

Institut für Visualisierung und Interaktive Systeme

Universität Stuttgart  
Universitätsstraße 38  
D-70569 Stuttgart

Bachelorarbeit

**Interaktive, vergleichende  
Visualisierung von  
georeferenzierten, historischen  
Daten**

Heiko Roggenbuck

**Studiengang:** Informatik

**Prüfer/in:** Prof. Dr. Thomas Ertl

**Betreuer/in:** Dr. Steffen Koch,  
M.Sc. Markus John,  
M.Sc. Robert Krüger

**Beginn am:** 6. April 2017

**Beendet am:** 28. September 2017

**CR-Nummer:** H.5.2, J.5



## **Kurzfassung**

Die Darstellung historischer Daten auf analogen, statischen Karten ist oft schwierig und manueller Aufwand wird benötigt, um verwendbare Ergebnisse zu erzielen. Dieses Vorgehen ist dann besonders problematisch, wenn eine Vielzahl komplexer Daten dargestellt werden soll.

Diese Bachelorarbeit konzentriert sich auf die interaktive Visualisierung von georeferenzierten, historischen Daten. Durch die Verwendung digitaler, interaktiver Karten und geeigneter Visualisierungstechniken und durch die Anwendung eines Multiple Coordinated Views-Ansatzes können Systeme entwickelt werden, welche die Darstellung, Erforschung und Analyse solcher Daten vereinfachen. Im Rahmen dieser Arbeit wurde ein entsprechender Ansatz entwickelt und grundlegende Designentscheidungen gegen alternative Umsetzungsmöglichkeiten abgewogen. Die Ergebnisse resultieren in einer prototypischen Implementierung des Ansatzes, welcher konzeptionell und detailliert dargestellt wird.

Der entwickelte Prototyp besteht aus verschiedenen Ansichten, welche jeweils einen Aspekt des zugrundeliegenden Datensatzes visualisieren und Interaktionsmöglichkeiten für die Selektion und Filterung der Daten zur Verfügung stellen. Durch die Verknüpfung der verschiedenen Ansichten können komplexe Selektionen getätigt werden, wodurch verborgene Eigenschaften des Datensatzes aufgedeckt und der Fokus auf die wichtigen Aspekte der Daten gelegt werden kann. Um die übersichtliche Darstellung der Daten zu gewährleisten, werden geeignete Aggregationsverfahren für die Strukturierung und Zusammenfassung der Daten verwendet. Mittels einer Expertenbefragung und der Diskussion ausgewählter Anwendungsszenarien wurde der entwickelte Ansatz evaluiert.



# Inhaltsverzeichnis

<b>1 Einführung</b>	<b>11</b>
1.1 Motivation . . . . .	11
1.2 Zielsetzung . . . . .	11
<b>2 Grundlagen</b>	<b>13</b>
2.1 Grundlagen der Visualisierung . . . . .	13
2.2 Grundlagen der Geovisualisierung . . . . .	14
2.3 Multiple Coordinated Views . . . . .	17
2.3.1 Einführung . . . . .	17
2.3.2 Definitionen und Begriffe . . . . .	18
2.3.3 Entwicklung . . . . .	18
2.3.4 Richtlinien/Regeln . . . . .	20
<b>3 Verwandte Arbeiten</b>	<b>23</b>
3.1 Multiple Coordinated Views und Geovisualisierung . . . . .	23
3.1.1 City’O’Scope . . . . .	23
3.1.2 VAiRoma . . . . .	25
3.1.3 VisGets . . . . .	27
3.1.4 DC Crime Visualization und ähnliche Ansätze . . . . .	29
3.1.5 GeoTemCo . . . . .	33
3.1.6 Threatwiki . . . . .	35
<b>4 Konzept</b>	<b>39</b>
4.1 Anforderungen . . . . .	39
4.1.1 Verschiedene Ansichten . . . . .	39
4.1.2 Visualisierung der Daten . . . . .	39
4.1.3 Interaktion . . . . .	40
4.1.4 Benutzeroberfläche . . . . .	40
4.2 Entwurf . . . . .	40
4.2.1 Datensätze . . . . .	40
4.2.2 Benutzeroberfläche . . . . .	42
<b>5 Umsetzung</b>	<b>45</b>
5.1 Übersicht und Technologie . . . . .	45
5.2 Aggregationsprozedur . . . . .	46
5.3 Ansichten . . . . .	47
5.3.1 Kartenansicht . . . . .	47
5.3.2 Quellenansicht . . . . .	55

5.3.3	Gruppenansicht . . . . .	57
5.3.4	Zeitleistenansicht . . . . .	61
5.3.5	Verbindungen . . . . .	67
<b>6</b>	<b>Ergebnisse</b>	<b>69</b>
6.1	Anwendungsszenarien . . . . .	69
6.1.1	Untersuchung einer Quelle . . . . .	69
6.1.2	Untersuchung einer Religion . . . . .	71
6.1.3	Untersuchung einer Zeitspanne . . . . .	72
6.1.4	Untersuchung einer Region . . . . .	74
6.2	Expertenbefragung . . . . .	75
6.2.1	Allgemeines . . . . .	76
6.2.2	Kartenansicht . . . . .	77
6.2.3	Quellenansicht . . . . .	81
6.2.4	Gruppenansicht . . . . .	81
6.2.5	Zeitleistenansicht . . . . .	81
6.3	Diskussion . . . . .	82
<b>7</b>	<b>Zusammenfassung und Ausblick</b>	<b>85</b>
7.1	Zusammenfassung . . . . .	85
7.2	Ausblick . . . . .	86
<b>Literaturverzeichnis</b>		<b>89</b>

# Abbildungsverzeichnis

---

2.1	Visuelle Variablen nach J. Bertin [Ber83] . . . . .	15
2.2	Information Visualization Pipeline [CR98] . . . . .	16
2.3	Dataflow Model [UFK <sup>+</sup> 89] [HM90] . . . . .	18
3.1	CityOScope [BG03] . . . . .	23
3.2	VAiRoma [CDW <sup>+</sup> 16] . . . . .	25
3.3	VisGets [DCCW08] . . . . .	27
3.4	truliaHindsight [RR09] . . . . .	29
3.5	AsthMap [RR09] . . . . .	30
3.6	SpatialKey [RR09] . . . . .	31
3.7	DC Crime Visualization [RR09] . . . . .	32
3.8	GeoTemCo [JHS13] . . . . .	33
3.9	Threatwiki . . . . .	36
4.1	Die linke Spalte, welche aus der Quellenansicht und der Gruppenansicht besteht, und die Zeitleistenansicht können jeweils mittels eines Buttons eingeklappt oder ausgeklappt werden, wobei sich die Kartenansicht gemäß des verfügbaren Platzes ausdehnt. . . . .	42
4.2	Entwurf der graphischen Benutzeroberfläche, bestehend aus der Quellenansicht (a), der Gruppenansicht (b), der Zeitleistenansicht (c) und der Kartenansicht (d). . . . .	43
5.1	Graphische Benutzeroberfläche, bestehend aus der Quellenansicht (a), der Gruppenansicht (b), der Kartenansicht (c) und der Zeitleistenansicht (d). . . . .	45
5.2	Visualisierung der Aktivitäten verschiedener religiöser Gruppen innerhalb der Kartenansicht . . . . .	48
5.3	Graphische Baumstrukturen für die Repräsentation einzelner, nicht aggregierter Datenobjekte . . . . .	48
5.4	Graphische Baumstrukturen für die Repräsentation aggregierter Datenobjekte . . . . .	49
5.5	Ein Fenster veranschaulicht die Zusammensetzung eines Baumknotens, welcher mehrere, aggregierte Datenobjekte repräsentiert. . . . .	50
5.6	Rote Kreise visualisieren die Koordinaten der Städte, welche von den aggregierten Datenobjekten referenziert werden. . . . .	51
5.7	Jedes Mal, wenn ein weiterer Punkt zu dem Polygonzug hinzugefügt wird, wird dieser neue Punkt automatisch mit dem ersten und dem letzten hinzugefügten Punkt verbunden, wodurch der Polygonzug zu jeder Zeit geschlossen ist. . . . .	52
5.8	Ergebnis der geographischen Selektion innerhalb der Kartenansicht . . . . .	53

5.9	Knoten, welche einzelne Datenobjekte repräsentieren, im selektierten (links) und nicht selektierten (rechts) Zustand . . . . .	54
5.10	Kuchen- und Donutstücke visualisieren selektierte und nicht selektierte Anteile von Knoten, welche mehrere, aggregierte Datenobjekte repräsentieren. . . . .	54
5.11	Quellenansicht, bestehend aus einer Liste aller Quellen und einem Abschnitt für die Anzeige des Textes einer ausgewählten Quelle . . . . .	55
5.12	Für jede Quelle visualisiert ein Prozentbalken den selektierten Anteil. Wird ein Knoten innerhalb der Kartenansicht ausgewählt, so werden alle Quellen, welche in den durch den Knoten repräsentierten Datenobjekten referenziert werden, mittels pinker Rechtecke hervorgehoben. . . . .	56
5.13	Gruppenansicht . . . . .	58
5.14	Der Inhalt deselekterter Knoten ist schwarz. . . . .	59
5.15	Werden Teilbäume eingeklappt, so dehnt sich der restliche Baum gemäß des verfügbaren Platzes aus. Ein eingeklappter Teilbaum wird durch einen kleinen, farbigen Punkt rechts neben dem entsprechenden Knoten angedeutet. . . . .	60
5.16	Für jede Gruppe visualisiert ein Prozentbalken den selektierten Anteil. Wird ein Knoten innerhalb der Kartenansicht ausgewählt, so werden die Gruppen, welche in den durch diesen Knoten repräsentierten Datenobjekten referenziert werden, durch pinke Umrundungen hervorgehoben. . . . .	61
5.17	Wird innerhalb eines selektierten Teilbaumes ein Teilbaum ausgewählt, so wird die Selektion verfeinert. Links ist der an dem Knoten 'CHR' hängende Teilbaum selektiert. Durch die Auswahl des Knotens 'CHAL' wird die Selektion dahingehend verfeinert, dass nicht mehr der gesamte an 'CHR' hängende Teilbaum selektiert ist, sondern nur noch der an 'CHAL' hängende Teilbaum. . . . .	62
5.18	Wird ein Teilbaum ausgewählt, welcher bereits teilweise selektiert ist, so wird die Selektion auf den gesamten Teilbaum ausgeweitet. Links sind nur die Knoten 'MELK', 'MARO' und 'SYR' selektiert. Wird nun der Knoten 'CHR' ausgewählt, wird der gesamte daran hängende Teilbaum selektiert (rechts). . . . .	62
5.19	Zeitleistenansicht . . . . .	63
5.20	Zeitleistenansicht mit selektierten und nicht selektierten Abschnitten . . . . .	63
5.21	Die Selektion wird durch ein transparentes Rechteck über den gestapelten Säulen visualisiert. Wird ein Knoten innerhalb der Kartenansicht ausgewählt, so werden die in den durch den Knoten repräsentierten Datenobjekten referenzierten Zeitspannen mittels transparenter, pinker Rechtecke über der Beschriftung der x-Achse hervorgehoben. . . . .	64
5.22	Überschneidet sich die durch ein Datenobjekt referenzierte Zeitspanne (a, b, c, d in 1, 2) mit der selektierten Zeitspanne (schwarzes Rechteck), so wird das Datenobjekt in die Selektion aufgenommen. Gibt es keine Überschneidung (e, f in 3), so wird das Datenobjekt nicht aufgenommen. . . . .	65
6.1	Alle Quellen, bis auf Quelle 5, wurden deselektert. . . . .	69
6.2	Islamische Gruppen sind stärker vertreten als christliche oder jüdische Gruppen. . . . .	70
6.3	Drei Selektionszustände: Erst sind alle christlichen, dann alle islamischen und schließlich alle jüdischen Gruppen selektiert. . . . .	71
6.4	Religionen, beschrieben durch Quelle 5, im zeitlichen Verlauf . . . . .	71

6.5 Geographische Verteilung religiöser Gruppen, welche durch Quelle 5 Beschrieben werden, auf niedriger Zoomstufe . . . . .	72
6.6 Geographische Verteilung religiöser Gruppen, welche durch Quelle 5 Beschrieben werden, auf höherer Zoomstufe . . . . .	73
6.7 Quellenansicht, während lediglich der jüdische Teilbaum in der Gruppenansicht selektiert ist . . . . .	73
6.8 Die jüdische Religion im zeitlichen Verlauf . . . . .	74
6.9 Einzelne, jüdische Gruppen im zeitlichen Verlauf . . . . .	74
6.10 Die Zeitspanne zwischen 600 und 700 n. Chr. ist in der Zeitleistenansicht selektiert. . . . .	75
6.11 Quellenansicht, während die Zeitspanne zwischen 600 und 700 n. Chr. in der Zeitleistenansicht selektiert ist . . . . .	75
6.12 In der Gruppenansicht wurden alle drei Religionsäste eingeklappt. . . . .	76
6.13 Gruppenansicht, während die Zeitspanne zwischen 600 und 700 n. Chr. in der Zeitleistenansicht selektiert ist . . . . .	77
6.14 Kartenansicht auf niedriger Zoomstufe, während die Zeitspanne zwischen 600 und 700 n. Chr. in der Zeitleistenansicht selektiert ist . . . . .	78
6.15 Eine geographische Region wurde durch das Zeichnen eines Polygons selektiert. . . . .	79
6.16 Quellenansicht, während eine geographische Region in der Kartenansicht selektiert ist . . . . .	79
6.17 Gruppenansicht, während eine geographische Region selektiert ist . . . . .	80



# 1 Einführung

## 1.1 Motivation

Geografische Karten eignen sich gut, um historische Daten anschaulich darzustellen. Dabei können verschiedenste Informationen mittels verschiedener, graphischer Darstellungsformen kodiert und visualisiert werden. Die Verwendung von analogen, statischen Karten führt allerdings in einigen Szenarien und Anwendungsfällen zu Problemen. Die Exploration der Daten kann hierbei nicht durch Interaktionsmöglichkeiten unterstützt werden. Selektion und Filterung der Daten kann dabei lediglich kognitiv erfolgen. Somit müssen komplexe und große Datenmengen statisch dargestellt werden, ohne, dass Informationen verloren gehen, die Übersichtlichkeit allerdings gewahrt bleibt und die Karte nicht zu sehr verdeckt wird (visual clutter).

Um die genannten Limitierungen zu überwinden, können digitale, interaktive geographische Karten verwendet werden. Durch die Zoomfunktionalität muss lediglich entschieden werden, welche Informationen in welchem Detailgrad auf welcher Zoomstufe visualisiert werden. Auf niedrigen Zoomstufen wird die Menge von komplexen Informationen geeignet zusammengefasst und abstrahiert, während höhere Zoomstufen eine detailliertere Darstellung der einzelnen Daten ermöglichen. Somit wird der Verlust an dargestellten Informationen auf ein Minimum reduziert. Durch die digitale Visualisierung von Informationen können geeignete Werkzeuge entwickelt werden, um den Datensatz und damit auch die visualisierten Daten nach Bedarf zu filtern. Dies ermöglicht dem Anwender die Konzentration auf ausgewählte Aspekte des Datensatzes, ohne diese in der Menge von visualisierten Informationen aufwändig suchen zu müssen. Außerdem lassen sich damit verborgene, komplexe Eigenschaften und Zusammenhänge finden und erforschen.

