & BMVI, 2014, S. 27). Ebenso sollen „Innovationspotenziale der Digitalisierung“ (BMW, BMI, & BMVI, 2014, S. 28) in einer „durchgängigen Wertschöpfungskette“ (BMW, BMI, & BMVI, 2014, S. 28) für Markterfolge genutzt und digitale Technologien und das Internet für Kulturgüter und mediale Inhalte verwendet werden (vgl. BMW, BMI, & BMVI, 2014, S. 28).

Der sechste Handlungsbereich behandelt die Themen „Sicherheit, Schutz und Vertrauen für Gesellschaft und Wirtschaft“ (BMW, BMI, & BMVI, 2014, S. 30). In diesem Bereich wird vor allem der Datenschutz und „sichere digitale Infrastrukturen“ (BMW, BMI, & BMVI, 2014, S. 32) aufgegriffen.

Im siebten und letzten Handlungsbereich „Europäische und internationale Dimension der Digitalen Agenda“ (BMW, BMI, & BMVI, 2014, S. 34) wird beschrieben, dass viele Themen, wie der Netzausbau oder das „Völkerrecht des Netzes“ (BMW, BMI, & BMVI, 2014, S. 36), auf europäischer Ebene beziehungsweise international besprochen werden müssen. Deshalb verpflichtet sich die Bundesregierung, an entsprechenden Diskussionen aktiv teilzunehmen (vgl. BMW, BMI, & BMVI, 2014, S. 35f.).

Diese sieben Handlungsbereiche der deutschen Bundesregierung zur Förderung der Digitalisierung zeigen, dass in Deutschland in allen gesellschaftlichen Lebensbereichen der Digitalisierungsprozess stattfindet und voranschreiten soll.

2.2.2 Studien der europäischen Union zur Digitalisierung in Deutschland

Im EDPR wird die quantitative Erhebung des DESI „mit qualitativen Informationen zur Politik des jeweiligen Landes verknüpft“ (European Union, 2017a, S. 1). Dazu werden fünf Bereiche untersucht. Diese lauten Konnektivität, Humankapital, Internetnutzung, Integration der Digitaltechnik und Digitale öffentliche Dienste (vgl. European Union, 2017a, S. 1 und 2017c, S. 1). Dem DESI zur Folge befindet sich Deutschland auf Platz 11 der 29 untersuchten europäischen Länder (vgl. European Union, 2017a, S. 1 und 2017c, S. 1). Das zeigt, dass sich Deutschland im europäischen Vergleich aktuell im vorderen Mittelfeld in der Umsetzung der Digitalisierung befindet.

Wie in Abbildung 2.2 zu sehen ist, liegt Deutschland mit seinem DESI-Wert von 2017 über dem europäischen Durchschnitt, erreicht jedoch nie den höchsten gemessenen Wert.

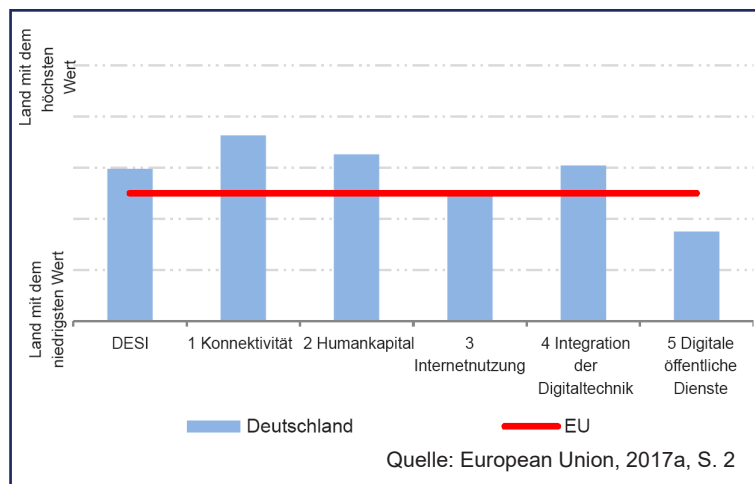


Abbildung 2.2: DESI 2017 - relative Leistung nach Bereichen

Im Bereich der „Konnektivität“, welcher „Festnetzbreitband, Mobilfunkbreitband, Breitbandgeschwindigkeit und -preise“ (European Union, 2017a, S. 1) umfasst, liegt in Deutschland weit über dem Durchschnitt von Europa. Im Bereich „Humankapital“, das sich auf die digitalen Kompetenzen der Einwohner bezieht (vgl. European Union, 2017a, S. 1) und im Bereich der „Integration der Digitaltechnik“, der den „Digitalisierungsgrad der Wirtschaft, [sowie den] Internethandel“ (European Union, 2017a, S. 1) umfasst, liegt Deutschland ebenfalls über dem Durchschnitt.

Währenddessen wird dagegen im Bereich „Internetnutzung“, der „Nutzung von Inhalten, Kommunikation und Online-Transaktionen durch Bürgerinnen und Bürger“ (European Union, 2017a, S. 1) erfasst, der europäische Durchschnitt erreicht, aber nicht überschritten. Am schlechtesten schneidet Deutschland im Bereich „Digitale öffentliche Dienste“, der das „eGovernment (elektronische Behördendienste)“ (European Union, 2017a, S. 1) einschließt, ab. Hier wird der europäische Durchschnitt deutlich unterschritten.

Zusammenfassung:

Deutschland befindet sich im vorderen Mittelfeld bei der Umsetzung der Digitalisierung im europäischen Vergleich. Vor allem der Bereich „Digitale öffentliche Dienste“ ist noch ausbaufähig. Insgesamt zeigt diese Studie, dass die Digitalisierung bereits viele, wenn nicht sogar alle, Lebensbereiche der Menschen in Europa und somit auch in Deutschland erreicht hat.

2.2.3 Deutsche Studien zum Stand der Digitalisierung

In Deutschland selbst werden auch einige Studien zum Stand der Digitalisierung in Deutschland erhoben. So gibt es unter anderem zum Beispiel die Studie „Deutschland-Index der Digitalisierung“¹³ (Opiela, et al., 2017a, S. Titel) oder die „ARD/ZDF-Onlinestudie“ (ARD/ZDF-Medienkommission, 2017, S. 1).

In die Studie „Deutschland-Index der Digitalisierung“ (Opiela, et al., 2017a, S. Titel) des Kompetenzzentrums Öffentliche IT (ÖFIT) fließen Ergebnisse mehrerer anderer Studien mit ein und ergeben gemeinsam die Werte für den Stand der Digitalisierung in Deutschland (vgl. Opiela, et al., 2017a, S. 5). So nutzt das ÖFIT Ergebnisse des von Ihnen ebenfalls herausgegebenen „ÖFIT-Atlas der Digitalisierung“ (Wiegand, et al., 2015, S. Titel), sowie „Detailanalysen zum E-Government in Deutschland“ (Opiela, et al., 2017a, S. 5.).

Der Deutschland-Index setzt sich für das Jahr 2017 aus fünf Bereichen, den Themenfeldern, zusammen. Diese Felder lauten „Infrastruktur“, „Digitales Leben“, „Wirtschaft und Forschung“, „Bürgerservices“ und „digitale Kommune“ (vgl. Opiela, et al., 2017a, S. 8).

Das erste Themenfeld des Deutschland-Indexes, „Infrastruktur und Versorgung“ (Opiela, et al., 2017a, S. 10), setzt sich aus den Punkten „Teilnetze und Domains“, „Mobilfunkabdeckung“, „Technologievielfalt“ und „Breitbandversorgung“ zusammen (vgl. Opiela, et al., 2017a, S. 10). Bei der Breitbandversorgung gibt es in Deutschland abhängig vom Standort Unterschiede. „Flächenländer, insbesondere in Ostdeutschland, sind im Mittel schlechter versorgt“ (Opiela, et al., 2017a, S. 11) als in den Stadtstaaten oder Ballungszentren. Dies spiegelt sich auch in der „Wahlfreiheit zwischen unterschiedlichen technologischen Lösungen“ (Opiela, et al., 2017a, S. 12) wieder. Hierbei fällt jedoch auf, dass die Flächenländer „Schleswig-Holstein und das Saarland sowie Bayern und Nordrhein-Westfalen“ (Opiela, et al., 2017a, S. 12) im Vergleich eine große Vielfalt an Technologien bieten. Insgesamt erreichen die Stadtstaaten direkt gefolgt von den „Flächenländer[n] Schleswig-Holstein, Hessen und Nordrhein-Westfalen sowie das Saarland“ (Opiela, et al., 2017a, S. 13) die besten Werte¹⁴ im ersten Themenfeld des Deutschland-Indexes. Im Durchschnitt erreicht Deutschland in diesem Themenfeld einen Gesamtwert von 70,8 Punkten¹⁵ (vgl. Opiela, et al., 2017a, S. 40).

Die Punktzahl der „Chaos-Computer-Club-Treffen“¹⁶, „Nutzung mobiler Endgeräte“, „Online-Shopping“, „FabLabs“¹⁷, „Nutzung sozialer Medien“, „Wikipedia-Artikel“ pro Einwohner und „tägliche Internetnutzung“ fallen in das zweite Themenfeld (vgl. Opiela, et al., 2017a, S. 14). In diesem Themenfeld unterscheiden sich die Bundesländer nicht so stark. „Beispielsweise liegen die Bundesländer mit der höchsten (62 Prozent) und der geringsten (46 Prozent) Nutzung sozialer Medien nur 16 Prozentpunkte auseinander“ (Opiela, et al., 2017a, S. 14). Dennoch führen auch hier wieder die Stadtstaaten¹⁸ (vgl. Opiela, et al.,

¹³ Im Folgenden Deutschland-Index genannt.

¹⁴ Hamburg erreicht beispielsweise eine 70 prozentige Abdeckung mit der Glasfaser-Verteilung, während in Bayern nur 10 Prozent abgedeckt sind (vgl. Opiela, et al., 2017a, S. 13).

¹⁵ Es können jeweils maximal 100 Punkte erreicht werden (vgl. Opiela, et al., 2017a, S. 13).

¹⁶ „Der Chaos Computer Club (CCC) ist die größte europäische Hackervereinigung und beschäftigt sich mit technischen und sozialen Entwicklungen“ (Opiela, et al., 2017a, S. 16).

¹⁷ „Bei FabLabs handelt es sich um Werkstätten, in denen Privatpersonen (computergesteuert) Produkte fertigen können“ (Opiela, et al., 2017a, S. 16).

¹⁸ „Bremen [erreicht] den Höchstwert von 98,5“ (Opiela, et al., 2017a, S. 16) Punkten.

2017a, S. 16). Der Durchschnitt von Deutschland liegt hier bei 85,3 Punkten (vgl. Opiela, et al., 2017a, S. 40).

Das Themenfeld „Wirtschaft und Forschung“ ist das dritte Thema und bewertet die Wettbewerbsfähigkeit der Bundesländer. Hier fließen die Punkte „IKT-Beschäftigte“¹⁹, „Auszubildende in IKT-Berufen“, „Fachkräftemangel im IKT-Bereich“, „Verdienst in der IuK“²⁰, „IuK-Betriebe“, „IKT-Forschungsförderung“, „IT-Neugründungen“ und „IuK-Patente“²¹ (vgl. Opiela, et al., 2017a, S. 17) in die Bewertung ein. In diesem Bereich spiegeln die erreichten Werte der einzelnen Bundesländer die „Zahlen zur Bruttowertschöpfung pro Kopf der Länder“ (Opiela, et al., 2017a, S. 19) wieder und Deutschland erreicht 55,9 Punkte (vgl. Opiela, et al., 2017a, S. 40).

Die „Anwendungen für den neuen Personalausweis“, das „E-Government-Gesetz auf Landesebene“, die „Teilnahme an GovData“, „erfolgreiche Informationsfreiheitsanfragen“ und die „elektronische Übermittlung von Formularen“ (vgl. Opiela, et al., 2017a, S. 20) werden im vierten Themenfeld, „Bürgerservices“ (Opiela, et al., 2017a, S. 20) bewertet. Dieser Bereich, bei dem Deutschland 68,7 Punkte erreicht (vgl. Opiela, et al., 2017a, S. 40), wird von „Rheinland-Pfalz, Bayern und Nordrhein-Westfalen“ (Opiela, et al., 2017a, S. 23) angeführt.

Das letzte Themenfeld, „Digitale Kommune“ (Opiela, et al., 2017a, S. 8) beziehungsweise „Kommunale Webportale“ (Opiela, et al., 2017a, S. 24) bewertet die „Verbindlichkeit“, die „Zusammenarbeit“, den „Nutzen“, die „Basiskomponenten“, den „Zugang“, die „Benutzbarkeit“ und die „Offenheit“ kommunaler Webportale (vgl. Opiela, et al., 2017a, S. 24). In diesem Bereich führen ebenfalls wieder die Stadtstaaten (vgl. Opiela, et al., 2017a, S. 30). Nordrhein-Westfalen ist das beste Flächenland und erreicht als einziges über 40 Punkte (vgl. Opiela, et al., 2017a, S. 30). Deutschland hat in diesem Bereich einen Wert von 38 Punkte erreicht (vgl. Opiela, et al., 2017a, S. 40), welcher demnach der Bereich mit den geringsten Punkten ist.

Das Gesamtergebnis des Deutschland-Indexes beträgt 62,3 Punkte für Deutschland (vgl. Opiela, et al., 2017a, S. 41), wie in Abbildung 2.3 zu sehen. Die besten Ergebnisse erzielen Hamburg und Berlin, danach folgen „Bremen, Hessen, Bayern und Nordrhein-Westfalen“ (Opiela, et al., 2017a, S. 41). In diese Berechnungen des Ergebnisses fließen die Themenfelder „In-

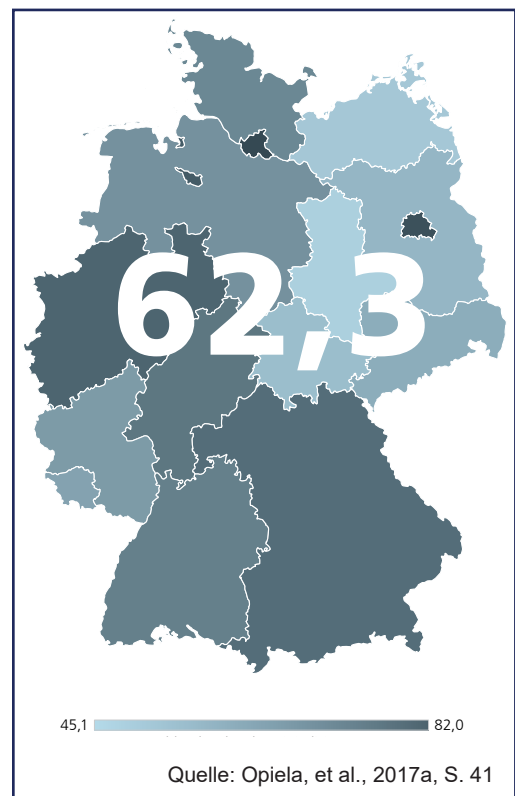


Abbildung 2.3: Deutschland-Index der Digitalisierung 2017

¹⁹ Beschäftigte in der Branche der „Informations- und Kommunikationstechnik (IKT)“ (Opiela, et al., 2017b, S. 5).

²⁰ „Durchschnittbruttomonatsverdienst [sic] in der Informations- und Kommunikationsbranche [(IuK)]“ (Opiela, et al., 2017b, S. 6).

²¹ „Anzahl der angemeldeten Patente im Bereich Information und Kommunikation gewichtet nach der Anzahl der Bevölkerung“ (Opiela, et al., 2017b, S. 6).

„Infrastruktur“ zu 25 Prozent, „Digitales Leben“ zu 20 Prozent, „Wirtschaft und Forschung“ ebenfalls zu 20 Prozent, „Bürgerservices“ zu 10 Prozent und „digitale Kommune“ zu 25 Prozent ein (vgl. Opiela, et al., 2017a, S. 39).

Im Gegensatz zum Deutschland-Index, der viele Bereiche der Digitalisierung untersucht, wird der Fokus in der Studie von ARD und ZDF ausschließlich auf die „Daten zur Internetnutzung in Deutschland“ (Frees & Koch, 2017, S. 434) gelegt. So hat sie zum Beispiel herausgefunden, dass die Deutschen, die zu 90 Prozent einen Internetzugang haben (vgl. ARD/ZDF-Medienkommission, 2017, S. 2), das Internet im Durchschnitt ca. zwei Stunden und 29 Minuten täglich nutzen (vgl. ARD/ZDF-Medienkommission, 2017, S. 3). Dabei werden 45 Minuten täglich für die „mediale Internetnutzung“ (ARD/ZDF-Medienkommission, 2017, S. 4), 59 Minuten für die „Individual-Kommunikation“ (ARD/ZDF-Medienkommission, 2017, S. 4) und eine Stunde 11 Minuten für die „sonstige Internetnutzung (Surfen, Shoppen, Spielen)“ (ARD/ZDF-Medienkommission, 2017, S. 4) verwendet²². Außerdem nutzen mittlerweile 30 Prozent der Befragten das Internet auch mobil (ARD/ZDF-Medienkommission, 2017, S. 7).

Eine weitere Studie ist die Studie „D21-Digital-Index“ (Initiative D21 e. V., 2016, S. Titel) der Initiative D21 e. V., die sich in „die vier unterschiedlich gewichteten Dimensionen Zugang, Nutzung, Kompetenz und Offenheit“ (Initiative D21 e. V., 2016, S. 7)²³ gliedert. Demzufolge nutzen 2016 „79 Prozent der deutschen Bevölkerung“ (Initiative D21 e. V., 2016, S. 8) einen Zugang zum Internet. Dennoch „sinkt [der Index] von 52 auf 51 Punkte (Skala von 0 bis 100)“ (Initiative D21 e. V., 2016, S. 9), da die „Kompetenz und der Offenheit“ (Initiative D21 e. V., 2016, S. 9) in Bezug auf die „Anforderungen der Digitalisierung“ (Initiative D21 e. V., 2016, S. 9) schlechtere Werte als im Jahr zuvor aufweisen.

Zusammenfassung:

Wie die Studie zum Deutschland-Index zeigt, ist die Digitalisierung in den einzelnen Bundesländern unterschiedlich stark vorangeschritten. Dennoch stellt diese Studie, ebenfalls dar, dass die Digitalisierung im Alltag der Deutschen stattfindet. Dies wird auch von der Studie von ARD und ZDF unterstützt, die darauf hinweist, dass die Deutschen das Internet nicht nur für Medien oder Aktivitäten, wie Einkäufe tätigen oder Spielen, nutzen, sondern etwas mehr als ein Drittel der Online-Zeit für die Kommunikation genutzt werden. Jedoch sinkt, laut der Studie „D21-Digital-Index“ (Initiative D21 e. V., 2016, S. Titel), die Offenheit und Kompetenz für die „Anforderungen der Digitalisierung“ (Initiative D21 e. V., 2016, S. 9) bei den Deutschen 2016, im Verhältnis zum Jahr zuvor.

²² Zu den einzelnen Bereichen der Internet-Nutzung sind noch weitere, vertiefende Umfragen von ARD und ZDF vorgenommen worden (ARD/ZDF-Medienkommission, 2017, S. 5ff), die hier zu weit führen würden.

²³ Eine Vorstellung der vertiefenden Ergebnisse würde in dieser Arbeit zu weit führen und wird daher nur zusammengefasst erläutert.

3 Soziale Medien im Digitalisierungsprozess

3.1 Definition „soziale Medien“

Der Begriff „soziale Medien“²⁴ (engl. Social Media) verliert „zunehmend an Klarheit“ (Scheffler, 2014, S. 13). Dabei gehören viele digitale Medien und Technologien zu den sozialen Medien (vgl. Zahn, 2013, S. 9).

Soziale Medien gelten als eine Fortsetzung in der Medienentwicklung, da auch sie auf der „medien- und informationstechnologischen Infrastruktur“ (Schmidt, 2017, S. 11) basieren. Dies gilt, weil das Internet, mit dem soziale Medien genutzt werden können, ebenfalls zu den Massenmedien, wie beispielsweise das Radio oder Zeitungen, gehört (vgl. Schmidt, 2017, S. 10f). Zusätzlich bieten soziale Medien jedoch auch Möglichkeit zur Veröffentlichung und Bearbeitung von Inhalten und zum Austausch mit anderen Menschen (vgl. Schmidt, 2017, S. 11). Dadurch schaffen sie einen Raum zwischen „der massenmedialen und der interpersonalen Kommunikation“ (Schmidt, 2017, S. 11), da sie sowohl zu den Massenmedien gehören, als auch Möglichkeiten zur Kommunikation und Interaktion zwischen Menschen bieten. So ist es zum Beispiel möglich, „Informationen, Meinungen, Eindruck[e] und Erfahrungen“ (Zahn, 2013, S. 9) auszutauschen oder gemeinsam neue Inhalte zu erstellen. Dadurch verschwimmt in sozialen Medien die „Grenze zwischen Produzent und Konsument“ (vgl. Zahn, 2013, S. 9).

Zu den sozialen Medien gehören unterschiedliche Formen von Plattformen, die meist unternehmensunabhängig, wie beispielsweise Facebook, sind (vgl. Scheffler, 2014, S. 14). Jedoch gibt es sie auch als unternehmenseigene Plattformen (vgl. Scheffler, 2014, S. 14). Auf den Plattformen wird mittels „Text und/oder Ton und/oder Bild und/oder Video und/oder ein einfacher Klick“ (Scheffler, 2014, S. 14) kommuniziert. So gibt es unter anderem soziale Netzwerke, beziehungsweise Netzwerkplattformen, Blogs, Micro-Blogs Video-, beziehungsweise Multimedia-Plattformen, Wikis und Instant-Messaging-Dienste (vgl. Scheffler, 2014, S. 14 und Schmidt, 2017, S. 12ff).

Soziale Netzwerke bilden so genannte „Online-Communities“, in denen registrierte Nutzer Angaben zu ihrer Person in einem Profil speichern und Beziehungen zu anderen Personen bestätigen können (vgl. Schmidt, 2017, S. 12). Dabei erfolgt die Kommunikation über „direkte Nachrichten, in thematischen Gruppen o.ä.“ (Schmidt, 2017, S. 12).

In Blogs werden dagegen „Online-Tagebücher (...) und kommentiert[e] Linklisten“ (Schmidt, 2017, S. 13) geführt. Dabei werden die einzelnen Einträge so sortiert, dass der neuste Eintrag oben steht (vgl. Schmidt, 2017, S. 13). Eine Unterkategorie der Blogs sind die Microblogs, wie zum Beispiel Twitter²⁵ (vgl. Schmidt, 2017, S. 13). Diese beinhalten nur kurze Beiträge einer jeweils festgelegten Zeichenlänge (vgl. Schmidt, 2017, S. 13).

Der Inhalt und nicht der einzelne Nutzer steht auf Multimedia-Plattformen, wie zum Beispiel YouTube, im Vordergrund (vgl. Schmidt, 2017, S. 13). Auf derartigen Plattformen können Fotos, Videos oder Musik ausgetauscht werden (vgl. Schmidt, 2017, S. 13).

²⁴ Der Begriff „soziale Medien“ wird oft auch mit den Begriffen „Web 2.0“, „Social Web“ oder „Social Software“ gleichgesetzt (vgl. Schmidt, 2017, S. 16).

²⁵ Siehe dazu auch Kapitel 4.1.

Instant-Messaging-Dienste verbinden „die Merkmale von Netzwerkplattformen und Chat- bzw. SMS-Systemen“ (Schmidt, 2017, S. 14). Sie werden oft auf Smartphones oder Tablets verwendet, wie beispielsweise WhatsApp, und die Nutzer „reagieren in der Regel sehr schnell aufeinander“ (Schmidt, 2017, S. 15).

Um problemlos „und ohne spezielle Programmierkenntnisse gemeinsam [eine Webseite] bearbeiten“ (Schmidt, 2017, S. 15) zu können, gibt es die Wiki-Plattformen. Zu den bekanntesten Wiki-Plattformen zählt Wikipedia (Schmidt, 2017, S. 15).

Zusammenfassung:

Soziale Medien dienen im Internet zum Austausch und zur Kommunikation, sowie zum Erstellen neuer Inhalte. Außerdem lassen sie sich in verschiedene Plattform-Arten unterteilen.

3.2 Soziale Medien als ein Indikator der Digitalisierung in Deutschland

Die Nutzung sozialer Medien ist ein Teil des Digitalisierungsprozesses, da sie Möglichkeiten zur digitalen Interaktion zwischen Menschen bieten. Sie sind auch in verschiedenen Studien, wie im DESI oder im Deutschland-Index, einer der Indikatoren für die Digitalisierung.

Der Deutschland-Index wertet unter anderem die Nutzung sozialer Medien aus. In seinem zweiten Themenfeld „Digitales Leben“ (Opiela, et al., 2017a, S. 14) fließt der Gebrauch sozialer Medien zu 15 Prozent ein (vgl. Opiela, et al., 2017b, S. 5). Eine Übersicht der Verwendung sozialer Medien in den einzelnen Bundesländern kann auf der Webseiten des ÖFIT über eine interaktive Karte, wie in Abbildung 3.1²⁶ zu sehen, erstellt werden.

Für das Jahr 2017 hat sich bei dieser Studie des Deutschland-Indexes heraus gestellt, dass „die Bundesländer mit der höchsten (62 Prozent) und der geringsten (46 Prozent) Nutzung sozialer Medien nur 16 Prozentpunkte auseinander“ (Opiela, et al., 2017a, S. 14) liegen. Das bedeutet, dass soziale Medien in den einzelnen Bundesländern sehr ähnlich stark verwendet werden.



Abbildung 3.1: Soziale Medien in Deutschland

Auch im DESI gibt es zwei Bereiche, in denen die sozialen Medien als Indikator der Digitalisierung eingesetzt werden (vgl. European Union, 2017b). Ein Bereich ist die Internetnutzung, in dem die digitale Kommunikation unter anderem durch die Nutzung von sozialen Netzwerken durch die Einwohner bewertet wird (vgl. European Union, 2017b). Soziale Netzwerke sind wie im vorherigen Kapitel beschrieben ein Teil sozialer Medien. In

²⁶ Diese Darstellung ist durch die Wahl der „Deutschlandkarte“ als Kartendarstellung, sowie der Auswahl „Digitales Leben“ und des Unterpunktes „soziale Medien“ erzeugt worden.

Deutschland nutzen 2017 56 Prozent der Internetnutzer soziale Netzwerke (vgl. European Union, 2017c, S. 3). Das sind sieben Prozent weniger als im europäischen Durchschnitt von 63 Prozent (vgl. European Union, 2017c, S. 3).

Der andere Bereich des DESI, in dem die Nutzung sozialer Medien als Indikator der Digitalisierung verwendet wird, lautet „Integration digitaler Technologien“ und bezieht sich auf die Digitalisierung in Unternehmen (vgl. European Union, 2017b). Dabei wird untersucht, wie viele Unternehmen mindestens zwei Formen²⁷ der sozialen Medien anwenden (vgl. European Union, 2017b). In Deutschland sind es mit 18 Prozent der Unternehmen zwei Prozent weniger als im europäischen Durchschnitt (vgl. European Union, 2017c, S. 3).

Zusammenfassung:

Das Einsetzen der Nutzung sozialer Medien als ein Bewertungskriterium in diesen Studien zeigt, dass soziale Medien in ihrer Funktion, der digitalen, menschlichen Interaktion, ein Teil der Digitalisierung sind. Somit kann die Nutzung sozialer Medien als einer von vielen Indikatoren für die Digitalisierung betrachtet werden.

²⁷ Siehe dazu Kapitel 3.1.

4 Twitter als Teil sozialer Medien im Digitalisierungsprozess

4.1 Twitter als Teil sozialer Medien

Twitter bedeutet im englischen „zwitchern“ und ist laut Dudenredaktion ein „System zur Versendung von Kurznachrichten [an eine große Zahl von Empfängern] über das Internet“ (Dudenredaktion, 2015b). Es ist eine Plattform der Firma Twitter, Inc.²⁸, die zum Konsumieren, Erstellen, Verteilen und Entdecken von Inhalten dient und über viele mobile Endgeräte, sowie über die Twitter-Webseite²⁹ oder per SMS aufgerufen werden kann (vgl. Twitter, Inc., 2017a, S. 5).

Die Kurznachrichten, die auf Twitter erstellt werden können, werden Tweets genannt (vgl. Twitter, Inc., 2017j) und umfassen maximal 280 Zeichen. Sie können auch Fotos oder Videos enthalten (vgl. Twitter, Inc., 2017j). Ebenso ist es den Nutzern möglich, Standortinformationen einem Tweet anzuhängen (vgl. Twitter, Inc., 2017i). Standardmäßig ist die Angabe des Standorts jedoch deaktiviert und muss vom Nutzer aktiviert werden, um einen Standort hinzuzufügen (vgl. Twitter, Inc., 2017i). Dabei können die Angaben sowohl Namen von Orten, wie Städtenamen, als auch geografische Koordinaten sein (vgl. Twitter, Inc., 2017i).

Tweets sind standardmäßig öffentlich zugänglich (vgl. Twitter, Inc., 2017a, S. 5 und Twitter, Inc., 2017k). Das bedeutet, dass keine Registrierung eines Nutzer von Twitter erforderlich ist, um Tweets zu lesen (vgl. Twitter, Inc., 2017k). Es ist Nutzern jedoch auch möglich einzustellen, dass die Tweets nicht öffentlich zugänglich sind (vgl. Twitter, Inc., 2017k).

Die Twitter-Plattform verzeichnet eigenen Angaben zur Folge im dritten Quartal 2017 330 Millionen Nutzer, die monatlich aktiv sind. Davon kommen 69 Millionen Nutzer aus den Vereinten Nationen und 261 Millionen aus dem Rest der Welt³⁰ (vgl. Twitter, Inc., 2017g, S. 1).

Zur Einordnung in den Bereich der sozialen Medien wird Twitter zu den Microblogs gezählt (vgl. Schmidt, 2017, S. 13 und Böck, et al., 2017, S. 2). Dennoch weist Twitter auch Ähnlichkeiten mit sozialen Netzwerken auf (vgl. Schmidt, 2017, S. 15). So können Nutzer die Tweets anderer Nutzer beispielsweise abonnieren, wodurch sie zum „Follower“ werden und sich soziale Beziehungen abzeichnen (vgl. Schmidt, 2017, S. 14).

Zusammenfassung:

Twitter ist ein Microblog mit zusätzlichen Eigenschaften sozialer Netzwerke und somit ein Teil von sozialen Medien. Eine Nachricht auf Twitter ist standardmäßig öffentlich zugänglich und wird Tweet genannt. Sie umfasst maximal 280 Zeichen und kann auch Fotos, Videos oder Standortinformationen enthalten.

4.2 Tweets als ein Indikator der Digitalisierung

Soziale Medien, zu denen auch Twitter gehört³¹, sind Teil der Digitalisierung und ihre Verwendung kann als einer der Indikatoren für die Digitalisierung angesehen werden³². Als

²⁸ Die Firma Twitter, Inc. hat ihren Hauptsitz in San Francisco, Kalifornien (vgl. Twitter, Inc., 2017a, S. 9).

²⁹ Verfügbar unter <https://twitter.com/> (abgerufen am: 12.11.2017).

³⁰ Twitter, Inc. macht weiter keine Angaben, woher die restlichen Nutzer stammen.

³¹ Siehe dazu Kapitel 4.1.

³² Siehe dazu Kapitel 3.2.

Rückschluss davon können Twitter und dementsprechend Tweets ebenfalls als kleiner Teilbereich der Digitalisierung betrachtet und die Twitter-Nutzung als ein Indikator der Digitalisierung herangezogen werden.

Die meisten Studien zur Digitalisierung untersuchen aber die Nutzung sozialer Medien insgesamt³³ und nicht aufgegliedert auf die einzelnen Plattformen. Dies wird jedoch in der Onlinestudie von ARD und ZDF untersucht, die sich mit der Internetnutzung der Menschen in Deutschland beschäftigt.

Bisher nutzen drei Prozent der Bevölkerung ab 14 Jahren Twitter mindestens wöchentlich und ein Prozent täglich (vgl. Frees & Koch, 2017, S. 444). Twitter ist damit hinter WhatsApp, Facebook, Instagram und Snapchat auf Platz fünf der meist genutzten Onlinecommunitys (vgl. Frees & Koch, 2017, S. 444). Allerdings zeigen Facebook und Twitter momentan „Stagnationstendenzen“ (Frees & Koch, 2017, S. 444) in ihrer Reichweite. In wie weit sich hier die Nutzung weiter entwickelt und somit die Digitalisierung der menschlichen Interaktion weiter fortschreitet, bleibt daher zu beobachten.

³³ Siehe dazu Kapitel 3.2.

5 Visualisierung von Daten

5.1 Definition „Informationsgrafik“ („Infografik“)

Informationsgrafiken werden auch als Infografiken bezeichnet (vgl. Lies, 2016b, S. 89) und dienen dazu, Informationen in komprimierter Form übersichtlich zu präsentieren (vgl. Lies, 2016a, S. 89). Für diese Visualisierung werden gestalterische und textliche Elemente miteinander kombiniert, um beispielsweise „Daten, Zusammenhänge [und beziehungsweise oder] Entwicklungen“ (Lies, 2016a, S. 89) in Form eines Schaubildes darzustellen. Die Aufgabe einer Infografik ist also „die Vermittlung von Informationen über ein bestimmtes Medium zu einem Empfänger“ (Schwochow, 2013, S. 147).

Es gibt jedoch verschiedene Arten von Infografiken. Es ist möglich Infografiken nach der Art ihrer Strukturierung von Informationen zu sortieren. Richard Saul Wurman (1997, S. 17) definiert dazu fünf mögliche Kategorien, die er als „LATCH“ bezeichnet. „LATCH“ steht für die Kategorien „Location“, „Alphabet“, „Time“, „Category“ und „Hierarchy“ (vgl. Wurman, 1997, S. 17). „Location“ bezeichnet die räumliche Ordnung, während „Alphabet“ die alphabetische Ordnung, wie beispielsweise in Wörterbüchern, darstellt (vgl. Wurman, 1997, S. 16f). „Time“ ist dagegen die zeitliche Ordnung und „Category“ die Ordnung nach Kategorien (vgl. Wurman, 1997, S. 17). Die letzte Kategorie, „Hierarchy“ stellt darüber hinaus die hierarchische Ordnung, also eine Ordnung nach einer Rangfolge, da (vgl. Wurman, 1997, S. 17).

Allerdings können auch Mischformen auftreten. So können Karten zum Beispiel mit Zeitleisten kombiniert werden (vgl. Rendgen, 2012, S. 96). Besonders gefördert werden diese Mischformen durch animierte und interaktive Infografiken, da „digitale Datensätze (...) heute flexibel nach Ort, Zeit oder anderen Kategorien sortiert werden“ (Rendgen, 2012, S. 96) können.

5.2 Interaktive, kartografische Infografik

Kartografische Infografiken beruhen auf Karten und visualisieren geografische Informationen (vgl. Rendgen, 2012, S. 99). Somit lassen sie sich in der Kategorie „Location“ einordnen.

Kartografische Infografiken können auch als interaktive Infografiken aufbereitet werden. Interaktive Infografiken sind Visualisierungen von Daten, die vom Betrachter beziehungsweise vom Nutzer selbst gesteuert und beeinflusst werden können (vgl. Schwochow, 2013, S. 151). Dies ist auch mit Karten möglich, denn auch Daten mit einem Raumbezug können „als dynamisch-interaktive oder animierte Karten zur Präsentation von Sachverhalten mit Raumbezug verwendet“ (Heidemann, 2013, S. 39) werden. Diese interaktiven, kartografischen Infografiken werden heute durch die „zunehmend frei[e] Verfügbarkeit von Geodaten (z. B. OpenStreetMap) und statistischen Daten sowie einfach zu nutzenden Werkzeugen zur Kartenherstellung“ (Heidemann, 2013, S. 39) immer häufiger verwendet.

5.3 Visualisierung von Daten aus sozialen Medien

Einige Plattformen sozialer Medien, wie Facebook³⁴ und Twitter³⁵, bieten die Möglichkeit ihre jeweiligen Daten abzurufen³⁶. Diese Daten können für ein besseres Verständnis visualisiert werden.

Bei der Visualisierung solcher Daten sollte jedoch grundsätzlich davon ausgegangen werden, dass „es (...) keine spezifische Social Media Visualisierung“ (Rahlf & Weller, 2014, S. 137) gibt und bekannte Techniken mit einer langen Tradition genutzt werden können. Dabei ist jedoch wichtig eine Form zu wählen, „die verständlich“ (Rahlf & Weller, 2014, S. 137) ist. So können beispielsweise Netzwerke abgebildet, Zeitreihen visualisiert oder Karten zur Darstellung der Daten verwendet werden (vgl. Rahlf & Weller, 2014, S. 143ff.).

Im Rahmen dieser Arbeit werden jeweils aktuelle, geografische Daten von Twitter³⁷ visualisiert, um die Digitalisierung zum jeweiligen Zeitraum in Deutschland beziehungsweise in den einzelnen Bundesländern darzustellen³⁸. Daher bietet sich die Verwendung einer kartografischen Infografik als Visualisierungsform an. Die zugrunde liegende Karte, wie in Abbildung 5.1³⁹ zu sehen, zeigt Deutschland, beziehungsweise das vom Nutzer ausgewählte Bundesland. Auf der Karte werden Punkte an den geografischen Koordinaten, an denen Tweets veröffentlicht worden sind, eingezeichnet. Die Farbe dieser Punkte kann ebenfalls vom Nutzer definiert werden.



Abbildung 5.1: Beispiel einer interaktiven, kartografischen Infografik in dieser Arbeit

Durch die Möglichkeit der Fokussierung des Kartenausschnitts auf ein Bundesland, die Entscheidungsfreiheit für die Farbe, sowie zum Ein- und Aus-Zoomen beziehungsweise Bewegungen in der Karte, ist die gewählte Darstellungsform nicht nur eine kartografische Infografik, sondern eine interaktive, kartografische Infografik, die einen aktiven Nutzer zur Erstellung benötigt.

Unterstützt wird die Infografik durch eine Auflistung der wichtigsten Daten unterhalb der Karte. Diese Auflistung gibt die Einwohnerzahl⁴⁰ von Deutschland und dem vom Nutzer ausgewählten Bundesland an. Darüber hinaus gibt sie die Anzahl der Tweets in ganz Deutschland und im gewählten Bundesland an. Diese Angaben können das Verständnis

34 Siehe <https://developers.facebook.com/> (abgerufen am: 16.11.2016).

35 Siehe <https://developer.twitter.com/en.html> (abgerufen am: 16.11.2016).

36 Siehe Kapitel 6.2.

37 Siehe Kapitel 4.3.2.

38 Siehe Kapitel 1.2.

39 Der Stream, der in Abbildung 5.1 visualisiert worden ist, ist am 06.01.2018 um 19.14 Uhr für eine Minute gelaufen.

40 Die Einwohnerzahlen stammen jeweils vom Statistischen Bundesamt: Genesis-Online, 05.11.2017; Datenlizenz by-2-0 (<https://www.govdata.de/dl-de/by-2-0>); Daten unter: <https://www-genesis.destatis.de/genesis/online?sequenz=tabelleErgebnis&selectionname=12411-0021®ionalschlüssel=>

des Nutzers dahingehend unterstützen, dass er nicht nur die punktförmigen Marker aller Tweets sieht, sondern eine genaue Vorstellung zur Anzahl der Tweets und ihrem Verhältnis innerhalb Deutschlands erlangt.

Zusammenfassung:

In dieser Arbeit wird eine interaktive, kartografische Form der Infografik gewählt, um den Ort der Digitalisierung in Deutschland anhand von Tweets darzustellen. Die Aussage dieser wird zusätzlich durch eine Auflistung der wichtigsten Zahlen gestützt.

6 Verwendete Werkzeuge

6.1 R und RStudio

RStudio ist eine „integrated development environment (IDE)“ (RStudio, Inc., 2017n), also eine Entwicklungsumgebung für R. R dient zum Programmieren vor allem im Bereich von Statistiken und Grafiken (vgl. The R Foundation, 2017).

Die Programmiersprache und -umgebung R ist nach den Bedingungen der „Free Software Foundation’s GNU General Public License“ (The R Foundation, 2017)⁴¹ kostenlos nutzbar und stellt eine Implementierung der Programmiersprache S von Rick Becker, John Chambers und Allan Wilks aus den Bell Laboratories dar (vgl. Venables, Smith & R Core Team, 2017, S. 2). Neben der Verarbeitung, der Berechnung und der grafischen Darstellung von Daten, bietet R beispielsweise auch Möglichkeiten zum Berechnen von Matrizen und für Datenanalysen (vgl. The R Foundation, 2017, S. 2). Darüber hinaus unterstützt diese Programmiersprache viele statistische und grafische Techniken (vgl. The R Foundation, 2017, S. 2). Somit kann sie auch als mathematische Programmiersprache betrachtet werden.

Grundsätzlich besteht sie aus 25 verschiedenen Paketen, die „Packages“ genannt werden, und kann durch weitere ergänzt werden (vgl. Venables, Smith & R Core Team, 2017, S. 2). Diese können beispielsweise über die CRAN-Webseiten⁴² erhalten werden (vgl. Venables, Smith & R Core Team, 2017, S. 2).

RStudio, als IDE für R, stellt viele praktische Funktionen bereit, um die R-Programmierung zu erleichtern. Neben „Syntax highlighting“ (RStudio, Inc., 2017l), also der farblichen Kennzeichnung der Sprachsyntax von R, liefert RStudio außerdem unter anderem die Möglichkeit Code automatisch zu komplettieren (vgl. RStudio, Inc., 2017l), indem beispielsweise der Anfang eines Funktions- oder Variablennamens eingegeben und der Rest des Namens automatisch angeboten wird. Darüber hinaus kann in RStudio ein „Debugger“ für die Fehlersuche im R-Code verwendet werden (vgl. RStudio, Inc., 2017l).

In dieser Bachelorarbeit wird die Webanwendung, die mittels R und Shiny entwickelt wird und die es Nutzern ermöglicht, den Ort der (gesellschaftlichen) Digitalisierung zu einem von ihnen bestimmten Zeitraum anhand von Twitter-Nachrichten in Deutschland beziehungsweise in den einzelnen Bundesländern zu visualisieren und mit anderen selbst gewählten Zeiträumen visuell zu vergleichen, in RStudio implementiert. Dabei wird die zur Zeit aktuelle R-Version R-3.4.2⁴³ und RStudio mit der aktuellen Version 1.0.153⁴⁴ auf einem Rechner mit dem 64-Bit-Betriebssystem „Windows 10 Home“⁴⁵ genutzt.

6.2 Das Twitter-API

Einige Plattformen von sozialen Medien, wie Facebook⁴⁶ und Twitter⁴⁷ ermöglichen einen Zugriff auf ihre Daten mittels einer jeweils von ihnen bereit gestellten technischen Schnitt-

41 Für weitere Informationen siehe <http://www.gnu.org/> (abgerufen am: 02.12.2017) und <https://www.r-project.org/COPYING> (abgerufen am: 02.12.2017).

