
Interaktive Namensverbreitungskarten für das Digitale Familiennamenwörterbuch Deutschlands (DFD)

Digital Philology |
Evolving Scholarship in Digital Philology - 04 | 2022

Herausgegeben von
Sabine Bartsch | Evelyn Gius | Marcus Müller | Andrea Rapp | Thomas Weitin

Beate Thull



TECHNISCHE
UNIVERSITÄT
DARMSTADT



Evolving Scholarship in Digital Philology

Impressum

Postadresse: Technische Universität Darmstadt

Institut für Sprach- und Literaturwissenschaft

Dolivostraße 15

64293 Darmstadt

Website: www.digital-philology.tu-darmstadt.de

Email: sprachli@linglit.tu-darmstadt.de



Zitierhinweis: Beate Thull: Interaktive Namensverbreitungskarten für das Digitale Familiennamenwörterbuch Deutschlands (DFD). In: Digital Philology | Evolving Scholarship in Digital Philology 04 | 2022. Darmstadt: TUPrints.

Vorbemerkung

Die nachfolgende Masterarbeit mit gleichem Titel wurde am Fachbereich Gesellschafts- und Geschichtswissenschaften, Institut für Sprach- und Literaturwissenschaft der TU Darmstadt angefertigt (Tag der Einreichung: 08. November 2021, Gutachterinnen: Prof. Dr. Andrea Rapp, Dr. Sabine Bartsch). Für die Publikation wurde der Text der Masterarbeit an einigen wenigen Stellen angepasst. Der in der vorliegenden Arbeit vorgestellte und diskutierte Code wird in GitLab-Repositories der RWTH Aachen vorgehalten und auf Anfrage von der Autorin bereitgestellt.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Problemlage	1
1.2	Ziel	2
1.3	Vorgehen	2
2	Problemstellung	3
2.1	Ausgangspunkt	3
2.2	Projektauftrag	4
2.2.1	Zweck	4
2.2.2	Zielgruppen	4
2.2.3	Ziele und Ergebnisse	6
2.2.4	Erfolgsmessung	6
2.2.5	Vorgehensweise	6
2.3	Recherche	8
2.4	Entwicklungsprozess	10
2.5	Anforderungen an die Visualisierung im Kartenmodul	11
2.5.1	Vision	11
2.5.2	Fragen- und Aufgabenkatalog	11
3	Ansatz	15
3.1	Informationsarchitektur	15
3.1.1	Organisationsschema	15
3.1.2	Site-Diagramm	16
3.1.3	URL-Design	16
3.1.4	Auslegung der Informationsarchitektur	17
3.2	Datenmodell	17
3.3	Visuelle Gestaltung und Implementierung der Anwendungsseiten	19
3.3.1	Datenfluss im Demonstrator	22
3.3.2	Technische Umsetzung der Interaktivität	23
4	Visualisierung der Namensverbreitungen im interaktiven Kartenmodul	26
4.1	Informationsvisualisierung	26
4.2	Visualisierungspipeline	31
4.3	Mapping-Modell für das Design von Visualisierungsideomen	33
4.4	Mapping-Modell für die Verbreitungskarten im DFD	36

4.5	Visualisierungsideom des interaktiven Kartenmoduls	38
4.5.1	Art der zugrunde liegenden Daten	38
4.5.2	Visuelle Kodierung der Daten	40
4.6	Wireframes	41
4.6.1	Startseite mit Namensliste	41
4.6.2	Namenartikel mit Vorschaukarte	42
4.6.3	Anwendungsseite mit interaktiver Verbreitungskarte	42
4.6.4	Parametereinstellungen und Berechnung der Kreise	45
4.6.5	Druckversion	45
4.6.6	Ein- und Ausblenden der Legende	48
4.6.7	Zoomen, Verschieben und Rückkehr zur Ausgangskarte	48
4.6.8	Ortssuche	48
4.6.9	Alternativer Kartenstil	51
4.6.10	Informationstext über einem Kreis und Zoom in die Region	51
4.6.11	Umschalten in den Vollbildmodus	56
5	Evaluation der Visualisierung im interaktiven Kartenmodul	57
6	Interaktive historische Namensverbreitungskarten für das DFD	61
6.1	Erweiterter Aufgabenkatalog für das Design der Visualisierungsideome	61
6.2	Art der verwendeten Daten	64
6.3	Visuelle Kodierung der Daten	70
6.3.1	Historische Namensverbreitungskarten mit Kreisen	70
6.3.2	Namensverbreitungskarte mit Kreisen und einblendbaren Postleitzahlenbereichen	75
6.3.3	Choropleth-Karte	77
6.3.4	Heatmaps	80
6.3.5	Markieren von geographischen Regionen	82
7	Diskussion und Ausblick	84
8	Zusammenfassung	93
A	Wireframes	95
A.1	Interaktive Karte mit veränderten Parametern	95
A.2	Ortssuche	95
A.3	Alternative Kartenstile	95
B	Interaktivität	103
B.1	Umschalten auf Absolute Zahlenwerte	104
B.2	Änderung der PLZ-Stelligkeit	105
B.3	Änderung der Farbe	105
B.4	Änderung der Farbsättigung der Kreisflächen	106
B.5	Legende ein- und ausblenden	106
B.6	Parametereinstellungen zurücksetzen	107

Literatur	109
Abbildungsverzeichnis	111
Tabellenverzeichnis	113
Listings	114

1 Einleitung

Das Forschungsprojekt „Digitales Familiennamenwörterbuch Deutschlands (DFD)“ wurde an der Mainzer Akademie der Wissenschaften und der Literatur in Kooperation mit der Technischen Universität Darmstadt und der Johannes Gutenberg-Universität Mainz im Jahr 2012 begonnen und hat eine voraussichtliche Laufzeit von 24 Jahren.¹ Im Rahmen des DFD wird der komplette Grundbestand der derzeit in Deutschland vorkommenden Namen, auch der sogenannten Fremdnamen, nach neuesten Erkenntnissen der Namenforschung lexikographisch erfasst, kartiert und unter Berücksichtigung der geographischen Verbreitung etymologisiert. Das Ziel des Forschungsprojekts ist die benutzerfreundliche und leicht zugängliche digitale Präsentation von rund 200.000 Namenartikeln. Die digitale Datenbasis wurde auf der Grundlage der Telefonbucheinträge der Deutschen Telekom aus dem Jahr 2005 erstellt. Sie wird dem DFD vom DFG-Projekt Deutscher Familiennamenatlas (DFA) zur Verfügung gestellt.² Jeder Familienname mit mindestens zehn Telefonanschlüssen wird im DFD in einem eigenen Artikel präsentiert, wobei häufige Namenvarianten wie „Meyer“ einen ausführlichen Hauptartikel und weniger frequente Namenvarianten oder Ableitungen wie „Maier“, „Meier“, „Mayerle“ etc. einen Verweisartikel mit Link auf den Hauptartikel erhalten. Der Großteil der im Wörterbuch veröffentlichten Namenartikel enthält eine Namensverbreitungskarte. Das derzeit verwendete Kartierungsprogramm zur Erstellung der Verbreitungskarten, im Folgenden „DFD-Kartierungstool“ genannt, wurde von der Firma Tivano Software GmbH in Neu-Isenburg ursprünglich für das DFA entwickelt. Das Tool wurde zuletzt im Jahr 2014 hinsichtlich einer bereinigten Datenbasis und für die Online-Präsentation der Verbreitungskarten im DFD angepasst (vgl. DFD 2014).

1.1 Problemlage

Die Autor*innen der Namenartikel für das DFD erstellen derzeit jede Verbreitungskarte, die im Wörterbuch präsentiert wird, mithilfe des DFD-Kartierungstools. Die Verbreitung eines Namens wird in der Karte mithilfe von Kreisen unterschiedlicher Größe dargestellt. Je nach intendierter Aussage über die geographische Verbreitung eines Familiennamens in Deutschland wählen die Artikelautor*innen unter mehreren im DFD-Kartierungstool zur Verfügung stehenden Optionen zur individuellen Gestaltung der entsprechenden Verbreitungskarte aus. Die fertige Karte wird als PDF sowie optional als SVG gespeichert und in der Datenbank des DFD abgelegt. Die so erzeugte Namensverbreitungskarte ist statisch

¹<https://www.namenforschung.net/dfd/woerterbuch/liste/>

²<https://www.namenforschung.net/dfa/projekt.html>

und enthält nur rudimentäre geographische Informationen. Benutzerbefragungen zeigen, dass diese statische Karte viele der Fragen, die die Nutzer*innen an die Visualisierung der Namensverbreitungen stellen, nicht beantworten kann.

1.2 Ziel

Das Ziel der vorliegenden Arbeit ist die Entwicklung und Dokumentation einer in die Website des DFD integrierbaren Software, die mithilfe von konsolidierten Technologien und basierend auf den Ergebnissen der aktuellen Forschung im Bereich der Informationsvisualisierung interaktive geographische Namensverbreitungskarten automatisch erzeugt und damit die derzeit im DFD verwendete Art der Kartierung ersetzt.³ Die Entwicklung der Visualisierung und der Interaktivität in den Karten orientiert sich dabei an den Fragen und Anforderungen der Nutzer*innen an die Präsentation der Namensverbreitungen.

1.3 Vorgehen

Das zentrale neue Element des in der vorliegenden Arbeit beschriebenen Software-Experiments ist die Interaktivität in den Namensverbreitungskarten des DFD. Die entsprechende Softwareentwicklung für das im Folgenden so bezeichnete „interaktive Kartenmodul“ für das DFD wird in den folgenden Kapiteln mit der Formulierung des Projektauftrags und der Beschreibung der Benutzerrecherche für das Software-Experiment sowie der Spezifikation und der technischen Umsetzung des Softwareprojekts ausführlich dokumentiert. Die Evaluation der Visualisierung und der Interaktivität im neuen Kartenmodul erfolgt nach den Methoden und Designkriterien der Informationsvisualisierung.⁴ Um die Aussagekraft der Visualisierung der Namensverbreitungen für ausgesuchte Familiennamen weiter zu erhöhen, werden mittelfristig auch interaktive historische Verbreitungskarten zur visuellen Darstellung von historischen Namenverschiebungen im Wörterbuch präsentiert. Es werden erste Studien solcher interaktiven historischen Verbreitungskarten mit im Vergleich zu den Karten im interaktiven Kartenmodul ergänzenden und alternativen Visualisierungs- und Interaktionsmöglichkeiten entwickelt. Die bisherige Softwareentwicklung für die historische Kartierung wird in der vorliegenden Arbeit nicht explizit dokumentiert. Es werden lediglich einige der Kartenstudien unter verschiedenen Gesichtspunkten wie den zusätzlichen Fragen der Nutzer*innen an die Visualisierungen, der Entwicklung des Designs der Visualisierungen und der Interaktivität sowie der Benutzerfreundlichkeit der Karten beschrieben.

³Der Begriff Website umfasst in diesem Zusammenhang alle Komponenten des DFD vom Frontend, in dem die statischen Karten durch interaktive Karten ersetzt werden, über das Web und die Datenbank des DFD bis zum Server-Backend, in dem die entsprechenden Software-Bibliotheken ausgetauscht werden.

⁴Aufgrund der Pandemie-Situation während der Softwareentwicklung war die für die Überprüfung der Funktionshypothese für das Software-Experiment begleitend erforderliche persönliche Beobachtung und Befragung von Nutzer*innen während ihrer Arbeit mit den interaktiven Karten leider nur sehr eingeschränkt möglich.

2 Problemstellung

Die besondere Herausforderung an die hier dokumentierte Softwareentwicklung liegt in der technischen Umsetzung eines neuen Kartenmoduls, bei der die bestehende Kartierungssoftware zur Erzeugung nicht geographischer statischer Namensverbreitungskarten so modifiziert und ausgebaut wird, dass sie den Nutzer*innen des DFD interaktive geographische Verbreitungskarten anbieten und die bisherige Kartierung ersetzen kann. Eine weitere Herausforderung ist die Entwicklung eines dem aktuellen Forschungsstand der Informationsvisualisierung entsprechenden Designs der Visualisierung und Interaktivität in den neuen Verbreitungskarten, das den Fragen und Anforderungen der Nutzer*innen des Wörterbuchs genügen kann. Die in der Art der derzeit im Wörterbuch verwendeten Kartierung begründete passive Rezeption der Darstellung der Namensverbreitung soll durch die Möglichkeit des aktiven Gestaltens der Verbreitungskarte je nach Interessenlage ersetzt werden.

2.1 Ausgangspunkt

Das von der Firma Tivano entwickelte DFD-Kartierungstool ermöglicht die kartographische Darstellung der geographischen Verbreitung einzelner Familiennamen in Deutschland. Mehrere im Tool zur Verfügung gestellte Parameter erlauben den einzelnen Autor*innen eines Namenartikels die Erstellung einer individuell erzeugten Version einer Verbreitungskarte für den bearbeiteten Namen. Die individuellen Parametereinstellungen der Autor*innen legen fest, ob zur Visualisierung der jeweiligen Namendaten die absolute oder die relative Darstellung der Namensvorkommen gewählt wird, ob die Vorkommen in den verschiedenen Postleitzahlenbereichen größerer Städte zu einem Wert zusammengefasst werden sollen, ob die Durchmesser der Kreise, die die Vorkommen repräsentieren, unabhängig von den generellen Größenberechnungen im Tool insgesamt größer oder kleiner dargestellt werden sollen u.a.m. Nach der Fertigstellung legen die Artikelautor*innen die Karte zur Präsentation im DFD-Frontend im PDF- und ggf. zusätzlich im SVG-Format ab (s. Abbildung 2.1). Die mithilfe des DFD-Kartierungstools erzeugte Verbreitungskarte zeigt eine weiße Fläche innerhalb der schwarz dargestellten Grenzen Deutschlands mit einigen durch kleine quadratische Marker mit schwarzem Rand und weißer Innenfläche repräsentierten großen Städten und einigen durch blau dargestellte Linien repräsentierte große Flüsse. Die Karte enthält weder topographische Informationen noch werden die Namen von Städten, Flüssen, Regionen o.ä. angegeben. Die Namensverbreitungen werden mithilfe von rot eingefärbten Kreisen

verschiedener Durchmesser, dargestellt.¹ Jeder Kreis ist im Mittelpunkt eines Postleitzahlenbereichs bzw. je nach vorgenommener Parametereinstellung im DFD-Kartierungstool im Mittelpunkt mehrerer zusammengefasster Postleitzahlenbereiche positioniert. Die Karte ist statisch und nicht interaktiv.

2.2 Projektauftrag

Die oben beschriebenen Betrachtungen der Ausgangslage und Ziele für das Softwareprojekt der Entwicklung eines modernen, interaktiven Kartenmoduls für das DFD führen unmittelbar zur Formulierung eines entsprechenden Projektauftrags. Im Projektauftrag wird der konzeptionelle Rahmen für die Softwareentwicklung festgelegt (vgl. Lynch und Horton 2016, S. 6-9).

2.2.1 Zweck

Mit der Entwicklung einer Funktionsbibliothek für ein interaktives Kartenmodul für das DFD sollen die derzeit im öffentlichen Frontend des Wörterbuchs angebotenen Namensverbreitungskarten durch interaktive geographische Karten ersetzt werden. Durch die Einführung der Interaktivität sollen die Nutzer*innen der Website in die Lage versetzt werden, eigene Verbreitungsszenarien eines Namens zu erzeugen und die angebotenen Daten zu explorieren, ohne dabei die von den einzelnen Autor*innen der Verbreitungskarten intendierten Aussagen über die Verbreitung eines Namens aus dem Auge zu verlieren. Der Vorgang der Kartierung soll automatisiert werden. Das interaktive Kartenmodul soll zu einem späteren Zeitpunkt vollständig in die Website des DFD integriert werden.

2.2.2 Zielgruppen

Die Nutzer*innen der DFD-Website gehören unterschiedlichen Gruppen von Interessierten an, die sich zu einem kleinen Teil aus wissenschaftlich im Umfeld der Namen-, Sprach- oder Historienforschung Arbeitenden, in der Hauptsache jedoch aus Personen ohne oder mit wenig fachlicher Expertise, die verschiedenen Altersgruppen und Nationalitäten angehören, zusammensetzen. Das interaktive Kartenmodul soll beide Zielgruppen gleichermaßen ansprechen. Die Bedienung und das Design der interaktiven Karten darf regelmäßige Besucher der Website, die bis dato nur die statischen Karten des Wörterbuchs kennen, weder abschrecken noch überfordern. Für beide Zielgruppen soll der Mehrwert der interaktiven Karten direkt erfahrbar sein.

¹Einige wenige ältere Verbreitungskarten enthalten grün eingefärbte Kreise. Die Farbe der Kreise ist nicht informationstragend.

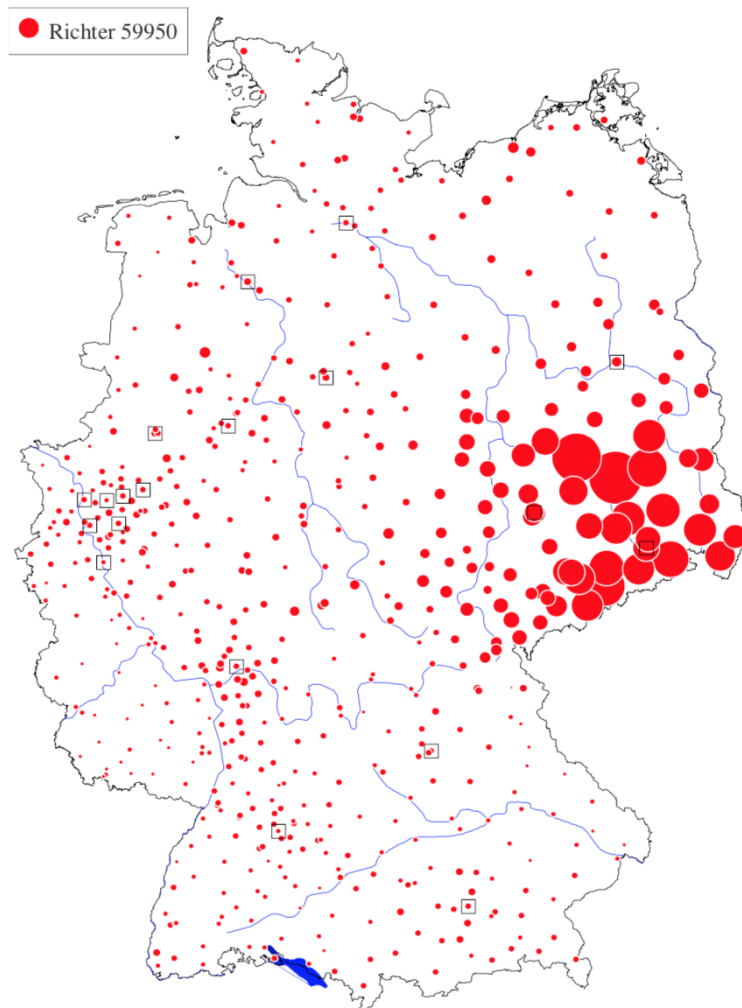


Abbildung 2.1: Beispielansicht einer derzeit im Wörterbuch verwendeten statischen Verbreitungskarte

2.2.3 Ziele und Ergebnisse

Auf der Grundlage der digitalen Datenbasis und der Software-Bibliothek des DFD-Kartierungstools wird mithilfe von konsolidierten Technologien und unter Wahrung der Kompatibilität mit den vorhandenen Softwarekomponenten des DFD eine produktive Software entwickelt, die unmittelbar in die bestehende Website integriert werden und das derzeit verwendete Kartierungstool ersetzen kann. Die nach aktuellem Forschungs- und Technologiestand veralteten Verbreitungskarten im PDF-Format werden durch automatisch erzeugte interaktive, geographische Karten ersetzt. Die Visualisierung und die Interaktivität in den neu entwickelten Namensverbreitungskarten orientieren sich an den Fragen und Anforderungen der Wörterbuchbenutzer*innen an die Visualisierung. Die Nutzer*innen erhalten diverse Möglichkeiten zur Interaktion mit den Verbreitungskarten und damit zur Exploration der Namensverbreitungen. Sie können die Darstellung der Verbreitung eines Namens durch Interaktionen verändern jedoch nicht endgültig überschreiben. Während der Arbeit mit den interaktiven Karten ist die jeweils automatisch erstellte Default-Karte jederzeit abrufbar. So wird sicher gestellt, dass mit der Umstellung auf interaktive Namensverbreitungskarten keinerlei Informationen verloren gehen, die bislang mithilfe der statischen Karten vermittelt wurden. Die Bedienung des Kartenmoduls ist weitgehend selbsterklärend und intuitiv bzw. leicht erlernbar. Während der Arbeit mit den interaktiven Namensverbreitungskarten werden die Nutzer*innen, wo nötig, durch kurze Informationstexte bei der Bedienung unterstützt. Die Präsentation moderner geographischer Karten und die Interaktivität im neuen Kartenmodul sollen einen deutlichen Mehrwert gegenüber der derzeitigen Kartierung im DFD sowohl im Hinblick auf die Fragen der Nutzer*innen ohne fachliche Expertise an die Visualisierung als auch im Hinblick auf die Anforderungen, die bei der Bearbeitung wissenschaftlicher Fragestellungen an die Visualisierungen gestellt werden, darstellen. Sie sollen das Interesse an den erweiterten Informationen über die Namensverbreitungen auf der DFD-Website wecken und zur Attraktivität des Wörterbuchs insgesamt beitragen.

2.2.4 Erfolgsmessung

Im Rahmen des DFD-Projekts werden Analysen zur Nutzung des Digitalen Wörterbuchs durchgeführt.² Die hier ermittelten Angaben zu den Besucherzahlen werden nach erfolgter Integration des interaktiven Kartenmoduls in die Website des DFD auch zur Erfolgsmessung für das Kartenmodul herangezogen. Darüber hinaus werden Rückmeldungen von Nutzer*innen u.a. via Email oder Twitter ausgewertet.³

2.2.5 Vorgehensweise

Bei der Softwareentwicklung für das interaktive Kartenmodul für das DFD liegt neben der Analyse der technischen Anforderungen im DFD-Projekt zunächst der Fokus auf der

²Die regelmäßige Analyse findet zur Zeit mithilfe der Open Source Webanalyse-Plattform PIWIK statt.

³Persönliche Benutzerbefragungen und -beobachtungen werden durchgeführt, sobald die Pandemielage es zulässt.

Kompatibilität mit der bestehenden Software der DFD-Website, um die spätere problemlose Integration des interaktiven Kartenmoduls in die Website des DFD zu gewährleisten. Für die automatisierte Generierung der interaktiven Namensverbreitungskarten wird daher die JavaScript-Bibliothek des DFD-Kartierungstools, die linear und im Batch-Betrieb eine Karte im PDF-Format produziert, nachvollzogen und auf die Anwendung angepasst. Hierbei werden die im DFD-Kartierungstool verwendete, bereits aufbereitete Datenbasis sowie die Art der visuellen Kodierung der Namensvorkommen mithilfe von Kreisen, die Größenberechnungen für die Kreise und deren geographische Verortung in der Karte übernommen. Die auf die Anwendung angepasste Software-Bibliothek (s. Datei `circles.js`) kann zusammen mit JavaScript Event-Handlern und mit Leaflet verwendet werden. Hilfreich für das Nachvollziehen der JavaScript-Bibliothek des DFD-Kartierungstools wäre eine Dokumentation der Softwareentwicklung von Seiten der Firma Tivano. Eine solche Dokumentation liegt jedoch nicht vor. Auch der Quellcode der JavaScript-Bibliothek des Tools liegt im Wesentlichen unkommentiert vor.⁴

Die zentrale Aufgabe des Softwareprojekts ist die Entwicklung einer interaktiven Namensverbreitungskarte für das DFD. Hierzu wird zunächst eine Demo-Version des interaktiven Kartenmoduls für das DFD gebaut. Dieser Demonstrator wird als eigene Anwendung entwickelt, die mit den Softwarekomponenten des DFD verbunden ist. Im Unterschied zur DFD-Website bietet die Demo-Anwendung lediglich eine kleine Auswahl von Familiennamen, die vom DFD-Team als im Hinblick auf Gesamtvorkommen und Namendeutung als repräsentativ bewertet wurde. Die Seiten der Anwendung erscheinen im unveränderten Design des Wörterbuchs. Die Inhalte der entsprechenden Namenartikel werden vom DFD übernommen. Für das interaktive Kartenmodul wird die bisherige Praxis, jede Namensverbreitungskarte mit von den Artikelautor*innen individuell im DFD-Kartierungstool festgelegten Parametereinstellungen zu veröffentlichen, aufgegeben. Dies geschieht aus zwei Gründen. Zum Einen wurden die Parametereinstellungen der Artikelautor*innen in der Vergangenheit nicht systematisch abgespeichert und stehen somit nicht als Berechnungsgrundlage für die automatische Erzeugung der interaktiven Verbreitungskarten zur Verfügung. Zum Anderen wird die Softwareentwicklung für das interaktive Kartenmodul im DFD als Anregung dafür genommen, die automatisierte Kartenerzeugung nach wenigen allgemeinen und einheitlichen Regeln durchführen zu lassen.⁵ Vor der Integration des Kartenmoduls in die Website des DFD werden diese Regeln entsprechend implementiert. Bis dahin wird im Demonstrator die Default-Version einer Karte zunächst mit für zusammengefasste dreistellige Postleitzahlenbereiche berechneten Kreisen in der Farbe Rot mit festgelegter Farbsättigung und einheitlicher Skalierung, d.h. ohne Anpassung der Kreisgrößen in Spezialfällen, automatisch erstellt. Ab einem Namensvorkommen, das größer als 1000 ist, werden die Kreise mit relativen Zahlenwerten berechnet. Die Default-Karte erscheint mit einer eingeblendeten Legende. Den Nutzer*innen werden vielfältige interaktive Funktionalitäten angeboten. Für die Entwicklung der Anwendung werden aktuelle Standards bzw. konsolidierte Technologien eingesetzt. Zur Erzeugung der interaktiven geographischen Karten im modernen Design wird die Open Source JavaScript-Bibliothek Leaflet⁶ eingesetzt. Es werden für den Zweck

⁴Die JavaScript-Bibliothek des DFD-Kartierungstools, `karte.js`, findet sich unter <http://kartierung.dfd.namenforschung.net> (Zugangsdaten erforderlich).

⁵Vgl. dazu die Ausführungen im Kapitel 7

⁶<https://leafletjs.com>, zuletzt abgerufen am 25.09.2021

und die Aussage einer Namensverbreitungskarte passende geographische Hintergrundkarten moderner Kartenanbieter ausgewählt. Bei der Auswahl der entsprechenden Kartenanbieter ist die Vorgabe zu beachten, dass, um Abhängigkeiten zu proprietären Lösungen zu vermeiden und auf aktuelle Entwicklungen flexibel reagieren zu können, das angebotene Kartenmaterial frei verfügbar sein soll und keine laufenden Kosten verursachen darf. Für die Entwicklung des Demonstrators für das interaktive Kartenmodul kommt außerdem die Programmiersprache Python zum Einsatz.⁷

2.3 Recherche

Die zentrale Rolle bei der Entwicklung und Gestaltung der interaktiven Verbreitungskarte für das DFD spielen die Nutzer*innen des Wörterbuchs. Die Nutzer*innen werden an der Formulierung der Aufgaben, die an die Entwicklung der Visualisierungen im Kartenmodul gestellt werden, beteiligt. Sie prüfen die Prototypen des Demonstrators, die während der verschiedenen Entwicklungsphasen entstehen, und liefern so wesentliche Informationen für die aufgabenangemessene Entwicklung der Software (vgl. Dahm 2006, S. 314-316). Für das Softwareprojekt des interaktiven Kartenmoduls muss hierbei unterschieden werden zwischen den (programmier-)technischen Anforderungen an die Softwareentwicklung für das Kartenmodul als Ersatz für das DFD-Kartierungstool - diese Anforderungen umfassen u.a. die Bedürfnisse der Artikelautor*innen mit z.T. speziellen wissenschaftlichen und technischen Fragestellungen - und den Anforderungen des Gros der alltäglichen Nutzer*innen des DFD-Frontends. Die hier und in den folgenden Abschnitten des vorliegenden Kapitels beschriebenen Anforderungen beziehen sich durchgängig auf die Fragen und Aufgaben, die von den Nutzer*innen des Wörterbuchs ohne oder mit wenig fachlicher Expertise an die Visualisierungen in den interaktiven Verbreitungskarten gestellt werden.⁸

Die Partizipation der späteren Anwender*innen der Software hat einen entscheidenden Einfluss beispielsweise auf die spätere schnelle Erlernbarkeit des Umgangs mit der Software, auf die Effizienz der Benutzerinteraktionen oder auf Fragen der Barrierefreiheit. Sie trägt damit zu positiven Benutzererfahrungen, der sogenannten *User Experience*, sowie zur Benutzerfreundlichkeit, der *Usability*, der entwickelten Anwendung bei. Die Demo-Version des interaktiven Kartenmoduls wird während der (Weiter-)Entwicklung regelmäßig den Artikelautor*innen des Wörterbuchs vorgestellt und deren kritische Bewertungen werden in die Softwareentwicklung mit aufgenommen. Die Autor*innen nehmen dabei auch die Rolle der Nutzer*innen des DFD-Frontends ohne fachliche Expertise an. Darüber hinaus werden auch die im Zeitraum der Entwicklung im DFD-Projekt beschäftigten Mitarbeiter*innen, die nicht direkt in den Softwareentwicklungsprozess involviert sind, einige mit dem DFD-Projekt vertraute Wissenschaftler*innen aus dem Bereich der digitalen Sprach- und Literaturwissenschaft sowie weitere Angehörige verschiedener Berufs- und Altersgruppen außerhalb des DFD-Projekts in die Recherche mit einbezogen. Es werden Informationen gesammelt über die Ziele, die die Nutzer*innen bei der Beschäftigung mit der Anwendung verfolgen,

⁷<https://www.python.org>

⁸Weitere Überlegungen zu den Anforderungen der verschiedenen Benutzergruppen des DFD finden sich in den Kapiteln 6 und 7

über die Fragen, die sie an die Visualisierung stellen und über ihre Interessen, Bedürfnisse und Verhaltensweisen. Diese Informationen sind wertvoll für die Softwareentwicklung im Hinblick auf die von den Nutzer*innen erwarteten Funktionalitäten, auf die Vorkenntnisse, die die Nutzer*innen mitbringen, darauf, was die Nutzer*innen bereit und in der Lage sind zu lernen, und darauf, welche technischen und organisatorischen Voraussetzungen jeweils gegeben sind (vgl. Dahm 2006, S. 23). Wöchentliche Besprechungen mit einer Wissenschaftlerin und einem Wissenschaftler, die das Projekt schon seit seinen Anfängen betreuen und die sowohl die Belange der Artikelautor*innen und die Benutzerbelange als auch den technischen Hintergrund des DFD gut kennen, sorgten bislang für die hier notwendige Fokussierung.

Aus den Ergebnissen der Benutzerrecherche werden fiktionale Nutzer*innen, sogenannte Personas, abgeleitet (vgl. Dahm 2006, S. 317). Diese Personas repräsentieren die Eigenschaften der typischen Nutzer*innen des Wörterbuchs. Die Reihenfolge, in der die Personas unten aufgeführt werden, orientiert sich in etwa an dem Anteil, den die entsprechenden Nutzer*innen jeweils an den Besuchen der DFD-Website haben. Die Eigenschaften des mit Abstand größten Teils der Nutzer*innen sind dabei in der erstgenannten Persona repräsentiert.

Johanna (44) Interessiert sich für die Herkunft ihres Nachnamens und besucht Websites, die ihr Auskunft nicht nur über den Ursprung und die Deutung ihres Namens geben können sondern auch darüber, in welchen Regionen Deutschlands der Name verbreitet ist. Auch die Namen von Verwandten und Freunden bezieht sie gerne in ihre Suche mit ein.

David (24) Germanistikstudent, sucht für eine Seminararbeit im Internet nach Informationen über deutsche Nachnamen und deren Verbreitung in verschiedenen Dialektgebieten.

Winfried (69) Widmet sich nach seinem Eintritt in den Ruhestand seinem langjährigen Hobby der Ahnenforschung, die er für seine Familie, aber auch für Freunde und Bekannte betreibt.

Anne (63) Enkelin jüdischer Emigranten nach Amerika, sucht in Deutschland nach ihren Urahnen und deren Nachkommen und findet immer wieder Informationsquellen und Websites, die sie in ihren Recherchen weiter bringen. Ihre Muttersprache ist Englisch, sie kann aber in Deutsch verfasste einfache Texte lesen und verstehen.

Elisabeth (49) Professorin für Germanistik, Forschungsschwerpunkt Namenforschung, interessiert sich für die Namenverschiebungen deutscher Familiennamen aufgrund von historischen Namenwanderungen.

Die Personas definieren durch die jeweilige Rolle, die sie bei der Arbeit mit dem interaktiven Kartenmodul einnehmen, die Fragen, auf die sie gerne eine Antwort hätten, die Aktionen, die sie gerne ausführen möchten, sowie die erhofften Ergebnisse dieser Aktionen und legen so die Grundlage für die sogenannten User Stories für das interaktive Kartenmodul (vgl. Lynch und Horton 2016, S. 81). User Stories sind in der (Alltags-)Sprache der Nutzer*innen formulierte konkrete Anforderungen an die Software wie z.B. „Als Nutzer*in des Digitalen Wörterbuchs möchte ich die Farbe der Kreise in der Karte anpassen, um trotz meiner Farb-Sehschwäche die

Verbreitung eines Namens gut ablesen zu können“. Die User Stories werden gesammelt, nach Themen gefiltert, die allen gemeinsam sind, und es wird nach Lösungsideen gesucht, die auf möglichst viele User Stories eingehen. Anschließend werden die User Stories in konkrete Einzelaufgaben für die Entwicklung der Anwendungssoftware übersetzt. Es ist hierbei weder möglich noch sinnvoll, jede einzelne Idee (sofort) zu verwirklichen. Stattdessen muss jede mögliche Lösung drei Filter durchlaufen. Erstens: Wie sehr wird die Lösung tatsächlich gewünscht? Zweitens: Kann die Lösung erfolgreich umgesetzt werden? Und drittens: Welcher Aufwand bzw. welche Kosten entstehen? (vgl. Lynch und Horton 2016, S. 39)

2.4 Entwicklungsprozess

Das oben beschriebene Konzept der User Stories ist Teil der im Projektmanagement zunehmend verwendeten Vorgehensweise „scrum“ zur agilen Softwareentwicklung, die für die hier dokumentierte Entwicklung der stark benutzerorientierten und auf Interaktivität ausgelegten Anwendung angewendet wird. „Scrum“ beschreibt eine Herangehensweise an ein Softwareentwicklungsprojekt, die der Beobachtung Rechnung trägt, dass die Entwicklung von Software ein fließender Prozess ist, der sich oftmals nur schwer in ein rigide vorgeplantes Gerüst von Anforderungen einpassen lässt. Die agile Softwareentwicklung steht für logische Planungsschritte bei gleichzeitiger Flexibilität, um sich ändernden Anforderungen begegnen zu können. Sie arbeitet mit iterativen, inkrementellen Methoden, legt den Fokus auf die Lösung realer, konkreter Bedürfnisse und Aufgabenstellungen durch die Nutzer*innen, erarbeitet rapide Softwarelösungen in einer Folge von kleinen Projekten und die entwickelte Software wird durch konstante Planungsanpassungen und ein kontinuierliches Fortschreiben der Anforderungen, die durch Anliegen, Bedürfnisse und zurückgemeldete Probleme der Nutzer*innen entstehen, fortlaufend verbessert (vgl. Lynch und Horton 2016, S. 77-83).

In Anlehnung an „scrum“ werden die Ideen und Designs für die Prototypen des Demonstrators mit ihren jeweils bis dahin ausgewählten und entwickelten Komponenten bzw. Features in wöchentlichen Besprechungen innerhalb der kleinen Fokusgruppe diskutiert und vorangebracht (vgl. Krug 2014, S. 113 und Lynch und Horton 2016, S. 44). Hier werden auch spezielle Aufgaben besprochen, die sich während des Entwicklungsprozesses beispielsweise aus der Art der vorhandenen Datenlage oder aus der Tatsache, dass für die Entwicklung des interaktiven Kartenmoduls benötigten Softwareressourcen keinerlei Dokumentation vorliegt, ergeben.

Die hier vorliegende Dokumentation der Softwareentwicklung für das interaktive Kartenmodul ist, wie in der Vorgehensweise „scrum“ üblich, sowohl während als auch im Anschluss an die Softwareentwicklungsphase entstanden. Anpassungen, Erweiterungen und Korrekturen, die nach der Integration des Kartenmoduls in die Website des DFD nötig werden bzw. bereits anvisiert sind, werden zum jeweiligen Zeitpunkt ergänzend dokumentiert.

In der Abbildung 2.2 werden schematisch die einzelnen während der Softwareentwicklung vollzogenen Prozessschritte, die in den folgenden Abschnitten und Kapiteln dokumentiert

werden, aufgezeigt.

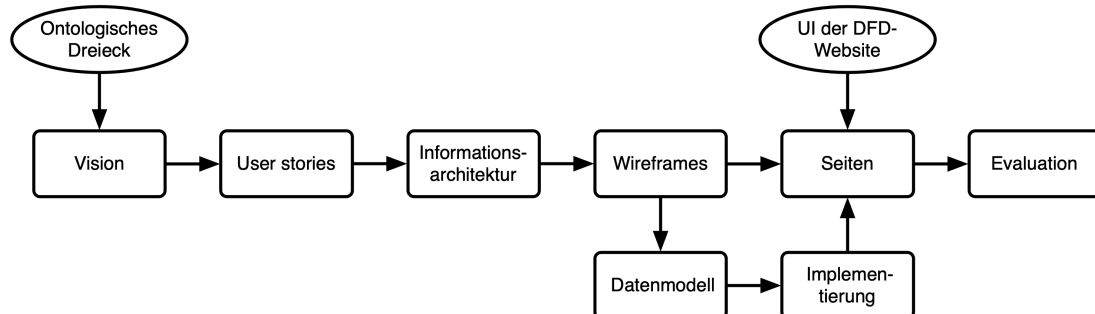


Abbildung 2.2: Entwicklungsprozess für das interaktive Kartenmodul. Rechteckige Kästchen beschreiben die Ergebnisse von Arbeitsschritten und Ellipsen Konzepte oder Ressourcen, die bereits gegeben sind.

2.5 Anforderungen an die Visualisierung im Kartenmodul

In der Vorgehensweise „scrum“ beschreibt die „Vision“ die Zielvorstellung, die durch die Entwicklung der Software erreicht werden soll (vgl. Rosenfeld, Morville und Arango 2015, S. 376). In Anlehnung an das in Abbildung 2.3 dargestellte ontologische Designdiagramm von Gui Bonsiepe enthält die Vision die drei wesentlichen Aspekte Nutzer*in, Aufgabe und Werkzeug (vgl. Bonsiepe 1996, S. 19).

2.5.1 Vision

Den Nutzer*innen des DFD werden interaktive, geographische Namensverbreitungskarten angeboten. Die Interaktivität der Karten bieten dem Gros der Nutzer*innen ohne oder mit wenig fachlicher Expertise, die aus bestimmten privaten Interessen mit den Karten arbeiten, die Möglichkeit, die dem Kartenmodul zugrundeliegenden Daten nach ihrem Interesse zu betrachten, zu explorieren und ggf. nach wissenschaftlichen Gesichtspunkten zu analysieren.

2.5.2 Fragen- und Aufgabenkatalog

Die Auswertung der oben beschriebenen User Stories führt zu einem Fragen- und Aufgabenkatalog, der unmittelbar aus der Vision für das interaktive Kartenmodul abgeleitet wird. Nach dem Filtern der User Stories im Hinblick auf konkrete Aufgaben, die die Softwareentwicklung für das interaktive Kartenmodul erfüllen soll, werden sogenannte Story Cards angelegt. Die Abbildung 2.4 zeigt den typischen Aufbau einer solchen Story Card (vgl. Lynch und Horton 2016, S. 80).

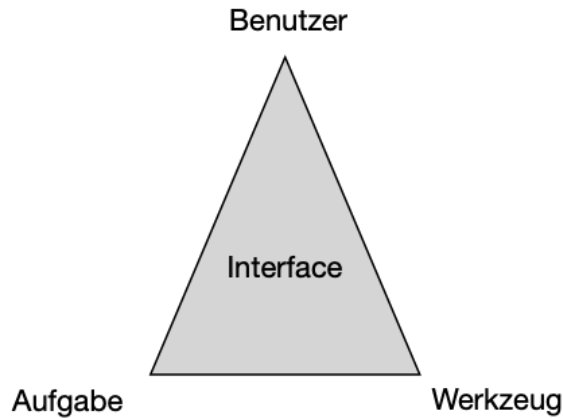


Abbildung 2.3: Das ontologische Designdiagramm nach Gui Bonsiepe. Das Modell setzt sich aus drei Elementen zusammen, die durch die Benutzungsschnittstelle (Interface) verbunden sind: der*die Nutzer*in, der Aufgabe, die bewältigt werden soll, und dem Werkzeug, das zur Lösung der Aufgabe benötigt wird.

User story card

<p>Title:</p> <hr/> <p>As: (User role, user persona)</p> <p>I want to: What action to take</p> <p>So that: What benefit from the action</p> <p>Priority Author Estimate</p>	<p>Title: <i>Create calendar item, all-day event</i></p> <hr/> <p>As: <i>Calendar site user</i></p> <p>I want to: <i>Create an all-day event that doesn't block the view of other shorter events</i></p> <p>So that: <i>I can create the event but not complicate the view of other shorter events on the same day</i></p> <p><i>2</i> <i>Pat</i> <i>6 hrs</i></p> <p>Priority Author Estimate</p>
---	--

Abbildung 2.4: Story Card in der Entwicklungsmethode „scrum“. Quelle: Lynch und Horton 2016, S. 80, CC BY 4.0

Für die Softwareentwicklung für das interaktive Kartenmodul liegen die im Folgenden aufgelisteten Story Cards vor.

Als Nutzer*in des interaktiven Kartenmoduls möchte ich ...

- A1** ... einen speziellen Familiennamen auswählen, um Informationen über den Namen zu erhalten.
- A2** ... neben den allgemeinen Informationen zum Namen eine Überblickskarte zur geographischen Verbreitung des Namens in Deutschland erhalten, um mir ein erstes Bild hiervon machen zu können.
- A3** ... die Möglichkeit haben, eine Verbreitungskarte für den von mir ausgesuchten Namen in vergrößerter Form aufzurufen, um detaillierte Informationen sehen zu können.
- A4** ... eine Auswahl von verschiedenen Hintergrundkarten im Stil der mir bekannten Karten von z.B. OpenStreetMap oder Google Maps erhalten, um mich auf der Karte gut orientieren zu können und die Informationen zur Verbreitung in verschiedenen Darstellungsvarianten zu prüfen.
- A5** ... den Kartenausschnitt verkleinern oder vergrößern und die Karte verschieben, um Details, die mich interessieren, zu finden.
- A6** ... jederzeit zur Karte mit der voreingestellten Zoomtiefe und Zentrierung zurückkehren, um zu jedem Zeitpunkt den Überblick zu behalten.
- A7** ... innerhalb der Karte mithilfe einer Suchfunktion nach bestimmten Orten und Regionen suchen, um die Verbreitung eines Namens an diesen Orten bzw. in diesen Regionen zu prüfen.
- A8** ... Informationen über die Postleitzahlenbereiche und die Anzahl der dort vorhandenen Telefonanschlüsse angezeigt bekommen, um mir ein Bild von der Verbreitung des Namens in den entsprechenden Regionen machen zu können.
- A9** ... die Möglichkeit haben, einen bestimmten Postleitzahlenbereich mit einem Klick heranzoomen, um die entsprechende Region in der Karte zu prüfen.
- A10** ... die Karte mitsamt den Kreisen und innerhalb der Karte angebotenen interaktiven Manipulationsmöglichkeiten in Bildschirmgröße aufrufen, um ausschließlich mit der eigentlichen Verbreitungskarte arbeiten zu können.
- A11** ... die Farbe der Kreise in der Karte anpassen, um trotz meiner Farb-Sehschwäche die Verbreitung eines Namens gut ablesen zu können.
- A12** ... zu jedem Zeitpunkt zu der automatisch erstellten Default-Karte zurückkehren, um die ursprünglich intendierte Aussage der Verbreitungskarte zu erhalten.
- A13** ... eine individuelle Druckversion erstellen, um die Karte als PDF zu exportieren oder auszudrucken.