## 1.2 Zielsetzung

Da die Möglichkeiten der Datenvisualisierung auf analogen, statischen Karten stark limitiert ist, müssen Ansätze entwickelt werden, um komplexe und hierarchische Daten auf digitalen, interaktiven geographischen Karten übersichtlich und ohne bedeutenden Verlust an Informationen zu visualisieren. Das Ziel dieser Arbeit ist die Entwicklung eines Ansatzes für die interaktive Visualisierung von historischen, georeferenzierten Daten. Dieser Ansatz wird im Rahmen einer prototypischen Implementierung umgesetzt und anschließend evaluiert. Eine übersichtliche und klar strukturierte Benutzeroberfläche soll entwickelt werden, um die effiziente und effektive Arbeit mit dem Ansatz zu gewährleisten. Dabei werden verschiedene, miteinander verknüpfte Ansichten benötigt, die verschiedene Aspekte des Datensatzes visuell darstellen. All diese Ansichten sollen interaktive Funktionalitäten zur Verfügung stellen, um den Datensatz gemäß der Bedürfnisse und Ziele des Benutzers zu filtern. Die Kartenansicht

## 1 Einführung

---

muss dabei alle gefilterten beziehungsweise selektierten Daten verständlich, übersichtlich, strukturiert und ohne signifikanten Verlust an Informationen darstellen. Außerdem wird eine Ansicht für die Betrachtung der Originaltexte, aus welchen die visualisierten Daten extrahiert wurden, benötigt, um den Ursprung der Informationen nachvollziehen zu können.

## 2 Grundlagen

In diesem Kapitel werden die Grundlagen für den entwickelten Ansatz beschrieben. Zunächst werden die Grundlagen der Visualisierung im Allgemeinen und der Geovisualisierung angerissen. Anschließend wird der Ansatz der Multiple Coordinated Views genauer betrachtet, da dieser in meinem Ansatz eine herausragende Rolle spielt.

### 2.1 Grundlagen der Visualisierung

Oft liegen Datenmengen vor, welche aufgrund ihrer Quantität und ihrer Komplexität nicht überschaubar und nicht einfach zu verstehen sind. Das Ziel von Visualisierung ist daher, solche Datensätze dahingehend aufzubereiten und visuell darzustellen, dass sie ohne großen Aufwand seitens des Betrachters analysiert und erfasst werden können. Der Anwender entsprechender Systeme soll dabei unterstützt werden, Daten in Informationen umzuwandeln, wobei 'Daten' Repräsentationen von Informationen sind, welche erst interpretiert werden müssen. Mittels solcher Ansätze soll der Anwender erwartete Eigenschaften der Datensätze bestätigen und unerwartete Zusammenhänge finden und erforschen können. Um diese Ziele zu erreichen, müssen große und komplexe Datenmengen geeignet zusammengefasst und aggregiert werden.

Nach Tamara Munzner [Mun14] sollte Visualisierung dann eingesetzt werden, wenn ein Problem nicht dahingehend beschrieben und definiert werden kann, so dass es von einem Rechner algorithmisch lösbar ist. Stattdessen wird durch Visualisierung versucht, die menschlichen Fähigkeiten des Anwenders zu erweitern, um die Bewältigung der Aufgaben zu ermöglichen.

Generell kann der Bereich Visualisierung in die Unterbereiche *Informationsvisualisierung*, *Wissenschaftliche Visualisierung* (*Scientific Visualization*) und *Visual Analytics* aufgeteilt werden. Es existieren viele verschiedene Definitionen für den Begriff Informationsvisualisierung, jedoch beschreiben viele davon die rechnergestützte, visuelle Repräsentation von abstrakten Daten. Die Interaktivität dieser visuellen Repräsentationen wird dabei oft betont. Nach der Definition von Friendly [FD01] beschäftigt sich die wissenschaftliche Visualisierung hauptsächlich mit dreidimensionalen, zum Beispiel architektonischen, medizinischen oder biologischen Phänomenen, wobei die räumliche Darstellung dieser Daten von großer Bedeutung ist. Unter dem Begriff *Geovisualisierung* werden Techniken und Werkzeuge verstanden, welche geographische und räumliche Daten visuell und interaktiv darstellen, um dadurch die Analyse dieser Daten zu unterstützen. Dieser Bereich vereint dabei Merkmale der beiden Bereiche Informationsvisualisierung und wissenschaftliche Visualisierung.

Für die Nutzbarkeit der Visualisierungen ist die Wahl geeigneter Repräsentationsformen essentiell. Nachfolgend wird kurz auf verschiedene Diagramme eingegangen, um die Grundlage für Diskussionen getroffener Designentscheidungen in den folgenden Kapiteln zu legen.

## 2 Grundlagen

---

Für die Visualisierung von prozentualen Anteilen eines Gesamten werden häufig Kuchendiagramme verwendet. Diese Diagramme wurden in meinem Ansatz für die Visualisierung der Informationen auf der geographischen Karte eingesetzt. Die Interpretation solcher Diagramme ist einfach und intuitiv, allerdings ist die genaue Abschätzung des durch ein Kuchenstück repräsentierten Anteils schwierig. Außerdem wird die Darstellung mit steigender Anzahl von Kuchenstücken und damit auch die Beschriftung dieser problematisch. Ich habe mich dennoch für diese Diagramme entschieden, weil die angesprochenen problematischen Aspekte im meinem Anwendungsfall nicht zum Tragen kommen.

Säulendiagramme werden verwendet, um Daten, bestehend aus Paaren von diskreten Einheiten und quantitativen Attributen, darzustellen. Diese Diagramme sind einfach zu verstehen und die repräsentierten Werte können gut unterschieden werden. Mittels gestapelter Säulen können die quantitativen Daten visuell in Segmente unterteilt werden. Diese Repräsentationsform wurde in meinem Ansatz für die Visualisierung der temporalen Daten verwendet. Problematisch hierbei ist lediglich die Beschriftung der x-Achse, falls viele Säulen dargestellt werden müssen. Da in meinem Ansatz die x-Achse den zeitlichen Verlauf repräsentiert, wobei jedem Jahr eine Säule beziehungsweise gestapelte Säulen zugewiesen werden, ist es nicht notwendig, jede einzelne dieser Säulen zu beschriften.

Die Wahl der Repräsentationsformen ist abhängig von den zu repräsentierenden Daten. S. S. Stevens [S<sup>+</sup>46] unterscheidet hier zwischen vier verschiedenen Datentypen:

- Skalen: Nominal, ordinal, Intervallskalen, Verhältnisskalen
- Komplexe Datentypen: Relational, hierarchisch, mehrdimensional, multivariat
- Zeitlich variierende Daten
- Räumliche Daten

J. Bertin definiert in [Ber83] das Konzept der *visuellen Variablen*, welche beschreiben, mittels welcher visueller Merkmale welche Eigenschaften der Daten dargestellt und verdeutlicht werden können, siehe Abbildung 2.1. Unter *visual mapping* wird hier die Abbildung von Daten auf visuelle Variablen verstanden.

Für das Grundverständnis des Visualisierungsprozesses ist die *Information Visualization Pipeline* (Abbildung 2.2) von Chi und Riedl [CR98] von großer Bedeutung. Sie beschreibt die verschiedenen Verarbeitungsschritte und die daraus resultierenden Zustände der Daten im Verlauf dieses Prozesses.

## 2.2 Grundlagen der Geovisualisierung

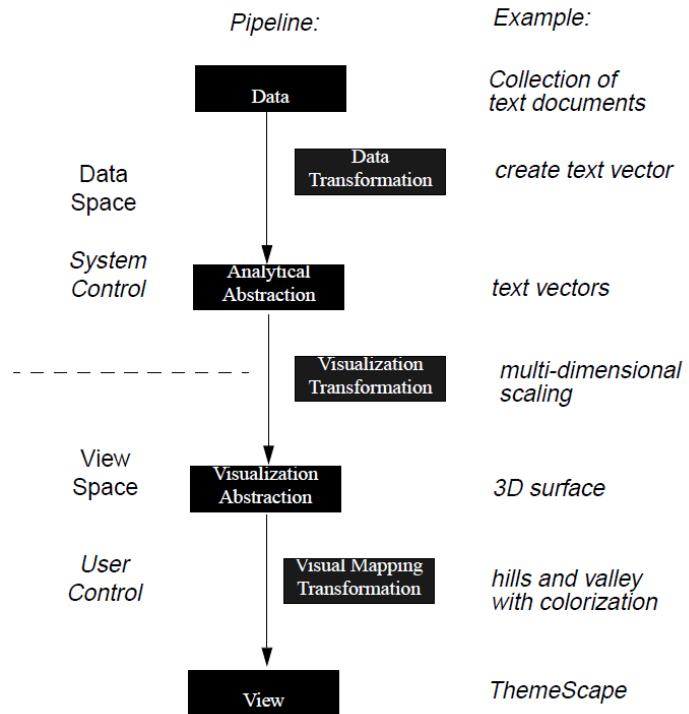
Bei der Geovisualisierung werden georeferenzierte, oft auch temporale Daten auf geographischen Karten visualisiert. Durch die Visualisierung auf geographischen Karten können Zusammenhänge und Eigenschaften der Datensätze, Ereignisse, Anhäufungen, Anomalien und Ausreißer erkannt und analysiert werden. Nach MacEachren [Mac04] werden in der Geovisualisierung Ansätze aus den Bereichen Kartographie, Geographie, wissenschaftliche und Informationsvisualisierung, explorative Datenanalyse (EDA) und Bilderkennung (Image Analysis) verwendet, um die visuelle Erforschung,

Position	
Größe	
Form	
Helligkeit (Value)	
Farbe	
Orientierung	
Textur	

Abbildung 2.1: Visuelle Variablen nach J. Bertin [Ber83]

Analyse und Darstellung georeferenzierter Daten zu ermöglichen. MacEachren beschrieb die drei Dimensionen der Geovisualisierung folgendermaßen: Geovisualisierung präsentiert das Bekannte oder hilft, das Unbekannte zu enthüllen, ist interaktiv oder statisch und wird entwickelt für die Öffentlichkeit oder für Experten.

Die Kartographie bildet eine wichtige Grundlage für die Geovisualisierung, da das Design der verwendeten Karte für die Verwendbarkeit der Anwendung von großer Bedeutung ist. Geographische Koordinaten werden verwendet, um Orte eindeutig und präzise zu spezifizieren. Diese werden anhand zweier Zahlen, für welche es wiederum einige verschiedene Notationen gibt, angegeben: *Längengrad (Longitude)* und *Breitengrad (Latitude)*. In meinem Ansatz referenziert jedes Datenobjekt in dem zu visualisierenden Datensatz ein Array, welches jeweils für diese beiden Zahlen einen Wert enthält. Auf Basis dieser Koordinaten werden die visuellen Objekte auf der Karte geographisch positioniert. Projektionen sind mathematische Formeln, welche die dreidimensionalen Daten der Erde, beschrieben als Rotationsellipsoid, zu einer planaren Oberfläche transformieren. Meistens handelt es sich dabei um eine Transformation von Längen- und Breitengrad zu kartesischen Koordinaten. Es gibt einige verschiedene Projektionen, welche wiederum bestimmte Vor- und Nachteile aufweisen, wobei die bekannteste Projektion die Mercatorprojektion ist. Diese Projektionen können anhand gewisser Merkmale, wie Winkeltreue, Flächentreue und Richtungstreue verglichen werden. Die Mercatorprojektion ist winkeltreu, nicht aber flächentreu und nicht richtungstreu. Ebenfalls können Projektionen bezüglich der verwendeten Hilfsfläche klassifiziert werden, wobei zwischen Azimutalprojektionen, Zylinderprojektionen und Kegelprojektionen unterschieden wird. Die Mercatorprojektion ist eine Variante der Zylinderprojektion.



**Abbildung 2.2:** Information Visualization Pipeline [CR98]

Wie bereits weiter oben erwähnt prägte J. Bertin [Ber83] den Begriff der *visuellen Variablen* (Abbildung 2.1) und beschäftigte sich mit der Überlegung, welche Art von Information mittels welcher Variablen intuitiv und eindeutig darstellbar ist. Dabei werden die Elemente Glyphen, Linien und Flächen in verschiedenen Gestalten bezüglich Position, Größe, Form, Helligkeit, Farbe, Orientierung und Texturierung diskutiert. Die Auswahl geeigneter visueller Variablen ist für die Verwendbarkeit der Visualisierung von entscheidender Bedeutung. Ein anderer wichtiger Aspekt ist die Wahl der Aggregationsmethoden, welche kollidierende Elemente auf der Karte zusammenfassen und somit Klarheit und Übersichtlichkeit schaffen und Interaktivität bewahren.

Grundsätzlich gibt es zwei Arten von Datenmodellen für die Entwicklung von digitalen, geographischen Daten. Raster-Datenmodelle verwenden regelmäßige Gitterzellen, was für die Visualisierung kontinuierlicher Daten vorteilhaft ist, während Vektor-Datenmodelle Punkte, Linien und Polygone nutzen, um diskrete Daten darzustellen. Sogenannte *Map-Server* rendern geographische Daten zu Kartenkacheln und bieten Dienste an, um diese abzurufen. Diese Kacheln werden hierarchisch strukturiert, so dass eine Kachel beispielsweise vier Kacheln der nächsthöheren Zoomstufe repräsentiert, wobei die Kacheln der verschiedenen Zoomstufen verschiedene Detailgrade aufweisen. Mittels entsprechender Bibliotheken werden nun die gemäß der aktuellen Zoomstufe und der Verschiebung der Karte benötigten Kartenkacheln vom Klienten angefragt und vom Server zur Verfügung gestellt. Oft erlauben solche Dienste die Transformation von geographischen Koordinaten in Pixelpositionen und

umgekehrt und die Visualisierung von Informationen auf verschiedenen Schichten der Karte. Für die Umsetzung meines Entwurfes wurde im Rahmen dieser Arbeit Leaflet.js<sup>1</sup> verwendet.

Gewöhnliche Typen von georeferenzierten Daten sind Punkt-Daten, Verläufe und kontinuierliche Werte mit diskreter oder mit kontinuierlicher Verteilung. Georeferenzierte Daten können bezüglich ihrer räumlichen Kontinuität und ihrer Glätte (Smoothness) kategorisiert werden. Räumliche Kontinuität beschreibt dabei, ob die Daten diskrete Orte oder kontinuierliche Regionen definieren, während Glätte aussagt, wie groß die Variation der Werte von Ort zu Ort ist.

Zwei der am häufigsten verwendeten Visualisierungsmethoden für georeferenzierte Daten sind *Scatterplots* und *Choropleth Maps*. Scatterplots stellen jeden einzelnen diskreten Ort als Punkt auf einer geographischen Karte dar, wodurch jeder dieser Punkte präzise selektiert werden kann. Choropleth Maps unterteilen den geographischen Raum in Regionen, wobei der Wert einer Region mittels einer Farbe aus einer Farbpalette für verschiedene Wertebereiche dargestellt wird.

## 2.3 Multiple Coordinated Views

### 2.3.1 Einführung

Falls große oder komplexe Mengen von Daten für die Erforschung visualisiert werden sollen, genügt es oft nicht, sich auf eine einzige Form der Darstellung und auf eine Ansicht zu beschränken. Besonders bei Datensätzen, die viele verschiedene Dimensionen aufweisen, ist es nur beschränkt möglich, diese in einer einzigen Repräsentationsform ohne Informationsverlust wiederzugeben. Mit der Komplexität der Daten wächst ebenfalls die Komplexität der visuellen Darstellung. Dies kann zu zweierlei Problemen führen. Zum einen besteht die Gefahr, dass es dem Anwender schwer fällt, die Komplexität zu durchdringen, den Überblick zu behalten und ein gutes Verständnis für den Datensatz zu erlangen. Zum anderen ist die Entwicklung und letztendlich auch die Bedienung von Interaktionsmöglichkeiten mit den visualisierten Objekten deutlich erschwert, wenn viele Aspekte des Datensatzes in einer Ansicht repräsentiert werden.

Diese Probleme können durch das Konzept der sogenannten *Multiple Coordinated Views* (MCV, CMV) gelöst werden. Dabei werden die Daten gemäß verschiedener Gesichtspunkte in verschiedenen, miteinander geeignet verknüpften Ansichten visualisiert. Durch diese Aufteilung in mehrere Ansichten können Darstellungen der Daten erstellt werden, die eine wesentlich geringere Komplexität aufweisen, als dies bei einer einzigen Ansicht, die die Gesamtheit der Daten umfasst, der Fall wäre. Gut umgesetzt kann dieser Ansatz dabei helfen, ein wesentlich besseres Verständnis für den zugrundeliegenden Datensatz zu erarbeiten.