42 Für weitere Informationen siehe <https://cran.r-project.org/> (abgerufen am: 02.12.2017).

43 Siehe dazu Anhang A.2.

44 Siehe dazu Anhang A.3.

45 Siehe dazu Anhang A.1.

46 Siehe <https://developers.facebook.com/> (abgerufen am: 16.11.2016).

47 Siehe <https://developer.twitter.com/en.html> (abgerufen am: 16.11.2016).

stelle. Um Daten von Twitter zu erhalten, bietet Twitter ein Application Programming Interface (API) an, das Twitter-API genannt wird (vgl. Burgess & Bruns, 2014, S. 192) und auf dem Hypertext Transfer Protocol (HTTP) zur Datenübertragung und dem Secure Sockets Layer (SSL) zur Datenverschlüsselung basiert (vgl. Twitter, Inc., 2017h). Auf diese API kann von vielen Bibliotheken, wie beispielsweise Java, Python, C++ oder vielen mehr, zugegriffen werden (vgl. Twitter, Inc., 2017l).

Mit dem API ist es möglich, selbst Anwendungen⁴⁸ zur Nutzung von Twitter zu entwickeln, sowie die Daten öffentlicher Tweets abzurufen, da es „die technischen Spezifikationen, wie eine Software-Anwendung“ (Burgess & Bruns, 2014, S. 194) auf Twitter zugreifen soll, bietet. Die mit dem API abgerufenen Daten werden dabei im JavaScript Object Notation (JSON) Format an die abrufende Software zurückgegeben (vgl. Kumar, Morstatter & Liu, 2014, S. 6 und Twitter, Inc., 2017h).

Zur Nutzung des API muss jedoch die jeweilige Anwendung bei Twitter⁴⁹ registriert werden (vgl. Twitter, Inc. 2017d). Durch die Registrierung erhält man von Twitter die Daten für die Authentifizierung⁵⁰. Ohne diese wird der Zugang zu den Twitter-Daten nicht gewährt.

Das Twitter-API bietet somit sowohl Möglichkeit für Unternehmen, als auch für wissenschaftliche Untersuchungen. Allerdings wird der Zugang zu den Twitter-Daten für die Wissenschaft erschwert, da Twitter einen vollen Zugriff nur „auf kommerziellem Preisniveau“ (Burgess & Bruns, 2014, S. 196) erlaubt. Dadurch enthält der kostenfreie Zugang nicht den vollständigen Datenumfang. Deshalb kann nur auf das zugegriffen werden, „was Twitters proprietäres und häufig wechselndes API zur Verfügung stellt“ (Burgess & Bruns, 2014, S. 197).

6.2.2 Twitter-Search-API versus Twitter-Streaming-API

Twitter unterscheidet sein API-Angebot in ein REST-API und ein Streaming-API (vgl. Kumar, Morstatter & Liu, 2014, S. 5 und Twitter, Inc., 2017h). Alle API-Schnittstellen, abgesehen von dem Streaming-API, folgen nämlich dem Design-Prinzip des Representational State Transfer (REST)⁵¹ und zählen somit zu den REST-APIs von Twitter (vgl. Twitter, Inc., 2017h).

Zum Erhalten von Tweets und ihrer Daten kann sowohl das Search-API, das Teil der REST-API ist, als auch das Streaming-API verwendet werden (vgl. Kumar, Morstatter & Liu, 2014, S. 14 und Twitter, Inc., 2017f). Allerdings gibt es Unterschiede zwischen diesen APIs, wie in Tabelle 6.1 auf der nächsten Seite dargestellt, die bei einer Entscheidung zu bedenken sind.

Der grundlegende Unterschied zwischen dem Search- und dem Streaming-API von Twitter liegt in seiner jeweiligen Funktion. Während das Search-API das Twitter-Archiv nach der jeweiligen Anfrage durchsucht und somit vergangene Tweets im JSON-Format zurückliefert (vgl. Twitter, Inc., 2017f), ruft das Streaming-API die Tweets in Echtzeit ab, die exakt zum Zeitpunkt der Abfrage veröffentlicht werden, und liefert diese im JSON-Format zurück (vgl. Twitter, Inc., 2017c).

48 Diese Anwendungen werden von Twitter auch „endpoint“ genannt (vgl. Twitter, Inc., 2017d).

49 Die Registrierung erfolgt unter <https://apps.twitter.com/> (abgerufen am: 16.11.2017).

50 Siehe dazu Kapitel 6.2.

51 Dieses Prinzip soll in dieser Arbeit aus Relevanz-Gründen nicht weiter erläutert werden, da es nur der internen Verarbeitung der Daten in den Software-Programmen, die diese API verwenden, dient.

Merkmale	Twitter-Search-API	Twitter-Streaming-API
Untergeordnete APIs	Standard Search API, Search Tweets: 30-Day API, 30-Day Search API, Full-Archive Search API (vgl. Twitter, Inc., 2017f)	POST statuses/filter, Realtime PowerTrack (vgl. Twitter, Inc., 2017c)
Funktion	Tweets aus dem Twitter-Archiv (vgl. Twitter, Inc., 2017f)	Tweets in Echtzeit abrufen (vgl. Twitter, Inc., 2017c)
Authentifizierung	Nutzer- / App-Authentifizierung (vgl. Twitter, Inc., 2017f)	Nutzer-Authentifizierung (vgl. Twitter, Inc., 2017c)
Mindestens benötigte Parameter	Suchbegriff wird benötigt, alle anderen Parameter sind optional (vgl. Twitter, Inc., 2017c)	Alle Parameter sind optional (vgl. Twitter, Inc., 2017c)
Antwort-Format	JSON (vgl. Twitter, Inc., 2017f)	JSON (vgl. Twitter, Inc., 2017c)
Rate Limit	180 Anfragen pro 15 Minuten bei Nutzer-Authentifizierung bzw. 450 Anfragen pro 15 Minuten bei App-Authentifizierung (vgl. Twitter, Inc., 2017f)	Wird die Verbindung zu oft nacheinander vom selben Client neu aufgebaut, wird sie für einige Minuten unterbrochen. Geschieht dies zu oft, wird die entsprechende IP für unbestimmte Zeit gesperrt. (vgl. Twitter, Inc., 2017c)
Kostenfreier Zugang	Zugang zum Archiv der letzten 7 Tage, fokussiert auf die Relevanz und nicht die Vollständigkeit (vgl. Twitter, Inc., 2017f)	Maximal 400 Stichwörter, 5.000 Nutzer-IDs & 25 Standort-Boxen (vgl. Twitter, Inc., 2017c)
Kostenpflichtiger Zugang	Verschiedene Modelle, voller Datenumfang, entweder für die letzten 30 Tage oder für das gesamte Archiv seit 2006 (vgl. Twitter, Inc., 2017f)	Pro Stream bis zu 250.000 Filter und jeweils bis zu 2.048 Zeichen (vgl. Twitter, Inc., 2017c)

Tabelle 6.1: Twitter-Search-API versus Twitter-Streaming-API

Quelle: eigene Darstellung

Darüber hinaus ist das Search-API immer an mindestens einen konkreten Begriff, der in den Tweets gesucht werden kann, als Parameter, gebunden, während alle anderen Parameter optional sind (vgl. Twitter, Inc., 2017f). Dagegen benötigt das Streaming-API keinen solchen Suchbegriff, sondern alle Parameter sind optional. Dadurch ist es möglich, dieses API auch völlig ohne Filter-Parameter zu nutzen, was das Search-API nicht ermöglicht. Allerdings erhält man ohne Filterung, nur mit den Standard-Vorgaben, lediglich eine zufällige, beispielhafte Sammlung von Tweets und nicht alle zu der Zeit veröffentlichten Tweets (vgl. Kearney, 2017a). Daher ist die Verwendung von mindestens einem Parameter als Filter sinnvoll. Dennoch können als Filter auch geografische Koordinaten, statt eines Suchbegriffs, genutzt werden (vgl. Twitter, Inc., 2017c).

Das Search-API von Twitter gliedert sich in vier APIs. Das erste API ist das kostenfreie „Standard Search API“ (Twitter, Inc., 2017f). Die anderen drei heißen „Search Tweets: 30-Day API“ (Twitter, Inc., 2017f), „30-Day Search API“ (Twitter, Inc., 2017f) und „Full-Archive Search API“ (Twitter, Inc., 2017f) und sind kostenpflichtig. Hierbei ist zu beachten, dass der kostenfreie Zugang nur Zugriff auf die Tweets der letzten sieben Tage gewährt, während der kostenpflichtige Zugang je nach Variante einen Zugriff auf die letzten dreißig Tage oder sogar auf das gesamte Archiv seit 2006 gewährt (vgl. Twitter, Inc., 2017f).

Auch das Streaming-API gliedert sich in zwei APIs. Dem kostenlosen „POST statuses/filter“ (Twitter, Inc., 2017c) und dem kostenpflichtigen „Realtime PowerTrack“ (Twitter, Inc., 2017c). Zusätzlich wird der Stream, den das Streaming-API empfängt, noch in den „Public Stream“, der die öffentlichen Tweets enthält, den „User Stream“, der jeweils die Tweets eines Nutzers beinhaltet, sowie den „Site Stream“, der Tweets von mehreren Nutzern streamt, unterschieden (vgl. Kumar, Morstatter & Liu, 2014, S. 5). Auch hier wird der kostenfreie Zugang eingeschränkt, da hier maximal 400 Stichwörter, 5.000 Nutzer-IDs und 25 Standort-Boxen gefiltert werden können, während in der kostenpflichtigen Variante pro Stream bis zu 250.000 Filter und jeweils bis zu 2.048 Zeichen eingesetzt werden können (vgl. Twitter, Inc., 2017c).

Ein weiterer Unterschied dieser beiden APIs liegt in den Angaben zum sogenannten „Rate Limit“ durch Twitter. Das „Rate-Limit“ bezeichnet dabei die maximale Anzahl an Anfragen an Twitter (vgl. Twitter, Inc., 2017e). Zum „Rate Limit“ des Streaming-API, welches nur mit einer Nutzer-Authentifizierung verwendet werden kann, macht Twitter keine genauen Angaben. Hier heißt es lediglich: Wird die Verbindung zu oft nacheinander vom selben Client neu aufgebaut, wird diese für einige Minuten unterbrochen. Geschieht dies zu oft, wird die entsprechende IP für unbestimmte Zeit gesperrt (vgl. Twitter, Inc., 2017c). Dagegen hat das Search-API ein festes „Rate Limit“ von 180 Anfragen pro 15 Minuten⁵², wenn die Nutzer-Authentifizierung verwendet wird, beziehungsweise 450 Anfragen pro 15 Minuten⁵³, wenn die App-Authentifizierung genutzt wird (vgl. Twitter, Inc., 2017f).

Zusammenfassung:

Das Search-API durchsucht anhand eines Suchbegriffs und gegebenenfalls weiterer Filter-Parameter das Twitter-Archiv, während die Streaming-API Tweets in Echtzeit abrufen, ohne zwingend Parameter zu benötigen. Beide APIs besitzen dabei sowohl einen kostenfreien, aber eingeschränkten, als auch einen kostenpflichtigen, vollständigen Zugang.

6.2.3 Verwendung des Twitter-API im Rahmen dieser Arbeit

Im Rahmen dieser Arbeit wird visualisiert, wo Tweets in Deutschland beziehungsweise in den einzelnen Bundesländern veröffentlicht werden⁵⁴. Um diese Daten über die Standorte der veröffentlichten Tweets zu erhalten, wird das Twitter-API benötigt. Dabei kommt auf den ersten Blick sowohl das Search-API als auch das Streaming-API in Frage. Theo-

⁵² Das heißt, es dürfen pro User-Access-Token innerhalb von 15 Minuten maximal 180 Anfragen an das Twitter-Archiv gestellt werden (vgl. Twitter, Inc., 2017e).

⁵³ Das heißt, es dürfen innerhalb von 15 Minuten maximal 450 Anfragen von der gesamten Anwendung an das Twitter-Archiv gestellt werden (vgl. Twitter, Inc., 2017e).

⁵⁴ Siehe dazu Kapitel 1.2.

retisch wäre es daher sinnvoll, das Streaming-API für aktuelle und das Search-API für vergangene Tweets zu nutzen.

Sowohl das Search-API als auch das Streaming-API bieten die Möglichkeit, nach geografischen Koordinaten zu filtern (vgl. Twitter, Inc., 2017c und Twitter, Inc., 2017f). Allerdings kann das Search-API, im Gegensatz zum Streaming-API, nicht alle Tweets des Twitter-Archivs nur nach den Koordinaten durchsuchen, sondern es benötigt einen konkreten Suchbegriff. Das bedeutet, dass alle Tweets, die diesen Begriff nicht beinhalten, werden nicht berücksichtigt, auch wenn sie die gesuchten Koordinaten besitzen (vgl. Twitter, Inc., 2017f). Da jedoch alle Tweets unabhängig vom Inhalt berücksichtigt werden sollen⁵⁵, kann das Search-API hier nicht verwendet werden.

Dementsprechend kann in dieser Arbeit nur das Streaming-API verwendet werden. Das bedeutet aber auch, dass immer nur Tweets vom jeweils aktuellen Zeitpunkt über einen bestimmbaren Zeitabschnitt, wie zum Beispiel für 60 Sekunden, zum Filtern nach den entsprechenden Koordinaten zur Verfügung stehen. Daher können keine bereits vergangenen Tweets abgerufen werden. Ein Vergleich mit anderen Zeiträumen ist daher nur möglich, indem die Daten gespeichert und mit anderen gespeicherten Daten, oder einige Zeit später⁵⁶, mit einem dann aktuellen Stream-Zeitraum verglichen werden.

Zu bedenken ist dabei jedoch auch, wie die Streaming-API die Standortdaten ermittelt. Enthält der jeweilige Tweet vom Ersteller freigegebene geografische Koordinaten, werden diese verwendet (vgl. Twitter, Inc., 2017c). Enthält er diese nicht, aber stattdessen eine Ortsangabe, wird dies genutzt (vgl. Twitter, Inc., 2017c). Ansonsten wird der Tweet bei dieser Filterung nicht berücksichtigt (vgl. Twitter, Inc., 2017c).

Zusammenfassung:

Innerhalb dieser Bachelorarbeit wird also die Streaming-API verwendet, um die geografischen Koordinaten der Tweets nach Tweets aus Deutschland zu filtern und die entsprechenden Daten abzurufen.

6.3 R-Package „rtweets“

Zum Abrufen der Daten von Twitter mittels R wird das R-Package „rtweet“ benötigt, das es ermöglicht, sowohl Daten über die REST-APIs, wie das Search-API⁵⁷, als auch über das Streaming-API⁵⁸ abzurufen (vgl. Kearney, 2017c, S.1).

Hauptziele des Packages ist es, Anfragen an die Twitter-APIs zu formulieren und zu stellen, Daten dieser abzurufen und sie in aufgeräumten Strukturen unterzubringen (vgl. Kearney, 2017c, S.1). Dazu bietet das Package viele Möglichkeiten, in Form verschiedener Funktionen, zum Abrufen, Filtern und Verarbeiten der Daten (vgl. Kearney, 2017c, S. 2ff).

Um dieses Ziel erfüllen zu können, wird eine Authentifizierung bei Twitter benötigt, da Twitter zur Benutzung seiner API eine „Open Authentication (OAuth)“ (Kumar, Morstatter & Liu, 2014, S. 6) benötigt. Dies ist eine Authentifizierung mittels eines sogenannten

⁵⁵ Siehe dazu Kapitel 1.2.

⁵⁶ Dies können beispielsweise Sekunden, Minuten, Stunden, Tage, Wochen oder Jahre später sein.

⁵⁷ Siehe Kapitel 4.3.1.

⁵⁸ Siehe Kapitel 4.3.1.

Drei-Wege-Handshake (vgl. Kumar, Morstatter & Liu, 2014, S. 6). Dadurch muss der Nutzer sein Passwort für Twitter keiner Dritten Partei, wie dem Anwendungsanbieter, mitteilen (vgl. Twitter, Inc., 2017b). Dazu gibt es im Package „rtweets“ die Funktion `create_token(app = appname, consumer_key = key, consumer_secret = secret)` (vgl. Kearney, 2017b). Diese benötigt einen Schlüssel und ein sogenanntes Geheimnis, das von Twitter erhalten werden kann, in dem auf apps.twitter.com⁵⁹ eine Anwendung eingetragen wird (vgl. Kearney, 2017b). Dies ist nötig, um die Funktionen des Packages „rtweet“ nutzen zu können.

Da, wie in Kapitel 6.2.2 bereits beschrieben, für diese Arbeit das Streaming-API verwendet wird, ist hier die Funktion `stream_tweets()` (vgl. Kearney, 2017a) die wichtigste für die Programmierung dieser Webanwendung. Mit dieser Funktion wird der aktuelle Twitter-Stream abgerufen. Sie bietet Möglichkeiten, um den Stream bereits während des Abrufens zu filtern (vgl. Kearney, 2017c, S. 49). Dabei kann sowohl nach einem Suchbegriff, als auch nach einer Nutzer-ID oder nach geografischen Koordinaten gefiltert werden (vgl. Kearney, 2017c, S. 49). Ohne Filterung erhält man, wie in Kapitel 6.2.2 erläutert, nur eine zufällige, beispielhafte Sammlung von Tweets und nicht alle zu der Zeit veröffentlichten Tweets (vgl. Kearney, 2017a). Deshalb sollte mindestens ein Filter, wie der Filter nach den geografischen Koordinaten, verwendet werden, um tatsächlich alle in Frage kommenden Daten zu empfangen.

In dieser Arbeit findet die Version 0.6.0⁶⁰ des R-Packages „rtweets“ Verwendung, um die gewünschten Daten von Twitter abzurufen.

6.4 R-Package „shiny“

Eine Shiny-App ist eine interaktive Web-Anwendung (App), die direkt aus R mit dem R-Package „shiny“ erstellt wird (vgl. Chang, et al., 2017, S. 1 und RStudio, Inc., 2017e). In dieser Arbeit wird daher die Version 1.0.5⁶¹ dieses Packages genutzt.

Es gibt zwei verschiedene Möglichkeiten, eine Shiny-App zu strukturieren. Die erste Möglichkeit ist, sie grundsätzlich als ein R-Script, namens „app.R“, in einem Ordner, der den Namen der App trägt, anzulegen (vgl. RStudio, Inc., 2017e). Dieses Script muss dann aus drei Funktionen bestehen. Diese sind eine Funktion für die Benutzeroberfläche (UI), eine Server-Funktion und die „shinyApp“-Funktion. Letztere Funktion dient dabei dazu, die anderen beiden als Objekt aufzurufen (vgl. Chang, et al., 2017, S. 150). Die andere Möglichkeit ist, die App stattdessen auf zwei Scripten aufzubauen (vgl. RStudio, Inc., 2017e). Dazu wird eine „ui.R“ für die Programmierung der UI und eine „server.R“ für die Server-Logik genutzt (vgl. RStudio, Inc., 2017e).

Bis zur Version 0.10.2 des Packages ist nur die Variante mit zwei Scripten möglich (vgl. RStudio, Inc., 2017e). Erst in den nachfolgenden Versionen kann die „app.R“ verwendet werden (vgl. RStudio, Inc., 2017e). Dennoch haben beide Varianten ihre Vor- und Nachteile. Existiert nur ein Script, ist es möglich, den Code zu kopieren und vollständig in die

⁵⁹ Abgerufen am: 03.12.2017

⁶⁰ Siehe dazu Anhang A.4.

⁶¹ Siehe dazu Anhang A.6.

R-Konsole einzufügen (vgl. RStudio, Inc., 2017e). Dies ist mit zwei Scripten nicht möglich. Je komplexer eine App jedoch ist, beziehungsweise je mehr Code für eine App benötigt wird, desto länger wird auch das Script. Um eine komplexe App schneller und einfacher überblicken zu können, bietet es sich daher an, die Variante mit den zwei Scripten zu nutzen. Aus diesem Grund wird auch in dieser Arbeit mit einer „ui.R“ und einer „server.R“ gearbeitet.

6.4.1 Die Benutzeroberfläche

Wie bereits erwähnt, wird die UI entweder im Script „ui.R“ oder im Script „app.R“ definiert. In der „app.R“ geschieht dies, indem der Variable „ui“ eine Funktion zum Erzeugen einer Seite zugewiesen wird (vgl. RStudio, Inc., 2017f). Ähnlich wird auch in der „ui.R“ eine Seite erzeugt. Hier wird die Funktion `shinyUI()` und in ihr dann eine Funktion zum Erzeugen einer Seiten aufgerufen (vgl. Chang, et al., 2017, S. 153f).

Eine dieser Funktionen zum Erzeugen einer Seite ist `fluidPage()`. Die damit erstellte Seite passt sich selbstständig an die Fenstergröße des genutzten Internetbrowsers an (vgl. RStudio, Inc., 2017f), und ermöglicht so auch die Ansicht auf Mobilgeräten.

Um der Seite Elemente hinzuzufügen, können in der Funktion wiederum sogenannte „Panels“, wie unter anderem das „sidebarPanel“ oder das „mainPanel“, eingesetzt werden (vgl. RStudio, Inc., 2017f). Dies können auch mehrere verschiedene Panels sein. In diesen Panel-Funktionen können dann die Inhalte, wie HTML-Elemente (vgl. RStudio, Inc., 2017f) oder Widgets, die beispielsweise Buttons oder Auswahlmöglichkeiten für Nutzerangaben bieten (vgl. RStudio, Inc., 2017g), implementiert werden. Die Eingaben in den Widgets, die der Nutzer im fertigen UI tätigt, können dann wiederum vom Server als „input“ verarbeitet werden, wenn der Nutzer die Auswahl ändert, beziehungsweise einen Button drückt (vgl. RStudio, Inc., 2017h).

Darüber hinaus kann ein Panel auch eine oder mehrere sogenannte „Output“-Funktionen, wie „plotOutput“ oder „textOutput“, enthalten. Diese Funktionen dienen der Ausgabe vom Server erzeugter Elemente, wie beispielsweise eine Infografik oder sich dynamisch anpassender Text (vgl. RStudio, Inc., 2017h).

6.4.2 Die Server-Funktion

Die Server-Funktion wird ebenfalls entweder in der „app.R“ als Zuweisung der Variable „server“ (`server <- function(input, output){func}`) (vgl. RStudio, Inc., 2017i), oder in der „server.R“ direkt, ohne Zuweisung als `shinyServer(func)` aufgerufen (vgl. Chang, et al., 2017, S. 152f). Diese enthält den gesamten Code für die Server-Logik (vgl. RStudio, Inc., 2017a). Der Code beinhaltet Funktionen, die der Nutzer nicht sehen kann, sondern im Hintergrund ablaufen, um die Ausgabe zu steuern.

Bei der Implementierung der Server-Logik ist zu beachten, dass es drei verschiedene Bereiche für den eigenen Code gibt (vgl. RStudio, Inc., 2017i). Das gilt sowohl für die Version mit „app.R“ (vgl. RStudio, Inc., 2017i), als auch für die Version mit der „server.R“ (vgl. Chang, et al., 2017, S. 153). In der Abbildung 6.1 auf der nächsten sind diese Bereiche dargestellt.

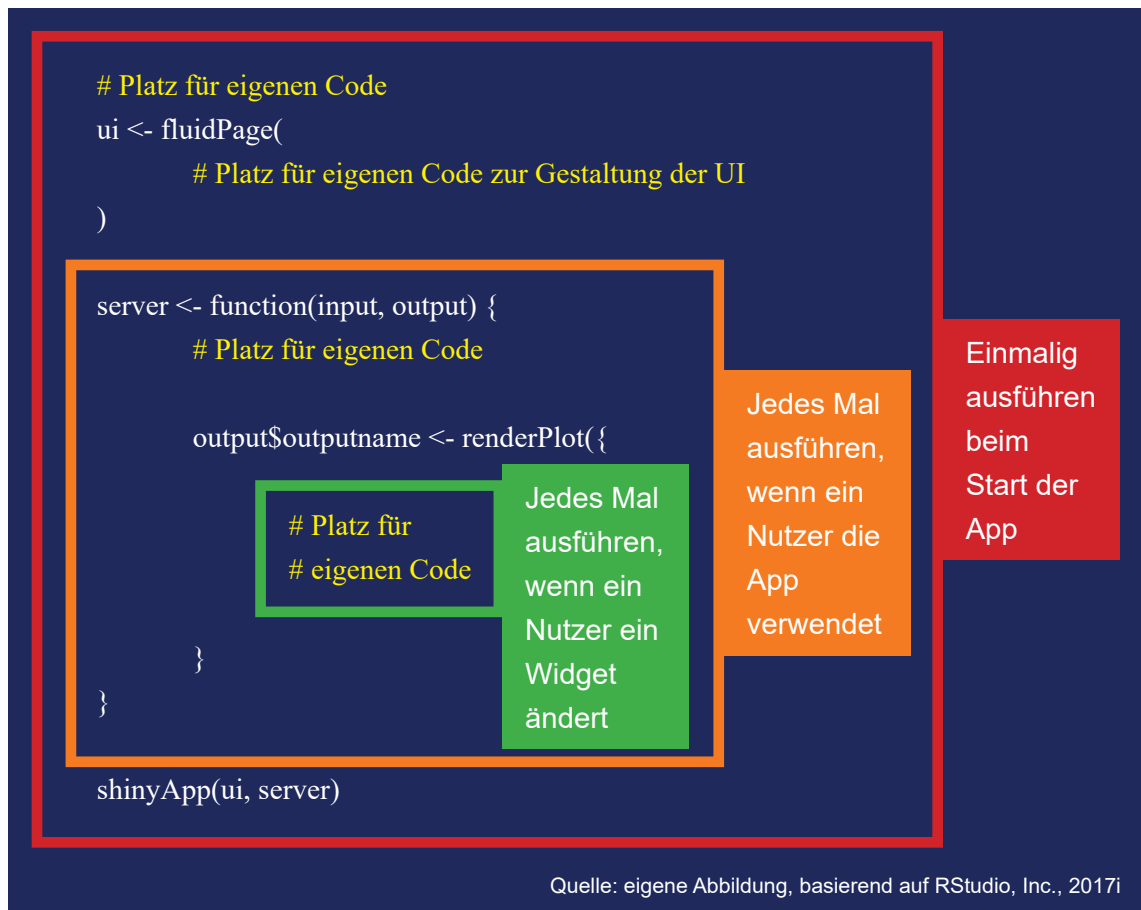


Abbildung 6.1: Aufbau einer Shiny-App - Die „app.R“

Der erste, äußerste Bereich umgibt die eigentliche Server-Funktion und wird nur einmalig durchgeführt, wenn die App gestartet wird (vgl. RStudio, Inc., 2017i). Dieser bietet sich daher dafür an, weitere Skripte, R-Packages oder auch Dateien einmalig einzuladen (vgl. RStudio, Inc., 2017i), so dass sie in der App durchgängig genutzt werden können.

Die Server-Funktion selbst bildet den zweiten Bereich, der immer dann ausgeführt wird, wenn ein Nutzer die App verwendet (vgl. RStudio, Inc., 2017i). In diesem Bereich können beispielsweise nutzerspezifische Variablen deklariert werden (vgl. RStudio, Inc., 2017i).

Der dritte Bereich ist der innerste Bereich (vgl. RStudio, Inc., 2017i). Dieser ist die Render-Funktion, die jedes Mal ausgeführt wird, wenn der Nutzer den Wert eines Widgets in der UI ändert⁶² (vgl. RStudio, Inc., 2017i). Hierfür gibt es, abhängig davon, welche Output-Funktion in der UI verwendet wird, entsprechende Render-Funktionen, wie beispielsweise `renderPlot({func})` oder `renderText({func})` (vgl. RStudio, Inc., 2017h).

6.4.3 Die Veröffentlichung einer Shiny-App

Da Shiny ein R-Package zur Erstellung interaktiver Web-Anwendungen direkt aus R ist, gibt es verschiedene Möglichkeiten, diese Apps anderen zugänglich zu machen, beziehungsweise sie zu veröffentlichen (vgl. RStudio, Inc., 2017k). Dabei wird unterschieden, ob sie als R-Script oder als Webseite bereit gestellt werden sollen (vgl. RStudio, Inc., 2017k).

⁶² Zum Beispiel durch Treffen einer Auswahl oder Klicken eines Buttons.

Wird eine App als R-Script veröffentlicht, muss der gesamte Ordner der App publiziert werden (vgl. RStudio, Inc., 2017k). Dieser kann beispielsweise per E-Mail versendet oder im Internet zugänglich gemacht werden (vgl. RStudio, Inc., 2017k). Wird der App-Ordner per E-Mail oder ähnlichem weitergegeben, kann der Empfänger diesen in seinem Arbeitspfad in R speichern und die App über `runApp(„App-Ordner-Name“)` ausführen, sofern Shiny installiert und geladen ist (vgl. RStudio, Inc., 2017k). Ähnlich funktioniert die Ausführung auch bei einer Bereitstellung des gesamten Ordners der App über das Internet. Abhängig davon, ob die App auf einem eigenen Server, einem GitHub-Server oder anonym auf einem GitHub-Server bereitgestellt wird, können alternativ zur Ausführung die Befehle `runURL(„www-Adresse“)`, `runGitHub(„Repository-Name“, „Nutzer-Name“)` oder `runGist(„Repository-Nummer“)` in die Konsole für R eingegeben werden (vgl. RStudio, Inc., 2017k). Nachteilig an der Veröffentlichung als R-Script ist jedoch, dass der Nutzer selbst R und Shiny benötigt, um die App auszuführen (vgl. RStudio, Inc., 2017k).

Soll die App jedoch auch von Nutzern verwendet werden, die R und Shiny nicht nutzen, ist eine Veröffentlichung als Webseite empfehlenswert (vgl. RStudio, Inc., 2017k). Es besteht dann die Möglichkeit, die App entweder auf einem eigenen Webserver zu veröffentlichen oder eines der vier Angebote von RStudio zu nutzen (vgl. RStudio, Inc., 2017k). RStudio bietet hier entweder die „Shinyapps.io“, den „Shiny Server“, den „Shiny Server Pro“ oder „RStudio Connect“ an (vgl. RStudio, Inc., 2017k). Eine als Webseite veröffentlichte App kann dann von jedem Interessierten, der einen Internetzugang besitzt, ausgeführt werden.

6.5 R-Package „leaflet“

Die Version 1.1.0⁶³ des R-Packages „leaflet“ wird ebenfalls in dieser Bachelorarbeit verwendet. Mit diesem Package können interaktive Karten in R erstellt und an die jeweiligen Anforderungen angepasst werden (vgl. Cheng, J., et al., 2017, S. 1), da es die Implementierung der freien JavaScript Bibliothek „Leaflet“⁶⁴ ist, die als eine der bekanntesten Bibliotheken in diesem Bereich gilt (vgl. RStudio, Inc., 2017a). Außerdem ist sie auch auf mobilen Endgeräten ausführbar (vgl. Agafonkin, 2017).

Eine mit dem R-Package „leaflet“ erzeugte Karte kann aus unterschiedlichem Kartenmaterial bestehen, die als „Basemaps“ bekannt sind (vgl. RStudio, Inc., 2017c). Dazu können Kartenmaterialien verschiedener Anbieter⁶⁵ oder über eine eigene URL bereitgestelltes Kartenmaterial ausgewählt werden (vgl. RStudio, Inc., 2017c). Ohne eine Spezifizierung des Kartenmaterials wird das Kartenmaterial von OpenStreetMap genutzt (vgl. RStudio, Inc., 2017c). Es ist außerdem möglich, verschiedene Kartenmaterialien miteinander zu kombinieren, indem diese als Ebenen übereinander gelegt werden (vgl. RStudio, Inc., 2017c). Allerdings ist dabei darauf zu achten, dass das Kartenmaterial, das ein anderes überlagert, entsprechende Transparenz aufweist (vgl. RStudio, Inc., 2017c), da ansonsten das untere Material lediglich verdeckt wird.

Des Weiteren können auch sogenannte Marker als Ebene zu einer Karte hinzugefügt werden (vgl. RStudio, Inc., 2017b). Der Standard-Marker ist ein Pin (vgl. RStudio, Inc.,

⁶³ Siehe dazu Anhang A.5.

⁶⁴ Für weitere Informationen siehe <http://leafletjs.com/> (abgerufen am: 05.12.2017).

⁶⁵ Ein Überblick mit einer Vorschau des Kartenmaterials ist zu finden unter: <http://leaflet-extras.github.io/leaflet-providers/preview/index.html> (abgerufen am: 05.12.2017)

2017b). Es können aber auch Kreise oder eigene Bilder als Marker verwendet werden (vgl. RStudio, Inc., 2017b). Außerdem sind die Farben der Marker anpassbar.

Neben den vielen Möglichkeiten zum Anpassen des Layouts, zum Beispiel durch die Auswahl der Karten oder die Farb- und Formänderung der Marker, bietet dieses Package eigene Funktionen für die Integration der „leaflet“-Karten in Shiny-Anwendungen⁶⁶ (vgl. RStudio, Inc., 2017d). So beinhaltet es die Funktion `leafletOutput("kartename")`, die in der „ui.R“ genutzt werden kann, und die Funktion `renderLeaflet({func})` für die Verwendung in der „server.R“ (vgl. RStudio, Inc., 2017d). Dadurch bietet sich dieses Package für die Nutzung seiner Funktionen in dieser Arbeit an.

⁶⁶ Ein Beispiel für die Verwendung von Leaflet in Shiny ist zu finden unter:
<http://shiny.rstudio.com/gallery/superzip-example.html> (abgerufen am: 05.12.2017)

7 Visualisierung der Digitalisierung in den Bundesländern

Wie bereits in Kapitel 6.4 erwähnt, wird für die Implementierung der Webanwendung zur Visualisierung der Digitalisierung in Deutschland beziehungsweise in den deutschen Bundesländern am Beispiel von Twitter-Nachrichten die Variante mit zwei Scripten für eine Shiny-App verwendet.

Diese Skripte und sämtliche Unterordner, Dateien und weitere R-Skripte⁶⁷ der entwickelten App liegen, wie in Abbildung 7.1, im Ordner „DeutschlandDigital“. Dementsprechend sind in diesem Hauptordner die Skripte „server.R“ und „ui.R“, die Datei „Einwohnerzahlen.csv“, die alle Einwohnerzahl von Deutschland und den einzelnen Bundesländern beinhaltet, und drei Unterordner zu finden.

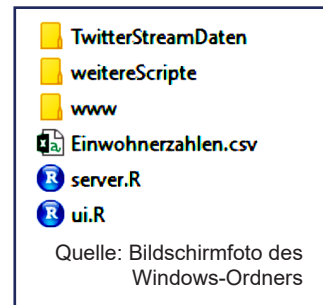


Abbildung 7.1: Ordnerstruktur der Shiny-App

Im Unterordner „TwitterStreamDaten“ wird jeweils der empfangene Stream durch die App abgespeichert. Der Unterordner „www“ ist ein Ordner, in den alle Dateien gehören, die der Webbrowser des Nutzers benötigt, wie beispielsweise Bilder oder Cascading Style Sheets (CSS) Dateien (vgl. RStudio, Inc., 2017f). Daher ist in dieser App die „style.css“ in diesem Ordner.

Darüber hinaus enthält der Unterordner „weitereSkripte“ alle zusätzlichen 22 R-Skripte⁶⁸. Diese werden von der „server.R“ verwendet. Außerdem rufen sie sich zum Teil auch gegenseitig auf. Eine grafische Darstellung, welche Skripte sich aufrufen, ist in Abbildung 7.2 zu sehen.

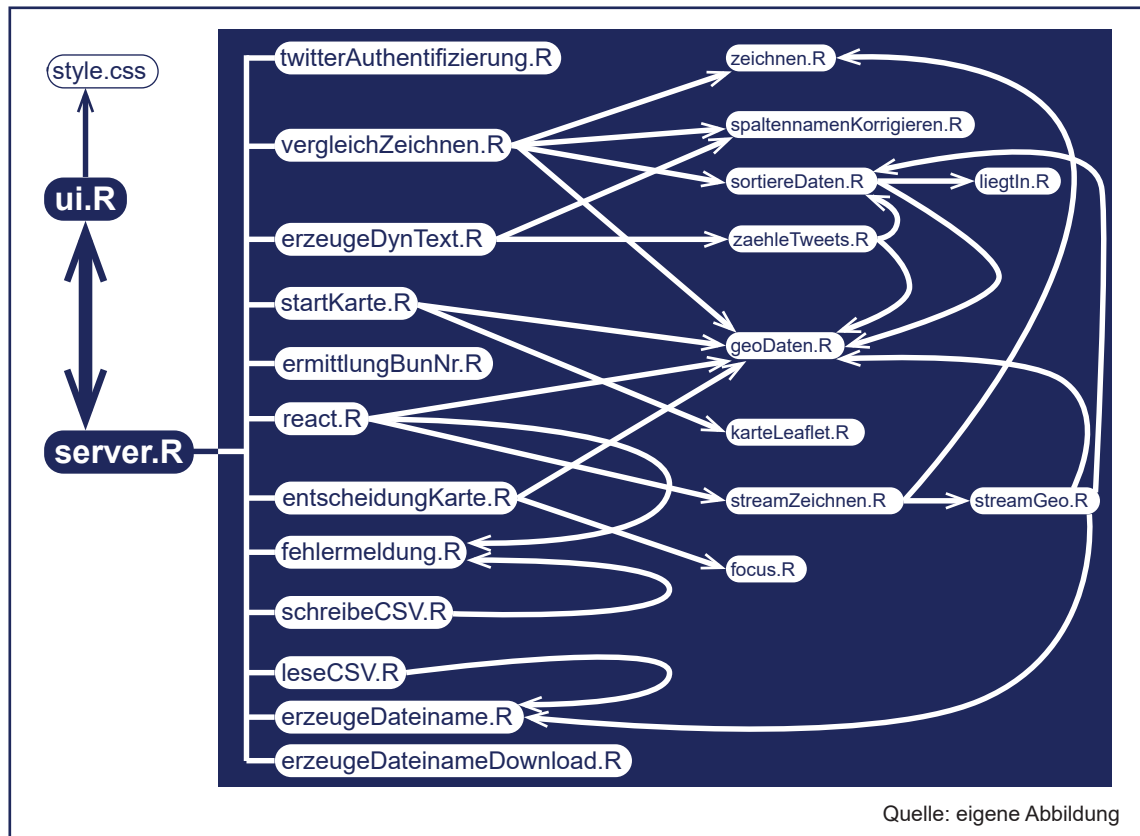


Abbildung 7.2: Übersicht über den strukturellen Aufbau der Shiny-App

67 Der vollständige Code aller R-Skripte der App sind in Anhang B einsehbar.

68 Siehe dazu Kapitel 7.3.

7.1 „ui.R“ und „style.css“ der Shiny-App

Die Benutzeroberfläche der Anwendung wird in der „ui.R“ festgelegt, während mit der „style.css“ beispielsweise Farben und die typografische Gestaltung veranlasst werden.

Die „ui.R“ lädt zu Anfang alle benötigten Packages jeweils mit dem Befehl `library(Packagename)` und alle Scripte jeweils mit dem Befehl `source(Dateipfad/Dateiname.R)`. Außerdem wird eine Variable definiert, die alle Farbnamen, die später für die Initialisierung der Farbauswahl dienen, als Array enthält. Dann erst wird die Webseite als `fluidPage(...)` initialisiert, damit sie sich dynamisch an die Breite des Fensters vom Nutzer anpasst (vgl. Chang, W., et al., 2017, S. 48).

Auch die „style.css“ wird hier eingebunden. Diese bestimmt die Breite der verschiedenen Elemente, sowie die Farbgebung. Sie legt darüber hinaus die Maße der Elemente fest und macht Angaben zu Schriftart, Schriftgröße und ähnlichen typografischen Elementen.

Die Webseite selbst besteht aus dem `titlePanel(„Gesellschaftliche Digitalisierung in Deutschland“)`, einem `sidebarPanel(...)`, einem Bereich mit Lizenzbedingungen von Twitter, fünf `conditionalPanel(...)` und einem `mainPanel(...)`. Der Bereich mit den Lizenzbedingungen von Twitter und den `conditionalPanel(...)` wird seinerseits innerhalb des `tags$div(...)` namens „Bedienung“ zusammengefasst, damit dieser Bereich gemeinsam als HTML-DIV-Box im CSS angesprochen werden kann, ohne die anderen Bereiche zu beeinflussen. Auf diese Weise ist es möglich, die Visualisierung später neben den Bedienelementen für die Einstellungen zu platzieren, sofern die Fensterbreite des Nutzers dies zulässt.

Das `sidebarPanel(...)` beinhaltet ein Auswahlfeld, `selectInput(...)`, mit fünf verschiedenen Auswahlmöglichkeiten. Dieses heißt „auswahl“ und ermöglicht dem Nutzer die Auswahl einer Tätigkeit. So kann der Nutzer entscheiden, ob er aktuelle Daten von Twitter empfangen und visualisieren will, einen bestehenden Datensatz anzeigen oder mit aktuell empfangenen Daten von Twitter vergleichen lassen will, oder zwei bestehende Datensätze miteinander vergleichen lassen will. Abhängig von der gewählten Tätigkeit wird dann das entsprechende `conditionalPanel(...)` in der UI abgebildet, das die jeweiligen Einstellungsmöglichkeiten für die jeweils ausgewählte Tätigkeit enthält.

Der Bereich mit Lizenzbedingungen von Twitter wird ebenfalls mit `tags$div(...)` erzeugt und enthält Check-Boxen mit den Verlinkungen zu den entsprechenden Bedingungen von Twitter. Nur wenn der Nutzer jede dieser Boxen aktiviert und mit dem `actionButton(„zustimmung“, „Ich habe die Bedingungen gelesen und stimme ihnen zu.“)` bestätigt hat, werden die `conditionalPanel(...)` angezeigt.

Die `conditionalPanel(...)` sind Bereiche, die abhängig von einer Bedingung ein- oder ausgeblendet werden. In dieser Webanwendung ist die Bedingung für das Anzeigen der einzelnen Bereiche, dass allen Lizenzbedingungen zugestimmt worden ist und jeweils eine der möglichen Tätigkeit ausgewählt worden ist. Das erste `conditionalPanel(...)` bietet Auswahlmöglichkeiten für die Bundesländer, um das ausgewählte Bundesland in der Karte darzustellen. Dieses Panel wird bei jeder Tätigkeit genutzt, sofern nicht „...“ und somit keine Tätigkeit ausgewählt ist. Die anderen `conditionalPanel(...)` werden

nur bei jeweils einer der Tätigkeiten dem Nutzer zur Verfügung gestellt. Innerhalb der `conditionalPanel(...)` werden nämlich Angaben, wie eine Farbauswahl für die Tweets in der Karte, mittels `colourInput(...)`, und der Angabe, wie lange ein Abruf der Twitter-Daten ausgeführt werden soll, mittels `numericInput(...)`, oder eine Funktion zum Laden der Datensätze als `fileInput(...)` zur entsprechenden Einstellung angeboten. Jede diese Einstellungsmöglichkeiten besitzt eine so genannte „id“, um ihre Werte jeweils mit der „server.R“ verarbeiten zu können. Darüber hinaus gibt es zu jeder einen entsprechenden Begleitsatz, der dem Nutzer erklärt, was eingestellt werden kann. Dieser wird mittels `helpText(...)` implementiert. Zusätzlich enthält jedes `conditionalPanel(...)` noch einen `actionButton(...)`, um die Visualisierung beziehungsweise den Abruf der Twitter-Daten und die Visualisierung der empfangenen Daten zu starten. Außerdem enthalten die Panels, die die Einstellungsmöglichkeiten für das Streamen der Twitter-Daten bereitstellen, einen `downloadButton(...)` zum Herunterladen der empfangenen Daten. Auf diese Weise bietet jedes der `conditionalPanel(...)` abhängig von der gewünschten Tätigkeit passende Einstellungsmöglichkeiten für den Nutzer. Mit diesen kann er die Visualisierung steuern.