A14 ... die Einstellungen in der Karte, wie z.B. die Stelligkeit der Postleitzahlenbereiche, ändern, um in anderen Kartenansichten weitergehende Informationen ablesen zu können.

A15 ... eine Legende einblenden können, um Informationen über die Relation zwischen Kreisgrößen und Namensvorkommen zu erhalten.

Die oben aufgeführten Fragen und Aufgaben gehen als Teil des Projektauftrags und der im Folgenden beschriebenen Spezifikation der Anwendung in die Umsetzung des Softwareprojekts mit ein.⁹

⁹Im Rahmen eines erweiterten Aufgabenkatalogs für die historische Kartierung des DFD im Kapitel 6 der vorliegenden Arbeit werden eine Reihe weiterer Aufgaben für die Softwareentwicklung des interaktiven Kartenmoduls aufgeführt. Diese betreffen zu einem Teil speziellere Fragestellungen der Artikelautor*innen an die Visualisierungen.

3 Ansatz

Die Spezifikation des Softwareprojekts erfordert zunächst die Herleitung der Informationsarchitektur für die zu entwickelnde Software (vgl. Lynch und Horton 2016, Kap. 4) sowie die Beschreibung des zugrunde liegenden Datenmodells (vgl. Rosenfeld, Morville und Arango 2015, S. 122). Die entsprechende Umsetzung erfolgt mit der visuellen Gestaltung und Implementierung der Anwendungsseiten und mit der Implementierung der Interaktivität (vgl. Lynch und Horton 2016, S. 230-243).

3.1 Informationsarchitektur

Die Informationsarchitektur der Anwendung, d.h. die Organisation und Kategorisierung von Inhalt und Struktur des interaktiven Kartenmoduls (vgl. Lynch und Horton 2016, S. 59) bildet die Grundlage für die Erstellung eines Interaktionsmodells, das den unterschiedlichen Benutzer*innen des Kartenmoduls eine möglichst intuitive und effiziente Benutzererfahrung ermöglicht. Die Informationsarchitektur zielt darauf ab, in Abhängigkeit von den Besonderheiten der Bedürfnisse der Zielgruppen Inhalte so zu organisieren, dass die Nutzer*innen die Funktionalität der Anwendung leicht erfassen und je nach intendierter Suche nach Informationen einen hohen Recall bzw. eine hohe Precision erwarten können.¹

3.1.1 Organisationsschema

Das Organisationsschema der Anwendung beschränkt sich auf die alphabetische Liste derjenigen im Wörterbuch veröffentlichten Namen, die im Demonstrator für das interaktive Kartenmodul durch elf Namenbeispiele repräsentiert werden (vgl. Rosenfeld, Morville und Arango 2015, S. 116). Die Namenartikel bzw. die interaktiven Verbreitungskarten sind jeweils fest mit einem Namen verbunden und stellen dem Namen zugeordnete Attribute dar. Die Namensliste und die beiden Attribute des ausgewählten Namens werden in jeweils eine Seite der Anwendung überführt. Damit liegen im Organisationsschema für das interaktive Kartenmodul die drei Knoten Namensliste, Namenartikel und Verbreitungskarte vor.

¹Die Wissenschaftler*innen unter den Nutzer*innen werden einen hohen Recall schätzen, weil sie so möglichst viele Daten erhalten. Die in den Personas erwähnte Enkelin jüdischer Emigranten, die auf der Suche nach ihren Vorfahren ist, benötigt eine hohe Precision, da sie ausschließlich an jüdischen Namen interessiert ist.

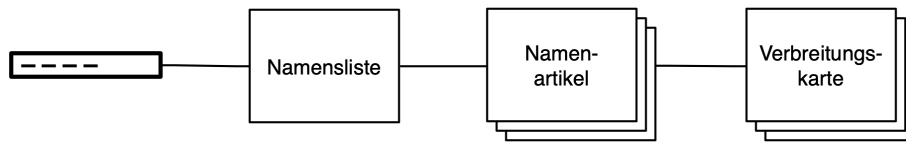


Abbildung 3.1: Site-Diagramm der Anwendung. Der Einstieg in das interaktive Kartenmodul erfolgt über die Seite „Namensliste“. Die Seite „Namenartikel“ wird mit einem Klick erreicht. Die Seite „Verbreitungskarte“ wird mit zwei Klicks erreicht. Innerhalb der Seite „Verbreitungskarte“ werden die Varianten der Karte und die Druckversion mit weiteren Klicks erstellt.

3.1.2 Site-Diagramm

Aus dem Organisationsschema der Anwendung leitet sich eine Informationsarchitektur ab, die den oben aufgeführten Fragen- und Aufgabenkatalog in Form eines Site-Diagramms abbildet (vgl. Lynch und Horton 2016, S. 60 und 129-130, Rosenfeld, Morville und Arango 2015, S. 393-407). Das Organisationsschema bedingt einen Klick-Pfad innerhalb der Anwendung, der einem „drill down“ entspricht, der von der Seite mit der Namensliste über den einzelnen Namenartikel mit Vorschaukarte zur interaktiven Verbreitungskarte einschließlich der Varianten dieser Karte, die durch die Aktionen der Nutzer*innen entstehen, führt. Dieser „drill down“ folgt dem von Ben Shneiderman entwickelten sogenannten Visual Information Seeking Mantra: „Overview first: Gain an overview of the entire collection“, „Zoom: Zoom in on items of interest“, „Filter: filter out uninteresting items“ und „Details-on-Demand: select an item or group and get details when needed“ (vgl. Shneiderman 1996, S. 1 und Shneiderman 1998). Im Site-Diagramm sind die folgenden Seiten dargestellt (s. Abbildung 3.1):

Namensliste Die Seite enthält einen Überblick über die im Wörterbuch veröffentlichten bzw. für den Demonstrator ausgewählten Familiennamen, alphabetisch geordnet in Form einer Liste.

Namenartikel Die Seite enthält die Informationen des DFD zu einem speziellen Namen sowie die Vorschaukarte zur Verbreitung dieses Namens.

Verbreitungskarte Die Seite bietet eine interaktive Karte zur Verbreitung eines Namens an. Die Karte erscheint in der Default-Version. Die Nutzer*innen können Varianten der Verbreitungskarte erzeugen sowie eine Druckversion der Karte erstellen.

3.1.3 URL-Design

Die Tabelle 3.1 zeigt die Auslegung der URLs für die Anwendung. Die URLs der Seiten „Namenartikel“ und „Verbreitungskarte“ enthalten eine ID in Form einer sechsstelligen Zahl. Diese ID wird bei der Abfrage des Namens im DFD-Kartierungstool vergeben.

Auswahl eines Familiennamens	http://.../dfd/
Namenartikel mit Vorschaukarte	http://.../dfd/175876
Interaktive Verbreitungskarte	http://.../dfd/maps/175876

Tabelle 3.1: URL-Design der Anwendung

3.1.4 Auslegung der Informationsarchitektur

Der Einstieg in das interaktive Kartenmodul erfolgt über die Seite „Namensliste“. Die Seite „Namenartikel“ kann über die Auswahl eines Familiennamens auf der Seite „Namensliste“ mit einem Klick erreicht werden. Durch das Anwählen der Vorschaukarte auf der Seite „Namenartikel“ wird die Seite „Verbreitungskarte“ mit insgesamt zwei Klicks erreicht. Innerhalb der Seite „Verbreitungskarte“ werden die Varianten der Karte und die Druckversion mit weiteren Klicks erstellt. Der Weg zu einer erneuten Auswahl eines Namens führt von der Seite „Verbreitungskarte“ über einen Link auf die Seite „Namenartikel“ und einen weiteren Link zurück zur Seite „Namensliste“. Damit sind für beide Wege jeweils zwei Klicks nötig.

Sämtliche Interaktionen zwischen den Nutzer*innen und der Anwendung, die auf der Seite „Verbreitungskarte“ innerhalb der interaktiven Verbreitungskarte oder in den Interaktionsfeldern rechts neben der Karte stattfinden, betreffen die Auslegung der Informationsarchitektur nicht. Sie sind jedoch notwendig, um den Großteil der verschiedenen Fragen und Aufgaben aus dem oben aufgeführten Katalog auf Basis des Site-Diagramms zu lösen und sind insofern ein wesentlicher Bestandteil der Anwendung. Sowohl die auf den drei Anwendungsseiten als auch alle auf der Seite „Verbreitungskarte“ möglichen Interaktionen werden im Abschnitt 4.6 dargestellt.

3.2 Datenmodell

Um die Inhalte der verschiedenen Knoten im oben hergeleiteten Organisationsschema der Anwendung darstellen zu können, sind eine Reihe von Daten notwendig, die sich aus den Inhalten der Anwendungsseiten sowie den verschiedenen Ansichten der interaktiven Verbreitungskarte ableiten lassen (s. Abschnitt 4.6). Da der Softwareentwicklung für das interaktive Kartenmodul aufgrund der Anpassung der DFD-Kartierungstool-Software an die Anwendung die bereits aufbereitete Datenbasis des DFD-Kartierungstools zugrunde liegt, wird das unten aufgeführte Datenmodell für die Anwendung nicht implementiert. Es stellt stattdessen eine Abbildung des Datenmodells des DFD-Kartierungstools dar. Für die hier dokumentierte Entwicklung der Anwendung werden Daten verwendet, die zum Teil dem Wörterbuch direkt entnommen werden. Bis auf einige Bilddaten, die für die Implementierung der interaktiven Funktionalitäten für die Verbreitungskarten benötigt werden, stammen alle weiteren Daten vom DFD-Kartierungstool. Hierbei handelt es sich zum Einen um Daten zu den ein-, zwei-, drei- und fünf-stelligen deutschen Postleitzahlenbereichen. Zum Anderen liegt für jeden abgefragten Namen sowohl eine bei der Abfrage des Namens erzeugte ID sowie

eine Datei mit den Datenwerten zur Größe des Namensvorkommens in allen fünfstelligen Postleitzahlenbereichen vor.

Die Listings 3.1 und 3.2 zeigen jeweils einen Auszug aus den Dateien `plz_5.csv` bzw. `plz_3.csv`. Die Einträge in den Spalten „unmerged“ und „merged“ stehen dabei im Bezug zur Kartierungsoption „Städte zusammenfassen“.

Listing 3.1: Auszug aus der Datei `plz_5.csv`

```
key, name, lon, lat, unmerged, merged
01445, 01445 Radebeul, 13.6432842540266, 51.1136319459259, true, true
01454, 01454 Radeberg, 13.9174778607087, 51.1322204297424, true, true
01458, 01458 Ottendorf-Okrilla, 13.8458463274789, 51.1806998157721, true, true
01462, 01462 Cossebaude, 13.6233072524789, 51.0666308149292, true, true
...
...
99427, 99427 Weimar, 11.3056283760774, 51.0127435034417, true, false
99428, 99428 Weimar, 11.2339667234831, 50.9821024119995, true, false
99438, 99438 Weimar, 11.2744111760922, 50.9119524018455, true, false
99441, 99441 Weimar, 11.4333514977135, 50.9412188048523, true, false
```

Listing 3.2: Auszug aus der Datei `plz_3.csv`

```
key, name, lon, lat, unmerged, merged, , , , ,
...
531**-53199, 531** Bonn, 7.10142739785236, 50.6965737131347, true, false, , , , ,
532**-53299, 532** Bonn, 7.1541577868705, 50.7404222195408, true, false, , , , ,
597**-59799, Arnsberg, 7.95185297006225, 51.4342855959492, true, false, , , , ,
603**-60399, PLZ-Gebiet 603**, 8.71580879252431, 50.1328776351954, true, false, , , , ,
604**-60499, 604** Frankfurt am Main, 8.65593990150028, 50.1667993819462, true,
false, , , , ,
605**-60599, 605** Frankfurt am Main, 8.62950428792459, 50.0696372227566, true,
false, , , , ,
651**-65199, Wiesbaden, 8.21479626567643, 50.0935470800927, true, false, , , , ,
652**-65299, PLZ-Gebiet 652**, 8.25440186932464, 50.1001884649613, true, false, , , , ,
659**-65999, 659** Frankfurt am Main, 8.54520750657277, 50.101568728816, true, false
', , , , ,
681**-68199, 681** Mannheim, 8.48239464833631, 49.4833461052544, true, false, , , , ,
682**-68299, 682** Mannheim, 8.5493557241793, 49.457892040948, true, false, , , , ,
...
```

Das Listing 3.3 zeigt einen Auszug aus der Datei `karte.csv` für den Familiennamen Richter. Jeder Datensatz in der Tabelle enthält die vollständigen Informationen zum zahlenmäßigen Vorkommen des ausgewählten Familiennamens und zum Vorkommen aller Familiennamen insgesamt in fünfstelligen Postleitzahlenbereichen. Die entsprechenden Datenwerte für die ein-, zwei- und dreistelligen Postleitzahlenbereiche werden jeweils mithilfe einer Aggregationsfunktion aus den in der Tabelle enthaltenen Daten für die fünfstelligen Postleitzahlenbereiche errechnet. Die Stern-Symbole in den Listings 3.1, 3.2 und 3.3 stehen - soweit sich die Daten ohne vorliegende Dokumentation korrekt entschlüsseln lassen - im Zusammenhang mit der Aggregation der Daten und der Option „Städte zusammenfassen“.

Listing 3.3: Auszug aus der Datei `karte.csv` für den Namen Richter

PLZ-Bezirk;Gesamt;Richter

```
01***;6126;34
01067;3581;37
01069;7874;77
01097;3529;32
01099;7094;42
01108;1672;25
```

Mithilfe der beschriebenen Daten werden alle Berechnungen für die Informationsvisualisierung in der interaktiven Verbreitungskarte durchgeführt. Die Daten bilden damit die Grundlage sowohl für die Erzeugung der Default-Version der interaktiven Verbreitungskarte als auch der durch die Benutzerinteraktionen erzeugten Varianten dieser Karte.

Das Listing 3.4 zeigt einen Auszug aus der Datei `karte.properties` für den Namen Richter. In dieser Datei sind die Default-Parametereinstellungen für die interaktive Karte für den Namen Richter zu sehen.

Listing 3.4: Auszug aus der Datei `karte.properties` für den Namen Richter

```
...
query.title = Richter
...
query.id = 173014
...
piechartLayer.visible = true
piechartLayer.mergeCities = true
piechartLayer.relativeStatistics = true
piechartLayer.level = 3
...
legendLayer.visible = true
...
```

Die Tabelle 3.2 zeigt die jeweiligen Seiten der Anwendung mit deren Inhalten und den zur Berechnung und Darstellung dieser Inhalte notwendigen Daten. Die dem Wörterbuch entnommenen Daten sind in der Tabelle mit dem Zusatz (w) gekennzeichnet. Alle Daten, die mithilfe des DFD-Kartierungstools bzw. im Rahmen der angepassten Kartierungssoftware zur Verfügung gestellt werden, sind mit dem Zusatz (k) versehen. Das Akronym „PLZ“ steht jeweils für den Begriff „Postleitzahlenbereich“.

Werden die Daten in Tabelle 3.2 im Hinblick auf Entitäten und die ihnen zugeordneten Attribute analysiert, so erhält man die Entitäten „Name“, „Häufigkeit“ und „PLZ“, denen die erforderlichen Daten als Attribute zugeordnet werden können. Die Abbildung 3.2 zeigt das aus den Vorüberlegungen resultierende Entity-Relationship-Diagramm für das interaktive Kartenmodul (vgl. Rosenfeld, Morville und Arango 2015, S. 122-126).

3.3 Visuelle Gestaltung und Implementierung der Anwendungsseiten

Im Hinblick auf die spätere Integration des interaktiven Kartenmoduls in die Website des DFD werden für die visuelle Gestaltung der Anwendungsseiten des Demonstrators das

Knoten	Inhalt	Daten
Namensliste	Alphabetisch geordnete Liste der Familiennamen	Familiennamen (w) Rang des Namens (w) Häufigkeit des Namens (w)
Namenartikel	Informationen zum Namen und Vorschaukarte	Familiennamen (w) Rang des Namens (w) Häufigkeit des Namens (w) Sprachvorkommen (w) Hauptverbreitung (w) Etymologie (w) Verbreitung mit Vorschaukarte (w, k) Verwandte Artikel (w)
Verbreitungskarte	Namensverbreitungen in der interaktiven Karte	Geodaten des PLZ-Mittelpunkts (k) Zahlencode des PLZ (k) Bezeichnung des PLZ (k) Namenvorkommen gesamt im PLZ (k) Namenvorkommen individuell im PLZ (k)

Tabelle 3.2: Knoten im Organisationsschema, deren Inhalte und hierfür notwendige Daten. Für die Attribute Hauptverbreitung, Sprachvorkommen und Verwandte Namen sind Mehrfachbenennungen möglich. Der Einfachheit halber wird hier auf eine entsprechende Differenzierung verzichtet.

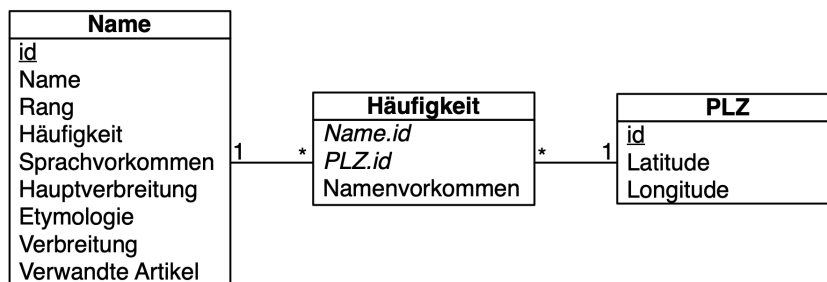


Abbildung 3.2: Entity-Relationship-Diagramm des interaktiven Kartenmoduls

Layout des Wörterbuchs mit dem Kopf- und Fußteil der Wörterbuchseiten, die Inhalte der Startseite und der Namenartikelseiten sowie der Stil des Wörterbuchs übernommen (s. Datei `layout.html`). Das im Wörterbuch in die Startseite integrierte Feld mit der alphabetischen Namensliste wird auf der Startseite des Demonstrators durch ein Feld mit den alphabetisch aufgelisteten elf für den Demonstrator ausgewählten Familiennamen ersetzt (s. Abbildung 4.5). Auf der Namenartikel-Seite tritt außerdem eine kleine, nicht interaktive Vorschaukarte der interaktiven Verbreitungskarte an die Stelle des derzeit in den Namenartikel eingebundenen Vorschaubildes der statischen Verbreitungskarte (s. Abbildung 4.6). Diese Vorschaukarte wird beim Aufruf der Seite mithilfe des Skripts `map.js` generiert. Ein Vorteil der mit einem Skript erzeugten Vorschaukarte gegenüber dem Vorschaubild im gif-Format ist, dass die Darstellung der Karte im Hinblick auf die Auswahl der Hintergrundkarte, die Strichbreite und Farbe der Umrisse der Kreise sowie die Farbe und Transparenz der Kreisinnenflächen mit wenigen Änderungen im JavaScript Code angepasst werden kann. Der Inhalt der Anwendungsseite, die die interaktive Verbreitungskarte enthält, setzt sich zusammen aus dem Kopf- und Fußteil der Wörterbuchseiten, einem Link zurück zum Namenartikel, einem Textfeld mit dem Familiennamen und dem Gesamtvorkommen des Namens sowie der interaktiven Verbreitungskarte mit den Interaktionsfeldern rechts neben der Karte (s. Abbildung 4.7).

In einigen Details werden die Anwendungsseiten und die interaktive Verbreitungskarte mithilfe der Stylesheet-Sprache Cascading Style Sheets (CSS) in einem externen Style Sheet zusätzlich gestaltet (s. Datei `maps.css`) (vgl. Lynch und Horton 2016, S. 138). Alle mit den Originalinhalten des Wörterbuchs übernommenen Eingabefelder und Links sind bis auf den Back-Link auf der Namenartikelseite in der Demo-Version des interaktiven Kartenmoduls inaktiv.

Für die Implementierung der Anwendungsseiten wird, wie oben bereits erwähnt, neben der angepassten JavaScript-Bibliothek des DFD-Kartierungstools die Open Source JavaScript-Bibliothek Leaflet zur Erzeugung interaktiver Karten für das Web verwendet. Leaflet ist eine aktuell weit verbreitete Bibliothek mit einer lebendigen Entwickler-Community. Leaflet Tile Provider² wie das Open Source Projekt OpenStreetMap stellen mit den Kartenanbietern Mapbox³ und Carto⁴ kostenfreies Kartenmaterial für Leaflet zur Verfügung.⁵ Für die Visualisierung der Namensverbreitungen werden mithilfe von Leaflet sogenannte Layer erstellt, die als zusätzliche Ebenen über die als Hintergrundkarten ausgewählten geographischen Karten der Provider gelegt werden. Leaflet bietet eine Fülle von Funktionen und ergänzenden Plugins für die Implementierung der Interaktivität.⁶ Neben den verwendeten JavaScript-Bibliotheken werden außerdem die Auszeichnungssprache HTML und die Programmiersprache Python für die Implementierung des Demonstrators eingesetzt.

²<https://leaflet-extras.github.io/leaflet-providers/preview/>

³<https://www.mapbox.com>

⁴<https://www.carto.com>

⁵Weitere Informationen zur Kostengestaltung der Anbieter Mapbox und Carto finden sich auf den jeweiligen Firmen-Websites.

⁶<https://leafletjs.com/plugins.html>

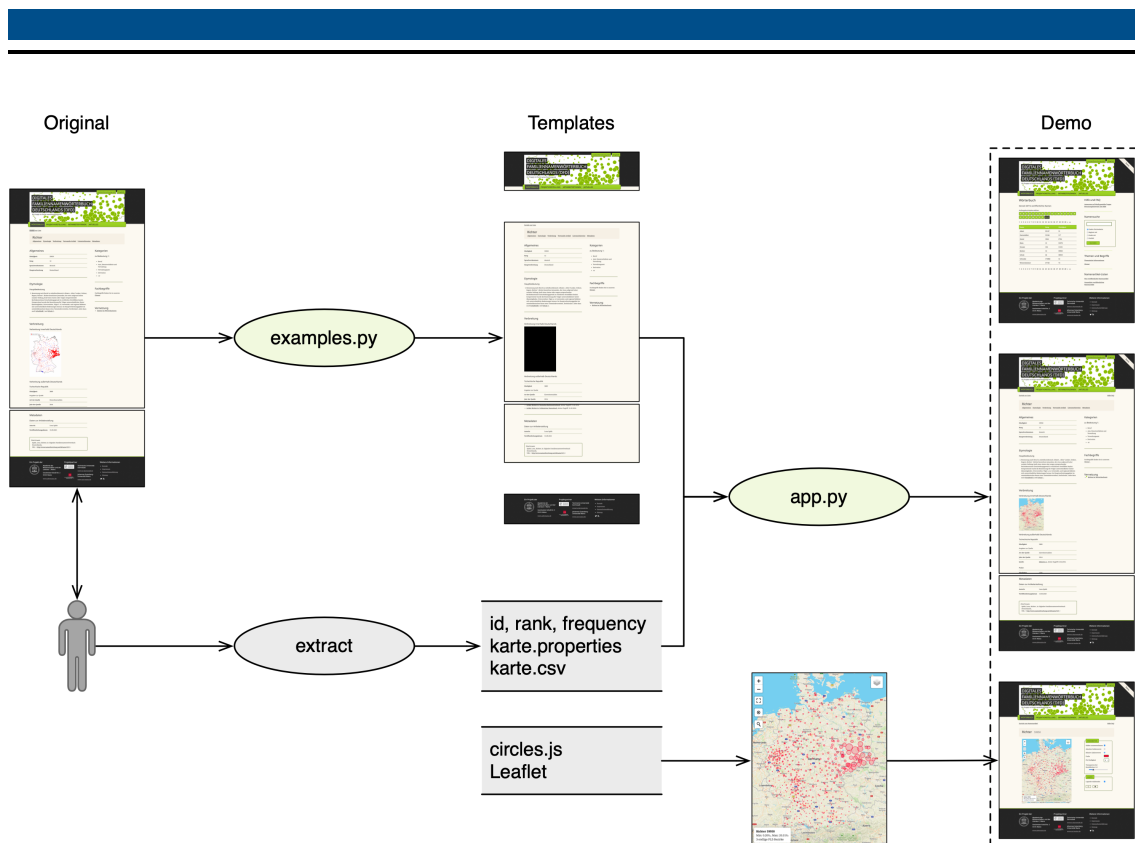


Abbildung 3.3: Datenfluss im Demonstrator für das interaktive Kartenmodul

3.3.1 Datenfluss im Demonstrator

Die Demo-Version des interaktiven Kartenmoduls für das DFD wird auf Basis des Python Webframeworks Flask gebaut (vgl. Dale 2016).⁷ Die Abbildung 3.3 zeigt den Datenfluss im Demonstrator mit der Vorbereitung der Ressourcen für die anschließende Echtzeit-Generierung der Anwendungsseiten.

Aus den original Namenartikeln der elf für den Demonstrator ausgewählten Namen wird in dem Python-Skript `examples.py` jeweils ein namenspezifisches Template generiert. Im Demonstrator werden aus den elf namenspezifischen Templates sowie den original Kopf- und Fußteilen des Wörterbuchs die Namenartikelseiten für die ausgewählten Namen zusammengesetzt. Das auf der Originalseite eingebundene Vorschaubild der Verbreitungskarte wird nicht übernommen. Aus dem Wörterbuch und dem DFD-Kartierungstool werden parallel hierzu die Namen-ID sowie der Rang und die Häufigkeit des Namens bzw. die für die Erzeugung der Karte (automatisch) vorgenommenen Parametereinstellungen und die Namensvorkommen in den Postleitzahlenbereichen händisch extrahiert (s. Dateien `karte.properties` und `karte.csv`). Aus den so gewonnenen Daten erzeugt das Python-Skript `app.py` die Anwendungsseiten des Demonstrators. In die Startseite wird die Namensliste bestehend aus den elf ausgesuchten Namen eingesetzt (s. Datei `index.html`). Für jeden gesuchten Namen wird der Namenartikel samt Vorschaukarte erstellt (s. z.B. Datei `173014.html` für

⁷<https://flask.palletsprojects.com/en/2.0.x/>

den Namen „Richter“). Mithilfe des aus der JavaScript-Bibliothek des DFD-Kartierungstools abgeleiteten Skripts `circles.js` und Leaflet wird u.a. in der Datei `karte.html` die interaktive Karte samt Interaktionsfeldern erzeugt und in die entsprechende Anwendungsseite eingesetzt.

3.3.2 Technische Umsetzung der Interaktivität

Die auf der Anwendungsseite Verbreitungskarte implementierte Interaktivität lässt sich mithilfe eines sogenannten Sequenzdiagramms darstellen.⁸ Bei der Anpassung der originalen JavaScript-Bibliothek des DFD-Kartierungstools wurden innerhalb des für das interaktive Kartenmodul abgeleiteten Skripts `circles.js` alle für die Darstellung der interaktiven Verbreitungskarten im DFD-Frontend relevanten Code-Bestandteile in einer Klasse `C` zusammengefasst. In den übernommenen Skriptanteilen wurden im Vergleich zum Originalcode lediglich marginale Änderungen vorgenommen, die die erforderliche Kompatibilität nicht gefährden. So wurde beispielsweise der `leaflet_scale_factor` von ursprünglich 100 auf den Wert 600 hochgesetzt, um die Größendarstellung der Kreise in den interaktiven Karten insgesamt anzupassen, und die Variable `msg`, mit deren Hilfe der Informationstext in den Popups über den Kreisen in den interaktiven Karten erzeugt wird, wurde ergänzt (s. Datei `circles.js`, Codezeilen 20 und 252).

Das Sequenzdiagramm in der Abbildung 3.4 zeigt das Verhalten des aus Nutzer*in, Applikation und `circles.js` bestehenden Systems. Innerhalb des Skripts `circles.js` enthält die Klasse `C` u.a. die Funktionen `addCircles` und `renderCircles`. Sobald durch die entsprechende Nutzeraktion die Anwendungsseite mit der interaktiven Verbreitungskarte aufgerufen wird, wird über den einmaligen Aufruf der Funktion `addCircles` die Anwendung (s. Datei `karte.html`) initialisiert. In der Funktion `addCircles` werden die Parametereinstellungen für die Defaultkarte, d.h. die Daten aus der Datei `karte.properties`, eingelesen und die in der Datei `karte.html` definierten Rückruf- oder Callback-Funktionen `circle`, `legend` und `info` werden hinterlegt. In der Applikation erzeugt die Funktion `circle` die Kreissymbole und legt sie in den Layer über der Karte. Die Funktion `legend` produziert den HTML-Inhalt der Legende. Die Funktion `info` stellt die Parametereinstellungen der Interaktionsfelder rechts neben der Karte ein, beim ersten Aufruf die Default- und nach jeder Benutzerinteraktion die aktualisierten Einstellungen. Die Funktion `addCircles` ruft ihrerseits zum ersten Mal die Funktion `renderCircles` auf, die wiederum durch den Aufruf der Callback-Funktionen die Daten für die Erzeugung der Kreise und den HTML-Inhalt der Legende erhält. In der ebenfalls in `circles.js` enthaltenen Funktion `getMap` werden alle Daten aus der Datei `karte.csv` und die von der Funktion `info` gelieferten Daten zu den Parametereinstellungen eingelesen und ausgewertet. Die Funktion `renderCircles` wird ausgeführt und die Darstellung der Default-Karte wird sichtbar. Jede Benutzerinteraktion bewirkt den Wiederaufruf der Funktion `renderCircles` mit den von den Callback-Funktionen gelieferten Daten für die Neuberechnung des Kreise-Layers und die Aktualisierung der Legende in der Karte. Die zur Ansicht bereitgestellte Verbreitungskarte wird entsprechend modifiziert. Die technische

⁸<https://www.uml.org>

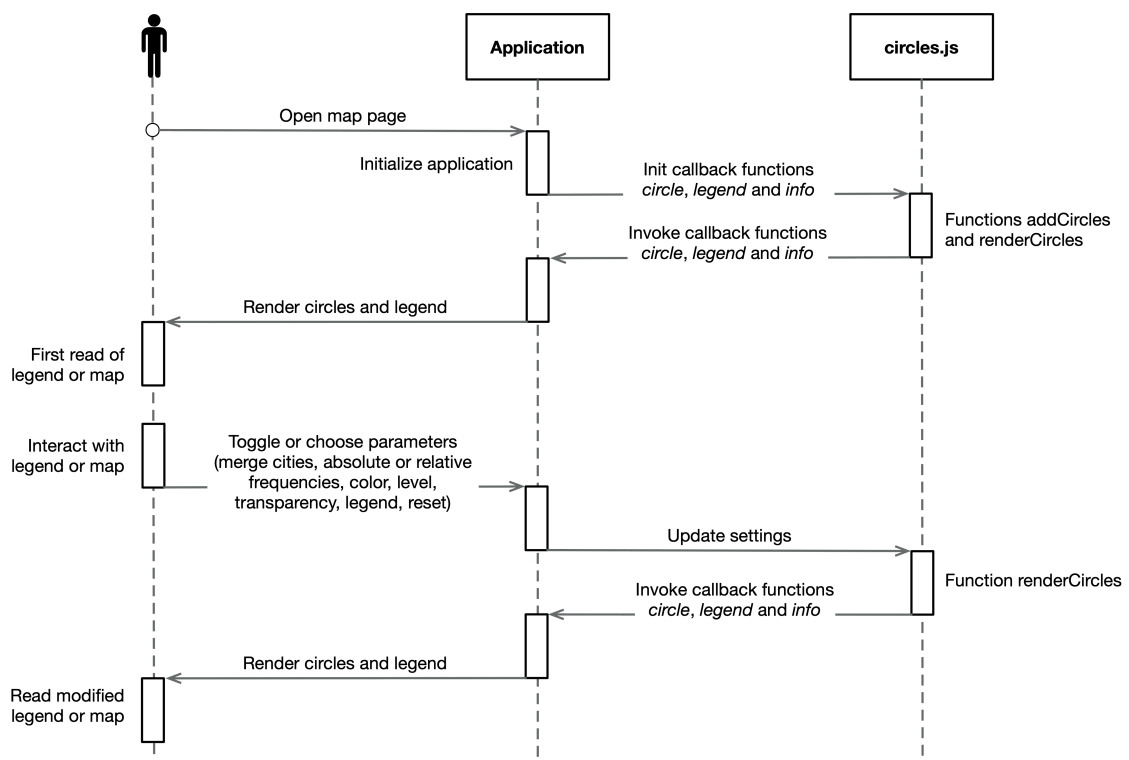


Abbildung 3.4: Interaktivität in der Verbreitungskarte des interaktiven Kartenmoduls

Umsetzung der Interaktivität in den Verbreitungskarten des interaktiven Kartenmoduls findet sich im Anhang B.

4 Visualisierung der Namensverbreitungen im interaktiven Kartenmodul

Für das Softwareprojekt der Entwicklung neuer interaktiver Namensverbreitungskarten für das DFD spielen das Design und die Qualität der Visualisierung der Namendaten in den Karten eine zentrale Rolle. Die Aufgabe der Visualisierung ist es, die in den Daten vorhandenen Informationen möglichst exakt und umfassend abzubilden. Einen wesentlichen Beitrag leistet hierbei die Interaktivität. Benutzerinteraktionen tragen zur effektiven und effizienten Exploration der der Visualisierung zugrunde liegenden Daten bei. Im besten Falle laden sie dazu ein sich eingehend mit den in der Visualisierung repräsentierten Daten zu beschäftigen.

4.1 Informationsvisualisierung

Die allgemeine Bedeutung des Begriffs Visualisierung ist die Veranschaulichung von abstrakten Sachverhalten durch eine optische oder bildliche Darstellung. Der Begriff der computer-gestützten Visualisierung bezeichnet die in der Regel interaktive graphische Repräsentation von Daten. Card et al. definieren die Visualisierung als „The use of computer-supported, interactive, visual representations of data to amplify cognition.“. Speziell für den Begriff der Informationsvisualisierung präzisieren sie diese Definition zu „The use of computer-supported, interactive, visual representations of abstract data to amplify cognition.“(vgl. Card, J. D. Mackinlay und Shneiderman 1999, S. 6, 7). In der Informationsvisualisierung werden digitale Daten auf visuelle Repräsentationen abgebildet, die wiederum von den Betrachter*innen interpretiert werden (vgl. Spence 2007, S. 5). Diese visuellen Repräsentationen sollen eine detaillierte Betrachtung der Daten und das Aufdecken innerer, verborgener Zusammenhänge oder unerwarteter Muster innerhalb der Datenmenge ermöglichen und so den Betrachter*innen bei der effektiven Auswertung einer Datenmenge helfen. (vgl. Schumann und Müller 2000, S. 5 und Munzner 2014, S. 1 und 7). Der Informatiker und Meerestechniker Colin Ware hat in diesem Zusammenhang den Begriff des „Visual Thinking“ geprägt, den er mit den folgenden Worten beschreibt: „[...] we should think about graphic designs as cognitive tools, enhancing and extending our brains. [...] Although we can, to some extent, form mental images in our heads, we do much better when those images are out in the world, on paper or computer screen. Diagrams, maps, web pages, information graphics, visual instructions, and technical illustrations all help us to solve problems through a process of visual thinking. [...] Visual thinking tools are especially important because they

harness the visual pattern finding part of the brain. Almost half the brain is devoted to the visual sense and the visual brain is exquisitely capable of interpreting graphical patterns, even scribbles, in many different ways. Often, to see a pattern is to understand the solution to a problem.“ (vgl. Ware 2008, S. ix).

Die Anfänge der Informationsvisualisierung liegen lange vor der Erfindung des Computers. Mithilfe von Stift und Papier sind Beispiele für die Visualisierung abstrakter statistischer und räumlicher Daten in zweidimensionalen Graphiken bzw. geographischen Karten entstanden. Sie alle haben gemeinsam, dass sie den Betrachter*innen durch die Visualisierung von quantitativen und geographischen Daten Informationen aus ebendiesen Daten zugänglich machen, die aus der reinen Auswertung von Tabellen und Zahlenreihen nicht oder nur schwer „ersichtlich“ wären. Ein Beispiel für eine solche Graphik ist die Cholera-Karte von Dr. John Snow (s. Abbildung 4.1). Wie Edward R. Tufte in seinem Buch „The Visual Display of Quantitative Information“ meint, wäre es mit viel Glück und Arbeit mit Berechnungen und Analysen auch ohne Graphik gelungen, den Zusammenhang zwischen der kontaminierten Wasserpumpe in der Broad Street und dem Cholera-Ausbruch von 1854 im Zentrum von London festzustellen (vgl. Tufte 1983, S. 24). Mit John Snows Graphik wird dieser Zusammenhang dagegen schnell und unmittelbar „sichtbar“.

Ein bedeutendes Beispiel für die Visualisierung komplexer Daten in einer Karte ist Charles Joseph Minards Karte zu Napoleons Russlandfeldzug aus dem Jahr 1861 (s. Abbildung 4.2). In seiner Karte gelingt Minard die Darstellung von insgesamt sechs verschiedenen Variablen: der Größe der Armee, der Position der Armee auf einer zweidimensionalen Oberfläche, der Bewegungsrichtung der Armee und der Temperaturen zu verschiedenen Zeitpunkten während des Rückzugs von Moskau (vgl. Tufte 1983, S. 40). Während Tufte Minards Karte als die möglicherweise beste statistische Graphik bezeichnet, die je gezeichnet wurde, überträgt Kyran Dale die Problematik, die Minard auf analoge Art angegangen ist, auf die digitale Informationsvisualisierung: „[...] the chief problem of modern interactive data-viz is exactly that faced by Minard: how to move beyond conventional one-dimensional bar-charts etc. (perfectly good for many things) and develop new ways to communicate cross-dimensional patterns effectively.“ (vgl. Dale 2016, S. xxxi).

Erst mit dem Zugang zu Computern und der rasanten Entwicklung von Hard- und Softwaretechnologien in den letzten Jahrzehnten hat das Forschungsfeld der computerbasierten Visualisierung von Informationen aus großen Datenmengen stark an Bedeutung gewonnen und in viele verschiedene Forschungsbereiche Einzug gehalten. Die Anfänge der Forschung zur Informationsvisualisierung gehen auf den französischen Kartographen Jaques Bertin, der als Pionier im Bereich der - damals noch analogen - Informationsvisualisierung gilt, zurück. In seinem grundlegenden Buch „Sémiologie graphique“, das bereits im Jahr 1967 veröffentlicht wurde, identifiziert Bertin vier gemeinsame Aufgaben für die Informationsvisualisierung und entwickelt passende Kodiermethoden mit sogenannten retinalen Variablen und geeigneten Signaturen (vgl. Abschnitt 4.3).¹ Bertins Theorien wurden u.a. weiterentwickelt von den Statistikern William S. Cleveland und Robert McGill, die sich in der Hauptsache mit der Kodierung quantitativer Daten beschäftigt haben und von Jock D. Mackinlay, der

¹Jaques Bertin. *Sémiologie Graphique: Les diagrammes—Les réseaux—Les cartes*. Gauthier-Villard, 1967



Abbildung 4.1: Cholera Karte mit Punktsymbolen von Dr. John Snow, 1854. Quelle: <https://upload.wikimedia.org/wikipedia/commons/c/c7/Snow-cholera-map.jpg>

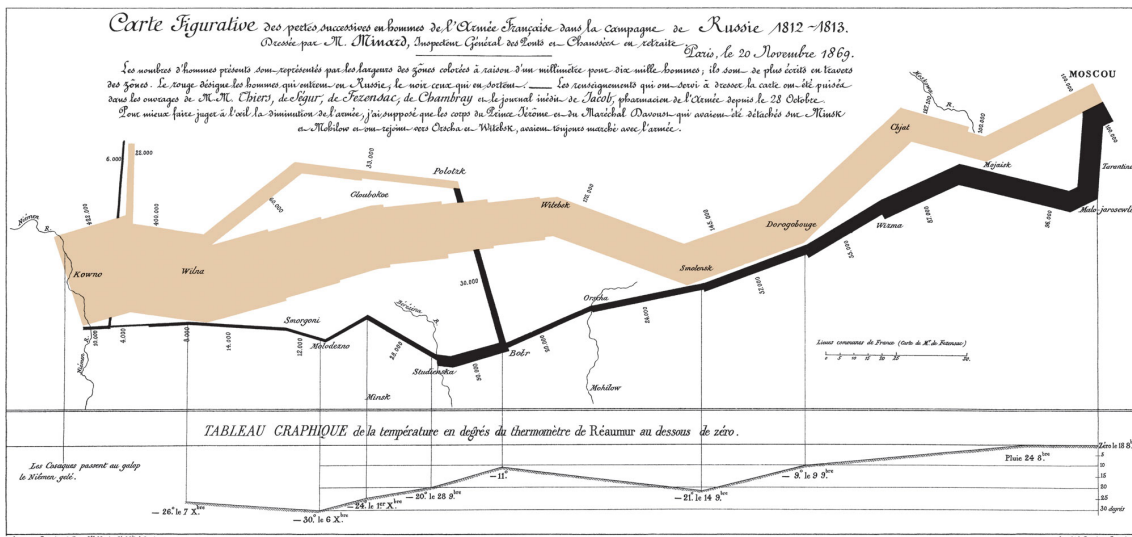


Abbildung 4.2: Charles Joseph Minards Karte zu Napoleons Russlandfeldzug, 1861. Quelle: <https://upload.wikimedia.org/wikipedia/commons/2/29/Minard.png>

als Computergraphiker und aus dem Blickwinkel der künstlichen Intelligenz Bertins Designtheorie formalisiert und zusammen mit Stuart K. Card und George G. Robertson weitere Techniken der Informationsvisualisierung entwickelt hat (vgl. Cleveland und McGill 1984 und J. Mackinlay 1986).

Mit den technischen Fortschritten im Bereich der Graphikhardware entstand im Laufe der Zeit eine neue Generation von Benutzungsschnittstellen mit Benutzerinteraktionen und großen Informationsmengen. Mittlerweile wird die Forschungslandschaft rund um die Informationsvisualisierung getragen von einer großen lebendigen Community, die in den verschiedensten Forschungsgebieten unterwegs ist. Es existieren große Mengen von spezialisierten Entwicklungen, Techniken und Standards, die oftmals nicht unmittelbar von dem einen auf den anderen Anwendungsbereich übertragbar sind. Jedes Forscherteam entwickelt seine eigene Sicht- und Herangehensweise. Das bedeutet, dass ein einheitlicher Stand der Forschung im Bereich der Informationsvisualisierung schwer zu beschreiben ist. Ein gewisses Zentrum rund um die Visualisierungsforschung stellt eine Gruppe von Wissenschaftler*innen rund um Ben Shneiderman, Colin Ware, Tamara Munzner u.a.m. dar, die versuchen, die technischen Möglichkeiten der Informationsvisualisierung u.a. mit Aspekten wie der Human-Computer-Interaction, den Prozessen der visuellen Perzeption im menschlichen Sehapparat und den Informationsbedürfnissen der Nutzer*innen in Einklang zu bringen.