<sup>1</sup><http://leafletjs.com/>

### 2.3.2 Definitionen und Begriffe

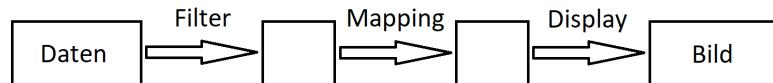
In wissenschaftlichen Arbeiten zu MCV wird eine Vielzahl von Begriffen verwendet, die auf den ersten Blick das gleiche bedeuten. Um Missverständnisse zu vermeiden, werden hier kurz grundlegende Begriffe erklärt [Sch08] [Rob07].

- Single View: Dabei handelt es sich um eine einzige Ansicht, die bestimmte Daten repräsentiert.
- Form: Dieser Begriff beschreibt die tatsächliche, gewählte Form der Darstellung beziehungsweise Repräsentation der Daten.
- Multiform: Gemeint ist ein System, welches aus mehreren Forms zusammengesetzt ist, die allesamt die selben Daten visualisieren.
- Distinct Views: Mehrere Ansichten repräsentieren verschiedene Aspekte des selben Datensatzes.
- Multiple (coordinated) Views: Hierbei werden die entsprechenden Daten in mehreren Ansichten dargestellt. Die verschiedenen Ansichten müssen aber nicht zwangsläufig verschiedene Aspekte darstellen. Es besteht eine Art der Koordination, also eine logische Verknüpfung, zwischen den Ansichten [Rob07].
- Multiple View System: Ein System, welches aus mehreren Distinct Views zusammengesetzt ist [WBWK00].

Zusätzlich zu den oben beschriebenen Begriffen beziehungsweise Konzepten existieren noch die sogenannten *Dual View Systems*. Diese bestehen im Normalfall aus genau zwei miteinander verbundenen Ansichten. Es wird zwischen folgenden Unterklassen unterschieden [Sch08] [Rob07]: Overview- und Detail-Views, Focus- und Context-Views, Difference-Views, Master-Slave-Views und World in Miniature.

### 2.3.3 Entwicklung

Bei der Entwicklung solcher MCV-Systeme müssen einige bedeutende Designentscheidungen getroffen und viele Faktoren berücksichtigt werden. Zunächst sollte geklärt werden, welche Klasse von MCV-Systemen für den zugrundeliegenden Datensatz am besten geeignet ist.



**Abbildung 2.3:** Dataflow Model [UFK<sup>+</sup>89] [HM90]

Eine wesentliche Entscheidung besteht in der Auswahl der Informationen und Aspekte, welche durch die verschiedenen Ansichten repräsentiert werden sollen. Das traditionelle *Dataflow-Modell* (Abbildung 2.3) [UFK<sup>+</sup>89] [HM90] beschreibt eine Pipeline, bestehend aus drei Phasen: Filterung der Daten, Mapping der Daten auf Repräsentation und Anzeigen der Repräsentation. Verschiedene Ansichten können nun durch Änderungen von Parametern oder Verfahren bezüglich der verschiedenen

Pipelinstufen generiert beziehungsweise entwickelt werden [Rob00]. Durch Parameteränderungen in der Filter-Stufe entstehen Repräsentationen verschiedener Teilmengen des zugrundeliegenden Datensatzes. In bestimmten Szenarien kann es zum Beispiel nützlich sein, durch entsprechende Filterkonfigurationen mehrere Ansichten zu generieren, welche hierarchisch verschiedene Filterresultate darstellen. In der Mapping-Stufe können verschiedene Ansichten mit verschiedenen Repräsentationen der selben Datenmenge generiert werden, indem die Visualisierungsverfahren verändert oder ausgetauscht werden. In der Anzeige-Stufe der Pipeline steht lediglich die Manipulation der Ansicht oder der Projektion der visualisierten Objekte für die Generierung verschiedener Ansichten zur Verfügung. Bezuglich einer 3D-Repräsentation kann beispielsweise zwischen paralleler und perspektivischer Ansicht oder verschiedenen Blickwinkeln auf das Objekt unterschieden werden.

Sicher ist die Wahl geeigneter Formen der Visualisierung und passender Mapping-Verfahren von entscheidender Bedeutung. Diese Formen können in die folgenden Klassen unterteilt werden [LBWR94]: Karten, Netzwerke, Diagramme & Graphen, Tabellen & Matrixlayouts und Symbole & Glyphen.

Bevor jedoch diese Entscheidungen getroffen werden, ist es sinnvoll und notwendig, sich Gedanken über eine etwaige Vorverarbeitung der Daten zu machen. Oft sind die zu visualisierenden Datensätze zu groß oder weisen eine zu hohe Komplexität auf, weshalb geeignete Aggregations-und/oder Abstraktionsverfahren gewählt werden müssen. Bevor die gewählten Visualisierungsformen umgesetzt und implementiert werden, sollten die Interaktionsmöglichkeiten, die dem Anwender zur Verfügung stehen sollen, spezifiziert und in das Konzept der Visualisierung integriert werden.

In MCV-Systemen kommen generell zwei Arten von Interaktionsmöglichkeiten zum Einsatz: *Linking* und *Navigation*. Dieser Sachverhalt wird auch *Koordination* genannt. Der Begriff *Linking* [BMMS91] beschreibt die logische Verknüpfung der visuellen Elemente verschiedener Ansichten. Das Grundkonzept besteht darin, dass die Manipulation von visuellen Repräsentationen in einer Ansicht entsprechende Änderungen in den anderen, verknüpften Ansichten zur Folge hat. Eine spezifische Variante, welche in vielen MCV-Systemen angewandt wird, ist das sogenannte *Brushing* [BC87]. Dabei selektiert der Anwender bestimmte, visuell repräsentierte Datenobjekte oder grafische Elemente, wodurch in den anderen verknüpften Ansichten entsprechende Objekte und Elemente hervorgehoben werden. Der Begriff *Dynamic querying* beschreibt die Möglichkeit der Filterung der repräsentierten Daten durch UI-Elemente wie Slider, Eingabefelder, Buttons oder Menüs. Manipulationsmöglichkeiten, welche durch Interaktion mit den visuellen Objekten und Elementen selber ausgeführt werden können, werden dementsprechend *direkte Manipulation* genannt. Werden dagegen Inhalte von Ansichten durch separate Interaktionselemente manipuliert, so spricht man von *indirekter Manipulation*. Somit handelt es sich beim Brushing um direkte und beim Dynamic Querying um indirekte Manipulation. Neben der Manipulation und Filterung der Daten beziehungsweise deren Repräsentation beim Linking, werden häufig Interaktionsmöglichkeiten für die koordinierte Navigation innerhalb der verschiedenen Ansichten umgesetzt. Die Idee dabei ist die selbe wie beim Linking: Navigationsoperationen in einer Ansicht werden in allen anderen, verknüpften Ansichten entsprechend der verschiedenen Visualisierungsformen ausgeführt. Der gängige Begriff hierfür lautet *Navigational slaving*. Weitere Interaktionsmöglichkeiten wären beispielsweise die Änderung der Datenverarbeitung, wie zum Beispiel die Aggregation von Datenobjekten, die Manipulation oder das Austauschen von Mappingverfahren für das Mapping von Daten auf Repräsentation, oder die manuelle Anpassung der Benutzeroberfläche, wie zum Beispiel das Verstecken bestimmter Ansichten oder die Anordnung dieser auf dem Bildschirm [Sch08] [Rob07].

## 2 Grundlagen

---

Ein wichtiger Aspekt bei der Entwicklung geeigneter Koordinationsoperationen ist die Frage, auf welche Art und Weise sich visuelle Repräsentationen von Daten in verknüpften Ansichten ändern, sobald Parameter durch die Interaktion mit den visuellen Objekten in einer Ansicht manipuliert werden. Roberts [Rob00] definierte hierfür drei verschiedene Vorgehensweisen: *Replace*, *Replicate* und *Overlay*.

- Replacement: Die Repräsentation der veralteten Parameter wird durch die neue, aktuelle Repräsentation ersetzt. Das hat zum Vorteil, dass der Anwender genau weiß, wo entsprechende Informationen aktualisiert werden. Allerdings gehen Informationen über den früheren Zustand verloren, weshalb ein Vergleich beider Zustände nur kognitiv möglich ist. Durch die Speicherung früherer Zustände und die Möglichkeit, jederzeit zu diesen zurück zu springen, kann der Vergleich vereinfacht werden. Außerdem kann der Vergleich von veralteter und aktueller Repräsentation durch animierte Übergänge erleichtert werden.
- Replication: Hierbei wird jedes Mal, wenn sich für eine Ansicht relevante Parameter ändern, eine neue Ansicht generiert, die den neuen Zustand repräsentiert. Die veraltete Ansicht bleibt dabei erhalten. Dadurch können die verschiedenen Zustände angenehm miteinander verglichen werden, ohne sich frühere Zustände im Gedächtnis behalten zu müssen. Problematisch ist hierbei die Anordnung und der Platzbedarf der verschiedenen, jeweils neu generierten Ansichten.
- Overlay: Verschiedene Zustände oder Repräsentationen werden in einer einzigen Ansicht repräsentiert. In bestimmten Szenarien kann dieses Vorgehen geeignet sein, da der Vergleich der visuellen Repräsentationen hierdurch unterstützt werden kann. Allerdings besteht dabei die Gefahr, dass zu viele verschiedene Repräsentationen der Daten in einer einzigen Ansicht ein zu komplexes Bild ergeben. Somit können der Vergleich, die Interaktion und Navigation und auch das Verständnis und die Interpretation der Repräsentationen durch diese Vorgehensweise erschwert werden.

Es ist ersichtlich, dass jede Vorgehensweise ihre Vor-und Nachteile mit sich bringt. Die Auswahl der Vorgehensweise sollte daher unbedingt vom Kontext, dem Szenario und den zugrundeliegenden Daten abhängig gemacht werden [Rob04] [Rob07].

### 2.3.4 Richtlinien/Regeln

Es ist ersichtlich, dass MCV-Systeme einige signifikante Vorteile gegenüber Single-View-Systemen bieten. Allerdings müssen dafür gewisse Trade-Offs und Nachteile in Kauf genommen werden. So sind beispielsweise die erforderliche Rechenleistung und der Platzbedarf auf dem Bildschirm erheblich höher als bei alternativen Systemen mit nur einer Ansicht. Auch sollte in Betracht gezogen werden, dass Entwicklung und Wartung eines solchen Systems mit höheren Kosten und einem höheren Ressourcenbedarf einhergehen. Mit der Komplexität des Systems steigt auch die erforderliche Zeit und der Aufwand für die Einarbeitung in das System. Somit entstehen auch für den Endanwender gewisse Herausforderungen. Daher ist eine der ersten, wichtigsten Entscheidungen bezüglich der Entwicklung eines Systems zur interaktiven Datenvisualisierung, ob überhaupt ein MCV-Ansatz umgesetzt werden soll. Dabei sind die Vorteile gegen die Nachteile beziehungsweise die entsprechenden Trade-Offs abzuwiegeln und es muss kritisch hinterfragt werden, welcher Ansatz für die Problemlösung geeigneter

ist. Bezuglich dieser Entscheidung und bezüglich der Fragestellung, wie ein MCV-System optimal umgesetzt werden kann, sobald man sich für diesen Ansatz entschieden hat, haben Baldonado et al. [WBWK00] jeweils vier Regeln formuliert, die bei der Entscheidungsfindung und der Entwicklung helfen sollen. Folgende Regeln sind anzuwenden, wenn es um die Entscheidung für oder gegen einen MCV-Ansatz geht:

- Regel der Vielfalt: Ein MCV-Ansatz sollte angewandt werden, wenn der Datensatz gewisse Typen der Vielfalt aufweist. Die konkreten Typen sind: Attribute, Modelle, Userprofile, Abstraktionsstufen und/oder Genres.
- Regel der Komplementarität: Wenn verschiedene Ansichten die Möglichkeit oder Wahrscheinlichkeit bieten, um Korrelationen und Ungleichheiten innerhalb der Daten aufzuzeigen, rechtfertigt dies die Anwendung eines MCV-Ansatzes.
- Regel der Zerlegung: Ein MCV-Ansatz ist nützlich, wenn der Datensatz komplex ist und einige Dimensionen aufweist, welche separat in verschiedenen Ansichten repräsentiert werden können.
- Regel der Sparsamkeit: Verschiedene Ansichten sind sparsam einzusetzen, denn jede weitere Ansicht erhöht den Bedarf an kognitiven und technischen Ressourcen.

Bezuglich des Designs und der Entwicklung von MCV-Systemen wurden folgende vier Regeln formuliert:

- Regel der Platz/Zeit-Ressourcenoptimierung: Werden verschiedene Ansichten nebeneinander oder sequentiell, also nacheinander, angezeigt? Diesbezüglich sollte geklärt werden, wie viel Bildschirmplatz und Rechenleistung beziehungsweise Rechenzeit zur Verfügung stehen. Falls der Anwender verschiedene Ansichten vergleichen muss, um sein Problem zu lösen, spricht dies beispielsweise für eine parallele Anzeige der Ansichten.
- Regel der Selbstverständlichkeit: Beziehungen zwischen verschiedenen Ansichten sollten mittels wahrnehmbarer, visueller oder auditiver Hinweise verdeutlicht werden. Beispiele für solche Hinweise sind Hervorhebung, Positionierung der Ansichten, visualisierte Beziehungen durch Pfeile und koordinierte Interaktion zwischen Ansichten.
- Regel der Konsistenz: Der Systemzustand, also der Datensatz und die Visualisierungen, sollten immer konsistent sein. Ebenfalls ist die Konsistenz der Benutzeroberfläche zu jedem Zeitpunkt relevant. Für letzteres ist beispielsweise wichtig, dass ähnliche Repräsentationen auch über ähnliche oder gleiche Interaktionsmöglichkeiten verfügen.
- Regel des Aufmerksamkeitsmanagements: Das System sollte die Aufmerksamkeit des Benutzers leiten und Ablenkungen vermeiden. Dabei können Animationen, Geräusche, visuelles Hervorheben und Bewegung von Elementen helfen.



# 3 Verwandte Arbeiten

## 3.1 Multiple Coordinated Views und Geovisualisierung

In diesem Kapitel werden einige verwandte Arbeiten kurz beschrieben und dabei mit meinem Ansatz verglichen und von meinem Ansatz abgegrenzt. Betrachtet werden hier hauptsächlich Arbeiten aus dem Bereich der Geovisualisierung, welche einen MCV-Ansatz verwenden.