Die Visualisierung selbst erfolgt im `mainPanel(...)`. Für die Ausgabe der Karte wird die Funktion `leafletOutput(„MapPlot1“)` verwendet, während der Text, der sich abhängig vom gewählten Bundesland und der Anzahl der Tweets⁶⁹ dynamisch anpasst, mit der Funktion `tags$div(id=„Text“, htmlOutput(„dynText“))` ausgegeben wird. Der Inhalt dieser Ausgaben wird dabei von der „server.R“ erzeugt.

7.2 „server.R“ der Shiny-App

Auch in der „server.R“ werden zunächst die benötigten Packages und weiteren R-Skripte, wie bereits in der „ui.R“, eingeladen. Erst dann beginnt die Programmierung des Servers.

Als Erstes findet die Authentifizierung der App bei Twitter mit Hilfe des Scripts „twitter-Authentifizierung.R“ statt. Dann wird die CSV-Datei mit den Einwohnerzahlen eingelesen, sowie eine Variable mit der Einwohnerzahl von ganz Deutschland belegt. Dies alles geschieht in dem Bereich, der nur einmal beim Start der App abläuft⁷⁰. Die Elemente stehen dann im weiteren Verlauf zur Verfügung.

Als Nächstes wird die Funktionalität des `shinyServer(function(input, output, session){...})` implementiert. Dieser begründet den Teil der Anwendung, der immer ausgeführt wird, wenn ein Nutzer die App verwendet, während der Inhalt jeder Funktion von `output$AusgabeNameInUI` immer durchlaufen wird, wenn für die Funktion benötigte Werte von Widgets verändert werden⁷¹. In diesem Bereich, der immer ausgeführt wird, wenn ein Nutzer die App verwendet, werden anfangs die drei booleschen Variablen, „gestreamtEinzel“, „gestreamtVergleich“ und „zustimmungErteilt“, als `reactiveVal(...)` mit „FALSE“ initialisiert. Eine `reactiveVal(...)` kann im Gegensatz zu einfachen Variablen in den nachfolgenden Funktionen verändert und auf Veränderungen hin überwacht werden (vgl. Chang, W., et al., 2017, S. 114). Diese drei Variablen werden benötigt, um festzustellen, ob der Nutzer den Lizenzbedingungen zugestimmt hat und ob der Abruf der Twitter-Daten beendet ist.

69 Siehe dazu Kapitel 5.3.

70 Siehe dazu Kapitel 6.4.2.

71 Siehe dazu Kapitel 6.4.2.

Ob der Nutzer den Lizenzbedingungen zugestimmt hat, wird mit dem nächsten Code-Abschnitt überwacht. Wenn der Nutzer den Button zur Zustimmung drückt, wird überprüft, ob auch alle vier Bedingungen bestätigt worden sind. Ist dies geschehen, wird der gesamte Block für die Zustimmung aus der UI entfernt und die `reactiveVal(...)` namens „zustimmungErteilt“ wird mit „TRUE“ belegt. Hat der Nutzer einer oder mehr der Bedingungen nicht zugestimmt, erhält „zustimmungErteilt“ den Wert „FALSE“ und der Button für die Zustimmung wieder auf „0“ zurückgesetzt, indem er entfernt und direkt wieder neu eingefügt wird. Dies ist notwendig, damit die `conditionalPanel(...)` erst dann angezeigt werden, wenn der Nutzer die Bedingungen akzeptiert. Ohne das Zurücksetzen des Buttons wäre dies nicht möglich, da der initiale Wert eines Buttons in Shiny „0“ lautet und bei jeder Verwendung um eins erhöht wird (vgl. Chang, W., et al., 2017, S. 8).

Auch die Verlinkungen der Bedingungen erfolgt in der „server.R“. Damit diese in der UI als Links dargestellt werden, wird ein entsprechendes HTML-Element in der „server.R“ erzeugt und an die UI geleitet. Dies wird beispielsweise durch `output$TwitterTermsOfService <- renderUI({tagList(a(„Twitter Terms of Service“, href=“https://twitter.com/tos“))})` erreicht.

Die Karte, die in der UI visualisiert wird, wird über den `output$MapPlot1` mit der Benutzeroberfläche verbunden und über die Funktion `renderLeaflet({...})` für die UI gerendert. Dabei wird die Karte über die Funktion `startKarte(input)` des Scripts „startKarte.R“ erstellt.

Die Fokussierung des Kartenausschnitts auf ein bestimmtes Bundesland ist abhängig davon, welches Bundesland der Nutzer auswählt. Mit `observeEvent(input$bundesland, {...})` wird dies vom Server überwacht. Zur Fokussierung wird dann das Script „entscheidungKarte.R“ mit `entscheidungKarte(input)` aufgerufen. Diese Fokussierung wird auch erbracht, wenn der Nutzer die Auswahl der gewünschten Aktivität ändert.

Die Punkte an den geografischen Koordinaten der Tweets, die als Marker bezeichnet werden, werden je nach gewählter Aktivität in die Karte eingetragen. Dazu wird der jeweilige Button für die Bestätigung mittels `observeEvent(input$ButtonName, {...})` beobachtet und jeweils als erstes alle eventuell bestehenden Marker gelöscht. Dann erfolgt die jeweilige Aktion und das Einfügen der zugehörigen Marker in der Karte.

Sollen nur aktuelle Daten von Twitter empfangen, also gestreamt, und in die Karte eingezeichnet werden, wird die Funktion des Scripts „react.R“ dazu aufgerufen. Anschließend wird die `reactiveVal(...)` namens „gestreamtEinzel“ als „TRUE“ deklariert.

Wenn der Stream mit Daten aus einer CSV-Datei verglichen werden soll, wird dagegen zunächst geprüft, ob eine Datei vom Nutzer hochgeladen worden ist. Ist keine Datei hochgeladen, gibt es eine Fehlermeldung in der UI mittels des R-Scripts „fehlermeldung.R“. Ansonsten werden durch das Script „react.R“ aktuelle Daten von Twitter empfangen und in die Karte eingezeichnet. Danach werden die Daten aus der Datei des Nutzers in der Karte mit dem Script „vergleichKarte.R“ eingetragen. Sollte es dabei zu Problemen durch die Datei kommen, wird ebenfalls eine Fehlermeldung mittels des R-Scripts „fehlermeldung.R“ an die UI weitergeleitet.

Für die Aktivitäten „Datensatz anzeigen lassen“ und „Zwei Datensätze vergleichen“ wird nach Prüfung, ob die jeweiligen Dateien hochgeladen worden sind, ebenfalls das Script „vergleichKarte.R“ verwendet, um die Daten aus der jeweiligen Datei des Nutzers in der Karte einzutragen. Ansonsten werden auch hier wieder die entsprechenden Fehlermeldungen erbracht.

Um es dem Nutzer zu ermöglichen, die empfangenen Twitter-Daten herunterzuladen, wird eine Funktion benötigt, die die Daten aus einer empfangenen Datei entsprechend einliest. Dies ist die Funktion `datasetInput1 <- reactive({leseCSV(input$minuten)})` für den einzelnen Stream und die Funktion `datasetInput2 <- reactive({leseCSV(input$minutenVergleich)})` für den Stream im Vergleich mit einer CSV-Datei. Die Funktionen werden im jeweiligen `DownloadHandler(filename = function() {...}, content = function() {...})` verwendet und über das Script „schreibeCSV.R“ wieder als CSV-Datei bereitgestellt. Der Dateiname wird dabei mit Hilfe des Scripts „erzeugeDateinameDownload.R“ und dem Anhängen von „.tweets.csv“ angelegt.

Zum Schluss wird der dynamische Text zusammengesetzt und an die UI gegeben. Dieser berechnet als erstes die Einwohnerzahl des ausgewählten Bundeslandes. Dafür wird mit dem Script „ermittlungBunNr.R“ ausgewertet, welche Spalte der am Anfang als Variable „einwohnerzahl“ eingeladenen CSV-Datei die Einwohnerzahl des Bundeslandes enthält und diese dann abgerufen. Der dynamische wird immer dann neu kreiert, wenn der Nutzer die Auswahl des Bundeslandes wechselt, da es eine Funktion mit `reactive({...})` ist und diese nur bei einer Änderung des Widgets ausgeführt wird (vgl. RStudio, Inc., 2017j).

Anschließend werden die fünf Variablen, „tweetsStream“, „tweetsDatei1“, „tweetsDatei2“, „tweetsDatei3“ und „tweetsDatei4“, jeweils mit „NULL“ initialisiert. Die Werte dieser Variablen werden abhängig von der ausgewählten Tätigkeit, sowie den Werten der `reactiveVal(...)` namens „gestreamtEinzel“ und „gestreamtVergleich“ und den hochgeladenen Dateien, unterschiedlich belegt. Wenn „gestreamtEinzel“ oder „gestreamtVergleich“ mit „TRUE“ belegt sind, bedeutet das, dass Daten von Twitter empfangen und in der Karte eingezeichnet worden sind. Dann werden alle für die Erzeugung des Textes benötigten Variablen, die nicht zu der vom Nutzer gewählten Aktivität gehören⁷², auf ihre initialen Werte zurückgesetzt und die Variable „tweetsStream“ mit den Daten der empfangenen Twitter-Daten belegt. Dazu wird der Dateiname und -pfad mit dem Script „erzeugeDateiname.R“ modelliert und diese Datei eingeladen.

Ähnlich wird auch für die hochgeladenen CSV-Dateien vorgegangen. Beim Anzeigen der Daten aus einer Datei wird die Variable „tweetsDatei1“ mit der zugehörigen Datei belegt, für den Vergleich zwischen Stream und Datei wird „tweetsDatei4“ mit der zugehörigen Datei belegt und „tweetsDatei2“ und „tweetsDatei3“ werden beim Vergleich zweier Dateien mit den zugehörigen Dateien belegt. Hat der Nutzer die Zustimmung zu den Bedingungen am Anfang erteilt und eine Auswahl für die Aktivität getroffen, wird nun der dynamische Text über das R-Script „erzeugeDynText.R“ ausgearbeitet und anschließend in der UI platziert.

⁷² Für den einzelnen Stream werden die Variablen „gestreamtVergleich“, „tweetsDatei1“, „tweetsDatei2“, „tweetsDatei3“ und „tweetsDatei4“ auf ihrer initialen Werte zurückgesetzt, während für den Vergleich eines Streams mit einer CSV-Datei dies mit den Variablen „gestreamtEinzel“, „tweetsDatei1“, „tweetsDatei2“ und „tweetsDatei3“ geschieht.

7.3 Weitere eingebundene R-Skripte in alphabetischer Reihenfolge

7.3.1 entscheidungKarte.R

Das Skript „entscheidungKarte.R“ ist für die Fokussierung des Kartenausschnitts zuständig und wird mit seiner Funktion `entscheidungKarte(input)` aufgerufen. Die benötigte Variable „input“ wird dazu vom Server weitergegeben und enthält alle Widget-Werte.

Am Anfang wird der String „Deutschland“ als Variable „fokussieren“ initialisiert. Hat der Nutzer jedoch ein Bundesland ausgewählt, wird diese Variable mit dem entsprechenden Namen des Bundeslandes belegt.

Im Anschluss werden über das Skript „geoDaten.R“ Koordinaten für den zu fokussierenden Bereich ermittelt und mit diesen der Kartenausschnitt über das Skript „focus.R“ festgelegt.

7.3.2 ermittlungBunNr.R

Mit der Funktion `ermittlungBunNr(bundesland)` wird das Skript „ermittlungBunNr.R“ verwendet. Sie ermittelt eine Zahl, die der Nummer der Spalte in der CSV-Datei der Einwohnerzahlen des ausgewählten Bundeslandes entspricht. Dazu benötigt sie den Namen des ausgewählten Bundeslandes als String und gibt die Nummer zurück.

7.3.3 erzeugeDateiname.R und erzeugeDateinameDownload.R

Die Skripte „erzeugeDateiname.R“ und „erzeugeDateinameDownload.R“ sind sehr ähnlich. Während „erzeugeDateiname.R“ mit `erzeugeDateiname(minuten)` aufgerufen wird, wird „erzeugeDateinameDownload.R“ mit `erzeugeDateinameDownload(minuten)` aufgerufen. Beide benötigen die Angabe, wie lange der Abruf der Twitter-Daten erfolgt ist und beide geben am Ende einen String zurück.

Zum Erstellen des Dateinamens wird unter anderem das Datum verwendet. Dieses wird mit `Sys.Date()` ermittelt und mit `as.Date()` in einen String für den Dateinamen formatiert. Dann wird der String des Dateinamens in der „erzeugeDateiname.R“ aus dem Dateipfad, dem Datum, einer textlichen Bezeichnung, sowie der Minuten für die Dauer des Streams zusammengesetzt. Dadurch entsteht die Form „TwitterStreamDaten/rtweet-stream-YYYY-MM-TT_Stream-Zeit-MINmin“. In der „erzeugeDateinameDownload.R“ wird dagegen der Dateipfad weggelassen und es entsteht die Form „YYYY-MM-TT_TwitterStreamDaten_Stream-Zeit-MINmin“. Der erzeugte String ist dann die jeweilige Rückgabe der Funktion.

7.3.4 erzeugeDynText.R

Der dynamische Text, der in der UI angezeigt wird, wird mit dem R-Skript „erzeugeDynText.R“ gebildet. Dieses benötigt die Twitter-Daten vom Stream beziehungsweise den verwendeten Dateien, das gewählte Bundesland und die Einwohnerzahlen des Bundeslandes, sowie von Deutschland, um eine Liste mit HTML-Elementen, die den dynamischen Text enthalten, zurückzugeben.

Zunächst wird ein String für die Einwohnerzahl in Deutschland deklariert und die Variablen, die später die Strings für die einzelnen Teile der Ausgabe enthalten sollen, mit „NULL“ initialisiert. Dann wird der String für die Einwohnerzahl des Bundeslandes erzeugt, sofern der Nutzer ein Bundesland ausgewählt hat.

Im Anschluss wird getestet, welche der Variablen mit Twitter-Daten, die an das Script übergeben worden sind, Daten enthalten oder „NULL“ sind. Ist beispielsweise gestreamt worden, ist die Variable „tweetsStream“ nicht „NULL“ sondern enthält Daten. Wird ein Stream mit einem Datensatz aus einer CSV-Datei des Nutzers verglichen, sind die Variablen „tweetsStream“ und „tweetsDatei4“ mit Daten belegt. Wird nur ein Datensatz angezeigt, sind diese Daten in der Variable „tweetsDatei1“ gespeichert, während die Variablen „tweetsDatei2“ und „tweetsDatei3“ Daten enthalten, wenn der Nutzer zwei Datensätze miteinander vergleicht.

Abhängig davon, in welcher der Variablen Daten enthalten sind, wird die Anzahl der Tweets für den entsprechenden Datensatz mit der Funktion `zaehleTweets(Daten, bundesland)` ermittelt. Daraus werden dann entsprechende Strings für die Tweet-Anzahl in Deutschland und im gewählten Bundesland erzeugt.

Diese erzeugten Strings, sowie die Strings mit den Einwohnerzahlen werden dann für die Rückgabe in HTML-Elemente eingefügt und mit diesen eine Liste erzeugt. Dann erfolgt die Rückgabe der erzeugten Liste an die aufrufende Variable im Script „server.R“, das diese an die UI weiterleitet. Dort erfolgt die Ausgabe des dynamischen Textes.

7.3.5 fehlermeldung.R

Das Script „fehlermeldung.R“ gibt eine Fehlermeldung in der UI für den Benutzer aus. Abhängig von einer Zahl, die als Fehlercode fungiert, wird dem Nutzer eine entsprechende Meldung angezeigt. Dazu bietet Shiny die Funktion `showNotification(„Text der Fehlermeldung“, duration= Anzeigedauer in Sekunden)` (vgl. Chang, W., et al., 2017, S. 155). Die Zahl für die Fehlermeldung, also der Fehlercode, wird in verschiedenen Scripten der App erzeugt, während der Text der Fehlermeldung den Fehler beschreibt und dieser Text für zwanzig Sekunden in der UI sichtbar ist. Dadurch kann dem Nutzer eine Meldung in Textform angezeigt werden und es erscheint kein für den Nutzer unverständlicher Fehlercode.

7.3.6 focus.R

Die Funktion `focus(geoBundesland)` des Scripts „focus.R“ wird vom Script „entscheidung-Karte.R“ aufgerufen, um den Ausschnitt der Karte auf das ausgewählte Bundesland einzustellen.

Dazu werden zunächst die geografischen Koordinaten des gewählten Bundeslandes, die diese Funktion bei ihrem Aufruf erhält, in vier einzelnen Variablen gespeichert, die die Ecken der Bounding-Box⁷³ für das Bundesland darstellen. Diese sind der nördliche und der südliche Breitengrad⁷⁴, sowie der östliche und der westliche Längengrad⁷⁵ (vgl. Cheng, J., et al., 2017, S. 48).

Dann wird für die Leaflet-Karte die eigens für Shiny entwickelte Leaflet-Funktion `leafletProxy(„MapPlot1“)` verwendet (vgl. Cheng, J., et al., 2017, S. 41), um mittels `fitBounds(westlon, southlat, eastlon, northlat)` den Ausschnitt der Karte auf das Bundesland zu fokussieren. Die Änderung der Karte wird dann zurückgegeben.

⁷³ Eine Bounding-Box ist hier ein Rechteck, dass so eng wie möglich um die Grenzen des Bundeslandes gezogen wird.

⁷⁴ Engl. Latitude.

⁷⁵ Engl. Longitude.

7.3.7 geoDaten.R

Die geografischen Koordinaten des vom Nutzer ausgewählten Bundeslandes werden im Script „geoDaten.R“ ermittelt. Dazu wird die Funktion `lookup_coords()` des Packages „rtweets“ verwendet (vgl. Kearney, M. W., 2017c, S. 29). Diese Funktion ermittelt mittels der Twitter-API Längen- und Breitengrade anhand eines Ortsnamens, wie der Name des gewählten Bundeslandes, und gibt diese Angaben als Objekt zurück (vgl. Kearney, M. W., 2017c, S. 29).

Es kommt jedoch vor, dass die Twitter-API keine Daten zurück liefert und die Anfrage erneut gestellt werden muss. Daher wird diese Funktion maximal zweihundert Mal ausgeführt, so lange keine Daten zurückgeliefert werden. Werden Daten erhalten, werden diese wieder vom Script „geoDaten.R“ an das aufrufende Script zurückgegeben. Andernfalls erhält der Nutzer die Fehlermeldung „Twitter ist zur Zeit nicht erreichbar“, da die App ohne Daten von Twitter nicht funktionstüchtig ist.

7.3.8 karteLeaflet.R

Die Erzeugung der Leaflet-Karte erfolgt über das Script „karteLeaflet.R“ anhand von geografischen Koordinaten.

Wie im Script „focus.R“ werden auch hier zu erst die geografischen Koordinaten in vier Variablen für die Längen- und Breitengrade gespeichert. Dann wird die Karte selbst mit der Funktion `leaflet()`, sowie dem Kartenmaterial „CartoDB.PositronNoLabels“ und „OpenMapSurfer.AdminBounds“ erzeugt und auf die geografischen Koordinaten fokussiert. Diese Karte wird dann an das aufrufende Script zurückgegeben.

7.3.9 leseCSV.R

Mit dem Script „leseCSV.R“ wird die CSV-Datei mit den empfangenen Twitter-Daten als Variable eingeladen. Dazu wird anfangs der Dateiname, der beim Speichern des Streams als Datei entsteht, als String erstellt.

Anschließend wird geprüft, ob die Datei existiert. Existiert sie, wird sie zunächst mit `readCSV()` eingeladen. Dann werden alle nicht benötigten Spalten, wie beispielsweise der Tweet-Inhalt, entfernt. Dies wird dann an das aufrufende Script übermittelt. Existiert die Datei jedoch nicht, wird ein Fehlercode zurückgegeben, mit dem über das Script „fehlermeldung.R“ eine entsprechende Fehlermeldung ausgegeben werden kann.

7.3.10 liegtIn.R

Wenn das Script „liegtIn.R“ aufgerufen wird, prüft es, ob bestimmte Längen- und Breitenkoordinaten innerhalb der Grenzen der Bounding-Box des ausgewählten Bundeslandes liegen. Kommt diese Überprüfung zu dem Ergebnis, dass die Werte innerhalb des Bundeslandes liegen, wird der boolesche Wert „TRUE“ und ansonsten der boolesche Wert „FALSE“ zurückgegeben.

7.3.11 react.R

Mit der Funktion `react(minuten, bundesland, farbe, farbeAußen)` des Scripts „react.R“ wird dem Nutzer eine Meldung über die Dauer bis zum Abschluss des Empfangs der Twit-

ter-Daten ausgegeben, der Stream gestartet und die Daten in der Karte eingezeichnet. Dazu wird die Zeit für den Stream in Sekunden umgerechnet und mittels `showNotification(„Bitte warten: Twitter-Daten werden empfangen.“, duration = streamZeit)` der Hinweis für den Nutzer erstellt.

Dann werden die geografischen Koordinaten des gewählten Bundeslandes mit dem Script „geoDaten.R“ abgerufen und das Empfangen und Einzeichnen der Twitter-Daten in der Karte über das Script „streamZeichnen.R“ gestartet. Sollte es dabei zu einem Fehler kommen, wird eine Fehlermeldung zurückgegeben, die dann mit Hilfe des Scripts „fehlermeldung.R“ dem Nutzer angezeigt wird.

7.3.12 schreibeCSV.R

Das Script „schreibeCSV.R“ erzeugt eine CSV-Datei für den Nutzer. Dazu benötigt sie die Rückgabe des Scripts „leseCSV.R“⁷⁶ und den Dateipfad⁷⁷ der zu erzeugenden Datei. Ist vom Scripts „leseCSV.R“ eine Datei fehlerfrei eingelesen worden, enthält die Rückgabe dieser den Fehlercode „0“ und die Datei kann mit `write.csv(datei[2], file)` erstellt werden. Ansonsten wird der Fehlercode mit Hilfe des Scripts „fehlermeldung.R“ in eine Fehlermeldung für den Nutzer umgewandelt und ausgegeben.

7.3.13 sortiereDaten.R

Um die Daten der Tweets danach zu sortieren, ob sie innerhalb oder außerhalb des gewählten Bundeslandes liegen, wird das Script „sortiereDaten.R“ verwendet.

Dabei müssen die drei Sonderfälle der Stadtstaaten Berlin, Bremen und Hamburg, berücksichtigt werden. Diese liegen nämlich innerhalb der Bounding-Boxen anderer Bundesländer. So liegt Berlin innerhalb von Brandenburg, Hamburg innerhalb der Bounding-Box Schleswig-Holsteins und Bremen und Hamburg innerhalb der Bounding-Box Niedersachsens. Wird also das sie umgebende Bundesland ausgewählt, müssen die Tweets innerhalb des jeweiligen Stadtstaats als außerhalb des ausgewählten Bundeslandes bewertet werden. Dazu werden in diesem Fall im ersten Schritt die Koordinaten der Stadtstaaten ermittelt.

Unabhängig von den Sonderfällen erfolgt dann die Vorbereitung zweier Variablen, die für die Sortierung nötig sind. Diese sind einerseits die Variable „bundeslandTweets“, die anfangs noch alle Tweets, genau wie die übergebene Variable „deutschlandTweets“, enthält, und andererseits die Variable „loeschen“, die als leeres Array initialisiert wird.

Anschließend wird in einer „for“-Schleife mit dem Script „liegtIn.R“ für jeden einzelnen Tweet des Datensatzes ermittelt, ob die Koordinaten innerhalb des Bundeslandes liegen, und als boolescher Wert gespeichert. Dabei werden auch die Sonderfälle über die entsprechenden „if“-Abfragen berücksichtigt. Der ermittelte boolesche Wert wird dann jeweils am Ende des Arrays „loeschen“ eingefügt.

Nach der „for“-Schleife werden durch die Werte des Arrays „loeschen“ dann alle Daten, die nicht innerhalb des gewählten Bundeslandes liegen, und somit im Array „loeschen“ mit „FALSE“ deklariert werden, aus der Variable „bundeslandTweets“ entfernt.

⁷⁶ Diese Rückgabe ist benannt mit dem Variablennamen „datei“.

⁷⁷ Der Dateipfad liegt in der Variable „file“.

Anschließend werden aus der Variable „deutschlandTweets“ dagegen alle Tweets entfernt, die innerhalb des gewählten Bundeslandes liegen. Zum Schluss werden diese beiden Variablen an das aufrufende Script zurückgegeben.

7.3.14 spaltennamenKorrigieren.R

Die Korrektur der Spaltennamen ist notwendig, weil die CSV-Datei, die vom Nutzer beim Streamen heruntergeladen werden kann, beim Wiedereinladen zum Anzeigen beziehungsweise Vergleichen des Datensatzes ein „daten.“ vor jedem Spaltennamen stehen haben. Dieses muss für die Verarbeitung in der App entfernt werden.

Dazu werden mit dem Script „spaltennamenKorrigieren.R“ diese Spaltennamen in ein String-Array eingeladen. Dasselbe wird mit den Spaltennamen gemacht, die für die Verarbeitung benötigt werden. Dann werden sie im Datensatz mit Hilfe der Funktion `setnames()` korrigiert und der korrigierte Datensatz zurückgegeben.

7.3.15 startKarte.R

Das Script „startKarte“ erzeugt die Ansicht der Karte bevor der Nutzer Eingaben innerhalb einer der Aktivitäten getätigt hat. Dazu prüft es, ob die Lizenzbedingungen akzeptiert worden sind und der Nutzer generell eine Aktivität ausgewählt hat.

Trifft beides zu, werden die geografischen Koordinaten mittels des Scripts „geoDaten.R“ ermittelt und die Karte mit dem Script „karteLeaflet.R“ erstellt.

7.3.16 streamGeo.R

Das Abrufen und Speichern der Twitter-Daten erfolgt durch das Script „streamGeo.R“. Dazu benötigt es die Informationen, welches Bundesland ausgewählt worden ist, wie lange der Stream laufen soll, einmal in Minuten⁷⁸ und einmal in Sekunden⁷⁹, und wie die geografischen Koordinaten des gewählten Bundeslandes lauten.

Als erstes wird dafür der Dateiname und -pfad für das Speichern der Daten mit der Funktion `erzeugeDateiname(minuten)` erzeugt. Zusätzlich werden für den Stream die geografischen Koordinaten von Deutschland mit dem Script „geoDaten.R“ abgerufen.

Dann wird der Stream selbst mit dem Befehl `stream_tweets(geoDeutschland, timeout = streamZeit, parse = TRUE, file_name = dateinameFuerStream)` gestartet. Dieser Stream wird nach Ablauf der Stream-Zeit wieder geschlossen und alle in dieser Zeit empfangenen Daten werden unter dem angegebenen Dateinamen und -pfad als JSON-Datei gespeichert.

Danach wird geprüft, ob der Stream auch tatsächlich Daten enthält, da es bei kurzen Stream-Zeiten manchmal vorkommen kann, dass keine Tweets in dieser Zeit veröffentlicht worden sind. Enthält der Stream keine Daten, wird ein Fehlercode zurückgegeben. Ansonsten erfolgt die Verarbeitung der Daten.

Bei der Verarbeitung der Daten müssen zunächst explizit die Längen- und Breitengrade der empfangenen Tweets von Twitter abgerufen werden, da diese nicht automatisch im

⁷⁸ Für die Dateibenennung.

⁷⁹ Für das Einstellen des Streams.

Stream enthalten sind. Danach werden alle Tweets entfernt, die nicht die Länderkennung⁸⁰ „DE“ haben, denn durch die Nutzung einer Bounding-Box kann es im Grenzbe-
reich dazu kommen, dass auch Tweets empfangen werden, die Nahe der deutschen
Grenze in den Nachbarländern veröffentlicht worden sind. Diese werden daher mit dem
Befehl `subset(deutschlandTweets, country_code == „DE“)` entfernt.

Danach wird der Dateiname und -pfad der gespeicherten JSON-Datei mit den
Stream-Daten in einer Variable gespeichert, um diesen später mit zurückgeben zu kön-
nen. Anschließend wird ein Dateiname und -pfad für eine CSV-Datei erzeugt. Diese
werden benötigt, um dann die Daten des Streams mit der Funktion `save_as_csv(deutsch-
landTweets, dateinameDerCSV)` als CSV-Datei zu sichern.

Nun werden noch die Tweets aus dem Stream mit dem Script „sortiereDaten.R“ danach
sortiert, ob sie innerhalb oder außerhalb des gewählten Bundeslandes liegen und an-
schließend erfolgt die Rückgabe aller relevanten Daten an das aufrufende Script.

7.3.17 streamZeichnen.R

Zum Zeichnen der Marker für die Tweets eines Streams in der Karte wird die Funk-
tion `streamZeichnen(bundesland, streamZeit, geoBundesland, farbe, farbeAußen, minuten)` der
„streamZeichnen.R“ verwendet. Diese ruft als erstes das Script „streamGeo.R“ auf, um
die Stream-Daten zu erhalten.

Sofern kein Fehler aufgetreten ist, der Stream also Daten enthält, werden mit dem Script
„zeichnen.R“ erst die Tweets für das ausgewählte Bundesland und dann die Tweets für
den Rest von Deutschland in der jeweils vom Nutzer gewählten Farbe eingezeichnet.
Außerdem wird ein Fehlercode an das aufrufende Script zurückgegeben, der „0“ lautet,
wenn kein Fehler vorliegt.

7.3.18 twitterAuthentifizierung.R

Die OAuth-Authentifizierung bei der Twitter-API erfolgt mit dem Script „twitterAuthenti-
fizierung.R“. Dazu werden die Daten, die bei der Registrierung der App von Twitter er-
halten werden, als einzelne Variablen gespeichert. Dies sind der Name mit dem die App
bei Twitter registriert ist, ein Schlüssel namens „consumer_key“ und ein Geheimnis, das
„consumer_secret“ genannt wird (vgl. Kearney, 2017c, S. 5).

Dann erfolgt die Authentifizierung mit der Funktion `create_token(app, consumer_key, consu-
mer_secret)` des R-Package „rtweet“. Läuft die App lokal und nicht im Internet funktioniert
die Authentifizierung mit dieser Funktion nur, wenn bei der Registrierung die Einstellung
der „Callback URL [auf] http://127.0.0.1:1410“ (Kearney, 2017c, S. 5) gesetzt worden
ist. Durch die Funktion erhält man ein sogenanntes „Token“ (vgl. Kearney, 2017c, S. 5),
das gespeichert werden kann. Diese Speicherung als Datei erfolgt im nächsten Schritt.
Außerdem wird dieses „Token“ an das aufrufende Script zurückgegeben.

⁸⁰ Spaltenname „country_code“.

7.3.19 vergleichZeichnen.R

Das Script „vergleichZeichnen.R“ dient dem Zeichnen der Marker von Tweets aus einer hochgeladenen CSV-Datei. Dazu wird über `tryCatch(..., error = function(e){...})` abgefangen, ob die Datei leer ist und in diesem Fall ein entsprechender Fehlercode an das aufrufende Script zurückgegeben. Ansonsten wird sie als Variable eingeladen und verarbeitet.

Sofern in der Datei Längen- und Breitengrade angegeben sind, wird kein Fehlercode zurückgegeben, sondern es erfolgt die Verarbeitung. Zunächst werden die geografischen Koordinaten mit dem Script „geoDaten.R“ ermittelt und die Spaltennamen mit dem Script „spaltennamenKorrigieren.R“ angepasst.

Anschließend werden auch hier die Tweets mit dem Script „sortiereDaten.R“ sortiert und erst die Marker für das gewählte Bundesland, sowie für den Rest von Deutschland mit dem Script „zeichnen.R“ in die Karte eingezeichnet. Danach wird der Fehlercode „0“ zurückgegeben, wenn kein Fehler aufgetreten ist.

7.3.20 zaehleTweets.R

Mit dem Script „zaehleTweets.R“ werden die Tweets gezählt. Dazu benötigt die Funktion eine CSV-Datei mit den Daten der Tweets⁸¹ und den Namen des gewählten Bundeslandes.

Sofern die Datei Daten enthält, werden die geografischen Koordinaten des Bundeslandes ermittelt und die Daten der Datei mit dem Script „sortiereDaten.R“ sortiert. Außerdem werden alle Daten entfernt, die nicht aus Deutschland⁸² stammen.

Anschließend wird die Länge des Datensatzes für das gewählte Bundesland ermittelt. Dies ist die Anzahl der Tweets für dieses Bundesland. Die Anzahl der Tweets für ganz Deutschland ergibt sich dann aus der Summe der Länge des Datensatzes der Tweets außerhalb des gewählten Bundeslandes und der Anzahl der Tweets für dieses Bundesland. Diese ermittelten Werte werden dann beide an das aufrufende Script zurückgegeben.

7.3.21 zeichnen.R

Die Funktion `zeichnen(lng, lat, farbe)` fügt die Marker in Form von Punkten an den gegebenen geografischen Koordinaten der Karte hinzu. Dazu benötigt sie den Längen- und den Breitengrad an dem der Marker gesetzt werden soll, sowie die gewählte Farbe für den Marker. Diese werden in die Funktion `addCircleMarkers(lng = lng, lat = lat, radius = 5, color = farbe)` eingesetzt, um über `leafletProxy(„MapPlot1“)` die Koordinaten in der Karte darzustellen.

81 Die CSV-Datei kann sowohl direkt vom Stream (ohne heruntergeladen zu sein) stammen, oder eine vom Nutzer hochgeladene Datei sein.

82 `country_code == „DE“`

8 Beispiele für die Verwendung und Auswertung der App

Die entwickelte Webanwendung kann verwendet werden, um unterschiedliche Fragestellungen zu den Orten der Digitalisierung in Deutschland durch Tweets, zu untersuchen und zu visualisieren. Eine Fragestellung kann zum Beispiel sein, zu welcher Tageszeit die Digitalisierung am stärksten ist, also wann und wo die meisten Tweets veröffentlicht worden sind, oder an welchem Tag innerhalb eines Zeitraumes die Digitalisierung am stärksten ist und wo sie stattfindet.

Diese Fragen werden im Folgenden für die Beispiele Berlin und Nordrhein-Westfalen (NRW) mit der Webanwendung visualisiert und untersucht. Dazu werden viermal täglich⁸³ an sieben aufeinander folgenden Tagen⁸⁴ die Twitter-Daten für eine Stunde⁸⁵ gestreamt und visualisiert⁸⁶, beziehungsweise visuell verglichen.

8.1 Tageszeiten mit der stärksten Digitalisierung durch Tweets

Zur Untersuchung der Frage, zu welcher Tageszeit der Tage vom 07.12.2017 bis zum 13.12.2017 die Digitalisierung durch Tweets am stärksten ist, werden die Streams der jeweiligen Tageszeit aller Tage zu einer CSV-Datei⁸⁷ zusammengefasst und anschließend mit der Webanwendung für Berlin und NRW grafisch aufbereitet. Währenddessen ist die Tabelle 8.1 eine Übersicht über die Anzahl aller Tweets⁸⁸, die visualisiert werden.

Tageszeit	Deutschland	Berlin	Nordrhein-Westfalen
morgens	8117	1459	1486
mittags	9444	1573	2024
nachmittags	10963	1711	2396
abends	12313	1887	2948

Tabelle 8.1: Anzahl der Tweets pro Tageszeit vom 07.12.2017 bis 13.12.2017

Quelle: eigene Darstellung

Da mit der Webanwendung jeweils nur zwei Dateien miteinander vergleichbar sind, werden der Morgen in blau und der Mittag in grün jeweils in einer Infografik zusammengefasst, sowie der Nachmittag in rot und der Abend in gelb⁸⁹ in einer weiteren Infografik. Dazu wird als erste Datei der Morgen beziehungsweise der Nachmittag und als zweite Datei der Mittag beziehungsweise der Abend eingeladen. Dabei werden die Punkte der ersten Datei grundsätzlich über die der zweiten Datei gezeichnet. Im Anschluss werden dann die Tageszeiten aus der ersten und zweiten Visualisierung, die jeweils mehr Tweets enthalten, gemeinsam in einer Karte dargestellt. Dabei werden die Daten der zuerst eingeladenen Datei immer als erstes in die Visualisierung eingezeichnet. Auf diese Weise kann aufgezeigt werden, zu welcher Tageszeit die Visualisierung am stärksten ist.

In den Abbildungen 8.1 und 8.2 sind die Tweets zu den jeweiligen Tageszeiten für Berlin zu sehen. Dabei wird deutlich, dass sich die Tweets zu jeder Tageszeit hauptsächlich auf das Zentrum von Berlin konzentrieren und nur wenige in den Randbereichen zu finden

⁸³ Der Scann startet jeweils morgens zwischen 08.30 Uhr und 08.45 Uhr, Mittags zwischen 11.30 Uhr und 11.45 Uhr, am späten Nachmittag zwischen 16.30 Uhr und 16.45 Uhr, sowie abends zwischen 20.30 Uhr und 20.45 Uhr.

⁸⁴ Vom Donnerstag den 07.12.2017 bis einschließlich zum Mittwoch den 13.12.2017.

⁸⁵ Dazu wird die Einstellung für die Stream-Zeit in der Anwendung auf 60 Minuten gestellt.

⁸⁶ Siehe Anhang C für die Visualisierungen jedes einzelnen Streams.

⁸⁷ Ein Beispiel für eine CSV-Datei eines einzelnen Tages zu einer einzigen Tageszeit ist in Anhang E zu finden.

⁸⁸ Die Daten stammen aus Visualisierungen der Streams, da diese Tabelle nicht mit der App erzeugt werden kann.

⁸⁹ Farbwerte: Morgen = #0000FF (blau), Mittag = #00FF00 (grün), Nachmittag = #FF0000 (rot), Abend = #FFF700 (gelb)

sind. Demnach ist die Digitalisierung im Zentrum von Berlin allgemein stärker als in den Randbereichen. Darüber hinaus gibt es in Abbildung 8.1 mehr grüne als blau Punkte, was durch den Text darunter bestätigt wird. Daher ist die Digitalisierung am Mittag stärker als am Morgen. In Abbildung 8.2 gibt es dagegen mehr gelbe Punkt. Dies bedeutet, dass die Digitalisierung am Abend stärker als am Nachmittag ist.

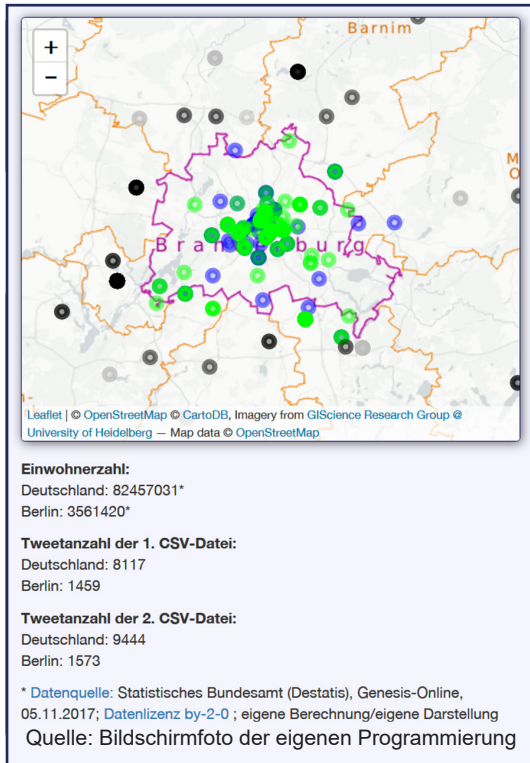


Abbildung 8.1: Tweets in Berlin morgens (blau) und mittags (grün)

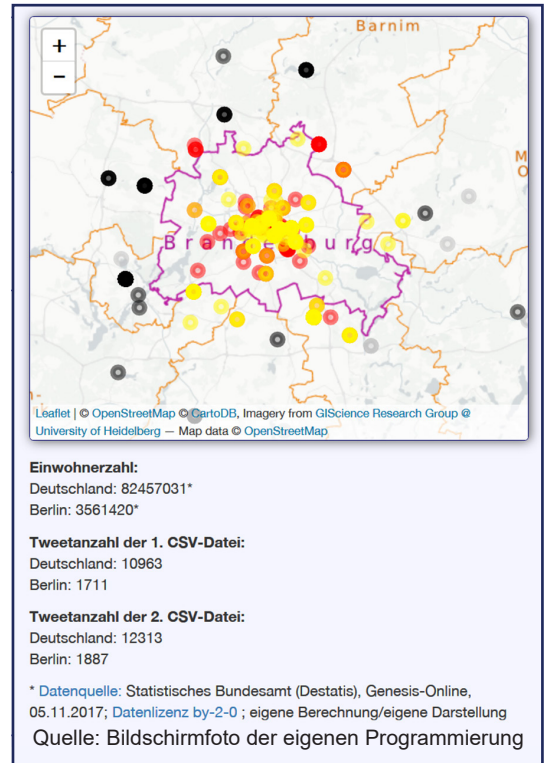


Abbildung 8.2: Tweets in Berlin nachmittags (rot) und abends (gelb)

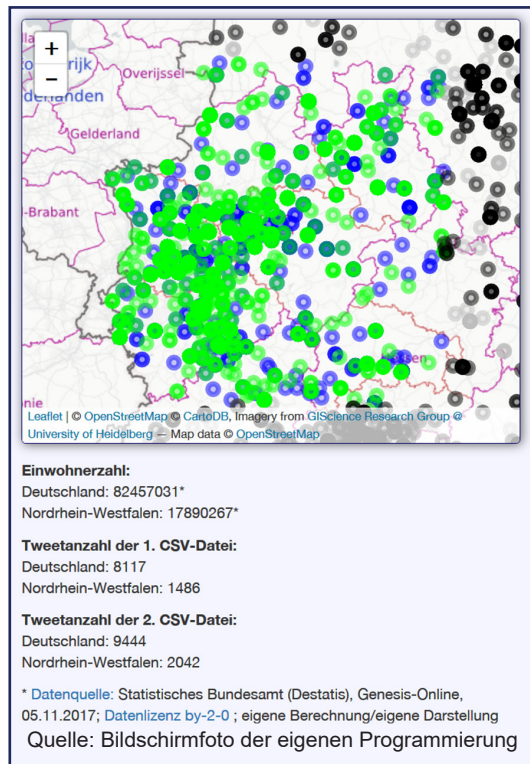


Abbildung 8.3: Tweets in NRW morgens (blau) und mittags (grün)

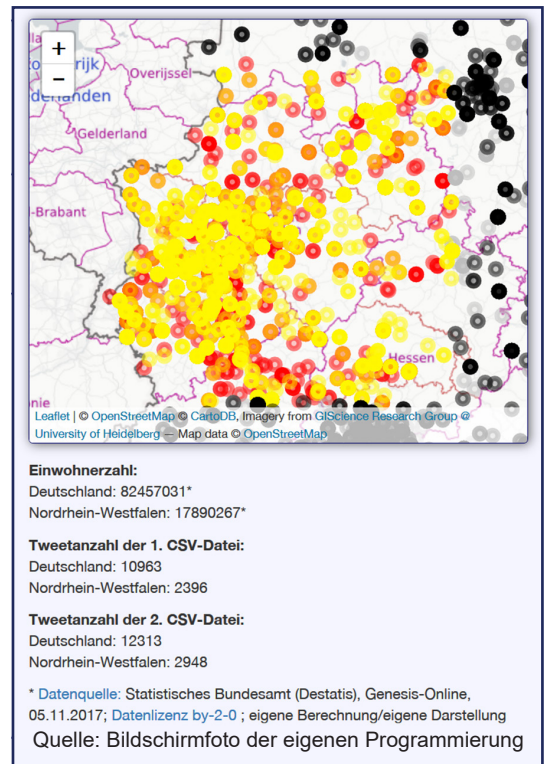


Abbildung 8.4: Tweets in NRW nachmittags (rot) und abends (gelb)

Ähnlich gilt dies auch für NRW. Auch hier ist, den Abbildungen 8.3 und 8.4 zur Folge, die Digitalisierung am Mittag stärker als am Morgen und am Abend stärker als am Nachmittag. Hier konzentrieren sich die Tweets jedoch nicht auf das Zentrum von NRW, sondern sie kommen häufiger auf der linken Seite des Bundeslandes vor. Im Allgemeinen sind die Tweets jedoch ansonsten ziemlich gleichmäßig verteilt. Demnach ist die Digitalisierung in NRW annähernd gleichstark und überwiegt lediglich etwas auf der linken Seite. In dem Bereich in dem die Tweets etwas mehr sind und somit die Digitalisierung etwas stärker ist, liegen viele Großstädte, wie Düsseldorf, Essen und Köln. Daher lässt sich daraus schließen, dass die Digitalisierung in den Großstädten stärker stattfindet, als in anderen Bereichen. Dies kann daran liegen, dass in Großstädten mehr Menschen auf engerem Raum leben.