R.B. Haber, D.A. McNabb und Tamara Munzner sprechen bei jedem individuellen Lösungsansatz zum Entwurf und zur Manipulation von visuellen Repräsentationen wissenschaftlicher Datenmengen von sogenannten Idiomen der Visualisierung. Haber und McNabb definieren ein Visualisierungsidiom als „a specific sequence of data enrichment and enhancement transformations, visualization mappings and rendering transformations that produce an

abstract display of a scientific data set“ (vgl. Haber und McNabb 1990). Für jede Visualisierungsaufgabe sind die Möglichkeiten, eine Datenmenge in einem einzigen Bild visuell zu kodieren, vielfältig und werden zudem noch erweitert, wenn dieses oder mehrere solcher Bilder durch Benutzerinteraktionen manipulierbar sind. Die in der Visualisierungsforschung formulierten gemeinsamen Grundsätze und Überlegungen zur Entwicklung einer qualitativ hochwertigen Visualisierung betreffen die konkreten Fragen und Anforderungen, die die Nutzer*innen an eine Visualisierung stellen, die Art und Struktur der Daten, die die Nutzer*innen sehen, und schließlich die Art und Weise, wie die visuelle Kodierung der Daten und die Benutzerinteraktionen in dem zugehörigen Visualisierungsideom umgesetzt werden (vgl. Munzner 2014, S. 17). Einer der Grundsätze, die für jede Visualisierung einer Datenmenge gelten, ist das Prinzip der Expressivität eines Visualisierungsideoms (vgl. J. Mackinlay 1986). Die Expressivität oder Ausdrucksfähigkeit einer Visualisierung betrifft die korrekte, vollständige und wahrheitsgetreue Darstellung der Datenmenge, die möglichst alle relevanten Aspekte der Daten umfasst. Das stellt eine Herausforderung für die Entwickler*in dar, da jede Abbildung einer Datenmenge in einem Idiom eine Abstraktion darstellt, bei der z.B. Entscheidungen darüber getroffen werden müssen, welche Aspekte hervorgehoben werden sollen und welche nicht. Es kann folglich für eine abzubildende Datenmenge durchaus mehrere Idiome geben, die das Kriterium der Expressivität erfüllen (vgl. Schumann und Müller 2000, S. 9 und Munzner 2014, S. 13). Um unter verschiedenen Visualisierungsideomen die Darstellung der Datenmenge, die den Bedürfnissen der Nutzer*innen, der Zielsetzung der Visualisierung und dem jeweiligen Anwendungskontext am besten gerecht wird, herauszusuchen, nennt Mackinlay als ein weiteres Kriterium das Prinzip der Effektivität der Visualisierung (vgl. J. Mackinlay 1986). Die Effektivität eines Visualisierungsideoms bemisst sich an der Beantwortung von Fragen wie den folgenden:

- Welche Ziele bzw. Aufgaben verfolgen die Nutzer*innen bei der Betrachtung bzw. Beschäftigung mit der Visualisierung und werden diese Ziele unterstützt? Nutzer*innen können z.B. auf der Suche nach neuen Erkenntnissen über die Datenmenge sein, sie können die Visualisierung heranziehen, um eine eigene Hypothese über die Daten zu verifizieren oder zu widerlegen, oder sie möchten vielleicht lediglich Informationen über die Datenmenge erhalten und die Visualisierung als solche genießen (vgl. Munzner 2014, Kap. 3).
- Werden die kognitiven Fähigkeiten und die visuellen Wahrnehmungsprozesse des Menschen generell und die der individuellen Nutzer*innen im Besonderen berücksichtigt? Nicht nur die Speicher- und Graphikkapazitäten des Computers sondern auch körperliche Einschränkungen wie spezielle Sehschwächen sowie die Grenzen des menschlichen Gedächtnisses, die kognitive Belastung und die aufgewendete Zeit stellen limitierende Faktoren dar (vgl. Munzner 2014, S. 15, Ware 2013 und Schumann und Müller 2000, S. 9).
- Soll das Visualisierungsideom zuvor festgelegte Informationen und Interpretationen über eine Datenmenge lediglich präsentieren oder soll den Nutzer*innen angeboten werden, Datenmanipulationen vorzunehmen? Eine einzelne statische Ansicht kann nur lediglich einen Aspekt einer Datenmenge zeigen. Eine Ansicht, die sich interaktiv verändern lässt, erlaubt den Nutzer*innen, eigene Anfragen zu stellen und neue

Ansichten zu erzeugen, die wiederum andere Aspekte der Datenmenge beleuchten können. Die angebotenen Möglichkeiten zur Datenmanipulation dürfen die Nutzer*innen jedoch nicht überfordern. Die Folge wäre eine ineffektive Visualisierung.

In der Forschungsliteratur führen die Überlegungen zu den Prinzipien der Expressivität und der Effektivität und damit der Benutzerfreundlichkeit einer Visualisierung zu verschiedenen Problemlösungsansätzen mit verschiedenen Schwerpunkten. Wie oben bereits erwähnt, verfolgt Ben Shneiderman einen Ansatz aus der Perspektive der Forschungen zur Human-Computer-Interaction und entwickelt das von ihm so genannte Visual Information Seeking Mantra: „Overview first: Gain an overview of the entire collection“, „Zoom: Zoom in on items of interest“, „Filter: filter out uninteresting items“ und „Details-on-Demand: select an item or group and get details when needed“. (vgl. Shneiderman 1996 und Shneiderman 1998). Colin Ware konzentriert sich auf die Eigenschaften und Besonderheiten des menschlichen Sehapparats und die visuelle Perzeption von Formen, Farben etc. und entwickelt seine Visualisierungen entlang der kognitiven und visuellen Wahrnehmungsverarbeitung der Nutzer*innen (vgl. Ware 2013). Der Geograph Alan MacEachren verfolgt einen ähnlichen kognitiv-semiotischen Ansatz in Bezug auf die mentale Perzeption und Verarbeitung von Informationen in geographischen Karten (vgl. MacEachren 1995). Rae A. Earshaw gibt als Ergebnis der Forschungen in der Informationsvisualisierung präzise formulierte Empfehlungen für die Entwicklung von Visualisierungen und Interaktivität (vgl. Earshaw 2005). Die Visualisierungsforscherin Tamara Munzner stellt sich in ihrem Buch „Visualization Analysis & Design“ der Aufgabe, die wesentlichen Grundlagen der Informationsvisualisierung aus den eigenen Forschungsergebnissen und den Erkenntnissen und Entwicklungen ihrer Forschungskollegen von den Anfängen bis zum Zeitpunkt ihrer Arbeit zusammenzutragen. Sie bindet die konsolidierten Forschungsergebnisse aus den unterschiedlichen Wissenschaftsgebieten in einem einzigen Framework zusammen (vgl. Munzner 2014). Dieses Framework schließt die Interaktivität und auch Aspekte der Geovisualisierung und die von Munzner als „the integration of non-spatial data with base spatial data“ definierte thematische Kartographie mit ein (vgl. Munzner 2014, S. 180).

4.2 Visualisierungspipeline

Für jede Entwicklung eines Visualisierungsideoms wird die sogenannte Visualisierungspipeline festgelegt. Während eines Visualisierungsprozesses werden mehrere Entwicklungsschritte durchlaufen, um abstrakte, nicht geometrische Daten in Form von Bildern zu veranschaulichen. Die drei wesentlichen, in der Visualisierungspipeline aufeinander folgenden Schritte sind die Datenaufbereitung mit der Entwicklung des Datenmodells für die Visualisierung, das Mapping auf visuelle Repräsentationen für die Daten und schließlich das Rendering der im Mapping erzeugten Daten (vgl. Schumann und Müller 2000, S. 15-21).

Bei der Datenaufbereitung handelt es sich grundsätzlich um eine Daten-zu-Daten-Abbildung von der Menge der Rohdaten auf eine vervollständigte oder auch reduzierte Menge von nach bestimmten Kriterien gefilterten Daten. Für den Visualisierungsprozess im interaktiven Kartenmodul werden die Rohdaten der Deutschen Telekom in ihrer im DFD-Kartierungstool

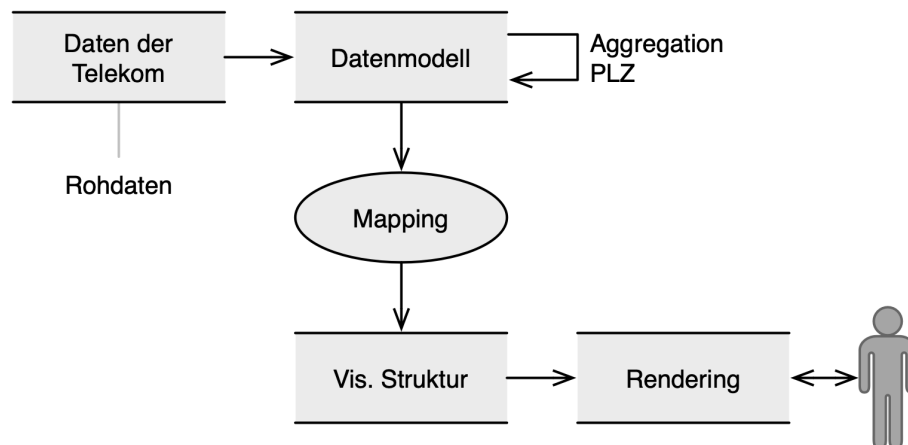


Abbildung 4.3: Datenfluss in der Visualisierungspipeline

aufbereiteten Form übernommen und auf das im Abschnitt 3.2 beschriebene Datenmodell abgebildet. Innerhalb dieses Datenmodells findet eine zusätzliche Aggregation der PLZ-Daten statt. Die entsprechenden Daten enthalten die zusammengefassten Informationen zu den Häufigkeiten eines Namens in ein-, zwei-, drei- bzw. fünfstelligen Postleitzahlenbereichen. Die aufbereiteten Daten sind der Ausgangspunkt für den nächsten Schritt des Visualisierungsprozesses, das Mapping. Beim Mapping werden die nicht-geometrischen Daten des DFD auf geometrische Daten wie Punkte, Linien oder Flächen mit zugehörigen visuellen Attributen wie Farbe, Anordnung, Größe etc. abgebildet. Das Mapping hat damit einen entscheidenden Einfluss auf die spätere visuelle Repräsentation der Daten. Es bildet den Übergang vom Datenmodell zur Visualisierungsstruktur des Kartenmoduls. Die eigentliche Bildgenerierung, das Rendering, ist der letzte Schritt in der Visualisierungspipeline. Hier werden die im Mapping erzeugten Geometriedaten auf digitale Bilddaten abgebildet. Diese Bilddaten werden den Nutzer*innen schließlich in den interaktiven Namensverbreitungskarten präsentiert. Die Abbildung 4.3 zeigt den Datenfluss in der Visualisierungspipeline.

Nachdem im Abschnitt 3.2 das Datenmodell für das interaktive Kartenmodul entwickelt wurde, wird in den folgenden Abschnitten der in der Visualisierungspipeline zentrale Schritt des Mappings für das Visualisierungsidiom der interaktiven Verbreitungskarten umfassend beschrieben und die einzelnen Mappingschritte werden jeweils im Hinblick auf die Prinzipien der Expressivität und Effektivität der Visualisierung ausführlich erläutert. Weitere Aspekte der Visualisierungen wie das Design von visuellen Feedback-Mechanismen, die Umsetzung des Prinzips Overview-first-Detail-on-Demand, die Frage der kognitiven Belastung der Betrachter*innen oder die Antwortzeiten für die visuelle Suche etc. werden an einigen Stellen angesprochen, aber nicht systematisch vertieft. Die für die Forschung im Bereich der Informationsvisualisierung grundlegenden Techniken und Theorien hierzu können in den Arbeiten von Card et al. und Ben Shneiderman für den Bereich der Human-Computer-Interaction und in denen von Colin Ware für die Forschung zum Zusammenhang zwischen den visuellen Wahrnehmungsprozessen beim Menschen und Visualisierungen nachgelesen werden (vgl. Card, J. D. Mackinlay und Shneiderman 1999, Shneiderman 1996 und Shneiderman 1998 sowie Ware 2008 und Ware 2013).

Name	Art	Bestand	Anteil (ger.)	lfd. Nummer
Buche	Laubbaum	1254	32,6	1
Kiefer	Nadelbaum	189	4,9	2
Ulme	Laubbaum	37	1,0	3
Fichte	Nadelbaum	2367	61,5	4
<i>Kategorial</i>	<i>Kategorial</i>	<i>Quantitativ</i>	<i>Quantitativ</i>	<i>Ordinal</i>

Tabelle 4.1: Beispiel für den Datenmengentyp Tabelle

4.3 Mapping-Modell für das Design von Visualisierungsidiomen

In jedem Visualisierungsprozess orientiert sich das Mapping an der sorgfältigen Analyse der Fragen und Anforderungen der zukünftigen Betrachter*innen an die Visualisierung (vgl. Abschnitt 2.5.2). Welche konkreten Möglichkeiten der Abbildung von Daten auf visuelle Repräsentationen bei der Entwicklung eines Visualisierungsidioms generell gegeben sind, hängt dabei von der Art der zugrunde liegenden *Datenmengen* ab. Die Datenmengen können in mehreren unterschiedlichen *Datenmengentypen* vorliegen. Der mit Abstand am häufigsten für Visualisierungen jeder Art vorliegende Datenmengentyp ist die *Tabelle*. Eine Tabelle besteht, wie in dem einfachen Beispiel der Tabelle 4.1 zu sehen ist, aus Spalten und Zeilen. Die Zeilen sind vom *Datentyp Datensatz* und bestehen jeweils aus einem Tupel der in den Tabellenzellen einer Zeile eingetragenen Werte. Die Spalten, in der Tabelle oben mit „Name“, „Art“, „Bestand“ usw. bezeichnet, sind vom *Datentyp Attribut*. Sie enthalten diejenigen Datenwerte, die beim späteren Mapping auf visuelle Repräsentationen abgebildet werden. Enthält ein Attribut Datenwerte, die geordnet sind bzw. die sich in eine Ordnung bringen lassen, so ist das Attribut entweder vom *Attributtyp Ordinal* und enthält Daten, die sequentiell von einem Minimum- zu einem Maximum-Wert geordnet sind, oder das Attribut ist vom *Attributtyp Quantitativ*. In diesem Attributtyp liegen die Daten in einer nicht sequentiellen Ordnung vor. Ein Attribut ist vom *Attributtyp Kategorial*, wenn die Einträge verschiedene, untereinander abgrenzbare Identitäten bezeichnen. Die Datenmenge, die einer geographischen Karte zugrundeliegt, ist vom Datenmengentyp *Geometrie*. Für diesen Datenmengentyp einzigartig ist die exakte Spezifikation von geographischen Orten. Diese spezifische Information ist vom *Datentyp Position*.

Für die konkrete Abbildung von Attributdaten auf visuelle Repräsentationen stehen drei *primitive graphische Elemente* zur Verfügung. Diese sind der nulldimensionale *Punkt*, die eindimensionale *Linie* und die zweidimensionale *Fläche*. Das Erscheinungsbild der primitiven graphischen Elemente in einer Visualisierung ist variabel und kann durch sogenannte *visuelle Attribute* wie z.B. Farbe oder Form gestaltet werden. Die oben beschriebenen Attributtypen Quantitativ, Kategorial etc. werden im Mapping-Schritt auf solche durch geeignete visuelle Attribute gestaltete graphische Elemente abgebildet. Um die Bezeichnungen der oben beschriebenen Datenmengen- und Attributtypen zu vereinfachen und Verwechslungen der Begriffe Attribut und visuelles Attribut im weiteren Text zu vermeiden, wird im Folgenden vereinfacht von *Tabellen*, von *Spalten* anstelle von Attributen, von *visuelle Attributen* und

von *ordinalen, quantitativen, kategorialen* und *geometrischen Daten* gesprochen.

Das Ziel des Mappings ist eine visuelle Repräsentation der in den Attributen enthaltenen Daten, die die Anforderungen der Expressivität mit Blick auf das Datenmaterial und der Effektivität mit Blick auf die Nutzer*innen erfüllt. Die Qualität der Expressivität der Visualisierung hängt davon ab, wie gut die Auswahl der visuellen Attribute für das Mapping zu den Charakteristika der Daten passt. Für die Abbildung kategorialer Daten sind das solche visuellen Attribute, die sich zur Identifizierung unterschiedlicher Identitäten eignen wie z.B. die Lage im Raum, die geometrische Form oder verschiedene Farbtöne. Für das Mapping quantitativer Daten wird ein graphisches Element durch visuelle Attribute gestaltet, die im weitesten Sinne nach Größen geordnet werden können. Das sind Attribute wie Länge, Flächengröße oder verschiedene Abstufungen in Farbsättigung und Farbhelligkeit. Das Maß der Effektivität der Visualisierung hängt wiederum ab von der Fähigkeit der graphischen Elemente und ihrer visuellen Attribute, die in den Daten vorhandenen Informationen so darzustellen, dass die Betrachter*innen der Visualisierung Antworten auf möglichst viele der Fragen erhalten, die sie an die Repräsentationen der sie interessierenden Datenmengen stellen.

Welche visuellen Attribute für die primitiven graphischen Elemente beim Mapping von Datenwerten auf visuelle Repräsentationen im Einzelnen zur Verfügung stehen bzw. nach welchen Kriterien diese visuellen Attribute im Hinblick auf die Expressivität und die Effektivität einer Visualisierung ausgewählt werden, ist Gegenstand der Forschung im Bereich der Informationsvisualisierung. In seiner Theorie der graphischen Semiotik unterscheidet Jaques Bertin zwischen sieben retinalen Variablen, d.h., übertragen auf die oben dargelegten Zusammenhänge sieben verschiedenen visuellen Attributen, für graphische Elemente: Position, Größe, (Farb-)Helligkeit, Textur, Farbton, Orientierung und Form (vgl. Spence 2007, S. 52-54). Diese visuellen Attribute unterteilt Bertin in selektive, ordinale und proportionale Attribute, die er folgendermaßen beschreibt. Mit selektiven Attributen wie der Farbe eines graphischen Elements kodierte Daten werden von den Betrachter*innen der Visualisierung spontan als unterschiedlich wahrgenommen und in Gruppen aufgeteilt. Mit ordinalen Attributen wie der (Farb-)Helligkeit eines graphischen Elements kodierte Daten werden spontan in eine Ordnung gebracht. Daten, die mithilfe von proportionalen Attributen wie der Größe eines graphischen Elements kodiert werden, werden als proportional zueinander wahrgenommen. Proportionale Attribute eignen sich laut Bertin besonders für die Kodierung quantitativer Daten. Bertin definiert außerdem vier Aufgaben, die an eine Visualisierung gestellt werden können. Er nennt diese Aufgaben Assoziation, Selektion, Ordnung und Quantität. Seine Liste der Attribute ordnet Bertin grob nach der Anzahl der Aufgaben, die die einzelnen visuellen Attribute jeweils übernehmen können:

Größe eines graphischen Elements Graphische Elemente können

- a) als ähnlich wahrgenommen und gruppiert werden (Assoziation),
- b) als geordnet wahrgenommen werden (Ordnung),
- c) als proportional zueinander wahrgenommen werden (Quantität).

Textur eines graphischen Elements Graphische Elemente können

-
- a) als ähnlich wahrgenommen und gruppiert werden (Assoziation),
 - b) als verschieden wahrgenommen werden (Selektion),
 - c) als geordnet wahrgenommen werden (Ordnung).

Helligkeitswert eines graphischen Elements Graphische Elemente können

- a) als verschieden wahrgenommen werden (Selektion),
- b) als geordnet wahrgenommen werden (Ordnung).

Farbe eines graphischen Elements Graphische Elemente können

- a) als ähnlich wahrgenommen und gruppiert werden (Assoziation)
- b) als verschieden wahrgenommen werden (Selektion).

Orientierung eines graphischen Elements Graphische Elemente können aufgrund der Abweichung von der Horizontalen bzw. Vertikalen

- a) als ähnlich wahrgenommen und gruppiert werden (Assoziation).

Geometrische Form eines graphischen Elements Graphische Elemente können als

- a) ähnlich wahrgenommen und gruppiert werden (Assoziation).

Bertins Theorie der graphischen Zeichen ist grundlegend für die Forschung im Bereich der Informationsvisualisierung. Sie wurde, wie oben bereits erwähnt, u.a. von William S. Cleveland und Robert McGill sowie Jock Mackinlay weiterentwickelt. Cleveland und McGill haben psychologische Untersuchungen zur Effektivität einzelner visueller Attribute speziell zur Visualisierung quantitativer Daten durchgeführt, deren Ergebnisse u.a. in die unten aufgeführte von Mackinlay vorgeschlagene Rangordnung von 1. hocheffektiv bis 8. wenig effektiv für insgesamt acht verschiedene visuelle Attribute eingeflossen sind.

1. Lage (in einer Skala bzw. auf einer Ebene angegeben durch x- und y-Koordinaten)
2. Länge
3. Winkel
4. Neigung
5. Fläche
6. Volumen
7. Dichte
8. geometrische Form

Auch für kategoriale Daten stellt Mackinlay eine nach Effektivität geordnete Rangliste von insgesamt 13 verschiedenen visuellen Attributen auf. Der Großteil dieser Attribute ist identisch mit den Attributen, die Mackinlay auch für die Kodierung quantitativer Daten vorschlägt. Tamara Munzner hält dagegen die Verwendung von für quantitative Daten geeigneten visuellen Attributen für die Abbildung von kategorialen Daten für einen Verstoß gegen das Prinzip der Expressivität einer Visualisierung. Sie schlägt stattdessen zehn verschiedene visuelle Attribute für das Mapping von quantitativen und ordinalen Daten und lediglich vier visuelle Attribute für die Abbildung kategorialer Daten vor. Diese beiden Mengen visueller Attribute haben keine direkten Überschneidungen. In ihrem Framework bezeichnet Munzner diejenigen visuellen Attribute, die für das Mapping von Daten, die sich in eine Ordnung bringen lassen, geeignet sind, als „magnitude channels“ und diejenigen visuellen Attribute, die für die Abbildung kategorialer Daten verwendet werden können, als „identity channels“. Sie schreibt: „The expressiveness principle dictates that the visual encoding should express all of, and only, the information in the dataset attributes. The most fundamental expression of this principle is that ordered data should be shown in a way that our perceptual system intrinsically senses as ordered. Conversely, unordered data should not be shown in a way that perceptually implies an ordering that does not exist.“ (vgl. Munzner 2014, S. 100). „While it is possible in theory to use a magnitude channel for categorical data or a identity channel for ordered data, that choice would be a poor one because the expressiveness principle would be violated.“ (vgl. Munzner 2014, S. 101). Wie Mackinlay ordnet auch Munzner ihre visuellen Attribute nach dem Rang ihrer Effektivität. Das Prinzip der Effektivität definiert sie dabei folgendermaßen: „The effectiveness principle dictates that the importance of the attribute should match the salience of the channel; that is, its noticeability. In other words, the most important attributes should be encoded with the most effective channels in order to be most noticeable, and then decreasingly important attributes can be matched with less effective channels.“ (vgl. Munzner 2014, S. 102). Die Ergebnisse ihrer Forschungen fasst Munzner in einem Mapping-Modell zusammen, das in der Abbildung 4.4 zu sehen ist.

4.4 Mapping-Modell für die Verbreitungskarten im DFD

Die Tabelle 4.2 zeigt ein in Anlehnung an das Mapping-Modell von Tamara Munzner zusammengestelltes vereinfachtes Mapping-Modell für das interaktive Kartenmodul. Die Anzahl der Einträge ist reduziert und beschränkt sich auf die für das Mapping im Kartenmodul relevanten visuellen Attribute. Die Zuordnung der visuellen Attribute zu den quantitativen und den ordinalen Daten in der linken Spalte der Tabelle sowie die Zuordnung der visuellen Attribute zu den kategorialen Daten in der rechten Spalte der Tabelle entsprechen jeweils dem Prinzip der Expressivität. Die Attribute sind nach Effektivitätsrang von oben nach unten geordnet, wobei die an jeweils oberster Stelle aufgeführten Attribute den höchsten Effektivitätswert besitzen.

Channels: Expressiveness Types and Effectiveness Ranks

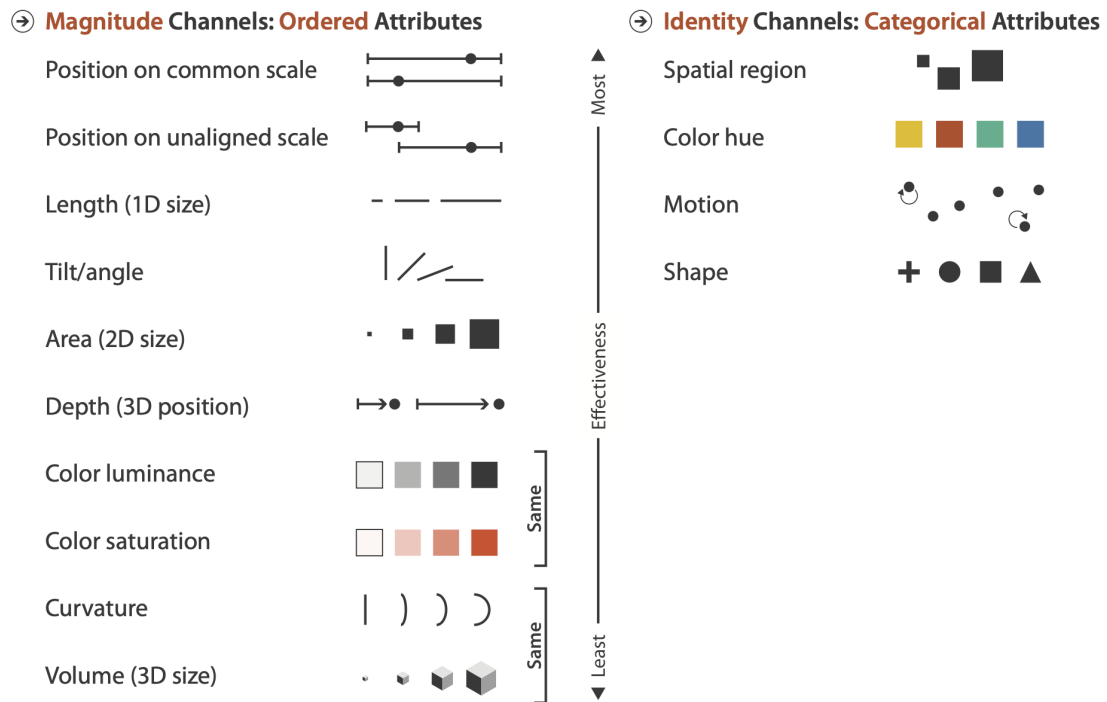


Abbildung 4.4: Mapping-Modell von Tamara Munzner. Quelle: Munzner 2014, S. 102, CC BY 4.0

Quantitative, Ordinale Daten	Kategoriale Daten
<i>Geordnete visuelle Attribute</i>	<i>Kategoriale visuelle Attribute</i>
Position auf einer nicht ausgerichteten Skala	Position im Raum
Flächengröße	Farbton
Farbhelligkeit, Farbsättigung	Geometrische Form
Volumen	

Tabelle 4.2: Mapping-Modell für die interaktiven Verbreitungskarten im DFD

4.5 Visualisierungsideiom des interaktiven Kartenmoduls

Die visuelle Kodierung der dem interaktiven Kartenmodul zugrunde liegenden Daten geschieht auf der Grundlage des in der Tabelle 4.2 festgelegten Mapping-Modells. Bei der Auswahl der visuellen Attribute für die Abbildungen der verschiedenartigen Daten steht als Maß für die Benutzerfreundlichkeit und damit die Qualität der Visualisierung die Frage im Vordergrund, inwieweit die jeweils vorgenommene Abbildung den Kriterien der Expressivität und der Effektivität der Visualisierung entsprechen. Obwohl die Qualität einer Visualisierung sich maßgeblich an diesen Prinzipien bemisst, spielen darüber hinaus auch weitere Faktoren eine Rolle. So kann die individuelle Wahrnehmung der Qualität der Visualisierung zusätzlich von dem Vorwissen oder den Vorlieben der Betrachter*innen, den speziellen Bearbeitungszielen, die die Betrachter*innen verfolgen, bis hin zu den körperlichen und technischen Voraussetzungen, unter denen sich die Betrachter*innen mit der Visualisierung beschäftigen, abhängen (vgl. Schumann und Müller 2000, S. 8-9). Zum Teil spiegeln sich diese Kriterien auch in den Fragen und Anforderungen der Nutzer*innen wider, die in dem Fragen- und Aufgabenkatalog im Abschnitt 2.5.2 und in dem erweiterten Aufgabenkatalog im Abschnitt 6.1 aufgelistet sind.

4.5.1 Art der zugrunde liegenden Daten

Wie aus dem Datenmodell für das interaktive Kartenmodul ersichtlich ist, liegen alle Datenmengen, die in den interaktiven Namensverbreitungskarten auf visuelle Repräsentationen abgebildet werden, in Form von Tabellen vor, die die aus der Datenbasis des DFD-Kartierungstools abgeleiteten Daten repräsentieren (s. Abschnitt 3.2). Es handelt sich hierbei um *statische* Datenmengen, d.h. alle benötigten Daten stehen für die Visualisierung gleichzeitig zur Verfügung, werden also nicht etwa erst während der Arbeit mit der Visualisierung dynamisch erzeugt. Um die für das Mapping der Daten auf passende graphische Elemente und die nach den Prinzipien der Expressivität und der Effektivität am besten geeigneten visuellen Attribute zur Gestaltung dieser graphischen Elemente zu finden, wird zunächst die Art der vorliegenden Daten genauer betrachtet (vgl. Munzner 2014 und Schumann und Müller 2000).

Jeder Datensatz in einer der Tabellen in den zugrunde liegenden plz-Dateien `plz_1.csv`, `plz_2.csv`, `plz_3.csv` und `plz_5.csv` bezeichnet einen individuellen Postleitzahlenbereich (s. Listings 3.1 und 3.2). Die sechs verschiedenen Spalten der Tabellen, „key“, „name“, „lon“, „lat“, „unmerged“ und „merged“, enthalten jeweils die zur Beschreibung eines Postleitzahlenbereichs benötigten Datenwerte.

Wie im Entity-Relation-Diagramm des interaktiven Kartenmoduls in der Abbildung 3.2 abzulesen ist, sind der Entität „PLZ“ die Attribute „Latitude“ und „Longitude“ zugeordnet. Dementsprechend sind die Tabellenspalten „lon“ und „lat“ relevant für das Mapping auf graphische Elemente und visuelle Attribute. Sie enthalten jeweils Zahlenwerte für Breiten- bzw. Längengrade.² Ein aus einem Breiten- und einem Längengrad zusammengesetztes

²Vgl. hierzu auch die Ausführungen im Kapitel 7

Wertepaar stellt im mathematischen Sinne die x- und y-Werte in einem erdumspannenden Koordinatensystem dar und legt genau einen speziellen Punkt auf der Erde fest. In jedem Datensatz in den Tabellen der plz-Dateien geben die Datenwerte in den Spalten „lon“ und „lat“ die Geokoordinaten des Mittelpunkts eines individuellen Postleitzahlenbereichs an. Bei den entsprechenden Daten handelt es sich um nicht sequentiell geordnete mathematische Größen und damit um *quantitative Daten*.

Wie ebenfalls im Abschnitt 3.2 bereits erläutert wurde, enthält jeder Datensatz in der Tabelle der Datei `karte.csv` die vollständigen Informationen erstens zum zahlenmäßigen Vorkommen des jeweils ausgewählten Familiennamens und zweitens zum Vorkommen aller Familiennamen insgesamt in fünfstelligen Postleitzahlenbereichen (s. Listing 3.3).

Im Entity-Relation-Diagramm in der Abbildung 3.2 ist abzulesen, dass der Entität „Häufigkeit“ das Attribut „Namenvorkommen“ zugeordnet ist. Dementsprechend sind für das Mapping auf graphische Elemente und visuelle Attribute die Spalten „Gesamt“ und die mit dem ausgewählten Namen bezeichnete Spalte in der Datei `karte.csv` relevant. Die Datenwerte in diesen Tabellenspalten geben jeweils Größenordnungen an. Damit liegen auch hier *quantitative Daten* vor.

Mit der Verwendung der JavaScript-Bibliothek Leaflet für das Visualisierungsidiom der interaktiven Verbreitungskarte werden im Gegensatz zur derzeit für das Wörterbuch eingesetzten Präsentation der Namensverbreitungskarten geographische Karten eingesetzt. Bei der Visualisierung der oben beschriebenen zugrunde liegenden Daten dienen diese geographischen Karten als Hintergrundkarten für die verschiedenen deckungsgleichen Layer mit Kreisen, die bei jedem Aufruf einer interaktiven Karte sowie nach bestimmten Benutzerinteraktionen erzeugt werden und in denen die Informationen über einen ausgewählten Familiennamen zusammengefasst sind. Die Leaflet Tile Provider der in den interaktiven Karten des Demonstrators eingesetzten Hintergrundkarten Mapbox und Carto liefern eine Fülle von ausgewählten geographischen Daten mit. Es handelt sich hierbei z.B. um Topographie-, Gewässer- Verkehrswege-, Orts-, und ähnliche geographisch-räumliche Daten, die als statische, durch Benutzerinteraktionen nicht veränderbare Layer deckungsgleich übereinander gelegt werden und sich so zu einer Hintergrundkarte zusammensetzen. Bei diesen Daten handelt es sich um *geometrische Daten*. Sie sind nicht nur für die Entscheidung für für das Wörterbuch geeignete Hintergrundkarten relevant. Sie erlauben vor allem die mithilfe von Geokoordinaten präzise Angabe von *Positionen* in der geographischen Karte.

Obwohl sie für die Abbildung der Daten auf visuelle Repräsentationen nur mittelbar relevant sind, werden die bisher nicht beschriebenen Spalten in den Tabellen der Dateien `plz_1.csv`, `plz_2.csv`, `plz_3.csv`, `plz_5.csv` und `karte.csv` kurz erläutert. Die Dateneinträge in diesen Spalten bestehen zum Einen aus metadatenähnlichen Informationen wie den verschiedenen (alpha)numerisch codierten Bezeichnungen für die Postleitzahlenbereiche in den Spalten „key“, „name“ bzw. „PLZ-Bezirk“, die z.T. in den Popup-Texten über den Kreisen erscheinen. Zum Anderen stellen sie Verknüpfungen unter den einzelnen Tabellen dar oder werden für die Neuberechnung der verschiedenen Layer während der Verarbeitung von Benutzeraktionen gebraucht. So sind die Boole'schen Ausdrücke `true` und `false` in den Spalten „unmerged“ und „merged“ für die Umrechnungen relevant, die je nach Auswahl der Checkbox „Städte zusammenfassen“ im Interaktionsfeld „Parameter“ rechts neben der

interaktiven Karte nötig sind. Die Wahrheitswerte dieser Boole'schen Ausdrücke beeinflussen jeweils die Ausdehnung eines individuellen Postleitzahlenbereichs. Wird in der interaktiven Karte die Parametereinstellung „Städte zusammenfassen“ gewählt, so werden beispielsweise für die Großstadt Berlin etliche Postleitzahlenbereiche der Stadt zu einer Region mit einem entsprechend berechneten geographischen Mittelpunkt zusammengefasst. Die Datensätze mit den Kodierungen und Benennungen für die resultierenden Postleitzahlenbereiche und den entsprechend berechneten Geokoordinaten für die Mittelpunkte dieser Postleitzahlenbereiche sind vollständig in den plz-Dateien enthalten.

4.5.2 Visuelle Kodierung der Daten

Mit der visuellen Kodierung der oben beschriebenen Daten, d.h. dem Mapping der quantitativen Daten auf graphische Elemente mit gestaltenden visuellen Attributen, sollen möglichst viele der Fragen, die die Betrachter*innen an die Visualisierung stellen, beantwortet werden. Für das Design der Visualisierung und der Interaktivität in den Verbreitungskarten bedeutet das im Wesentlichen, dass für jeden einzelnen Layer, der über eine Hintergrundkarte gelegt wird, aussagekräftige und effektive visuelle Repräsentationen der in den Spalten der Tabellen enthaltenen Datenwerte gefunden werden müssen. Der Zusammenhang zwischen den zugrunde liegenden Zahlenwerten der individuellen Namensvorkommen und den Ausprägungen der durch die visuellen Attribute gestalteten graphischen Elemente muss für die Betrachter*innen nicht nur klar erkennbar sondern möglichst spontan erfassbar sein. Die visuellen Repräsentationen sollten so ausgewählt werden, dass Korrelationen zwischen den Namensvorkommen und den räumlichen Positionen in den Karten unmittelbar herausgelesen werden können (vgl. Schumann und Müller 2000, S. 222).

In den geographischen Hintergrundkarten, die für die interaktiven Verbreitungskarten verwendet werden, kann mithilfe der Angabe von Geokoordinaten die Position eines jeden Ortspunkts innerhalb der Karte präzise festgelegt werden. Wie oben beschrieben, enthalten die quantitativen Daten des interaktiven Kartenmoduls die berechneten Zahlenwerte für die Längen- und Breitengrade der Mittelpunkte von Postleitzahlenbereichen. Im Einklang mit dem Prinzip der Expressivität lassen sich diese Daten auf das graphische Element *Punkt* in Verbindung mit dem für quantitative Daten passenden visuellen Attribut *Position auf einer nicht ausgerichteten Skala*, kurz *Position*, abbilden. Im Mapping-Modell für das interaktive Kartenmodul ist das visuelle Attribut *Position* im Ranking der effektivsten visuellen Attribute am höchsten eingestuft (s. Tabelle 4.2). Dem gleichen graphischen Element *Punkt* können zur weiteren Ausgestaltung außer der *Position* weitere visuelle Attribute zugeordnet werden. Für die Zuordnung visueller Attribute zur Repräsentation der quantitativen Daten mit den Zahlenwerten der Namensvorkommen in der Datei *karte.csv* kommen im Einklang mit dem Expressivitätsprinzip die visuellen Attribute *Flächengröße*, *Farbhelligkeit* und *-sättigung* und *Volumen* in Frage. Mit der Auswahl eines der in der Tabelle 4.2 an vorletzter Stelle aufgeführten visuellen Attribute *Farbhelligkeit* bzw. *Farbsättigung* würden die in den Mittelpunkten der Postleitzahlenbereiche positionierten Punkte durch die Zuordnung unterschiedlicher Farbabstufungen, die verschieden große Namensvorkommen repräsentieren, weiter ausgestaltet. Folgt man den Forschungsergebnissen zu den visuellen Wahrnehmungsprozessen im menschlichen Gehirn, so kommen diese Attribute nur dann in Betracht, wenn in der

Farbhelligkeit nicht mehr als zwei bis vier verschiedene Abstufungen vor einem homogenen Hintergrund bzw. nicht mehr als drei Abstufungen in der Farbsättigung benutzt werden (vgl. Munzner 2014, Kap. 10). Für die visuelle Kodierung der Zahlenwerte der Namensvorkommen werden jedoch in den meisten Fällen deutlich mehr Abstufungen gebraucht. Zudem bilden die verwendeten geographischen Karten keinen homogenen Hintergrund. Die Verwendung dieser Attribute zur weiteren Ausgestaltung der Punkte würde also hier dem Prinzip der Effektivität widersprechen. Ähnliches gilt auch für das rangniedrigste visuelle Attribut Volumen. Bei seiner Verwendung würden die Namensvorkommen beispielsweise in Form von verschieden hohen dreidimensionalen Säulen über den Punkten repräsentiert. Diese Repräsentation der Namensvorkommen in der geographischen Karte würde insbesondere bei einer größeren Anzahl von Datenwerten schnell zu einer unübersichtlichen Darstellung führen, die die Wahrnehmungsfähigkeit der Betrachter*innen überfordern würde und damit ineffektiv wäre. Damit fällt die Wahl auf die Ausgestaltung der Punkte durch die Zuordnung des im Effektivitätsrang höher stehenden visuellen Attributs *Flächengröße*. Mit dem visuellen Attribut Flächengröße wird dem Punkt zusätzlich zur Position im Mittelpunkt eines Postleitzahlenbereichs eine Ausdehnung zugeordnet, die sich jeweils aus den Zahlenwerten der beiden die Namensvorkommen beziffernden Tabellenspalten in der Datei `karte.csv` errechnen lässt. Gleichzeitig erhalten alle Punkte sowohl eine einheitliche Farbe als auch eine einheitliche Form, auf die sich der Wert der Ausdehnung anwenden lässt. Da in den interaktiven Verbreitungskarten im Kartenmodul weder die Farbe noch die Form der Punkte informationstragend sind, können zunächst die Farbe und die für die Unterscheidbarkeit verschiedener Layer über der Hintergrundkarte relevante Farbsättigung der Punkte beliebig gewählt werden. U.a. festgelegt durch die entsprechenden Berechnungen in der Software-Bibliothek des DFD-Kartierungstools erhält schließlich der Punkt die Form eines Kreises, dem über das visuelle Attribut Flächengröße ein konkreter Durchmesser zugeordnet wird.

4.6 Wireframes

Analog zu dem in der Abbildung 2.2 dargestellten Schema der Prozessschritte der Anwendungsentwicklung werden nach der Benutzerrecherche, der Festlegung der Informationsarchitektur, der Entwicklung des Datenmodells und der Beschreibung und Umsetzung des Mapping-Modells für die visuelle Kodierung in den interaktiven Karten die Wireframes, d.h. die prototypischen Ansichten der Anwendungsseiten, entworfen (vgl. Lynch und Horton 2016, S. 130). Die Wireframes entstehen analog zu der Liste von Aufgaben, die sich aus dem im Abschnitt 2.5.2 aufgeführten Fragen- und Aufgabenkatalog ergeben, und schließen alle durch die Interaktionen der Nutzer*innen erzeugten Ansichten der Namensverbreitungskarten mit ein (vgl. Lynch und Horton 2016, S. 230-243).

4.6.1 Startseite mit Namensliste

Die Abbildung 4.5 zeigt das Wireframe der Startseite des Demonstrators für das interaktive Kartenmodul. Auf dieser Seite wird den Nutzer*innen eine Liste von elf unterschiedlichen Familiennamen präsentiert. Der Klick auf einen der Namen führt auf die Seite mit dem

entsprechenden Namenartikel und der Vorschaukarte. Dieser Teil der Spezifikation erfüllt die Anforderung A1 im Fragen- und Aufgabenkatalog im Abschnitt 2.5.2.

4.6.2 Namenartikel mit Vorschaukarte

Die Abbildung 4.6 zeigt das Wireframe für die Seite mit dem Namenartikel und der neu gestalteten Vorschaukarte am Beispiel des Familiennamens Richter. Die Kreise auf der Vorschaukarte sind rot und halb transparent. Die Vorschaukarte erscheint im Kartenstil Mapbox Streets des Leaflet Tile Providers Mapbox. Durch die entsprechende Konfiguration der Software können als alternative Hintergrundkarten auch die Mapbox-Karte Mapbox Light, die Karte CartoDB Positron des Leaflet Tile Providers Carto oder die eigens erzeugte schwarz-weiße Variante der Karte DB Positron eingesetzt werden. Nähere Informationen zu den Auswahlkriterien für die verwendeten Karten finden sich im Kapitel 7. Von der Seite mit dem Namenartikel führt ein Link zurück zur Seite mit der Namensliste. Mit dem Klick auf die Vorschaukarte öffnet sich die Seite mit der interaktiven Verbreitungskarte. Dieser Teil der Spezifikation erfüllt die Anforderung A2 im Fragen- und Aufgabenkatalog im Abschnitt 2.5.2.