### 3.1.1 City'OScope

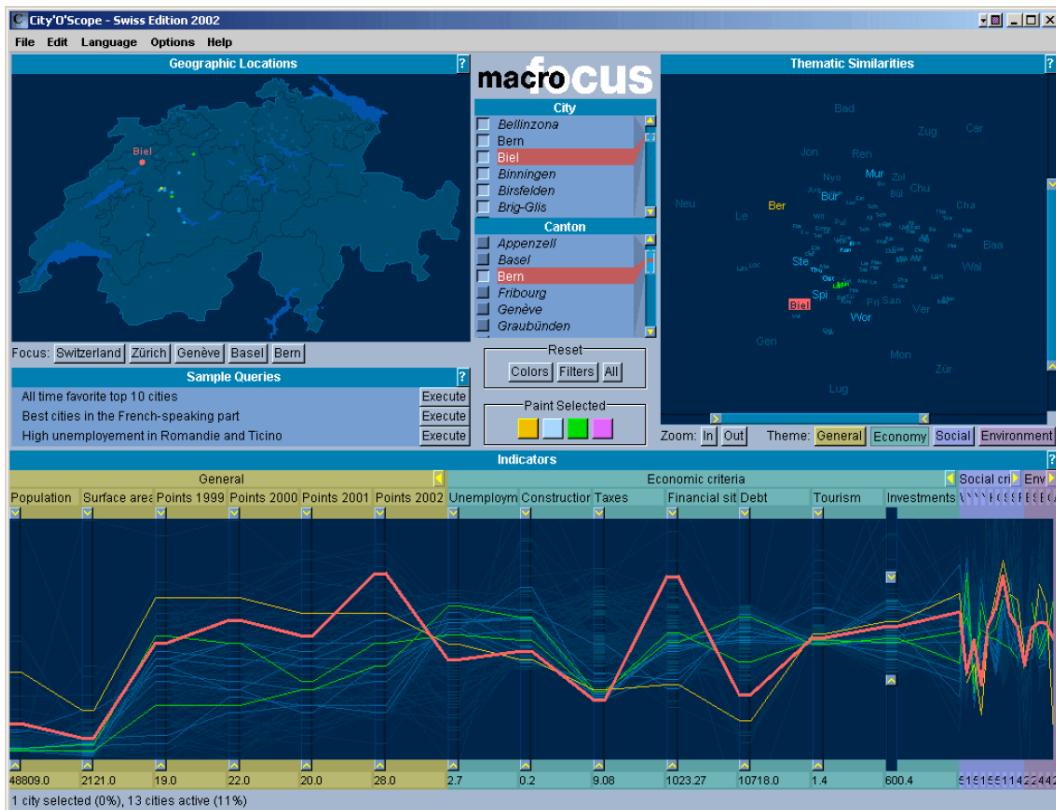


Abbildung 3.1: CityOScope [BG03]

Brodbeck und Girardin stellen in [BG03] einen Ansatz für die Analyse georeferenzierter Daten, welche eine hohe Anzahl an Dimensionen aufweisen, vor. Die Benutzeroberfläche ist in Abbildung 3.1 darge-

### 3 Verwandte Arbeiten

---

stellt. Bei den Daten und den verschiedenen Dimensionen handelt es sich um Städte in der Schweiz und um relevante Kennzahlen und Indikatoren für die Bewertung der Attraktivität dieser Städte. Dieser Ansatz soll den Anwender durch geeignete Interaktionsmöglichkeiten dabei unterstützen, ein besseres Verständnis für den Datensatz zu erlangen, als dies durch eine statische Repräsentation dieser großen Datenmengen möglich wäre.

Das grundlegende Konzept für den gewählten Ansatz ist die Verwendung von koordinierten Ansichten. Für verschiedene Informationen oder verschiedene Aspekte derselben Informationen wurden verschiedene Ansichten entwickelt.

Die interaktive, geographische Karte stellt die selektierten Städte mittels farbiger Punkte und entsprechender Labels dar. Dabei dient die Karte hier lediglich der Veranschaulichung der geographischen Lagen der Städte, ohne weitere Informationen zur Verfügung zu stellen. Über die selektierten Städte können weitere Informationen bezogen werden durch die Betrachtung und Analyse der Zustände der anderen verbundenen Ansichten. In meinem Ansatz ist die geographische Karte die bedeutendste Komponente. Hier werden nicht nur die geographischen Lagen der selektierten Städte und Orte visualisiert, sondern auch komplexe, hierarchische Informationen inklusive selektierter und nicht selektierter Anteile, welche diese Städte referenzieren. Dabei werden die Städte und Orte und die dazugehörigen Daten bei Bedarf automatisch aggregiert. In diesem Ansatz kann die Karte nach Bedarf verschoben und gezoomt werden, während in City’O’Scope vordefinierte Optionen zur Verfügung stehen, um bestimmte geographische Regionen durch geeignete Zoom-Parameter und Fisheye-Methoden [SB94] zu fokussieren.

Innerhalb einer Similarity Map werden die (abgekürzten) Namen der Städte so angeordnet, dass Städte, welche ähnliche Werte in bestimmten Dimensionen aufweisen, nahe beieinander dargestellt werden. In meinem Ansatz ist eine solche Komponente nicht enthalten, es wäre aber denkbar, diese Methode auf den von mir verwendeten Datensatz anzuwenden. So könnten Städte näher beieinander dargestellt werden, innerhalb welcher die selben religiösen Gruppen angesiedelt oder aktiv waren. Dadurch wäre es beispielsweise möglich, das gemeinsame Auftreten verschiedener Gruppen innerhalb von Städten zu erforschen.

Für die Darstellung der verschiedenen Kennzahlen oder Indikatoren für jede Stadt wurde eine Parallel Koordinaten-Ansicht entwickelt, welche für jeden Indikator eine vertikale Achse enthält. Innerhalb jeder Achse sind die Werte der verschiedenen Städte bezüglich dieser Dimension markiert und achsenübergreifend sind alle Werte derselben Stadt durch Linien miteinander verbunden. Jede dieser Achsen verfügt über zwei vertikale Slider, über welche ein Wertebereich innerhalb dieser Dimension selektiert werden kann. Dabei werden automatisch alle Städte deseletktiert, welche in entsprechenden Dimensionen Werte aufweisen, die außerhalb der selektierten Bereiche liegen.

Dieser Ansatz enthält keine Ansicht, über welche eine Zeitspanne oder verschiedene Zeitpunkte ausgewählt werden können, was allerdings bezüglich der Entwicklung der einzelnen Städte und der dazugehörigen Indikatoren von Interesse wäre. Mein Ansatz stellt eine graphische Zeitleiste zur Verfügung, welche über den zeitlichen Verlauf hinweg die Entwicklung bestimmter Daten visualisiert und mittels welcher der Anwender eine Zeitspanne selektieren kann.

In City’O’Scope wurden drei verschiedene Modi der Methode *Brushing and Linking* umgesetzt. Mittels dieser drei Modi können Elemente ausgewählt und entsprechende Detailinformationen angezeigt oder in anderen Ansichten hervorgehoben, ausgewählte Elemente selektiert und letztendlich selektierte

### 3.1 Multiple Coordinated Views und Geovisualisierung

Elemente permanent farblich markiert werden. Mein Ansatz enthält sowohl den ersten Modus als auch einen Modus, welcher als Kombination der beiden letzten Modi interpretiert werden kann.

#### 3.1.2 VAIroma

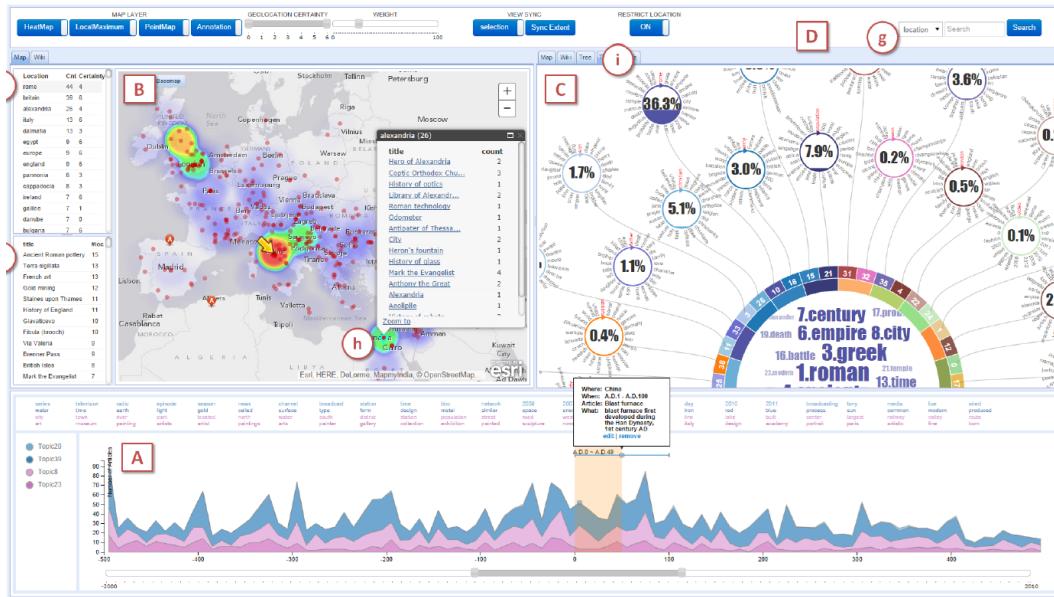


Abbildung 3.2: VAIroma [CDW<sup>+16</sup>]

Der von Cho et al. entwickelte *Visual Analytics*-Ansatz [CDW<sup>+16</sup>] soll den Anwender dabei unterstützen, neues Wissen über die römische Geschichte zu erwerben und zu entwickeln. Dafür werden Informationen aus einer Sammlung von Wikipedia-Artikeln automatisch extrahiert und über eine interaktive Benutzeroberfläche graphisch dargestellt. Die Informationen beschreiben hauptsächlich Ereignisse, Orte und Zeit. Somit besteht bezüglich der Daten eine Ähnlichkeit zu meinem Ansatz, in dem visualisierte Datenobjekte jeweils eine Zeitspanne, eine Stadt inklusive geografischer Lage und Informationen über die Ansiedlung oder Aktivitäten religiöser Gruppen beschreiben. Letztere Information kann mit den Ereignissen in VAIroma verglichen werden.

Für die Umsetzung der Benutzeroberfläche, welche in Abbildung 3.2 dargestellt wird, wurde ebenfalls ein MCV-Ansatz gewählt. Die enthaltenen Ansichten sind dabei vergleichbar mit den Ansichten meines Ansatzes.

Die geographische Karte verfügt über drei Schichten, welche verschiedene Informationen visualisieren. Die erste Schicht besteht aus einer Heatmap, welche farblich kodiert die Häufigkeit der Erwähnungen der Orte und Regionen in den Wikipedia-Artikeln wiedergibt. Im meiner Arbeit wurde eine andere Art der Aggregation und eine andere Form der Repräsentation von häufig vorkommenden Orten gewählt: Datenobjekte, welche die selbe Stadt oder entsprechend der Zoomstufe nahe beieinander liegende Städte beschreiben, werden aggregiert und das Resultat wird als Baumstruktur dargestellt, wobei die Größe der Knoten visualisiert, wie viele Objekte hierbei zusammengefasst wurden. Auf

### 3 Verwandte Arbeiten

---

Anfrage können die exakten geographischen Lagen der aggregierten Städte durch kleine, rote Punkte visualisiert werden. Durch die Kombination dieser beiden Ansätze wird weniger Fläche der zugrundeliegenden Karte verdeckt, wodurch wichtige Informationen trotz der Visualisierung auf der Karte weiterhin zur Verfügung stehen. Die zweite Schicht visualisiert die exakten Lagen der einzelnen Städte, dargestellt durch rote Kreise. Diese sind transparent gestaltet, so dass Überschneidungen gut sichtbar sind. Überschneidungen wurden in meinem Ansatz komplett eliminiert durch eine geeignete Aggregation der visualisierten, georeferenzierten Daten. Durch Anklicken einer dieser Kreise wird ein Fenster angezeigt, welches eine Liste relevanter, entsprechender Wikipedia-Artikel darstellt. Die dritte Schicht markiert selektierte oder aus einer Suchanfrage resultierende Orte, also rote Kreise, mit Nadeln, welche ebenfalls angeklickt werden können, um eine Liste anzeigen zu lassen, die die Namen von Artikeln enthält, die den entsprechenden Ort referenzieren. Diese Interaktionsmöglichkeiten in den letzten beiden Schichten ermöglichen dem Anwender, abzufragen, wie die visuellen Elemente zu Stande kommen und welche Informationen diesen zugrunde liegen. Eine ähnliche Funktionalität ist in meinem Ansatz enthalten: Wählt der Anwender ein visuelles Objekt in der Kartenansicht aus, so erscheint ein Fenster, welches darstellt, welche religiösen Gruppen und wie viele Datenobjekte, welche diese referenzieren, hierbei aggregiert wurden.

Ebenfalls verfügt die Benutzeroberfläche über eine Ansicht, welche eine Zeitleiste darstellt. Die x-Achse stellt den zeitlichen Verlauf, die y-Achse die Anzahl der entsprechenden Wikipedia-Artikel dar. Der Verlauf ist durch einen Graphen repräsentiert. Der Anwender kann dabei eines oder mehrere Themen auswählen, welche dann in der Zeitleiste visualisiert werden. Werden mehrere Themen selektiert, so werden diese mittels gestapelter Graphen dargestellt. Die y-Achse beschreibt dabei, wie viele Wikipedia-Artikel in der Sammlung existieren, welche den entsprechenden Zeitpunkt (x-Achse) und das entsprechende Thema enthalten. Die Zeitleiste in meinem Ansatz beschreibt ebenfalls anhand der x-Achse den zeitlichen Verlauf, angegeben in Jahren, und anhand der y-Achse die Anzahl der Datenobjekte, welche eine Zeitspanne referenzieren, in denen das entsprechende Jahr enthalten ist. Dabei werden jederzeit mehrere Religionen dargestellt, welche mit den oben erwähnten Themen verglichen werden können. Für die Unterscheidung der verschiedenen Religionen ist die Zeitleiste in meinem Ansatz als gestapeltes Säulendiagramm umgesetzt. Wird in VAiRoma innerhalb der Zeitleiste eine Zeitspanne selektiert, so werden nur entsprechende Orte in der geographischen Karte visualisiert. Da die Visualisierung in der Kartenansicht in meinem Ansatz aggregierte Datenobjekte darstellt, verschwinden keine visuellen Elemente durch die Selektion einer Zeitspanne durch die Zeitleiste, stattdessen visualisieren die Elemente den aktuell selektierten und den nicht selektierten Anteil.

VAiRoma enthält außerdem Ansichten für die Visualisierung der hierarchischen Themen-Struktur, mittels welcher Themen selektiert werden können. Eine der beiden alternativen Darstellungen ist eine Darstellung als visuelle Baumstruktur. Diese kann mit der Visualisierung innerhalb der Gruppenansicht innerhalb meines Ansatzes verglichen werden. Ein wichtiger Unterschied ist, dass die selektierten Gruppen in meinem Ansatz direkt als solche innerhalb der geographischen Karte dargestellt werden, während in VAiRoma lediglich rote Kreise für die Darstellung der Orte verwendet werden.

Außerdem verfügt VAiRoma über eine Ansicht, mittels welcher der Anwender ausgewählte Wikipedia-Artikel anzeigen lassen und betrachten kann.

Zwischen den verschiedenen Ansichten wurden verschiedene Verknüpfungen umgesetzt. Durch die Selektion einer Zeitspanne werden entsprechende Orte in der geographischen Karte und in

### 3.1 Multiple Coordinated Views und Geovisualisierung

einer Liste, welche relevante Orte anzeigt, und entsprechende Artikel in einer Liste von relevanten Wikipedia-Artikeln, angezeigt. Ebenfalls können zwei Zeitspannen gleichzeitig selektiert werden, wodurch die Karte dupliziert wird und die beiden Visualisierungen verglichen werden können. Durch entsprechende Interaktionsmöglichkeiten kann der Anwender nach Themen, Zeit und geographischen Regionen filtern.