Wie bereits erwähnt gibt es in Berlin und in NRW jeweils am Mittag und am Abend mehr Tweets. Daher werden diese jeweils mit der Webanwendung noch einmal visuell verglichen. Die Ergebnisse davon sind für Berlin in Abbildung 8.5 und für NRW in Abbildung 8.6 zu sehen. Dabei wird deutlich, dass es jeweils abends die meisten Tweets gibt. Somit ist die Digitalisierung abends am stärksten.

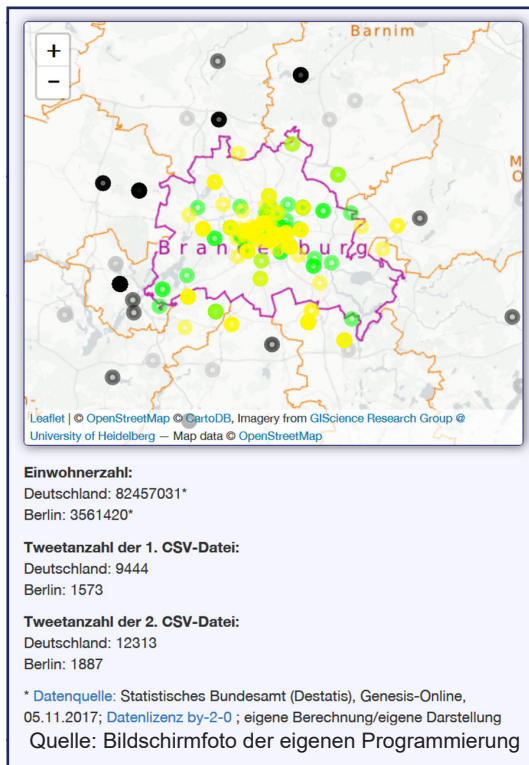


Abbildung 8.5: Tweets in Berlin
mittags (grün) und abends (gelb)

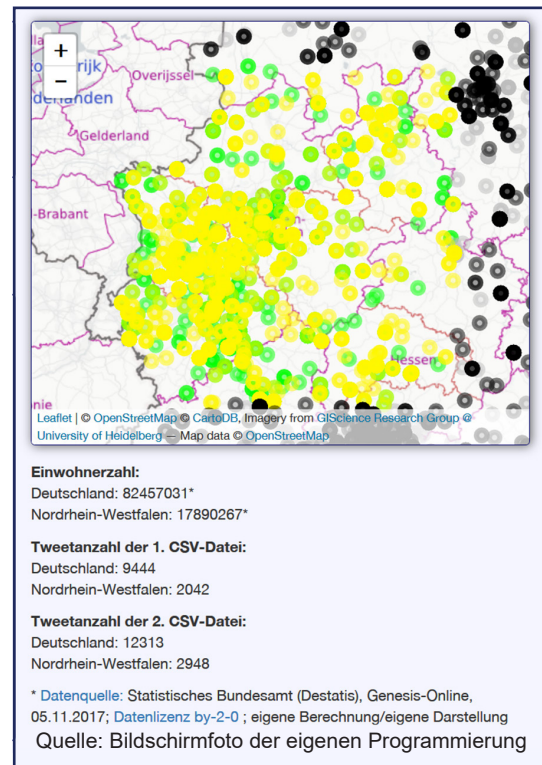


Abbildung 8.6: Tweets in NRW
mittags (grün) und abends (gelb)

8.2 Tag mit der stärksten Digitalisierung durch Tweets

Die Frage, an welchem der Tage vom 07.12.2017 bis zum 13.12.2017 die Digitalisierung in Berlin beziehungsweise in NRW durch Tweets am stärksten ist, kann mit einer ähnlichen Vorgehensweise, wie in Kapitel 8.1 beantwortet und visualisiert werden.

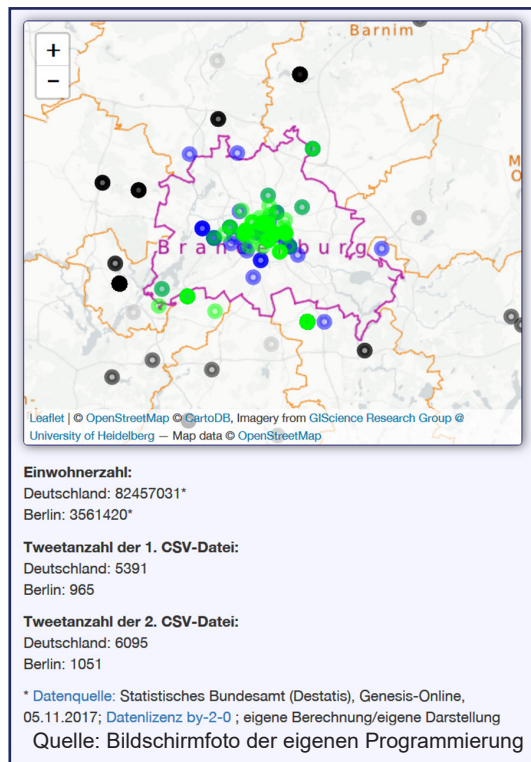
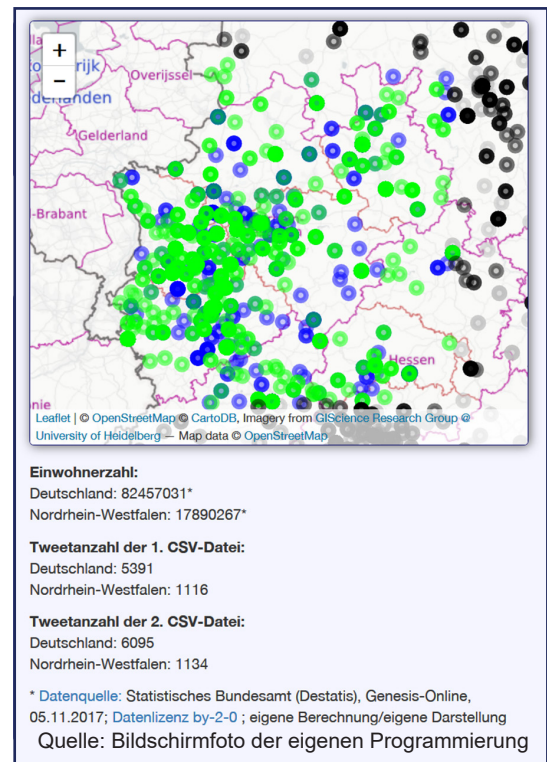
In diesem Fall werden alle Tageszeiten pro Tag zu einer CSV-Datei zusammen gefasst. Dadurch entstehen sieben Dateien, die in einem Infografik verglichen werden. Auch hier ist eine generelle Übersicht zur Anzahl der Tweets an den jeweiligen Tagen in Tabelle 8.2 aufgeführt.

Tag	Deutschland	Berlin	Nordrhein-Westfalen
07.12.2017	5391	965	1116
08.12.2017	6095	1051	1134
09.12.2017	5643	889	1252
10.12.2017	7024	920	1670
11.12.2017	5588	855	1188
12.12.2017	5436	964	1219
13.12.2017	5660	986	1293

Tabelle 8.2: Anzahl der Tweets pro Tag vom 07.12.2017 bis 13.12.2017

Quelle: eigene Darstellung

Es können auch hier immer nur zwei Dateien in einer Infografik mit Hilfe der Webanwendung dargestellt werden. Daher erfolgt die Visualisierung der Daten in drei Stufen. Zu erst wird der 07.12.2017 mit dem 08.12.2017, der 09.12.2017 mit dem 10.12.2017, der 11.12.2017 mit dem 12.12.2017 und der 13.12.2017 mit dem 07.12.2017 in jeweils unterschiedlichen Farben⁹⁰ mit der angegebenen Reihenfolge visualisiert. Dabei entstehen für Berlin und NRW jeweils die vier Visualisierung, die in den Abbildungen 8.7 bis 8.14 zu sehen sind. Aus diesen Visualisierungen ergibt sich, dass es jeweils am 08.12.2017, am 10.12.2017, am 12.12.2017 und am 13.12.2017 mehr Tweets gibt als an den Tagen mit denen sie verglichen werden. Somit ist die Digitalisierung an diesen Tagen stärker. Dies gilt sowohl für Berlin als auch für NRW. Es fällt jedoch auch auf, dass sich die jeweilige Anzahl der Tweets in Berlin am 07.12.2017 und am 13.12.2017 nur wenig unterscheiden. Mit der Infografik allein ist dieser Unterschied kaum sichtbar. Dadurch wird deutlich, dass die Auflistung unterhalb der Infografik wichtig ist, um die Visualisierung richtig auswerten zu können. Durch diese wird ersichtlich, dass es am 07.12.2017 965 Tweets gibt, während es am 13.12.2017 986 Tweets sind und somit ist die Digitalisierung an dem Tag etwas stärker ist.

**Abbildung 8.7:** Tweets in Berlin am 07.12.2017 (blau) und am 08.12.2017 (grün)**Abbildung 8.8:** Tweets in NRW am 07.12.2017 (blau) und am 08.12.2017 (grün)

90 Farbwerte: 07.12.2017 = #0000FF (blau), 08.12.2017 = #00FF00 (grün), 09.12.2017 = #FF0000 (rot), 10.12.2017 = #FFF700 (gelb), 11.12.2017 = #6808D6 (lila), 12.12.2017 = #FF00C4 (pink), 13.12.2017 = #FFAA20 (orange)

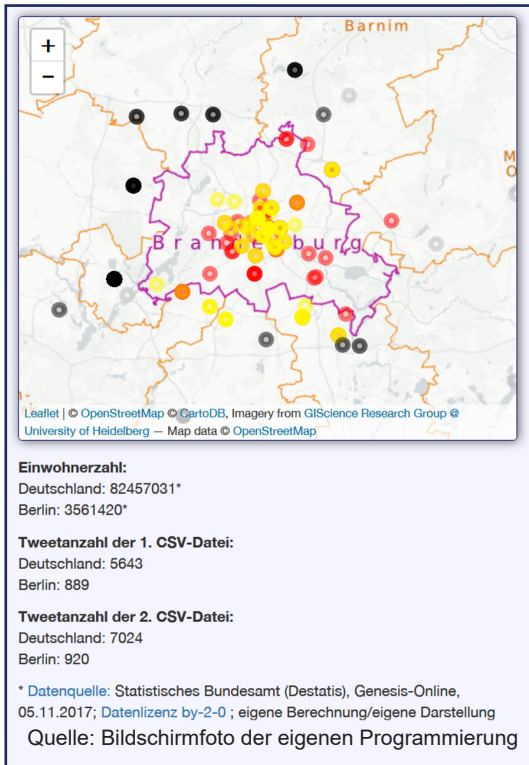


Abbildung 8.9: Tweets in Berlin am 09.12.2017 (rot) und am 10.12.2017 (gelb)

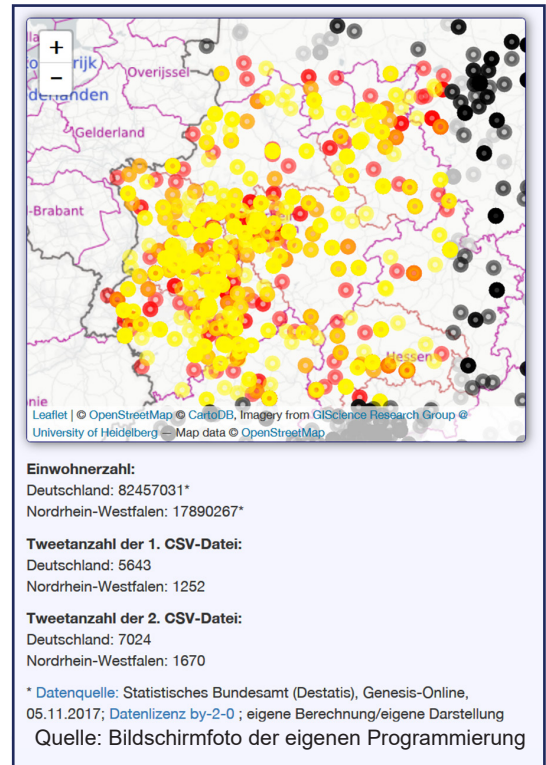


Abbildung 8.10: Tweets in NRW am 09.12.2017 (rot) und am 10.12.2017 (gelb)

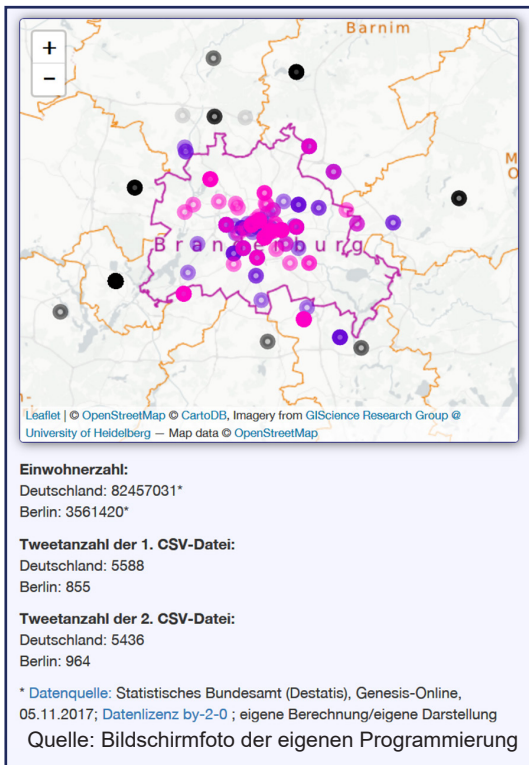


Abbildung 8.11: Tweets in Berlin am 11.12.2017 (lila) und am 12.12.2017 (pink)

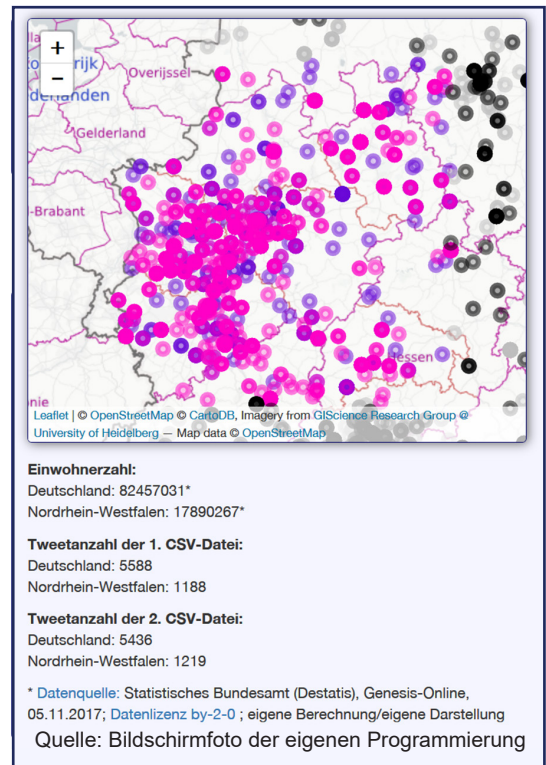


Abbildung 8.12: Tweets in NRW am 11.12.2017 (lila) und am 12.12.2017 (pink)

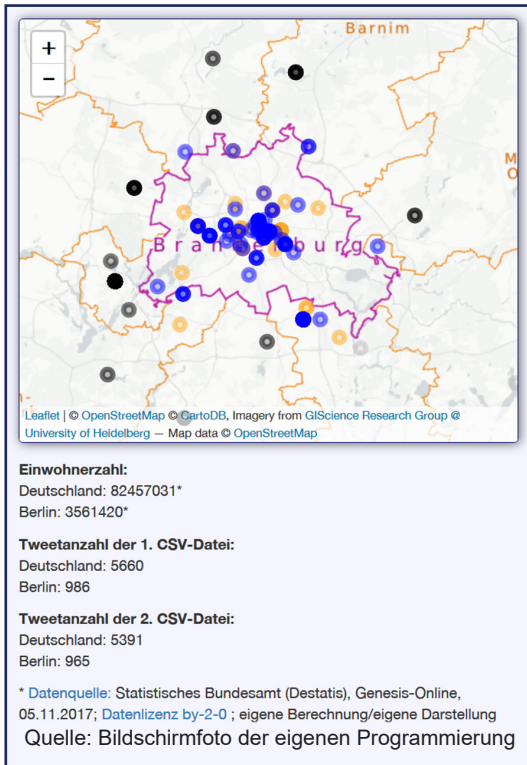


Abbildung 8.13: Tweets in Berlin am 13.12.2017 (orange) und am 07.12.2017 (blau)

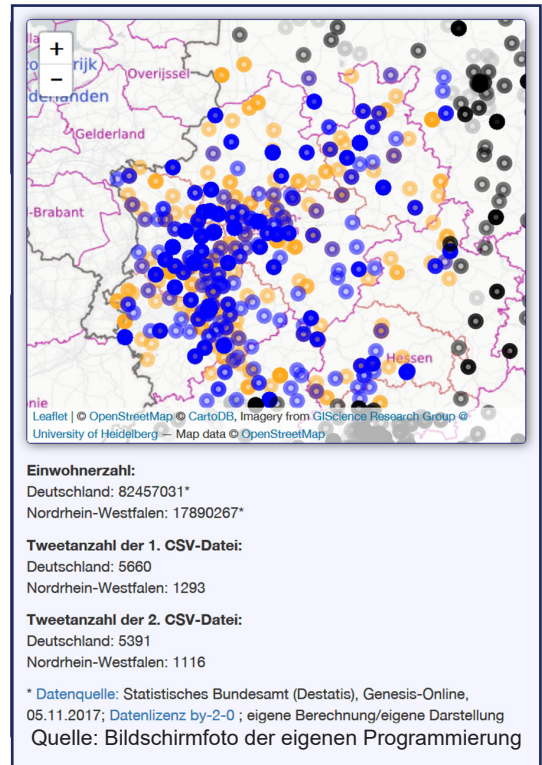


Abbildung 8.14: Tweets in NRW am 13.12.2017 (orange) und am 07.12.2017 (blau)

Im nächsten Schritt werden die Daten mit der jeweils größeren Tweet-Anzahl, wiederum miteinander verglichen. Daher ergibt sich, dass jeweils der 08.12.2017 mit dem 10.12.2017 und der 12.12.2017 mit dem 13.12.2017 verglichen werden. Dabei entstehen jeweils zwei Visualisierungen, wie in den Abbildungen 8.15 bis 8.18 dargestellt. Aus diesen Abbildungen ergibt sich, dass die Digitalisierung in Berlin am 08.12.2017 und am 12.12.2017 am stärksten ist, während es in NRW der 10.12.2017 und der 13.12.2017 sind.

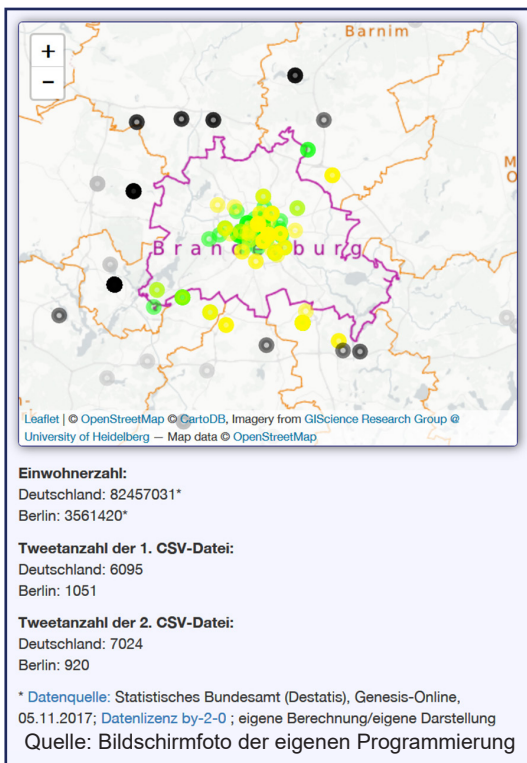


Abbildung 8.15: Tweets in Berlin am 08.12.2017 (grün) und am 10.12.2017 (gelb)

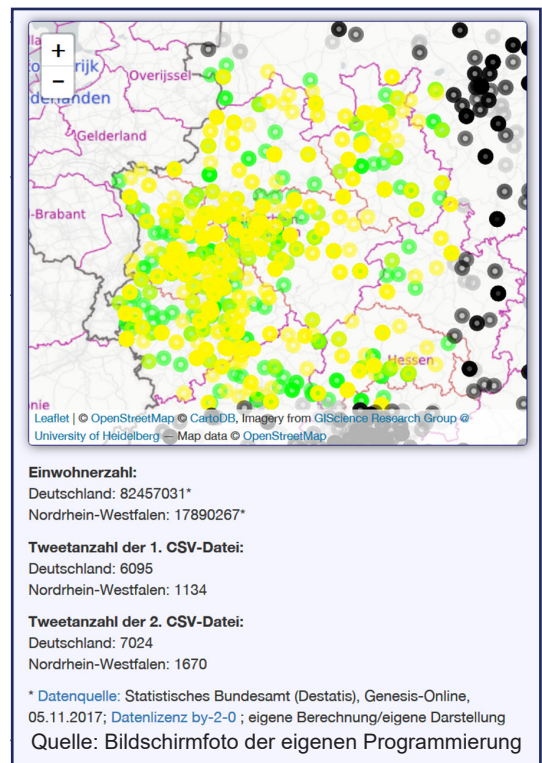


Abbildung 8.16: Tweets in NRW am 08.12.2017 (grün) und am 10.12.2017 (gelb)

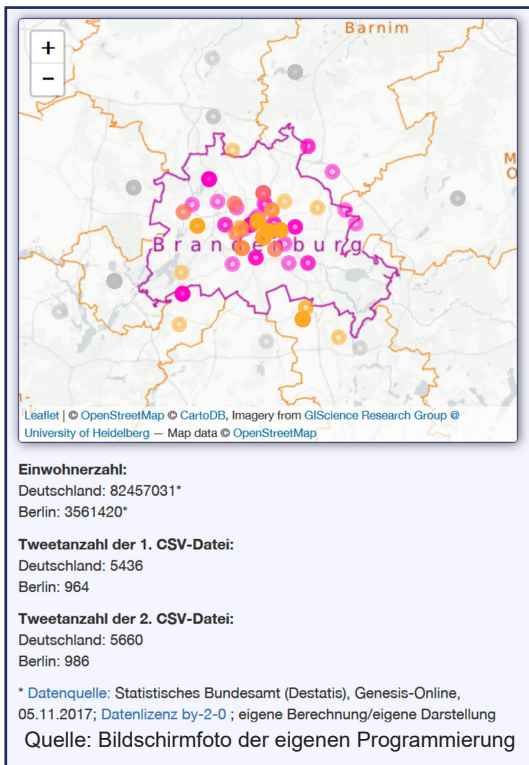


Abbildung 8.17: Tweets in Berlin am 12.12.2017 (pink) und am 13.12.2017 (orange)

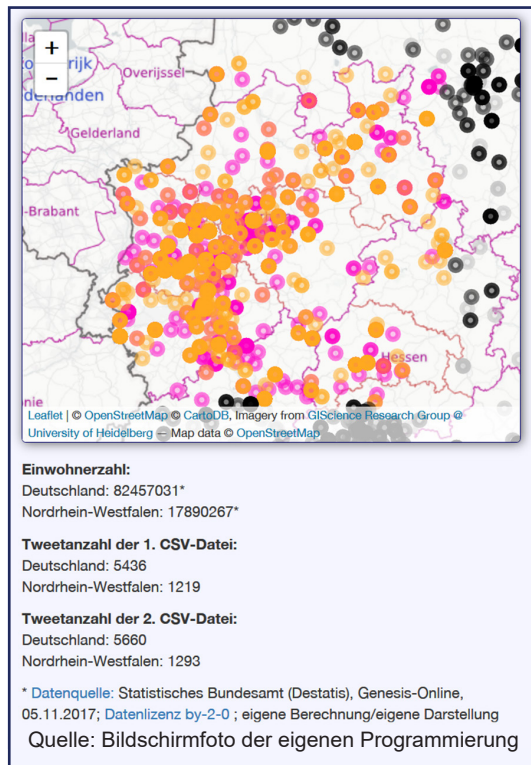


Abbildung 8.18: Tweets in NRW am 12.12.2017 (pink) und am 13.12.2017 (orange)

Anschließend werden wieder die Daten mit der jeweils größeren Tweet-Anzahl dieser Visualisierungen miteinander verglichen. Daraus entsteht dann das Ergebnis, dass in den Abbildungen 8.19 und 8.20 dargestellt ist. In Berlin gibt es, wie Abbildung 8.19 zeigt, am 08.12.2017 die meisten Tweets. Demnach ist an diesem Tag die Digitalisierung im beobachteten Zeitraum am stärksten. In NRW ist dagegen die Digitalisierung am 10.12.2017 am stärksten, da es dann die meisten Tweets in NRW gibt.

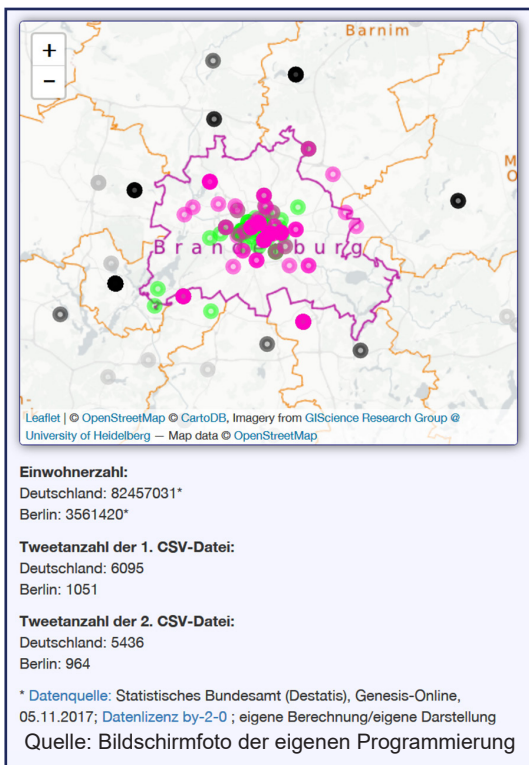


Abbildung 8.19: Tweets in Berlin am 08.12.2017 (grün) und am 12.12.2017 (pink)

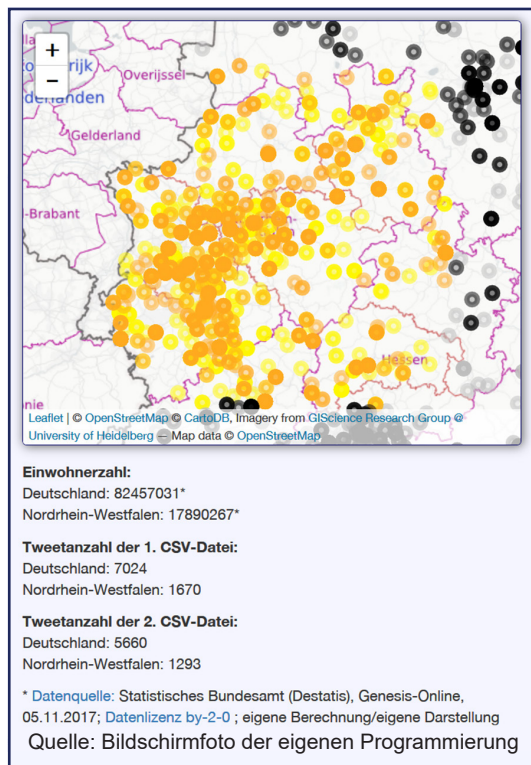


Abbildung 8.20: Tweets in NRW am 10.12.2017 (gelb) und am 13.12.2017 (orange)

9 Abschlussbetrachtung

9.1 Kritische Betrachtung

Im vorherigen Kapitel ist deutlich geworden, dass sich die mit R und Shiny entwickelte Webanwendung zur Visualisierung der Digitalisierung in Deutschland am Beispiel von Twitter-Nachrichten zur Untersuchung verschiedenster Fragestellungen dieses Bereichs nutzen lässt. Dennoch gibt es einige Kritikpunkte, die es zu bedenken gilt.

Ist es für die Untersuchung einer Frage wichtig, dass der Stream täglich auf die Sekunde genau gestartet wird, spielt der Faktor Mensch eine wichtige Rolle. Der Nutzer müsste täglich exakt zur selben Sekunde den Button zum Starten des Streams drücken. Daher wäre hier eine Automatisierung zum Starten des Streams zu empfehlen.

Darüber hinaus wird immer eine durchgängige Internetverbindung beziehungsweise Verbindung zum Twitter-Stream benötigt. Bricht diese Verbindung ab oder ist der Stream der Twitter-API nicht erreichbar, können die Daten nicht mehr nachträglich empfangen werden. Stattdessen werden neue Daten eines nachfolgenden Zeitraums empfangen.

Ein weiterer Kritikpunkt ist die Nutzung einer sogenannten „Bounding-Box“ zur Entscheidung, ob ein Tweet innerhalb oder außerhalb des ausgewählten Bundeslandes liegt. Eine „Bounding-Box“ ist ein Rechteck, das das Bundesland umgibt. Liegt der Tweet innerhalb dieses Rechtecks, wird er auch als innerhalb des Bundeslandes bewertet. Da die Grenzen der Bundesländer aber nicht rechteckig sind, werden manche Tweets auch so farblich visualisiert, als lägen sie innerhalb der Grenzen, obwohl sie außerhalb liegen.

Auch die Auswahl der Farbe, in der die Tweets visualisiert werden, sowie die Reihenfolge in der zwei Dateien mit Tweet-Daten eingeladen werden, sollte bei einem Vergleich von zwei Datensätzen gut durchdacht sein. Die gleichen Datensätze können, je nach Einstellung der App, unterschiedliche Eindrücke beim Betrachter hinterlassen. Lediglich die Auflistung der wichtigsten Zahlen durch den dynamischen Text verschafft Klarheit über die tatsächliche Menge der Tweets. Die Abbildungen 9.1 bis 9.4 zeigen beispielsweise jeweils die abendlichen Daten vom Samstag, den 09.12.2017, und Sonntag, den 10.12.2017 ohne diese Auflistung. Dabei wirkt jede Abbildung anders auf den Betrachter.

Bei den Abbildungen 9.1 und 9.2 sind erst die Daten vom Samstag und dann die Daten vom Sonntag visualisiert worden. In den Abbildungen 9.3 und 9.4

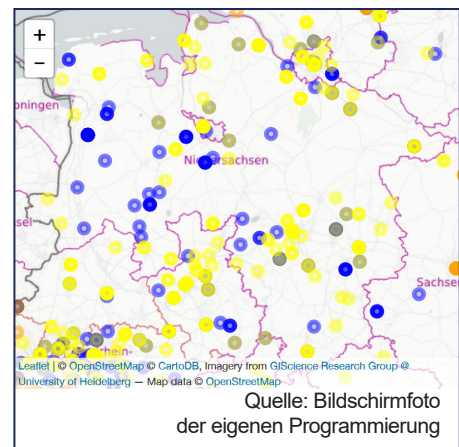


Abbildung 9.1: 1. Vergleich: 09.12.2017 (blau) und 10.12.2017 (gelb)

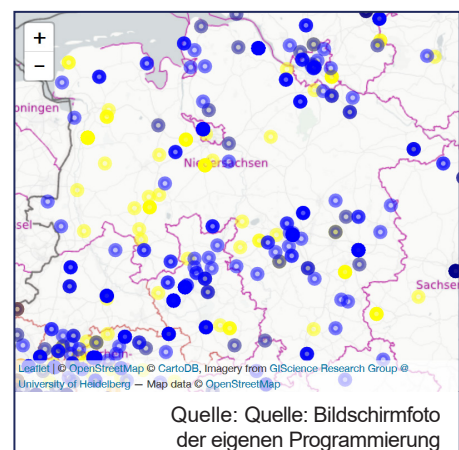


Abbildung 9.2: 2. Vergleich: 09.12.2017 (gelb) und 10.12.2017 (blau)

ist dagegen erst der Sonntag, dann der Samstag eingeladen worden. Dabei ist der Samstag in den Abbildungen 9.1 und 9.4 blau und in den Abbildungen 9.2 und 9.3 gelb dargestellt. In Abbildung 9.1 wirkt es, als habe es Sonntags mehr Tweets gegeben. Dies wird in Abbildung 9.2 noch verstärkt und die Tweets vom Samstag treten noch weiter in den Hintergrund. In Abbildung 9.3 wirkt die Tweet-Anzahl am Samstag und am Sonntag dagegen ziemlich gleich, während in Abbildung 9.4 sogar die Anzahl am Samstag größer zu sein scheint. Das zeigt, dass es viel ausmacht, mit welchen Farben und in welcher Reihenfolge die Tweets visualisiert werden, da dadurch der erste Eindruck über die Ergebnisse manipuliert werden kann. Dennoch ist positiv anzumerken, dass es deutlich sichtbar ist, wenn mehrere Tweets an den selben Koordinaten liegen, da diese durch die Überlagerung der Kreise dunkler wirken als einzelne Tweets.

Ebenfalls zu berücksichtigen ist, dass die Tweets nur dann empfangen und visualisiert werden können, wenn sie auch Standortinformationen enthalten. Dies muss der Ersteller des einzelnen Tweets jedoch selbst freigegeben⁹¹ (vgl. Twitter, Inc., 2017i). Es kann demnach durchaus mehr Tweets im Stream-Zeitraum geben als visualisiert werden, weil nur die Tweets mit Standortdaten von der App erfasst werden können. Darüber hinaus können die Standortdaten als Ortsnamen und nicht als geografische Koordinaten angegeben sein. Dann wird für den entsprechenden Tweet der Mittelpunkt des angegebenen Ortes ermittelt und der Marker an den entsprechenden Koordinaten platziert. Dies müssen jedoch nicht die Koordinaten sein, von wo aus der Tweet tatsächlich veröffentlicht worden ist, da der Ersteller lediglich den Namen, wie beispielsweise „Prenzlauer Berg, Berlin“ (Twitter, Inc., 2017i), angeben muss. Diese Angabe kann theoretisch auch von überall in der Welt aus gemacht werden. Es ist daher nicht nachweisbar, ob der entsprechende Tweet tatsächlich an den visualisierten Koordinaten veröffentlicht worden ist.

Es bleibt daher die Frage offen, wie repräsentativ die Tweets mit Standortinformationen für die Digitalisierung sind, zumal Twitter nur von drei Prozent der Bevölkerung ab 14 Jahren mindestens wöchentlich und ein Prozent täglich genutzt wird (vgl. Frees & Koch, 2017, S. 444). Außerdem ist Twitter nur ein sehr kleiner Teil der sozialen Medien. Die Repräsentativität von Twitter-Daten als Indikator für die Digitalisierung ist daher fragwürdig und ebenfalls ein Kritikpunkt dieser Arbeit, zumal auch darauf vertraut werden muss, dass die Daten, die das Twitter-API liefert, vollständig und korrekt sind, ohne dass dies

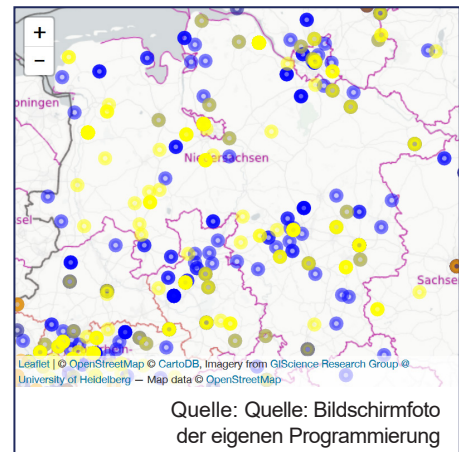


Abbildung 9.3: 3. Vergleich: 10.12.2017 (blau) und 09.12.2017 (gelb)

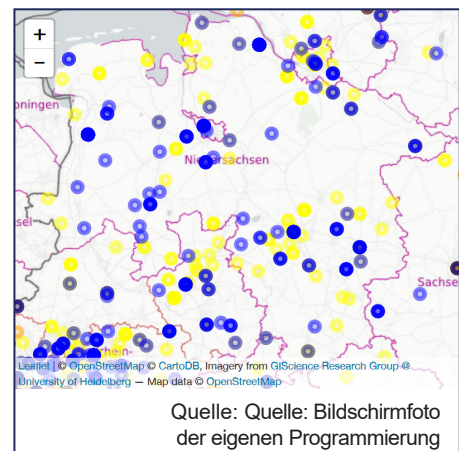


Abbildung 9.4: 4. Vergleich: 10.12.2017 (gelb) und 09.12.2017 (blau)

91 Siehe Kapitel 4.

nachprüfbar ist. Es ist daher repräsentativer, wenn auch Daten aus anderen sozialen Medien mit in die Visualisierung einbezogen werden.

Zusammenfassung:

An dieser Arbeit gibt es einige Kritikpunkte. Dazu gehört, dass der Nutzer den Stream manuell starten muss und dass die Visualisierung durch den Nutzer beeinflussbar ist, da die gewählte Reihenfolge und Farbe der Visualisierung den ersten Eindruck des Betrachters manipulieren kann. Es wird auch kritisiert, dass Rechtecke als Ländergrenzen verwendet werden und dass immer eine durchgängige Verbindung benötigt wird. Die größte Kritik ist jedoch die Frage nach der Repräsentativität der Tweets für die Digitalisierung.

9.2 Ausblick für die Forschung

Die entwickelte Anwendung kann als Werkzeug zur Untersuchung und vor allem zur Visualisierung der Digitalisierung in Deutschland am Beispiel von Tweets verwendet werden. Aber auch darüber hinaus ergeben sich noch Fragen für die weitere Forschung.

Wie bereits mehrfach erwähnt und in Kapitel 8 beispielhaft gezeigt, lassen sich mit der App als Werkzeug viele Fragestellungen in diesem Bereich visualisieren. Dennoch ist es sinnvoll, wenn diese Fragestellungen in Zukunft über einen längeren Zeitraum als sieben Tage erhoben und visualisiert werden. Nur bei einer Untersuchung über einen längeren Zeitraum, wie beispielsweise über ein ganzes Jahr hinweg, können statistische Aussagen getroffen werden. Sieben Tagen sind lediglich eine Stichprobe. Die Repräsentativität dieser Stichprobe für allgemein gültige Aussagen kann nur über einen längeren Zeitraum hinweg bewiesen werden. Aus diesem Grund ist in Zukunft eine längere Untersuchung verschiedener Fragestellungen zur Digitalisierung in Deutschland am Beispiel von Tweets empfehlenswert.

Auch gibt es Potential, die App weiter zu entwickeln. Es können beispielsweise noch weitere soziale Medien mit einbezogen werden, um mehrere soziale Medien miteinander vergleichbar zu machen. Durch einen solchen Vergleich kann auch überprüft werden, ob die Verwendung von Twitter repräsentativ für die Verwendung sozialer Medien in Deutschland und somit für die Digitalisierung ist. Außerdem können so auch mehr soziale Medien für allgemein gültige Aussagen zur Digitalisierung in Deutschland hinzugezogen werden und somit die Repräsentativität erhöht werden.

Außerdem ist es für größere Untersuchungen sinnvoll, mehr als zwei Dateien miteinander vergleichen zu können. Dadurch muss die Visualisierung nicht, wie im Kapitel 8 beschrieben, in mehreren Schritten erfolgen, sondern alle Datensätze könnten in einer Infografik dargestellt werden. Zusätzlich ist auch eine Funktion sinnvoll, um nicht nur ein Bundesland, sondern mehrere gleichzeitig zu betrachten. Das ermöglicht, dass zum Beispiel zwei oder mehr Karten verschiedener Bundesländer nebeneinander visualisiert werden können. Daher gibt es auch noch Möglichkeiten für die Forschung, die Visualisierung in der Anwendung zu erweitern.

Darüber hinaus ist eine Funktion zum Speichern der Visualisierung für den Nutzer ebenfalls sinnvoll. Bisher ist es nötig ein Bildschirmfoto der Visualisierung zu erzeugen, um

diese als Bild speichern und weiter verwenden zu können. Dieser Umweg würde mit einer entsprechenden Funktion zum Herunterladen einer Bilddatei der Visualisierung entfallen. Daher ist auch eine Weiterentwicklung der Anwendung in diesem Bereich hilfreich für die Forschung.

Neben der Arbeit mit der App, sowie der Weiterentwicklung der App ist auch zu untersuchen, ob R in Verbindung mit Shiny die beste Möglichkeit zur Umsetzung der Anwendung als Webseite ist. Da auf das Twitter-API mit vielen Bibliotheken, wie beispielsweise Java, Python, C++ oder vielen mehr, zugegriffen werden (vgl. Twitter, Inc., 2017I), sollte untersucht werden, ob es nicht eine Bibliothek gibt, die für diese Anwendung besser geeignet ist. Obwohl es mit R und Shiny möglich ist, eine Webanwendung aus R zu erzeugen, kann es beispielsweise sein, dass die Performance mit anderen Bibliotheken schneller ist. Außerdem ist es mit R und Shiny nicht möglich, die Marker während des Empfangs der Twitter-Daten einzeln zu visualisieren, sondern sie können erst nach dem Empfang aller Daten visualisiert werden. Hier ist es interessant zu untersuchen, ob eine der anderen Bibliotheken dies ermöglicht.