4.6.3 Anwendungsseite mit interaktiver Verbreitungskarte

Auf der Anwendungsseite mit der interaktiven Verbreitungskarte werden der ausgewählte Name und das Gesamtvorkommen des Namens in Deutschland angezeigt. Die interaktive Verbreitungskarte selbst nimmt zusammen mit zwei Interaktionsfeldern rechts neben der Karte den größten Raum auf der Seite ein. Nach dem Öffnen der Seite erscheint die interaktive Verbreitungskarte in der Default-Darstellung und wie die Vorschaukarte im Kartenstil Mapbox Streets. Die Zoomstufe der Karte ist für alle Familiennamen so voreingestellt, dass Deutschland auf der Karte zentriert erscheint. Die Kreise, die die Verbreitung des Namens repräsentieren, werden als Layer über die Hintergrundkarte gelegt. Die Innenflächen der Kreise sind transparent. In der linken unteren Ecke der Karte ist eine Legende eingeblendet, die neben dem Familiennamen, dem Gesamtvorkommen des Namens sowie der Angabe der eingestellten Stelligkeit der Postleitzahlenbereiche statistische Angaben zu den Häufigkeiten enthält, die jeweils durch den kleinsten bzw. den größten Kreis in der Karte visualisiert werden. Alle interaktiven Bedienelemente sowohl in der linken oberen Ecke der Verbreitungskarte als auch der Reset-Button im Interaktionsfeld „Karte“ rechts neben der Verbreitungskarte sind zur Erklärung ihrer Funktion mit sogenannten Tooltips versehen (vgl. Lynch und Horton 2016, S. 240). Tooltips sind beim Schweben mit der Maus über einem Button eingeblendete kurze Texte, die Informationen zur Funktion des Buttons geben.³ Entgegen der Konvention für Webdesign und Interaktivität (vgl. Lynch und Horton 2016, S. 239) sind die Checkboxes in den Interaktionsfeldern „Parameter“ und „Karte“ rechts neben der Verbreitungskarte per Default angewählt, da hier das festgelegte einheitliche Schema für die automatische Erzeugung der Default-Karte maßgeblich ist. Die

³Einige der Buttons wurden mit Icons aus dem Angebot des Open Source Projekts Font Awesome (<https://fontawesome.com/icons>) gestaltet.

namenforschung.net
Demo

DIGITALES FAMILIENNAMENWÖRTERBUCH DEUTSCHLANDS (DFD)

Ein Projekt im Portal namenforschung.net

WÖRTERBUCH
PROJEKTVORSTELLUNG
MITARBEITER/INNEN
AKTUELLES

Wörterbuch

Derzeit 34712 veröffentlichte Namen

Anfangsbuchstabe wählen:

A B C D E F G H I J K L M N O P Q R
S T U V W X Y Z ALLE

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 » »»

Name	Rang	Häufigkeit
Abbel	93147	31
Darmstädter	31544	117
Demir	1064	2704
Klein	15	53378
Nowak	156	11551
Richter	12	59950
Scholz	44	30010
Schwalie	174982	12
Winterheimer	47728	73

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 » »»

Hilfe und FAQ

Antworten auf häufig gestellte Fragen
Benutzungshinweise und Hilfe

Namensuche

Exakte Zeichenkette
 Beginnt mit
 Endet mit
 Enthält

SUCHEN

Themen und Begriffe

Thematische Informationen

Glossar

Namenartikel-Listen

Neu veröffentlichte Namenartikel

Gesamtliste veröffentlichter
Namenartikel

Ein Projekt der

Akademie der
Wissenschaften und der
Literatur | Mainz
Geschwister-Schöll-Str. 2
55131 Mainz
www.adwmainz.de

Projektpartner

Technische Universität
Darmstadt
www.tu-darmstadt.de

Johannes Gutenberg-
Universität Mainz
www.uni-mainz.de

Weitere Informationen

- [Kontakt](#)
- [Impressum](#)
- [Datenschutzerklärung](#)
- [Sitemap](#)

Abbildung 4.5: Startseite des Demonstrators für das interaktive Kartenmodul mit Namensliste

DIGITALES FAMILIENNAMENWÖRTERBUCH DEUTSCHLANDS (DFD)
 Ein Projekt der Leibniz-Universität Hannover

Zurück zur Liste DFD-DFD

Richter
 Abgebildet in: DFB-Atlas, DFB-Atlas, DFB-Atlas, DFB-Atlas, DFB-Atlas

Allgemeines

Herkunft:	9993
Baum:	12
Spezialnamen:	Deutsch
Ursprungsform:	Deutschland

Kategorien

- in Abteilung 1
- mit Ortsnennungen und -bezeichnungen
- mit Ortsnamen
- mit

Etimologie

Die Bedeutung des Namens ist unklar. Ein möglicher Ursprung ist der Name 'Richter', der von dem althochdeutschen 'rihter' (Richter) abstammt. Ein weiterer möglicher Ursprung ist der Name 'Richter', der von dem althochdeutschen 'rihter' (Richter) abstammt.

Verbreitung

Verbreitung innerhalb Deutschlands

Verbreitung außerhalb Deutschlands

Technische Republik:	
Herkunft:	9993
Angenommen von Quelle:	
Art der Quelle:	Einzelnamenliste
Jahr der Quelle:	2012
Quelle:	Wikipedia, letzter Zugriff: 13.03.2014

Publ.

Herkunft:	9993
Angenommen von Quelle:	
Art der Quelle:	Einzelnamenliste
Jahr der Quelle:	2009
Quelle:	Wikipedia, letzter Zugriff: 13.03.2014

Historische Verbreitung

Zur historischen Verbreitung siehe Namensverbreitungstabelle, letzter Zugriff: 13.03.2014

Verwandte Artikel (Auswahl)

- [Richter](#)
- [Richter](#)
- [Richter](#)

Literaturhinweise

Literatur

- [Richter](#), in: *Die Deutschen*, herausgegeben von Gerd Gammeter, 1994, S. 114.
- [Richter](#), in: *Die Deutschen*, herausgegeben von Gerd Gammeter, 1994, S. 114.
- [Richter](#), in: *Die Deutschen*, herausgegeben von Gerd Gammeter, 1994, S. 114.
- [Richter](#), in: *Die Deutschen*, herausgegeben von Gerd Gammeter, 1994, S. 114.

Weblogs

- [Richter](#), in: *Die Deutschen*, herausgegeben von Gerd Gammeter, 1994, S. 114.
- [Richter](#), in: *Die Deutschen*, herausgegeben von Gerd Gammeter, 1994, S. 114.
- [Richter](#), in: *Die Deutschen*, herausgegeben von Gerd Gammeter, 1994, S. 114.

Metadaten

Open zur Ansicht

DFD-DFD

DFD-DFD

DFD-DFD

DFD-DFD

Abbildung 4.6: DFD-Namenartikel mit neuer Vorschaukarte für den Namen Richter

Abbildung 4.7 zeigt das Wireframe für die Seite mit der interaktiven Verbreitungskarte für den Familiennamen Richter. Die Default-Karte zeigt die Verbreitung des Namens mit den automatisch vorgenommenen Parametereinstellungen. Für den Namen Richter sind dies für einige Städte zusammengefasste Postleitzahlenbereiche, relative Zahlenwerte, dreistellige Postleitzahlenbereiche und Kreise in der Farbe Rot. Ein Link führt zurück auf die Seite mit dem Namenartikel. Dieser Teil der Spezifikation erfüllt die Anforderung A3 im Fragen- und Aufgabenkatalog.

4.6.4 Manipulation der Parametereinstellungen und Berechnung der Kreise

Mit den Auswahlmöglichkeiten im Interaktionsfeld „Parameter“ können die Nutzer*innen die Parametereinstellungen „Städte zusammenfassen“, „Absolute“ oder „Relative Zahlenwerte“, „PLZ-Stelligkeit“ sowie „Farbe“ und „Transparenz der Kreisflächen“ interaktiv verändern. Die entsprechenden Wireframes zeigen die Ansichten der Verbreitungskarte, nachdem jeweils eine der möglichen Manipulationen an der Default-Karte vorgenommen wurde. Grundsätzlich können beliebige Änderungen von Parametern auch unmittelbar nacheinander ausgeführt und miteinander kombiniert werden. Jede Benutzeraktion erzeugt einen neu berechneten Layer mit Kreisen über der Karte. Jeweils nach der Manipulation der Parameter „Städte zusammenfassen“, „Absolute“ oder „Relative Zahlenwerte“ und „PLZ-Stelligkeit“ zeigt die Legende unten links die jeweils angepasste Statistik für die geänderte Darstellung der Verbreitung. Unmittelbar nach der Änderung eines Parameters der Default-Karte wird zudem der Reset-Button mit dem Tooltip „Parametereinstellungen zurücksetzen“ in dem Interaktionsfeld „Karte“ aktiv. Über diesen Button ist die Rückkehr zur Default-Karte zu jedem Zeitpunkt möglich. Die Abbildung 4.8 zeigt das Wireframe für die interaktive Verbreitungskarte für den Namen Richter nach dem Umschalten auf „Absolute Zahlenwerte“. Die Legende unten links zeigt die angepasste Statistik für die geänderte Darstellung der Verbreitung. Der Reset-Button für das Zurücksetzen der Parametereinstellungen findet sich unten links in dem Interaktionsfeld „Karte“. Dieser Teil der Spezifikation erfüllt die Anforderungen A11, A12 und A14 im Fragen- und Aufgabenkatalog. Alle weiteren Wireframes, die die Karte nach der Änderung eines der oben genannten Parameter zeigen, finden sich im Anhang A.1.

4.6.5 Druckversion

Nach dem Betätigen des „Drucken“-Buttons im Interaktionsfeld „Karte“ wird eine Druckversion der Karte erstellt. Die Druckversion enthält weder Kontrollelemente noch Tasten. Da beim Druck oder Export eines herangezoomten Ausschnitts der Verbreitungskarte die statistischen Angaben in der Legende zu den Häufigkeiten, die jeweils durch den kleinsten bzw. den größten Kreis visualisiert werden, irreführend wären, ist auch die Legende in der Druckversion ausgeblendet. Dargestellt wird der aktuelle Kartenausschnitt in dem aktuell gewählten Kartenstil mit den individuell berechneten und gestalteten Kreisen sowie ggf. gesetzten Markern. Ein Informationstext oberhalb der Karte gibt den Familiennamen und das Gesamtvorkommen des Namens an. Die Abbildung 4.9 zeigt eine Druckversion der

namenforschung.net English

DIGITALES FAMILIENNAMENWÖRTERBUCH DEUTSCHLANDS (DFD)

Ein Projekt im Portal namenforschung.net

WÖRTERBUCH
PROJEKTVORSTELLUNG
MITARBEITER/INNEN
AKTUELLES

[Zurück zum Namenartikel](#) [Hilfe FAQ](#)

Richter 59950

PARAMETER

Städte zusammenfassen

Absolute Zahlenwerte

Relative Zahlenwerte

Farbe

PLZ-Stelligkeit 3

Transparenz der Kreisflächen: 0.2

KARTE

Legende einblenden

↺
🔒

Ein Projekt der

Akademie der Wissenschaften und der Literatur | Mainz
Geschwister-Scholl-Str. 2
55131 Mainz
www.adwmainz.de

Projektpartner

Technische Universität Darmstadt
www.tu-darmstadt.de

Johannes Gutenberg-Universität Mainz
www.uni-mainz.de

Weitere Informationen

- Kontakt
- Impressum
- Datenschutzerklärung
- Sitemap

Abbildung 4.7: Anwendungsseite mit interaktiver Verbreitungskarte für den Namen Richter. Rechts neben der Karte werden Interaktionsfelder angeboten.

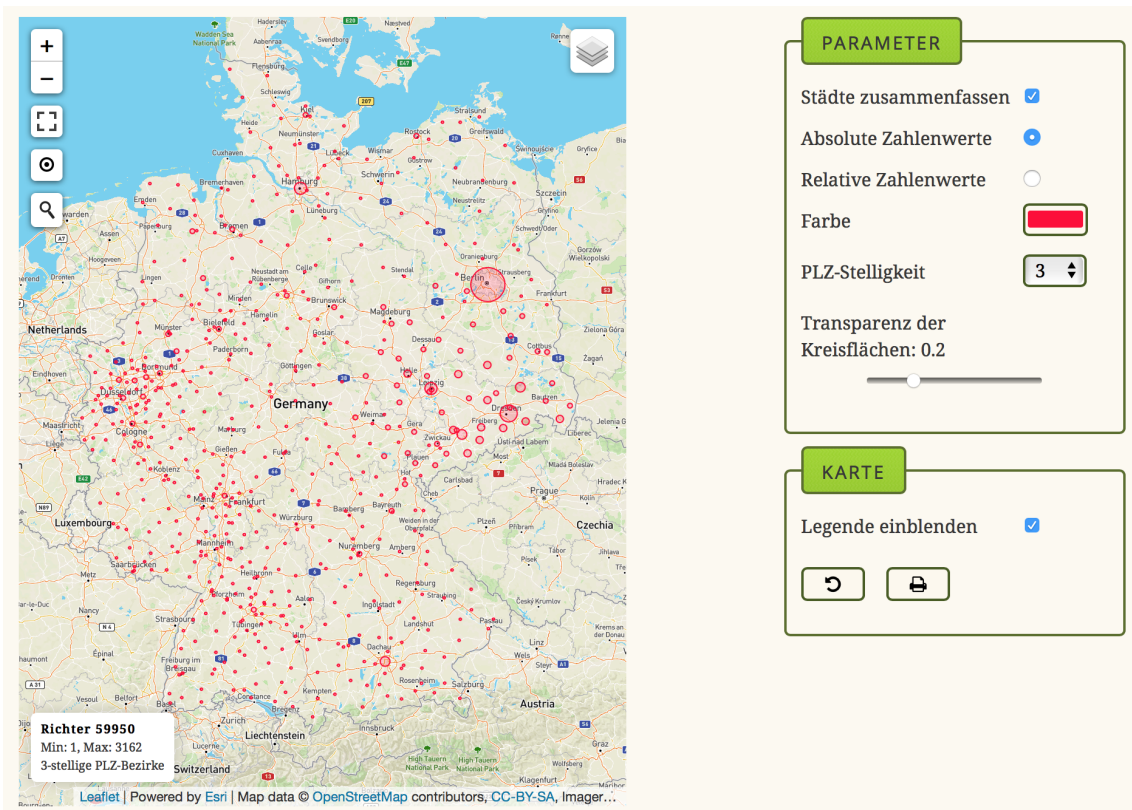


Abbildung 4.8: Interaktive Verbreitungskarte für den Namen Richter im Kartenstil Mapbox Streets nach dem Umschalten auf „Absolute Zahlenwerte“

interaktiven Verbreitungskarte für den Namen Richter. Dieser Teil der Spezifikation erfüllt die Anforderung A13 im Fragen- und Aufgabenkatalog.

4.6.6 Ein- und Ausblenden der Legende

Die Nutzer*innen können durch die entsprechende Auswahl im Interaktionsfeld „Karte“ die Legende aus- und wieder einblenden. Aus lizenzrechtlichen Gründen ist die Anzeige der konkreten Zahl der absoluten bzw. relativen Häufigkeit eines Namens innerhalb der Karte, der Legende oder der Informationstexte über den Kreisen nicht möglich. Obwohl die Legende in dieser Form wenig aussagekräftig ist, wird in Abstimmung mit den Artikelautor*innen des DFD als Orientierung für die Nutzer*innen in der einblendbaren Legende der kleinste bzw. der größte Zahlenwert, der jeweils durch den kleinsten bzw. größten Kreis in der Karte repräsentiert wird, angezeigt. Das Wireframe in der Abbildung 4.10 zeigt einen Kartenausschnitt mit ausgeblendeter Legende. Dieser Teil der Spezifikation erfüllt die Anforderung A15 im Fragen- und Aufgabenkatalog.

4.6.7 Zoomen, Verschieben und Rückkehr zur Ausgangskarte

Der Kartenausschnitt der interaktiven Verbreitungskarte kann verschoben und gezoomt werden. Hierbei sind die Zoomstufen aus Datenschutzgründen nach oben begrenzt, d.h. es ist nicht möglich, etwa einzelne Wohngebiete oder Straßen beliebig nah heran zu zoomen, um mögliche Rückschlüsse von den visualisierten Daten auf konkrete Wohnorte zu unterbinden. Der Button „Ausgangskarte“ führt zurück zur voreingestellten Zentrierung und Zoomstufe. Um zu verhindern, dass beim Heranzoomen die dem Zoom entsprechende Vergrößerung der Kreise dazu führt, dass u.U. eine gesamte Region durch die farbige, transparente Innenfläche eines oder mehrerer Kreise verdeckt wird, werden beim Zoom in die Karte die Größen der Kreise durch entsprechende Berechnungen linear nach unten korrigiert. Die Abbildung 4.10 zeigt das Wireframe für einen vergrößerten Ausschnitt der interaktiven Verbreitungskarte für den Namen Richter. In diesem Wireframe ist die Legende ausgeblendet. Dieser Teil der Spezifikation erfüllt die Anforderungen A5 und A6 im Fragen- und Aufgabenkatalog.

4.6.8 Ortssuche

Die Abbildungen 4.11 und 4.12 zeigen die Wireframes für die Ortssuche in der interaktiven Verbreitungskarte für den Namen Richter. Nach dem Anwählen des Buttons „Ortssuche“ und der Eingabe eines Ortsnamens, einer Region oder einer Postleitzahl öffnet sich ein Dropdown Menü mit einer Auswahl von bis zu fünf Treffern. Nach der Anwahl eines der Treffer wird an der entsprechenden Position in der Karte ein wieder entfernbare Marker gesetzt und es wird ein automatischer Zoom in die Region ausgelöst. Alternativ zur Anwahl eines der Treffer im Dropdown-Menü können die Nutzer*innen auch die Ortssuche insgesamt bestätigen. Das

Richter 59950

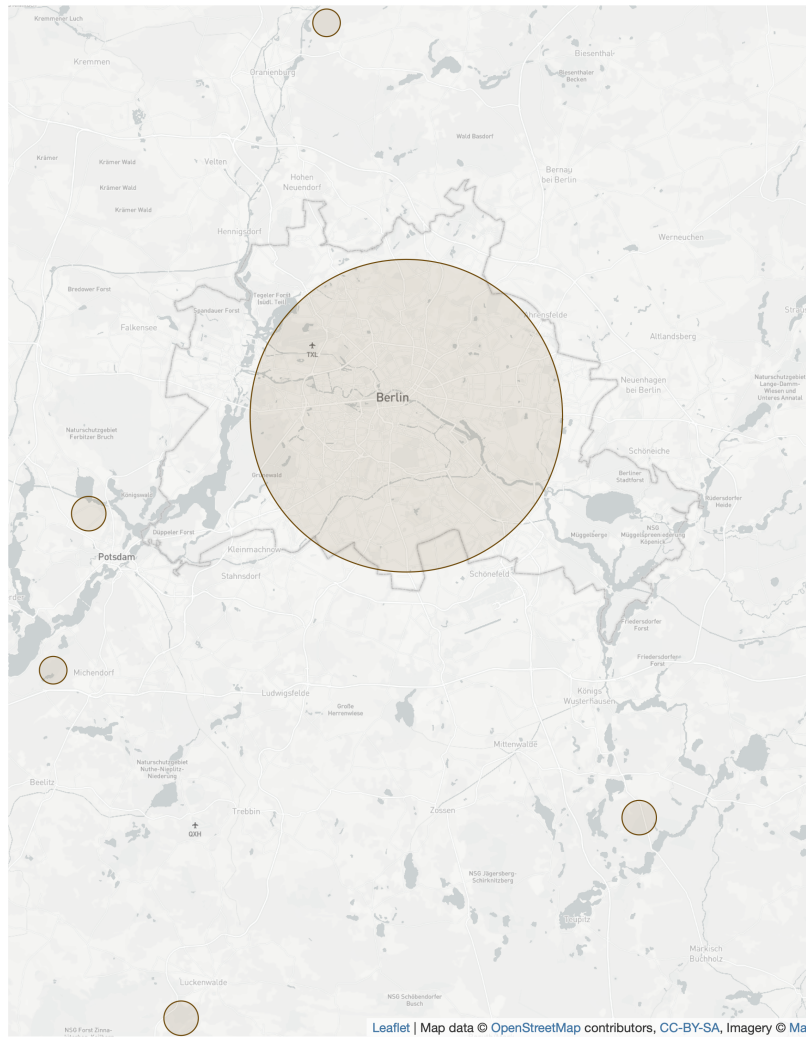


Abbildung 4.9: Druckversion der interaktiven Verbreitungskarte für den Namen Richter im Kartenstil Mapbox Light und mit „Absoluten Zahlenwerten“

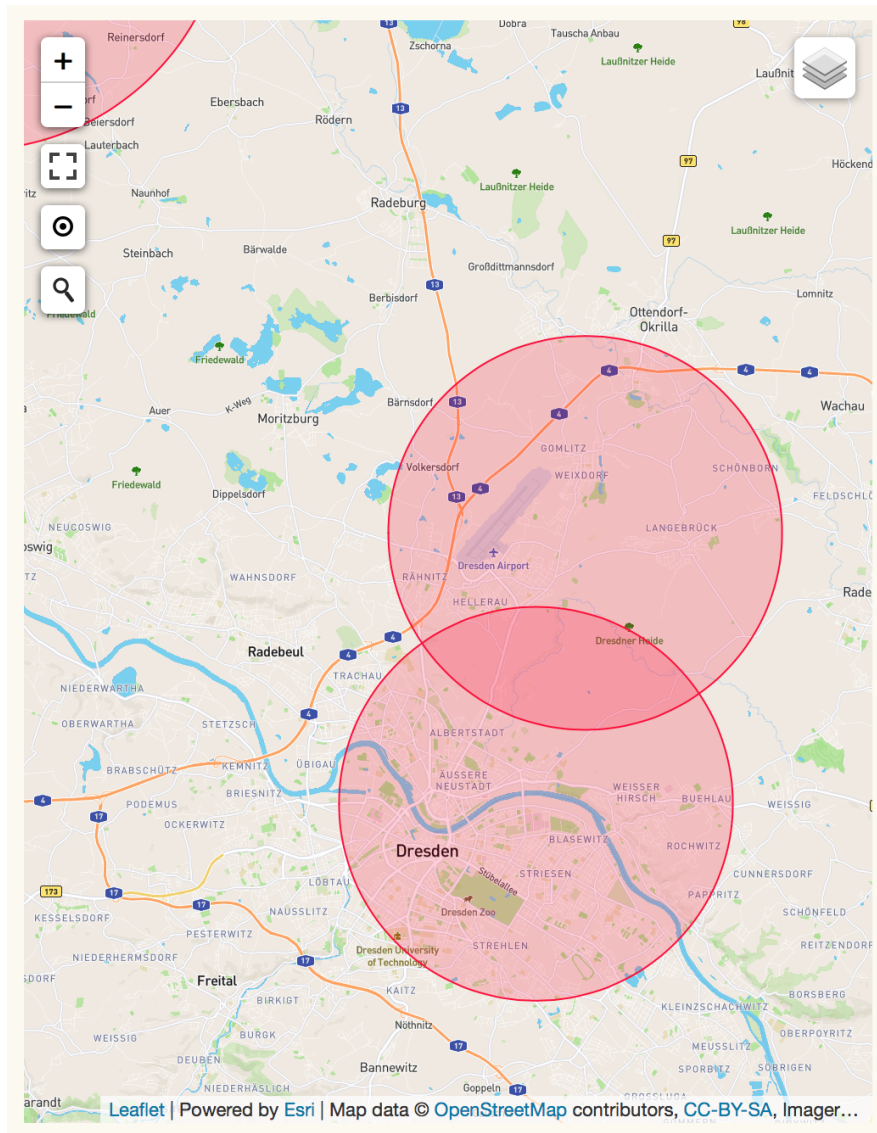


Abbildung 4.10: Interaktive Verbreitungskarte für den Namen Richter im Kartenstil Mapbox Streets nach dem Zoom in eine Region

Wireframe für diese Aktion findet sich im Anhang A.2. Dieser Teil der Spezifikation erfüllt die Anforderung A7 im Fragen- und Aufgabenkatalog.⁴

4.6.9 Alternativer Kartenstil

Die Abbildung 4.13 zeigt das Wireframe für die interaktive Verbreitungskarte für den Namen Richter in dem alternativen Kartenstil CartoDB Positron des Kartenanbieters Carto. In der rechten oberen Ecke der interaktiven Verbreitungskarte können die Nutzer*innen per Mouseover und anschließendem Klick zwischen den farbigen Hintergrundkarten Mapbox Streets bzw. CartoDB Positron wählen. Als Reminiszenz an die derzeit im Wörterbuch veröffentlichten Verbreitungskarten und auf Wunsch einiger Artikelautor*innen werden darüber hinaus die zwar geographischen, zoombaren und interaktiven, jedoch schlicht in schwarz-weiß gehaltenen Varianten der oben genannten Stile angeboten. Die Wireframes für den Kartenstil Mapbox Light bzw. für die eigens mithilfe eines Leaflet-Plugins für die Anwendung erzeugte schwarz-weiße Variante der Hintergrundkarte CartoDB Positron finden sich im Anhang A.3. Dieser Teil der Spezifikation erfüllt die Anforderung A4 im Fragen- und Aufgabenkatalog.

4.6.10 Informationstext über einem Kreis und Zoom in die Region

Jeder Kreis in der Verbreitungskarte kann von den Nutzer*innen angewählt werden um einen Popup-Text mit der Angabe der Bezeichnung für den angewählten Postleitzahlenbereich zu erhalten.⁵ Aus lizenzrechtlichen Gründen ist, wie oben bereits erwähnt, auch im Informationstext über den Kreisen die Angabe exakter Zahlen für die Namensvorkommen in den Postleitzahlenbereichen nicht möglich. Lizenzrechtlich möglich wäre die Angabe von Größenordnungen in Form von Intervallen wie „zwischen 20 und 40 Vorkommen“. Diese Option wurde aus zwei Gründen verworfen. Zum Einen geht wichtige Information verloren und zum Anderen könnte bei den Nutzer*innen der Eindruck entstehen, dass dem DFD keine exakten Daten zur Verbreitung eines Namens vorliegen. In den Popup-Texten über den Kreisen wird den Nutzer*innen zusätzlich der Button „Zoomen“ angeboten. Ein Klick auf diesen Button löst den automatischen Zoom in die Region um den betreffenden Kreis aus. Die Abbildung 4.14 zeigt das Wireframe für die interaktive Verbreitungskarte für den Namen Richter mit dem Informationstext über einem Kreis und nach dem Heranzoomen der durch den Kreis repräsentierten Region in der Karte. Dieser Teil der Spezifikation erfüllt die Anforderung A9 und m.E. die Anforderung A8 im Fragen- und Aufgabenkatalog.

⁴Zur Zeit der Implementierung der Ortssuche-Funktion wurde das Leaflet-Plugin Esri Leaflet Geocoder unentgeltlich angeboten (<https://github.com/Esri/esri-leaflet-geocoder>). In der Zwischenzeit ist ein kostenpflichtiger API-Key für die Verwendung erforderlich. In der Folge funktioniert die Ortssuche in der interaktiven Karte z.Zt. nicht. Unter den dieser Arbeit beigelegten Dateien findet sich u.a. ein während der Softwareentwicklung aufgenommenes Video, in dem alle Interaktionen, auch die Ortssuche, in der Karte demonstriert werden (s. Datei `maps.mov`).

⁵In den Popup-Texten über den Kreisen erscheinen z.T. unklar formulierte Angaben wie „**xy“. Mit den Daten des DFD-Kartierungstools werden auch die Stern-Symbole in den Bezeichnungen für die Postleitzahlenbereiche übernommen. Wie bereits erwähnt, stehen diese im Zusammenhang mit der Aggregation der PLZ-Daten bzw. mit dem Zusammenfassen einzelner Postleitzahlenbereiche im DFD-Kartierungstool.

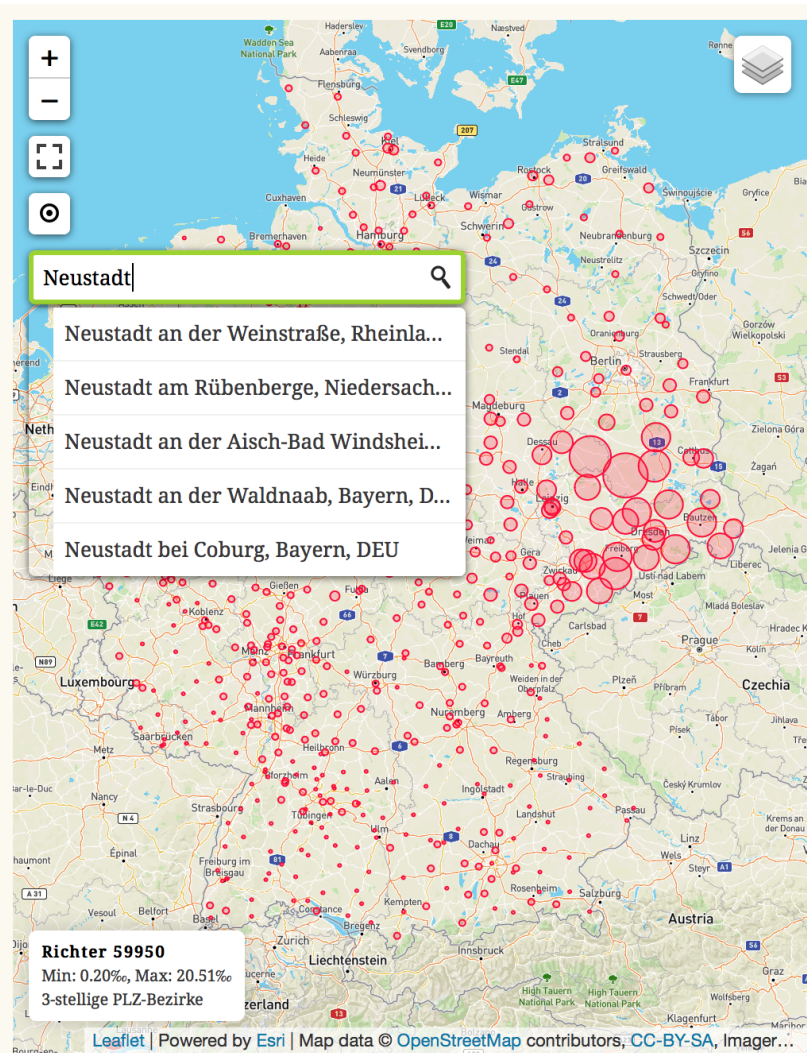


Abbildung 4.11: Ortssuche in der interaktiven Verbreitungskarte für den Namen Richter. Die Anwahl eines der Treffer der Ortssuche löst den Zoom in die Region um den entsprechenden Ort aus.

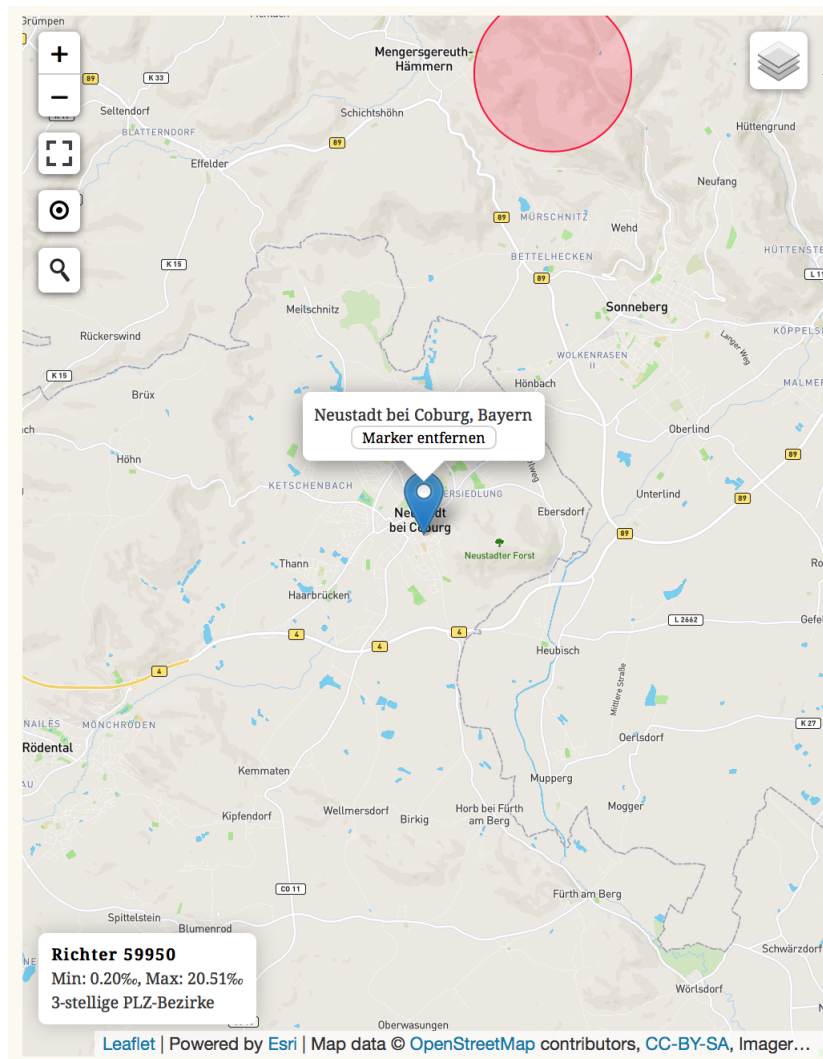


Abbildung 4.12: Ortssuche in der interaktiven Verbreitungskarte für den Namen Richter. An dem gefundenen Ort wird ein löscher Marker gesetzt.

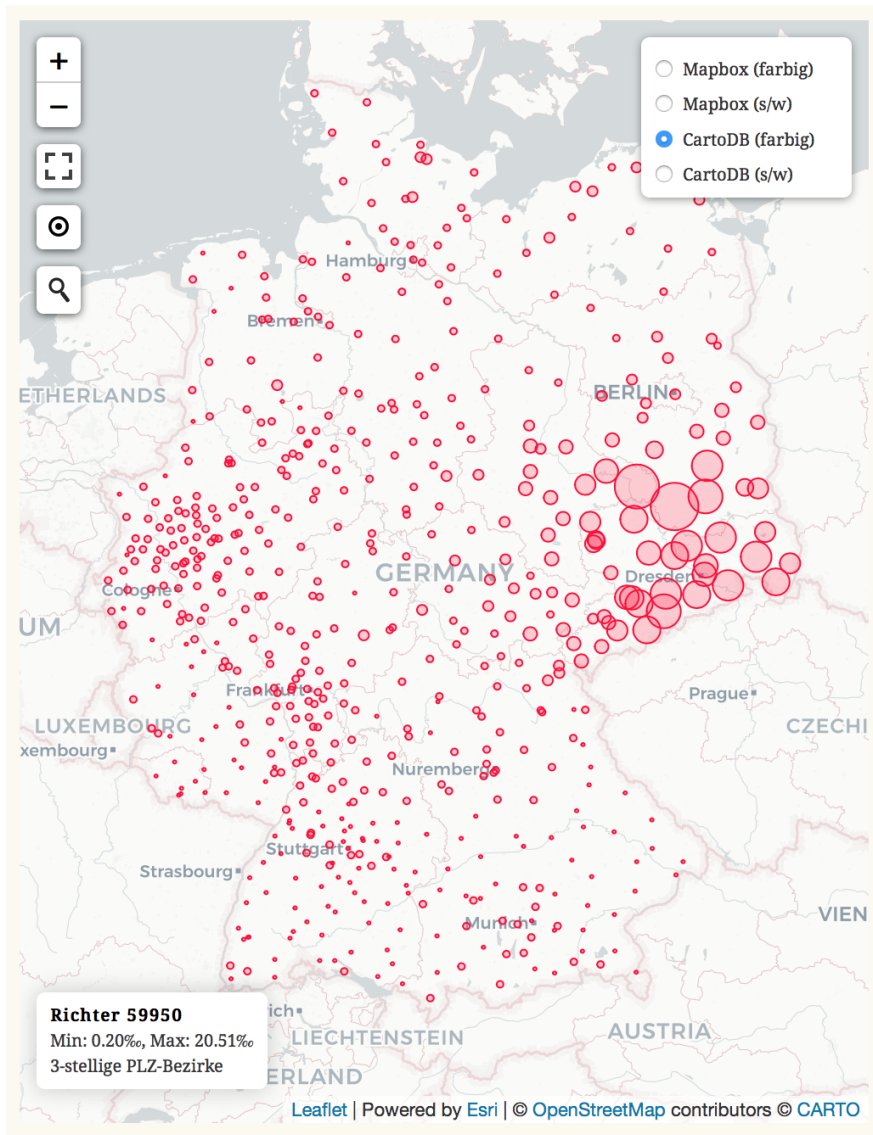


Abbildung 4.13: Interaktive Verbreitungskarte für den Namen Richter in dem alternativen Kartenstil CartoDB Positron

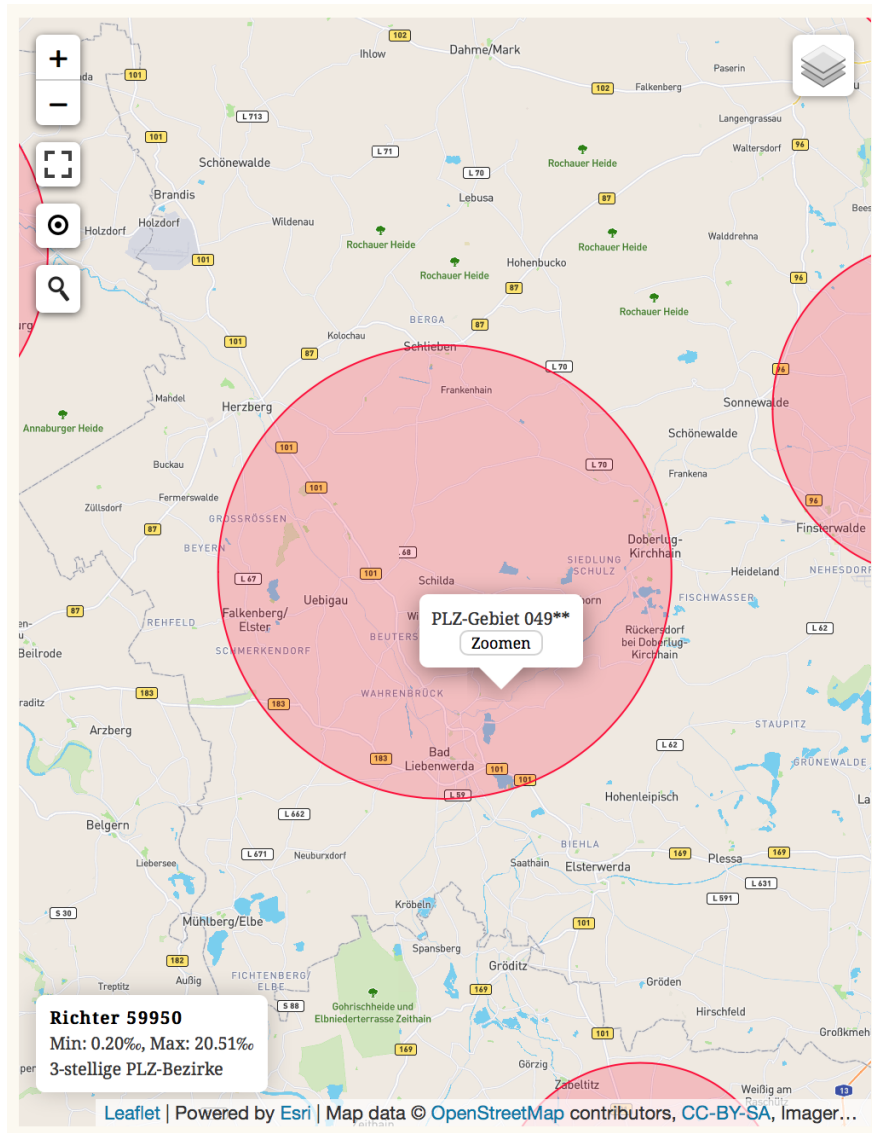


Abbildung 4.14: Interaktive Verbreitungskarte für den Namen Richter im Kartenstil Mapbox Streets mit dem Informationstext über einem Kreis und nach dem Zoom in die Region

4.6.11 Umschalten in den Vollbildmodus

Die interaktive Namensverbreitungskarte erscheint nach dem Klick auf den Button „Vollbild anzeigen“ in Bildschirmgröße (ohne Abbildung). Sämtliche Interaktionsmöglichkeiten innerhalb der Karte stehen zur Verfügung, nicht jedoch die Interaktionsmöglichkeiten in den Feldern rechts neben der Karte. Mit dem Betätigen des Buttons „Vollbild schließen“ wird der Fullscreen-Modus beendet. Dieser Teil der Spezifikation erfüllt die Anforderung A10 im Fragen- und Aufgabenkatalog.

5 Evaluation der Visualisierung im interaktiven Kartenmodul

Ein besonderes Augenmerk bei der Evaluation der Visualisierung im interaktiven Kartenmodul liegt auf der Bewertung der vom DFD-Kartierungstool übernommenen visuellen Repräsentation der Namensvorkommen durch verschieden große Kreise. Das Mapping der Geokoordinaten des Mittelpunkts eines Postleitzahlenbereichs auf das graphische Element Punkt, dem das hoch effektive visuelle Attribut Position und das in der Rangfolge der Effektivität an zweiter Stelle eingeordnete visuelle Attribut Flächengröße zugeordnet wird, ist eine Abbildung, über die Tamara Munzner schreibt: „In the context of visual encoding, point marks intrinsically convey information only about position and are exactly the vehicle for conveying additional information through area and shape.“ (vgl. Munzner 2014, S. 98). Was die Form der zugeordneten Fläche angeht, so könnte bei der visuellen Kodierung der Daten dem Punkt anstatt der Form eines Kreises beispielsweise auch die Form eines Rechtecks oder eines Dreiecks, dem über das visuelle Attribut Flächengröße jeweils eine konkrete Ausdehnung zugeordnet wird, eingesetzt werden. Nach den Erkenntnissen der Informationsvisualisierung werden Größenunterschiede bei Rechtecken mit gleich langer Grundseite, wie sie beispielsweise in Balkendiagrammen verwendet werden, vom menschlichen Sehapparat tatsächlich besser wahrgenommen als die Größenunterschiede bei Kreisflächen (vgl. Schumann und Müller 2000, S. 127-131). Beim Positions- und Größenvergleich in einer geographischen Karte kann diese „eckige“ Variante jedoch schnell zu einer unübersichtlichen und damit ineffektiven Darstellung führen. Ähnliches gilt auch für Dreiecke oder andere Flächenformen. Für die visuelle Kodierung der Namensvorkommen in den Verbreitungskarten des interaktiven Kartenmoduls ist demnach der mit dem visuellen Attribut Flächengröße ausgestaltete Kreis eine gute Wahl.

In seinem Buch „Visual Thinking for Design“ schreibt Colin Ware über die Wahrnehmungsprozesse, die sich in bestimmten Regionen des primären visuellen Cortex, einem Teil des menschlichen Sehapparates, abspielen: „These cortical areas are parallel computers because they process every part of the visual image simultaneously, computing *local orientation information*, *local color difference information*, *local size information*, and *local motion information*.“¹ (vgl. Ware 2008, S. 26). Zusammengefasst und übertragen auf die Visualisierung im interaktiven Kartenmodul heißt das für die Betrachter*innen, dass sie die Anordnung der farbigen Kreise, ggf. Gruppierungen von Kreisen oder einzelne Ausreißer vor der Hintergrundkarte besonders gut wahrnehmen und die Größenordnungen der Namensvorkommen an den verschiedenen Durchmesser der Kreise spontan ablesen und vergleichen können.