#### 3.1.3 VisGets

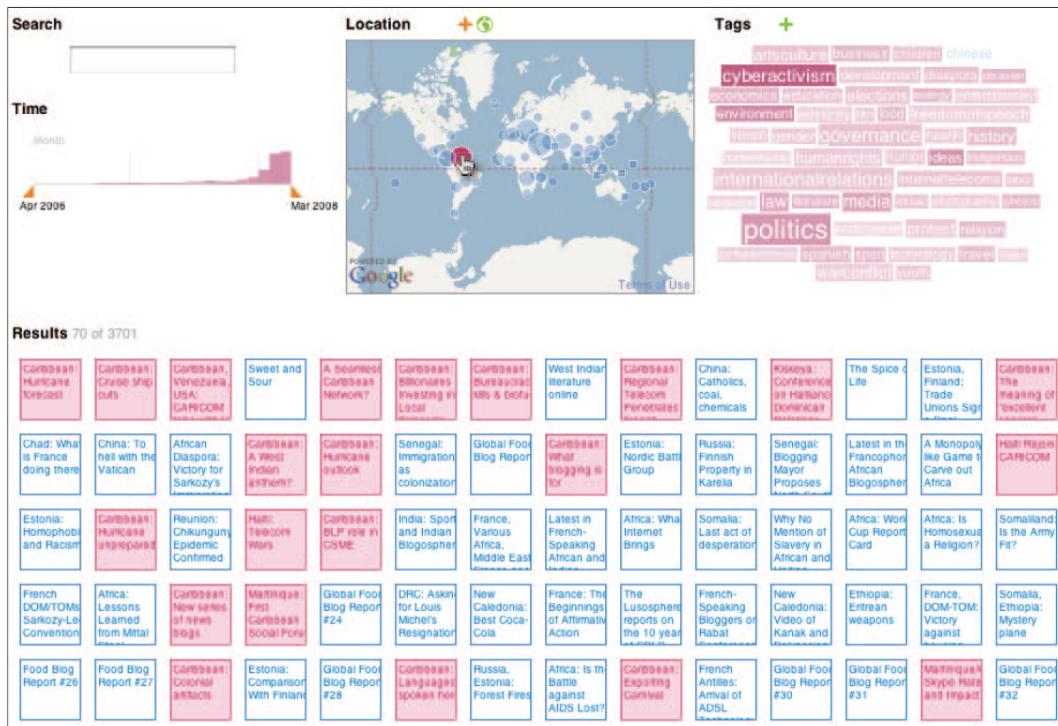


Abbildung 3.3: VisGets [DCCW08]

Die Autoren Dörk et al. präsentieren in [DCCW08] VisGets, einen Ansatz für die Verwendung von koordinierten Ansichten für die Erforschung und Entdeckung web-basierter Inhalte. Die Benutzeroberfläche ist in Abbildung 3.3 dargestellt. Dieser Ansatz zielt darauf ab, eine effizientere Möglichkeit der Informationssuche im Internet anzubieten, als die gewöhnliche Textsuche, wie sie heute noch standardmäßig verwendet wird. Der Begriff VisGets beschreibt hierbei Widgets, nachfolgend einfach 'Ansichten' genannt, welche interaktive Informationsvisualisierungen enthalten. Durch die Interaktion des Anwenders mit diesen Ansichten wird die Selektion der Informationen und damit die Visualisierung dieser verändert. Diese Interaktion aktualisiert nicht nur die Visualisierungen, sondern liefert als Resultat auch eine Liste von gefundenen Informationsobjekten, welche den Selektionsparametern entsprechen.

Bezüglich der visualisierten Dimensionen der verschiedenen Ansichten beschränken sich die Autoren zunächst auf die Aspekte Zeit, Ort und Thema, da diese Informationen zunehmend für Internetinhalte

### 3 Verwandte Arbeiten

---

zur Verfügung stehen. Diese drei Aspekte bilden ebenfalls in meiner Arbeit den Kern des präsentierten Ansatzes. Der dritte Aspekt, also die Themen, kann dabei mit dem Aspekt der religiösen Gruppen in meiner Arbeit verglichen werden.

Die Zeit-Ansicht stellt eine interaktive, graphische Zeitleiste zur Verfügung. Die x-Achse repräsentiert den zeitlichen Verlauf in Monaten, während die y-Achse darstellt, wie viele Informationsobjekte in dem entsprechenden Monat veröffentlicht wurden. Eine Zeitspanne kann durch die Verwendung von Slidern selektiert werden. Die Umsetzung der Zeitleiste in meinem Ansatz basiert ebenfalls auf diesen grundlegenden Merkmalen. In VisGets kann ein einzelner Monat durch die Auswahl des entsprechenden Balkens selektiert werden. Daraufhin erscheint unter der Zeitleiste eine zweite Zeitleiste, welche die einzelnen Tage des selektierten Monats für die Skalierung der x-Achse verwendet. Eine solche detailliertere Darstellung einer bestimmten Zeiteinheit ergibt im Kontext von VisGets Sinn. Da meine Arbeit einen zeitlichen Bereich von mehreren Jahrhunderten abdeckt und die Dichte der Informationen relativ gering ist, ist in diesem Kontext eine solche Detailansicht nicht sinnvoll. Die Balkendiagramme unterscheiden sich dahingehend deutlich, dass die Balken in meiner Arbeit segmentiert beziehungsweise gestapelt sind und damit detaillierter visualisieren, wie die Informationen bezüglich einer Zeiteinheit zusammengesetzt sind.

Die Ort-Ansicht stellt eine interaktive, geographische Karte zur Verfügung. Hier werden die in den Informationsobjekten referenzierten Orte dargestellt. Individuelle Objekte werden durch Rechtecke, aggregierte Objekte durch Kreise repräsentiert. Die Größe der Kreise entspricht dabei der Anzahl der aggregierten Objekte. Auch in meinem Ansatz wird diese Anzahl durch die Größe der Kreise beziehungsweise der Baumknoten visualisiert. Knoten, welche mehrere, also aggregierte, Objekte repräsentieren, sind dabei jeweils durch ein Kuchen- und ein Donutdiagramm zu erkennen, welche entsprechend den aktuell selektierten und nicht selektierten Anteil darstellen. Ein weiterer, bedeutender Unterschied zwischen beiden Ansätzen ist, dass in meinem Ansatz die Kreise wichtige Informationen über die hier visualisierten Datenobjekte enthalten, nämlich die durch diese referenzierten religiösen Gruppen, während die Elemente in dem andern Ansatz lediglich geographische Koordinaten und die Anzahl der repräsentierten Datenobjekte visualisieren. In der Kartenansicht in VisGets werden alle Objekte selektiert, welche durch den aktuellen Zustand der Karte sichtbar sind. Wird das Zoom-Level oder die Verschiebung der Karte verändert, so ändert sich automatisch die Selektion. Durch das Zeichnen eines Polygons innerhalb der Karte kann in meinem Ansatz eine Region ausgewählt werden, wodurch alle Objekte, welche Koordinaten enthalten, die innerhalb dieses Polygons liegen, selektiert werden. Alle anderen Objekte werden automatisch deseletiert. Somit kann eine wesentlich feinere Selektion getätigter werden. Außerdem bietet mein Ansatz den Vorteil, dass neben den selektierten auch deseletierten Elementen weiterhin sichtbar sind. Wird in VisGets ein visuelles Objekt angeklickt, welches mehrere Datenobjekte repräsentiert, so wird automatisch das Zoom-Level erhöht und die aggregierten Orte werden individuell dargestellt. In meinem Ansatz wird dies durch manuelles Zoomen und Verschieben der Karte bewerkstelligt.

Außerdem wurde eine Tag-Ansicht entwickelt, welche eine Tag Cloud darstellt, die alphabetisch sortiert die in den Informationsobjekten enthaltenen Tags darstellt, wobei die Schriftgröße die Häufigkeit des Vorkommens eines Tags visualisiert. Durch Anklicken eines Tags wird dieser selektiert und die Ergebnisliste wird dahingehend aktualisiert, so dass diese nur Objekte auflistet, welche diesen selektierten Tag enthalten. Die Ergebnisliste stellt alle Informationsobjekte dar, welche den aktuellen Selektionsparametern der anderen Ansichten entsprechen. Die dargestellten Informationsobjekte

### 3.1 Multiple Coordinated Views und Geovisualisierung

enthalten jeweils einen Titel, einen Hyperlink zur tatsächlichen Informationsquelle im Internet und gegebenenfalls eine kurze Beschreibung mit Bild.

Die verschiedenen Ansichten in VisGets sind mittels zweier Arten von Koordinierung verbunden. Fährt der Anwender mit dem Mauszeiger über ein visuelles Element in einer Ansicht, so werden entsprechende Elemente in allen anderen Ansichten hervorgehoben. Die Verbindungen zwischen dem ausgewählten Elementen und den hervorgehobenen Elementen werden durch den Einsatz von Farbe und Opazität gewichtet dargestellt. In meinem Ansatz können Knoten in der Kartenansicht ausgewählt werden, wodurch entsprechende visuelle Elemente oder Informationen in den anderen Ansichten farblich hervorgehoben werden. Eine Gewichtung ergibt dabei im Kontext dieser Arbeit keinen Sinn. Werden die Selektionsparameter in einer Ansicht verändert, so werden die Visualisierungen in den anderen Ansichten und in der Ergebnisliste aktualisiert. Dies entspricht der koordinierten Selektion in meinem Ansatz.

#### 3.1.4 DC Crime Visualization und ähnliche Ansätze

Die Autoren Roth und Ross präsentieren in [RR09] vier verschiedene Ansätze, welche geographische und temporale Daten auf geographischen Karten visualisieren und animieren. Der vierte Ansatz wurde mittels einer von den Autoren selber entwickelten Erweiterung der Google Maps API, welche das Kernthema deren Arbeit ist, erstellt. Nachfolgend werden diese vier Ansätze kurz erklärt und mit meinem Ansatz verglichen.

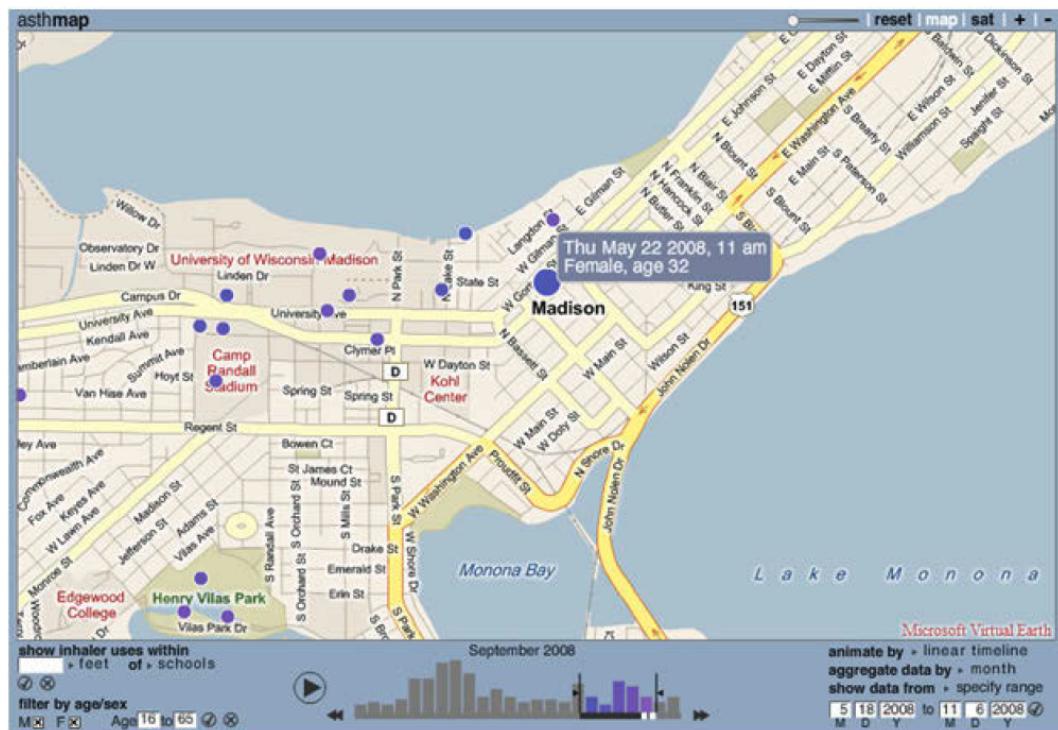


Abbildung 3.4: truliaHindsight [RR09]

### 3 Verwandte Arbeiten

---

Trulia Hindsight, dargestellt in Abbildung 3.4, visualisiert Daten bezüglich des Wohnungsbaus auf einer geographischen Karte, um die Entwicklung von Wohngebieten historisch nachvollziehbar zu machen. Diese Ereignisse werden als einfache, farbige Kreise dargestellt, welche auf der Karte gemäß den geographischen Koordinaten des entsprechenden Ortes positioniert sind. Die Kreise sind gemäß einer zeitlichen Legende eingefärbt, welche auch als interaktives Histogramm zur Verfügung steht. Die x-Achse gibt dabei den zeitlichen Verlauf, angegeben in Jahren, wieder, während die y-Achse die Frequenz der Ereignisse für das entsprechende Jahr darstellt. Der Verlauf der Jahre ist mit einem kontinuierlichen Farbschema gefärbt, wodurch jedem Balken des Histogramms eindeutig eine Farbe zugewiesen wird. Die Kreise, welche auf der Karte die Ereignisse visualisieren, sind entsprechend des jeweiligen Datums gefärbt. Der Anwender kann durch Interaktion mit dem Histogramm ein einzelnes Jahr oder eine Teilmenge der Jahre für die Visualisierung auswählen, wodurch lediglich die entsprechenden Ereignisse auf der Karte angezeigt werden. Außerdem kann der zeitliche Verlauf der Ereignisse visuell animiert werden. Die Zeitleiste ist bezüglich der x- und y-Dimensionen und der Animationsfunktionalität durchaus mit der Zeitleiste in meinem Ansatz vergleichbar. Dagegen sind die visualisierten Daten und damit auch die visuellen Elemente auf der Karte in Trulia Hindsight wesentlich primitiver. Eine Aggregation der Daten im Hintergrund oder der visuellen Elemente in der Visualisierung findet nicht statt.

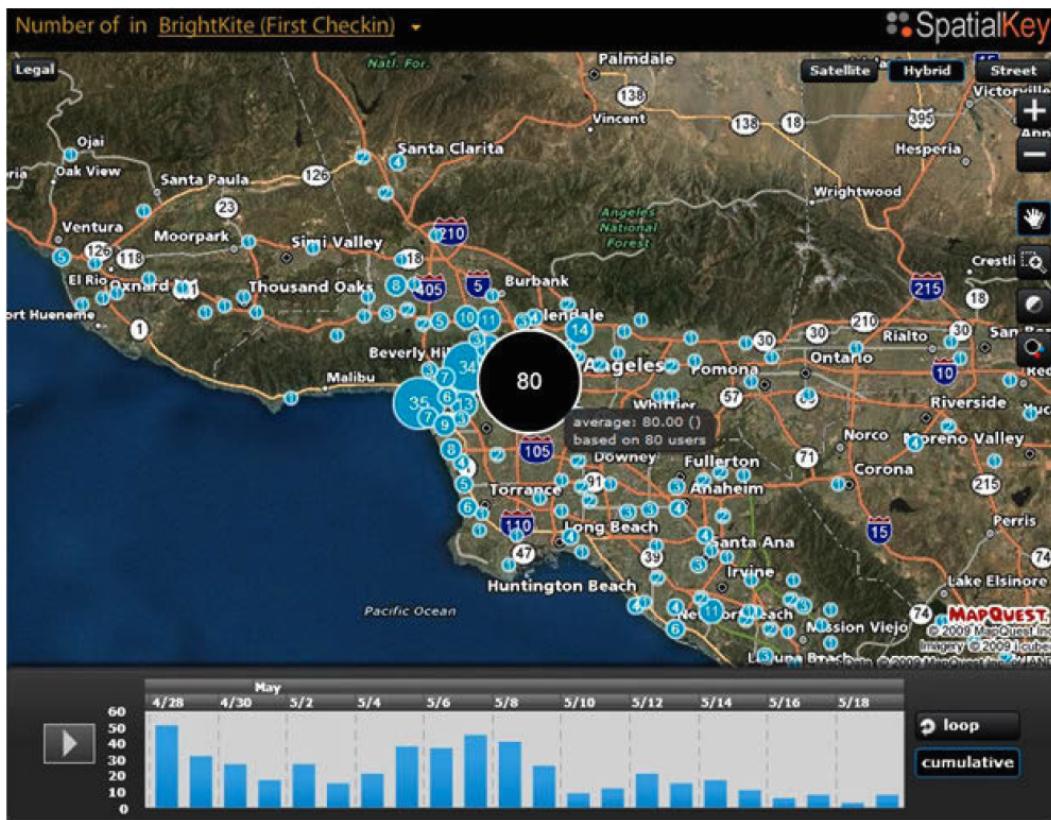


**Abbildung 3.5:** AsthMap [RR09]

AsthMap, dargestellt in Abbildung 3.5, visualisiert Daten bezüglich der Verschlimmerung der Krankheitsbilder von Asthmaerkrankungen geographisch auf einer Karte und im zeitlichen Verlauf. Grundlage für diesen Ansatz war das oben beschriebene Trulia Hindsight, weshalb beide Dienste hohe Ähnlichkeit aufweisen. Die Ereignisse werden hier ebenfalls als farbige Kreise dargestellt, wobei die

### 3.1 Multiple Coordinated Views und Geovisualisierung

Farbe wieder dem Datum der Ereignisse entspricht. Bei Bedarf werden mehrere Kreise zu größeren Kreisen aggregiert, so dass keine Kollisionen auftreten. Außerdem sind diese Kreise nun interaktiv, so dass durch Interaktion Detailinformationen über das Ereignis abgefragt werden können. Sowohl die Aggregation von mehreren Datenobjekten oder visuellen Elementen als auch die Darstellung von Detailinformationen auf Anfrage durch Interaktion mit diesen Elementen spielen in meinem Ansatz wichtige Rollen, wobei die Aggregation aufgrund der hierarchischen Struktur und der Vielfalt der Daten in meinem Ansatz wesentlich aufwändiger und komplexer ist. Über das Histogramm, welches ähnlich gestaltet ist wie in Trulia Hindsight, kann eine Zeitspanne für die Visualisierung auf der Karte selektiert werden. Außerdem können die Daten nach Raum, Zeit und Geschlecht und Alter der von den Ereignissen betroffenen Personen gefiltert werden. Die Animation des zeitlichen Verlaufs kann nach Bedarf konfiguriert werden. Auch in meinem Ansatz kann der Anwender einige Parameter und Optionen verändern, um die Animation der Zeitleiste anzupassen.