Unabhängig von der entwickelten Anwendung gibt es ebenfalls noch offene Fragen, die es noch zu erforschen gilt. Einerseits sollte in Zukunft untersucht werden, ob die kostenpflichtige Version des Twitter-API exaktere beziehungsweise mehr oder vollständigere Daten bereitstellt, als die kostenfreie Version und wenn ja, welche. Dies sollte untersucht werden, da es laut Twitter, Inc. (2017c) in der kostenpflichtigen Version mehr als eine Möglichkeit gibt, Standortinformationen abzurufen. Ob die anderen Filteroptionen auch andere Ergebnisse liefern, wird nicht angegeben.

Darüber hinaus sollte ebenfalls erforscht werden, wie viel Prozent der Nutzer von Twitter in Deutschland ihre Standortinformationen frei geben. Hierzu macht Twitter ebenfalls keine Angaben. Daher ist fraglich, wie viel Prozent der Tweets in Deutschland mit der entwickelten Anwendung tatsächlich empfangen werden können, da alle Tweets ohne Standortinformationen nicht berücksichtigt werden. Es gilt daher zu untersuchen, ob die meisten Nutzer Standortinformationen öffentlich zugänglich machen oder nicht.

Zusammenfassung:

Es gibt viele Möglichkeiten für weitere Forschungen. Einerseits bietet die entwickelte Anwendung eine Möglichkeit, Fragestellungen zur Digitalisierung in Deutschland am Beispiel von Tweets zu untersuchen und zu visualisieren. Andererseits gibt es noch einige Möglichkeiten, die Anwendung weiterzuentwickeln und zu untersuchen, ob die Verwendung von R und Shiny die beste Umsetzungsmöglichkeit ist oder eine andere Bibliothek besser geeignet ist. Darüber hinaus sollte untersucht werden, ob die kostenpflichtige Version des Twitter-API exaktere beziehungsweise mehr oder vollständigere Daten bereitstellt, als die kostenfreie Version und wie viel Prozent der Nutzer von Twitter in Deutschland ihre Standortinformationen öffentlich zugänglich machen.

9.3 Fazit

Mit dieser Bachelorarbeit ist eine Webanwendung mittels R und Shiny geschaffen worden, die es Nutzern ermöglicht, den Ort der gesellschaftlichen Digitalisierung zu einem

von ihnen bestimmten Zeitraum anhand von Tweets in Deutschland beziehungsweise in den einzelnen Bundesländern zu visualisieren und mit anderen selbst gewählten Zeiträumen visuell zu vergleichen.

Die Umwandlung von bisher analogen gesellschaftlichen Bereichen ins Digitale ist als gesellschaftliche Digitalisierung definiert. In Deutschland betrifft die Digitalisierung alle Lebensbereiche, wie die sieben Handlungsbereiche in der „Digitale(n) Agenda 2014 - 2017“ (BMW, BMI, & BMVI, 2014, S. Titel) der Bundesregierung, sowie die europäischen Studien, wie der EDPR (vgl. European Union, 2017a, S. 1) und der DESI (vgl. European Union, 2017c, S. 1), und die deutschen Studien, wie der „Deutschland-Index der Digitalisierung“ (Opiela, et al., 2017a, S. Titel) und die „ARD/ZDF-Onlinestudie“ (ARD/ZDF-Medienkommission, 2017, S. 2), belegen. Der Studie „D21-Digital-Index“ (Initiative D21 e. V., 2016, S. Titel) zufolge besitzen die Deutschen 2016 jedoch weniger Offenheit und Kompetenz für die „Anforderungen der Digitalisierung“ (Initiative D21 e. V., 2016, S. 9) im Vergleich zum Jahr 2015.

Ein Indikator der Digitalisierung ist die Nutzung sozialer Medien. Sie dienen im Internet dem Austausch, der Kommunikation und dem Erstellen neuer Inhalte und können in verschiedene Plattform-Arten unterteilt werden. Außerdem wird die Nutzung sozialer Medien auch in den Studien zur Digitalisierung mit zur Bewertung hinzugezogen.

Twitter gehört als Plattform-Art „Microblog“, zu sozialen Medien. Zusätzlich weist Twitter Eigenschaften sozialer Netzwerke auf, die ebenfalls zu den sozialen Medien zählen. Deshalb können Twitter-Nachrichten, genannt Tweets, selbst als Indikator der Digitalisierung betrachtet werden. Tweets sind meistens öffentlich zugänglich und können, neben bis zu 280 Zeichen, auch Fotos, Videos und weitere Informationen, wie Standortinformationen, enthalten. Standortinformationen sind jedoch nur dann öffentlich zugänglich, wenn derjenige, der den Tweet veröffentlicht, diese explizit freigibt.

Die Visualisierung der Orte, an denen Tweets veröffentlicht worden sind, erfolgt als interaktive, kartografische Informationsgrafik. Informationsgrafiken, beziehungsweise Infografiken, stellen komprimierte Informationen durch die Kombination von gestalterischen und textlichen Elementen dar. Bei interaktiven, kartografischen Infografiken liegt eine vom Betrachter beeinflussbare Visualisierung von geografischen Informationen, in Form einer Karte, vor. Diese interaktive, kartografische Infografik wird im Rahmen dieser Arbeit mit einer Auflistung der wichtigsten Zahlen unterstützt.

Für die Implementierung der App wird die mathematische Programmiersprache R in der Entwicklungsumgebung RStudio eingesetzt. Außerdem werden die Daten von Twitter mit der Streaming-API von Twitter abgerufen. Dazu wird in R das Package „rtweets“ benötigt. Damit die App als Webseite aus R heraus implementiert werden kann, wird außerdem das R-Package „shiny“ mit der Variante bestehend aus den beiden Scripten „server.R“ und „ui.R“ genutzt. Darüber hinaus gibt es noch 22 weitere Scripte, um die Funktionalität der App zu implementieren und die Datei „style.css“ wird für die optische Gestaltung der UI-Elemente genutzt. Die Darstellung der Karte und der Marker für die Positionen der Tweets erfolgt dabei mit dem R-Package „leaflet“.

Außerdem ist die Verwendung und Auswertung der entwickelten App an zwei beispielhaften Fragestellungen für Berlin und NRW demonstriert worden. Eine Fragestellung lautet, zu welcher Tageszeit die Digitalisierung am stärksten ist, also wann und wo die meisten Tweets veröffentlicht worden sind. Die andere lautet, an welchem Tag innerhalb eines Zeitraumes die Digitalisierung am stärksten ist und wo sie stattfindet. Dabei ist vom 07.12.2017 bis zum 13.12.2017 der Twitter-Stream viermal täglich für eine Stunde empfangen worden. Im Vergleich hat sich durch die Visualisierungen herausgestellt, dass sowohl in NRW, als auch in Berlin abends die meisten Tweets veröffentlicht worden sind. In NRW hat es außerdem am 10.12.2017 die meisten Tweets und in Berlin am 08.12.2017 gegeben. Zu diesen Zeiten ist demnach die Digitalisierung am stärksten gewesen.

Trotz der Verwendbarkeit der App gibt es einige Kritikpunkte. Die größte Kritik ist dabei die Frage nach der Repräsentativität der Tweets für die Digitalisierung. Einerseits wegen der Eigenschaften der Standortinformationen, andererseits weil Twitter nur von einer geringen Prozentzahl der Bevölkerung verwendet wird und nur ein kleiner Teil der sozialen Medien ist. Darüber hinaus ist die Vollständigkeit und Korrektheit der Daten, die über das Twitter-API erhalten werden, nicht nachweisbar. Eine bessere Repräsentativität der Tweets für die Digitalisierung kann erreicht werden, wenn auch Daten aus anderen sozialen Medien mit in die Visualisierung einbezogen werden.

Für die Forschung bleiben noch viele offene Möglichkeiten. So können mit der App längere Untersuchungen zur Digitalisierung am Beispiel von Tweets durchgeführt werden. Außerdem gibt es noch Potential die App weiter zu entwickeln, wie das Einbinden weiterer sozialer Medien oder Möglichkeit zu schaffen, mehr Dateien gleichzeitig in einer Karte zu visualisieren beziehungsweise mehrere Karten für einen direkten Vergleich zwischen den Bundesländern zu implementieren. Auch eine Funktion zum Speichern der Visualisierung als Bild könnte hilfreich sein. Darüber hinaus sollte untersucht werden, ob die Umsetzung dieser App besser mit einer anderen Programmiersprache als R und Shiny erfolgen sollte und ob die kostenpflichtige Twitter-API mehr und exaktere Daten liefert, als die kostenfreie. Auch sollte erforscht werden, wie viele Twitter-Nutzer ihre Standortinformationen tatsächlich freigegeben.

Zusammenfassung:

Mit der App zur Visualisierung der Digitalisierung in Deutschland am Beispiel von Twitter-Nachrichten können Fragestellungen dieses Themas untersucht und visualisiert werden. Die App besteht aus 24 R-Scripten. Sie nutzt, neben R, das R-Package „rtweet“ für die Verwendung des Streaming-API von Twitter und die R-Packages „shiny“, und „leaflet“. An dieser App gibt es einige Kritikpunkte, jedoch ermöglicht sie auch weitere Forschungen.

Literaturverzeichnis

Adoptitude Analytics, Hrsg. (2017) Tweet Analyzer, Verfügbar unter: <http://socialdash.adoptitude.com/> (abgerufen am: 18.11.2017)

Agafonkin, V. (2017) Leaflet - a JavaScript library for interactive maps, Verfügbar unter: <http://leafletjs.com/index.html> (abgerufen am: 05.12.2017)

ARD/ZDF-Medienkommission, Hrsg. (2017) Kern-Ergebnisse der ARD/ZDF-Onlinestudie 2017, Verfügbar unter: http://www.ard-zdf-onlinestudie.de/files/2017/Artikel/Kern-Ergebnisse_ARDZDF-Onlinestudie_2017.pdf (abgerufen am: 27.10.2017)

Bendel, O., Prof. Dr. (2015) Gabler Wirtschaftslexikon, Stichwort: Digitalisierung, Springer Gabler Verlag, Verfügbar unter: <http://wirtschaftslexikon.gabler.de/Archiv/-2046143105/digitalisierung-v3.html> (abgerufen am 27.10.2017)

Böck, M., et al., Hrsg. (2017) Social-Media-Analyse – mehr als nur eine Wordcloud : HMD Best Paper Award 2016, Wiesbaden: Springer Fachmedien Wiesbaden, Verfügbar unter: <http://www.springerlink.com/content/978-3-658-19802-2> (abgerufen am: 03.11.2016)

Bundesministerium für Wirtschaft und Energie (BMWi), Bundesministerium des Innern (BMI) und Bundesministerium für Verkehr und digitale Infrastruktur (BMVI), Hrsg. (2014) Digitale Agenda 2014 – 2017, Bundesministerium für Wirtschaft und Energie (BMWi), Bundesministerium des Innern (BMI), Bundesministerium für Verkehr und digitale Infrastruktur (BMVI), Verfügbar unter: <http://www.bmwi.de/Redaktion/DE/Publikationen/Digitale-Welt/digitale-agenda.html> (abgerufen am: 27.10.2017)

Burgess, J. und Bruns, A. (2014) Twitter-Archive und die Herausforderungen von »Big Social Data« für die Medien- und Kommunikationswissenschaft, In: Reichert, R., Big Data : Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie, Bielefeld: transcript Verlag, S. 191-202

Chang, W., et al. (2017) shiny: Web Application Framework for R, R Package Version 1.0.5, Verfügbar unter: <https://cran.r-project.org/web/packages/shiny/shiny.pdf> (abgerufen am: 22.11.2017)

Cheng, J., et al. (2017) leaflet: Create Interactive Web Maps with the JavaScript ,Leaflet' Library, R Package Version 1.1.0, Verfügbar unter: <https://cran.r-project.org/web/packages/leaflet/leaflet.pdf> (abgerufen am: 22.11.2017)

Dudenredaktion, Hrsg. (2015a) Duden - Deutsches Universalwörterbuch: Das umfassende Bedeutungswörterbuch der deutschen Gegenwartssprache, Stichwort: digitalisieren, S. 424, 8., überarb. und erw. Aufl. , Berlin: Bibliographisches Institut GmbH

Dudenredaktion, Hrsg. (2015b) Duden - Deutsches Universalwörterbuch: Das umfassende Bedeutungswörterbuch der deutschen Gegenwartssprache, Stichwort: Twitter, S. 1806, 8., überarb. und erw. Aufl. , Berlin: Bibliographisches Institut GmbH

European Union, Hrsg. (2017a) Bericht über den Stand der Digitalisierung in Europa 2017 – Länderprofil Deutschland, European Union, Verfügbar unter: http://ec.europa.eu/newsroom/document.cfm?doc_id=44307 (abgerufen am: 27.10.2017)

European Union, Hrsg. (2017b) DESI 2017 List of indicators, European Union, Verfügbar unter: http://ec.europa.eu/newsroom/document.cfm?doc_id=43049 (abgerufen am: 04.11.2017)

European Union, Hrsg. (2017c) Digital Economy and Society Index 2017 - Germany, European Union, Verfügbar unter: http://ec.europa.eu/newsroom/document.cfm?doc_id=43012 (abgerufen am: 27.10.2017)

Frees, B. und Koch, W. (2017) ARD/ZDF-Onlinestudie 2017: Neun von zehn Deutschen online, Media Perspektiven, MP 9 (9/2017), S. 434-446, Verfügbar unter: http://www.ard-werbung.de/fileadmin/user_upload/media-perspektiven/pdf/2017/917_Koch_Frees.pdf (abgerufen am: 27.10.2017)

Gadatsch, A. und Mangiapane, M. (2017) IT-Sicherheit : Digitalisierung der Geschäftsprozesse und Informationssicherheit, Wiesbaden: Springer Fachmedien Wiesbaden, Verfügbar unter: <http://www.springerlink.com/content/978-3-658-17713-3> (abgerufen am: 23.10.2017)

Gillmann, B. (2017) Deutschland und die Digitalisierung: Der digitale Nachzügler, Online: Handelsblatt GmbH, veröffentlicht am: 24.07.2017, Verfügbar unter: <http://www.handelsblatt.com/politik/deutschland/deutschland-und-die-digitalisierung-der-digitale-nachzuegler-/20098982.html> (abgerufen am: 18.11.2017)

Heidmann, F. (2013) Interaktive Karten und Geovisualisierungen, In: Weber, W., Burmester, M. und Tille, R., Hrsg., Interaktive Infografiken, Berlin, Heidelberg: Springer Berlin Heidelberg, S. 39-69, Verfügbar unter: https://doi.org/10.1007/978-3-642-15453-9_3 (abgerufen am: 03.11.2017)

Himmelberg, C. (2017) Global Digital Report 2017: So digital ist Deutschland, Verfügbar unter: <https://wearesocial.com/de/Special-Reports/global-digital-report-2017-digital-ist-deutschland> (abgerufen am: 10.01.2018)

Initiative D21 e. V., Hrsg. (2016) 2016 D21-DIGITAL-INDEX - Jährliches Lagebild zur Digitalen Gesellschaft, Verfügbar unter: <http://initiated21.de/app/uploads/2017/01/studie-d21-digital-index-2016.pdf> (abgerufen am: 27.10.2017)

Kearney, M. W. (2017a) Live streaming tweets, R Package Version 0.6.0, Verfügbar unter: <https://cran.r-project.org/web/packages/rtweet/vignettes/stream.html> (abgerufen am: 03.12.2017)

Kearney, M. W. (2017b) Obtaining and using access tokens, R Package Version 0.6.0, Verfügbar unter: <https://cran.r-project.org/web/packages/rtweet/vignettes/auth.html> (abgerufen am: 03.12.2017)

Kearney, M. W. (2017c) rtweet: Collecting Twitter Data, R Package Version 0.6.0, Verfügbar unter: <https://cran.r-project.org/web/packages/rtweet/rtweet.pdf> (abgerufen am: 22.11.2017)

Kollmann, T. und Schmidt, H. (2016) Deutschland 4.0 : Wie die Digitale Transformation gelingt, Wiesbaden: Springer Fachmedien Wiesbaden, Verfügbar unter: <http://www.springerlink.com/content/978-3-658-13145-6> (abgerufen am: 23.10.2017)

Kompetenzzentrum Öffentliche IT (ÖFIT), Hrsg. (2017) Der Deutschland Index - Digitales Leben - Soziale Medien, Verfügbar unter: <http://www.oeffentliche-it.de/digitalindex> (abgerufen am: 11.11.2017)

Kumar, S., Morstatter, F. und Liu, H. (2014) Twitter Data Analytics, New York, NY: Springer New York, Verfügbar unter: <http://www.springerlink.com/content/978-1-4614-9372-3> (abgerufen am: 23.10.2017)

Landesregierung Nordrhein-Westfalen, Hrsg. (2015) „NRW 4.0“: Digitaler Wandel in Nordrhein-Westfalen - Fortschrittsbericht der Landesregierung, Landesregierung Nordrhein-Westfalen, Verfügbar unter: https://www.land.nrw/sites/default/files/asset/document/digitaler_wandel_in_nrw_-_fortschrittsbericht_der_landesregierung.pdf (abgerufen am: 10.01.2017)

Lies, J. (2016a) Kompakt-Lexikon PR : 2.000 Begriffe nachschlagen, verstehen, anwenden, Stichwort: Infografik, Wiesbaden: Springer Fachmedien Wiesbaden, Verfügbar unter: <http://www.springerlink.com/content/978-3-658-08742-5> (abgerufen am 03.11.2017)

Lies, J. (2016b) Kompakt-Lexikon PR : 2.000 Begriffe nachschlagen, verstehen, anwenden, Stichwort: Informationsgrafik, Wiesbaden: Springer Fachmedien Wiesbaden, Verfügbar unter: <http://www.springerlink.com/content/978-3-658-08742-5> (abgerufen am 03.11.2017)

Opiela, N., et al. (2017a) Deutschland-Index der Digitalisierung 2017; ÖFIT-Whitepaper, Berlin: Kompetenzzentrum Öffentliche IT (ÖFIT), Verfügbar unter: <https://cdn2.scrvt.com/fokus/fd557e2f812ea594/cdd7fb5c67ea/Deutschland-Index-der-Digitalisierung.pdf> (abgerufen am: 27.10.2017)

Opiela, N., et al. (2017b) Zum Deutschland-Index der Digitalisierung 2017, Berlin: Kompetenzzentrum Öffentliche IT (ÖFIT), Verfügbar unter: <https://www.oeffentliche-it.de/documents/10181/14412/Begleitheft+Deutschland-Index/> (abgerufen am: 07.11.2017)

Presse- und Informationsamt der Bundesregierung (BPA), Hrsg. (2015) Unsere Digitale Agenda für Deutschland, Bundesregierung

Rahlf, T. und Weller, K. (2014) Visualisierung großer Datenmengen aus Social Media Diensten, In: König, C. und Stahl, M., Hrsg., Soziale Medien: Gegenstand und Instrument der Forschung - Schriftenreihe der ASI – Arbeitsgemeinschaft Sozialwissenschaftlicher Institute, Wiesbaden: Springer Fachmedien Wiesbaden, S. 137-159, Verfügbar unter: <https://link.springer.com/book/10.1007%2F978-3-658-05327-7> (abgerufen am: 26.10.2017)

Rendgen, S. (2012) LATCH, In: Rendgen, S., et al., Information Graphics, Italien: Taschen, S. 96

RStudio, Inc., Hrsg. (2017a) Leaflet for R - Introduction, Verfügbar unter: <https://rstudio.github.io/leaflet/> (abgerufen am: 05.12.2017)

RStudio, Inc., Hrsg. (2017b) Leaflet for R - Markers, Verfügbar unter: <https://rstudio.github.io/leaflet/markers.html> (abgerufen am: 05.12.2017)

RStudio, Inc., Hrsg. (2017c) Leaflet for R - Using Basemaps, Verfügbar unter: <https://rstudio.github.io/leaflet/basemaps.html> (abgerufen am: 05.12.2017)

RStudio, Inc., Hrsg. (2017d) Leaflet for R - Using Leaflet with Shiny, Verfügbar unter: <https://rstudio.github.io/leaflet/shiny.html> (abgerufen am: 05.12.2017)

RStudio, Inc., Hrsg. (2017e) Lesson 1 - Welcome to Shiny, Verfügbar unter: <http://shiny.rstudio.com/tutorial/written-tutorial/lesson1/> (abgerufen am: 06.12.2017)

RStudio, Inc., Hrsg. (2017f) Lesson 2 - Build a user interface, Verfügbar unter: <http://shiny.rstudio.com/tutorial/written-tutorial/lesson2/> (abgerufen am: 06.12.2017)

RStudio, Inc., Hrsg. (2017g) Lesson 3 - Add control widgets, Verfügbar unter: <http://shiny.rstudio.com/tutorial/written-tutorial/lesson3/> (abgerufen am: 06.12.2017)

RStudio, Inc., Hrsg. (2017h) Lesson 4 - Display reactive output, Verfügbar unter: <http://shiny.rstudio.com/tutorial/written-tutorial/lesson4/> (abgerufen am: 06.12.2017)

RStudio, Inc., Hrsg. (2017i) Lesson 5 - Use R scripts and data, Verfügbar unter: <http://shiny.rstudio.com/tutorial/written-tutorial/lesson5/> (abgerufen am: 06.12.2017)

RStudio, Inc., Hrsg. (2017j) Lesson 6 - Use reactive expressions, Verfügbar unter: <http://shiny.rstudio.com/tutorial/written-tutorial/lesson6/> (abgerufen am: 06.12.2017)

RStudio, Inc., Hrsg. (2017k) Lesson 7 - Share your apps, Verfügbar unter: <http://shiny.rstudio.com/tutorial/written-tutorial/lesson7/> (abgerufen am: 06.12.2017)

RStudio, Inc., Hrsg. (2017l) RStudio IDE features, Verfügbar unter: <https://www.rstudio.com/products/rstudio/features/> (abgerufen am: 02.12.2017)

RStudio, Inc., Hrsg. (2017m) Shiny from RStudio, Verfügbar unter: <http://shiny.rstudio.com/> (abgerufen am: 18.11.2017)

RStudio, Inc., Hrsg. (2017n) Take control of your R code, Verfügbar unter: <https://www.rstudio.com/products/rstudio/> (abgerufen am: 02.12.2017)

SAS Lightstream, Hrsg. (2017) Tweetping - Realtime Twitter Activity, Verfügbar unter: <https://tweetping.net/#/> (abgerufen am: 18.11.2017)

Scheffler, H. (2014) Soziale Medien: Einführung in das Thema aus Sicht der Marktforschung, In: König, C. und Stahl, M., Hrsg., Soziale Medien: Gegenstand und Instrument der Forschung - Schriftenreihe der ASI – Arbeitsgemeinschaft Sozialwissenschaftlicher Institute, Wiesbaden: Springer Fachmedien Wiesbaden, S. 13-27, Verfügbar unter: <https://link.springer.com/book/10.1007%2F978-3-658-05327-7> (abgerufen am: 26.10.2017)

Schmidt, J.-H. (2017) Social Media, 2. Auflage 2018, aktualisierte und erw., Wiesbaden: Springer Fachmedien Wiesbaden, Verfügbar unter: <https://link.springer.com/book/10.1007%2F978-3-658-19455-0> (abgerufen am: 03.11.2017)

Schwochow, J. (2013) Im Vordergrund steht immer die Information. Erfahrungen eines Infografikers, In: Weber, W., Burmester, M. und Tille, R., Hrsg., Interaktive Infografiken, Berlin, Heidelberg: Springer Berlin Heidelberg, S. 147-160, Verfügbar unter: https://doi.org/10.1007/978-3-642-15453-9_8 (abgerufen am: 03.11.2017)

The R Foundation, Hrsg. (2017) What is R?, Verfügbar unter: <https://www.r-project.org/about.html> (abgerufen am: 18.11.2017)

Twitter, Inc., Hrsg. (2017a) Annual Report 2017, Verfügbar unter: http://files.shareholder.com/downloads/AMDA-2F526X/5553139176x0x935049/05E6E71E-D609-4A17-A8BD-B621324A950D/TWTR_2016_Annual_Report.pdf (abgerufen am: 11.11.2017)

Twitter, Inc., Hrsg. (2017b) Authentication, Verfügbar unter: <https://developer.twitter.com/en/docs/basics/authentication/overview/oauth> (abgerufen am: 15.11.2017)

Twitter, Inc., Hrsg. (2017c) Filter realtime Tweets, Verfügbar unter: <https://developer.twitter.com/en/docs/tweets/filter-realtime> (abgerufen am: 10.12.2017)

Twitter, Inc., Hrsg. (2017d) Getting started, Verfügbar unter: <https://developer.twitter.com/en/docs/basics/getting-started> (abgerufen am: 15.11.2016)

Twitter, Inc., Hrsg. (2017e) Rate Limiting, Verfügbar unter: <https://developer.twitter.com/en/docs/basics/rate-limiting> (abgerufen am: 15.11.2017)

Twitter, Inc., Hrsg. (2017f) Search Tweets, Verfügbar unter: <https://developer.twitter.com/en/docs/tweets/search> (abgerufen am: 15.11.2017)

Twitter, Inc., Hrsg. (2017g) Selected Company Metrics and Financials, Verfügbar unter: http://files.shareholder.com/downloads/AMDA-2F526X/5553139176x0x961126/1C3B5760-08BC-4637-ABA1-A9423C80F1F4/Q317_Selected_Company_Metrics_and_Financials.pdf (abgerufen am: 11.11.2017)

Twitter, Inc., Hrsg. (2017h) Things every developer should know, Verfügbar unter: <https://developer.twitter.com/en/docs/basics/things-every-developer-should-know> (abgerufen am: 15.11.2017)

Twitter, Inc., Hrsg. (2017i) Twitter Hilfe-Center - Hinzufügen deines Standorts zu einem Tweet, Verfügbar unter: <https://support.twitter.com/articles/314849#> (abgerufen am: 12.11.2017)

Twitter, Inc., Hrsg. (2017j) Twitter Hilfe-Center - Über verschiedene Arten von Tweets, Verfügbar unter: <https://help.twitter.com/de/using-twitter/types-of-tweets> (abgerufen am: 13.12.2017)

Twitter, Inc., Hrsg. (2017k) Twitter Hilfe-Center - Über öffentliche und geschützte Tweets, Verfügbar unter: <https://support.twitter.com/articles/334631#> (abgerufen am: 12.11.2017)

Twitter, Inc., Hrsg. (2017l) Twitter libraries, Verfügbar unter: <https://developer.twitter.com/en/docs/developer-utilities/twitter-libraries> (abgerufen am: 26.12.2017)

Venables, W. N., Smith, D. M. und R Core Team (2017) An Introduction to R - Notes on R: A Programming Environment for Data Analysis and Graphics, Version 3.4.2, Verfügbar unter: <https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf> (abgerufen am: 22.11.2017)

Wiegand, F., et al. (2015) ÖFIT-Atlas der Digitalisierung, Berlin: Kompetenzzentrum Öffentliche IT (ÖFIT) und Fraunhofer-Institut für Offene Kommunikationssysteme (FOKUS), Verfügbar unter: <http://www.oeffentliche-it.de/documents/10181/14412/Atlas+der+Digitalisierung> (abgerufen am: 07.10.2017)

Windhagen, E., et al. (2017) Das digitale Wirtschaftswunder - Wunsch oder Wirklichkeit?, Online: McKinsey Global Institute (MGI), Verfügbar unter: https://www.mckinsey.de/files/mgi_das_digitale-wirtschaftswunder.pdf (abgerufen am: 10.01.2018)

Wittpahl, V. (2017) Vorwort, In: Wittpahl, V., iit-Themenband - Digitalisierung : Bildung | Technik | Innovation, Berlin, Heidelberg: Springer Berlin Heidelberg, S. 5-7, Verfügbar unter: <http://www.springerlink.com/content/978-3-662-52854-9> (abgerufen am: 31.10.2017)

Wurman, R.S. (1997) Introduction: Richard Saul Wurman, In: Bradford, P. und Wurman, R.S., Hrsg., Information Architects, New York, NY: Graphis, S. 15-18

Zahn, A.-M. (2013) Status Quo: Wo steht die Social Media Forschung heute? Überblick zu Instrumenten, Anforderungen und Anwendungspotentiale der Social Media Forschung, Berufsverband Deutscher Markt- und Sozialforscher e.V, Frankfurt am Main: 25. NEON-Plenum mit dem Thema „Social Media Guidelines - Richtlinien und rechtliche Rahmenbedingungen der Social Media-Forschung verstehen und anwenden“, Verfügbar unter: https://bvm.org/fileadmin/images/NEON/25._Plenum/2013-02-18_NEON_Vortrag_Zahn.pdf (abgerufen am: 09.11.2017)

Zeit Online, Hrsg. (2015) Digitalisierung: Deutschland im Internet nur Mittelmaß, Online: ZEIT ONLINE GmbH, veröffentlicht am: 24.02.2015, Verfügbar unter: <http://www.zeit.de/digital/internet/2015-02/internet-europaeische-kommission-digitalisierung-studie> (abgerufen am: 18.11.2017)

Anhang

Anlage A: Spezifikation der verwendeten Versionen der Werkzeuge

A.1 Technische Spezifikationen des verwendeten Rechners

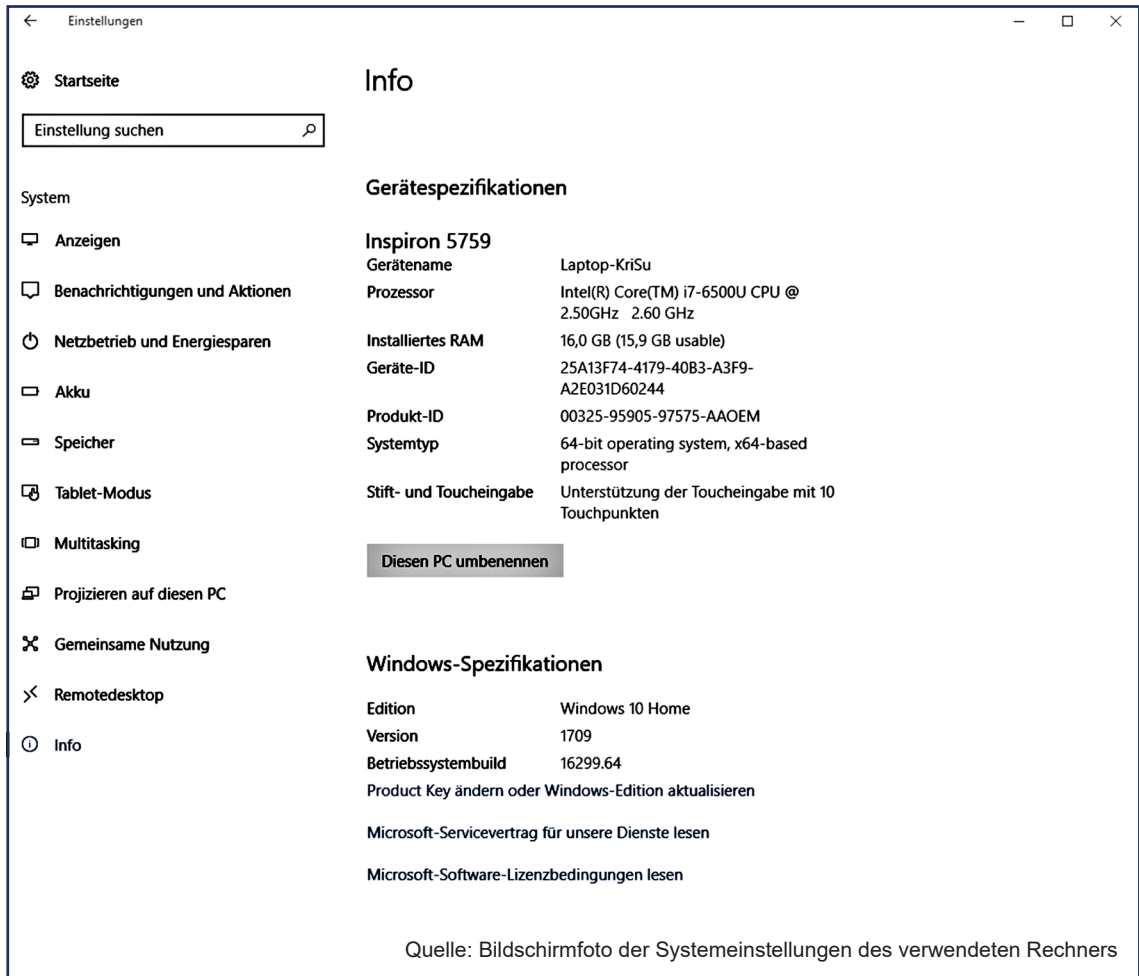


Abbildung A.1: Die technischen Spezifikationen des verwendeten Rechners

A.2 R-Version

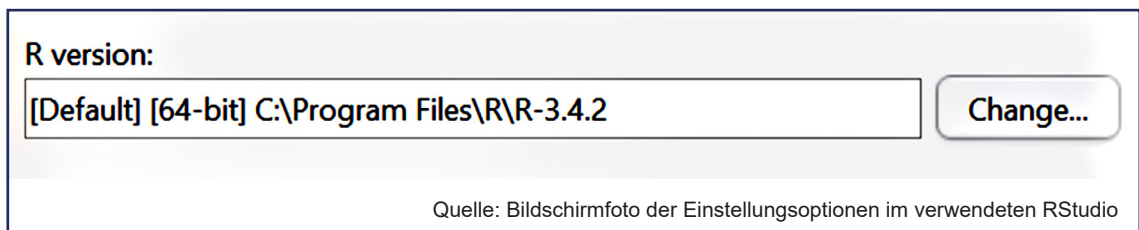


Abbildung A.2: Die verwendete Version von R

A.3 RStudio-Version



Abbildung A.3: Die verwendete Version von RStudio

A.4 Rtweet-Version

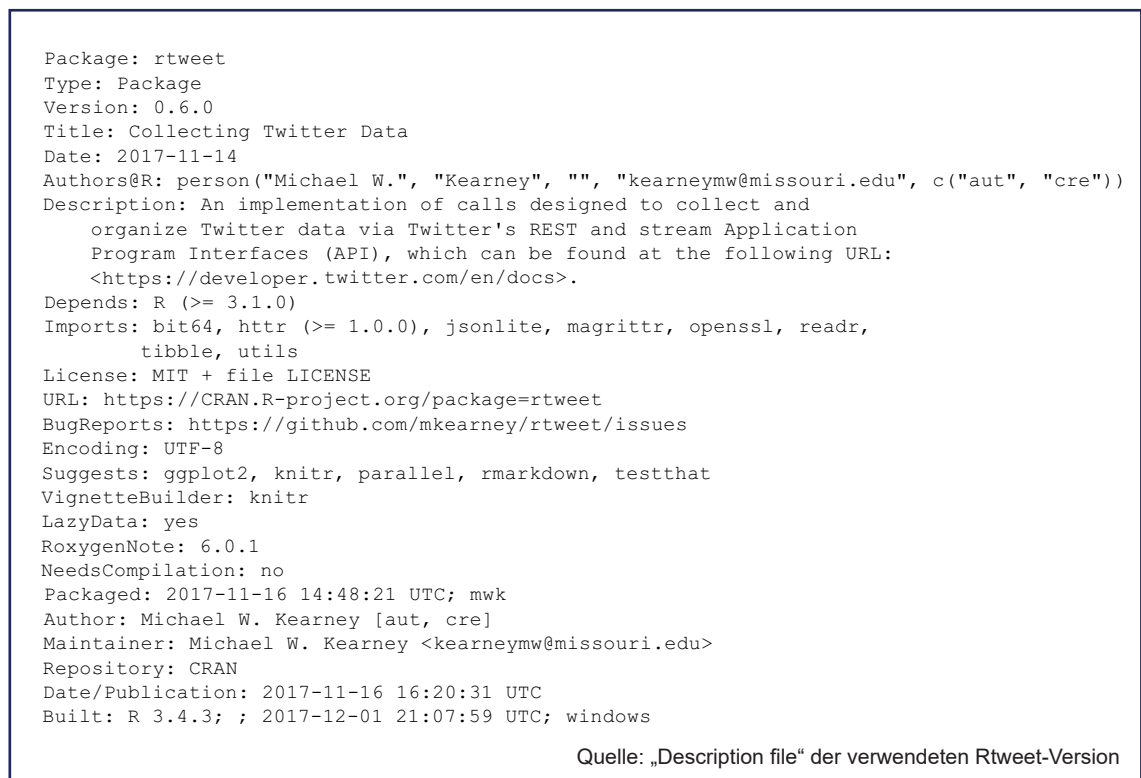


Abbildung A.4: Die verwendete Version von Rtweet

A.5 Leaflet-Version

```

Package: leaflet
Type: Package
Title: Create Interactive Web Maps with the JavaScript 'Leaflet'
      Library
Version: 1.1.0
Date: 2017-02-17
Authors@R: c(
  person("Joe", "Cheng", email = "joe@rstudio.com", role = c("aut", "cre")),
  person("Bhaskar", "Karambelkar", role = c("aut")),
  person("Yihui", "Xie", role = c("aut")),
  person("Hadley", "Wickham", role = c("ctb")),
  person("Kenton", "Russell", role = c("ctb")),
  person("Kent", "Johnson", role=c("ctb")),
  person("jQuery Foundation and contributors", role = c("ctb", "cph"), comment = "jQuery library"),
  person("Vladimir", "Agafonkin", role = c("ctb", "cph"), comment = "Leaflet library"),
  person("CloudMade", role = c("cph"), comment = "Leaflet library"),
  person("Leaflet contributors", role = c("ctb"), comment = "Leaflet library"),
  person("Leaflet Providers contributors", role = c("ctb", "cph"), comment = "Leaflet Providers plugin"),
  person("Brandon Copeland", role = c("ctb", "cph"), comment = "leaflet-measure plugin"),
  person("Jacob Toye", role = c("ctb", "cph"), comment = "Leaflet.label plugin"),
  person("Joerg Dietrich", role = c("ctb", "cph"), comment = "Leaflet.Terminator plugin"),
  person("Benjamin Becquet", role = c("ctb", "cph"), comment = "Leaflet.MagnifyingGlass plugin"),
  person("Norkart AS", role = c("ctb", "cph"), comment = "Leaflet.Minimap plugin"),
  person("L. Voogdt", role = c("ctb", "cph"), comment = "Leaflet.awesome-markers plugin"),
  person("Daniel Montague", role = c("ctb", "cph"), comment = "Leaflet.EasyButton plugin"),
  person("Kartena AB", role = c("ctb", "cph"), comment = "Proj4Leaflet plugin"),
  person("Robert Kajic", role = c("ctb", "cph"), comment = "leaflet-locationfilter plugin"),
  person("Mapbox", role = c("ctb", "cph"), comment = "leaflet-omnivore plugin"),
  person("Michael Bostock", role = c("ctb", "cph"), comment = "topojson"),
  person("RStudio", role = c("cph"))
)
Maintainer: Joe Cheng <joe@rstudio.com>
Description: Create and customize interactive maps using the 'Leaflet'
  JavaScript library and the 'htmlwidgets' package. These maps can be used
  directly from the R console, from 'RStudio', in Shiny apps and R Markdown
  documents.
License: GPL-3
URL: http://rstudio.github.io/leaflet/
BugReports: https://github.com/rstudio/leaflet/issues
Depends: R (>= 3.1.0)
Imports: base64enc, crosstalk, htmlwidgets, htmltools, magrittr,
  markdown, methods, png, RColorBrewer, raster, scales (>=
  0.2.5), sp, stats, viridis
Suggests: knitr, maps, sf, shiny, testit (>= 0.4), rgdal, rgeos, R6,
  RJSONIO, purrr, testthat
RoxygenNote: 5.0.1
LazyData: true
NeedsCompilation: no
Packaged: 2017-02-21 16:56:25 UTC; jcheng
Author: Joe Cheng [aut, cre],
  Bhaskar Karambelkar [aut],
  Yihui Xie [aut],
  Hadley Wickham [ctb],
  Kenton Russell [ctb],
  Kent Johnson [ctb],
  jQuery Foundation and contributors [ctb, cph] (jQuery library),
  Vladimir Agafonkin [ctb, cph] (Leaflet library),
  CloudMade [cph] (Leaflet library),
  Leaflet contributors [ctb] (Leaflet library),
  Leaflet Providers contributors [ctb, cph] (Leaflet Providers plugin),
  Brandon Copeland [ctb, cph] (leaflet-measure plugin),
  Jacob Toye [ctb, cph] (Leaflet.label plugin),
  Joerg Dietrich [ctb, cph] (Leaflet.Terminator plugin),
  Benjamin Becquet [ctb, cph] (Leaflet.MagnifyingGlass plugin),
  Norkart AS [ctb, cph] (Leaflet.Minimap plugin),
  L. Voogdt [ctb, cph] (Leaflet.awesome-markers plugin),
  Daniel Montague [ctb, cph] (Leaflet.EasyButton plugin),
  Kartena AB [ctb, cph] (Proj4Leaflet plugin),
  Robert Kajic [ctb, cph] (leaflet-locationfilter plugin),
  Mapbox [ctb, cph] (leaflet-omnivore plugin),
  Michael Bostock [ctb, cph] (topojson),
  RStudio [cph]
Repository: CRAN
Date/Publication: 2017-02-21 19:08:12
Built: R 3.4.2; ; 2017-10-20 16:25:27 UTC; windows