¹Hervorhebungen im Text wurden geändert bzw. hinzugefügt.

Es kann also an dieser Stelle festgehalten werden, dass die vom DFD-Kartierungstool für das interaktive Kartenmodul übernommene Repräsentation der Namendaten durch Kreise in Bezug auf die Expressivität und die Effektivität der entsprechenden visuellen Kodierung als insgesamt sehr effizient und benutzerfreundlich bezeichnet werden kann. Damit erfüllt das Visualisierungsideom des interaktiven Kartenmoduls ein wesentliches Qualitätskriterium.

Eine deutliche Verbesserung im Hinblick auf die Expressivität des Visualisierungsideoms der interaktiven Verbreitungskarten stellt die im Vergleich zur derzeitigen Praxis der Kartenerstellung veränderte Erzeugung der Default-Karte sowie die Implementierung der Interaktivität dar. Die visuelle Darstellung der Verbreitung eines Familiennamens wird derzeit durch die von den einzelnen Artikelautor*innen individuell festgelegten Parametereinstellungen im DFD-Kartierungstool definiert. Diese Praxis stellt eine Verletzung des Expressivitätskriteriums dar. Die Parametereinstellungen werden zwar mit Expertenwissen, letztlich aber doch nach subjektivem Verständnis vorgenommen. Die Visualisierung der Daten kann insofern nicht immer möglichst unverfälscht und vollständig sein und entspricht damit nicht dem Prinzip der Expressivität. Zudem kann die visuelle Repräsentation in der derzeit für das Wörterbuch erzeugten statischen Karte immer nur einen Aspekt der gesamten Datenmenge zeigen. Diesen Problemen wird im interaktiven Kartenmodul zum Einen mit einem einheitlichen Schema für die Erzeugung einer Default-Karte und zum Anderen mit der Implementierung der Interaktivität begegnet. Grundsätzlich definiert das Visualisierungsideom für das Kartenmodul exakt das, was die Nutzer*innen bei der Arbeit mit der Visualisierung auf ihrem Bildschirm sehen. Als Bestandteil des Visualisierungsideoms kontrolliert das sogenannte Interaktionsideom dagegen die Änderungen dessen, was die Nutzer*innen auf ihrem Bildschirm sehen (vgl. Munzner 2014, S. 71). Im Unterschied zu der Arbeit mit der derzeit im Wörterbuch angebotenen statischen Verbreitungskarte, können die Nutzer*innen innerhalb der automatisch erstellten Default-Karte über eigene Interaktionen die Änderung der Ansicht bewirken und so etliche weitere visuelle Informationen über die vergleichsweise komplexen Zusammenhänge in den Daten abrufen (vgl. Munzner 2014, S. 9 und 254). Im Einklang mit den Prinzipien der Expressivität und der Effektivität erhalten sie damit die Möglichkeit, aktiv Antworten auf die tiefergehenden Fragen, die sie an die Visualisierung stellen, zu suchen.

Mit der Implementierung der Interaktivität lässt sich noch ein weiterer problematischer Aspekt der visuellen Kodierung der Daten in den interaktiven Verbreitungskarten entschärfen. Bei der Auswahl eines geeigneten Farbtons für die Kreise über der geographischen Karte spielen neben visualisierungstechnischen Kriterien wie der Unterscheidbarkeit der Farbe der Kreise vor der Hintergrundkarte vor allem die Frage nach der Barrierefreiheit eine Rolle. In der derzeitigen Präsentation der Namensverbreitungen im Wörterbuch erscheinen in der überwiegenden Zahl der statischen Verbreitungskarten die Kreise in der Farbe Rot. Diese Farbwahl wurde auf Wunsch der Artikelautor*innen auch für die als Default-Karte angebotene interaktive Verbreitungskarte übernommen. Der Rotton der Kreise unterscheidet sich von allen anderen Farben in den Hintergrundkarten und ist damit für die Betrachter*innen gut erkenn- und unterscheidbar (vgl. Ware 2008, S. 33). In der interaktiven Default-Karte werden die Kreise zudem halb transparent dargestellt und geben damit zumindest teilweise den Blick auf die darunter liegende geographische Karte frei. In der Regel nimmt das menschliche Auge Farbe in drei verschiedenen Spektren auf. Das eine ist das Spektrum der

Farben von Rot bis Grün, das zweite das Spektrum der Farben von Blau bis Gelb und das dritte die Helligkeitsstufen von Schwarz und Weiß. Für das Design des Visualisierungsideoms der interaktiven Karte kann die Tatsache problematisch werden, dass bei rund 8% aller Männer eine Störung der Wahrnehmung im ersten Farbspektrum Rot bis Grün vorliegt (vgl. Munzner 2014, S. 220). Um dieser Sehschwäche im Sinne der Barrierefreiheit zu begegnen, sollte folgerichtig die Farbe Rot in den Karten durch eine andere Farbe ersetzbar sein. Die Interaktivität in den neu entwickelten Namensverbreitungskarten ermöglicht den Betrachter*innen die Auswahl eines für sie passenden Farbtons aus dem gesamten zur Verfügung stehenden Farbspektrum und erhöht auch hiermit die Expressivität und die Effektivität der Visualisierung.

Ein in der Verwendung der angepassten Software-Bibliothek des DFD-Kartierungstools begründeter Aspekt, der nicht ganz im Einklang mit den aktuellen Erkenntnissen der Informationsvisualisierung steht, ist die fehlende Angabe der Geltungsbereiche für die visuell kodierten Zahlenwerte in den interaktiven Verbreitungskarten (vgl. Schumann und Müller 2000, S. 223). In den Karten des interaktiven Kartenmoduls werden die visuellen Kodierungen der Namensvorkommen in durch verschiedene Benutzeraktionen erzeugten Layern vor dem Hintergrund einer geographischen Deutschlandkarte präsentiert. Damit ist der globale Geltungsbereich der in den zugrunde liegenden Daten enthaltenen Datenwerte veranschaulicht. Allerdings fehlt für die durch die einzelnen Kreise in den Mittelpunkten der Postleitzahlenbereiche repräsentierten Datenwerte die genaue visuelle Angabe ihres jeweiligen regionalen Geltungsbereichs. Zwar sind aufgrund der Verwendung von geographischen Karten als Hintergrund für die überblendeten Layer die Positionen der Kreise schon bei der Betrachtung der Default-Karte deutlich besser einzuordnen als in den derzeit im Wörterbuch eingesetzten Verbreitungskarten. Auch haben die Nutzer*innen die Möglichkeit, in eine Region, die sie interessiert, interaktiv hinein zu zoomen, um detailliertere Informationen zu erhalten. Für die vollständige Angabe der Geltungsbereiche für die Datenwerte fehlt jedoch auch in den interaktiven Verbreitungskarten die visuelle Präsentation der Ausdehnung des jeweils betrachteten Postleitzahlenbereichs. Nur mit der visuellen Repräsentation des Postleitzahlenbereichs können die Betrachter*innen der Visualisierung sehen, für welche Region genau ein durch den Kreis repräsentiertes Namensvorkommen gültig ist. Eine Erklärung dafür, aus welchem Grund bisher die Angabe der Geltungsbereiche in den interaktiven Verbreitungskarten fehlt, findet sich im Kapitel 7.

Ein weiterer Aspekt der visuellen Repräsentation der Namensverbreitungen im interaktiven Kartenmodul, der den Forschungsergebnissen der Informationsvisualisierung nicht entspricht, ist die Größenberechnung der Kreise. Während die oben beschriebene Auswahl des graphischen Elements Punkt mit den visuellen Attributen Position und Flächengröße sowohl für die derzeitige Präsentation der Verbreitungskarten im Wörterbuch als auch für die interaktiven Verbreitungskarten im Kartenmodul den Kriterien der Expressivität und der Effektivität genügen, gilt das nicht für die benutzerfreundliche Darstellung der Flächengrößen der Kreise. In der JavaScript-Bibliothek des DFD-Kartierungstools wird für das Mapping der Namensvorkommen auf verschieden große Kreise jeweils die Änderung des Durchmessers eines Kreises in direkter Abhängigkeit vom darzustellenden Zahlenwert des Namensvorkommens berechnet. Durch die Anpassung der Software-Bibliothek des Kartierungstools auf die hier beschriebene Anwendung wird diese Berechnung auch auf die

Visualisierung der Namensvorkommen in den interaktiven Verbreitungskarten angewendet (s. Datei `circles.js`). Laut Schuhmann und Müller ist „beim visuellen Vergleich von zwei Kreisen [...] jedoch der Flächeninhalt relevant und nicht der Durchmesser. Eine solche Darstellung verfälscht somit die zugrunde liegenden Daten und verstößt gegen das Expressivitätskriterium.“ Da darüber hinaus der wahrgenommene Flächeninhalt eines Kreises nicht direkt mit dem objektiven Flächeninhalt übereinstimmt, schlagen Schuhmann und Müller für das Mapping der Namensvorkommen auf die Flächengröße des Kreises die Anwendung der unten aufgeführten Steven’schen psychophysischen Potenzfunktion zur Berechnung des Kreisradius aus dem Namensvorkommen x vor (vgl. Schumann und Müller 2000, S. 129-130).

$$r = \frac{x^{\frac{5}{7}}}{\sqrt{2 * \pi}}$$

Im Kapitel 6 werden Studien historischer Namensverbreitungskarten vorgestellt, die unabhängig von der Datenlage und der Software des DFD-Kartierungstools entwickelt werden. Hierdurch können einige einschränkende Aspekte, die aus der Abhängigkeit von der Datenbasis und der Software des DFD-Kartierungstools resultieren, neu angegangen und überwunden werden. Außerdem können neue Akzente im Hinblick auf Visualisierungstechniken und Interaktivität gesetzt werden. Neben einer Kartenstudie, in der die Begrenzungen einzelner Postleitzahlenbereiche auf Wunsch interaktiv eingeblendet und damit die Geltungsbereiche für die kodierten Datenwerte visualisiert werden können, wird in allen im Kapitel 6 vorgestellten Studien historischer Verbreitungskarten mit Kreisen die oben genannte Formel zur Berechnung der Kreisradien eingesetzt. Die Umstellung auf die Steven’sche psychophysische Potenzfunktion ist hier ein Beispiel dafür, wie eine Möglichkeit der visuellen Kodierung so genutzt werden kann, dass diese sich den menschlichen Fähigkeiten anpasst und auf diese Art der Visualisierungsaufgabe gerecht wird (vgl. Munzner 2014, S. 13). Erste Rückmeldungen aus dem Kreis des Mitarbeiterteams des DFD bestätigen den positiven Effekt der veränderten Größendarstellung und Proportionalität der Kreise, der sich vor allem in der verbesserten Wahrnehmung von Größenunterschieden insbesondere bei sehr kleinen und sehr großen Kreisen zeigt. Weitere Überlegungen zur Problematik der Größenberechnung der Kreise finden sich im Kapitel 7.

6 Interaktive historische Namensverbreitungskarten für das DFD

Für die visuelle Präsentation der Namensverbreitungen im DFD ist neben der möglichst genauen Veranschaulichung der aktuellen Verbreitungen individueller Familiennamen auf der Grundlage der oben beschriebenen Datenmengen insbesondere die Darstellung historischer Namensverbreitungen vor 1945 wertvoll. Sie ist nicht nur für die Nutzer*innen des DFD-Frontends interessant. Historische Namensverbreitungskarten unterstützen die Artikelautor*innen auch bei der Namendeutung im DFD. Sie können Auskunft geben über die Herkunft eines Familiennamens, sie können zu einer präzisen Namendeutung und Sprachzuordnung beitragen und sie zeigen Namenwanderungen auf, die vor allem nach dem Zweiten Weltkrieg stattgefunden haben. Aus diesem Grund werden im Rahmen eines weiteren Softwareprojekts erste Studien interaktiver historischer Verbreitungskarten entwickelt, in denen Namensverbreitungen zur Zeit des Deutschen Kaiserreichs im Jahr 1890 und Namensverbreitungen in der Bundesrepublik Deutschland im Jahr 1996 visualisiert werden.¹ Die historischen Karten sollen nach erfolgter Integration des interaktiven Kartenmoduls in die Website des DFD ebenfalls in das Modul integriert werden. Erste Überlegungen hierzu sind bereits erfolgt. In der Zwischenzeit wurden einige Studien interaktiver historischer Karten verschiedenen Nutzer*innen des DFD, u.a. dem Team der Artikelautor*innen, vorgestellt. In der Folge wurden insbesondere von der Benutzergruppe der Autor*innen eine Reihe weiterer während der Beschäftigung mit den Namensverbreitungskarten beider Softwareprojekte aufgekommener Anforderungen an die Visualisierungen geäußert.

6.1 Erweiterter Aufgabenkatalog für das Design der Visualisierungssidiome

Für die Aufstellung der Erfordernisse, Bearbeitungsziele und Fragen der Nutzer*innen, die für die Entwicklung der Visualisierungssidiome der interaktiven Verbreitungskarten des Kartenmoduls insgesamt maßgeblich sind, zeichnet sich die mögliche Unterscheidung zweier verschiedener Benutzergruppen für das DFD ab. Die Gruppe der Nutzer*innen ohne oder

¹In den in dieser Arbeit beschriebenen Vorstudien für das Softwareprojekt der historischen Kartierung für das DFD sind neben diversen Verbreitungskarten mit historischen Namensvorkommen auch zwei Kartenstudien ausschließlich mit (nicht historischen) Daten aus dem Jahr 1996 sowie eine datenunabhängige Karte enthalten. Da alle diese Karten auf der gleichen Datengrundlage entstanden sind, werden sie im Folgenden unter dem Begriff „historische Verbreitungskarten“ zusammengefasst.

mit wenig fachlicher Expertise, die im Wesentlichen die Benutzergruppe des interaktiven Kartenmoduls im DFD-Frontend stellt, beschäftigt sich aus bestimmten privaten Interessen mit den Karten. Diese Nutzer*innen erwarten von der Visualisierung ausschließlich die leicht zugängliche Präsentation der Verbreitung eines Familiennamens in Deutschland. Sie möchten auf einen Blick sehen, ob ein gesuchter Name hochfrequent oder selten ist, in welchen Regionen er besonders häufig vorkommt, ob er an bestimmten Orten überhaupt nicht vertreten ist etc. Für die Entwickler*in der Visualisierung heißt das, dass die Informationen, die in den Daten enthalten sind, möglichst vollständig und intuitiv abrufbar kommuniziert werden sollten. Ein zweites Benutzerprofil umfasst die Nutzer*innen mit Expertenwissen, zu denen auch das Mitarbeiterteam und die Artikelautor*innen des DFD zählen. Für sie ist nicht nur die Präsentation konkreter, detaillierter Informationen rund um die Namensvorkommen im geographischen Kontext relevant sondern auch die Möglichkeit, diese Informationen weiter explorieren und neue Information generieren zu können. Sie möchten vielleicht eigene Hypothesen zu aktuellen Namensverbreitungen oder historischen Namenverschiebungen aufstellen und diese verifizieren oder widerlegen. Sie interessieren sich für Abhängigkeiten und Korrelationen innerhalb der visualisierten Daten. Für sie steht nicht nur die Arbeit mit den Karten im DFD-Frontend im Vordergrund, sondern auch alle Fragen rund um die Kartierung im DFD-Backend. Die in der Liste unten aufgeführten zusätzlichen Anforderungen der Nutzer*innen an die Visualisierungen sowohl im interaktiven Kartenmodul als auch in der historischen Kartierung stellen eine Sammlung von potenziellen Aufgaben für die weitere Softwareentwicklung dar, die sich zum Einen aus diversen Gesprächen mit der Benutzergruppe der Mitarbeiter*innen des DFD-Projekts, insbesondere der Artikelautor*innen und zum Anderen aus ersten Überlegungen zu den möglichen Personas, die für das Design speziell der Visualisierungsideome historischer Namensverbreitungskarten maßgeblich sein könnten, ergeben.²

Wie im folgenden Abschnitt ausführlich beschrieben wird, unterscheidet sich die Datenlage für das Design der Visualisierungsideome der Studien historischer Verbreitungskarten grundlegend von der des interaktiven Kartenmoduls. Wie oben bereits erwähnt, erfolgt die Softwareentwicklung für die historische Kartierung vollkommen unabhängig von der Datenbasis und der JavaScript-Bibliothek des DFD-Kartierungstools. Daraus und aus den erweiterten Abfragemöglichkeiten der eigens für die Entwicklung der Studien historischer Verbreitungskarten eingerichteten Datenbank ergeben sich neue Möglichkeiten der Visualisierung der Namensverbreitungen, die auch spezialisiertere Anforderungen erfüllen können. An den unten aufgeführten Aufgaben für die weitere Entwicklung der Visualisierungs- und Interaktionsideome der Verbreitungskarten für das DFD zeichnet sich derzeit eine Verschiebung des Benutzerprofils von den Nutzer*innen ohne fachliche Expertise hin zu wissenschaftlich mit Namensverbreitungen und historisch belegten Namenverschiebungen arbeitenden Nutzer*innen ab.

H1 Spontane Rezeption historischer Namensverbreitungen und Namenverschiebungen

H2 Anzeige der exakten Zahlenwerte der Namensvorkommen

H3 Anzeige der Geltungsbereiche für die Datenwerte

²Aufgrund der Pandemiesituation während der Entstehung dieser Arbeit war die eingehendere persönliche Beobachtung und Befragung von Vertretern der verschiedenen Benutzergruppen leider nicht möglich.

-
-
- H4** Überblenden der Verbreitungen von mehr als einem Namen, z.B. verwandte Namen oder Namen aus Verweis- und zugehörigen Hauptartikeln bzw. Überblenden der Verbreitungen eines Namens in verschiedenen Zeiträumen
 - H5** Speicherung der Suchhistorie
 - H6** Anzeige von Sprach- und Dialektgebieten
 - H7** Erweiterung der interaktiven Anpassungsmöglichkeiten der Visualisierung, z.B. für die individuelle Erstellung, den Druck oder die Veröffentlichung einer Karte
 - H8** Interaktive Änderung der Größe aller Kreise insgesamt
 - H9** Für das menschliche Auge effektiv unterscheidbare Größenunterschiede der Kreise
 - H10** Clustering von Namensvorkommen in Karten mit großen Anzahlen überlagerter Kreise
 - H11** Zu der Darstellung mit Kreisen alternative Darstellung der Namensverbreitungen
 - H12** Eigene individuell festgelegte Markierung ausgesuchter Regionen in einer Karte
 - H13** Suche über einen Ort oder eine Region anstatt über einen Namen
 - H14** Kartierung von Berufsnamen, Wohnstättennamen u.ä.
 - H15** 'Pro'-Version aller Karten für DFD-Mitarbeiter*innen und für Wissenschaftler*innen mit legitimem Forschungsinteresse, u.a. mit der Freigabe und Anzeige konkreter Zahlen zu den Namensvorkommen
 - H16** Erklärende Informationstexte zur Datenlage, zu Default-Parametereinstellungen und zur Anwendung selbst
 - H17** Eigens für das DFD erstellte Tiles für geographische Hintergrundkarten
 - H18** Ergänzung von erklärenden Graphiken und Diagrammen zur Statistik der Namensvorkommen
 - H19** Implementierung einer Teilen-Funktion

Einige der oben aufgeführten Aufgaben für die Softwareentwicklung wie z.B. die Anforderung H2 mit der Frage nach exakten Zahlenwerten der Namensvorkommen, die in der Anforderung H3 festgehaltene Angabe der Geltungsbereiche für Datenwerte, die in der Anforderung H9 angefragte Anwendung der wahrnehmungskonformen Größenberechnung der Kreise oder auch die Anforderung H11 mit dem Wunsch nach einer alternativen Darstellung der Namensvorkommen spiegeln die Reaktionen von Nutzer*innen auf die aufgrund von bestehenden Lizenzen und der Abhängigkeit vom DFD-Kartierungstool eingeschränkten Möglichkeiten des interaktiven Kartenmoduls wider. Bei der Darstellung und Analyse der Visualisierungsideome für die Studien historischer Verbreitungskarten in den folgenden Abschnitten liegt der Fokus u.a. auch auf der Demonstration von Lösungsansätzen und Visualisierungsdetails, die diese Anforderungen durch den Einsatz aktueller Techniken der Informationsvisualisierung teilweise bereits erfüllen können.

6.2 Art der verwendeten Daten

Wie die Verbreitungskarten im interaktiven Kartenmodul werden auch die Studien historischer Verbreitungskarten mithilfe der JavaScript-Bibliothek Leaflet erzeugt. Da die entsprechende Softwareentwicklung unabhängig von der Software-Bibliothek und der Datenbasis des DFD-Kartierungstools erfolgt, wird der für die Erzeugung der Layer über den ausgewählten Hintergrundkarten des Leaflet Tile Providers Carto erforderliche JavaScript-Code neu entwickelt. Die der historischen Kartierung zugrunde liegenden Daten stammen ursprünglich aus deutschen Verlustlisten des Ersten Weltkriegs. Diese Verlustlisten wurden im Rahmen eines Crowdsourcing-Projekts vom Verein für Computergenealogie erfasst, indiziert und aufbereitet.³ Die entsprechenden digitalen Daten enthalten rund acht Millionen Einträge und werden der Öffentlichkeit kostenfrei übers Web zur Verfügung gestellt (vgl. Zedlitz 2021). Sie stellen die umfangreichste und flächendeckendste namenkundliche Datenbasis dar, die derzeit für die historische Kartierung zur Verfügung steht (vgl. Flores und Gilles 2020). In den Verlustlisten finden sich die Familiennamen und Heimatorte von im Ersten Weltkrieg gefallenen, verwundeten, in Gefangenschaft geratenen oder auf andere Art zu Schaden gekommenen Soldaten. Das errechnete mittlere Geburtsdatum der Soldaten liegt im Jahr 1890. Die Heimatorte der Soldaten wurden im Rahmen der Datenaufbereitung georeferenziert. Neben den aufbereiteten Daten der Verlustlisten stellt der Verein für Computergenealogie außerdem aus den Rohdaten einer Telefonbuch-CD der Deutschen Telekom aufbereitete Daten zu den Namensvorkommen in Deutschland im Jahr 1996 zur Verfügung. Die unten aufgeführten Listings 6.1 und 6.2 zeigen jeweils Auszüge aus den entsprechenden Dateien des Vereins für Computergenealogie.

Listing 6.1: Auszug aus der Datei `familyname2coordinate_1890`

```
name longitude latitude
Aa 6.823 52.035
Aa 7.110 52.845
Aa 7.110 52.845
Aa 7.110 52.845
Aa 8.578 53.544
Aa 8.578 53.544
Aab 7.2 51.27
Aab 7.2 51.27
Aab 7.2 51.27
Aab 8.410 49.002
...
```

Listing 6.2: Auszug aus der Datei `familyname2coordinate_1996`

```
name longitude latitude
...
Abada 13.4454518 52.45362659999999
Abada 13.4454518 52.45362659999999
Abaday 13.374087838803598 52.5186437505076
Abaday 13.40535 52.52468
Abadayev 13.37879 52.48545333333333
Abadi 13.424756000000002 52.495422999999995
```

³<https://nvk.genealogy.net>

```
Abadi 13.339609621693933 52.50165113061706
Abadi 13.431545666666667 52.51400433333333
Abadi 13.400666666666666 52.45420333333333
...
```

In den Datensätzen der oben aufgeführten Tabellen wird jedes einzelne Vorkommen eines Namens zusammen mit den Geodaten des Geburtsortes des Namenträgers festgehalten. Das führt dazu, dass oftmals ein und derselbe Familienname am gleichen Ort mehrfach aufgeführt wird. Um die in den Daten enthaltenen Informationen über die Summe der Vorkommen eines individuellen Namens an jeweils einem Ort zu erhalten, werden die Daten in eine SQLite Datenbank geschrieben und mithilfe des Python-Skripts `db.py` für alle für die historischen Kartenstudien ausgewählten Namen entsprechend abgefragt. Zur Einbindung in den JavaScript-Code der Leaflet-Karten werden anschließend für jeden Namen mithilfe des Python-Skripts `name.py` jeweils zwei sogenannte Feature Collections, d.h. speziell in einzelne Dictionaries mit Key-Value-Paaren strukturierte GeoJSON-Dateien, mit den jeweils aufsummierten Namensvorkommen für das Jahr 1890 bzw. für das Jahr 1996 geschrieben. Das Listing 6.3 zeigt einen Ausschnitt aus der Feature Collection für den Namen Nowak im Jahr 1890.

Listing 6.3: Auszug aus der Datei `Nowak_1890.js`

```
var dataNowak1890 = [
  {
    "features": [
      {
        "geometry": {
          "coordinates": [
            50.034,
            17.88
          ],
          "type": "Point"
        },
        "popupContent": "Nowak: 102 Vorkommen",
        "properties": {
          "count": 102
        },
        "type": "Feature"
      },
      {
        "geometry": {
          "coordinates": [
            52.341,
            17.122
          ],
          "type": "Point"
        },
        "popupContent": "Nowak: 100 Vorkommen",
        "properties": {
          "count": 100
        },
        "type": "Feature"
      },
      ...
    ]
  },
  ...
]
```

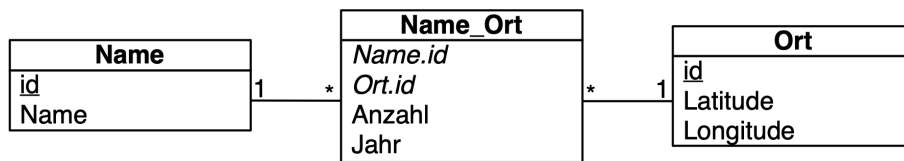


Abbildung 6.1: Entitäten und Attribute in den aus den Verlustlisten gewonnenen Daten

Die Feature Collections werden jeweils in Form von JavaScript-Dateien abgelegt (s. Dateien `Nowak_1890.js`, `Nowak_1996.js`, `Luedtke_1890.js` etc.). Die so aufbereiteten Daten bilden die Datengrundlage für die Erzeugung aller in den Abschnitten 6.3.1 und 6.3.4 beschriebenen interaktiven historischen Kartenstudien.

Das Datenmodell für die beschriebenen Daten und ihre Beziehungen untereinander wird in dem Entity-Relation-Diagramm in der Abbildung 6.1 dargestellt. Wie im Entity-Relation-Diagramm abzulesen ist, sind der Entität „Name“ das Attribut „Name“, der Entität „Name_Ort“ die Attribute „Anzahl“ und „Jahr“ und der Entität „Ort“ die Attribute „Latitude“ und „Longitude“ zugeordnet. Dementsprechend sind in der Feature Collection im Listing 6.3 die Datenwerte der Schlüssel „coordinates“ und „count“ und die Angaben zum „Namen“ und zum „Jahr“ relevant für die Abbildung der Daten auf graphische Elemente mit visuellen Attributen. Die Datenwerte des Schlüssels „coordinates“ bestehen jeweils aus einem Tupel mit den beiden Komponenten „Longitude“ und „Latitude“. Diese Komponenten sind die Längen- und Breitengrade der Heimatorte der Soldaten im Jahr 1890 bzw. der Wohnorte der entsprechenden Namensinhaber im Jahr 1996. Die Geokoordinaten werden für die für Leaflet-Karten passende Mercator-Projektion berechnet und liegen folglich als *quantitative* Daten vor. Die Datenwerte des Schlüssels „count“ bilden ein sogenanntes abgeleitetes Attribut, das durch das Aufaddieren der in den Originaldaten enthaltenen Anzahlen der Vorkommen eines individuellen Familiennamens an jeweils einem Ort entsteht. Diese Daten sind ebenfalls *quantitativ*. Die jeweiligen Angaben zu den individuellen Namen und zum Jahr der Verbreitung, z.Zt. im Dateinamen festgehalten, werden als *kategoriale* Daten unter den Bezeichnungen „name“ und „year“ eingeordnet. Auf die Abbildung dieser Daten auf graphische Elemente mit visuellen Attributen, die mit dem Prinzip der Expressivität der Visualisierung im Einklang steht, wird im folgenden Abschnitt näher eingegangen.

Für zwei der in den nächsten Abschnitten vorgestellten Kartenstudien, der Namensverbreitungskarte mit Kreisen und einblendbaren Postleitzahlenbereichen im Abschnitt 6.3.2 und der Choropleth-Karte im Abschnitt 6.3.3, werden die Summen der Vorkommen des Namens „Richter“ in den deutschen Postleitzahlenbereichen im Jahr 1996 benötigt. Die entsprechenden Daten werden von Amaru Flores Flores, einem Mitarbeiter der Akademie der Wissenschaften und der Literatur in Mainz, für die Softwareentwicklung der historischen Kartierung im DFD zur Verfügung gestellt.⁴ Es handelt sich hierbei um einen PostgreSQL Datenbank-Dump von aus den Verlustlisten aufbereiteten Daten. Der Dump wird in eine lokale PostgreSQL Datenbank eingelesen und die Daten werden entsprechend der Visualisierungsaufgabe abgefragt. Das Listing 6.4 zeigt das Ergebnis der Abfrage für den Familiennamen Richter.

⁴Vielen Dank an Amaru Flores Flores für die Bereitstellung der Daten.

Listing 6.4: Auszug aus dem Ergebnis der PostgreSQL-Datenbankabfrage für den Namen Richter

```
name,anzahl1996,namentraeger1996,plz,jahr,anzahl1890,namentraeger1890
...
...
Richter,958,99820,010,1996,4,1432
Richter,332,33320,011,1996,1,8210
Richter,133,14216,013,1996,15,42384
Richter,225,19854,014,1996,1,7011
Richter,234,18948,015,1996,2,9069
Richter,311,28925,016,1996,4,24088
...
...
```

Für die Visualisierungsaufgabe in der Kartenstudie Namensverbreitungskarte mit Kreisen und einblendbaren Postleitzahlenbereichen werden zusätzlich zu den in der Spalte „plz“ enthaltenen Zahlencodes für die Postleitzahlenbereiche auch die Geokoordinaten der Mittelpunkte dieser Postleitzahlenbereiche gebraucht. Die Daten mit den Angaben zu den Längen- und Breitengraden dieser Mittelpunkte sind im Web in der Mercator-Projektion und im GeoJSON-Format frei verfügbar.⁵ Sie werden heruntergeladen und als JavaScript-Datei abgelegt. Mithilfe des Python-Skripts `add_properties_centroids.py` werden die Daten aus der Datei `Richter1996_plz3.csv` in eine Feature Collection geschrieben und die Geokoordinaten der Mittelpunkte werden hinzugefügt. Bei diesen Geokoordinaten-Daten handelt es sich wieder um *quantitative* Daten. Das Listing 6.5 zeigt einen Ausschnitt aus der Feature Collection `Richter1996_centroids.js`.

Listing 6.5: Auszug aus der Datei `Richter1996_centroids.js`

```
var richterCentroids = {
  "features": [
    {
      "geometry": {
        "coordinates": [
          51.08973191304776,
          13.808835680236237
        ],
        "type": "Point"
      },
      "popupContent": "PLZ-Gebiet 010: 958 Vorkommen",
      "properties": {
        "count1996": "958",
        "einwohner": 80263,
        "name": "Richter",
        "plz": "010",
        "qkm": 74.02
      },
      "type": "Feature"
    },
    {
      "geometry": {
        "coordinates": [
```

⁵<https://www.suche-postleitzahl.org>, zuletzt abgerufen am 27.08.2021

```

        51.06280134077906,
        13.646456930734022
    ],
    "type": "Point"
},
"popupContent": "PLZ-Gebiet 011: 332 Vorkommen",
"properties": {
    "count1996": "332",
    "einwohner": 182367,
    "name": "Richter",
    "plz": "011",
    "qkm": 121.25
},
"type": "Feature"
},
...
...

```

Für die Abbildung der Daten auf graphische Elemente mit visuellen Attributen in der Karte relevant sind unmittelbar die Datenwerte der Schlüssel „coordinates“ mit den entsprechenden Tupeln aus Längen- und Breitengraden und „count1996“ mit den Namensvorkommen sowie mittelbar die Datenwerte des Schlüssels „plz“ mit den Zahlencodes der Postleitzahlenbereiche. Es handelt sich hierbei durchweg um *quantitative* Daten.

Die zur Darstellung der Begrenzungen der einblendbaren Postleitzahlenbereiche benötigten Daten sind ebenfalls unter dem oben genannten Link im Web frei verfügbar. Die entsprechenden Feature Collections enthalten Sammlungen von Listen von Geokoordinatenpaaren in der Mercator-Projektion für alle dreistelligen deutschen Postleitzahlenbereiche. Da für die Erzeugung der im Abschnitt 6.3.3 vorgestellten Choropleth Karte die gleichen Daten, jedoch angereichert mit den oben genannten Daten für den Namen Richter, benötigt werden, werden den Geodaten in der Feature Collection mithilfe des Python-Scripts `add_properties.py` die entsprechenden Schlüssel und Datenwerte hinzugefügt. Bei den Geokoordinaten in der Feature Collection, d.h. den Datenwerten des Schlüssels „coordinates“, handelt es sich wieder um *quantitative* Daten. Das Listing 6.6 zeigt die resultierende Feature Collection in der Datei `Richter_plz3.js`.

Listing 6.6: Auszug aus der Datei `Richter_plz3.js`

```

var RichterPlz3 = {
  "features": [
    {
      "geometry": {
        "coordinates": [
          [
            [
              13.730314,
              51.0217019
            ],
            [
              13.7298675,
              51.0207366
            ]
          ]
        ]
      }
    }
  ]
}

```

```

    [
      13.72964,
      51.0202346
    ],
    ...
    ...
    [
      13.8321743,
      51.1140021
    ],
    [
      13.827665000000001,
      51.1168476
    ],
    [
      13.8247763,
      51.114297
    ]
  ]
],
"type": "MultiPolygon"
},
"properties": {
  "count1996": "958",
  "einwohner": 80280,
  "name": "Richter",
  "plz": "010",
  "qkm": 74
},
...
...

```

Neben den namenbezogenen Daten werden auch für die Entwicklung der Visualisierungsidiome der historischen Kartenstudien die *geometrischen* Daten der geographischen Hintergrundkarten und der mit ihnen verbundene Datentyp *Position* eingesetzt. Für einige der Kartenstudien werden die geometrischen Daten durch einen zusätzlichen Layer ergänzt. In diesem Layer werden die Grenzen des Deutschen Kaiserreichs von 1871-1918 dargestellt. Die hierzu benötigten Daten im GeoJSON-Format sind ebenfalls im Web frei verfügbar.⁶ Sie enthalten im Wesentlichen zu geordneten Listen zusammengefasste Geokoordinaten in der Mercator-Projektion. Bei diesen Daten handelt es sich wieder um *quantitative* Daten. Auch auf dieses für die Kartenstudien zusätzlich entwickelte Visualisierungsdetail wird im folgenden Abschnitt näher eingegangen. Wie für das interaktive Kartenmodul liegen auch für die historische Kartierung ausschließlich *statische* Datenmengen vor.

⁶<https://maps.princeton.edu/catalog/harvard-ghgis1914germanempire>, zuletzt abgerufen am 03.10.2021

6.3 Visuelle Kodierung der Daten

Wie für die Verbreitungskarten im interaktiven Kartenmodul gilt auch für die Studien interaktiver historischer Verbreitungskarten, dass sich die visuelle Kodierung der Daten und die Interaktivität in den Karten an den Anforderungen und Fragen orientieren, die die Betrachter*innen an die Visualisierung stellen. Die visuellen Repräsentationen und die z.T. umfangreichen Möglichkeiten der Benutzerinteraktionen sollen auch hier möglichst vollständige Informationen über die Zusammenhänge zwischen Namensvorkommen und Positionen und über die Proportionalität zwischen den zugrunde liegenden Zahlenwerten liefern. Das Mapping der Werte in den zugrunde liegenden Datenmengen auf passende visuelle Attribute für die graphischen Elemente erfolgt wieder nach den Prinzipien der Expressivität und Effektivität und orientiert sich wie das Mapping im interaktiven Kartenmodul an dem im Abschnitt 4.4 beschriebenen Mapping-Modell für die Verbreitungskarten im DFD. Um das Design der Visualisierungsideome so gut wie möglich entlang der Bedürfnisse der Wörterbuchnutzer*innen zu entwickeln, wurden auch hier die verschiedenen Designvorschläge und Prototypen der interaktiven Karten während der Softwareentwicklung in regelmäßigen Abständen mit den Mitarbeiter*innen des DFD-Projekts besprochen.⁷

Obwohl die Verlustlisten nur etwa 15 Prozent der damaligen Gesamtbevölkerung repräsentieren und nur ein geringer Prozentsatz der derzeitigen deutschen Bevölkerung einen der in den Verlustlisten vorkommenden Familiennamen tragen, sind die Daten der Namensvorkommen im Jahr 1890 und deren visuelle Repräsentationen in den geographischen Leaflet-Karten ausreichend um im Vergleich mit den visualisierten Daten des Jahres 1996 Tendenzen in Bezug auf Namenverschiebungen aufgrund von Vertreibung und Flucht im und nach dem Zweiten Weltkrieg zu veranschaulichen (vgl. Flores und Gilles 2020). In den folgenden Abschnitten werden einige Beispiele historischer Kartenstudien gezeigt, in denen die Verbreitungen eines Namens in den Jahren 1890 und 1996 und damit die entsprechenden Namenverschiebungen innerhalb ein und derselben Karte veranschaulicht werden.

6.3.1 Historische Namensverbreitungskarten mit Kreisen

Den historischen Namensverbreitungskarten mit Kreisen liegen u.a. die Daten in den wie oben für den Namen Nowak beschriebenen Dateien mit den Datenwerten der Schlüssel „count“ und „coordinates“ sowie den Informationen zum Namen und zum Zeitraum der Verbreitung „name“ und „year“ zugrunde. Im interaktiven Kartenmodul wird mithilfe des vom DFD-Kartierungstool abgeleiteten Skripts `circles.js` z.Zt. nur jeweils die Verbreitung eines einzelnen Namens im Jahr 2005 kartiert. Die Datenlage und die unabhängig entwickelte Software für die historische Kartierung lässt dagegen die Kartierung mehrerer Namen, ggf. auch in verschiedenen Zeiträumen, innerhalb ein und derselben Karte zu. In dem in der Abbildung 6.2 gezeigten ersten Beispiel der Studie einer historischen Verbreitungskarte

⁷Persönliche Gespräche mit weiteren Nutzer*innen außerhalb des Projekts und/oder Beobachtungen im Sinne einfacher Usability Tests nach der Methode von Steve Krug (vgl. Krug 2014, Kap. 9) waren während des Entwicklungszeitraums aufgrund der andauernden Coronapandemie leider nicht möglich.

Nowak 1890 und 1996

Nowak 1890 (rot): 4364 Vorkommen an 1499 Orten, Max 102, Min 1; Nowak 1996 (grün): 12006 Vorkommen an 3386 Orten, Max 103, Min 1

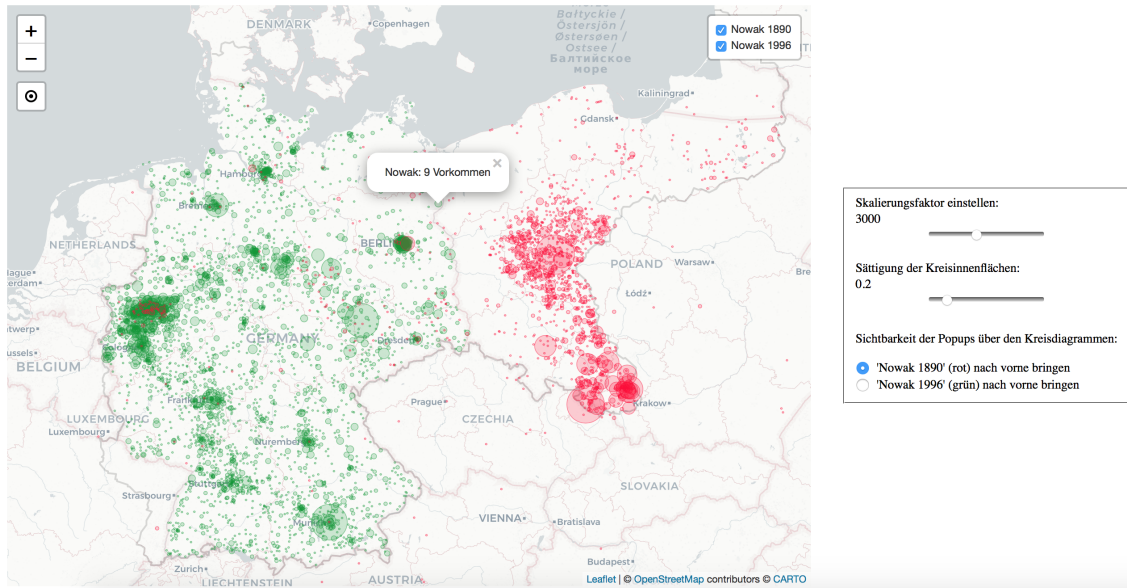


Abbildung 6.2: Interaktive historische Verbreitungskarte mit Kreisen für den Namen Nowak in den Jahren 1890 und 1996

für den Namen Nowak wird die Verbreitung des Namens in den Jahren 1890 und 1996 innerhalb einer Karte dargestellt und wie im interaktiven Kartenmodul mithilfe von Kreisen veranschaulicht. Zudem werden, da die zugrunde liegenden Daten hier keinem Lizenzrecht unterliegen, die in den Datenwerten von „count“ enthaltenen konkreten Zahlenwerte in den Informationstexten über den Kreisen angezeigt.

Das Mapping der quantitativen Daten aus den Schlüsseln „count“ und „coordinates“ auf graphische Elemente mit visuellen Attributen erfolgt in dieser und in allen weiteren im vorliegenden Abschnitt vorgestellten Kartenstudien mit Kreisen im Wesentlichen so wie im Abschnitt 4.5.2 für das interaktive Kartenmodul beschrieben. Im Einklang mit den Expressivitäts- und Effektivitätsprinzipien wird dem graphischen Element *Punkt* zunächst das visuelle Attribut *Position* zugeordnet. Die Punkte werden in der Karte an den durch die entsprechenden Geokoordinaten festgelegten Orten positioniert. Neben der Position ist das zweite visuelle Attribut für die Punkte wieder die *Flächengröße*. Da dem Punkt auch hier die geometrische Form Kreis zugeordnet wird, definiert sich die Flächengröße über die Berechnung der Kreisgröße. Im Unterschied zu den im interaktiven Kartenmodul durchgeführten Berechnungen von Kreisdurchmessern wird in sämtlichen hier beschriebenen Studien historischer Verbreitungskarten mit Kreisen der Radius der Kreise nach der Steven'schen psychophysischen Potenzfunktion berechnet.⁸ Wie im Kapitel 5 erläutert wurde, orientiert sich diese Art der Größenberechnung für die Kreise in einer geographischen Karte an den visuellen Wahrnehmungsprozessen beim Menschen. Die Spezifikation der

⁸Die entsprechende Implementierung findet sich jeweils in der Code-Zeile `var radius = ...` in den `.html`-Dateien der interaktiven historischen Karten mit Kreisen.

Kartenstudie in der Abbildung 6.2 erfüllt damit bereits die im Abschnitt 6.1 aufgeführte Anforderung H9.