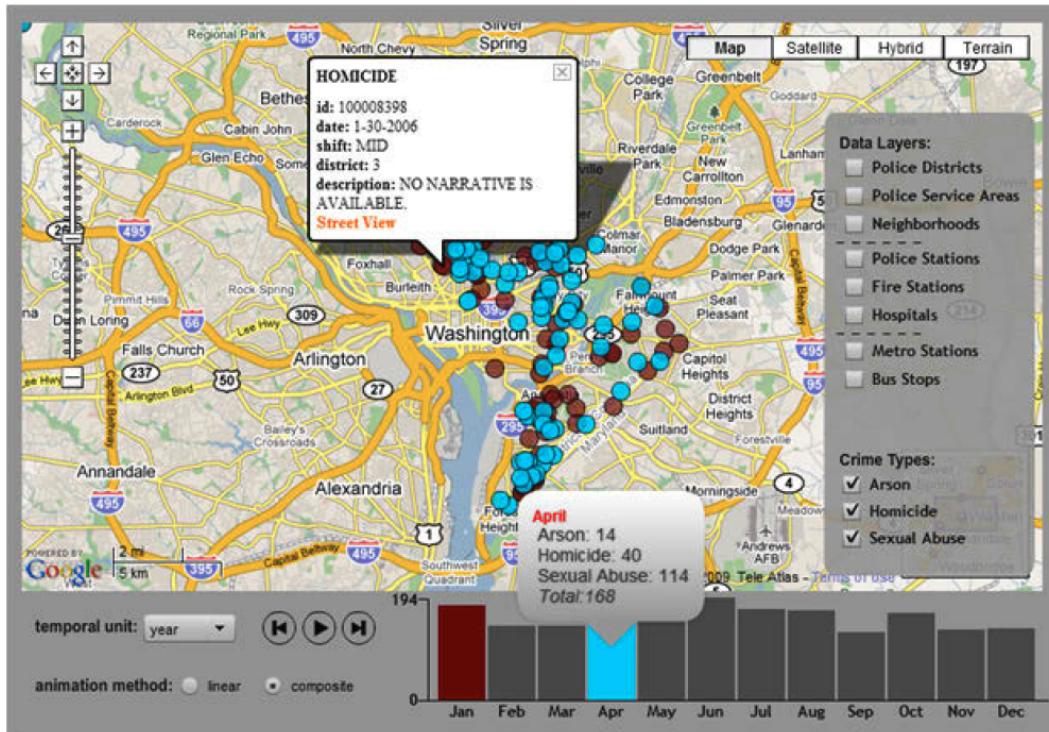


**Abbildung 3.6:** SpatialKey [RR09]

SpatialKey, entwickelt von Universal Mind und dargestellt in Abbildung 3.6, ist grundsätzlich sehr ähnlich wie die beiden oben beschriebenen Ansätze, allerdings können die Anwender hier ihre eigenen Daten hochladen, welche dann entsprechend auf der geographischen Karte und durch ein interaktives Histogramm visualisiert werden. Außerdem stehen drei verschiedene Verfahren für die Aggregation und die Darstellung der aggregierten Daten zur Verfügung: Eine Ansicht mit Kreisen, welche gemäß der Anzahl der aggregierten Daten verschieden groß sind, eine Heat Map-Ansicht und eine Heat Grid-Ansicht.

### 3 Verwandte Arbeiten

---



**Abbildung 3.7:** DC Crime Visualization [RR09]

DC Crime Visualization, dargestellt in Abbildung 3.7, wurde mittels der von den Autoren von [RR09] entwickelten Erweiterung der Google Maps API erstellt und weist ähnliche Merkmale und Komponenten auf wie die oben beschriebenen drei Ansätze.

Dabei werden Gewaltverbrechen im District of Columbia als Kreise auf der geographischen Karte dargestellt. Über die Interaktion mit den Kreisen können, wie auch bei AsthMap, Detailinformationen bezüglich des repräsentierten Ereignisses bezogen werden, wobei auch ein Link zu Google Street View angefügt ist.

Die Zeitleiste ist ebenfalls als interaktives Histogramm gestaltet, wie in den anderen vorgestellten Ansätzen. Dabei kann der Anwender die Zeiteinheit selber wählen. Durch die Auswahl eines Balkens innerhalb des Histogramms werden die entsprechenden Ereignisse auf der Karte visualisiert. Es können auch zwei Balken gleichzeitig ausgewählt werden, wobei die Ereignisse der beiden Balken durch verschiedene gefärbte Kreise auf der Karte dargestellt werden. Eine Aggregation der Daten beziehungsweise der visuellen Elemente, wie in AsthMap, SpatialKey oder wie in meinem Ansatz, findet nicht statt. Auch hier kann der zeitliche Verlauf animiert werden.

### 3.1.5 GeoTemCo

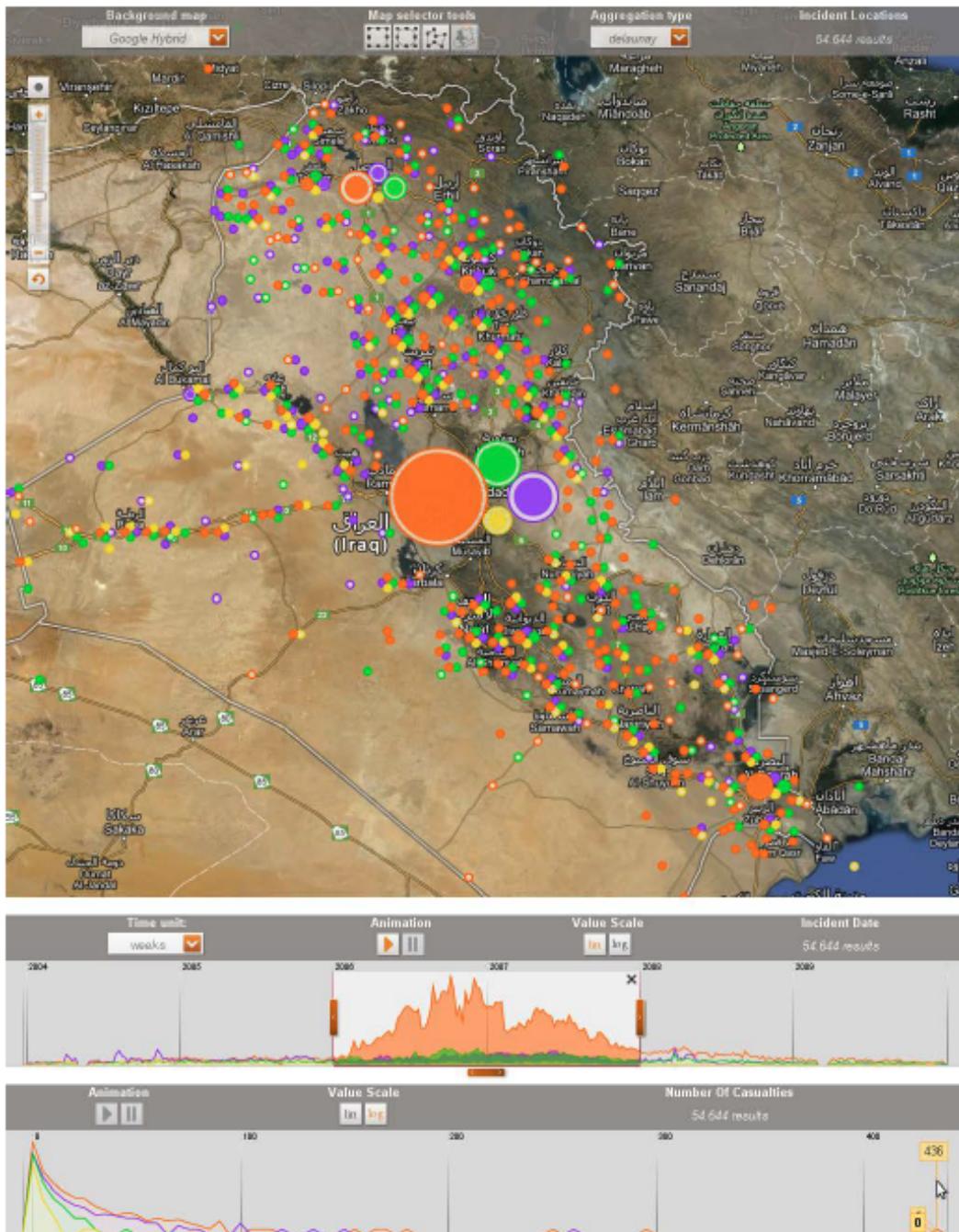


Abbildung 3.8: GeoTemCo [JHS13]

Jänicke et al. präsentieren in [JHS13] einen Geovisualisierungsansatz für die vergleichende Visualisierung von geographischen und temporalen Daten, dargestellt in Abbildung 3.8. Die Entwicklung

### 3 Verwandte Arbeiten

---

des Ansatzes wurde inspiriert durch VisGets, eine Arbeit von Dörk et al., welche in Abschnitt 3.1.3 diskutiert wurde. Die entwickelte Webanwendung erlaubt die Formulierung verschiedener Suchanfragen im räumlichen und temporalen Kontext und den visuellen Vergleich mehrerer Datensätze. Durch eine geeignete Aggregationsprozedur, welche auf dynamischer Delaunay-Triangulierung basiert, funktioniert der Ansatz auch für eine hohe Anzahl von Objekten gut.

Die Umsetzung basiert auf dem Ansatz der Multiple Coordinated Views. Dabei wurden eine Kartenansicht, eine Zeitansicht und eine abrufbare Tag Cloud implementiert. Jede der Ansichten verfügt über koordinierte Funktionen für die Navigation und Selektion, welche automatisch die anderen Ansichten aktualisieren.

Die verschiedenen Datensätze, welche durch den Ansatz vergleichbar visuell dargestellt werden, sind mit verschiedenen Farben assoziiert, die in allen Ansichten einheitlich verwendet werden. Die selektierten Daten werden hier mittels gesättigteren Versionen der entsprechenden Farben visuell hervorgehoben. Diese Hervorhebung wird auf Tabelleneinträge, Anteile von Glyphen innerhalb der Kartenansicht und Anteile von Graphen innerhalb der Zeitleiste angewandt. Für meine Arbeit wurde ein ähnlicher Ansatz gewählt, um die selektierten und nicht selektierten Anteile von Kreisen beziehungsweise Knoten innerhalb der Kartenansicht hervorzuheben. Die selektierten Anteile sind dabei in einer der entsprechenden religiösen Gruppe oder der entsprechenden Religion zugewiesenen Farbe gefärbt, wobei die nicht selektierten Anteile dieselbe Farbe verwenden, allerdings wurde dabei die Opazität deutlich verringert und eine transparente, graue Fläche darübergelegt. In GeoTemCo werden die georeferenzierten Daten generell durch farbige Kreise auf der geographischen Karte dargestellt. Der selektierte Anteil wird dabei durch einen kleineren Kreis innerhalb des Kreises dargestellt, welcher, wie oben beschrieben, mit einer gesättigteren Version der Farbe gefärbt ist. Dieser Ansatz der visuellen Unterscheidung des selektierten und des nicht selektierten Anteils stand für meine Arbeit ebenfalls zur Debatte, allerdings entschied ich mich für eine andere Umsetzung, da hier nicht intuitiv und eindeutig erkennbar ist, ob das Verhältnis der Flächen oder das Verhältnis der Radien der beiden konzentrisch angeordneten Kreise das Verhältnis zwischen den beiden Anteilen angibt. Bei beiden Varianten ist die Abschätzung des Verhältnisses zwischen den beiden Anteilen schwierig. Mein Ansatz visualisiert diese beiden Anteile durch jeweils ein Kuchendiagramm und ein Donutdiagramm, wobei die selektierten und nicht selektierten Kuchen- und Donutstücke wie oben erläutert voneinander unterscheidbar gestaltet sind. Generell sind Kuchen- und Donutdiagramme ebenfalls nicht einfach zu interpretieren, wie in Abschnitt 2.1 angesprochen, allerdings erschienen mir diese Diagramme die beste Wahl für meinen Ansatz zu sein, wie in Abschnitt 5.3.1 erläutert.

Für die Vermeidung von Kollisionen zwischen den Kreisen wurde ein Verfahren für die Aggregation dieser entwickelt, welches auf der Delaunay-Triangulierung und dem Vergleich der Kreise bezüglich Größe und Distanz basiert. Die Aggregation in meinem Ansatz ist weniger komplex und basiert lediglich auf dem Vergleich der Größe und Distanz zwischen einzelnen Knoten und dem abgeschätzten Platzbedarf von komplexeren Bäumen mit mehr als zwei Knoten. In beiden Ansätzen wird die Anzahl der aggregierten Objekte beziehungsweise Kreise durch die Größe der resultierenden Kreise oder Knoten repräsentiert.

Wenn in GeoTemCo mehrere Datensätze vergleichend visualisiert werden sollen, wird für jeden Datensatz und für jedes georeferenzierte Datenobjekt darin ein Kreis mit der entsprechenden Färbung erstellt. Die Kreise der verschiedenen Datensätze, welche dieselben geographischen Koordinaten referenzieren, werden dabei zu sogenannten Kreisgruppen zusammengefasst, was bedeutet, dass die

### 3.1 Multiple Coordinated Views und Geovisualisierung

---

Kreise nebeneinander um die geographische Position herum angeordnet werden. Die Visualisierung innerhalb der Kartenansicht in meinem Ansatz ergibt ein ähnliches Bild, wobei dabei nicht verschiedene Datensätze, sondern verschiedene Religionen oder verschiedene religiöse Gruppen durch verschieden gestaltete Kreise repräsentiert werden. Diese Kreise sind als Baumstruktur angeordnet, welche durch ein kräftebasiertes Layout angeordnet wird.

Durch das Anklicken visueller Elemente, das Zeichnen eines Kreises, Rechtecks oder Polygons oder das Anklicken einer geographischen Region können Elemente innerhalb der Kartenansicht selektiert werden. In meinem Ansatz kann ebenfalls ein Polygon gezeichnet werden, um eine geographische Region zu spezifizieren und um alle Datenobjekte, welche darin liegende Koordinaten referenzieren, zu selektieren. Wird in GeoTemCo ein Kreis innerhalb der Kartenansicht ausgewählt, so erscheint in der Ansicht eine Tag Cloud, welche die Namen der Städte, welche durch den Kreis repräsentiert werden, darstellt. Wird in meinem Ansatz ein Knoten durch den Mauszeiger ausgewählt, so erscheinen auf der Karte kleine, rote Punkte, welche die exakten geographischen Lagen der Städte, welche durch den Knoten repräsentiert werden, visualisieren.