```

Quelle: „Description file“ der verwendeten Leaflet-Version

Abbildung A.5: Die verwendete Version von Leaflet

A.6 Shiny-Version

```

Package: shiny
Type: Package
Title: Web Application Framework for R
Version: 1.0.5
Authors@R: c(
  person("Winston", "Chang", role = c("aut", "cre"), email = "winston@rstudio.com"),
  person("Joe", "Cheng", role = "aut", email = "joe@rstudio.com"),
  person("JJ", "Allaire", role = "aut", email = "jj@rstudio.com"),
  person("Yihui", "Xie", role = "aut", email = "yihui@rstudio.com"),
  person("Jonathan", "McPherson", role = "aut", email = "jonathan@rstudio.com"),
  person(family = "RStudio", role = "cph"),
  person(family = "jQuery Foundation", role = "cph",
    comment = "jQuery library and jQuery UI library"),
  person(family = "jQuery contributors", role = c("ctb", "cph"),
    comment = "jQuery library; authors listed in inst/www/shared/jquery-AUTHORS.txt"),
  person(family = "jQuery UI contributors", role = c("ctb", "cph"),
    comment = "jQuery UI library; authors listed in inst/www/shared/jqueryui/AUTHORS.txt"),
  person("Mark", "Otto", role = "ctb",
    comment = "Bootstrap library"),
  person("Jacob", "Thornton", role = "ctb",
    comment = "Bootstrap library"),
  person(family = "Bootstrap contributors", role = "ctb",
    comment = "Bootstrap library"),
  person(family = "Twitter, Inc", role = "cph",
    comment = "Bootstrap library"),
  person("Alexander", "Farkas", role = c("ctb", "cph"),
    comment = "html5shiv library"),
  person("Scott", "Jehl", role = c("ctb", "cph"),
    comment = "Respond.js library"),
  person("Stefan", "Petre", role = c("ctb", "cph"),
    comment = "Bootstrap-datepicker library"),
  person("Andrew", "Rowls", role = c("ctb", "cph"),
    comment = "Bootstrap-datepicker library"),
  person("Dave", "Gandy", role = c("ctb", "cph"),
    comment = "Font-Awesome font"),
  person("Brian", "Reavis", role = c("ctb", "cph"),
    comment = "selectize.js library"),
  person("Kristopher Michael", "Kowal", role = c("ctb", "cph"),
    comment = "es5-shim library"),
  person(family = "es5-shim contributors", role = c("ctb", "cph"),
    comment = "es5-shim library"),
  person("Denis", "Ineshin", role = c("ctb", "cph"),
    comment = "ion.rangeSlider library"),
  person("Sami", "Samhuri", role = c("ctb", "cph"),
    comment = "Javascript strftime library"),
  person(family = "SpryMedia Limited", role = c("ctb", "cph"),
    comment = "DataTables library"),
  person("John", "Fraser", role = c("ctb", "cph"),
    comment = "showdown.js library"),
  person("John", "Gruber", role = c("ctb", "cph"),
    comment = "showdown.js library"),
  person("Ivan", "Sagalaev", role = c("ctb", "cph"),
    comment = "highlight.js library"),
  person(family = "R Core Team", role = c("ctb", "cph"),
    comment = "tar implementation from R")
)
Description: Makes it incredibly easy to build interactive web
  applications with R. Automatic "reactive" binding between inputs and
  outputs and extensive prebuilt widgets make it possible to build
  beautiful, responsive, and powerful applications with minimal effort.
License: GPL-3 | file LICENSE
Depends: R (>= 3.0.2), methods
Imports: utils, httpuv (>= 1.3.5), mime (>= 0.3), jsonlite (>= 0.9.16),
  xtable, digest, htmltools (>= 0.3.5), R6 (>= 2.0), sourcetools,
  tools
Suggests: datasets, Cairo (>= 1.5-5), testthat, knitr (>= 1.6),
  markdown, rmarkdown, ggplot2, magrittr
URL: http://shiny.rstudio.com
BugReports: https://github.com/rstudio/shiny/issues
Collate: 'app.R' 'bookmark-state-local.R' 'stack.R' 'bookmark-state.R'
  'bootstrap-layout.R' 'conditions.R' 'map.R' 'globals.R'
  'utils.R' 'bootstrap.R' 'cache.R' 'diagnose.R' 'fileupload.R'
  'graph.R' 'reactives.R' 'reactive-domains.R' 'history.R'
  'hooks.R' 'html-deps.R' 'htmltools.R' 'image-interact-opts.R'
  'image-interact.R' 'imageutils.R' 'input-action.R'
  'input-checkbox.R' 'input-checkboxgroup.R' 'input-date.R'
  'input-daterange.R' 'input-file.R' 'input-numeric.R'
  'input-password.R' 'input-radiobuttons.R' 'input-select.R'
  'input-slider.R' 'input-submit.R' 'input-text.R'
  'input-textarea.R' 'input-utils.R' 'insert-tab.R' 'insert-ui.R'
  'jqueryui.R' 'middleware-shiny.R' 'middleware.R' 'modal.R'
  'modules.R' 'notifications.R' 'priorityqueue.R' 'progress.R'
  'react.R' 'render-plot.R' 'render-table.R' 'run-url.R'
  'serializers.R' 'server-input-handlers.R' 'server.R'
  'shiny-options.R' 'shiny.R' 'shinyui.R' 'shinywrappers.R'

```

Quelle: „Description file“ der verwendeten Shiny-Version

Abbildung A.6: Die verwendete Version von Shiny

Anlage B: R-Scripte

B.1 ui.R

```

#für shiny
library(shiny)
library(colourpicker)

#für die Karte
library(leaflet)

#für Twitter
library(rtweet)

#für csv
library(readr)
library(data.table)

#für json
library(jsonlite)

#Einladen und einmaliges Aufrufen der verwendeten Scripte
source(„weitereScripte/twitterAuthentifizierung.R“)
source(„weitereScripte/streamGeo.R“)
source(„weitereScripte/karteLeaflet.R“)
source(„weitereScripte/zeichnen.R“)
source(„weitereScripte/streamZeichnen.R“)
source(„weitereScripte/vergleichZeichnen.R“)
source(„weitereScripte/geoDaten.R“)
source(„weitereScripte/fehlermeldung.R“)
source(„weitereScripte/leseCSV.R“)
source(„weitereScripte/react.R“)
source(„weitereScripte/schreibeCSV.R“)
source(„weitereScripte/entscheidungKarte.R“)
source(„weitereScripte/startKarte.R“)
source(„weitereScripte/focus.R“)
source(„weitereScripte/sortiereDaten.R“)
source(„weitereScripte/erzeugeDateiname.R“)
source(„weitereScripte/erzeugeDateinameDownload.R“)
source(„weitereScripte/zaehleTweets.R“)
source(„weitereScripte/ermittlungBunNr.R“)
source(„weitereScripte/erzeugeDynText.R“)
source(„weitereScripte/spaltennamenKorrigieren.R“)
source(„weitereScripte/liegtIn.R“)

# Variable für Farbe
farbeStart <- c(Grün=„green“, Blau=„blue“, Schwarz=„black“, Grau =„gray“)

# Define UI for application that draws a histogram
shinyUI(fluidPage(theme = „style.css“,

  # Titel
  titlePanel(„Gesellschaftliche Digitalisierung in Deutschland“),

  sidebarPanel(
    # Panel-Auswahl
    selectInput(„auswahl“, „Was möchten Sie tun?“,
      c(„...“,„Aktuelle Twitter- Daten streamen“,
        „Datensatz anzeigen lassen“,
        „Datensatz mit aktuellen Twitter-Daten vergleichen“,
        „Zwei Datensätze vergleichen“),
      selected = „...“)
  ),

  #----- Twitter-Lizenzen (Anfang) -----
  tags$div(id=„Bedienung“,
    tags$div(id = „Lizenzen“,
      helpText(„Wenn Sie die diesen Dienst nutzen möchten, müssen Sie folgendem zustimmen“),

      checkboxInput(„TwitterLizenzbedingung1“,
        label = htmlOutput(„TwitterTermsOfService“)
      ),
      checkboxInput(„TwitterLizenzbedingung2“,
        label = htmlOutput(„TwitterPrivacyPolicy“)
      ),
      checkboxInput(„TwitterLizenzbedingung3“,
        label = htmlOutput(„TwitterDeveloperAgreement“)
      ),
      checkboxInput(„TwitterLizenzbedingung4“,
        label = htmlOutput(„TwitterDeveloperPolicy“)
      ),
      actionButton(„zustimmung“, „Ich habe die Bedingungen gelesen und stimme ihnen zu.“)
    ),
    #----- Twitter-Lizenzen (Ende) -----

    conditionalPanel(condition =„input.auswahl != ‚...‘ &&
      input.TwitterLizenzbedingung1 &&
      input.TwitterLizenzbedingung2 &&
      input.TwitterLizenzbedingung3 &&
      input.TwitterLizenzbedingung4 &&
      input.zustimmung > 0“,
      tags$div(id=„streaming“,
        helpText(„Für welches Bundesland interessieren Sie sich?“),
        selectInput(„bundesland“, „Bundesland:“,
          c(„...“,„Berlin“, „Baden-Württemberg“, „Bayern“, „Brandenburg“,
            „Bremen“, „Hamburg“, „Hessen“, „Mecklenburg-Vorpommern“,
            „Niedersachsen“, „Nordrhein-Westfalen“, „Rheinland-Pfalz“, „Saarland“,
            „Sachsen“, „Sachsen-Anhalt“, „Schleswig-Holstein“, „Thüringen“),

```

```

        selected = „...“
    )),
#----- Panel 1: Aktuelle Twitter- Daten streamen (Anfang) -----
conditionalPanel(condition = „input.auswahl == ‚Aktuelle Twitter- Daten streamen‘ &&
input.TwitterLizenzbedingung1 &&
input.TwitterLizenzbedingung2 &&
input.TwitterLizenzbedingung3 &&
input.TwitterLizenzbedingung4 &&
input.zustimmung > 0“,
tags$div(id=„streaming“,
tags$div(id=„farbe“,
helpText(„In welcher Farbe sollen die Tweets innerhalb des gewählten Bundeslandes auf der Karte gezeigt werden?“),
colourInput(„farbe1“,
„Farbe der Tweets:“,
farbeStart[2])
),
tags$div(id=„farbeAußen“,
helpText(„In welcher Farbe sollen die Tweets außerhalb des gewählten Bundeslandes auf der Karte gezeigt werden?“),
colourInput(„farbeAußen1“,
„Farbe der Tweets:“,
farbeStart[4])
),
helpText(„Für wie viele Minuten soll der Stream zum Empfang der Twitter-Daten laufen? (max. 1440 Minuten, entspricht 24 Stunden).“),
numericInput(„minuten“, „Minuten“, 1, min = 1, max = 1440, step = 1)
),
actionButton(„aktuell“, „Stream starten“),
helpText(„Bitte erst Streamen. Vorher stehen keine Daten zum Herunterladen zur Verfügung“),
downloadButton(„TwitterStream“, „Stream-Daten herunterladen“)
),
#----- Panel 1: Aktuelle Twitter- Daten streamen (Ende) -----
#----- Panel 2: Daten mit aktuellen Twitter-Daten vergleichen (Anfang) -----
conditionalPanel(condition = „input.auswahl == ‚Datensatz mit aktuellen Twitter-Daten vergleichen‘ &&
input.TwitterLizenzbedingung1 &&
input.TwitterLizenzbedingung2 &&
input.TwitterLizenzbedingung3 &&
input.TwitterLizenzbedingung4 &&
input.zustimmung > 0“,
tags$div(id=„streaming“,
tags$div(id=„farbe“,
helpText(„In welcher Farbe sollen die Tweets des aktuellen Streams innerhalb des gewählten Bundeslandes auf der Karte gezeigt werden?“),
colourInput(„farbe2“,
„Farbe der aktuellen Tweets:“,
farbeStart[2])
),
tags$div(id=„farbeAußen“,
helpText(„In welcher Farbe sollen die Tweets des aktuellen Streams außerhalb des gewählten Bundeslandes auf der Karte gezeigt werden?“),
colourInput(„farbeAußen2“,
„Farbe der aktuellen Tweets:“,
farbeStart[4])
),
helpText(„Für wie viele Minuten soll der Stream zum Empfang der Twitter-Daten laufen? (max. 1440 Minuten, entspricht 24 Stunden).“),
numericInput(„minutenVergleich“, „Minuten“, 1, min = 1, max = 1440, step = 1)
),
tags$div(id=„dateiupload“,
helpText(„Laden Sie eine CSV-Datei hoch“),
tags$div(id=„Erleutern“,
fileInput(„dateiVergleich“, „CSV-Datei“,
accept = c(
„text/csv“,
„text/comma-separated-values“,
„csv“)
),
helpText(„Zum Vergleichen wird eine Datei benötigt, die Sie über einen Twitter-Stream erhalten haben. Diese können Sie über den Download
des jeweils aktuellen Twitter-Streams erhalten.“)
),
tags$div(id=„farbe“,
helpText(„In welcher Farbe sollen die Tweets aus der CSV-Datei innerhalb des gewählten Bundeslandes auf der Karte gezeigt werden?“),
colourInput(„farbe3“,
„Farbe der Tweets aus der CSV-Datei:“,
farbeStart[1])
),
tags$div(id=„farbeAußen“,
helpText(„In welcher Farbe sollen die Tweets aus der CSV-Datei außerhalb des gewählten Bundeslandes auf der Karte gezeigt werden?“),
colourInput(„farbeAußen3“,
„Farbe der Tweets aus der CSV-Datei:“,
farbeStart[3])
),
actionButton(„vergleich“, „Stream starten und mit Daten vergleichen“),
helpText(„Bitte erst Streamen. Vorher stehen keine Daten zum Herunterladen zur Verfügung“),
downloadButton(„aktuellerTwitterStream“, „aktuelle Stream-Daten herunterladen“)
),
#----- Panel 2: Daten mit aktuellen Twitter-Daten vergleichen (Ende) -----
#----- Panel 3: Daten anzeigen lassen (Anfang) -----
conditionalPanel(condition = „input.auswahl == ‚Datensatz anzeigen lassen‘ &&
input.TwitterLizenzbedingung1 &&
input.TwitterLizenzbedingung2 &&
input.TwitterLizenzbedingung3 &&
input.TwitterLizenzbedingung4 &&
input.zustimmung > 0“,
tags$div(id=„dateiupload“,

```

```
        helpText(„Laden Sie eine CSV-Datei hoch“),
        tags$div(id=„Erleutern“,
          fileInput(„dateiAnzeige“, „CSV-Datei“,
            accept = c(
              „text/csv“,
              „text/comma-separated-values“,
              „csv“)
          ),
          helpText(„Es wird eine Datei benötigt, die Sie über einen Twitter-Stream erhalten haben. Diese können Sie über den Download des jeweils aktuellen Twitter-Streams erhalten.“)
        ),
        tags$div(id=„farbe“,
          helpText(„In welcher Farbe sollen die Tweets aus der CSV-Datei innerhalb des gewählten Bundeslandes auf der Karte gezeigt werden?“),
          colourInput(„farbe6“,
            „Farbe der Tweets aus der CSV-Datei:“,
            farbeStart[2]
          ),
        ),
        tags$div(id=„farbeAußen“,
          helpText(„In welcher Farbe sollen die Tweets aus der CSV-Datei außerhalb des gewählten Bundeslandes auf der Karte gezeigt werden?“),
          colourInput(„farbeAußen6“,
            „Farbe der Tweets aus der CSV-Datei:“,
            farbeStart[4]
          ),
        )
      ),
      actionButton(„dateiAnzeigeStart“, „Daten anzeigen“
    ),
    #----- Panel 3: Daten anzeigen lassen (Ende) -----
    #----- Panel 4: Zwei Dateien, die Twitter-Daten enthalten, vergleichen (Anfang) -----
    conditionalPanel(condition = „input.auswahl == ‚Zwei Datensätze vergleichen‘ &&
      input.TwitterLizenzbedingung1 &&
      input.TwitterLizenzbedingung2 &&
      input.TwitterLizenzbedingung3 &&
      input.TwitterLizenzbedingung4 &&
      input.zustimmung > 0“,
      helpText(„Zum Vergleichen werden zwei Datei benötigt, die Sie jeweils über einen Twitter-Stream erhalten haben. Diese können Sie über den Download des jeweils aktuellen Twitter-Streams erhalten.“),
      tags$div(id=„dateiupload“,
        helpText(„Laden Sie die 1. CSV-Datei hoch“),
        fileInput(„datei1Vergleich“, „1. CSV-Datei“,
          accept = c(
            „text/csv“,
            „text/comma-separated-values“,
            „csv“)
        ),
        tags$div(id=„farbe“,
          helpText(„In welcher Farbe sollen die Tweets aus der 1. CSV-Datei innerhalb des gewählten Bundeslandes auf der Karte gezeigt werden?“),
          colourInput(„farbe4“,
            „Farbe der Tweets aus der 1. CSV-Datei:“,
            farbeStart[2]
          ),
        ),
        tags$div(id=„farbeAußen“,
          helpText(„In welcher Farbe sollen die Tweets aus der 1. CSV-Datei außerhalb des gewählten Bundeslandes auf der Karte gezeigt werden?“),
          colourInput(„farbeAußen4“,
            „Farbe der Tweets aus der 1. CSV-Datei:“,
            farbeStart[4]
          ),
        )
      ),
      tags$div(id=„dateiupload“,
        helpText(„Laden Sie die 2. CSV-Datei hoch“),
        fileInput(„datei2Vergleich“, „2. CSV-Datei“,
          accept = c(
            „text/csv“,
            „text/comma-separated-values“,
            „csv“)
        ),
        tags$div(id=„farbe“,
          helpText(„In welcher Farbe sollen die Tweets aus der 2. CSV-Datei auf der Karte gezeigt werden?“),
          colourInput(„farbe5“,
            „Farbe der Tweets aus der 2. CSV-Datei:“,
            farbeStart[1]
          ),
        ),
        tags$div(id=„farbeAußen“,
          helpText(„In welcher Farbe sollen die Tweets aus der 2. CSV-Datei außerhalb des gewählten Bundeslandes auf der Karte gezeigt werden?“),
          colourInput(„farbeAußen5“,
            „Farbe der Tweets aus der 2. CSV-Datei:“,
            farbeStart[3]
          ),
        )
      ),
      actionButton(„dateienVergleich“, „Daten der beiden Dateien vergleichen“)
    )
  ),
  # Ausgabe der erzeugten Visualisierung
  mainPanel(
    #Karte
    leafletOutput(„MapPlot1“),

    #Dynamischer Text
    tags$div(id=„Text“, htmlOutput(„dynText“))
  )
)
```

B.2 server.R

```

#für shiny
library(shiny)
library(colourpicker)

#für die Karte
library(leaflet)

#für Twitter
library(rtweet)

#für csv
library(readr)
library(data.table)

#für json
library(jsonlite)

#Einladen und einmaliges Aufrufen der verwendeten Scripte
source(„weitereScript/twitterAuthentifizierung.R“)
source(„weitereScript/streamGeo.R“)
source(„weitereScript/karteLeaflet.R“)
source(„weitereScript/zeichnen.R“)
source(„weitereScript/streamZeichnen.R“)
source(„weitereScript/vergleichZeichnen.R“)
source(„weitereScript/geoDaten.R“)
source(„weitereScript/fehlermeldung.R“)
source(„weitereScript/leseCSV.R“)
source(„weitereScript/react.R“)
source(„weitereScript/schreibeCSV.R“)
source(„weitereScript/entscheidungKarte.R“)
source(„weitereScript/startKarte.R“)
source(„weitereScript/focus.R“)
source(„weitereScript/sortiereDaten.R“)
source(„weitereScript/erzeugeDateiname.R“)
source(„weitereScript/erzeugeDateinameDownload.R“)
source(„weitereScript/zaehleTweets.R“)
source(„weitereScript/ermittlungBunNr.R“)
source(„weitereScript/erzeugeDynText.R“)
source(„weitereScript/spaltennamenKorrigieren.R“)
source(„weitereScript/liegtIn.R“)

# Ausführung einmalig, wenn die App gestartet wird.
# Authentifizierung bei der Twitter-API
token <- twitterAuthentifizierung()

#Ermitteln der Einwohnerzahl in Deutschland mittels Daten aus CSV
einwohnerzahl <- read.csv2(„Einwohnerzahlen.csv“, check.names=FALSE)
deuEZ <- einwohnerzahl$Deutschland

shinyServer(function(input, output, session) {
  # Ausführung, wenn ein Nutzer die App besucht.
  # Variablen initialisieren, die in allen Funktionen abrufbar sind
  gestreamtEinzel <- reactiveVal(FALSE)
  gestreamtVergleich <- reactiveVal(FALSE)
  zustimmungErteilt <- reactiveVal(FALSE)

  #----- Twitter-Lizenzen (Anfang) -----
  observeEvent(input$zustimmung, {
    #Nach Zustimmung den Block mit Twitter-Lizenzen entfernen
    if(input$TwitterLizenzbedingung1 &&
       input$TwitterLizenzbedingung2 &&
       input$TwitterLizenzbedingung3 &&
       input$TwitterLizenzbedingung4){
      removeUI(selector = „div#Lizenzen“)
      zustimmungErteilt(TRUE)
    }else {
      #Button auf 0 zurücksetzen durch entfernen und neu einfügen
      removeUI(selector = „button#zustimmung“)
      insertUI(
        selector = „div#Lizenzen“,
        where = „beforeEnd“,
        ui = actionButton(„zustimmung“, „Ich habe die Bedingungen gelesen und stimme ihnen zu.“)
      )
      zustimmungErteilt(FALSE)
    }
  })

  # Twitter-Lizenzen verlinken
  output$TwitterTermsOfService <- renderUI({
    tagList(a(„Twitter Terms of Service“, href=„https://twitter.com/tos“))
  })
  output$TwitterPrivacyPolicy <- renderUI({
    tagList(a(„Twitter Privacy Policy“, href=„https://twitter.com/privacy“))
  })

  output$TwitterDeveloperAgreement <- renderUI({
    tagList(a(„Twitter Developer Agreement“, href=„https://developer.twitter.com/en/developer-terms/agreement“))
  })

  output$TwitterDeveloperPolicy <- renderUI({
    tagList(a(„Twitter Developer Policy“, href=„https://developer.twitter.com/en/developer-terms/policy“))
  })
  #----- Twitter-Lizenzen (Ende) -----

  #----- Ausgabe der Karte (Anfang) -----
  output$MapPlot1 <- renderLeaflet({
    # Erzeugen der Karte abhängig von den einzelnen Aktivitäten (siehe „Auswahl“)
    startKarte(input)
  })
  #----- Ausgabe der Karte (Ende) -----

```

```

#----- Fokus-Wechsel (Anfang) -----
observeEvent(input$bundesland, {
  entscheidungKarte(input)
})
observeEvent(input$auswahl, {
  entscheidungKarte(input)
})

#----- Fokus-Wechsel (Ende) -----

#----- Zeichnen in Karte (Anfang) -----

# Stream zeichnen (siehe Auswahl: Aktuelle Twitter- Daten streamen), wenn der Button betätigt wird
observeEvent(input$aktuell, {
  leafletProxy(„MapPlot1“) %>%
    clearMarkers()

  react(input$minuten, input$bundesland, input$farbe1, input$farbeAußen1)

  gestreamEinzel(TRUE)

})

# Stream + Daten aus Datei zeichnen (siehe Auswahl: Daten mit aktuellen Twitter-Daten vergleichen), wenn der Button betätigt wird
observeEvent(input$vergleich, {
  if(!is.null(input$dateiVergleich)){
    leafletProxy(„MapPlot1“) %>%
      clearMarkers()

    react(input$minutenVergleich, input$bundesland, input$farbe2, input$farbeAußen1)

    karteVergleich <- vergleichZeichnen(input$dateiVergleich, input$farbe3, input$farbeAußen3, input$bundesland)
    fehlermeldung(fehler = karteVergleich)

    gestreamVergleich(TRUE)

  } else {
    #Datei fehlt, daher Fehlermeldung (siehe fehlermeldung.R)
    fehler <- list(fehler=2)
    fehlermeldung(fehler = fehler)
  }
})

# Daten aus Datei zeichnen (siehe Auswahl: Daten anzeigen lassen), wenn der Button betätigt wird
observeEvent(input$dateiAnzeigeStart, {
  if(!is.null(input$dateiAnzeige)){
    leafletProxy(„MapPlot1“) %>%
      clearMarkers()
    karteDatei <- vergleichZeichnen(input$dateiAnzeige, input$farbe6, input$farbeAußen6, input$bundesland)
    fehlermeldung(fehler = karteDatei)
  } else {
    #Datei fehlt, daher Fehlermeldung (siehe fehlermeldung.R)
    fehler <- list(fehler=2)
    fehlermeldung(fehler = fehler)
  }
})

# Daten aus zwei Dateien zeichnen (siehe Auswahl: Zwei Dateien, die Twitter-Daten enthalten, vergleichen), wenn der Button betätigt wird
observeEvent(input$dateienVergleich, {
  if(!is.null(input$datei1Vergleich) && !is.null(input$datei2Vergleich)){
    leafletProxy(„MapPlot1“) %>%
      clearMarkers()
    karte1Vergleich <- vergleichZeichnen(input$datei1Vergleich, input$farbe4, input$farbeAußen4, input$bundesland)
    karte2Vergleich <- vergleichZeichnen(input$datei2Vergleich, input$farbe5, input$farbeAußen5, input$bundesland)
    fehlermeldung(fehler = karte1Vergleich)
    fehlermeldung(fehler = karte2Vergleich)
  } else {
    #Datei fehlt, daher Fehlermeldung (siehe fehlermeldung.R)
    fehler <- list(fehler=2)
    fehlermeldung(fehler = fehler)
  }
})

#----- Zeichnen in Karte (Ende) -----

#----- Download (Anfang) -----
# Daten des Streams einlesen
# Wird benötigt um CSV auszugeben
datasetInput1 <- reactive({
  leseCSV(input$minuten)
})
datasetInput2 <- reactive({
  leseCSV(input$minutenVergleich)
})

# CSV erzeugen und Download starten nach Button Betätigung (siehe Auswahl: Aktuelle Twitter- Daten streamen)
output$TwitterStream <- downloadHandler(
  filename = function() {
    dateinameDerCSV <- erzeugeDateinameDownload(input$minuten)
    paste(dateinameDerCSV, „.tweets“, „.csv“, sep=““)
  },
  content = function(file) {
    datei <- datasetInput1()
    schreibeCSV(datei, file)
  }, contentType = „csv“
)

# CSV erzeugen und Download starten nach Button Betätigung (siehe Auswahl: Daten mit aktuellen Twitter-Daten vergleichen)
output$aktuellerTwitterStream <- downloadHandler(
  filename = function() {
    dateinameDerCSV <- erzeugeDateinameDownload(input$minutenVergleich)

```

```

paste(dateinameDerCSV, „tweets“, „.csv“, sep=““)
},
content = function(file) {
  datei <- datasetInput2()
  schreibeCSV(datei, file)
}, contentType = „csv“
)
#----- Download (Ende) -----

#----- dynamischen Text ausgeben (Anfang) -----
# Ermitteln der Einwohnerzahl im Bundesland, wenn ein Bundesland ausgewählt worden ist
bunEZ <- reactive({
  bunNr <- ermittlungBunNr(input$bundesland)
  einwohnerzahl[bunNr]
})

# Dynamischen Text ausgeben
output$dynText <- renderUI({
  #Variablen initialisieren + Einwohnerzahl abrufen
  bunEZ <- bunEZ()
  tweetsStream <-NULL
  tweetsDate1 <-NULL
  tweetsDate2 <-NULL
  tweetsDate3 <-NULL
  tweetsDate4 <-NULL

  #Variablen abhängig vom Input (Stream/Datensätze) und der ausgewählten Aktivität
  if((gestreamtEinzel() == TRUE) && input$auswahl==“Aktuelle Twitter- Daten streamen“){
    gestreamtVergleich(FALSE)
    tweetsDate1 <-NULL
    tweetsDate2 <-NULL
    tweetsDate3 <-NULL
    tweetsDate4 <-NULL
    tweetsStream <-read.csv(paste(erzeugeDateiname(input$minuten),
      „tweets“, „.csv“, sep=““))
  }
  else if((gestreamtVergleich() == TRUE && !is.null(input$dateiVergleich)) && input$auswahl==“Datensatz mit aktuellen Twitter-Daten vergleichen“){
    gestreamtEinzel(FALSE)
    tweetsDate1 <-NULL
    tweetsDate2 <-NULL
    tweetsDate3 <-NULL
    tweetsDate4 <-read.csv(input$dateiVergleich$datapath)
    tweetsStream <-read.csv(paste(erzeugeDateiname(input$minutenVergleich),
      „tweets“, „.csv“, sep=““))
  }
  else if(!is.null(input$dateiAnzeige) && input$auswahl==“Datensatz anzeigen lassen“){
    gestreamtEinzel(FALSE)
    gestreamtVergleich(FALSE)
    tweetsStream <-NULL
    tweetsDate2 <-NULL
    tweetsDate3 <-NULL
    tweetsDate4 <-NULL
    tweetsDate1 <-read.csv(input$dateiAnzeige$datapath)
  }
  else if(!is.null(input$datei1Vergleich) && !is.null(input$datei2Vergleich) && input$auswahl==“Zwei Datensätze vergleichen“){
    gestreamtVergleich(FALSE)
    gestreamtEinzel(FALSE)
    tweetsStream <-NULL
    tweetsDate1 <-NULL
    tweetsDate4 <-NULL
    tweetsDate2 <-read.csv(input$datei1Vergleich$datapath)
    tweetsDate3 <-read.csv(input$datei2Vergleich$datapath)
  }
  }

  # Aufrufen der Funktion zum Erzeugen des dynamischen Textes aus den Variablen
  # Aber erst, wenn die Lizenzbedingungen bestätigt und eine Aktivität ausgewählt worden ist
  if(zustimmungErteilt() == TRUE && input$auswahl != „...“){
    ausgabe <- erzeugeDynText(tweetsStream, tweetsDate1, tweetsDate2, tweetsDate3, tweetsDate4, input$bundesland, deuEZ, bunEZ)
  }

})

#----- dynamischen Text ausgeben (Ende) -----
})

```

B.3 entscheidungKarte.R

```

entscheidungKarte <- function(input){
  fokussieren <- „Deutschland“

  #Prüft, ob ein Bundesland gewählt worden ist
  if(input$bundesland != „...“){
    fokussieren <- input$bundesland
  }
  #Geo-Daten (Bounding-Box) ermitteln
  geoBundesland <- geoDaten(fokussieren)
  #Abruf der Karte anhand der Geo-Daten (Bounding-Box)
  karte <- focus(geoBundesland)
}

```

B.4 ermittlungBunNr.R

```

ermittlungBunNr <- function(bundesland){
  # Abhängig von der Auswahl des Bundeslandes wird eine Zahl zurückgegeben
  # Zahl entspricht der Nummer der Spalte in der CSV der Einwohnerzahlen
  nummer <- switch(bundesland,
    „...“ = 1,
    „Baden-Württemberg“ = 2,
    „Bayern“ = 3,
    „Berlin“ = 4,

```

```

    „Brandenburg“ = 4,
    „Bremen“ = 6,
    „Hamburg“ = 7,
    „Hessen“ = 8,
    „Mecklenburg-Vorpommern“ = 9,
    „Niedersachsen“ = 10,
    „Nordrhein-Westfalen“ = 11,
    „Rheinland-Pfalz“ = 12,
    „Saarland“ = 13,
    „Sachsen“ = 14,
    „Sachsen-Anhalt“ = 15,
    „Schleswig-Holstein“ = 16,
    „Thüringen“ = 17
  )
  return(nummer)
}

```

B.5 erzeugeDateiname.R

```

erzeugeDateiname <- function(minuten){
  #Dateiname und Pfad der Datei mit dem Datensatz des Streams wird als String angelegt und zurückgegeben
  aktuellesDatum <- Sys.Date()
  datum <- as.Date(aktuellesDatum)
  dateinameDerCSV <- paste(„TwitterStreamDaten/rtweet-stream-“, datum, „_Stream-Zeit-“, minuten, „min“, sep=““)
  return(dateinameDerCSV)
}

```

B.6 erzeugeDateinameDownload.R

```

erzeugeDateinameDownload <- function(minuten){
  #Dateiname für den Download der Datei mit dem Datensatz des Streams wird als String angelegt und zurückgegeben
  aktuellesDatum <- Sys.Date()
  datum <- as.Date(aktuellesDatum)
  dateinameDerCSV <- paste(datum, „_TwitterStreamDaten“, „_Stream-Zeit-“, minuten, „min“, sep=““)
  return(dateinameDerCSV)
}

```

B.7 erzeugeDynText.R

```

erzeugeDynText <- function(tweetsStream, tweetsDatei1, tweetsDatei2, tweetsDatei3, tweetsDatei4, bundesland, deuEZ, bunEZ){
  #Initialisierung der Variablen zum Bilden des dynamischen Textes auf der UI
  EZdeutschland <- paste(„Deutschland: „, deuEZ, „“, sep=““)
  EZbundesland <- NULL
  TweetZahlDatei1 <- NULL
  TweetZahlDatei2 <- NULL
  TweetZahlDatei3 <- NULL
  TweetZahlDatei4 <- NULL
  TweetZahlStreamText <- NULL

  #Einwohnerzahl des gewählten Bundeslandes als String-Variablen definieren
  if(bundesland != „“){
    EZbundesland <- paste(bundesland, „: „, bunEZ, „“, sep=““)
  }

  #Dynamischen Text abhängig von der Aktivität in der UI zusammen setzen
  #Der dynamische Text wird als HTML erstellt
  #Dynamischer Text für Stream
  if(!is.null(tweetsStream) && !is.null(tweetsDatei4)){
    TweetZahlStream <- zaehleTweets(tweetsStream, bundesland)
    TweetZahlStreamText1 <- paste(„Deutschland: „, TweetZahlStream$tweetsDeuAnz, sep=““)
    TweetZahlStreamText2 <- paste(bundesland, „: „, TweetZahlStream$tweetsBunAnz, sep=““)

    rueckgabe <- list(tags$span(„Einwohnerzahl:“),
      tags$span(EZdeutschland, tags$br(), EZbundesland),
      tags$span(„Tweetaanzahl des Streams:“),
      tags$span(TweetZahlStreamText1, tags$br(), TweetZahlStreamText2),
      tags$span(„“, tagList(a(„Datenquelle:“, href=“https://www-genesis.destatis.de/genesis/online?sequenz=tabelleErgebnis&selectionname=12411-0021&regionalschlus-
        esel=“),
        „Statistisches Bundesamt (Destatis), Genesis-Online, 05.11.2017;“,
        tagList(a(„Datenlizenz by-2-0“, href=“https://www.govdata.de/dl-de/by-2-0“)),
        „: eigene Berechnung/eigene Darstellung“))
    )
  }
  #Dynamischer Text für Stream und Datensatz im Vergleich
  else if(!is.null(tweetsStream) && !is.null(tweetsDatei4)){
    TweetZahlStream <- zaehleTweets(tweetsStream, bundesland)
    TweetZahlStreamText1 <- paste(„Deutschland: „, TweetZahlStream$tweetsDeuAnz, sep=““)
    TweetZahlStreamText2 <- paste(bundesland, „: „, TweetZahlStream$tweetsBunAnz, sep=““)

    tweetsDatei4 <- spaltennamenKorrigieren(tweetsDatei4)
    TweetZahl4 <- zaehleTweets(tweetsDatei4, bundesland)
    TweetZahl4Text1 <- paste(„Deutschland: „, TweetZahl4$tweetsDeuAnz, sep=““)
    TweetZahl4Text2 <- paste(bundesland, „: „, TweetZahl4$tweetsBunAnz, sep=““)

    rueckgabe <- list(tags$span(„Einwohnerzahl:“),
      tags$span(EZdeutschland, tags$br(), EZbundesland),
      tags$span(„Tweetaanzahl des Streams:“),
      tags$span(TweetZahlStreamText1, tags$br(), TweetZahlStreamText2),
      tags$span(„Tweetaanzahl der CSV-Datei:“),
      tags$span(TweetZahl4Text1, tags$br(), TweetZahl4Text2),
      tags$span(„“, tagList(a(„Datenquelle:“, href=“https://www-genesis.destatis.de/genesis/online?sequenz=tabelleErgebnis&selectionname=12411-0021&regionalschlus-
        esel=“),
        „Statistisches Bundesamt (Destatis), Genesis-Online, 05.11.2017;“,
        tagList(a(„Datenlizenz by-2-0“, href=“https://www.govdata.de/dl-de/by-2-0“)),
        „: eigene Berechnung/eigene Darstellung“))
    )
  }
  #Dynamischer Text für Datensatz
  else if(!is.null(tweetsDatei1)){
    tweetsDatei1 <- spaltennamenKorrigieren(tweetsDatei1)
    TweetZahl1 <- zaehleTweets(tweetsDatei1, bundesland)
    TweetZahl1Text1 <- paste(„Deutschland: „, TweetZahl1$tweetsDeuAnz, sep=““)
    TweetZahl1Text2 <- paste(bundesland, „: „, TweetZahl1$tweetsBunAnz, sep=““)
  }
}

```

```

rueckgabe <- list(tags$span(„Einwohnerzahl:“),
  tags$p(EZdeutschland, tags$br(), EZbundesland),
  tags$span(„Tweetanzahl der CSV-Datei:“),
  tags$p(TweetZahl1Text1, tags$br(), TweetZahl1Text2),
  tags$p(„*“, tagList(a(„Datenquelle:“, href=“https://www-genesis.destatis.de/genesis/online?sequenz=tabelleErgebnis&selectionname=12411-0021&regionalschlu-
essel=“)),
    „Statistisches Bundesamt (Destatis), Genesis-Online, 05.11.2017;“,
    tagList(a(„Datenlizenz by-2-0“, href=“https://www.govdata.de/dl-de/by-2-0“)),
    „; eigene Berechnung/eigene Darstellung“)
  )
}
#Dynamischer Text für 2 Datensätze
else if(!is.null(tweetsDatei2) && !is.null(tweetsDatei3)){
  tweetsDatei2 <- spaltennamenKorrigieren(tweetsDatei2)
  TweetZahl2 <- zaehleTweets(tweetsDatei2, bundesland)
  TweetZahl2Text1 <- paste(„Deutschland: „, TweetZahl2$tweetsDeuAnz, sep=““)
  TweetZahl2Text2 <- paste(bundesland, „: „, TweetZahl2$tweetsBunAnz, sep=““)

  tweetsDatei3 <- spaltennamenKorrigieren(tweetsDatei3)
  TweetZahl3 <- zaehleTweets(tweetsDatei3, bundesland)
  TweetZahl3Text1 <- paste(„Deutschland: „, TweetZahl3$tweetsDeuAnz, sep=““)
  TweetZahl3Text2 <- paste(bundesland, „: „, TweetZahl3$tweetsBunAnz, sep=““)

  rueckgabe <- list(tags$span(„Einwohnerzahl:“),
    tags$p(EZdeutschland, tags$br(), EZbundesland),
    tags$span(„Tweetanzahl der 1. CSV-Datei:“),
    tags$p(TweetZahl2Text1, tags$br(), TweetZahl2Text2),
    tags$span(„Tweetanzahl der 2. CSV-Datei:“),
    tags$p(TweetZahl3Text1, tags$br(), TweetZahl3Text2),
    tags$p(„*“, tagList(a(„Datenquelle:“, href=“https://www-genesis.destatis.de/genesis/online?sequenz=tabelleErgebnis&selectionname=12411-0021&regionalschlu-
essel=“)),
      „Statistisches Bundesamt (Destatis), Genesis-Online, 05.11.2017;“,
      tagList(a(„Datenlizenz by-2-0“, href=“https://www.govdata.de/dl-de/by-2-0“)),
      „; eigene Berechnung/eigene Darstellung“)
    )
  }
#Dynamischer Text ohne Aktion
else {
  if(is.null(EZbundesland)){
    rueckgabe <- list(tags$span(„Einwohnerzahl:“),
      tags$p(EZdeutschland),
      tags$p(„*“, tagList(a(„Datenquelle:“, href=“https://www-genesis.destatis.de/genesis/online?sequenz=tabelleErgebnis&selectionname=12411-0021&regionalschlu-
essel=“)),
        „Statistisches Bundesamt (Destatis), Genesis-Online, 05.11.2017;“,
        tagList(a(„Datenlizenz by-2-0“, href=“https://www.govdata.de/dl-de/by-2-0“)),
        „; eigene Berechnung/eigene Darstellung“)
      )
  } else {
    rueckgabe <- list(tags$span(„Einwohnerzahl:“),
      tags$p(EZdeutschland, tags$br(), EZbundesland),
      tags$p(„*“, tagList(a(„Datenquelle:“, href=“https://www-genesis.destatis.de/genesis/online?sequenz=tabelleErgebnis&selectionname=12411-0021&regionalschlu-
essel=“)),
        „Statistisches Bundesamt (Destatis), Genesis-Online, 05.11.2017;“,
        tagList(a(„Datenlizenz by-2-0“, href=“https://www.govdata.de/dl-de/by-2-0“)),
        „; eigene Berechnung/eigene Darstellung“)
      )
  }
}
return(rueckgabe)
}

```

B.8 fehlermeldung.R

```

fehlermeldung <- function(fehler){
  # Anhand der Fehler-Codes, die zurück geliefert werden, eine Fehlermeldung an den Nutzer geben
  if (fehler$fehler == 1){
    showNotification(„Es wurden keine Daten gefunden. Bitte erneut Streamen.“, duration = 20)
  } else if (fehler == 2){
    showNotification(„Bitte CSV mit Geo-Daten hochladen“, duration = 20)
  } else if (fehler == 3){
    showNotification(„Bitte erst streamen. Die Datei beinhaltet sonst keine Daten.“, duration = 20)
  } else if (fehler == 4){
    showNotification(„Die CSV-Datei ist leer“, duration = 20)
  }
}

```

B.9 focus.R

```

focus <- function(geoBundesland){
  # Geo-Koordinaten entsprechenden Variablen zuweisen
  northlat <- as.double(geoBundesland$box[4])
  eastlon <- as.double(geoBundesland$box[3])
  southlat <- as.double(geoBundesland$box[2])
  westlon <- as.double(geoBundesland$box[1])

  #Karte erzeugen
  karte <- leafletProxy(„MapPlot1“) %>%
    fitBounds(westlon, southlat, eastlon, northlat)

  return(karte)
}

```

B.10 geoDaten.R

```

geoDaten <- function(bundesland){
  #Geo-Koordinaten mittels Twitter-API abrufen
  geoBundesland <- lookup_coords(bundesland, „country:DE“)
  i <- 1

  #Da die Daten oft leer (NULL) zurück geliefert werden, muss die Funktion öfter aufgerufen werden, bis man diese erhält
  #Erhält man nach 200 Versuchen immer noch keine Daten, antwortet die Twitter-API nicht
  while(is.null(geoBundesland$box) && i <= 200){

```

```

geoBundesland <- lookup_coords(bundesland, „country:DE“)
if(i==200){
  showNotification(„Twitter ist zur Zeit nicht erreichbar.“, duration = 20)
}
i <- i+1
}

#Geo-Daten zurückgeben
return(geoBundesland)
}

```

B.11 karteLeaflet.R

```

karteLeaflet <- function(geoBundesland){
  # Geo-Koordinaten entsprechenden Variablen zuweisen
  northlat <- as.double(geoBundesland$box[4])
  eastlon <- as.double(geoBundesland$box[3])
  southlat <- as.double(geoBundesland$box[2])
  westlon <- as.double(geoBundesland$box[1])

  #Karte erzeugen
  karte <- leaflet() %>%
    addProviderTiles(„CartoDB.PositronNoLabels“) %>%
    addProviderTiles(„OpenMapSurfer.AdminBounds“) %>%
    fitBounds(westlon, southlat, eastlon, northlat)
  return(karte)
}

```

B.12 leseCSV.R

```

leseCSV <- function(minuten){

  #Dateiname erstellen
  dateinameDerCSV <- erzeugeDateiname(minuten)
  dateinameDerCSV <- paste(dateinameDerCSV, „,tweets“, „,csv“, sep=““)

  #Prüfen ob die Datei existiert, sonst Fehlermeldung
  if(file.exists(dateinameDerCSV)){
    #Datei einlesen
    dateiCSV <- read.csv(dateinameDerCSV)

    #Alles außer country_code, bbox_coords, lng und lat entfernen
    dateiCSV <- subset(dateiCSV, select= c(country_code, bbox_coords, lng, lat))

    #Rückgabe
    daten <- list(dateiCSV)
    fehler <- 0
    rueckgabe <- list(fehler = fehler, daten = daten)
    return(rueckgabe)
  } else {
    fehler <- 3
    rueckgabe <- list(fehler = fehler)
    return(rueckgabe)
  }
}