In der Karte in der Abbildung 6.2 werden die Verbreitungen eines ausgewählten Familiennamens in zwei verschiedenen Zeiträumen innerhalb ein und derselben Karte visualisiert. Den Betrachter*innen dieser Visualisierung muss eine Darstellung angeboten werden, die es ihnen ermöglicht, die entsprechenden Namensverbreitungen spontan zu erfassen und genau zu unterscheiden. Für das Mapping auf visuelle Attribute heißt das, dass auch der Datenwert „year“ auf ein passendes visuelles Attribut abgebildet werden muss. Nach dem Prinzip der Expressivität kommen für *kategoriale* Daten die in der Tabelle 4.2 in der rechten Spalte aufgeführten visuellen Attribute zur Unterscheidung von Identitäten in Frage. Für die vorliegende Visualisierungsaufgabe fällt die Wahl auf das im Effektivitätsrang an zweiter Stelle stehende visuelle Attribut *Farbton*. Für die beiden Datenwerte „1890“ und „1996“ werden zwei unterschiedliche, möglichst konträre Farbtöne ausgewählt. In einem sogenannten mehrdimensionalen Mapping werden diese beiden Farbtöne den zuvor bereits durch die visuellen Attribute Position und Flächengröße zu Kreisen ausgestalteten graphischen Elemente Punkt, die die Verbreitungen in den Jahren 1890 und 1996 repräsentieren, als weitere visuelle Attribute zugeordnet. Für jeden der beiden Zeiträume wird ein eigener Layer mit den entsprechend unterschiedlich gefärbten Kreisen erzeugt und über die Hintergrundkarte gelegt. Welche Farbtöne sich am besten für die Darstellung vor der geographischen Hintergrundkarte bzw. für die Unterscheidung von zwei oder mehreren Identitäten eignen, hängt von den komplexen Zusammenhängen in den Prozessen der menschlichen Farbwahrnehmung ab und kann im Detail in den Arbeiten von Colin Ware (vgl. Ware 2013, Kap. 4 und Ware 2008, Kap. 4) und Tamara Munzner (vgl. Munzner 2014, Kap. 10) oder bei Alan MacEachren und Ben Fry nachgelesen werden (vgl. MacEachren 1995 und Fry 2008, S. 39).

In dem Interaktionsfeld rechts neben der Karte kann die Farbsättigung der Kreise individuell eingestellt werden. Außerdem lässt sich hier auch die Skalierung verändern. Die allgemeine Größendarstellung der Kreise in den Layern über der Karte ist auf einer nach unten und nach oben begrenzten Skala frei wählbar. Um außerdem den ggf. unerwünschten Nebeneffekt abzufangen, dass die Kreise des einen Layers wegen des darüber liegenden anderen Kreiselayers nicht mehr anwählbar sind und damit die Informationen in den Popups nicht auf Wunsch zur Verfügung stehen, kann der betreffende Layer mit einem Klick nach oben geholt werden. Die Spezifikation dieser Kartenstudie erfüllt neben der Anforderung H9 auch die Anforderungen H1, H2, H3, H4 und H8.

In der Karte in der Abbildung 6.2 wird die Visualisierung der Namensverbreitung ein und desselben Familiennamens in zwei verschiedenen Zeiträumen demonstriert. Die Karte in der Abbildung 6.3 zeigt eine weitere historische Kartenstudie, in der die Verbreitungen mehrerer unterschiedlicher sogenannter Humanistennamen innerhalb ein und desselben Zeitraums visualisiert werden. Bei dem mehrdimensionalen Mapping für diese Darstellung werden den Datenwerten von „name“ verschiedene für diese Art der visuellen Darstellung geeignete Farbtöne zugeordnet. Für die visuelle Repräsentation der verschiedenen Namensvorkommen in dieser Beispielkarte werden außerdem bei der Datenbankabfrage verschiedene Familiennamen gruppiert und mithilfe eines Python-Skripts in eine Feature Collection geschrieben. Die Verbreitung dieser Namensgruppen wird mit den aufsummierten Vorkommen an den

Fabricius-Fabritius-Fabrizius (rot) & Pistorius-Praetorius-Prätorius-Sartorius-Sutorius (grün)

Fabricius, Fabritius, Fabrizio: 174 Vorkommen an 96 Orten, Max 8, Min 1 & Pistorius, Praetorius, Prätorius, Sartorius, Sutorius: 415 Vorkommen an 209 Orten, Max 16, Min 1

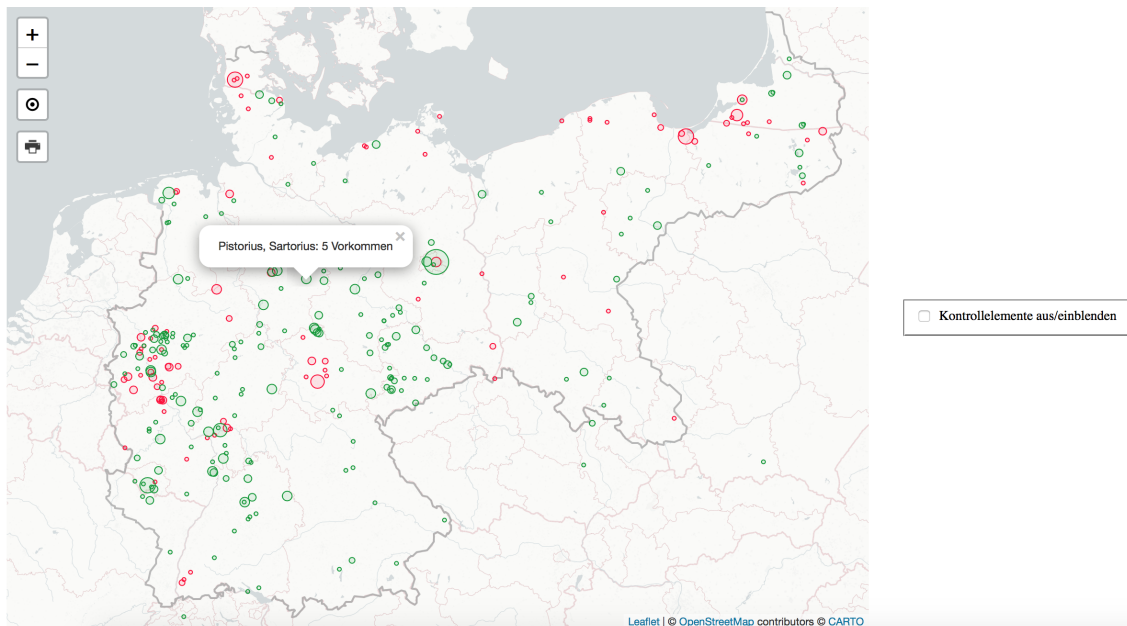


Abbildung 6.3: Interaktive historische Verbreitungskarte mit Kreisen für die Namensgruppen Fabricius-Fabritius-Fabrizius und Pistorius-Praetorius-Prätorius-Sartorius-Sutorius im Jahr 1890

jeweiligen Orten visualisiert (s. `name_group.py` und `Fabric-t-z-ius1890.js` bzw. `5-x-torius1890.js`). Die Spezifikation dieser Kartenstudie erfüllt die Anforderungen H1, H2, H3, H4 und H9.

Die Abbildung 6.4 zeigt die Studie einer historischen Verbreitungskarte, in der beide oben beschriebenen mehrdimensionalen Mappings miteinander verbunden werden. In dieser Karte erhält jede Kombination aus einem Namen und einem Zeitraum einen jeweils anderen Farbton. Die Verwendung der JavaScript-Bibliothek Leaflet geht einher mit der Entscheidung, die verschiedenen Verbreitungen eines Namens mithilfe von deckungsgleich über eine geographische Karte gelegten Layern darzustellen, die durch Benutzerinteraktionen erzeugt werden. Am Beispiel der hier gezeigten Weiterentwicklungen interaktiver Verbreitungskarten wird der Mehrwert dieser Art der visuellen Präsentation gegenüber separaten Ansichten, die neben- oder untereinander auf einer Seite präsentiert werden, deutlich. Alle Layer mit Informationen zu den Namensverbreitungen sind bis auf die Regionen, in denen die graphischen Elemente positioniert sind, transparent. Werden eine begrenzte Anzahl Layer, die jeweils durch die gezielte Auswahl visueller Attribute einer vom anderen gut unterscheidbar sind, übereinander gelegt, so werden den Betrachter*innen alle kodierte Informationen zur gleichen Zeit an einem Ort präsentiert. So sind neben der spontanen Wahrnehmung von Größenordnungen und Positionen auch direkte Vergleiche der Informationen in verschiedenen Layern möglich. Den Betrachter*innen werden signifikante Unterschiede in den Verbreitungen mehrerer Namen oder offensichtliche Namenverschie-

Naujoks - Naujokat 1890 und 1996

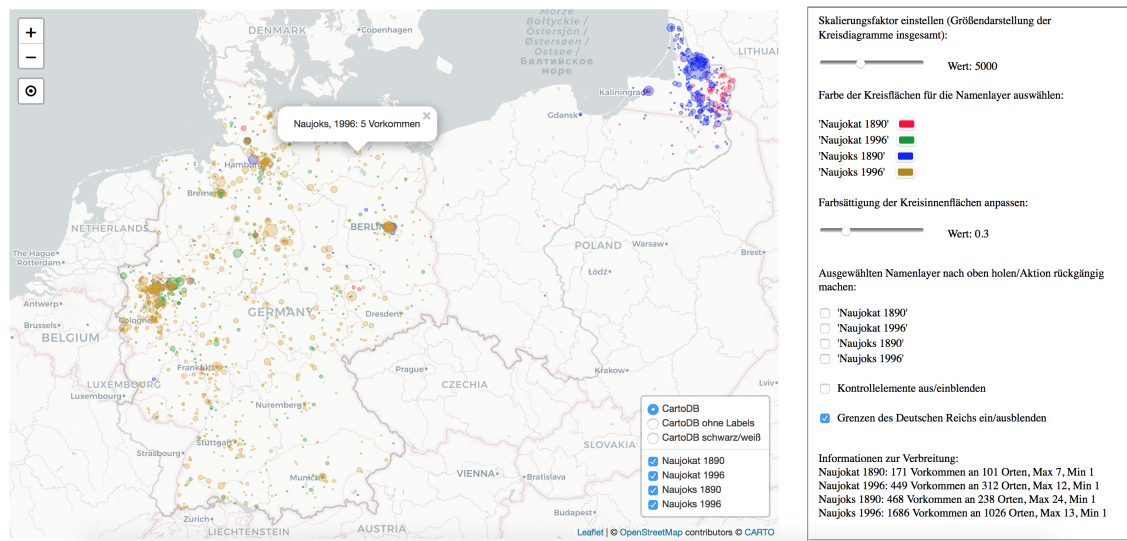


Abbildung 6.4: Interaktive historische Verbreitungskarte mit Kreisen für die Namen Naujoks und Naujokat in den Jahren 1890 und 1996

bungen vom einem dargestellten Zeitraum zum anderen direkt zugänglich gemacht. Auch das auf wenige Sekunden begrenzte Kurzzeitgedächtnis und die Aufmerksamkeitsspanne der Betrachter*innen werden so nicht überfordert, ein Kriterium, das wiederum der Effektivität der Visualisierung zugute kommt (vgl. Ware 2013, S. 22-23 und Kap. 11).

Für die Karte in der Abbildung 6.4 wurde zudem ein vergleichsweise umfangreiches Interaktionsidom entwickelt, das neben den für die Kartenstudie in der Abbildung 6.2 beschriebenen Interaktionsmöglichkeiten wie beispielsweise das nach oben Holen eines Layers auch die freie Farbwahl der Kreise für jeden einzelnen der vier Layer oder die Möglichkeit, alle Kontrollelemente innerhalb der Karte sowie die Grenzen des Deutschen Kaiserreichs ein- und ausblenden zu können, anbietet. Die verschiedenen zusätzlichen Interaktionsmöglichkeiten sind in dem Feld rechts neben der Karte zu sehen. Den Betrachter*innen werden damit diverse Möglichkeiten angeboten, die Visualisierung je nach ihren Fragen, Bedürfnissen, Bearbeitungszielen oder Verwendungszwecken individuell abzufragen bzw. zu gestalten. Die Spezifikation dieser Kartenstudie erfüllt die Anforderungen H1, H2, H3, H4, H7, H8 und H9.

In jeder der bis hierhin vorgestellten Studien historischer Verbreitungskarten mit Kreisen sowie in den im Abschnitt 6.3.4 beschriebenen sogenannten Heatmaps sind in einem z.T. interaktiv zuschaltbaren Layer die Grenzen des Deutschen Kaiserreichs von 1871-1918 zu sehen. Zur Darstellung dieser Grenzen werden, wie oben bereits erwähnt, zusätzlich zu den in den Hintergrundkarten vorhandenen geometrischen Daten weitere, in Form von Feature Collections bereitgestellte Daten eingesetzt, in denen Paare von Längen- und Breitengraden in festgelegten Reihenfolgen zusammengefasst sind (s. Datei `germanEmpire_1890.js` bzw. `reich.js`). Jedes in den Daten enthaltene Geokoordinatenpaar wird auf das graphische Element *Punkt* mit dem zugeordneten visuellen Attribut *Position* abgebildet. Für die visuelle

Repräsentation des Grenzverlaufs wird als ein weiteres graphisches Element die *Linie* eingesetzt. Das graphische Element der Linie kann gerade oder gekrümmt sein. Für das hier beschriebene Mapping bildet die gerade Linie die Verbindung zwischen jeweils zwei in den zugrunde liegenden geordneten Daten unmittelbar aufeinanderfolgenden Punkten. Damit ist die Position der Linie in der geographischen Karte festgelegt. Jede Linie hat jeweils einen festen Anfangs- und Endpunkt und eine feste Länge. Durch die Aneinanderreihung der aus den Daten hervorgehenden Linien mit gemeinsamen End- bzw. Anfangspunkten entsteht eine sogenannte Polylinie. Da weder die Breite der Linie noch ihre Farbe, Farbhelligkeit oder Farbsättigung für die Darstellung des Grenzverlaufs informationstragend sind, kann die Polylinie durch entsprechende visuelle Attribute passend zum Hintergrund und der gewünschten visuellen Aussage ausgestaltet werden (vgl. Ware 2013 und Munzner 2014). Durch die visuelle Repräsentation der Daten für den Grenzverlauf des deutschen Kaiserreichs lassen sich die Anordnungen der Kreise für die Betrachter*innen in Bezug setzen zu den damaligen geographischen Begebenheiten. Die Hinzunahme dieses zusätzlichen Layers kommt daher der Expressivität der Visualisierung zugute.

6.3.2 Namensverbreitungskarte mit Kreisen und einblendbaren Postleitzahlenbereichen

Bei der Evaluation des Visualisierungsideoms des interaktiven Kartenmoduls wurden die von dem DFD-Kartierungstool übernommenen Berechnungen der Kreisgrößen sowie die fehlende Angabe der jeweiligen Geltungsbereiche für die Datenwerte als Kritikpunkte beschrieben. Mit den bis hierhin gezeigten Studien historischer Karten liegen bereits mehrere Beispiele vor, in denen die Berechnung der Kreisgrößen im Einklang mit den aktuellen Erkenntnissen der Informationsvisualisierung nach der Steven'schen psychophysischen Potenzfunktion mit den beschriebenen Vorteilen für die Betrachter*innen erfolgt. Auch die Angabe der Geltungsbereiche für die zugrunde liegenden Datenwerte stellt in diesen Karten kein Problem dar. Der globale Geltungsbereich der Gesamtheit der visuell kodierten Daten wird in den historischen Karten durch die Überblendung der Grenzen des Deutschen Kaiserreichs von 1871 bis 1918 angegeben. Auch die regionalen Geltungsbereiche sind unmittelbar erkennbar, da in den aufbereiteten Verlustlisten die Summen der Namensvorkommen jeweils für einen durch ein entsprechendes Geokoordinatenpaar klar positionierten geographischen Ort gelten. Die gleiche Problematik der fehlenden Angaben von Geltungsbereichen, wie sie in den interaktiven Verbreitungskarten des Kartenmoduls vorliegt, besteht jedoch zunächst auch in der im Folgenden beschriebenen Namensverbreitungskarte für den Namen Richter im Jahr 1996. Hier werden, wie in den Verbreitungskarten des interaktiven Kartenmoduls, die Namensvorkommen innerhalb von dreistelligen Postleitzahlenbereichen visualisiert, wobei die verschieden großen Kreise jeweils in den Mittelpunkten dieser Regionen positioniert sind. Die vom DFD-Kartierungstool unabhängige Datenlage und Softwareentwicklung erlauben es hier jedoch, auch die Geltungsbereiche für die Datenwerte der Namensvorkommen unmittelbar visuell zu repräsentieren. Mit der in der Abbildung 6.5 gezeigten Kartenstudie liegt ein Vorschlag für die Darstellung des Geltungsbereichs der jeweils in den Kreisen kodierten Datenwerte vor, der unter bestimmten Voraussetzungen auch für die Namens-

Richter 1996

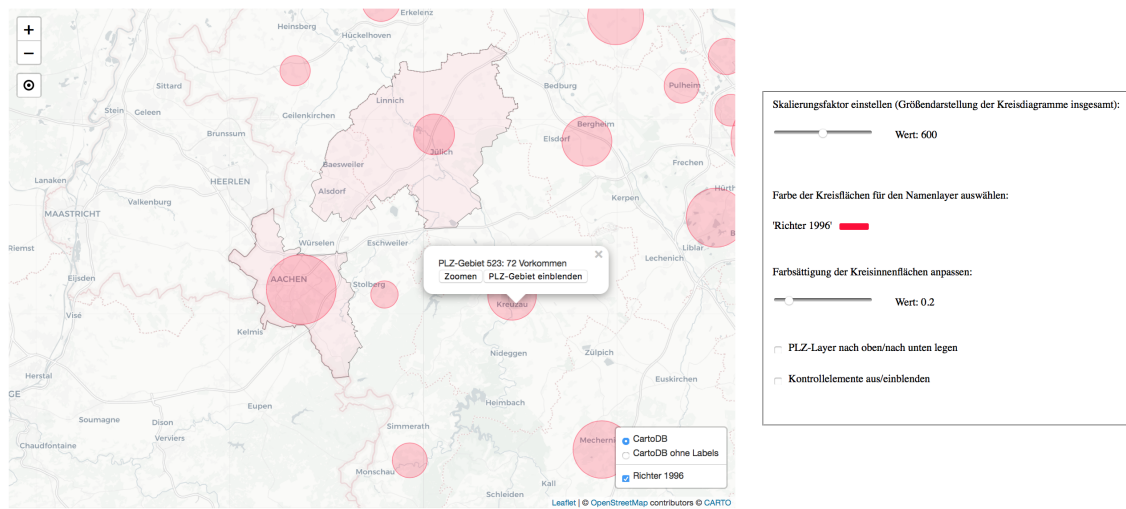


Abbildung 6.5: Interaktive Verbreitungskarte für den Namen Richter im Jahr 1996 mit Kreisen und einblendbaren Postleitzahlenbereichen

verbreitungskarten im interaktiven Kartenmodul in Betracht gezogen werden könnte.⁹ In dieser Kartenstudie können die Betrachter*innen für jeden in der Karte dargestellten Kreis interaktiv die Region in der geographischen Karte einblenden lassen, für die die jeweils durch den Kreis visuell kodierten Namensvorkommen gültig sind.¹⁰ Damit folgt die Implementierung der Interaktivität auch der Idee des Visual Information Seeking Mantras Overview-First-Details-on-Demand von Ben Shneiderman (vgl. Shneiderman 1996).

Der in der Abbildung 6.5 gezeigten Namensverbreitungskarte liegen u.a. die Daten in der Feature Collection `Richter1996_centroids.js` mit den Datenwerten der Schlüssel „name“, „count“, „plz“ und „coordinates“ zugrunde. Die visuelle Kodierung der Namensvorkommen mithilfe von Kreisen erfolgt analog zu der im vorhergehenden Abschnitt beschriebenen Auswahl visueller Attribute unter Beachtung der Expressivität und Effektivität. Die Daten für die Darstellung der Begrenzungen der dreistelligen deutschen Postleitzahlenbereiche liegen in der im Abschnitt 6.2 beschriebenen Feature Collection `Richter_plz3.js` vor (s. Listing 6.6). Jedes einzelne in der Feature Collection enthaltene Geokoordinatenpaar lässt sich wieder jeweils auf das graphische Element *Punkt* mit dem visuellen Attribut *Position* abbilden. Die einzelnen Punkte in der Karte werden wieder mithilfe des graphischen Elements *Linie* verbunden, das wiederum mit zusätzlichen visuellen Attributen ausgestaltet werden kann. Für die Darstellung der Begrenzungen eines Postleitzahlenbereichs charakteristisch ist die Aneinanderreihung der Linien zu einer Polylinie, die jeweils einen identischen Anfangs- und Endpunkt besitzt und sich so zu einem sogenannten Polygon zusammenschließt. Jedes Polygon umschließt jeweils eine geographische Region, deren Ausdehnung in der geographischen Karte durch die Positionszuordnungen zu den durch Linien verbundenen Punkten

⁹Ausführlichere Überlegungen hierzu finden sich im Kapitel 7

¹⁰Hinweis: Derzeit ist die Funktion für einige Postleitzahlenbereiche noch fehlerhaft.

festgelegt ist. In der Karte in der Abbildung 6.5 sind die Positionen der Punkte, Linien und Polygone der eingeblendeten Postleitzahlenbereiche erkennbar.

Zur vollständigen visuellen Darstellung der Postleitzahlenbereiche werden die Datenwerte des Schlüssels „plz“ auf das graphische Element *Fläche* und damit auf die Innenflächen der jeweiligen Polygone abgebildet. Auch dem graphischen Element Fläche können gestaltende visuelle Attribute zugeordnet werden. Die visuelle Gestaltung der Bestandteile des Polygons und die Ausgestaltung der Innenfläche des Polygons sind nicht informationstragend. Das graphische Element Fläche erhält hier jeweils die gleiche, interaktiv frei wählbare Farbe wie die Kreise, die in den Mittelpunkten der abgebildeten Postleitzahlenbereiche positioniert sind. Über der geographischen Hintergrundkarte können beliebig viele der dreistelligen deutschen Postleitzahlenbereiche in einem Layer dargestellt werden. Die Farbsättigung für die Fläche wird dabei so gewählt, dass die Hintergrundkarte gut lesbar bleibt und die Fläche sich von den überblendeten Kreisen abhebt. Über das Interaktionsfeld rechts neben der Karte ist u.a. wieder der einfache Austausch der Layer möglich, damit ggf. zuvor von einer Fläche verdeckte Kreise angewählt werden können. Die Spezifikation dieser Kartenstudie erfüllt die Anforderungen H1, H2, H3 und H9.

6.3.3 Choropleth-Karte

Eine von den bisherigen Darstellungen mit Kreisen grundlegend verschiedene visuelle Kodierung der zugrunde liegenden Daten inklusive der Geodaten für die dreistelligen Postleitzahlenbereiche ist in dem Visualisierungs- und Interaktionsidiom der Kartenstudie in der Abbildung 6.6 zu sehen. Diese Art der Visualisierung kartographischer Daten und der interaktive Zugang zu den gewünschten Informationen ist einem großen Teil der Bevölkerung durch die im Dashboard des Robert Koch-Instituts bereits über viele Monate fortlaufend aktualisierten COVID-19-Karte bekannt.¹¹ Die Visualisierung hat damit möglicherweise bereits einen gewissen Grad an Konvention erreicht, der auf die Interpretation und das Verständnis der dargestellten Zusammenhänge durch die Betrachter*innen durchaus Einfluss nehmen kann (vgl. Schumann und Müller 2000, S. 9).

Wie oben bereits erwähnt, liegen der hier beschriebenen Kartenstudie die Daten in der Feature Collection `Richter_plz3.js` zugrunde (s. Listing 6.6). In dieser sogenannten Choropleth-Karte spielt die visuelle Kodierung der zugrunde liegenden Daten mithilfe des visuell gestalteten graphischen Elements Fläche die zentrale Rolle. Im Unterschied zu der Kartenstudie in der Abbildung 6.5 wird die Fläche in der Choropleth-Karte nicht nur zur Darstellung des Geltungsbereichs kodierter Zahlenwerte benutzt, sondern wird unmittelbar zur visuellen Kodierung der quantitativen Daten eingesetzt. In einem Layer über der Deutschlandkarte werden zunächst die Umrisse aller mithilfe der oben beschriebenen Daten positionierten Polygone der Postleitzahlenbereiche dargestellt. Dies geschieht wieder, wie oben beschrieben, durch die Zuordnung von Positionen für die Punkte, die Hinzunahme der frei gestalteten Verbindungslinien sowie der Abbildung der Datenwerte des Schlüssels „plz“ auf das graphische Element Fläche und damit auf die Innenflächen der die Postleitzahlenbereiche begrenzenden Polygone.

¹¹<https://experience.arcgis.com/experience/478220a4c454480e823b17327b2bf1d4>

Richter 1996

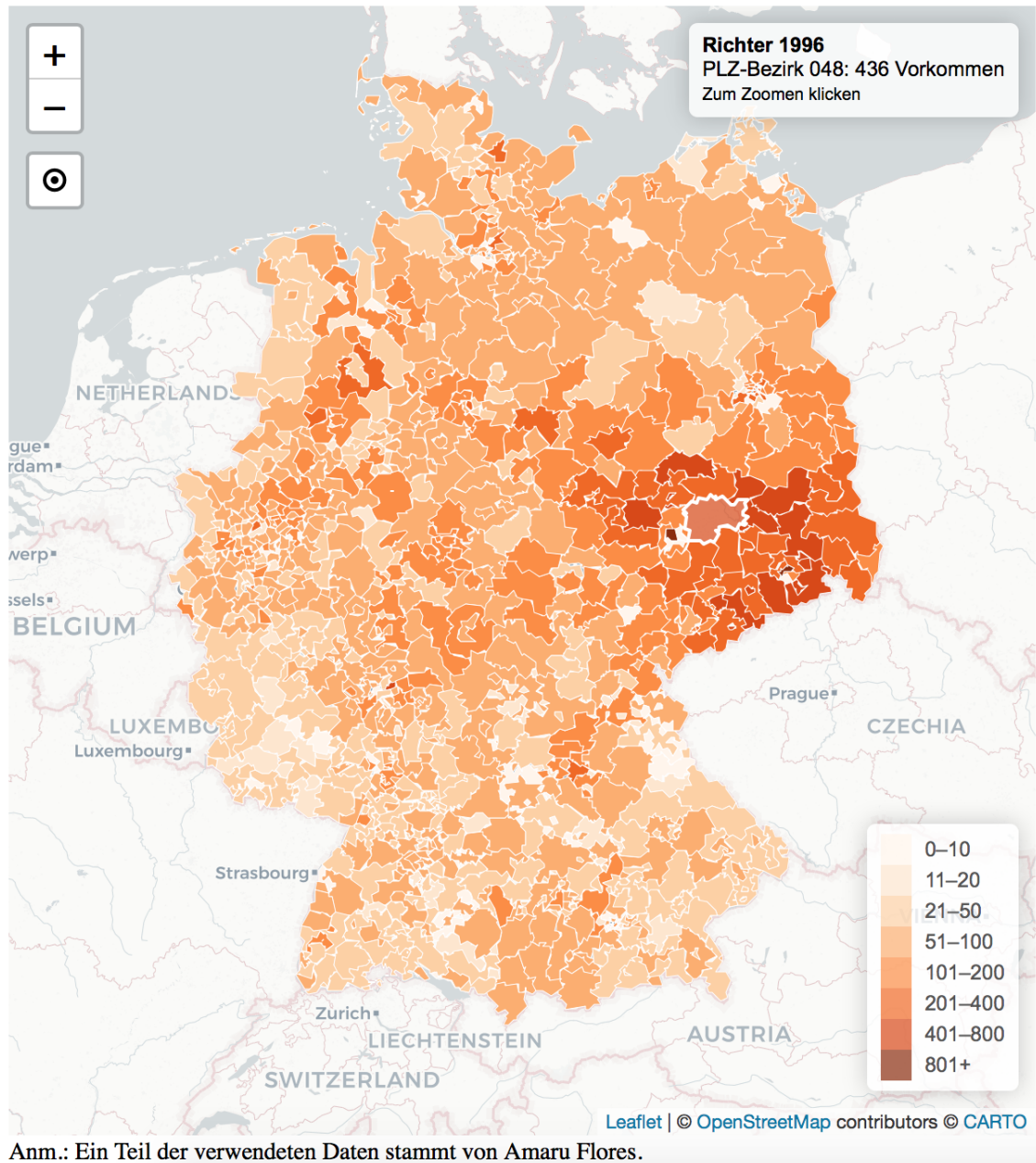


Abbildung 6.6: Interaktive Verbreitungskarte für den Namen Richter im Jahr 1996 in Flächendarstellung

Bei den Datenwerten des Schlüssels „anzahl1996“ mit den Angaben zu den Namensvorkommen in den Postleitzahlenbereichen handelt es sich um quantitative Daten, die sich nach ihrer mathematischen Größe ordnen lassen. Insbesondere lässt sich ein Minimum und ein Maximum der Werte bestimmen. Für das Visualisierungsideom der Choropleth-Karte werden jeweils abgeschlossene Intervalle gebildet, die sequentiell angeordnet sind und vom zuvor bestimmten Minimum bis zum Maximum der Zahlenwerte des Schlüssels „anzahl1996“ reichen. Diese Intervalle werden in ein neues, sogenanntes abgeleitetes Attribut „intervals“ geschrieben (s.u.).

```
intervals
[0-10]
[11-20]
[21-50]
...
[401-800]
[801-max]
```

Das abgeleitete Attribut „intervals“ enthält *ordinale* Daten. Diese ordinalen Daten werden in einem mehrdimensionalen Mapping auf das gleiche graphische Element Fläche wie die Datenwerte des Schlüssels „plz“ abgebildet. Dem graphischen Element Fläche kann anschließend eines der nach dem Prinzip der Expressivität für ordinale Daten passenden visuellen Attribute in der linken Spalte der Tabelle in der Abbildung 4.2 zugeordnet werden. Die Wahl für das Mapping fällt auf das im Effektivitätsrang an der dritten Stelle stehende visuelle Attribut *Farbsättigung* bzw. *Farbhelligkeit*. Die Anzahl der Farbabstufungen entspricht der Anzahl der verschiedenen Intervalle in dem abgeleiteten Attribut. Den einzelnen Intervallen in „intervals“ werden von klein nach groß jeweils die Farbabstufungen wenig gesättigt bis stark gesättigt zugeordnet. Die Zuordnungen sind also so gewählt, dass die Betrachter*innen die unterschiedlich farbgesättigten Flächen der einzelnen Postleitzahlenbereiche und damit die Größenordnungen und die Häufungen der visualisierten Namensvorkommen unmittelbar wahrnehmen und gut unterscheiden können (vgl. Ware 2013, Kap. 4 und Munzner 2014, Kap. 10). Bei der Auswahl der für die Farbkodierung der Flächen aussagekräftigsten Grundfarbe hilft die speziell für die Visualisierung mit Farben in der Kartographie zur Verfügung gestellte Webanwendung Colorbrewer.¹²

Interaktionsmöglichkeiten für die Nutzer*innen werden in der Choropleth-Karte nur innerhalb der Karte selbst angeboten. Beim Schweben mit der Maus über einem Postleitzahlenbereich erscheinen Informationen über das entsprechende Namensvorkommen. Beim Klick auf den ausgewählten Postleitzahlenbereich wird die Farbhelligkeitsstufe für die Zeit der Betrachtung so verändert, dass die geographischen Informationen der Region in der Hintergrundkarte sichtbar werden. Die Choropleth-Karte ist ein weiteres Beispiel für eine Visualisierung der zugrunde liegenden quantitativen Daten, die Auskunft über die Geltungsbereiche der Datenwerte gibt. Die Spezifikation dieser Kartenstudie erfüllt die Anforderungen H1, H2, H3 und H11.

¹²<https://colorbrewer2.org/#type=sequential&scheme=Reds&n=8>, zuletzt abgerufen am 03.09.2021. Eine weitere mögliche Quelle zur Auswahl von Farben ist die JavaScript-Bibliothek d3.js (<https://d3js.org>).

In ersten Gesprächsrunden über die Vor- und Nachteile der visuellen Kodierung der Daten mit größenkodierten Kreisen auf der einen und mit farbkodierten Flächen auf der anderen Seite wurde die Visualisierung mit Kreisen besser angenommen. Die Choropleth-Karte wurde im Gespräch mit Nutzer*innen als durchaus interessant, aber weniger ausdruckskräftig wahrgenommen - einer der Kritikpunkte war z.B., dass Regionen, in denen keinerlei Namensvorkommen verzeichnet werden, in der Choropleth-Karte nur schwer auszumachen sind, während sie in der Karte mit Kreisen unmittelbar erkennbar sind. An diesem Punkt verletzt das Visualisierungsideom der Choropleth-Karte das Prinzip der Expressivität. Die Bevorzugung der Karte mit Kreisen steht zudem eher im Einklang mit dem Effektivitätsprinzip, da das visuelle Attribut Flächengröße im Effektivitätsrang höher steht als die visuellen Attribute Farbsättigung bzw. Farbhelligkeit. Ein weiteres Beispiel für die Visualisierung der zugrunde liegenden quantitativen Daten durch Farbkodierung ist die sogenannte Heatmap. Das Visualisierungsideom der Heatmap wurde noch kontroverser diskutiert als das der Choropleth-Karte, soll aber dennoch hier kurz vorgestellt werden.

6.3.4 Heatmaps

Die Abbildungen 6.7 und 6.8 zeigen Visualisierungen der Verbreitungen des Familiennamens Nowak mithilfe von Heatmaps. Den für den Namen Nowak erzeugten historischen Heatmaps liegen u.a. die Daten in den Feature Collections `Nowak_1890.js` und `Nowak_1996.js` zugrunde. Die Feature Collections mit den quantitativen Daten der jeweiligen Namensvorkommen werden bei der Datenbankabfrage mithilfe des Python-Skripts `name_heat.py` für die Visualisierung in den Heatmaps angepasst (s. Dateien `heatData_Nowak1890.js` bzw. `heatData_Nowak1996.js`).

Die Geokoordinaten in den Feature Collections werden in der Heatmap wieder auf das graphische Element *Punkt* mit zugeordnetem visuellen Attribut *Position* abgebildet. Außerdem wird, ähnlich wie bei der Choropleth-Karte, auch bei der Heatmap ein abgeleitetes Attribut mit *ordinalen* Daten erzeugt. Diese Daten entsprechen einer abgestuften Temperaturskala von Heiß bis Kalt. Die verschiedenen Temperaturabstufungen werden assoziativ auf das visuelle Attribut *Farbton* abgebildet, wobei der Farbton Rot für Heiß und über verschiedene Zwischenstufen aus Gelb- und Grüntönen der Farbton Blau für Kalt steht. Anschließend werden die Farbtöne den Zahlenwerten in den Datenwerten des Schlüssels „count“ zugeordnet. Der höchste, heiße Zahlenwert erhält dementsprechend den Farbton Rot und der niedrigste, kalte Zahlenwert über die diversen Zwischenstufen den Farbton Blau.

Jedem graphischen Element Punkt in der Karte wird einmal oder mehrmals das visuelle Attribut *Flächengröße* und in einem mehrdimensionalen Mapping die Farben Rot bis Blau von Heiß nach Kalt zugeordnet. Das Ergebnis der Mappings ist in den Karten der Abbildungen 6.7 und 6.8 zu sehen. Im Gegensatz zu den Mappings in den Karten mit Kreisen lassen sich die farbkodierten Flächen der Heatmap keiner geometrischen Form zuordnen. Ihre Begrenzungen sind auch nicht wie in der Choropleth-Karte durch Geodaten festgelegt sondern resultieren aus für diese Art der Visualisierung angewandten mathematischen Berechnungen und zeigen vom Heißen zum Kalten hin zunehmend unscharf begrenzte Flächendarstellungen in der geographischen Karte. Sie geben insofern auch keine klar

Nowak 1890

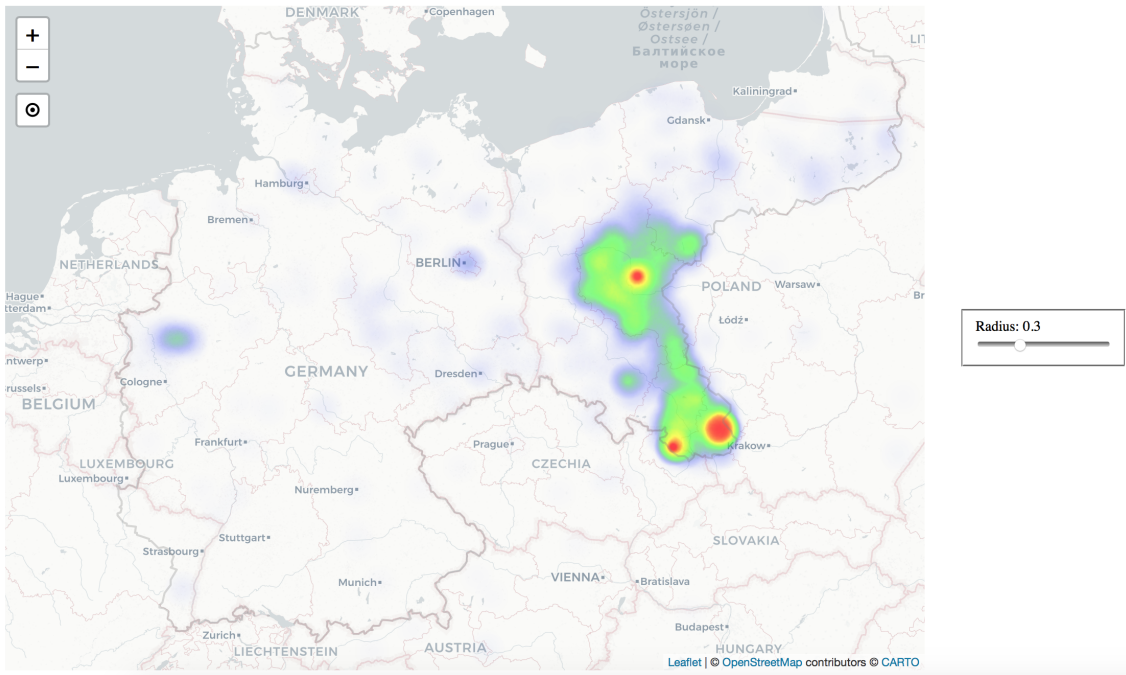


Abbildung 6.7: Heatmap für den Namen Nowak im Jahr 1890

Nowak 1996

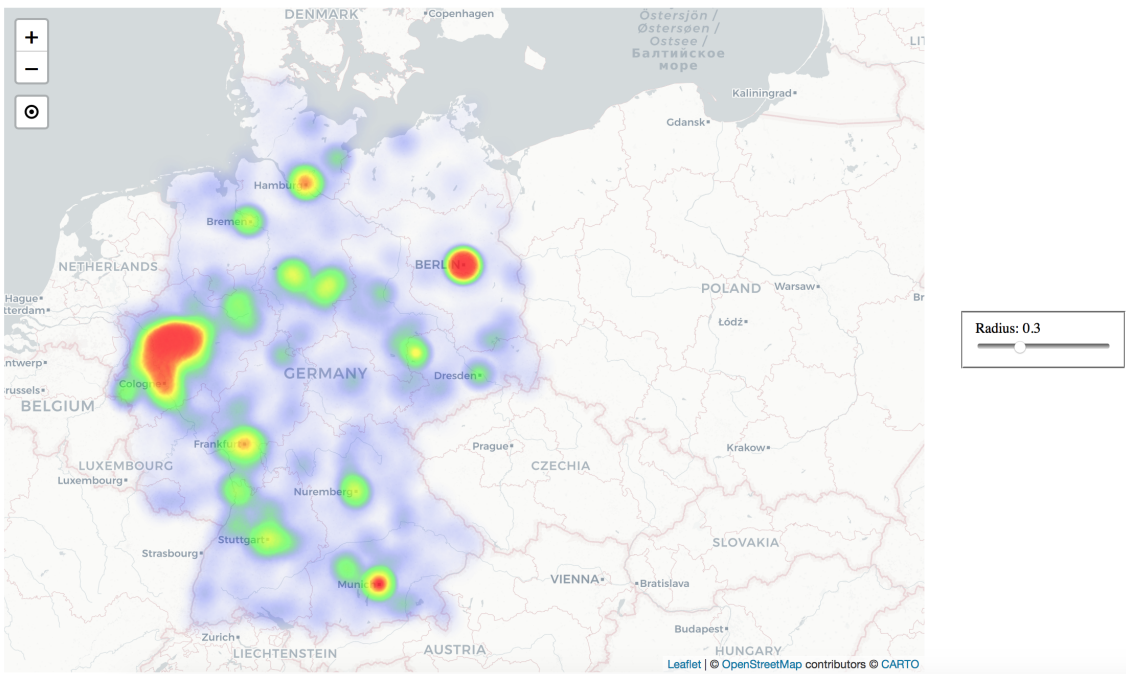


Abbildung 6.8: Heatmap für den Namen Nowak im Jahr 1996

umrissenen Geltungsbereiche für die der Visualisierung zugrunde liegenden Datenwerte an. In Bezug auf das zugrunde gelegte Mapping-Modell für die Verbreitungskarten im DFD erfüllen die Heatmaps zwar mit der Auswahl möglichst hoch geranker visueller Attribute die Kriterien der Effektivität (s. Tabelle 4.2). Mit dem Mapping *ordinaler* Daten auf das für *kategoriale* Daten passende visuelle Attribut *Farbton* für die Ausgestaltung des graphischen Elements Punkt bricht das Visualisierungsidiom jedoch streng genommen mit dem Prinzip der Expressivität, da Farbtöne im Gegensatz zu den visuellen Attributen Farbsättigung bzw. Farbhelligkeit, die jeweils geordnete Abstufungen von Farben darstellen, in der Regel nicht geordnet sind. Erst über die assoziative Verknüpfung eines bestimmten Farbtons mit einer bestimmten Temperatur werden die Farbtöne Rot für Heiß über die geordneten Zwischentöne bis hin zu Blau für Kalt in eine feste Ordnung gebracht und eignen sich daher hier für das beschriebene Mapping.

Ungeachtet einer vergleichsweise geringen Akzeptanz innerhalb der befragten Benutzergruppe zeigen die beiden Heatmaps in den Abbildungen 6.7 und 6.8, dass auch sie m.E. geeignet sind, um den Betrachter*innen die Häufungen von Namensvorkommen in bestimmten geographischen Regionen sowie von Namenverschiebungen von der Zeit um das Jahr 1890 bis ins Jahr 1996 zu demonstrieren. Die visuelle Kodierung der Daten ist jedoch als insgesamt weniger effektiv einzustufen als die Kodierung mithilfe von Kreisen. Zudem müssen für eine möglichst spontane Rezeption der Namensverbreitungen in den beiden verschiedenen Zeiträumen 1890 und 1996 die beiden entsprechend kleiner dargestellten Karten neben- oder untereinander auf einer Anwendungsseite präsentiert werden. Denn ein Umschalten von einer Karte zur anderen auf zwei verschiedenen Seiten würde wegen der auf Sekunden beschränkten Kapazität des menschlichen Kurzzeitgedächtnisses der Anforderung der spontanen Erfassung der Namenverschiebungen nicht genügen (vgl. Ware 2013, Kap. 11). Die Spezifikation dieser Kartenstudie erfüllt die Anforderungen H1 (m.E.), H3, H8 und H11.