Die Zeitleiste ist als überlappendes, segmentiertes Flächendiagramm umgesetzt, wobei die x-Achse den zeitlichen Verlauf und die y-Achse die Anzahl der Objekte darstellt, welche den entsprechenden Wert der x-Achse referenzieren. Die Zeitleiste ist dabei in Intervalle gleicher Länge eingeteilt. Selektiert werden kann hier entweder ein vordefiniertes Zeitintervall, oder eine mittels Slidern spezifizierte Zeitspanne. Die selektierte Zeitspanne kann animiert werden. Die Wahl der Dimensionen in diesem Ansatz entspricht den Dimensionen in meiner Arbeit, wobei in letzterer für die Darstellung der Zeitleiste ein gestapeltes Säulendiagramm verwendet wurde.

#### 3.1.6 Threatwiki

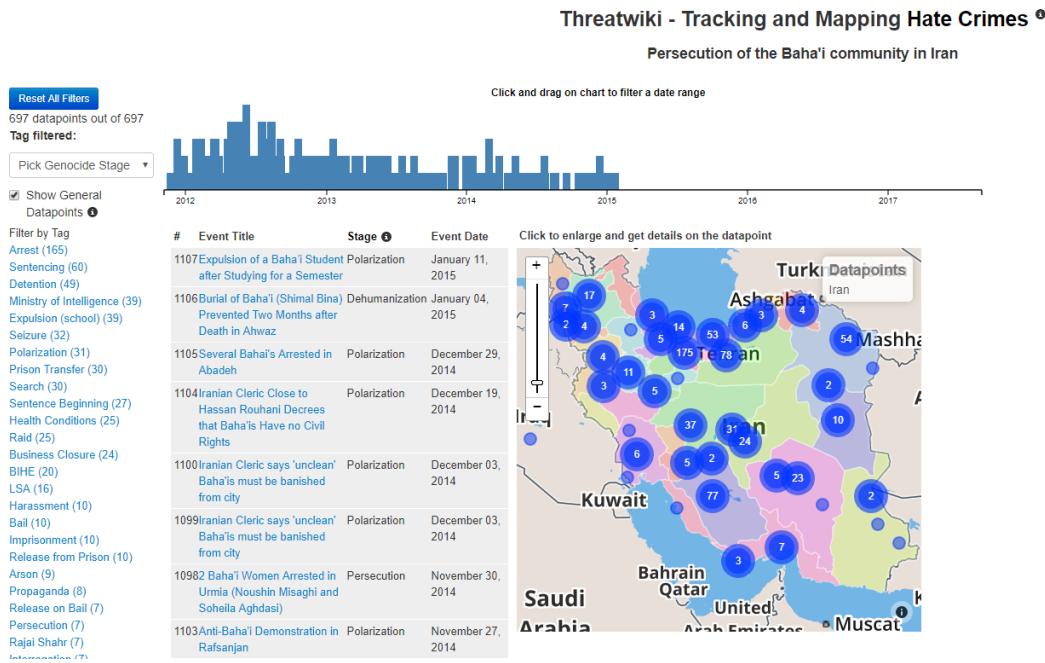
Die Webanwendung 'Threatwiki - Tracking and Mapping Hate Crimes'<sup>2</sup>, entwickelt von 'THE SENTINEL PROJECT FOR GENOCIDE PREVENTION' und dargestellt in Abbildung 3.9, visualisiert geografisch und temporal referenzierte Daten bezüglich Hassverbrechen im Iran. Die Anwendung basiert auf dem Konzept der Multiple Coordinated Views. Dabei wurden unter anderem eine Kartenansicht, eine Zeitleiste, eine Liste der visualisierten Events und eine Tagliste entwickelt.

Die geographische Karte kann manuell verschoben und gezoomt werden. Die Ereignisse werden mittels blauer Kreise visualisiert, welche entsprechend der referenzierten geographischen Koordinaten auf der Karte positioniert werden. Falls sich mehrere Kreise zu weit überschneiden, werden diese automatisch aggregiert. Diese Aggregation geschieht dynamisch zur Laufzeit basierend auf den Kreisen, welche gemäß der aktuellen Selektion in allen Ansichten auf der Karte vorhanden sind. In meiner Arbeit werden nicht selektierte Informationen nicht entfernt, sondern entsprechende Anteile an visuellen Elementen transparent und ausgegraft gestaltet, wodurch die Aggregation der Knoten nur ein Mal zu Beginn der Laufzeit berechnet werden muss. Kreise, welche das Ergebnis von Aggregation sind, sind wesentlich größer als Kreise, welche lediglich ein Ereignis visualisieren. Dabei

<sup>1</sup><http://vast-journey-7849.herokuapp.com/iranvisualization>

<sup>2</sup><http://vast-journey-7849.herokuapp.com/iranvisualization>

### 3 Verwandte Arbeiten



**Abbildung 3.9:** Threatwiki<sup>1</sup>

hängt, im Gegensatz zu meinem Ansatz, die Größe nicht von der Anzahl der aggregierten Kreise ab. Diese Anzahl wird allerdings als Zahl in der Mitte der Kreise dargestellt.

Während die Knoten innerhalb der Kartenansicht in meiner Arbeit Aufschluss über die hierbei repräsentierten religiösen Gruppen oder Religionen geben, können an den blauen Kreisen keine weiteren Informationen über die entsprechenden Ereignisse direkt abgelesen werden. Wird ein Kreis, welcher aggregierte Kreise repräsentiert, angeklickt, so zoomt die Anwendung automatisch weiter in die Karte hinein, so dass ersichtlich wird, aus welchen Kreisen, welche gegebenenfalls wieder Resultate von Aggregationen sind, dieser Kreis zusammengesetzt ist. Zoomt der Anwender manuell in die Karte hinein oder aus ihr heraus, so werden Knoten weiter aggregiert oder spalten sich in mehrere Knoten auf, was mittels einer Animation visualisiert wird.

Wird ein Kreis, welcher nicht das Ergebnis von Aggregation ist, angeklickt, so werden Detailinformationen über das entsprechende Ereignis, wie beispielsweise das Datum, der Ort und eine kurze Beschreibung, in einem Popup-Fenster angezeigt.

Die Zeitleiste ist als Säulendiagramm umgesetzt, wobei die x-Achse den zeitlichen Verlauf und die y-Achse die Anzahl der Ereignisse des entsprechenden Zeitpunktes darstellt. Die Balken stellen dabei lediglich die Anzahl der Ereignisse, nicht aber weitere Informationen über diese dar. Im Gegensatz dazu wurde in meinem Ansatz für die Visualisierung des zeitlichen Verlaufs ein gestapeltes Säulendiagramm verwendet, um mittels der gestapelten Säulen Aufschluss über die Zusammensetzung des dargestellten y-Wertes zu geben. Wie auch in meiner Arbeit kann mittels eines Sliders eine Zeitspanne selektiert werden.

### 3.1 Multiple Coordinated Views und Geovisualisierung

Die Tagliste, die Zeitleiste und die Kartenansicht sind hierbei koordiniert, so dass die Selektion innerhalb einer Ansicht die Inhalte der anderen Ansichten entsprechend aktualisiert.



# 4 Konzept

In diesem Kapitel wird das Konzept des entwickelten Ansatzes beschrieben. Zunächst werden die Anforderungen an einen solchen Ansatz dargestellt. Anschließend wird der konkrete Entwurf anhand der zugrundeliegenden Daten und der Benutzeroberfläche erläutert.

## 4.1 Anforderungen

In dieser Arbeit wurde ein MCV-Ansatz zur interaktiven Visualisierung und Erforschung georeferenzieller, historischer und hierarchischer Daten entwickelt. Dieses Kapitel skizziert die Anforderungen an diesen Ansatz.

### 4.1.1 Verschiedene Ansichten

Das System soll einen MCV-Ansatz implementieren. Um die Komplexität des Datensatzes geeignet darstellen zu können, soll für jede relevante Dimension des Datensatzes eine Ansicht mit entsprechender visueller Repräsentation der Daten existieren. Diese Ansichten sollen miteinander logisch verknüpft beziehungsweise koordiniert sein, um den Anwender dabei zu unterstützen, ein tiefes Verständnis für den Datensatz zu erarbeiten und um das Auffinden von komplexen Zusammenhängen zu ermöglichen.

### 4.1.2 Visualisierung der Daten

Für die verschiedenen Ansichten beziehungsweise für die verschiedenen Dimensionen müssen geeignete Visualisierungsformen gefunden werden, welche die entsprechenden Daten intuitiv und verständlich darstellen. Außerdem müssen passende Abstraktions- und Aggregationsverfahren angewandt werden, um die Quantität und Komplexität des Datensatzes zu vereinfachen, ohne, dass damit ein zu großer Informationsverlust einhergeht. Dabei ist im Rahmen dieser Arbeit vor allem die hierarchische Struktur der religiösen Gruppen zu erwähnen. Die gewählten Visualisierungsformen müssen so gestaltet werden, dass der Zustand des Systems jederzeit über entsprechende visuelle Merkmale repräsentiert ist. Der Zustand des Systems entspricht hier den einzelnen Selektionen in den verschiedenen Ansichten und der damit einhergehenden Filterkonfiguration im Hintergrund. Wichtig ist dabei die grafische Darstellung der selektierten und nicht selektierten Anteile aller Merkmale der einzelnen Dimensionen und Ansichten. Für die korrekte Interpretation des Ansatzes durch den Anwender ist von großer Bedeutung, dass die Repräsentationen in den verschiedenen Ansichten untereinander und bezüglich der gefilterten Daten zu jedem Zeitpunkt konsistent sind. Andernfalls

## 4 Konzept

---

könnte der Anwender falsche Schlussfolgerungen ziehen, da in einem MCV-Ansatz die verschiedenen Ansichten immer auch kombiniert betrachtet und interpretiert werden müssen, um den Datensatz zu erforschen.

### 4.1.3 Interaktion

Für jede Ansicht müssen geeignete und intuitive Interaktionsmöglichkeiten existieren. In dieser Arbeit werden hauptsächlich Funktionen für die Selektion und das Hervorheben visueller Elemente und die Filterung der Daten im Hintergrund benötigt. Dadurch soll die interaktive Erforschung der Daten ermöglicht werden. Über bestimmte Ansichten und direkte und indirekte Interaktion soll der Datensatz im Hintergrund gemäß der verschiedenen Dimensionen gefiltert werden können. Dabei sind die Selektionen der verschiedenen Ansichten miteinander zu verknüpfen, so dass der Filter letztendlich die Schnittmenge aller selektierten Teilmengen zurückliefert. Durch die Selektion innerhalb einer Ansicht sollen alle anderen Ansichten automatisch gemäß der neuen Filterkonfiguration aktualisiert werden. Außerdem sollen bestimmte Ansichten über Interaktionsmöglichkeiten verfügen, um entsprechende Daten beziehungsweise visuelle Elemente in anderen Ansichten hervorzuheben.

### 4.1.4 Benutzeroberfläche

Die Benutzeroberfläche muss so gestaltet sein, dass sie die Analysearbeit unterstützt und fehlerhafte Aktionen seitens des Anwenders vermieden werden. Alle Ansichten und grafischen Repräsentationen sollen hierfür bezüglich bestimmter Aspekte, wie zum Beispiel der Auswahl und Zuweisung von Farben für bestimmte Daten, einheitlich beziehungsweise ähnlich gestaltet sein. Außerdem sollen Funktionen zum Verstecken und zum Vergrößern von Ansichten entwickelt werden, um die Oberfläche an die Bedürfnisse des Anwenders anpassen zu können. Allgemein muss die Benutzeroberfläche eines Ansatzes zur interaktiven Datenvisualisierung sicherstellen, dass die Aufmerksamkeit des Benutzers jederzeit auf die zu diesem Zeitpunkt relevanten Ansichten gelenkt wird.

## 4.2 Entwurf

Im vorangehenden Abschnitt wurden die Anforderungen und die gewünschten funktionalen Aspekte des Ansatzes beschrieben. Dieser Abschnitt skizziert den Entwurf des Ansatzes, während das folgende Kapitel die konkrete Umsetzung dieses Entwurfs detailliert beschreibt.

### 4.2.1 Datensätze

Das System liest zunächst drei Datensätze ein, welche als JSON-Dateien (JavaScript Object Notation<sup>1</sup>) vorliegen.

<sup>1</sup><http://json.org/>

Der erste Datensatz besteht aus den zu visualisierenden Daten, während die anderen beiden Datensätze von diesem referenziert werden und zusätzliche Daten und Beziehungen zwischen Elementen beschreiben. Dieser Datensatz besteht aus einem Array folgendermaßen strukturierter JavaScript-Objekte:

```
{
  Stadt: Name,
  Koordinaten: [Breitengrad, Längengrad],
  Gruppe: Name,
  Zeitspanne: [Beginn, Ende],
  Quelle: Name,
  ObjektID: ID
}
```

Nachfolgend werden die einzelnen Eigenschaften kurz erklärt:

- Stadt: Der Name der Stadt, in der die entsprechende religiöse Gruppe angesiedelt war.
- Koordinaten: Die Koordinaten der entsprechenden Stadt, angegeben in geographischer Breite (Breitengrad, Latitude) und Länge (Längengrad, Longitude).
- Gruppe: Der Name der religiösen Gruppe, welche in der entsprechenden Stadt angesiedelt war.
- Zeitspanne: Die Zeitspanne, in der die entsprechende religiöse Gruppe in der entsprechenden Stadt angesiedelt war, angegeben in Jahren.
- Quelle: Der Name der literarischen Quelle, aus welcher die Informationen dieses Datenobjektes extrahiert wurden.
- ObjektID: Die eindeutige ID dieses Datenobjektes. Diese ist lediglich für Implementierungsdetails relevant.

Somit beschreiben die enthaltenen Datenobjekte jeweils den Ort und die Zeitspanne der Ansiedlung beziehungsweise Aktivität einer bestimmten religiösen Gruppe und die literarische Quelle dieser Informationen.

Der zweite Datensatz beschreibt die hierarchischen Beziehungen zwischen den referenzierten religiösen Gruppen. Er enthält hierarchisch verschachtelte Objekte folgender Struktur:

```
{
  Gruppe: Name,
  Abkürzung: Abkürzung,
  Untergruppen: [
    {...},
    {...},
    ...
  ]
}
```

## 4 Konzept

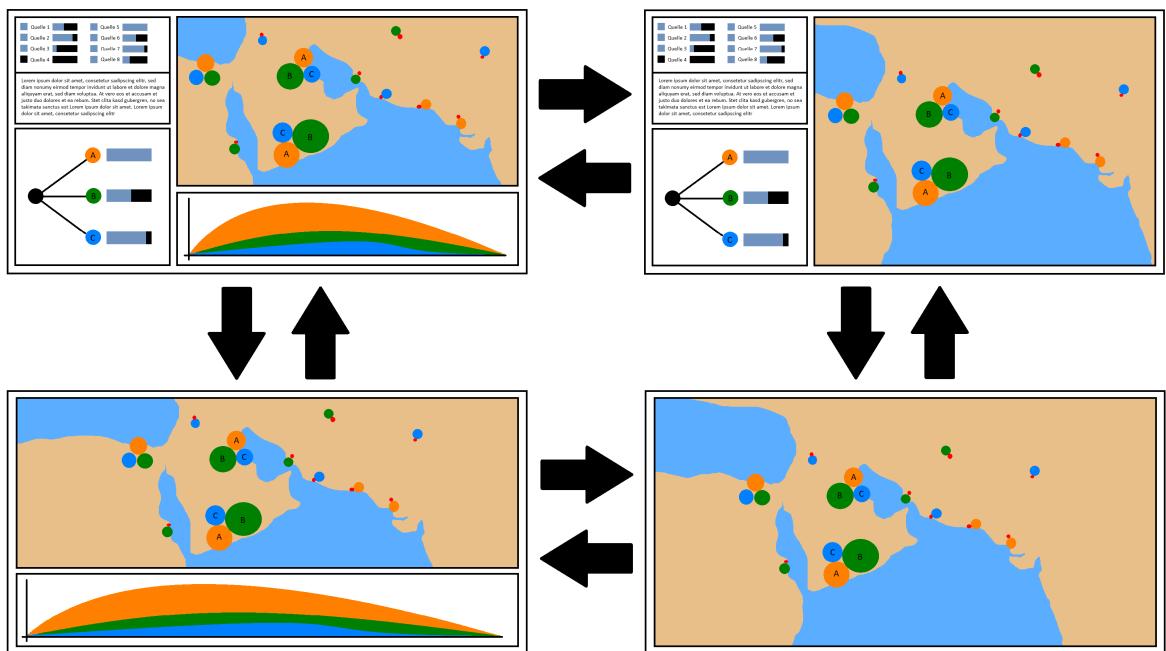
Jedes dieser Objekte beschreibt also den Namen der entsprechenden religiösen Gruppe, den abgekürzten Namen und deren Untergruppen, welche wiederum Objekte derselben Struktur sind.