```

B.13 liegtIn.R

```

liegtIn <- function(latKoordinate, lngKoordinate, geoBundesland){
  # Prüfen ob die gegebenen Koordinaten innerhalb des gewählten Bundeslandes (Box) liegen
  if((latKoordinate > geoBundesland$box[2] &&
    latKoordinate < geoBundesland$box[4] &&
    (lngKoordinate > geoBundesland$box[1] &&
    lngKoordinate < geoBundesland$box[3]))){
    return(TRUE)
  }
  else{
    return(FALSE)
  }
}

```

B.14 react.R

```

react <- function(minuten, bundesland, farbe, farbeAußen){
  #Zeit berechnen (Shiny benötigt Sekunden)
  streamZeit <- minuten * 60

  #Hinweis für den Nutzer anzeigen
  showNotification(„Bitte warten: Twitter-Daten werden empfangen.“, duration = streamZeit)

  #Geo-Daten abrufen (siehe geoDaten.R)
  geoFuerStream <- geoDaten(bundesland)

  #Erzeugen der Karte (siehe streamZeichnen.R)
  DieseKarte <- streamZeichnen(bundesland, streamZeit, geoFuerStream, farbe, farbeAußen, minuten)

  #Fehlermeldung aufrufen (siehe fehlermeldung.R)
  fehlermeldung(fehler = DieseKarte)
}

```

B.15 schreibeCSV.R

```

schreibeCSV <- function(datei, file){
  if(datei[1] == 0){
    csv <- write.csv(datei[2], file)
  } else {
    csv <- write.csv(NULL, file)
    fehlermeldung(datei[1])
  }
  return(csv)
}

```

B.16 sortiereDaten.R

```

sortiereDaten <- function(deutschlandTweets, geoBundesland, bundesland){
  #Sortieren von Tweets aus dem Bundesland und alle anderen

  # Geo-Koordinaten ermitteln, wenn Sonderfälle (siehe unten) eintreten
  if(bundesland == „Brandenburg“ || bundesland == „Niedersachsen“ || bundesland == „Schleswig-Holstein“){
    berlin <- geoDaten(„Berlin“)
    bremen <- geoDaten(„Bremen“)
    hamburg <- geoDaten(„Hamburg“)
  }

  # Variablen vorbereiten für die Sortierung
  bundeslandTweets <- deutschlandTweets
  loeschen <- c()

  for(i in 1:length(bundeslandTweets$lat)){
    #boolesche Variable, ob die Koordinaten innerhalb des Bundeslandes (Box) liegen
    bool <- liegtIn(bundeslandTweets$lat[i], bundeslandTweets$lng[i], geoBundesland)

    # Sonderfälle: Stadtstaaten liegen innerhalb eines anderen Bundeslandes und müssen daraus entfernt werden
    if(bundesland == „Brandenburg“ && bool){
      bool <- !liegtIn(bundeslandTweets$lat[i], bundeslandTweets$lng[i], berlin)
    }
    if(bundesland == „Niedersachsen“ && bool){
      bool <- (!liegtIn(bundeslandTweets$lat[i], bundeslandTweets$lng[i], bremen) &&
        !liegtIn(bundeslandTweets$lat[i], bundeslandTweets$lng[i], hamburg))
    }
    if(bundesland == „Schleswig-Holstein“ && bool){
      bool <- !liegtIn(bundeslandTweets$lat[i], bundeslandTweets$lng[i], hamburg)
    }
    # Speichern der booleschen Variable in einem Array
    loeschen <- c(loeschen, bool)
  }

  # Entfernen der Tweets die außerhalb des Bundeslandes liegen aus der Variable für die Tweets im Bundesland
  bundeslandTweets <- subset(bundeslandTweets, loeschen)

  # Entfernen der Tweets die innerhalb des Bundeslandes liegen aus der Variable für die Tweets im Rest von Deutschland
  deutschlandTweets <- subset(deutschlandTweets, !loeschen)

  #Rückgabe der sortierten Daten
  rueckgabe <- list(bundeslandTweets = bundeslandTweets, deutschlandTweets = deutschlandTweets)
  return(rueckgabe)
}

```

B.17 spaltennamenKorrigieren.R

```

spaltennamenKorrigieren <- function(daten){
  falscheNamen <- c(„daten.country_code“, „daten.bbox_coords“, „daten.lng“, „daten.lat“)
  richtigeNamen <- c(„country_code“, „bbox_coords“, „lng“, „lat“)

  setnames(daten, old = falscheNamen, new = richtigeNamen)
  return(daten)
}

```

B.18 startKarte.R

```

startKarte <- function(input){
  #Prüft, ob die Lizenzen akzeptiert und welches Panel (Aktivität) ausgewählt worden sind
  Lizenzbedingung <- FALSE
  if(input$TwitterLizenzbedingung1 &&
    input$TwitterLizenzbedingung2 &&
    input$TwitterLizenzbedingung3 &&
    input$TwitterLizenzbedingung4){
    Lizenzbedingung <- TRUE
  }

  #Prüft, welches Panel (Aktivität) ausgewählt worden ist, sofern alle Lizenzen akzeptiert worden sind
  if(Lizenzbedingung && input$auswahl != „...“){
    geoBundesland <- geoDaten(„Deutschland“)
    #Abruf der Karte anhand der Geo-Daten (Bounding-Box)
    karte <- karteLeaflet(geoBundesland)
  }
}

```

B.19 streamGeo.R

```

streamGeo <-function(bundesland, streamZeit, geoBundesland, minuten){

  #Dateiname und -pfad erzeugen
  dateinameFuerStream <- erzeugeDateiname(minuten)

  #Geo-Koordinaten für Deutschland ermitteln
  geoDeutschland <- geoDaten(„Deutschland“)

  #Stream-Daten für ganz Deutschland abrufen
  deutschlandDaten <- stream_tweets(geoDeutschland, timeout = streamZeit, parse = TRUE, file_name = dateinameFuerStream)

  # Falls der Stream leer ist, Fehler zurück melden
  if (is.null(deutschlandDaten)){
    datenRueckgabe <- list(fehler = 1)
    return(datenRueckgabe)
  }else{
    #exakte Koordinaten abrufen und den Daten anhängen
    deutschlandTweets <- lat_lng(deutschlandDaten)

    #Filtern aller Daten, die aus Deutschland stammen
    deutschlandTweets <- subset(deutschlandTweets, country_code == „DE“)

    #Dateiname und -pfad speichern
    dateinameDerJSON <- paste(dateinameFuerStream, „.json“, sep=„“)
    dateinameDerCSV <- paste(dateinameFuerStream, „.csv“, sep=„“)
  }
}

```

```

#Als CSV speichern
save_as_csv(deutschlandTweets, dateinameDerCSV)

#Sortieren von Tweets aus dem Bundesland und alle anderen
tweets <- sortiereDaten(deutschlandTweets, geoBundesland, bundesland)

#Rückgabe der Daten
datenRueckgabe <- list(fehler = 0, geoBundesland = geoBundesland, bundeslandTweets = tweets$bundeslandTweets, dateinameDerJSON = dateinameDerJSON,
deutschlandTweets = tweets$deutschlandTweets)
return(datenRueckgabe)
}

}

```

B.20 streamZeichnen.R

```

streamZeichnen <- function(bundesland, streamZeit, geoBundesland, farbe, farbeAußen, minuten){

#Den Stream aufrufen (siehe streamGeo.R)
wahlBundesland <- streamGeo(bundesland, streamZeit, geoBundesland, minuten)

if (wahlBundesland$fehler == 1){
  rueckgabe <- list(fehler = wahlBundesland$fehler)
  return(rueckgabe)
}else{
  #Geo-Punkte zeichnen (siehe zeichnen.R)
  zeichnen(wahlBundesland$bundeslandTweets$lng,
    wahlBundesland$bundeslandTweets$lat,
    farbe)
  zeichnen(wahlBundesland$deutschlandTweets$lng,
    wahlBundesland$deutschlandTweets$lat,
    farbeAußen)

  #Rückgabe der Geo-Punkte und des Fehler-Codes
  rueckgabe <- list(fehler = wahlBundesland$fehler)
  return(rueckgabe)
}

}

```

B.21 twitterAuthentifizierung.R

```

streamZeichnen <- function(bundesland, streamZeit, geoBundesland, farbe, farbeAußen, minuten){

#Den Stream aufrufen (siehe streamGeo.R)
wahlBundesland <- streamGeo(bundesland, streamZeit, geoBundesland, minuten)

if (wahlBundesland$fehler == 1){
  rueckgabe <- list(fehler = wahlBundesland$fehler)
  return(rueckgabe)
}else{
  #Geo-Punkte zeichnen (siehe zeichnen.R)
  zeichnen(wahlBundesland$bundeslandTweets$lng,
    wahlBundesland$bundeslandTweets$lat,
    farbe)
  zeichnen(wahlBundesland$deutschlandTweets$lng,
    wahlBundesland$deutschlandTweets$lat,
    farbeAußen)

  #Rückgabe der Geo-Punkte und des Fehler-Codes
  rueckgabe <- list(fehler = wahlBundesland$fehler)
  return(rueckgabe)
}

}

```

B.22 vergleichZeichnen.R

```

vergleichZeichnen <- function(datei, farbe, farbeAußen, bundesland){
#mit Try-Catch abfangen ob CSV leer ist
tryCatch({
  #CSV einlesen
  daten <- read.csv(datei$datapath)

  #Prüfen ob Koordinaten in der CSV enthalten sind
  if (is.null(daten$daten.lng) || is.null(daten$daten.lat)){
    #Keine Koordinaten = Fehlermeldung (siehe fehlermeldung.R)
    rueckgabe <- list(fehler = 2)
    return(rueckgabe)
  }else{
    geoBundesland <- geoDaten(bundesland)
    daten <- spaltennamenKorrigieren(daten)

    tweets <- sortiereDaten(daten, geoBundesland, bundesland)

    #Geo-Punkte zeichnen
    zeichnen(tweets$bundeslandTweets$lng,
      tweets$bundeslandTweets$lat,
      farbe)
    zeichnen(tweets$deutschlandTweets$lng,
      tweets$deutschlandTweets$lat,
      farbeAußen)

    #Rückgabe des Fehler-Codes (siehe fehlermeldung.R)
    rueckgabe <- list(fehler= 0)
    return(rueckgabe)
  }
},
error = function(e){
  #Rückgabe des Fehler-Codes (siehe fehlermeldung.R)

```

```

rueckgabe <- list(fehler = 4)
return(rueckgabe)
})
}

```

B.23 zaehleTweets.R

```

zaehleTweets <- function(tweetsDatei, bundesland){
# Zählt die Daten innerhalb eines Datensatzes
if(!is.null(tweetsDatei)){
#Geo-Daten des gewählten Bundeslandes ermitteln
geoBundesland <- geoDaten(bundesland)

# Tweets sortieren nach innerhalb und außerhalb des Bundeslandes
tweets <- sortiereDaten(tweetsDatei, geoBundesland, bundesland)
tweetsBun <- subset(tweets$bundeslandTweets, country_code == „DE“)
tweetsDeu <- subset(tweets$deutschlandTweets, country_code == „DE“)

# Anzahl der jeweiligen Tweets ermitteln
tweetsBunAnz <- length(tweetsBun$lng)
tweetsDeuAnz <- length(tweetsDeu$lng) + tweetsBunAnz

#Jeweilige Anzahl zurück geben
rueckgabe <- list(tweetsBunAnz = tweetsBunAnz, tweetsDeuAnz = tweetsDeuAnz)
return(rueckgabe)
}
}

```

B.24 zeichnen.R

```

zeichnen <- function(lng, lat, farbe){
#Der Karte Marker in Punktform anhand der Geo-Koordinaten hinzufügen
leafletProxy(„MapPlot1“) %>%
addCircleMarkers(lng = lng,
lat = lat,
radius = 5, color = farbe)
}

```

B.25 style.css

```

body {
background-color: #f4f5ff;
}

/* Auswahl: Was möchten Sie tun? */
.col-sm-4 {
float: left;
width: 100%;
min-width: 33%;
max-width: 400px;
padding-left: 0px;
padding-right: 0px;
}

.well {
background-color: #000169;
color: #fff;
border: none;
border-radius: none;
-webkit-box-shadow: none;
box-shadow: none;
}

.well .control-label {
color: #fff;
font-weight: 700;
}

/* Bedienelemente */
#Bedienung {
min-width: 33%;
max-width: 400px;
float: left;
clear: left;
}

#streaming, #dateiupload {
border: 1px solid #000169;
padding: 19px;
margin-bottom: 19px;
border-radius: 4px;
}

.help-block {
color: #000;
font-weight: 700;
}

label {
color: #737373;
font-weight: 400;
}

#Erleutern .help-block {
font-weight: 400;
color: #737373;
font-size: 10px;
clear: both;
}

.progress {
margin-bottom: 0px;
}

```

```

}

.form-group {
  margin-bottom: 0px;
}

.form-control {
  border: 1px solid #000169;
}
.selectize-input {
  border: 1px solid #000169;
}

.btn-default {
  color: #fff;
  background-color: #000169;
  border-color: #000169;
}

.btn-default:hover {
  background-color: #fff;
  color: #000169;
  border-color: #000169;
}

.selectize-dropdown-content {
  max-height: 8em;
}

/* Karte */
.col-sm-8 {
  max-width: 100%;
  width: 500px;
}

#MapPlot1 {
  border: 1px solid #000169;
  box-shadow: 1px 1px 10px #5f5e5e;
  border-radius: 4px;
}

/* Meldung */
#shiny-notification-panel {
  position: fixed;
  bottom: 30%;
  left: 10px;
  min-width: 33%;
  max-width: 500px;
  width: 100%;
}

.shiny-notification {
  background-color: #960000;
  color: #fff;
  border: 1px solid #ccc;
  border-radius: 5px;
  opacity: 1;
  height: 100px;
  padding-top: 40px;
  font-weight: 700;
  text-align: center;
  box-shadow: 1px 1px 10px #5f5e5e;
  border: none;
}

.shiny-notification-close {
  color: #fff;
}

/*dynamischer Text*/
#dynText span{
  font-weight: 700;
}
#Text {
  margin-top: 15px;
}

```

B.26 Einwohnerzahlen.csv

Deutschland;Baden-Württemberg;Bayern;Berlin;Brandenburg;Bremen;Hamburg;Hessen;Mecklenburg-Vorpommern;Niedersachsen;Nordrhein-Westfalen;Rheinland-Pfalz;Saarland;Sachsen;Sachsen-Anhalt;Schleswig-Holstein;Thüringen
 82457031;10943532;12916927;3561420;2492103;677572;1805316;6190664;1608241;7949413;17890267;4066712;997236;4080771;2237425;2880898;2158534

Anlage C: Visualisierungen der Streams vom 07. bis 13.12.2017 in Berlin

C.1 Daten vom 07.12.2017

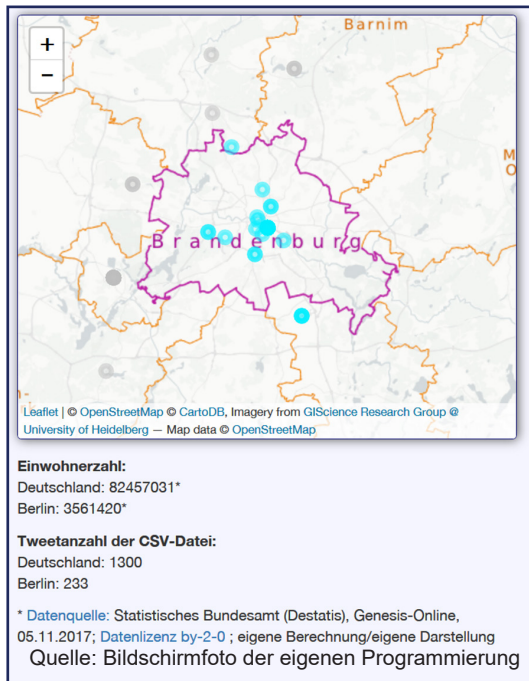


Abbildung C.1: Tweets in Berlin
morgens am 07.12.2017

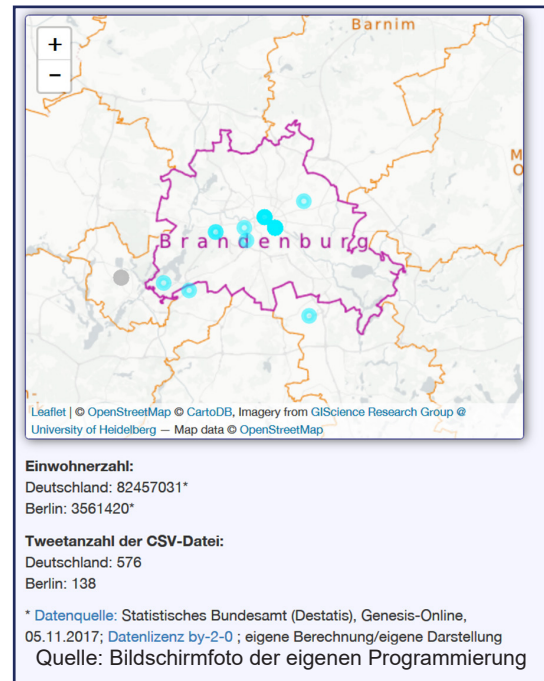


Abbildung C.2: Tweets in Berlin
mittags am 07.12.2017

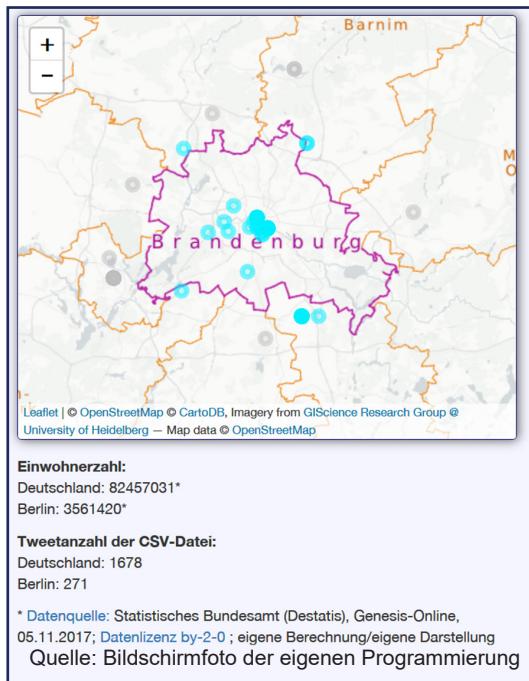


Abbildung C.3: Tweets in Berlin
nachmittags am 07.12.2017

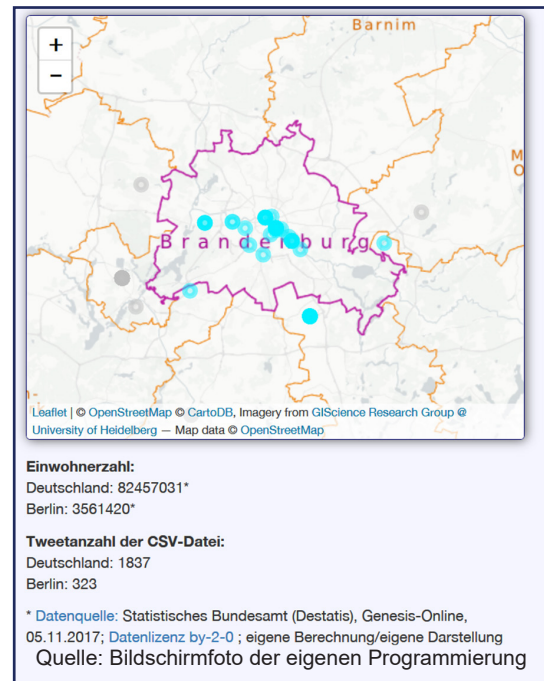


Abbildung C.4: Tweets in Berlin
abends am 07.12.2017

C.2 Daten vom 08.12.2017

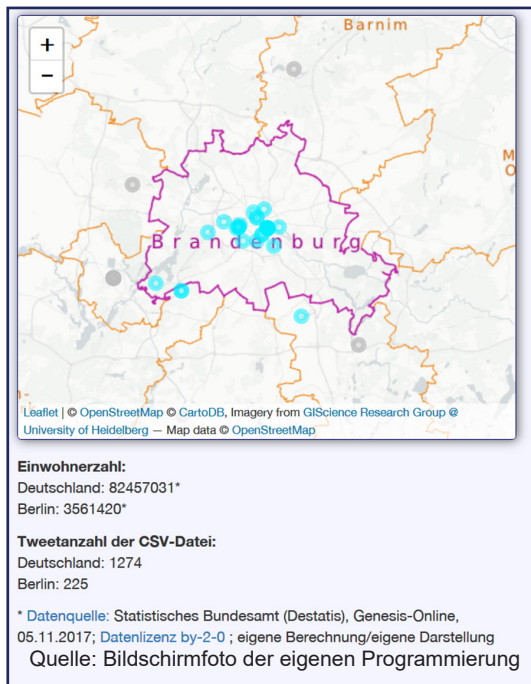


Abbildung C.5: Tweets in Berlin
morgens am 08.12.2017

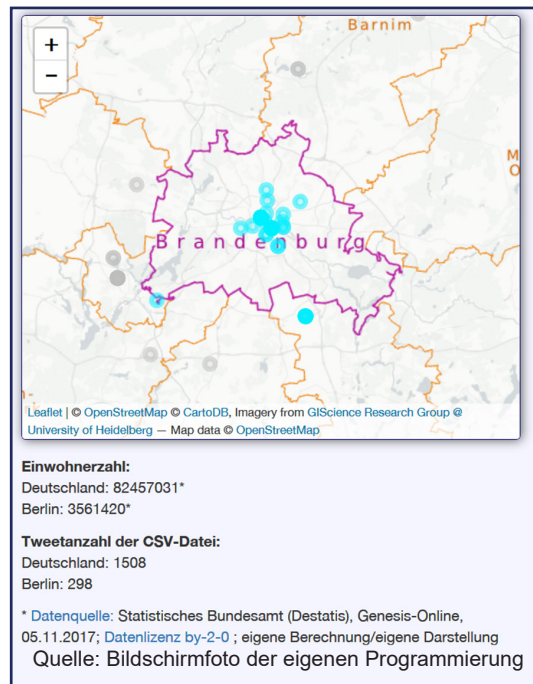


Abbildung C.6: Tweets in Berlin
mittags am 08.12.2017

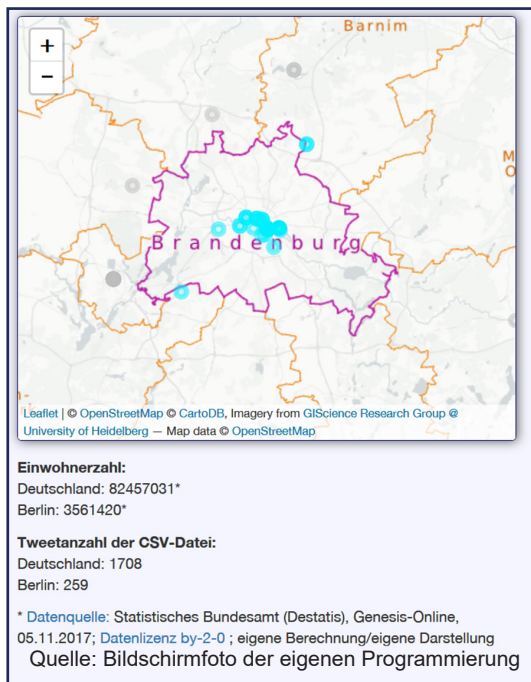


Abbildung C.7: Tweets in Berlin
nachmittags am 08.12.2017

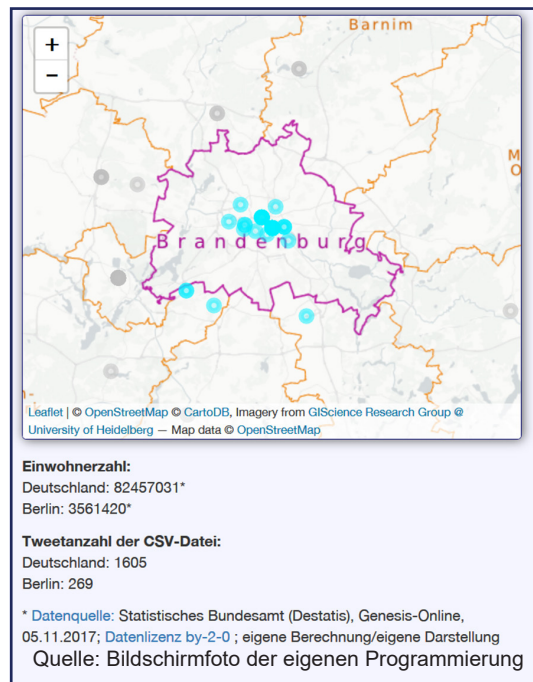


Abbildung C.8: Tweets in Berlin
abends am 08.12.2017

C.3 Daten vom 09.12.2017

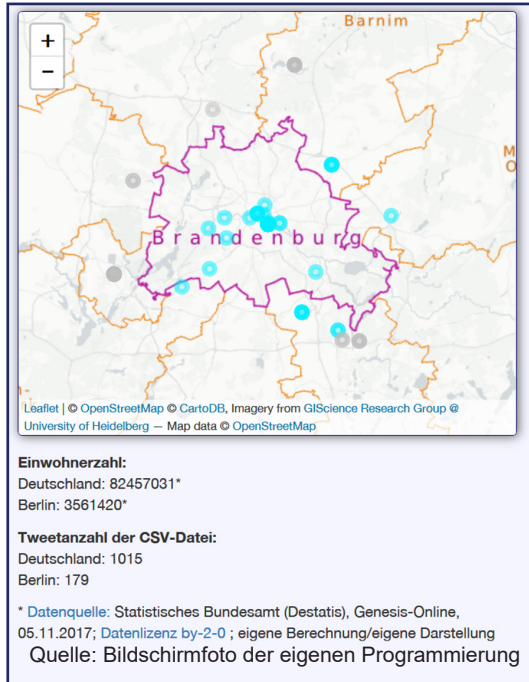


Abbildung C.9: Tweets in Berlin
morgens am 09.12.2017

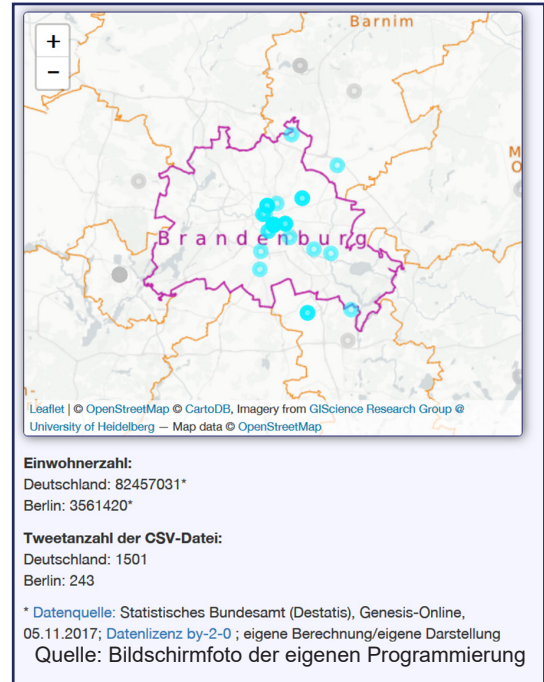


Abbildung C.10: Tweets in Berlin
mittags am 09.12.2017

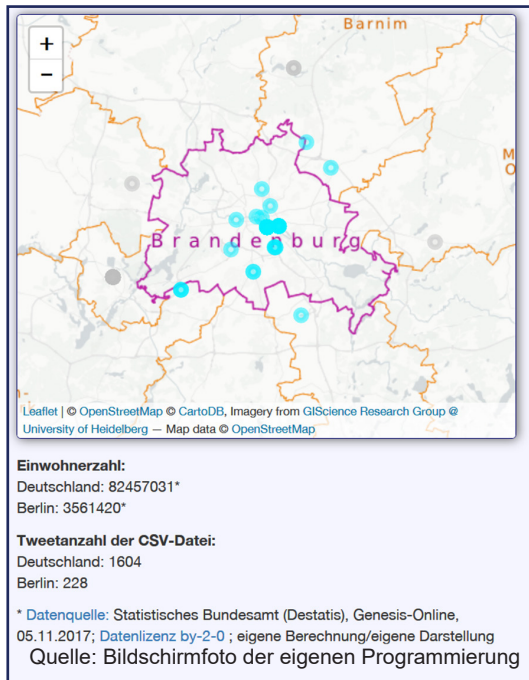


Abbildung C.11: Tweets in Berlin
nachmittags am 09.12.2017

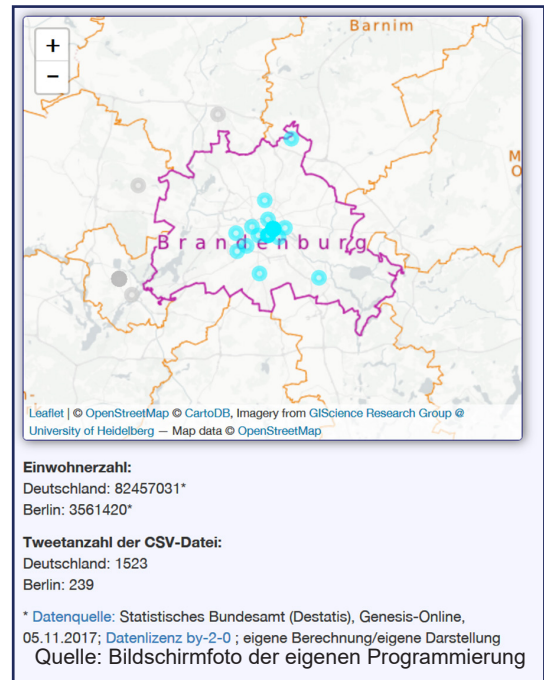


Abbildung C.12: Tweets in Berlin
abends am 09.12.2017

C.4 Daten vom 10.12.2017

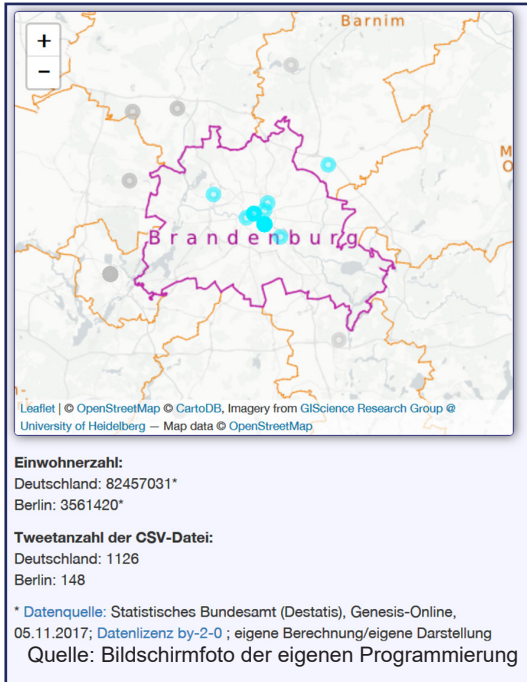


Abbildung C.13: Tweets in Berlin
morgens am 10.12.2017

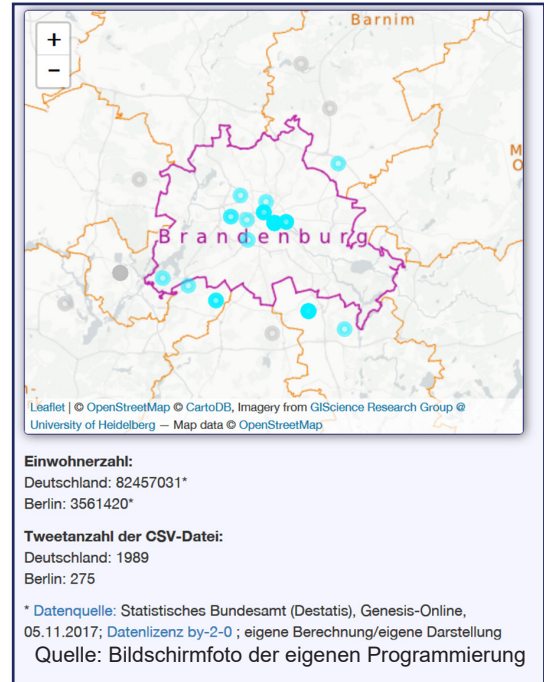


Abbildung C.14: Tweets in Berlin
mittags am 10.12.2017

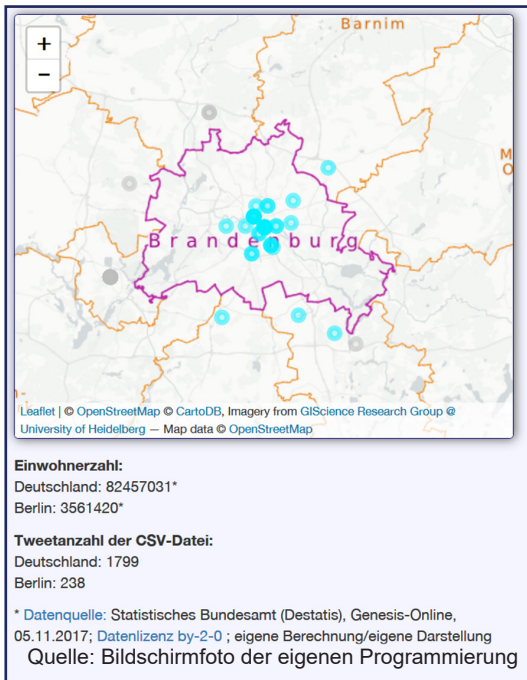


Abbildung C.15: Tweets in Berlin
nachmittags am 10.12.2017

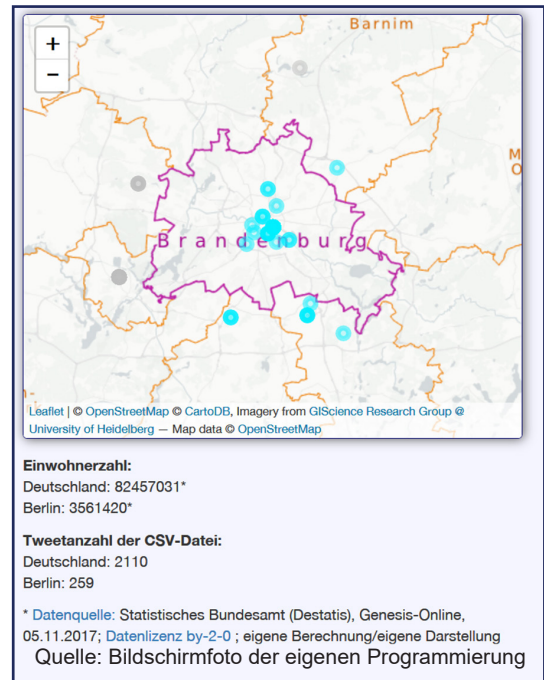
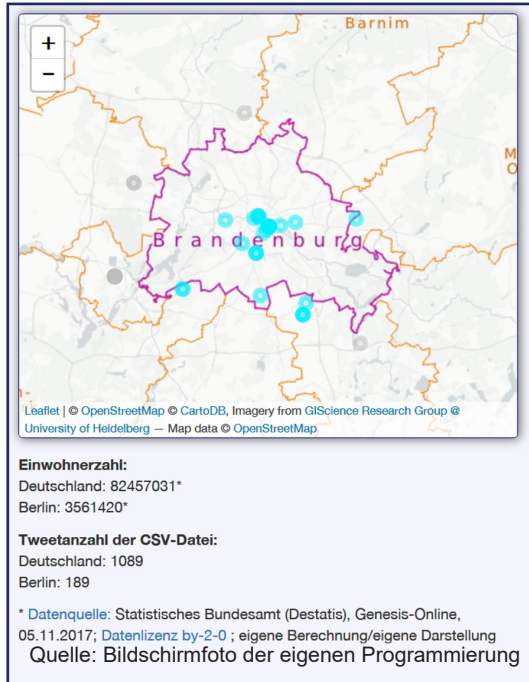
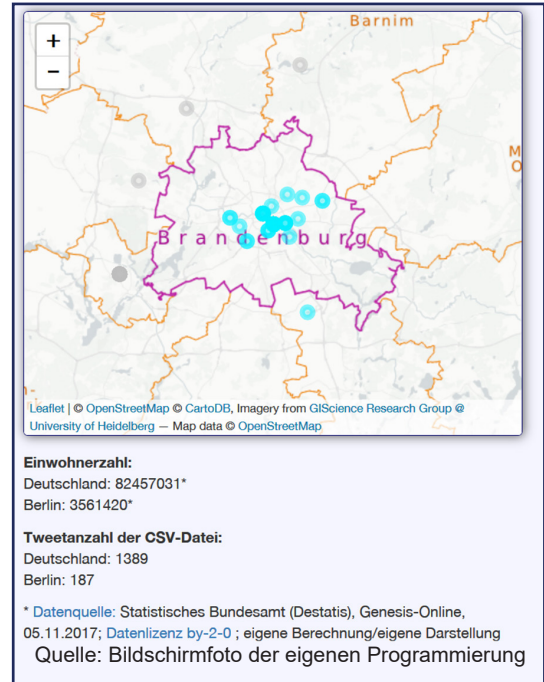
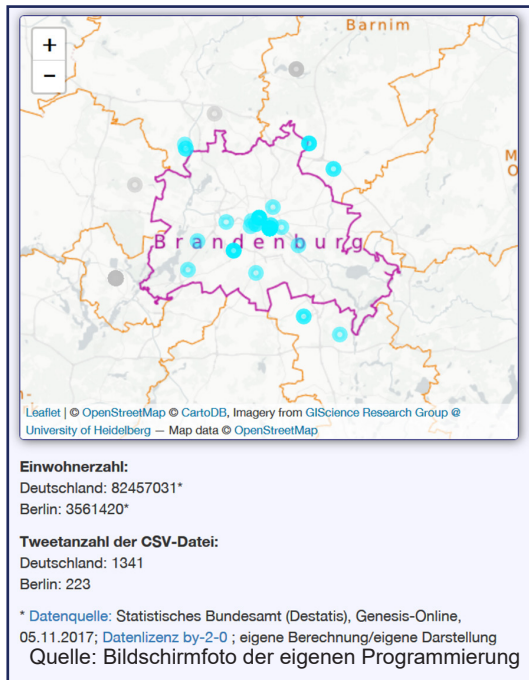
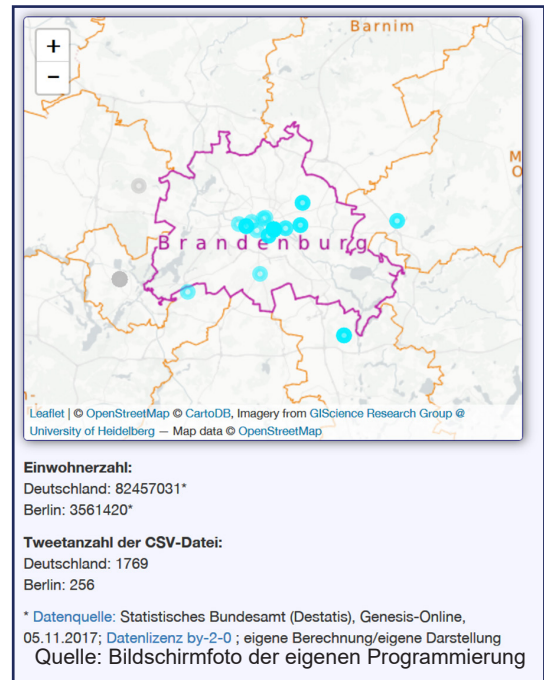


Abbildung C.16: Tweets in Berlin
abends am 10.12.2017

C.5 Daten vom 11.12.2017

Abbildung C.17: Tweets in Berlin
morgens am 11.12.2017Abbildung C.18: Tweets in Berlin
mittags am 11.12.2017Abbildung C.19: Tweets in Berlin
nachmittags am 11.12.2017Abbildung C.20: Tweets in Berlin
abends am 11.12.2017