6.3.5 Markieren von geographischen Regionen

In der letzten hier vorgestellten Kartenstudie steht nicht die Kartierung der Verbreitungen von Familiennamen im Vordergrund. Mithilfe dieser Karte wird vielmehr eine Möglichkeit aufgezeigt, in einer geographischen Karte ausgesuchte Regionen interaktiv zu markieren. Von einigen Mitarbeiter*innen bzw. Artikelautor*innen des DFD wurde der Wunsch geäußert, in den Namensverbreitungskarten die Ausdehnung verschiedener historisch interessanter Sprachregionen anzuzeigen, unterschiedliche Dialektgebiete zu kennzeichnen oder Häufungen von Namensvorkommen an bestimmten Orten zu unterstreichen. Das für die JavaScript-Bibliothek Leaflet entwickelte Plugin Leaflet.draw erlaubt es, Vektoren und Marker zu zeichnen, zu editieren und in Layern über eine Leaflet-Karte zu legen. Es können hierbei beliebig viele Polylinien, Polygone, Rechtecke und Kreise in beliebiger Größe erzeugt, nach Bedarf angepasst und wieder gelöscht werden. Für die Anwendung des Plugins in der in der Abbildung 6.9 gezeigten interaktiven Kartenstudie wurde aus den verschiedenen Markierungsmöglichkeiten die Option, Polygone in beliebigen Formen zu zeichnen und zu editieren, ausgewählt. Für das Einzeichnen der einzelnen Polygone wird wieder die *Position* als Bestandteil der *geometrischen* Daten der geographischen Hintergrundkarte

Farbige Markierung von Regionen durch Einzeichnen von Polygonen. (Löschen eines Polygons per Doppelclick)

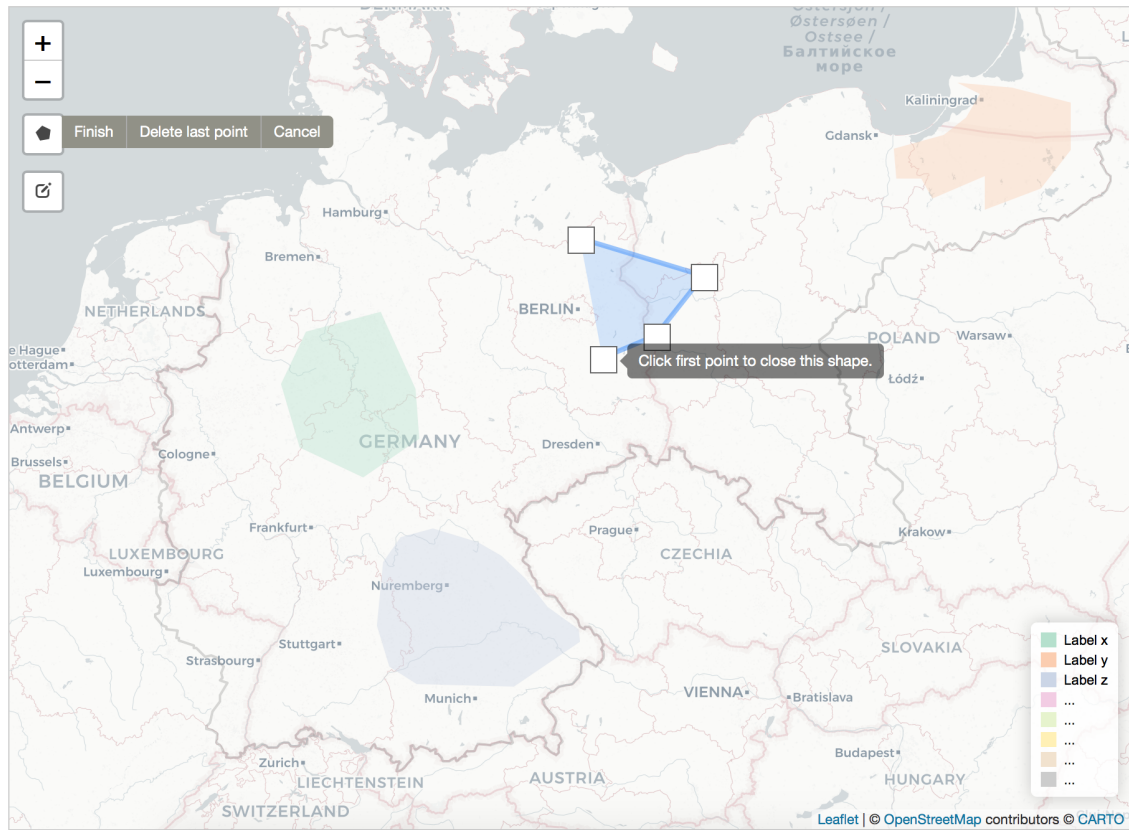


Abbildung 6.9: Interaktives Zeichnen von Polygonen in der geographischen Karte

genutzt. Die für das Mapping auf die graphischen Elemente *Punkt* benötigten Geokoordinaten liegen in diesem Fall nicht in einer statischen Datenmenge vor sondern werden durch Benutzeraktionen erzeugt. Ergänzend zum Plugin wurde zusätzlicher Code geschrieben, um eingezeichnete Polygone unterschiedlich einfärben zu können und der Karte wurde eine Legende zur Erläuterung der Bedeutung der entsprechenden Farben beigefügt. Die Abbildung 6.9 zeigt eine noch rudimentäre Version dieser Anwendung, in der mehrere Regionen in einer geographischen Karte mithilfe von Polygonen in unterschiedlichen Farben markiert sind.

Über die im Plugin vorgesehene Anwendung hinaus wäre auch die Speicherung der während der Arbeit mit der Karte von den Nutzer*innen interaktiv erzeugten Geodaten, ggf. angereichert durch Metadaten, denkbar. Die auf diese Weise erzeugten Datenmengen könnten dann als externe Datenmengen für die Erzeugung von entsprechenden Layern über beliebigen weiteren Leaflet-Karten eingesetzt werden. Die beschriebene Technik könnte sich m.E. auch zur Markierung von Geltungsbereichen für Datenwerte eignen. Die Spezifikation dieser Kartenstudie erfüllt die Anforderungen H12 und ggf. H6.

7 Diskussion und Ausblick

Zu Beginn der Softwareentwicklung für das interaktive Kartenmodul für das DFD waren einige Schwierigkeiten zu überwinden und Entscheidungen zu treffen, die sowohl die vorhandene Datenlage und die Rolle des DFD-Kartierungstools als auch die verwendeten Technologien, die Gestaltung der Anwendungsseiten und das Design des Visualisierungs- und Interaktionsidioms der interaktiven Verbreitungskarte betreffen. Unter den Mitarbeiter*innen des DFD-Projekts besteht bereits seit geraumer Zeit Konsens darüber, dass die mithilfe des DFD-Kartierungstools für die Namenartikel im Wörterbuch produzierten statischen Karten im PDF- bzw. SVG-Format durch an aktuelle Standards angepasste Verbreitungskarten mit Möglichkeiten der Interaktion durch die Nutzer*innen ersetzt werden sollen, wobei die Veranschaulichung der Namensverbreitungen mithilfe von Kreissymbolen unter den Artikelautor*innen auch für die interaktiven Verbreitungskarten favorisiert wird. Die konkrete Umsetzung dieser Aufgabe in Form eines neu entwickelten interaktiven Kartenmoduls für das DFD erfordert jedoch zunächst, dass die Mitarbeiter*innen sich von einigen bislang gewohnten Konzepten lösen (vgl. Lynch und Horton 2016, S. 2). Mit der Einführung interaktiver Namensverbreitungskarten müssen die Artikelautor*innen beispielsweise auf die derzeit übliche Darstellung der Namensverbreitungen in einer stilisierten Deutschlandkarte verzichten und sich stattdessen auf verschiedene Stile geographischer Karten einlassen. Das Lösen vom sehr zurückhaltenden Design der bisherigen Verbreitungskarten mit wenigen Gestaltungselementen fällt z.T. schwer. In gemeinsamen Besprechungen wurde daher die Auswahl der Kartenstile, die im interaktiven Kartenmodul angeboten werden sollen, intensiv diskutiert. Einige grundsätzliche Kriterien konnten benannt werden. Demnach sollen die Karten in erster Linie von einem nichtkommerziellen Anbieter stammen und kostenfrei bereitgestellt werden. Die Tiles, aus denen eine Karte zusammengesetzt ist, sollen fehlerfrei und in hoher Auflösung vorliegen.¹ Weitere diskutierte Details betreffen beispielsweise die möglichst wenig prominente Darstellung der für die Visualisierungsaufgabe weniger wichtigen Autobahnen, die der Intention der Präsentation der Namensverbreitungen angemessenen Farbgebungen innerhalb der geometrischen Daten der Karten oder die aussagekräftige Darstellung der Topographie als Hilfestellung für Namenverortungen und -deutungen. Einige moderne Kartenanbieter bieten geographisches Kartenmaterial an, das sich im Hinblick auf die technischen und gestalterischen Details sehr gut für das Projekt eignen würde. Diese Karten werden jedoch von kommerziellen Providern angeboten und ihr Einsatz in der Webanwendung wäre mit laufenden Kosten verbunden. Die Anforderung der Kostenfreiheit für das DFD hat dementsprechend zur Folge, dass Abstriche für die Hintergrundkarten im interaktiven Kartenmodul gemacht und Kompromisse eingegangen werden

¹Unter den verschiedenen möglichen Anbietern von Hintergrundkarten weisen z.B. die Karten des Anbieters Wikimedia beim Heranzoomen Fehler bei den Übergängen zwischen den einzelnen Kacheln auf.

müssen, nicht nur in Bezug auf die geographischen Details der Tiles, sondern auch was die Beschriftungen in den Karten angeht. So finden sich beispielsweise bei einem Großteil der Anbieter englische Labels zur Bezeichnung von Ländern, Städten etc. oder es wird eine überwiegend deutsche Beschriftung für kleinere Städte und Orte gewählt und einige wenige deutsche Großstädte werden mit englischen Labels versehen. Die Stadt „München“ erscheint dann beispielsweise mit dem Label „Munich“ und auch die Staatsbezeichnung lautet „Germany“ anstatt „Deutschland“. Ein Vergleich diverser Karten von unterschiedlichen kostenfreien Anbietern führt zur Auswahl der geographischen Karten der an das OpenStreetMap-Projekt angebotenen Leaflet Tile Provider Mapbox und Carto, deren Tiles die oben beschriebenen Kriterien weitgehend erfüllen.

Über die Umstellung auf moderne Hintergrundkarten hinaus sollten die Artikelautor*innen auch bereit sein, Interaktionen innerhalb der Verbreitungskarten zuzulassen, mit denen die Nutzer*innen die Karte in ihrem Sinne manipulieren können. Der Sorge, dass mit den Benutzerinteraktionen innerhalb der Verbreitungskarte die Informationen zur Verbreitung eines Namens, die die im Wörterbuch veröffentlichte Karte vermitteln soll, verfälscht oder verloren gehen könnten, konnte bei der Entwicklung der interaktiven Karten durch die konsequente Vorrangstellung der Default-Karte begegnet werden. Die Default-Einstellungen sind während der interaktiven Arbeit mit der Verbreitungskarte zu jedem Zeitpunkt „nur einen Klick weit entfernt“.

Während der Entwicklung des interaktiven Kartenmoduls wurde im Kreis der Artikelautor*innen beschlossen, die bisherige Praxis, die Verbreitungskarten mit individuell ausgewählten Parametereinstellungen im Kartierungsprogramm zu erzeugen, aufzugeben. Diese Praxis folgt weder einheitlichen Regeln noch wurden die individuell vorgenommenen Parametereinstellungen in der Vergangenheit systematisch abgespeichert. Die Automatisierung der Kartierung wird dadurch erheblich erschwert. Im Demonstrator für das interaktive Kartenmodul werden zur Zeit alle Verbreitungskarten einheitlich nach den für den Demonstrator festgelegten Regeln zur Parametereinstellung automatisch erstellt. Nach gemeinsamen Gesprächen und Überlegungen zu den Parametereinstellungen für automatisch erstellte aussagekräftige Namensverbreitungskarten im interaktiven Kartenmodul haben die Mitarbeiter*innen des DFD-Projekts in der Zwischenzeit drei differenzierte Regeln formuliert, nach denen jede Verbreitungskarte, die im Wörterbuch veröffentlicht wird, erzeugt werden soll.

1. Liegt die Anzahl der Namensvorkommen eines Namens insgesamt unter dem Wert 800, so werden „absolute“ Karten erzeugt, ansonsten „relative“ Karten.
2. Die Parametereinstellung „Städte zusammenfassen“ wird aktiviert, wenn das größte Namensvorkommen, d.h. der größte Kreis in der Visualisierung, in einer Großstadt wie z.B. Berlin, Hamburg oder München liegt.
3. Sind in einer „absoluten“ Karte alle Kreise gleich groß, d.h. alle Kreise repräsentieren jeweils nur einen Telefonanschluss, dann wird die Kreisgröße insgesamt von 35 auf 20 verringert.²

²Die Dimension der Werte 35 bzw. 20 ist im DFD-Kartierungstool nicht dokumentiert.

Mit der Implementierung dieser Regeln im JavaScript-Code der interaktiven Anwendung - „von 35 auf 20 verringern“ bedeutet z.B. die entsprechend berechnete Anpassung des Skalierungsfaktors für die Darstellung der Kreise in den interaktiven Verbreitungskarten - sind die endgültigen Parametereinstellungen für jede Karte festgelegt. Damit ist ein einheitliches Schema für die automatische Erzeugung der interaktiven Verbreitungskarten für alle veröffentlichten Namenartikel mit Verbreitungskarte festgelegt und die bisherige Aufgabe der einzelnen Autor*innen, für jeden zukünftig bearbeiteten Namen eine individuelle Karte zusammen zu stellen, entfällt. Zudem ist mit der Umstellung auf einheitliche Regeln zur Kartenerstellung die Transparenz der vorgenommenen redaktionellen Einstellungen nicht nur für den Kreis der Mitarbeiter*innen des DFD-Projekts sondern auch für die Nutzer*innen des Wörterbuchs gewährleistet. Als ein willkommener Nebeneffekt werden so auch veraltete Verbreitungskarten längst veröffentlichter Namen aus der Anfangszeit des DFD durch nach dem implementierten Schema erzeugte aktuelle Karten ersetzt.

Eine Herausforderung für die Entwicklung der interaktiven Karten stellte zunächst die in Bezug auf die Anordnung und Größe korrekte Darstellung der Kreise in den interaktiven Verbreitungskarten dar. Neben den mithilfe des DFD-Kartierungstools erstellten Verbreitungskarten im PDF-Format können seit der Weiterentwicklung des Tools im Jahr 2014 auch SVGs der Karten erzeugt werden. Anfängliche Überlegungen, den XML-Code dieser SVGs mithilfe der Python-Bibliothek *ElementTree*³ zu parsen und so die erforderlichen Daten zur Größe und Verortung der Kreise in der Karte zu erhalten, wurden verworfen. Zum Einen hätten für eine große Menge von Namen die entsprechenden SVGs nachproduziert werden müssen, zum Anderen stellte sich diese Arbeitsweise des sogenannten Reverse Engineering als viel zu aufwendig und nicht angemessen heraus. Stattdessen wurde, wie aus der Dokumentation der Softwareentwicklung ersichtlich ist, die für das DFD-Kartierungstool entwickelte JavaScript-Bibliothek der Softwarefirma Tivano nachvollzogen und auf die Anwendung angepasst. Um eine problemlose spätere Integration des interaktiven Kartenmoduls in die Website des DFD zu gewährleisten, wird auf die Anpassung der mittlerweile sieben oder mehr Jahre alten JavaScript-Bibliothek des DFD-Kartierungstools an aktuellere Technologien, wie z.B. neue Funktionalitäten in JavaScript und D3 oder das modernere JSON-Dateiformat als Ersatz für csv-Dateien, verzichtet und stattdessen darauf geachtet, nur solche Modifikationen vorzunehmen, die die Kompatibilität des Moduls mit den übrigen Softwarekomponenten des DFD weitgehend sicher stellen. Das für das interaktive Kartenmodul geschriebene Skript `circles.js` nutzt die gleichen Daten auf die gleiche Weise wie die Software-Bibliothek `karte.js` des DFD-Kartierungstools. Ein Funktionstest mit dem Einsatz der Datei `circles.js` anstelle von `karte.js` im DFD-Backend ist geplant jedoch bislang nicht erfolgt. Wegen der fehlenden Dokumentation sowohl der Datenbasis als auch der JavaScript-Bibliothek des DFD-Kartierungstools erwies sich die Anpassung der Software an die hier beschriebene Anwendung als vergleichsweise aufwendig. Der Vorteil der erwarteten problemlosen Integration des interaktiven Kartenmoduls in die Website des DFD als Ersatz für das bisher verwendete DFD-Kartierungstool wiegt diesen Aufwand jedoch wieder auf.

Wie für die technische Umsetzung des DFD generell gilt auch für die gesamte für das interaktive Kartenmodul entwickelte Software, dass moderne, offene Standards bzw. offene

³<https://docs.python.org/3/library/xml.etree.elementtree.html>

Standardsoftware verwendet werden. Es werden konsolidierte Technologien wie ausgesuchte Teile von JavaScript, D3, Python und Leaflet mit ausgesuchten Leaflet Plugins eingesetzt. Die Skripte und ggf. die Stylesheets der Leaflet Plugins werden dabei, soweit möglich, lokal installiert. Für die Generierung der interaktiven Karten für die Webanwendung bot sich neben der JavaScript-Bibliothek Leaflet auch die ebenfalls frei verfügbare JavaScript-Bibliothek OpenLayers an.⁴ Ein Vergleich der beiden Bibliotheken sowohl im Hinblick auf die Anforderungen an die Karten-Komponente des DFD als auch auf die aktuelle Verbreitung der jeweiligen Bibliothek und die Lebendigkeit der Entwickler-Communities fiel zugunsten von Leaflet aus. Leaflet bietet zudem die für die Anwendung ausgewählten Hintergrundkarten der Tile Provider Mapbox und Carto an. Im Vergleich zur derzeit verwendeten Kartierung entsprechen die mit einem stets in der Weiterentwicklung befindlichen Werkzeug wie Leaflet generierten interaktiven Karten modernen Anforderungen beispielsweise auch an die Barrierefreiheit und das Responsive Design.⁵

Der Einsatz der angepassten JavaScript-Bibliothek des DFD-Kartierungstools für die Erzeugung der Kreise-Layer in den interaktiven Leaflet-Karten wird u.a. auch deswegen möglich, weil für die Erstellung der statischen Verbreitungskarte in der Tivano Software eine einfache Rektangularprojektion für die Berechnung aller in der Datenbasis vorhandenen Geokoordinaten eingesetzt wird.⁶ Die Rektangularprojektion wird in Verbindung mit Transformationsberechnungen benutzt, eine Vorgehensweise, die es laut Tivano erlaubt, bei der Entwicklung der Kartierungssoftware direkt mit den gängigen Geokoordinaten-Angaben in Längen- und Breitengraden für alle geographischen Orte der Erde zu arbeiten. Damit entstünde kein Konflikt mit der in den Leaflet-Karten verwendeten Web Mercator Projektion. Jedoch ist weder die Aufbereitung der Telekom-Rohdaten noch die JavaScript Software-Bibliothek durch die Firma Tivano ausreichend dokumentiert. Da damit auch der Ursprung bzw. die Berechnung der in den aufbereiteten Daten enthaltenen Geokoordinaten unklar ist, wurde zur Überprüfung der in den Daten für das DFD-Kartierungstool enthaltenen Breiten- und Längengrade der Mittelpunkte der Postleitzahlenbereiche eine weitere Leaflet-Karte erzeugt (s. Datei `cities.html`). Zu den mitgelieferten Daten des DFD-Kartierungstools gehört die Datei `cities.csv`, in der die dort verwendeten Geodaten deutscher Städte gespeichert sind. Diese Daten werden auf Positionen in der geographischen Karte abgebildet und mithilfe von Punktsymbolen in einem Layer über der Leaflet-Karte dargestellt. Die Abbildung 7.1 zeigt, dass die Positionen der Punktsymbole und die tatsächliche Verortung der Städte in der Leaflet-Karte nicht deckungsgleich sind. Stattdessen werden in dem mit den Geokoordinaten des DFD-Kartierungstools erzeugten Layer für nahezu alle Städte unsystematisch abweichende Positionen festgestellt.

Damit liegt die Vermutung nahe, dass auch die Geokoordinaten der Mittelpunkte der Postleitzahlenbereiche in den von Tivano erstellten Daten nicht genau mit den tatsächlichen Geokoordinaten dieser Punkte übereinstimmen. Abgesehen von der aufgrund dessen

⁴<https://openlayers.org>

⁵Wie bei anderen Softwarelösungen gibt es auch hier typische Klick-Touch Unterschiede wie Tooltips oder Pinch.

⁶Bei einer Kartenprojektion werden mathematische Berechnungen durchgeführt, um das auf der gekrümmten Erdoberfläche verwendete Koordinatensystem in ein Koordinatensystem für eine flache Oberfläche zu konvertieren. Es gibt eine Reihe verschiedener Kartenprojektionen, die unterschiedliche Eigenschaften bereitstellen.

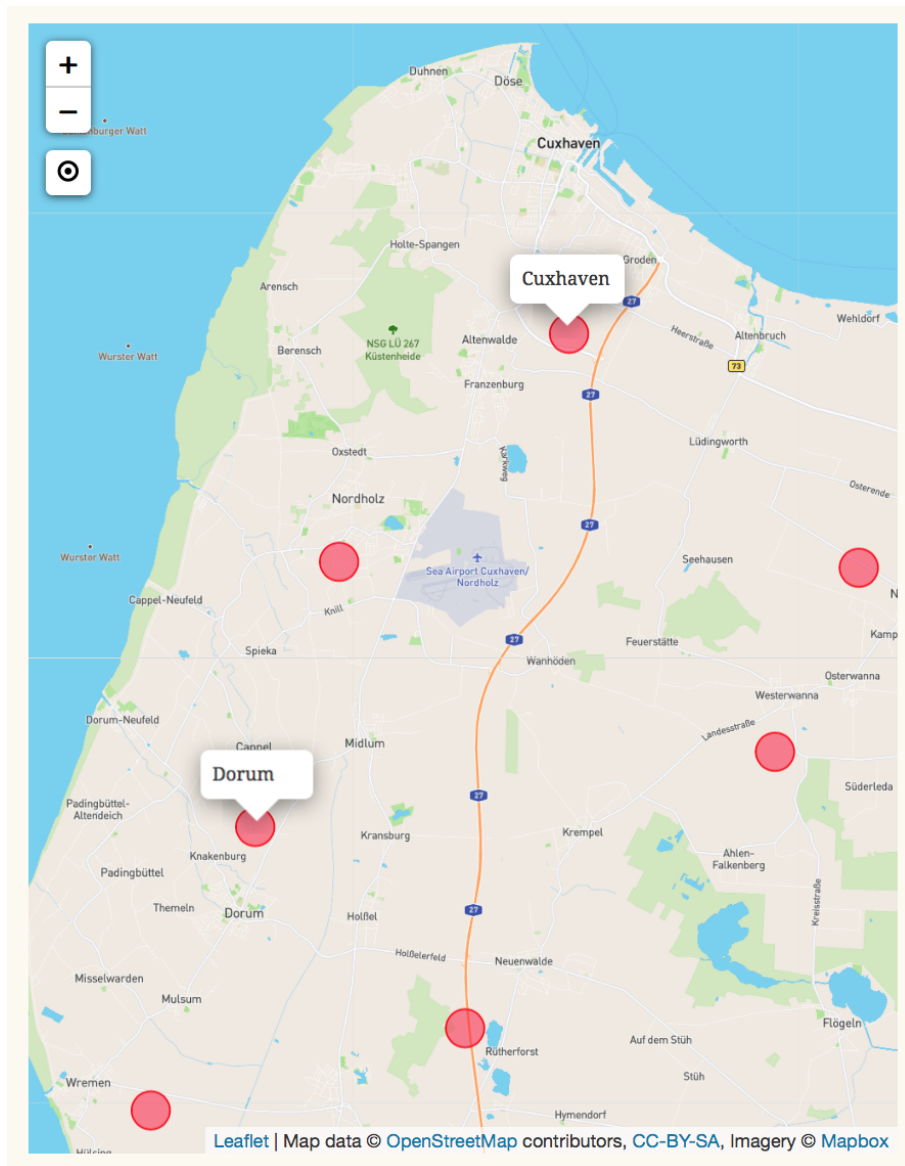


Abbildung 7.1: Abweichende Geokoordinaten deutscher Städte

möglicherweise nicht in jedem Fall exakten Verortung der Kreise in der Verbreitungskarte würde auch die in der Evaluation des Visualisierungsideoms für das interaktive Kartenmodul angemahnte Angabe der Geltungsbereiche für die visualisierten Datenwerte zusätzlich erschwert. Außerdem würde das Prinzip der Expressivität der Visualisierung an dieser Stelle verletzt. Im Zweifelsfall steht das manuelle Ersetzen aller Geokoordinaten in den Daten des DFD-Kartierungstools durch die entsprechenden Daten in der Mercator-Projektion an. Die Problematik der abweichenden Geokoordinaten wurde mit der Fokusgruppe im DFD-Projekt besprochen mit dem Ergebnis, dass dieser Umstand für die Softwareentwicklung des Demonstrators für das interaktive Kartenmodul zunächst keine vorrangige Rolle spielen sollte.

Weder die im Sinne der Forschungsergebnisse der Informationsvisualisierung benutzerfreundliche Implementierung der wahrnehmungskonformen Größenberechnungen der Kreisradien noch die der Geltungsbereiche für die repräsentierten Datenwerte ist also unmittelbar auf die Visualisierungsideome der Namensverbreitungskarten im interaktiven Kartenmodul übertragbar. Diese Implementierungen würden nur möglich durch die vollständige Lösung der Datenbasis und der Software-Bibliothek des interaktiven Kartenmoduls von der Datenbasis und der Software-Bibliothek des DFD-Kartierungstools. Eine Lösung vom DFD-Kartierungstool würde jedoch bedeuten, dass ein von Grund auf neues Kartierungsprogramm basierend auf den Telekom-Rohdaten von 2005 entwickelt werden müsste. Im DFD-Server wäre die Einrichtung einer eigenen Projektdatenbank erforderlich, die ggf. auch mit der Datenbank der historischen Karten zusammengeführt würde. Für die Visualisierung der Namensverbreitungen stünde die Neuentwicklung einer eigenen JavaScript-Bibliothek an. Es müsste geklärt werden, ob bzw. in wie weit das Migrieren auf neue Technologien im DFD-Projekt vorbereitet ist. Das Vorhaben würde u.U. eine Umstellung der gesamten Website auf neuere Technologien erfordern, es wäre damit sehr komplex und zeitintensiv. Dem Vorteil eines an moderne technologische Standards und an die Erkenntnisse der Informationsvisualisierung angepassten interaktiven Kartenmoduls für das DFD stünde die ungleich schwierigere und langwierigere Einbindung des Moduls in die aktuelle Softwarearchitektur des DFD entgegen.

Die Überprüfung der im Abschnitt 6.1 aufgeführten Liste von zusätzlichen Aufgaben für die Softwareentwicklung sowohl für das interaktive Kartenmodul als auch für die historische Kartierung zeigt Möglichkeiten zur Erweiterung und Verbesserung des Angebots sowohl für einzelne Nutzer*innen ohne oder mit wenig fachlicher Expertise als auch für die Benutzergruppe mit Expertenwissen mit den entsprechend spezialisierten wissenschaftlichen Anwendungen und Fragestellungen. In der Aufgabenliste bleiben für das interaktive Kartenmodul beispielsweise die Anforderungen der Suche über einen Ort oder eine Region anstatt über einen Namen (H13) oder die Überblendung verwandter Namen in ein und derselben Karte (H4) offen. Da bei der Erstellung der Namenartikel für das DFD derzeit verwandte Namen nur unsystematisch erfasst werden, kann der Wunsch nach einer Überblendung verwandter Namen in ein und derselben Karte nicht zeitnah erfüllt werden. Was die Suche über einen Ort oder eine Region anstatt über einen Namen angeht, so hat Saul Wurman in seinem Buch „Information Anxiety“ das sogenannte LATCH-Prinzip für die Organisation von Information entwickelt (vgl. Wurman 1989). Das Akronym LATCH steht für Location, Alphabet, Time, Category und Hierarchy. Im Organisationsschema des interaktiven Karten-

moduls ist die Information in einer alphabetischen Namensliste (Alphabet) organisiert (s. Abschnitt 3.1.1). Aus dem Organisationsschema leitet sich die Informationsarchitektur und damit die weitere Auslegung des Softwareprojekts ab. Ausgehend von der Datenlage für die Anwendung könnte das Organisationsschema alternativ auch nach dem Rang (Hierarchy), der Häufigkeit (Hierarchy) oder der Verteilung im Raum (Location) organisiert werden. Eine aus dem Organisationsschema des interaktiven Kartenmoduls abgeleitete alternative Auslegung der Informationsarchitektur nach der Verteilung im Raum würde den Einstieg in das Namenkorpus über die Geographie ermöglichen und damit die in der Aufgabe H13 formulierte Anforderung der Suche über einen Ort oder eine Postleitzahlenregion mit den entsprechenden Informationen über die Namendichte in den betreffenden geographischen Regionen erfüllen.

Weitere offene Punkte für das interaktive Kartenmodul sind die Überblendung der Namensverbreitungen aus Verweisartikeln mit der Verbreitung des Namens im zugehörigen Hauptartikel bzw. die Darstellung der Verbreitungen von mehr als einem Namen in ein und derselben Karte (H4), die Kartierung von Berufsnamen, Wohnstättennamen etc. (H14) oder die Einrichtung einer ‚Pro‘-Version aller Verbreitungskarten für die Mitarbeiter*innen des DFD-Projekts und für interessierte Wissenschaftler*innen, u.a. mit der im Einklang mit den Lizenzrechten stehenden Freigabe und Anzeige konkreter Zahlen zu den Namensvorkommen (H15). Lösungsansätze für einige der oben genannten Aufgaben lassen sich durch die weitere Erhebung und Aufbereitung von externen Daten bzw. mithilfe der bereits vorhandenen Daten und durch die Weiterentwicklung des JavaScript-Codes für die Generierung der Leaflet-Karten erarbeiten. Das Gleiche gilt auch für die Implementierung weiterer interaktiver Funktionalitäten für die historischen Namensverbreitungskarten mit Kreisen wie beispielsweise die Erzeugung eines Layers mit der Repräsentation relativer Zahlenwerte oder dem Clustering von Namensvorkommen in Karten mit großen Anzahlen überlagerter Kreise (H10).

Ein von Seiten der Artikelauteur*innen geäußertes Wunsch nach eigens auf das DFD zugeschnittenen geographischen Hintergrundkarten (H17) stellt dagegen ein eigenes Softwareprojekt dar. Um den Kompromissen, die mit der Verwendung der kostenfreien Karten der Leaflet Tile Provider Mapbox und Carto eingegangen werden, zu begegnen und in Zukunft unabhängig von diesen Kartenanbietern arbeiten zu können, müssen zunächst präzise Anforderungen an das Aussehen der für das interaktive Kartenmodul und die historische Kartierung geeigneten Kacheln formuliert werden. Die anschließende Entwicklung eigener geographischer Hintergrundkarten auf Basis des von OpenStreetMap als OpenSource-Angebot zur Verfügung gestellten Kartenmaterials sowie mithilfe der ebenfalls vom OpenStreetMap-Projekt angebotenen Software zur Erzeugung von an den eigenen Bedarf angepassten Tiles könnte dann eine lohnende Investition für die Präsentation des Wörterbuchs im DFD-Frontend sein.

Während die bis hierhin diskutierten offenen Benutzeranforderungen sich ausschließlich auf die Präsentation der Verbreitungskarte und die Interaktionen innerhalb dieser Karte beziehen, stellt die Frage nach erklärenden Graphiken und Diagrammen zur Statistik der Namensvorkommen eine weitere Herausforderung dar, da diese zusätzlich zur Namensverbreitungskarte und unter Beachtung der Expressivitäts- und Effektivitätsprinzipien präsentiert werden müssen (H2). Eine Lösung dafür stellt die sogenannte Juxtaposition, d.h.

die Nebeneinanderstellung verschiedener Datenvisualisierungen auf ein und derselben Seite, dar. Diese Seite enthält dann beispielsweise eine der oben beschriebenen interaktiven Verbreitungs- oder historischen Karten. Ganz im Sinne des semantischen Zoomens als Bestandteil des Visual Information Seeking Mantras Overview-First-Detail-on-Demand von Ben Shneiderman erscheint in einem kleinen Ausschnitt innerhalb der Karte eine durch eine Benutzeraktion speziell ausgewählte Region in der Karte in vergrößerter Darstellung mit angereicherten Informationen. Im unteren Bereich der Seite geben zusätzliche Graphiken und Diagramme weitere Auskünfte über die Daten. Die Stärke der Juxtaposition ist, dass verschiedene Aspekte einer Datenmenge den Betrachter*innen gleichzeitig und am gleichen Ort präsentiert werden, ohne dass zusätzliche Zeit für Interaktionen aufgewendet werden muss. Von den Betrachter*innen wird keine erhöhte Gedächtnisleistung gefordert und ihre Aufmerksamkeitsspanne wird nicht ausgereizt (vgl. Ware 2013, Kap. 11). Zudem lässt sich die Verteilung der verschiedenen visuellen Repräsentationen der Datenmenge, die auf der zweidimensionalen Seite in unmittelbarer räumlicher Nähe zueinander stehen, als eine Abbildung auf das im Effektivitätsrang an oberster Stelle stehende visuelle Attribut *Position im (zweidimensionalen) Raum* interpretieren (vgl. Munzner 2014, Kap. 12).

An dieser Stelle wird darauf hingewiesen, dass viele der bisher nicht erfüllten Anforderungen nicht zu den Fragen und Aufgaben gehören, die von den alltäglichen Nutzer*innen des DFD-Frontends an die Visualisierungen gestellt werden. Es handelt sich stattdessen um spezialisierte Anforderungen der Artikelautor*innen des DFD, deren Arbeit der Namendeutung und Artikelerstellung im Backend des DFD unterstützt und erleichtert werden würde und die im Ersetzen des DFD-Kartierungstools durch das interaktive Kartenmodul Chancen erkennen, um zukünftig eigene Ideen und Vorschläge umzusetzen. Die entsprechend kritische Durchsicht der beiden Anforderungskataloge in den Abschnitten 2.5.2 bzw. 6.1 ergibt, dass die Summen der Anforderungen, die von den beiden Benutzerprofilen „Nutzer*innen ohne fachliche Expertise“, also dem Gros der Nutzer*innen des DFD-Frontend, und „Artikelautor*innen und weitere Nutzer*innen mit Expertenwissen“ u.U. nicht innerhalb einer Anwendung gleichermaßen zu erfüllen sind. Zieht man zudem noch in Betracht, dass während der Weiterentwicklung des interaktiven Kartenmoduls nach der Einbindung des Moduls in die Website des DFD und mit der zusätzlichen Integration historischer Namensverbreitungskarten in das interaktive Kartenmodul weitere Anforderungen an die Visualisierungen hinzu kommen werden, dann könnte es sinnvoll sein, in Vokabular und Angebot unterschiedliche Versionen der Anwendung zu entwickeln und anzubieten. Gemeinsame Überlegungen dazu, ob eine Trennung der beiden Benutzergruppen innerhalb der Softwareentwicklung sinnvoll wäre, und wie ggf. die Ausrichtung der Visualisierungen auf verschiedene Benutzergruppen realisiert werden könnte, stehen aus.

In der Zwischenzeit wurde beiden Softwareprojekten, sowohl dem interaktiven Kartenmodul als auch der bislang lediglich in Form von Kartenstudien existierenden historischen Kartierung, von Nutzer*innen ohne Expertenwissen, den Artikelautor*innen des DFD und im Rahmen der regelmäßig durchgeführten Evaluation des DFD-Projekts auch von einer wissenschaftlichen Kommission bereits jetzt ein deutlicher Mehrwert gegenüber der derzeitigen Kartierung bzw. eine zu erwartende Aufwertung des Wörterbuchs insgesamt zugesprochen.

Letztendlich entscheiden jedoch die Nutzer*innen des Wörterbuchs über die Qualität der Visualisierungen. Auch wenn eine sorgfältig entwickelte und dokumentierte Software und

ihre Evaluation im Hinblick auf die visuelle Repräsentation der Daten des DFD bereits eine hohe Aussagekraft hinsichtlich ihrer Benutzerfreundlichkeit besitzt, würden sich zum Abschluss des Softwareprojekts des interaktiven Kartenmoduls und der historischen Kartierung für das DFD weitere Usability Tests anbieten, um valide Ergebnisse einer solchen Bewertung zu erhalten. Erste zeitnahe Rückmeldungen zur Benutzerfreundlichkeit des Demonstrators des interaktiven Kartenmoduls könnten auch mit einer Evaluation nach der vereinfachten Methode von Steve Krug (vgl. Krug 2014, Kap. 9) eingeholt werden.

8 Zusammenfassung

In der derzeit verwendeten Fassung des Digitalen Familiennamenwörterbuchs Deutschlands (DFD) werden für das Gros der bisher veröffentlichten Namenartikel statische, nicht geographische Namensverbreitungskarten im PDF- bzw. SVG-Format angeboten. Die Karten werden mithilfe des in den Anfangsjahren des DFD von der Firma Tivano Software GmbH entwickelten DFD-Kartierungstools für jeden Namen individuell erzeugt. Viele Fragen der Nutzer*innen an die Visualisierung der Namensverbreitungen im DFD-Frontend werden von der Darstellung in einer statischen Verbreitungskarte nicht beantwortet. Die Aufgabe und das Ziel der in der vorliegenden Arbeit dokumentierten Softwareentwicklung für ein interaktives Kartenmodul für das DFD besteht darin, den Nutzer*innen des DFD-Frontends eine interaktive Namensverbreitungskarte mit geographischer Information anzubieten und das derzeit im DFD eingesetzte DFD-Kartierungstool durch ein neu entwickeltes Kartierungsprogramm zur Erzeugung interaktiver Namensverbreitungskarten zu ersetzen. Zu diesem Zweck wurde die Software-Bibliothek des DFD-Kartierungstools unter Beachtung der erforderlichen Kompatibilität konsequent an die Softwareentwicklung für das interaktive Kartenmodul angepasst. Die Software des Kartenmoduls kann damit unmittelbar in die Website des DFD eingebettet werden. Die interaktiven Namensverbreitungskarten werden mithilfe der JavaScript-Bibliothek Leaflet nach festgelegten einheitlichen Regeln für jeden im Wörterbuch veröffentlichten Familiennamen automatisch erzeugt.

Im Vordergrund der Entwicklung sowohl der visuellen Repräsentationen der Namensvorkommen als auch der Interaktivität innerhalb der Karten stehen die Fragen und Anforderungen, die die Betrachter*innen aus verschiedenen Benutzergruppen mit verschiedenen Zielen und Interessen an die Visualisierung stellen. Die Verwendung der JavaScript-Bibliothek Leaflet erlaubt die nachhaltige Entwicklung von entsprechenden Visualisierungssidiomen mit frei verfügbarem Kartenmaterial und entlang der Erkenntnisse der wissenschaftlichen Forschung im Bereich der Informationsvisualisierung. Auf der Grundlage eines auf existierende Forschungsergebnisse der Informationsvisualisierung begründeten Mapping-Modells kann gezeigt werden, dass das für das interaktive Kartenmodul entwickelte Visualisierungs- und Interaktionsidiom im Hinblick auf die Prinzipien der Expressivität und der Effektivität den Ansprüchen an eine qualitativ gute Visualisierung genügen kann. Das schließt insbesondere die Überprüfung der vom DFD-Kartierungstool übernommenen Visualisierung der Namensvorkommen mit Kreisen im Hinblick auf ihre Benutzerfreundlichkeit mit ein.

Die Evaluation der für das interaktive Kartenmodul entwickelten Namensverbreitungskarte entlang der Erkenntnisse der Visualisierungsforschung offenbart jedoch auch grundsätzliche Defizite wie die fehlende Angabe von Geltungsbereichen für die Datenwerte und die nicht mit

den Wahrnehmungsprozessen des menschlichen Sehapparats konformen Größenberechnungen der Kreise, beides Defizite, die in der Abhängigkeit von der Datenlage und der Software des DFD-Kartierungstools begründet sind. Erste Studien interaktiver historischer Namensverbreitungskarten für das DFD wurden unabhängig vom DFD-Kartierungstool mit einer eigens aufgebauten Datenbank und neu entwickeltem JavaScript-Code erstellt. Anhand der Visualisierungsideome für diese Kartenstudien können zwar Lösungsmöglichkeiten für die oben genannten Defizite aufgezeigt werden. Eine unmittelbare Übertragbarkeit der entsprechenden Implementierungen auf das Visualisierungsideom der Karte im interaktiven Kartenmodul wäre jedoch nur durch die Überarbeitung der Datenbasis des DFD-Kartierungstools im einen Fall bzw. durch die vollständige Lösung der Software des interaktiven Kartenmoduls von der des DFD-Kartierungstools im anderen Fall möglich. Dies würde einen erheblichen Mehraufwand in der Entwicklungsarbeit bedeuten. Etliche von den Benutzer*innen des DFD angeregte zukünftige Weiterentwicklungen des interaktiven Kartenmoduls können dagegen ohne die im DFD-Kartierungstool begründeten Beeinträchtigungen erfolgen.

Mit dem Fortschreiten der Entwicklung des interaktiven Kartenmoduls und der historischen Kartierung für das DFD zeichnen sich indes immer deutlicher die verschiedenartigen Bedürfnisse und Anforderungen der einzelnen Benutzergruppen des DFD an die entwickelte Software ab. Den vergleichsweise moderaten Fragen und Aufgaben vonseiten der alltäglichen Nutzer*innen des DFD-Frontends, die bereits beim jetzigen Stand der Softwareentwicklungen zu einem großen Teil beantwortet bzw. erfüllt werden können, stehen spezialisierte herausforderndere Anforderungen aus der Gruppe der Artikelautor*innen entgegen. Überlegungen dazu, ob eine zukünftige Trennung der Nutzergruppen mit der Entwicklung von in Vokabular und Angebot unterschiedlichen Versionen der Anwendung sinnvoll erscheint, stehen aus.

A Wireframes

A.1 Interaktive Karte mit veränderten Parametern

Die Abbildung 1 zeigt die interaktive Verbreitungskarte für den Namen Richter nach dem Umschalten der Checkbox „Städte zusammenfassen“.

Die Abbildung 2 zeigt die interaktive Verbreitungskarte für den Namen Richter nach der Änderung der Parametereinstellung „PLZ-Stelligkeit“. Die Legende unten links zeigt die angepasste Statistik für die geänderte Darstellung der Verbreitung.

Die Abbildung 3 zeigt die interaktive Verbreitungskarte für den Namen Richter nach dem Ändern der farblichen Darstellung der Kreise.

Die Abbildung 4 zeigt die interaktive Verbreitungskarte für den Namen Richter nach dem Ändern der Transparenz der Innenflächen der Kreise.

A.2 Ortssuche

Die Abbildung 5 zeigt die Ortssuche in der interaktiven Verbreitungskarte für den Namen Richter mit alternativer Benutzeraktion.

A.3 Alternative Kartenstile

Die Abbildung 6 zeigt die interaktive Verbreitungskarte für den Namen Richter in dem alternativen Kartenstil Mapbox Light.

Die Abbildung 7 zeigt die interaktive Verbreitungskarte für den Namen Richter in dem alternativen Kartenstil CartoDB Positron in der Variante schwarz-weiß.

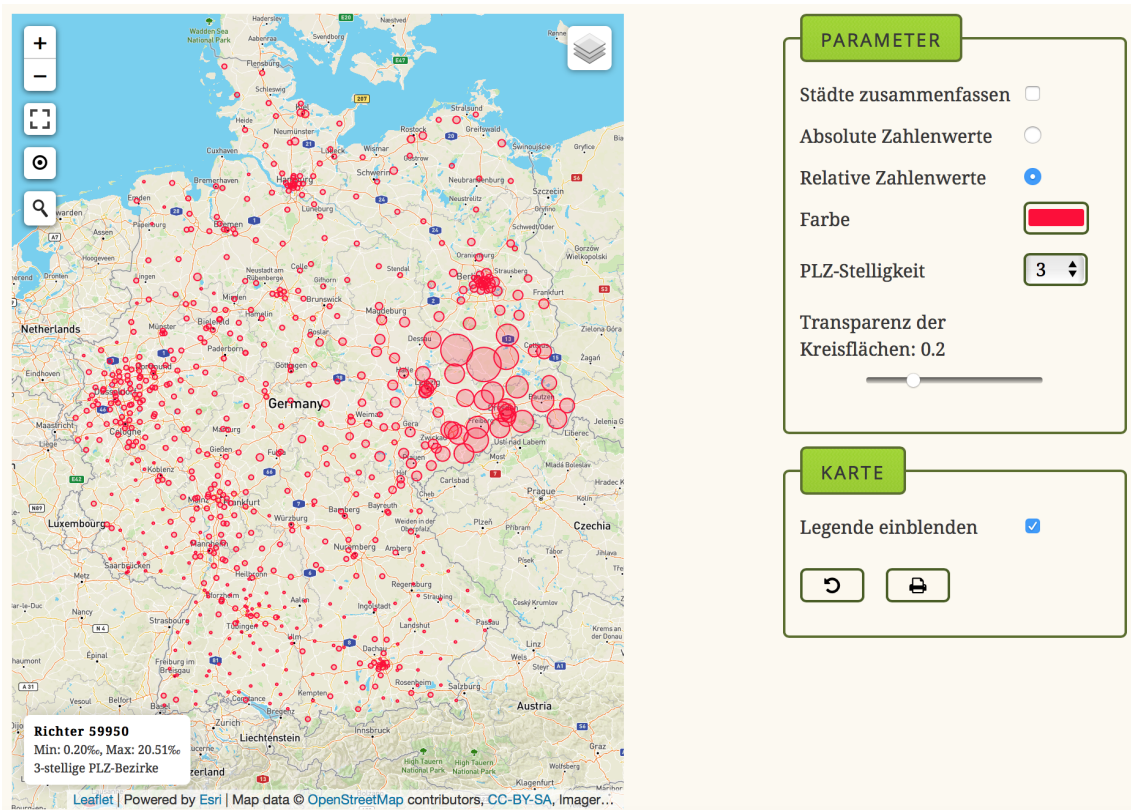


Abbildung 1: Interaktive Verbreitungskarte für den Namen Richter im Kartenstil Mapbox Streets nach der Änderung der Parametereinstellung „Städte zusammenfassen“

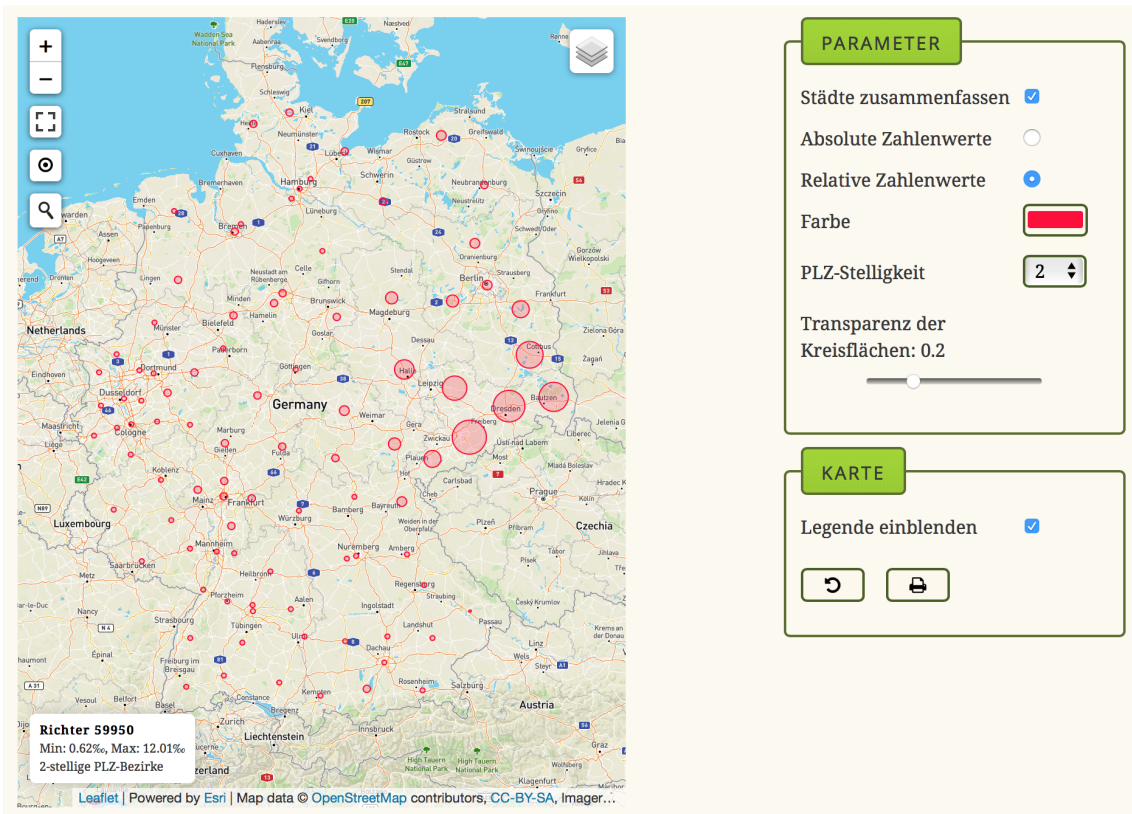


Abbildung 2: Interaktive Verbreitungskarte für den Namen Richter im Kartenstil Mapbox Streets nach der Änderung der Parametereinstellung „PLZ-Stelligkeit“ auf zweistellige Postleitzahlenbereiche

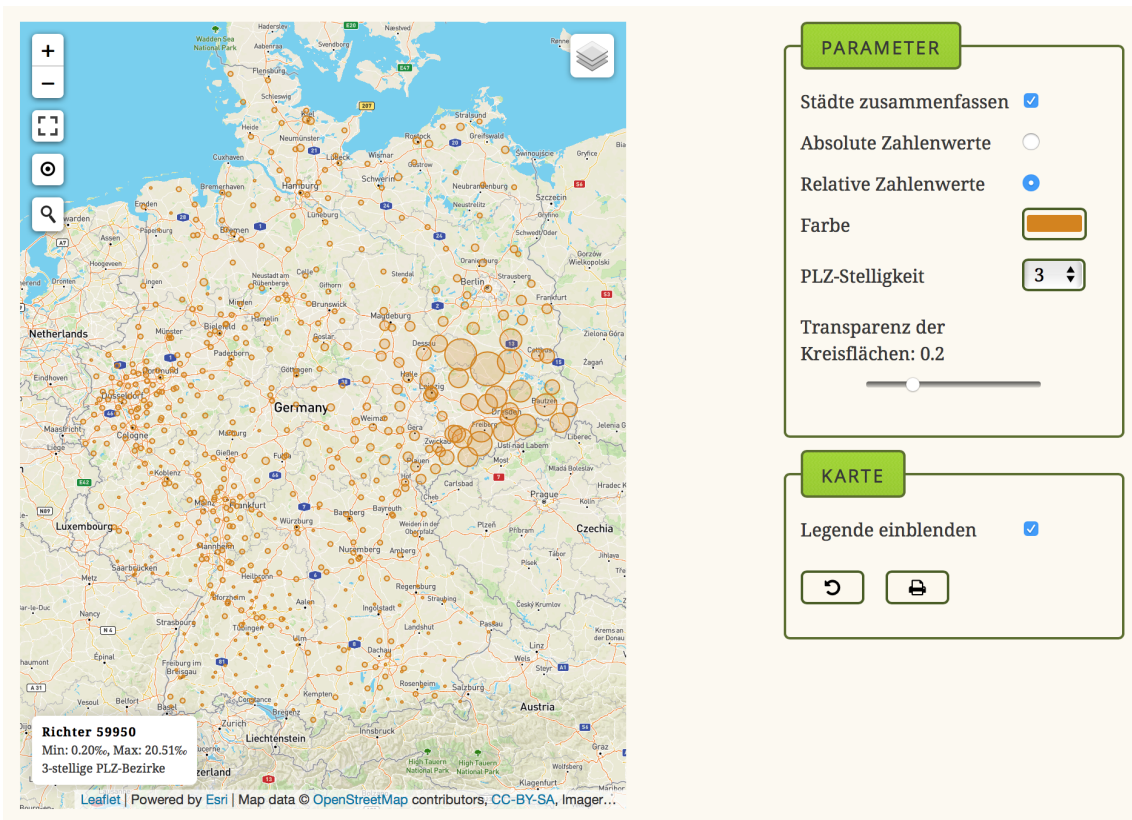


Abbildung 3: Interaktive Verbreitungskarte für den Namen Richter im Kartenstil Mapbox Streets nach dem Ändern der farblichen Darstellung der Kreise

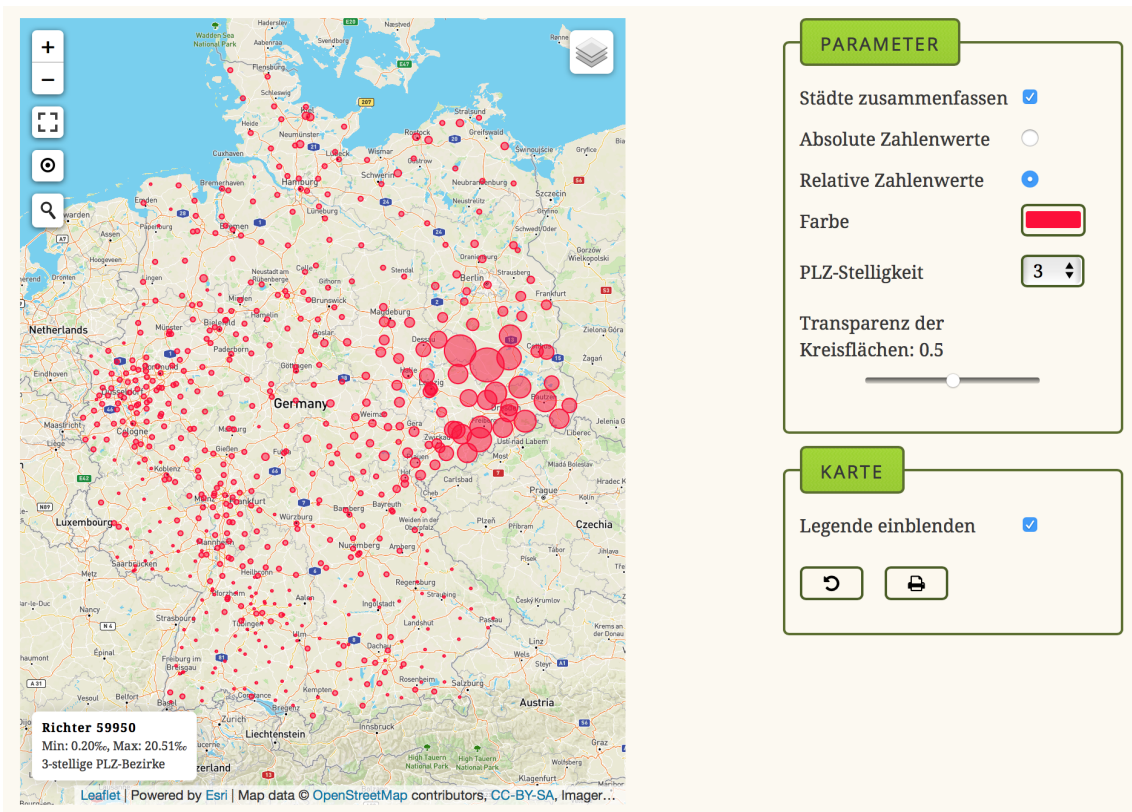


Abbildung 4: Interaktive Verbreitungskarte für den Namen Richter im Kartenstil Mapbox Streets nach dem Ändern der Transparenz der Innenflächen der Kreise

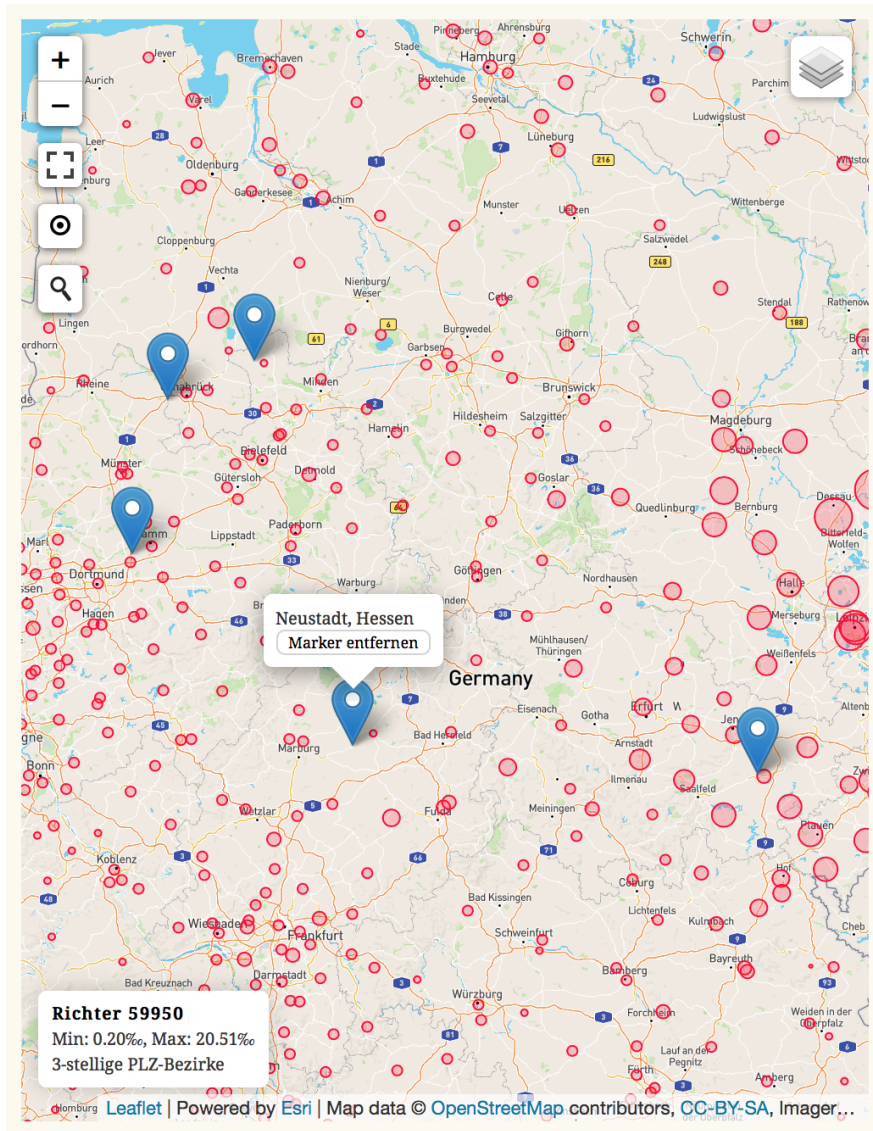


Abbildung 5: Ortssuche in der interaktiven Verbreitungskarte für den Namen Richter. Die Bestätigung der Ortssuche „Neustadt“ löst den Zoom in die Region um die Gesamtheit aller fünf Treffer der Ortssuche aus. An allen diesen Orten werden Marker gesetzt. Die Anwahl des Buttons „Marker entfernen“ in einem Popup über einem beliebigen Treffer bewirkt das Löschen aller in dieser Suche gesetzten Marker.

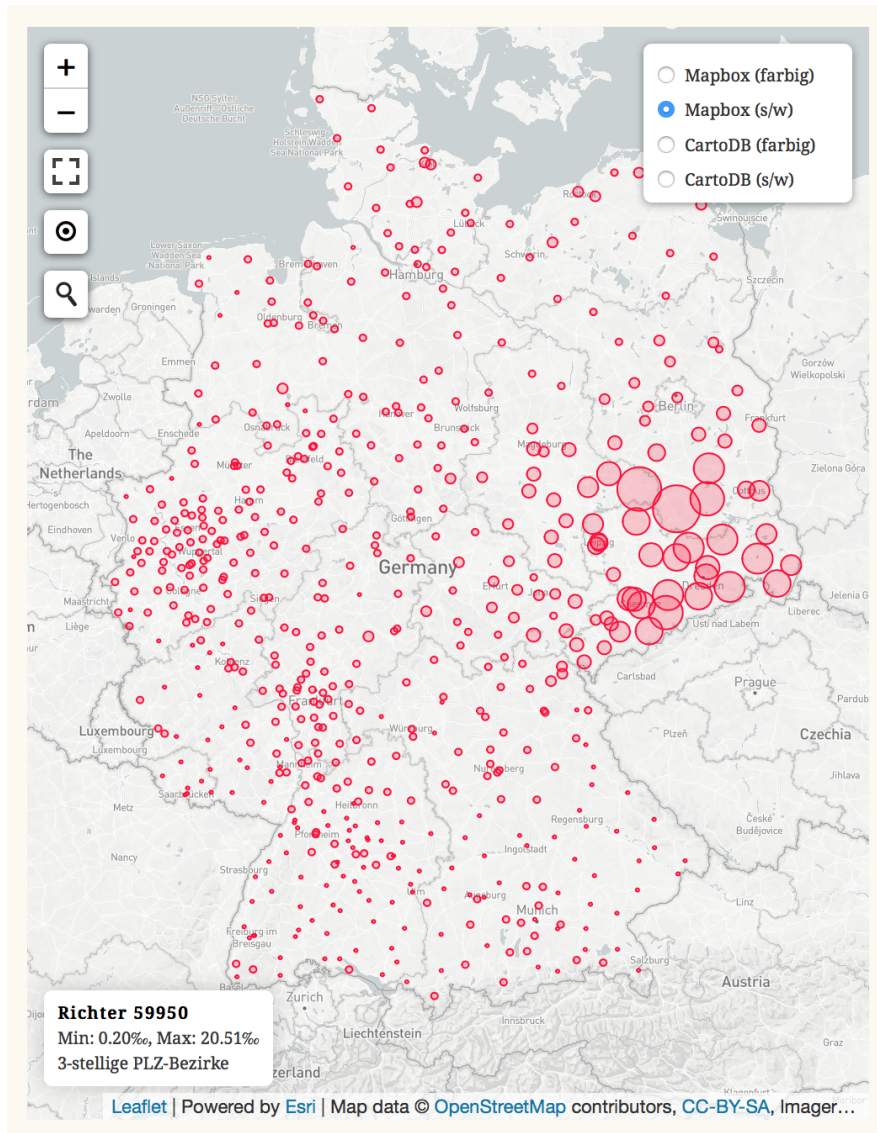


Abbildung 6: Interaktive Verbreitungskarte für den Namen Richter im Kartenstil Mapbox Light, der schwarz-weißen Variante des Kartenstils Mapbox Streets

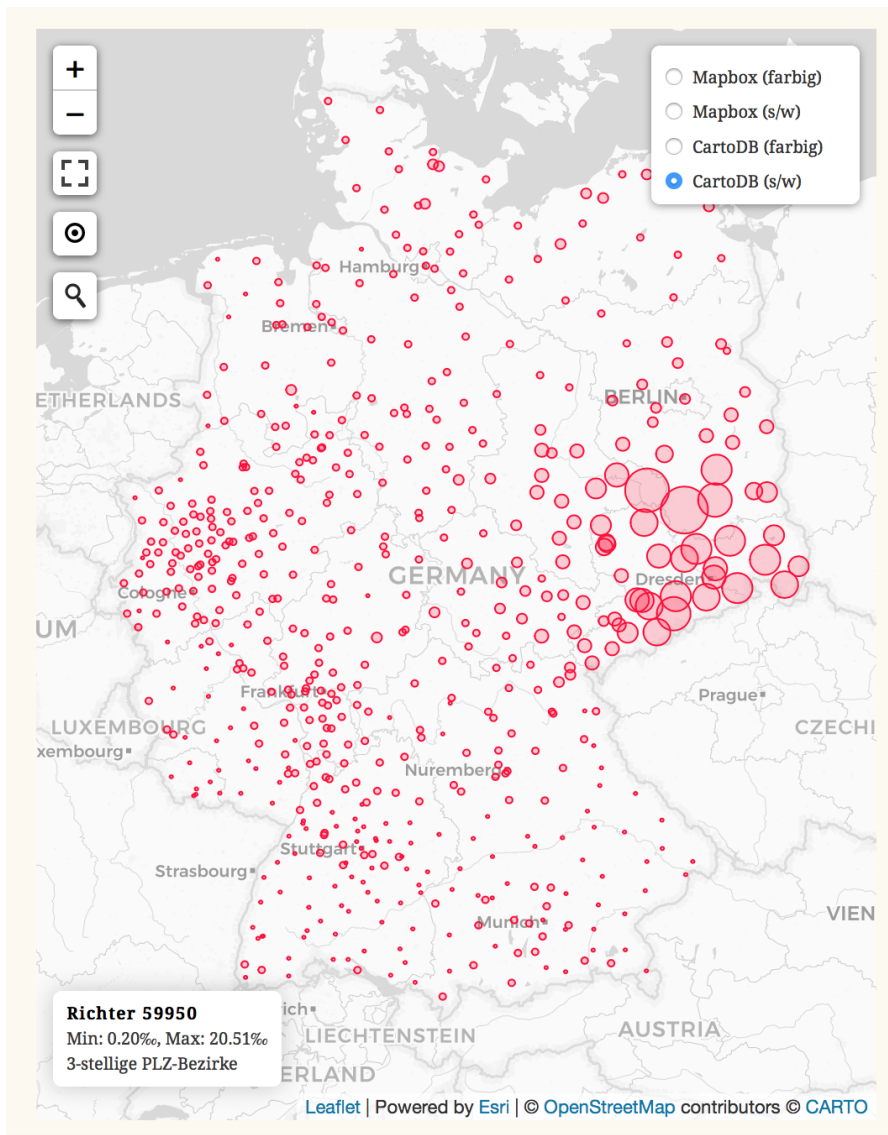


Abbildung 7: Interaktive Verbreitungskarte für den Namen Richter im Kartenstil CartoDB Positron in der Variante schwarz-weiß

B Interaktivität

In den Abschnitten B.1 bis B.4 wird die Implementierung der Interaktivität in den Interaktionsfeldern „Parameter“ und „Karte“ rechts neben der interaktiven Verbreitungskarte beschrieben. Die entsprechenden Codeausschnitte sind im Zusammenhang mit dem gesamten Code der Datei `karte.html` zu betrachten.

Die interaktive Änderung einer der Parameter „Städte zusammenfassen“, „Absolute Zahlenwerte“, „Relative Zahlenwerte“, „PLZ-Stelligkeit“ sowie „Farbe“ und „Farbsättigung der Kreisflächen“ erfordert jeweils die Erzeugung der Kreise für die Darstellung in der Karte. Dies geschieht innerhalb der Datei `karte.html` in der Callback-Funktion `circle` (s. Listing B.1).

Listing B.1: Auszug aus der Datei `karte.html` mit der Callback-Funktion `circle`

```
// Define a function to render the circles and reduce size on zoom in
// Bind an info popup to each circle and add the option to zoom into the map
var circle = function (latlng, size, msg) {
    var newSize = size * 6 / mymap.getZoom();
    {#console.log(size, newSize, mymap.getZoom());#}
    L.circle(latlng, newSize, {
        weight: 1,
        color: currentColor,
        fillColor: currentColor,
        fillOpacity: currentOpacity
    }).addTo(group).bindPopup(msg + "<br><button class='btn'
        onclick=zoom(" + latlng + ")>Zoomen</button>");
};
```

Die in den Abschnitten B.1 und B.2 beschriebenen Implementierungen der Interaktionen „Städte zusammenfassen“, „Absolute Zahlenwerte“, „Relative Zahlenwerte“ und „PLZ-Stelligkeit“ umfassen zusätzlich die Aktualisierung des HTML-Inhalts der Legende. Dies geschieht innerhalb der Datei `karte.html` in der Callback-Funktion `legend` (s. Listing B.2).

Listing B.2: Auszug aus der Datei `karte.html` mit der Callback-Funktion `legend`

```
// Update the legend based on feature properties passed
legend.update = function (min_max) {
    var content = '<h6><b>{{ name.label }} {{ name.frequency }}</b></h6>
        >';
    content += 'Min: ' + min_max.min + ', Max: ' + min_max.max;
    content += '<br>' + C.settings["piechartLayer.level"] + '-stellige
        PLZ-Bezirke';
    legend._div.innerHTML = content;
};
```

Die Implementierung der Interaktivität in dem Interaktionsfeld „Karte“, ebenfalls rechts neben der interaktiven Verbreitungskarte, wird in den Abschnitten B.5 und B.6 beschrieben.

Bei allen Codeauszügen in den Abschnitten B.1 bis B.6 handelt es sich um Auszüge aus der Callback-Funktion `info`, in der die interaktiv veränderten Parameter- und Karteneinstellungen aktualisiert werden.

B.1 Umschalten auf Absolute Zahlenwerte

Die Interaktionen „Städte zusammenfassen“, „Absolute Zahlenwerte“ und „Relative Zahlenwerte“ werden ähnlich implementiert. Die entsprechende Implementierung wird am Beispiel der Auswahl „Absolute Zahlenwerte“ gezeigt.

Die Interaktion erfolgt durch Umschalten über einen Radio-Button (bzw. im Falle der Interaktion „Städte zusammenfassen“ über eine Checkbox). Das Listing B.3 zeigt den HTML-Code für die Definition des Radio-Buttons für die Auswahl „Absolute Zahlenwerte“.

Listing B.3: Auszug aus der Datei `karte.html`, Radio-Button „Absolute Zahlenwerte“

```
<label for="absoluteStatistics">Absolute Zahlenwerte</label>
<input type="radio" id="absoluteStatistics" name="statistics" value="Absolute
    Zahlenwerte" checked>
```

Die Aktualisierung der Parametereinstellung in `C.settings["piechartLayer.relativeStatistics"]` in der Klasse `C` wird je nach der interaktiven Benutzerauswahl durch die Zuordnung der Boole'schen Variablen `True` oder `False` vorgenommen. Mit dem Aufruf der Funktion `renderCircles` in `circles.js` wird der Kreise-Layer entsprechend neu berechnet und die Darstellung der Karte aktualisiert. Der Reset-Button für die Rückkehr zur Default-Karte wird aktiviert. Das Listing B.4 zeigt einen Auszug aus der Datei `karte.html` mit dem JavaScript Event-Handler für die Auswahl „Absolute Zahlenwerte“.

Listing B.4: Auszug aus der Datei `karte.html`, Auswahl „Absolute Zahlenwerte“

```
document.getElementById("absoluteStatistics").addEventListener("change"
, function (e) {
    document.getElementById("default").disabled = false;
    C.settings["piechartLayer.relativeStatistics"] = (C.settings["
        piechartLayer.relativeStatistics"] === "false") ? "true" : "
        false";
    group.remove();
    group = L.layerGroup().addTo(mymap);
    C.renderCircles();
});
```

B.2 Änderung der PLZ-Stelligkeit

Die interaktive Änderung der PLZ-Stelligkeit erfolgt über ein Dropdown-Menü. Das Listing B.5 zeigt den HTML-Code mit der Definition des Dropdown-Menüs für die Auswahl „PLZ-Stelligkeit“.

Listing B.5: Auszug aus der Datei `karte.html`, Dropdown-Menü „PLZ-Stelligkeit“

```
<label for="level">PLZ-Stelligkeit</label>
    <select id="level" aria-label="level">
        <option value="1">1</option>
        <option value="2">2</option>
        <option value="3">3</option>
        <option value="5">5</option>
    </select>
```

Die entsprechende Änderung der Darstellung in der Karte wird wie in dem im vorhergehenden Abschnitt beschriebenen Beispiel implementiert. Da es sich bei der Benutzerauswahl im Dropdown-Menü nicht um die Wahl zwischen lediglich zwei alternativen Zuordnungen zu den `settings` in der Klasse `C` handelt, erfolgt die Aktualisierung des Parameterwertes anstatt über Boole'sche Variablen über die Zuordnung von `this.value` je nach der konkreten Nutzerauswahl. Das Listing B.6 zeigt einen Auszug aus der Datei `karte.html` mit dem JavaScript Event-Handler für die Auswahl der „PLZ-Stelligkeit“.

Listing B.6: Auszug aus der Datei `karte.html`, Auswahl „PLZ-Stelligkeit“

```
document.getElementById("level").addEventListener("change", function (e) {
    document.getElementById("default").disabled = false;
    C.settings["piechartLayer.level"] = this.value;
    group.remove();
    group = L.layerGroup().addTo(mymap);
    C.renderCircles();
});
```

B.3 Änderung der Farbe

Die interaktive Änderung der Farbe erfolgt über den durch die Verwendung von `input type="color"` zur Verfügung gestellten *color picker*. Das Listing B.7 zeigt den HTML-Code für die Definition der Auswahlbox für die „Farbe“ der Kreise.

Listing B.7: Auszug aus der Datei `karte.html`, Auswahlbox „Farbe“

```
<label for="color">Farbe</label>
<input type="color" id="color">
```

Die Implementierung erfolgt analog zu der Implementierung der PLZ-Stelligkeit mit dem Unterschied, dass hier der ausgewählte Farbwert nicht unmittelbar den `settings` in `C` zugeordnet wird. Stattdessen wird der zuvor in der Callback-Funktion `circle` in `karte.html`

den Kreisen zugeordnete Farbwert `currentColor` durch den Wert `this.value` ersetzt. Das Listing B.8 zeigt den Javascript Event-Handler für die Auswahl der „Farbe“.

Listing B.8: Auszug aus der Datei `karte.html`, Auswahl „Farbe“

```
document.getElementById("color").addEventListener("change", function (e) {
    document.getElementById("default").disabled = false;
    currentColor = this.value;
    group.remove();
    group = L.layerGroup().addTo(mymap);
    C.renderCircles();
});
```

B.4 Änderung der Farbsättigung der Kreisflächen

Die interaktive Änderung der Farbsättigung der Kreisflächen erfolgt über einen Schieber. Das Listing B.9 zeigt den HTML-Code für die Definition des Schiebers für die Auswahl der „Farbsättigung“.

Listing B.9: Auszug aus der Datei `karte.html`, Schieber „Farbsättigung“

```
<label for="opacity">Transparenz der Kreisflächen: <span id="choice"></span></label>
    <div class="slidecontainer" style="display: flex; align-items: center; justify-content: center;">
        <input type="range" id="range" min="0" max="1" step="0.1" value="0.2">
    </div>
```

Die Implementierung erfolgt analog zu der Implementierung der Farbänderung. Das Listing B.10 zeigt den Javascript Event-Handler für die Auswahl der „Farbsättigung“.

Listing B.10: Auszug aus der Datei `karte.html`, Auswahl „Farbsättigung“

```
document.getElementById("range").addEventListener("input", function (e) {
    document.getElementById("default").disabled = false;
    currentOpacity = this.value;    ///this.max
    group.remove();
    group = L.layerGroup().addTo(mymap);
    C.renderCircles();
});
```

B.5 Legende ein- und ausblenden

Das interaktive Ein- und Ausschalten der Legende erfolgt über eine Checkbox. Das Listing B.11 zeigt den HTML-Code für die Definition der Checkbox für die Auswahl „Legende einblenden“.

Listing B.11: Auszug aus der Datei `karte.html`, Checkbox „Legende einblenden“

```
<label for="showLegend">Legende einblenden</label>
<input type="checkbox" id="showLegend" checked>
```

Je nach Schaltung der Checkbox wird die Sichtbarkeit der Legende mithilfe der Stylesheet-Sprache CSS auf `visible` bzw. `hidden` gesetzt. Das Listing B.12 zeigt den Javascript Event-Handler für die Auswahl „Legende einblenden“.

Listing B.12: Auszug aus der Datei `karte.html`, Auswahl „Legende einblenden“

```
document.getElementById("showLegend").addEventListener("change",
    function (e) {
        if(e.target.checked === true) {
            legend._div.style.visibility = 'visible';
        }
        else {
            legend._div.style.visibility = 'hidden';
        }
    });
```

B.6 Parametereinstellungen zurücksetzen

Die interaktive Rückkehr zur Default-Karte erfolgt über einen Button. Das Listing B.13 zeigt den HTML-Code für die Definition des Buttons für das Zurücksetzen der Parametereinstellungen.

Listing B.13: Auszug aus der Datei `karte.html`, Reset-Button

```
<button class="mapbutton" id="default" title="Parametereinstellungen zurücksetzen" disabled><span class="fa fa-undo" aria-hidden="true"></span></button>
<button class="mapbutton" id="download"><span class="fa fa-print" aria-hidden="true"></span></button>
```

Mit jeder interaktiven Änderung einer der Parametereinstellungen in dem Interaktionsfeld „Parameter“ wird, wie erwähnt, der Reset-Button mit dem Tooltip „Parametereinstellungen zurücksetzen“ aktiviert. Die Betätigung des Buttons bewirkt das Zurücksetzen aller Parameter auf die Default-Einstellungen.¹ Das Listing B.14 zeigt den Javascript Event-Handler für das Zurücksetzen der Parametereinstellungen.

Listing B.14: Auszug aus der Datei `karte.html`, Reset der Parameter

```
document.getElementById("default").addEventListener("click", function (e) {
    document.getElementById("default").disabled = true;

    // Set the parameters to default
    C.settings["piechartLayer.mergeCities"] = defaultSettings['mergeCities'];
    C.settings["piechartLayer.relativeStatistics"] = defaultSettings['relativeStatistics'];
});
```

¹Das für den Reset-Button verwendete Icon stammt von <https://fontawesome.com/icons>.

```
C.settings["piechartLayer.level"] = defaultSettings['level'];

// Set the controls
document.getElementById("mergeCities").checked = (defaultSettings["
mergeCities"] === "true");
document.getElementById("absoluteStatistics").checked = (
defaultSettings["relativeStatistics"] === "false");
document.getElementById("relativeStatistics").checked = (
defaultSettings["relativeStatistics"] === "true");
document.getElementById("level").value = defaultSettings["level"];

// Reset the color
currentColor = defaultSettings["color"];
document.getElementById("color").value = currentColor;

// Reset the opacity
currentOpacity = defaultSettings["fillOpacity"];
document.getElementById("range").value = 0.2; //currentOpacity
output.innerHTML = 0.2;

// Render the circles
group.remove();
group = L.layerGroup().addTo(mymap);
C.renderCircles();
});
```

Literatur

- [1] Gui Bonsiepe. *Interface – Design neu begreifen*. Kommunikation & neue Medien. Bollmann, 1996. ISBN: 9783927901841.
- [2] Stuart K. Card, Jock D. Mackinlay und Ben Shneiderman. „Information Visualization“. In: *Readings in Information Visualization: Using Vision to Think*. Hrsg. von Stuart K. Card, Jock D. Mackinlay und Ben Shneiderman. Morgan Kaufmann Pub., 1999, S. 1–34.
- [3] William S. Cleveland und Robert McGill. „Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods“. In: *Journal of the American Statistical Association* 79.387 (1984), S. 531–554.
- [4] Markus Dahm. *Grundlagen der Mensch-Computer-Interaktion*. Pearson Studium, 2006.
- [5] Kyran Dale. *Data Visualization with Python and Javascript*. O’Reilly Media Inc., 2016.
- [6] Projekt Digitales Familiennamenwörterbuch Deutschlands DFD. *Arbeitsbericht 2012-2014*. Techn. Ber. Akademie der Wissenschaften und der Literatur Mainz, 2014.
- [7] Rae A. Earshaw. „Illuminating the Path: the Research and Development Agenda for Visual Analytics“. In: IEEE Computer Society, 2005, S. 69–104. ISBN: 0769523234.
- [8] W. Amaru Flores Flores und Peter Gilles. „Die Verlustlisten des Ersten Weltkriegs als historisches namengeographisches Korpus“. In: *Beiträge zur Namenforschung* 55.2-3 (2020), S. 127–167. URL: <https://bnf.winter-verlag.de/article/BNF/2020/2-3/5> (besucht am 02. 10. 2021).
- [9] Ben Fry. *Visualizing Data*. O’Reilly Media Inc., 2008.
- [10] Robert B. Haber und David A. McNabb. „Visualization Idioms: A Conceptual Model for Scientific Visualization Systems“. In: *Visualization in Scientific Computing*. Hrsg. von G. M. Nielson, B. Shriver und L.J. Rosenblum. IEEE Computer Society Press, 1990, S. 74–93.
- [11] Steve Krug. *Don’t make me think! – Web & Mobile Usability*. 3rd Edition. mitp, 2014.
- [12] Patrick Lynch und Sarah Horton. *Web Style Guide: Foundations of User Experience Design*. 4th Edition. Yale University Press, 2016. URL: info.med.yale.edu/caim/manual/contents.html (besucht am 11.09.2021).
- [13] Alan M. MacEachren. *How Maps Work: Representation, Visualization, and Design*. Guilford Press, 1995.
- [14] Jock Mackinlay. „Automating the design of graphical presentations of relational information“. In: *ACM Transactions on Graphics* 5.2 (1986), S. 100–141.

-
-
- [15] Tamara Munzner. *Visualization Analysis and Design*. A K Peters Visualization Series. CRC Press, 2014.
 - [16] Louis Rosenfeld, Peter Morville und Jorge Arango. *Information Architecture for the Web and Beyond*. 4th Edition. O'Reilly, 2015.
 - [17] Heide Schumann und Wolfgang Müller. *Visualisierung – Grundlagen und allgemeine Methoden*. Springer, 2000.
 - [18] Ben Shneiderman. *Designing the user interface : strategies for effective human-computer interaction*. 3rd Edition. Addison-Wesley, 1998.
 - [19] Ben Shneiderman. „The Eyes Have It: A Task by Data Type Taxonomy for Information, Visualizations“. In: *Proceedings of the IEEE Conference on Visual Languages*. 1996, S. 336–343.
 - [20] Robert Spence. *Information Visualization: Design for Interaction*. 2nd Edition. Prentice Hall, 2007.
 - [21] Edward R. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, 1983.
 - [22] Colin Ware. *Information Visualization: Perception for Design*. 3rd Edition. Morgan Kaufmann, 2013.
 - [23] Colin Ware. *Visual Thinking for Design*. Morgan Kaufmann, 2008.
 - [24] Richard Saul Wurman. *Information Anxiety*. 1st Edition. Doubleday, 1989.
 - [25] Jesper Zedlitz. *Namensverbreitungskarten auf Basis der Verlustlisten des 1. Weltkriegs*. URL: <https://gitlab.genealogy.net/project/namensverbreitung> (besucht am 23.09.2021).

Abbildungsverzeichnis

2.1	Statische Verbreitungskarte im DFD	5
2.2	Entwicklungsprozess für das interaktive Kartenmodul	11
2.3	Das ontologische Designdiagramm nach Gui Bonsiepe	12
2.4	Story Card in der Entwicklungsmethode „scrum“	12
3.1	Site-Diagramm der Anwendung	16
3.2	Entity-Relationship-Diagramm	20
3.3	Datenflussdiagramm	22
3.4	Sequenzdiagramm	24
4.1	Dr. John Snows Cholera Karte	28
4.2	Charles Joseph Minards Karte zu Napoleons Russlandfeldzug	29
4.3	Datenfluss in der Visualisierungspipeline	32
4.4	Mapping-Modell von Tamara Munzner	37
4.5	Startseite des Demonstrators	43
4.6	Anwendungsseite mit Namenartikel und Vorschaukarte	44
4.7	Anwendungsseite mit interaktiver Verbreitungskarte	46
4.8	Umschalten auf „Absolute Zahlenwerte“	47
4.9	Druckversion der Verbreitungskarte	49
4.10	Zoom in die Verbreitungskarte	50
4.11	Ortssuche für den Namen Richter: Schritt 1	52
4.12	Ortssuche für den Namen Richter: Schritt 2	53
4.13	Alternativer Kartenstil CartoDB Positron	54
4.14	Informationstext und Zoom in die Region	55
6.1	Entity-Relationship-Diagramm	66
6.2	Historische Verbreitungskarte Nowak 1890 und 1996	71
6.3	Historische Verbreitungskarte Namengruppen 1890	73
6.4	Historische Verbreitungskarte Naujoks, Naujokat 1890 und 1996	74
6.5	Interaktive Verbreitungskarte mit einblendbaren PLZ-Bereichen	76
6.6	Choropleth-Karte	78
6.7	Heatmap für den Namen Nowak im Jahr 1890	81
6.8	Heatmap für den Namen Nowak im Jahr 1996	81
6.9	Interaktives Zeichnen von Polygonen	83
7.1	Abweichende Geokoordinaten deutscher Städte	88

1	Änderung der Parametereinstellung „Städte zusammenfassen“	96
2	Änderung der PLZ-Stelligkeit	97
3	Änderung der Kreisfarbe	98
4	Änderung der Transparenz der Kreisflächen	99
5	Ortssuche in der Verbreitungskarte mit alternativer Benutzeraktion	100
6	Alternativer Kartenstil Mapbox Light	101
7	Alternativer Kartenstil CartoDB Positron schwarz-weiß	102

Tabellenverzeichnis

3.1	URL-Design der Anwendung	17
3.2	Knoten im Organisationsschema	20
4.1	Datenmengentyp Tabelle	33
4.2	Mapping-Modell für die interaktiven Verbreitungskarten im DFD	37

Listings

3.1	Auszug aus der Datei plz_5.csv	18
3.2	Auszug aus der Datei plz_3.csv	18
3.3	Auszug aus der Datei karte.csv für den Namen Richter	18
3.4	Auszug aus der Datei karte.properties für den Namen Richter	19
6.1	Auszug aus der Datei familyname2coordinate_1890	64
6.2	Auszug aus der Datei familyname2coordinate_1996	64
6.3	Auszug aus der Datei Nowak_1890.js	65
6.4	Auszug aus dem Ergebnis der PostgreSQL-Datenbankabfrage für den Namen Richter	67
6.5	Auszug aus der Datei Richter1996_centroids.js	67
6.6	Auszug aus der Datei Richter_plz3.js	68
B.1	Auszug aus der Datei karte.html mit der Callback-Funktion circle . . .	103
B.2	Auszug aus der Datei karte.html mit der Callback-Funktion legend . . .	103
B.3	Auszug aus der Datei karte.html, Radio-Button „Absolute Zahlenwerte“ .	104
B.4	Auszug aus der Datei karte.html, Auswahl „Absolute Zahlenwerte“ . . .	104
B.5	Auszug aus der Datei karte.html, Dropdown-Menü „PLZ-Stelligkeit“ . . .	105
B.6	Auszug aus der Datei karte.html, Auswahl „PLZ-Stelligkeit“	105
B.7	Auszug aus der Datei karte.html, Auswahlbox „Farbe“	105
B.8	Auszug aus der Datei karte.html, Auswahl „Farbe“	106
B.9	Auszug aus der Datei karte.html, Schieber „Farbsättigung“	106
B.10	Auszug aus der Datei karte.html, Auswahl „Farbsättigung“	106
B.11	Auszug aus der Datei karte.html, Checkbox „Legende einblenden“	107
B.12	Auszug aus der Datei karte.html, Auswahl „Legende einblenden“	107
B.13	Auszug aus der Datei karte.html, Reset-Button	107
B.14	Auszug aus der Datei karte.html, Reset der Parameter	107