C.6 Daten vom 12.12.2017

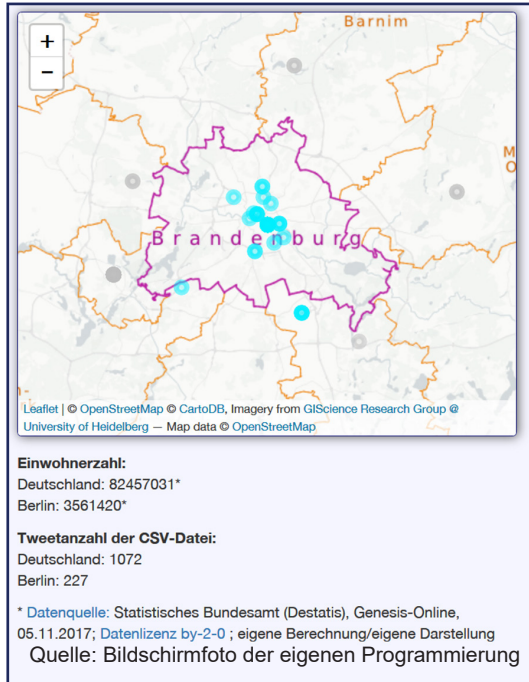


Abbildung C.21: Tweets in Berlin
morgens am 12.12.2017

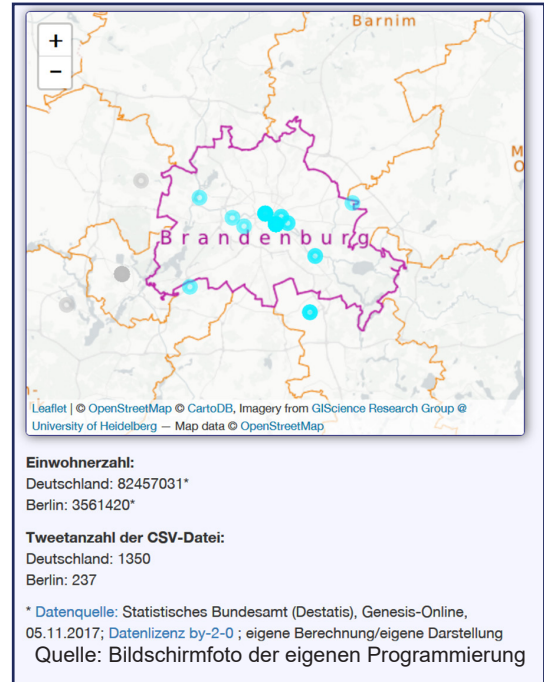


Abbildung C.22: Tweets in Berlin
mittags am 12.12.2017

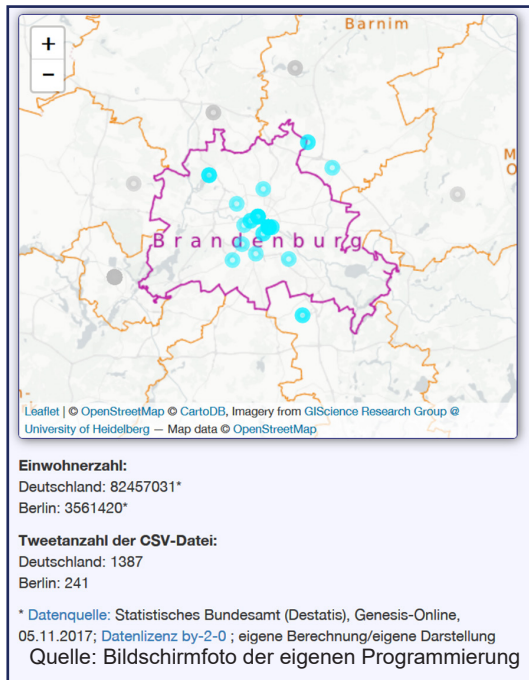


Abbildung C.23: Tweets in Berlin
nachmittags am 12.12.2017

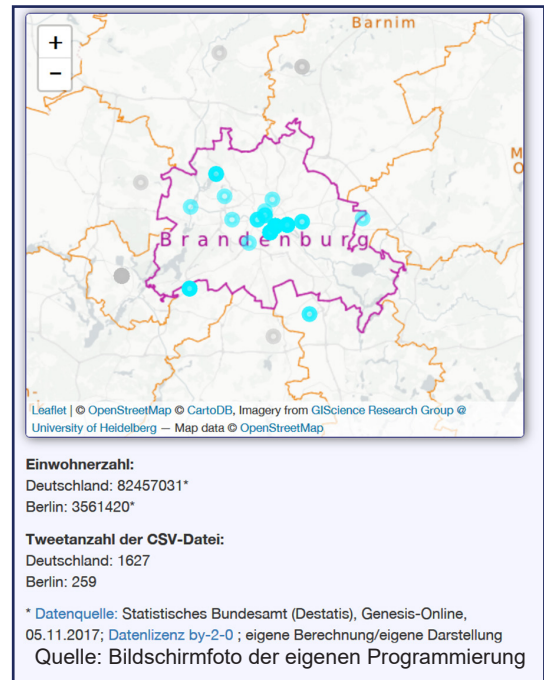


Abbildung C.24: Tweets in Berlin
abends am 12.12.2017

C.7 Daten vom 13.12.2017

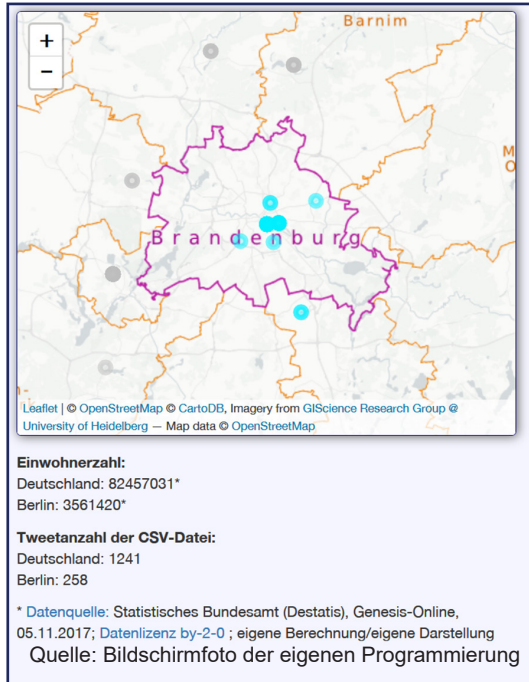


Abbildung C.25: Tweets in Berlin
morgens am 13.12.2017

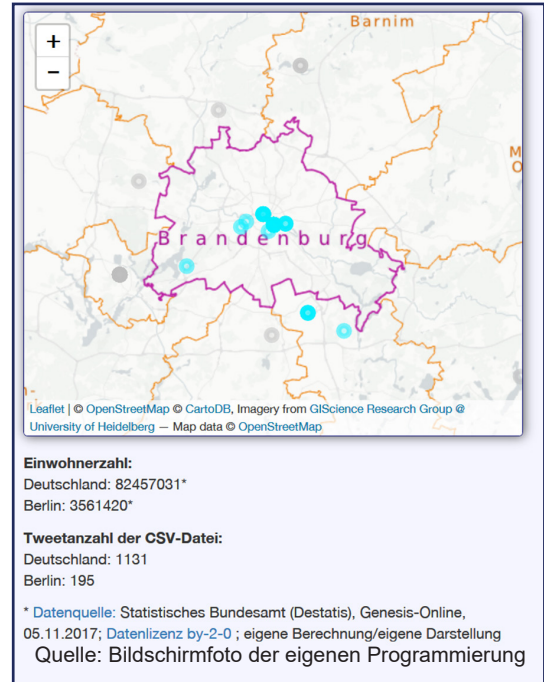


Abbildung C.26: Tweets in Berlin
mittags am 13.12.2017

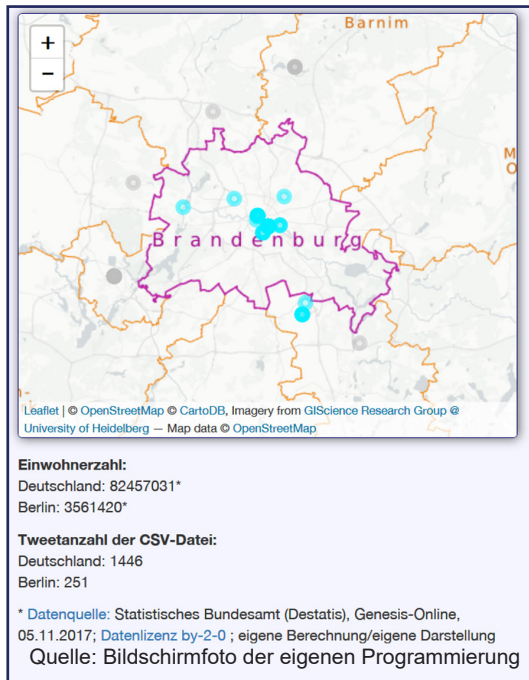


Abbildung C.27: Tweets in Berlin
nachmittags am 13.12.2017

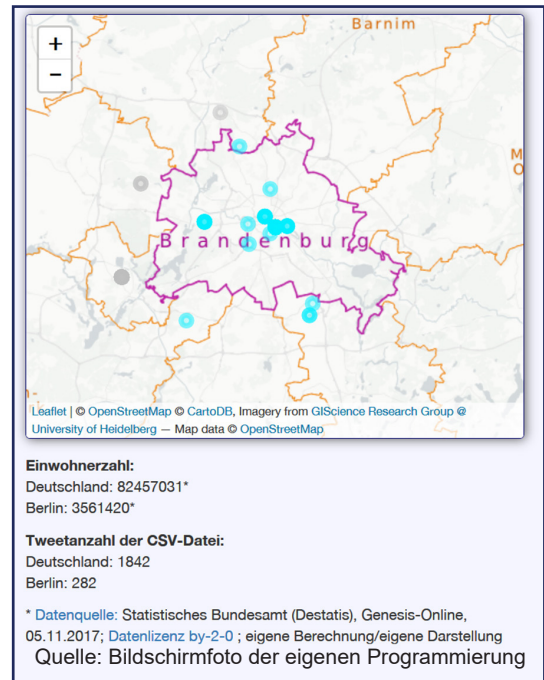


Abbildung C.28: Tweets in Berlin
abends am 13.12.2017

Anlage D: Visualisierungen der Streams vom 07. bis 13.12.2017 in NRW

D.1 Daten vom 07.12.2017

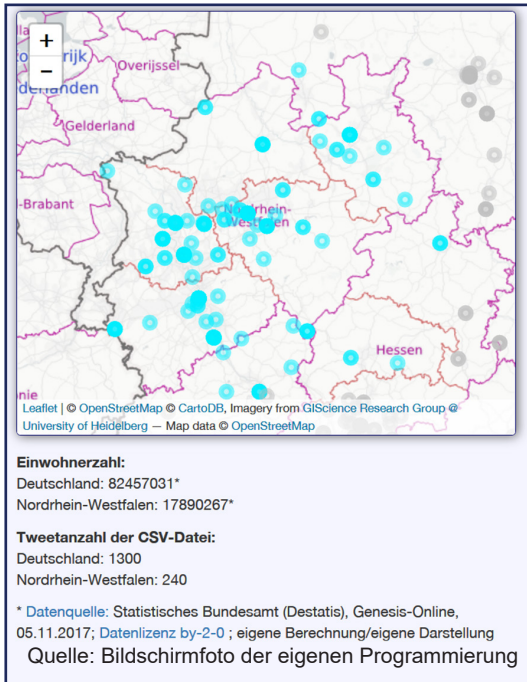


Abbildung D.1: Tweets in NRW morgens am 07.12.2017

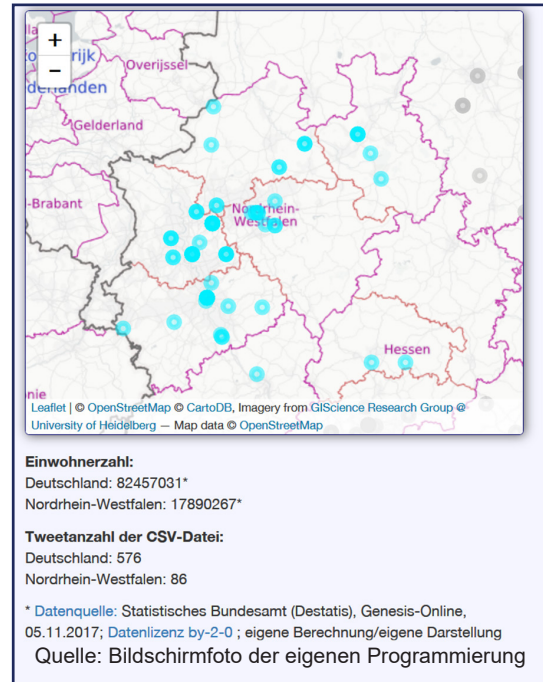


Abbildung D.2: Tweets in NRW mittags am 07.12.2017

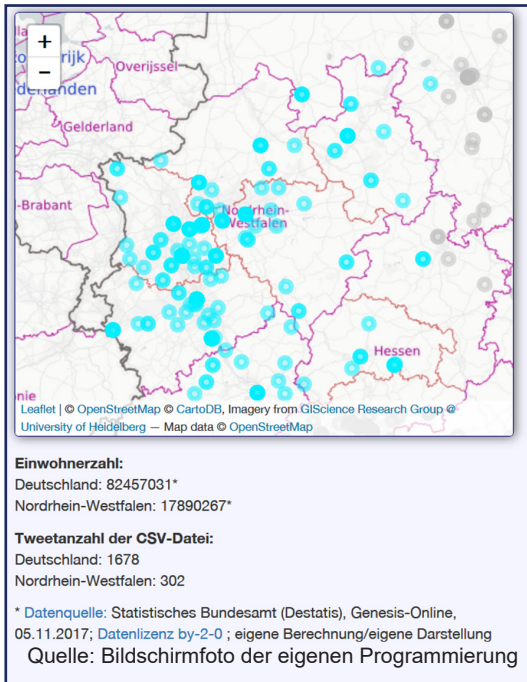


Abbildung D.3: Tweets in NRW nachmittags am 07.12.2017

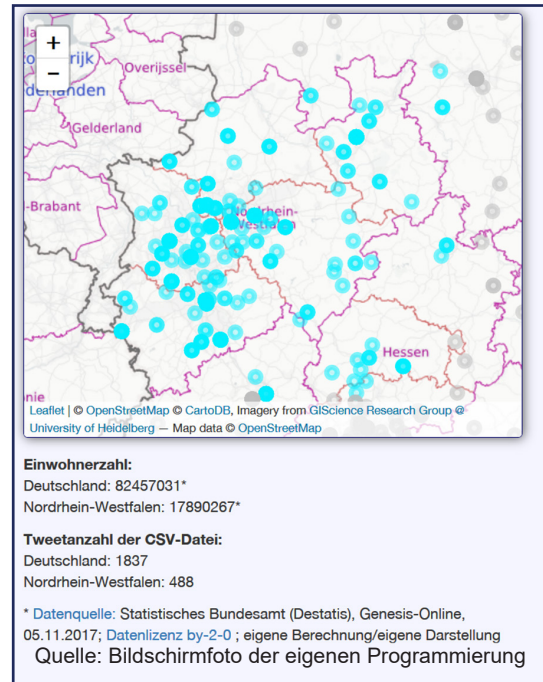


Abbildung D.4: Tweets in NRW abends am 07.12.2017

D.2 Daten vom 08.12.2017

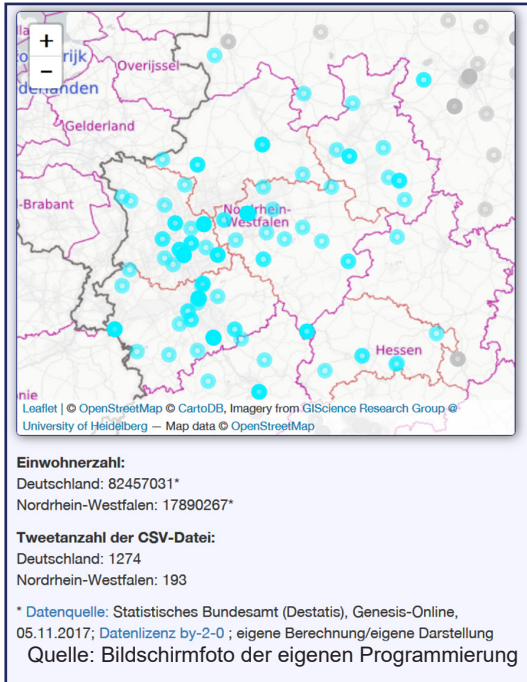


Abbildung D.5: Tweets in NRW
morgens am 08.12.2017

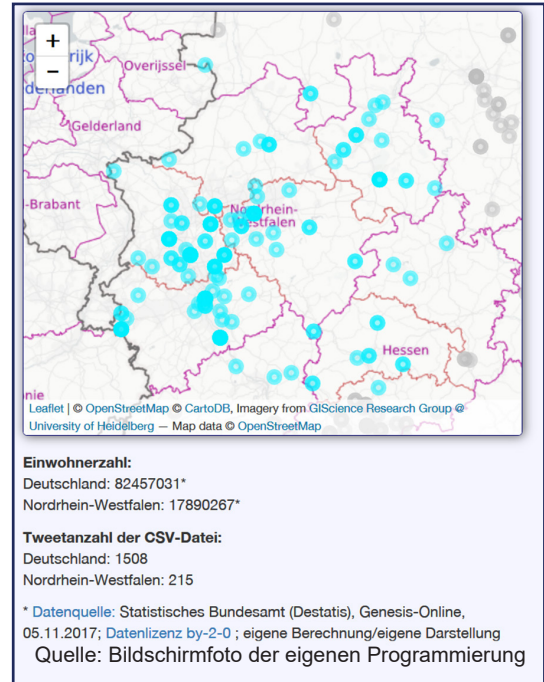


Abbildung D.6: Tweets in NRW
mittags am 08.12.2017

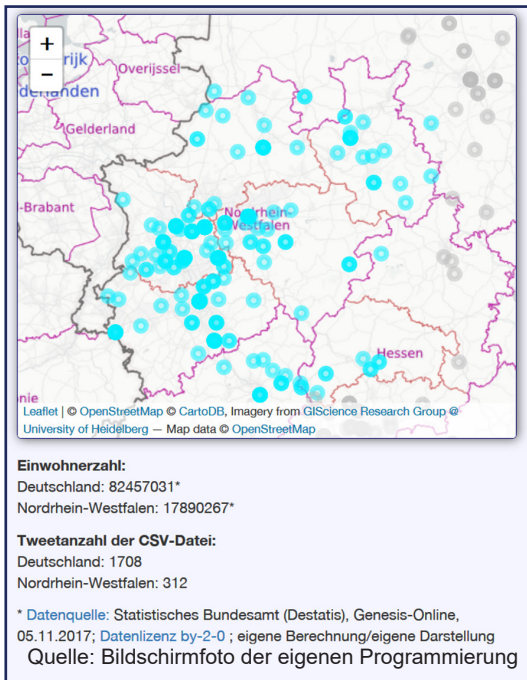


Abbildung D.7: Tweets in NRW
nachmittags am 08.12.2017

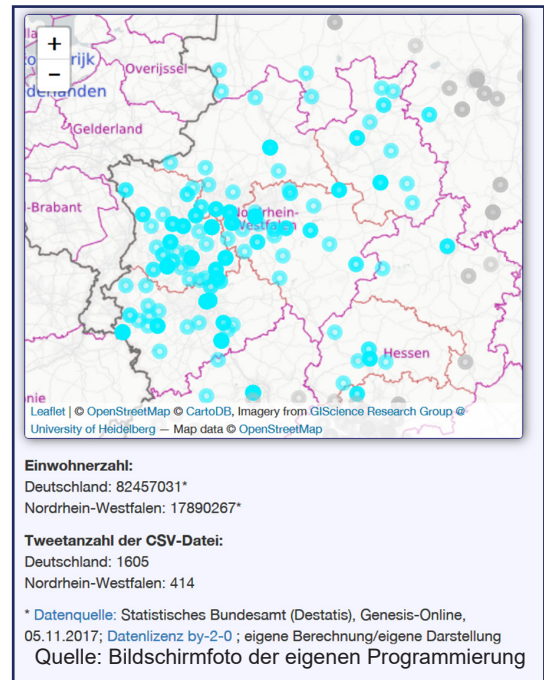


Abbildung D.8: Tweets in NRW
abends am 08.12.2017

D.3 Daten vom 09.12.2017

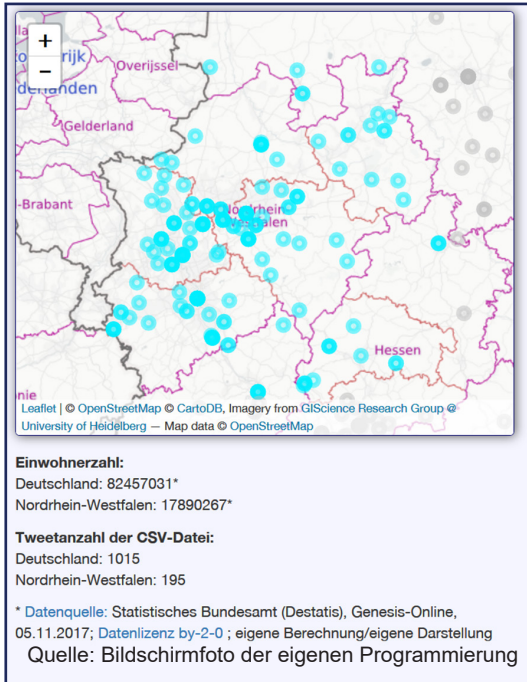


Abbildung D.9: Tweets in NRW
morgens am 09.12.2017

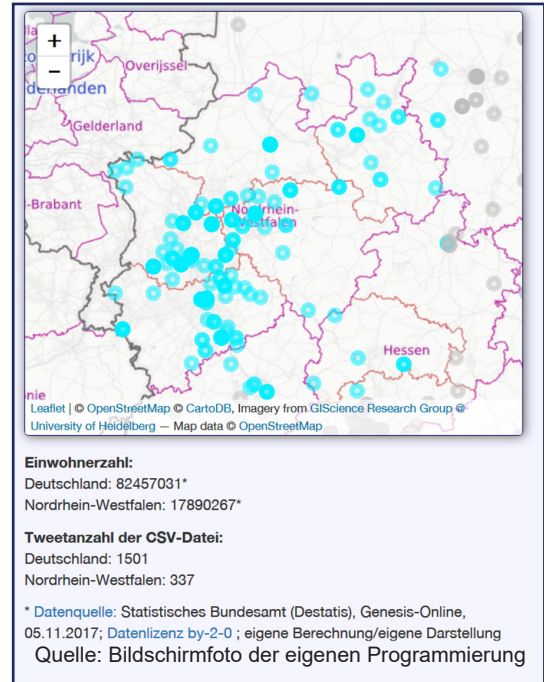


Abbildung D.10: Tweets in NRW
mittags am 09.12.2017

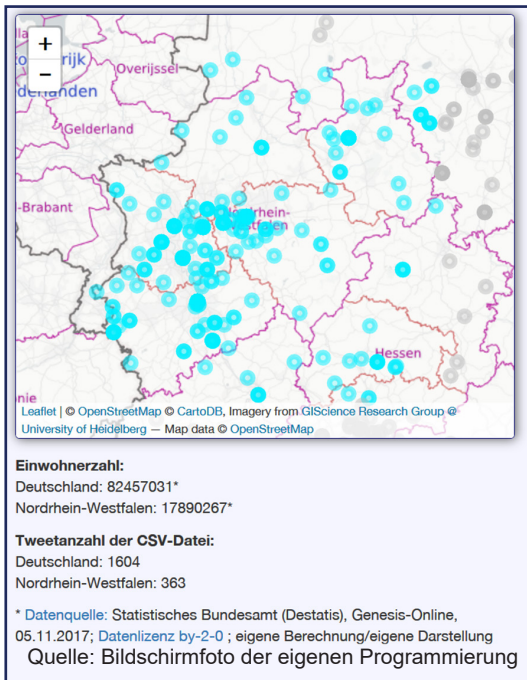


Abbildung D.11: Tweets in NRW
nachmittags am 09.12.2017

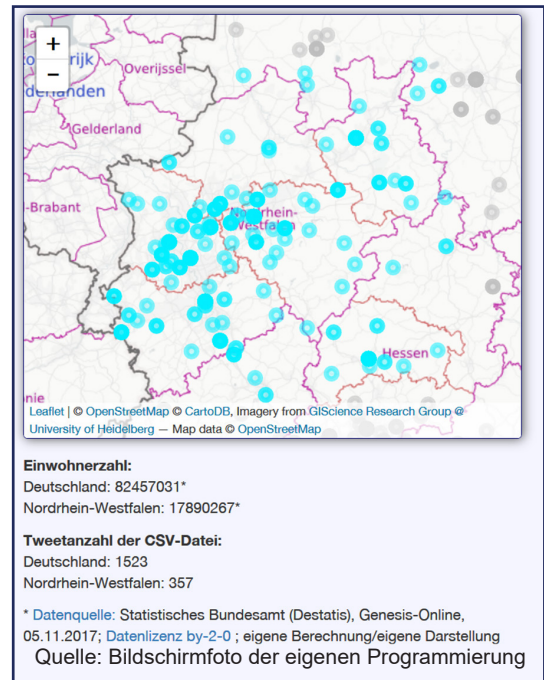


Abbildung D.12: Tweets in NRW
abends am 09.12.2017

D.4 Daten vom 10.12.2017

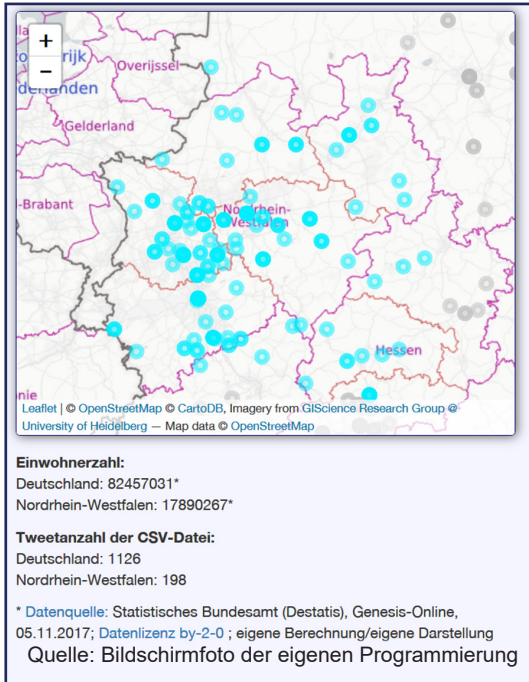


Abbildung D.13: Tweets in NRW
morgens am 10.12.2017

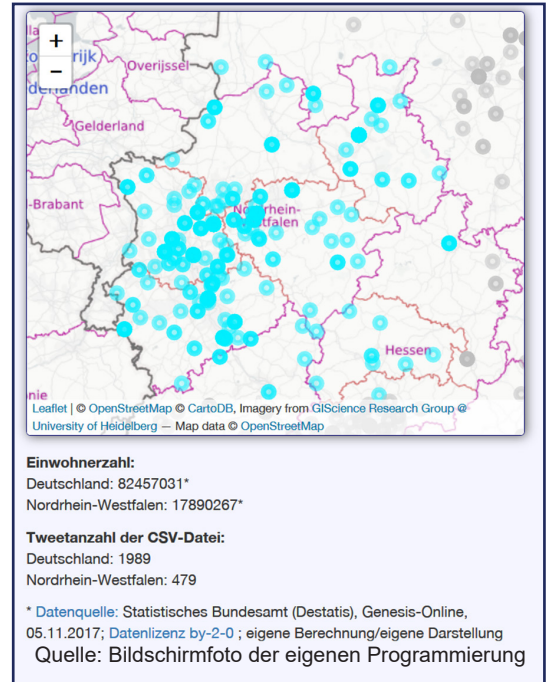


Abbildung D.14: Tweets in NRW
mittags am 10.12.2017

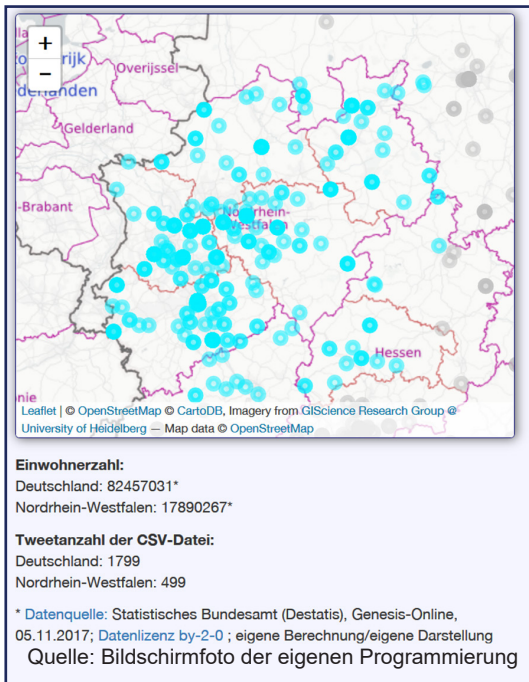


Abbildung D.15: Tweets in NRW
nachmittags am 10.12.2017

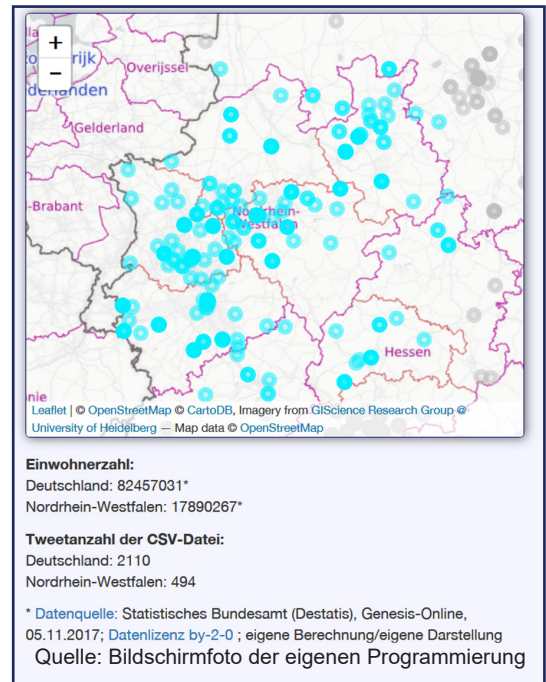


Abbildung D.16: Tweets in NRW
abends am 10.12.2017

D.5 Daten vom 11.12.2017

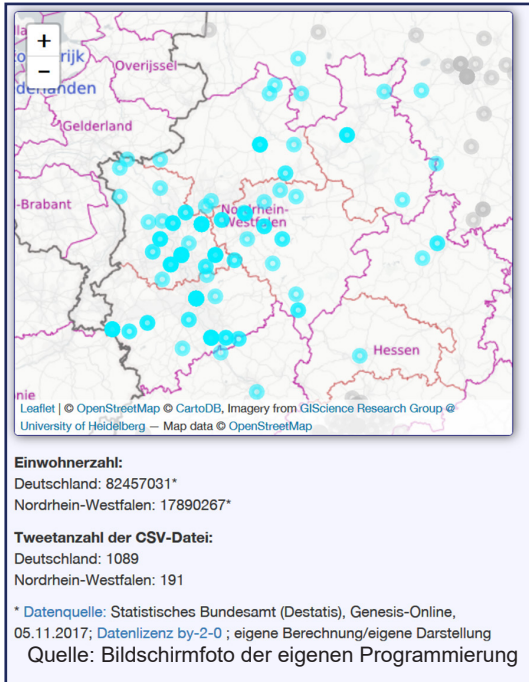


Abbildung D.17: Tweets in NRW
morgens am 11.12.2017

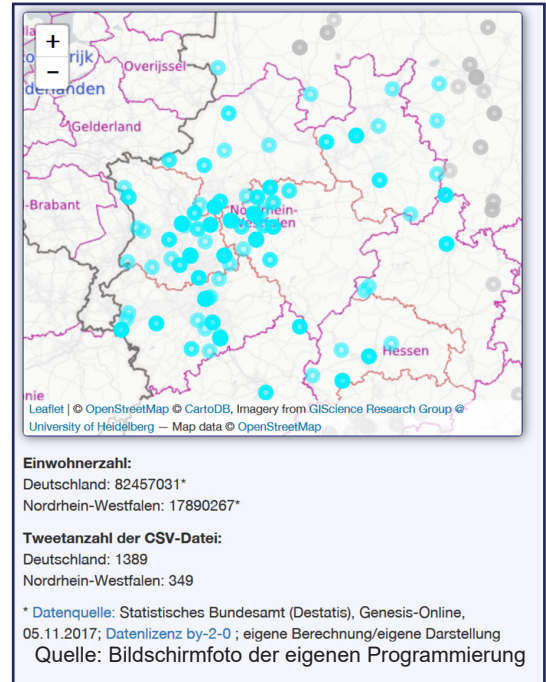


Abbildung D.18: Tweets in NRW
mittags am 11.12.2017

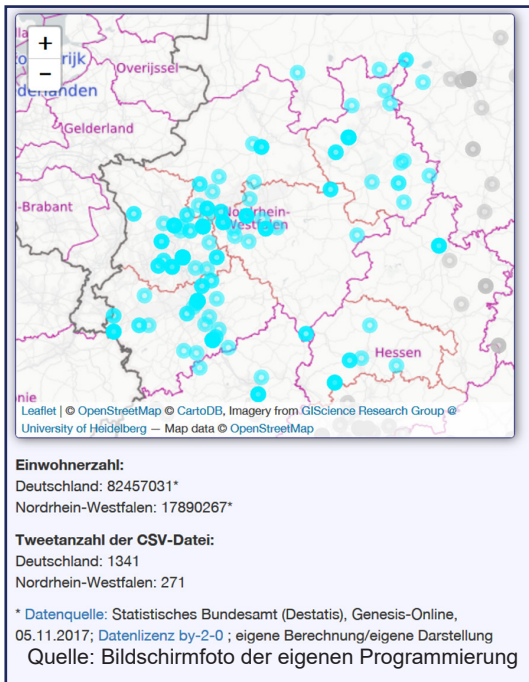


Abbildung D.19: Tweets in NRW
nachmittags am 11.12.2017

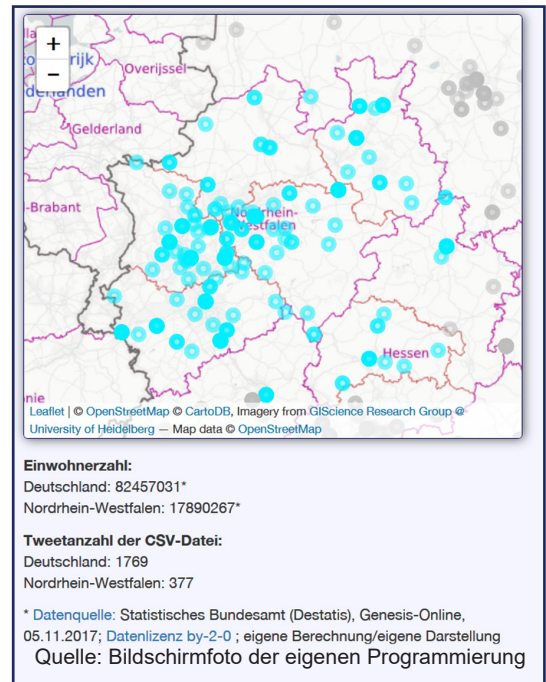


Abbildung D.20: Tweets in NRW
abends am 11.12.2017

D.6 Daten vom 12.12.2017

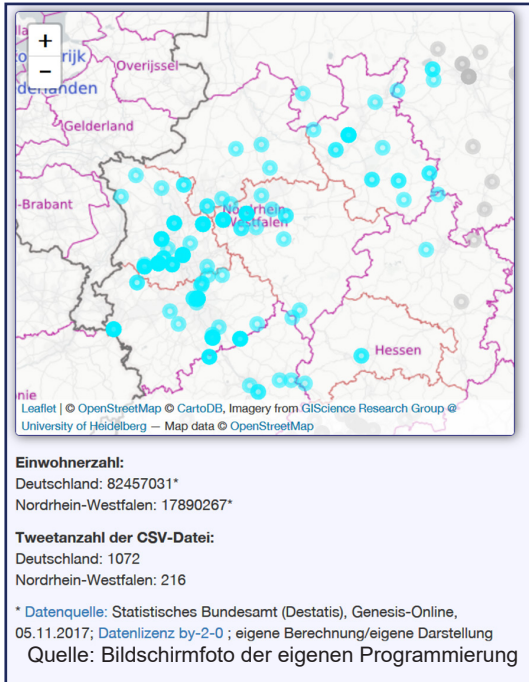


Abbildung D.21: Tweets in NRW
morgens am 12.12.2017

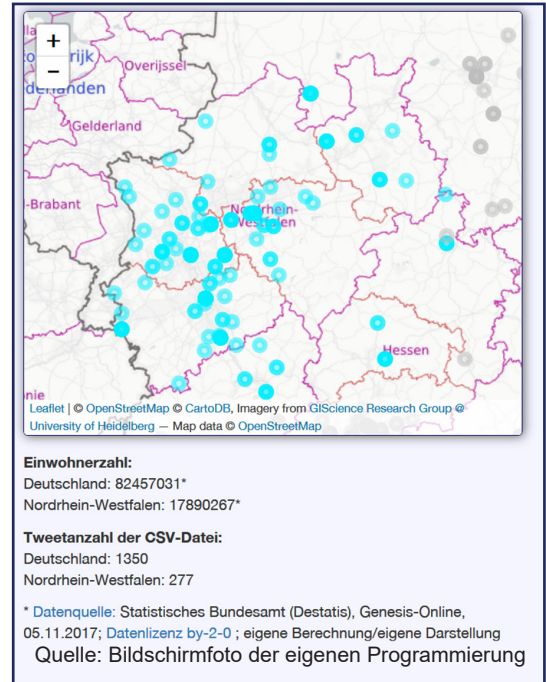


Abbildung D.22: Tweets in NRW
mittags am 12.12.2017

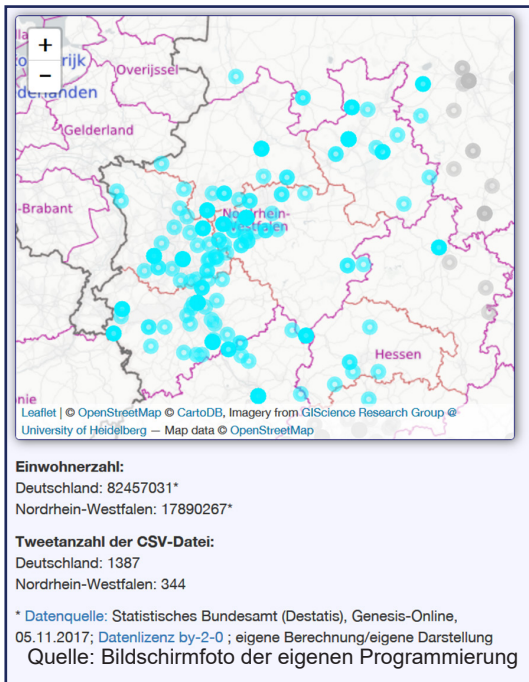


Abbildung D.23: Tweets in NRW
nachmittags am 12.12.2017

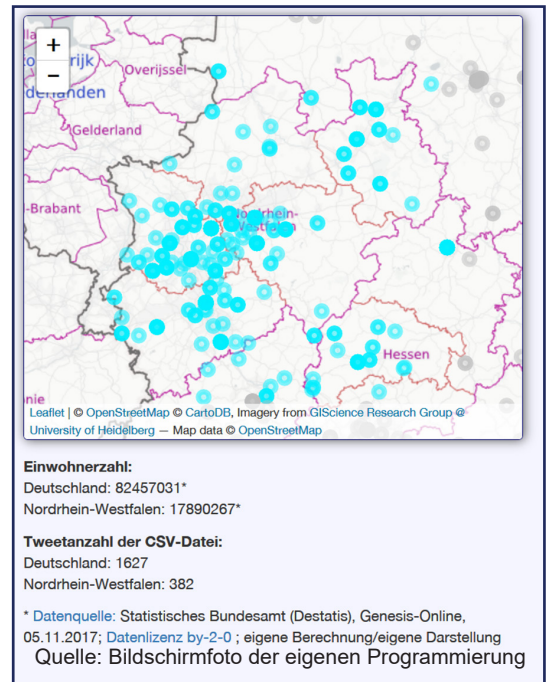


Abbildung D.24: Tweets in NRW
abends am 12.12.2017

D.7 Daten vom 13.12.2017

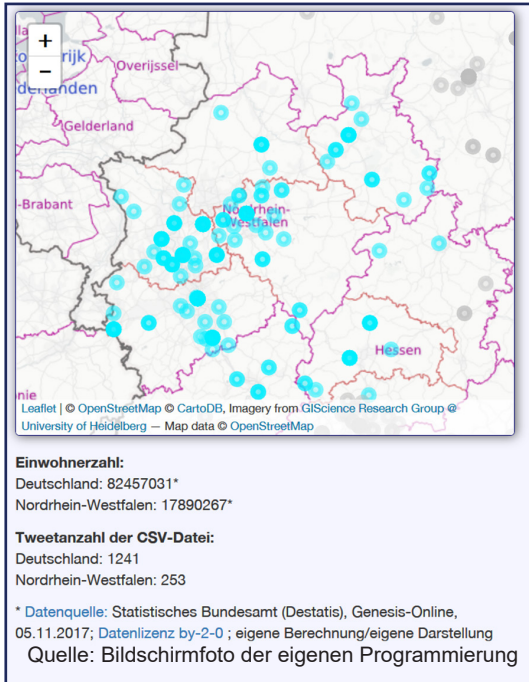


Abbildung D.25: Tweets in NRW
morgens am 13.12.2017

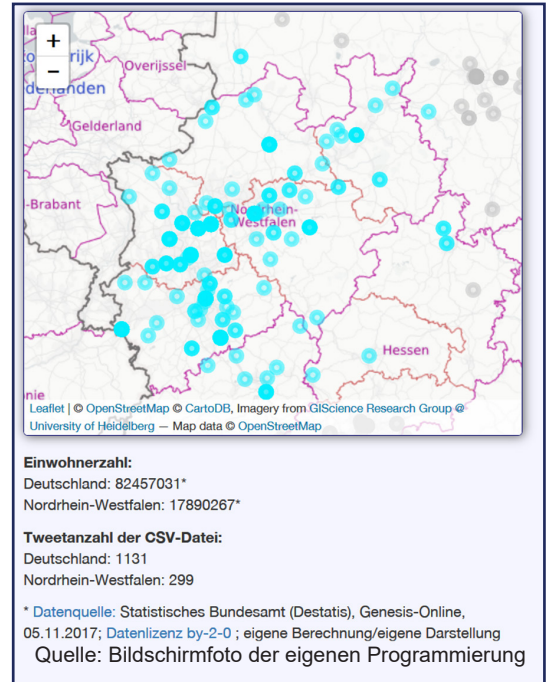


Abbildung D.26: Tweets in NRW
mittags am 13.12.2017

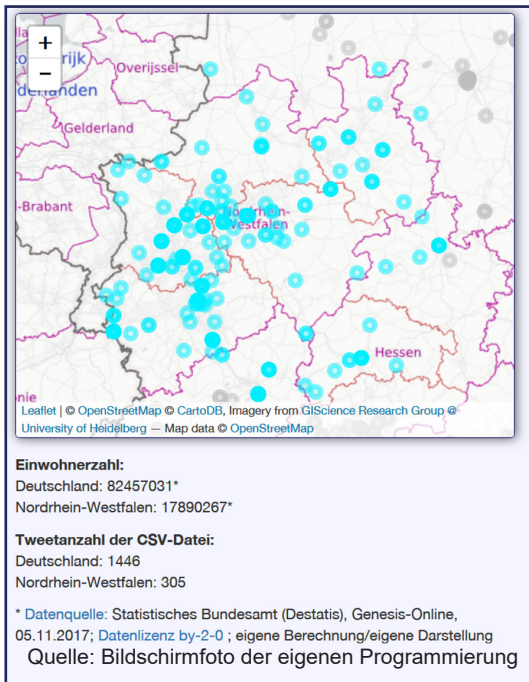


Abbildung D.27: Tweets in NRW
nachmittags am 13.12.2017

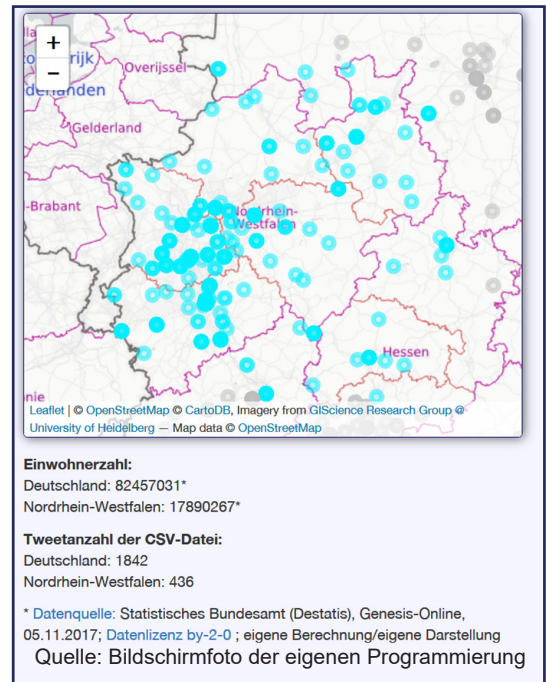


Abbildung D.28: Tweets in NRW
mittags am 13.12.2017

Selbstständigkeitserklärung

Hiermit erkläre ich, Kristina-Susann Baudach, dass ich die hier vorliegende Arbeit selbstständig und ohne unerlaubte Hilfsmittel angefertigt habe. Informationen, die anderen Werken oder Quellen dem Wortlaut oder dem Sinn nach entnommen sind, habe ich kenntlich gemacht und mit exakter Quellenangabe versehen. Sätze oder Satzteile, die wörtlich übernommen wurden, wurden als Zitate gekennzeichnet. Die hier vorliegende Arbeit wurde noch an keiner anderen Stelle zur Prüfung vorgelegt und weder ganz noch in Auszügen veröffentlicht. Bis zur Veröffentlichung der Ergebnisse durch den Prüfungsausschuss werde ich eine Kopie dieser Studienarbeit aufbewahren und wenn nötig zugänglich machen.

Datum

Unterschrift