Der dritte Datensatz enthält die Informationen bezüglich der im ersten Datensatz referenzierten literarischen Quellen. Für jede Quelle existiert ein Datenobjekt folgender Struktur, wobei diese Objekte in einem Array enthalten sind:

```
{  
    QuellenID: ID,  
    Titel: Titel,  
    Autoren: [...],  
    Text: Text  
}
```

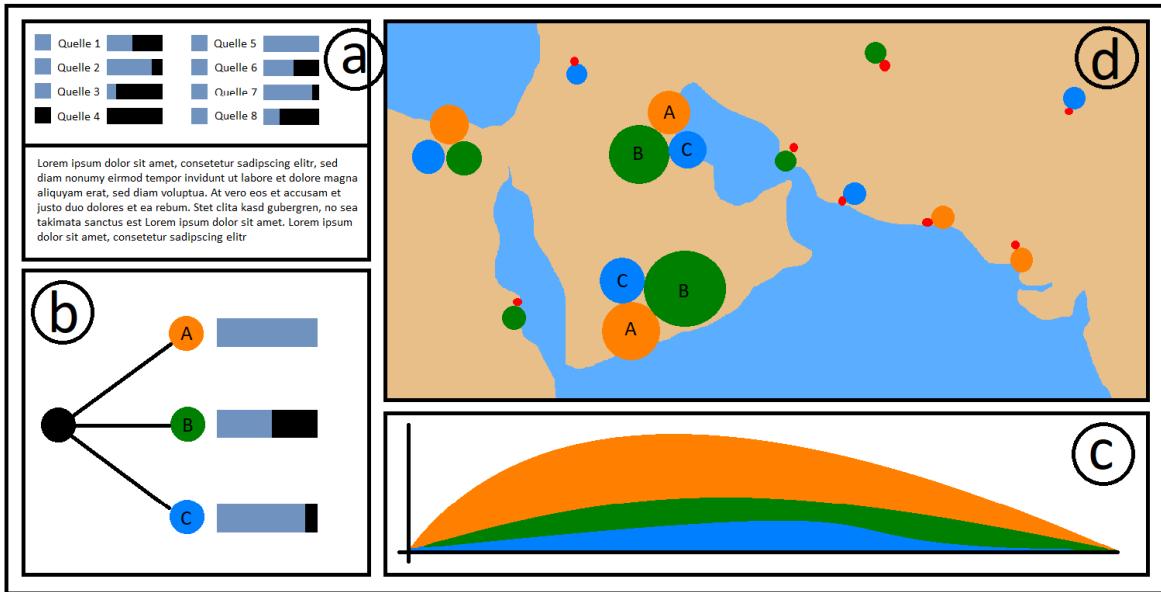
Ein solches Datenobjekt besteht also aus einer einzigartigen ID für die jeweilige Quelle, dem Titel, einer Liste der Autoren und dem tatsächlichen Textinhalt der Quelle.

### 4.2.2 Benutzeroberfläche



**Abbildung 4.1:** Die linke Spalte, welche aus der Quellenansicht und der Gruppenansicht besteht, und die Zeitleistenansicht können jeweils mittels eines Buttons eingeklappt oder ausgeklappt werden, wobei sich die Kartenansicht gemäß des verfügbaren Platzes ausdehnt.

Die Benutzeroberfläche besteht aus einer Menüleiste, welche Buttons für verschiedene Konfigurationsaktionen zur Verfügung stellt, und dem Hauptteil, der die verschiedenen Ansichten enthält. So kann durch entsprechende Buttons die linke Spalte, welche zwei Ansichten enthält, und die untere Ansicht in der rechten Spalte, jeweils manuell versteckt oder angezeigt werden. Werden diese Ansichten versteckt, so nimmt die Kartenansicht den vollen verfügbaren Platz ein. Dieser Sachverhalt ist in Abbildung 4.1 dargestellt.



**Abbildung 4.2:** Entwurf der graphischen Benutzeroberfläche, bestehend aus der Quellenansicht (a), der Gruppenansicht (b), der Zeitleistenansicht (c) und der Kartenansicht (d).

Alle Ansichten enthalten visuelle Repräsentationen von Daten, mit welchen direkt oder indirekt interagiert werden kann, um bestimmte visualisierte Daten zu selektieren. Durch diese Selektion wird der zugrundeliegende Datensatz gefiltert und die anderen Ansichten werden entsprechend aktualisiert. Werden visuelle Elemente in der Kartenansicht ausgewählt, so werden entsprechende Elemente in den anderen Ansichten hervorgehoben.

Nachfolgend werden die einzelnen Ansichten kurz skizziert. Die detaillierte Beschreibung ist in Abschnitt 5.3 enthalten. Es stehen folgende vier Ansichten für die Selektion von Daten zur Verfügung:

- **Quellenansicht:** Die Quellenansicht, skizziert in Abbildung 4.2 (a), listet die in dem ersten Datensatz referenzierten literarischen Quellen auf. Jede Quelle ist durch eine Checkbox selektierbar. Der selektierte und nicht selektierte Anteil einer Quelle ist durch einen entsprechenden Prozentbalken dargestellt. Durch das Anklicken des Quellennamens wird der Quellentext, welcher im dritten Datensatz gespeichert ist, im unteren Abschnitt der Ansicht angezeigt.
- **Gruppenansicht:** Die Hierarchie der religiösen Gruppen, welche im zweiten Datensatz beschrieben ist, wird in dieser Ansicht als horizontale Baumstruktur dargestellt. Die Gruppenansicht ist skizziert in Abbildung 4.2 (b) dargestellt. Für jede Gruppe existiert ein Baumknoten, welcher repräsentativ für den jeweiligen, anhängenden Teilbaum selektiert werden kann. Wie auch bei

## 4 Konzept

---

den Quellen zeigt hier für jeden Baumknoten ein Prozentbalken den selektierten und nicht selektierten Anteil aller Objekte an, welche die entsprechende religiöse Gruppe referenzieren.

- Zeitleistenansicht: Der Zeitaspekt wird durch die Zeitleiste, welche durch ein gestapeltes Säulenlendiagramm realisiert ist, dargestellt. Diese Ansicht ist skizziert in Abbildung 4.2 (c) dargestellt. Die zwei Dimensionen repräsentieren die Zeitspanne mit jeweils einem Wert für jedes Jahr, die Anzahl der Objekte, welche einen Zeitabschnitt referenzieren, der das entsprechende Jahr enthält, und die Segmentierung dieser Datenobjekte in Abschnitte für jede referenzierte Religion. Die Selektion funktioniert hier über die Auswahl einer Zeitspanne. Der zeitliche Verlauf kann animiert werden.
- Kartenansicht: Die Kartenansicht, skizziert in Abbildung 4.2 (d), bildet das Herzstück des gesamten Ansatzes und dient der geographischen Darstellung der Daten. Die Städte und Gruppen werden hierbei durch Knoten kleiner, simpler Baumstrukturen dargestellt, welche entsprechend der zugrundeliegenden Koordinaten geographisch platziert werden. Bei Bedarf werden mehrere Knoten und Bäume aggregiert und die Gruppenrepräsentation gemäß der im zweiten Datensatz enthaltenen Hierarchie abstrahiert. Durch Interaktion mit den visualisierten Objekten werden entsprechende Informationen in den anderen Ansichten hervorgehoben. Durch das Zeichnen eines Polygons kann eine geographische Region selektiert werden.

Der Selektionsmechanismus läuft bei allen Ansichten nach folgendem Schema ab:

1. Der Anwender selektiert durch direkte oder indirekte Interaktion mit den visuellen Elementen bestimmte Datenobjekte.
2. Diese neue Selektion wird nun mit den anderen Selektionen innerhalb der selben Ansicht kombiniert. Mehrere unabhängige Selektionen innerhalb einer Ansicht sind in der Quellenansicht und in der Gruppenansicht möglich.
3. Die Selektion innerhalb dieser Dimension wird an den Filter übermittelt, welcher nun die Selektionen aller Dimensionen miteinander verknüpft, indem er die Schnittmenge daraus bildet.
4. Die Visualisierung wird nun in allen Ansichten gemäß der neuen Filterdaten aktualisiert.

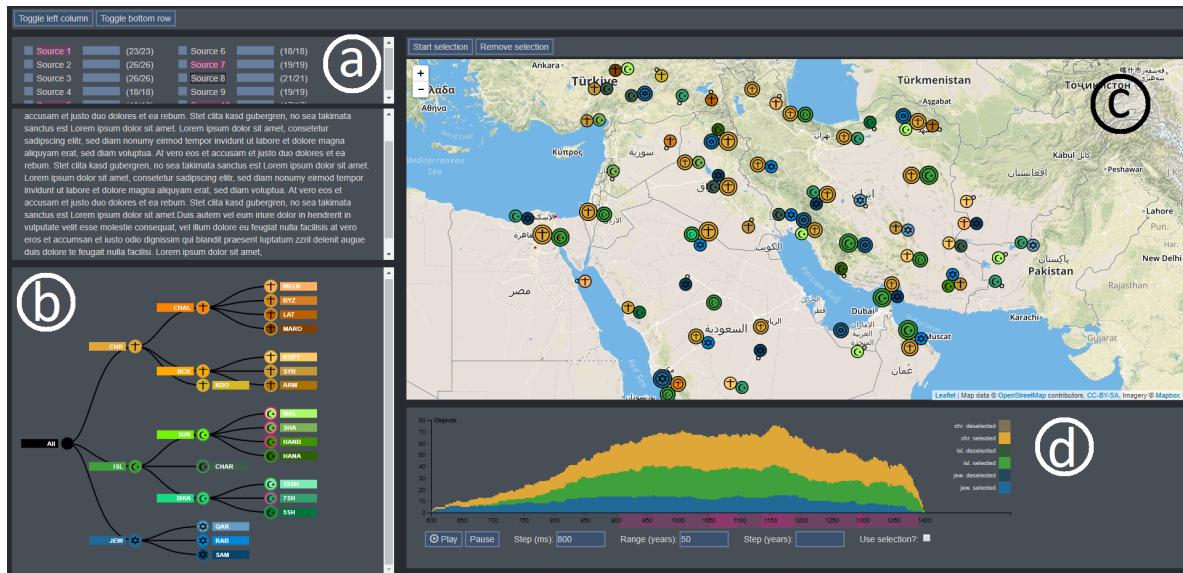
Die Interaktion mit visualisierten Objekten in der Kartenansicht läuft in zwei Schritten ab:

1. Der Anwender wählt ein Objekt mit dem Mauszeiger aus.
2. Die entsprechenden Elemente werden in allen Ansichten hervorgehoben.

# 5 Umsetzung

In diesem Kapitel wird detailliert die Umsetzung des im vorangegangenen Kapitel spezifizierten Ansatzes beschrieben. Zunächst werden eine kurze Übersicht skizziert und die verwendeten Technologien aufgeführt. Anschließend wird der entwickelte Ansatz im Detail anhand der verschiedenen Ansichten und deren Verbindungen untereinander erläutert.

## 5.1 Übersicht und Technologie



**Abbildung 5.1:** Graphische Benutzeroberfläche, bestehend aus der Quellenansicht (a), der Gruppenansicht (b), der Kartenansicht (c) und der Zeitleistenansicht (d).

Als Einleitung soll hier eine grobe Übersicht des Ansatzes skizziert und eine Übersicht der Verwendeten Technologien gegeben werden. Der Ansatz ist als Single-Page-Anwendung umgesetzt. Die Struktur, das Verhalten und das Styling der Anwendung wurden mittels HTML, JavaScript und CSS umgesetzt. Für die responsive Designmerkmale wurde das Grid System von Bootstrap<sup>1</sup> eingesetzt. Dieses ermöglicht die Strukturierung des Layouts mittels Spalten und Zeilen, welche sich automatisch der

<sup>1</sup><http://getbootstrap.com/>

## 5 Umsetzung

---

Größe des Displays anpassen. Mittels JQuery.js<sup>2</sup> wurden primitive Funktionen zur Manipulation der Struktur der Anwendung durch den Anwender umgesetzt. Für die einzelnen Grafiken und Diagramme wurde D3.js (Version 4)<sup>3</sup> verwendet. Alle vier Ansichten dienen der Darstellung und der interaktiven Selektion und Filterung verschiedener Aspekte des selben Datensatzes. Die Benutzeroberfläche ist in Abbildung 5.1 dargestellt.

- Die Quellenansicht, dargestellt in Abbildung 5.1 (a), stellt eine Liste der referenzierten Quellen zur Verfügung, über welche einzelne Quellen selektiert und gefiltert werden können.
- Die Gruppenansicht, dargestellt in Abbildung 5.1 (b), verfügt über einen graphischen Baum, welcher die Hierarchie der religiösen Gruppen darstellt. Anhand der einzelnen Baumknoten können Gruppen selektiert und gefiltert werden.
- Die Zeitleistenansicht, dargestellt in Abbildung 5.1 (d), stellt eine Zeitleiste zur Verfügung, welche den zeitlichen Verlauf und die Anwesenheit der verschiedenen Religionen für jedes einzelne Jahr darstellt. Der Anwender kann innerhalb der Zeitleiste eine Zeitspanne selektieren. Diese Ansicht ist rechts unten im Layout positioniert.
- Über der Zeitleistenansicht befindet sich die Kartenansicht, dargestellt in Abbildung 5.1 (c). Diese Ansicht visualisiert die Daten auf einer geographischen Karte. Durch das Zeichnen eines Polygons kann eine geographische Region selektiert werden. Für die Einbindung der interaktiven Karte wurde Leaflet.js<sup>4</sup> verwendet.

Jede der vier Ansichten steht in Beziehung zu jeder anderen Ansicht des Ansatzes. Die dafür relevante Filterung der Datensätze wurde mittels Crossfilter.js<sup>5</sup> realisiert.

### 5.2 Aggregationsprozedur

Um die Überladung der Karte mit visuellen Elementen und deren visuelle Überlappung auf der Karte zu vermeiden, wurde eine primitive Aggregationsprozedur entwickelt und eingesetzt. Ist in anderen Abschnitten dieser Arbeit die Rede von Aggregation, so sind damit diese Prozedur und deren Ergebnisse gemeint. Nachfolgend soll diese Prozedur beschrieben werden.

Die entwickelte Aggregationsprozedur wird nacheinander für jede dem Anwender zur Verfügung stehenden Zoomstufe durchgeführt. Der folgende Abschnitt beschreibt die Durchführung für eine dieser Zoomstufen.

Zunächst liegen die einzelnen Datenobjekte vor, aufgelistet in einem Array. Die Struktur dieser Datenobjekte ist in Abschnitt 4.2.1 beschrieben. Zu Beginn der Prozedur wird jedes einzelne dieser Objekte in einer eigenen Liste gespeichert. Diese Listen, welche zunächst jeweils nur ein Objekt enthalten, werden hier *Cluster* genannt. Anschließend wird jedes einzelne Datenobjekt mit jedem

<sup>2</sup><https://jquery.com/>

<sup>3</sup><https://d3js.org/>

<sup>4</sup><http://leafletjs.com/>

<sup>5</sup><http://square.github.io/crossfilter/>

anderen verglichen basierend auf den jeweils referenzierten geographischen Koordinaten. Liegt die Distanz der beiden Koordinaten unter einem bestimmten Schwellenwert, so werden die beiden Cluster, in welchen die beiden Objekte enthalten sind, miteinander vereint. Der hier verwendete Schwellenwert wurde heuristisch ermittelt und entsprechend festgelegt. Nach dieser Prozedur liegen mehrere Cluster vor, welche jeweils mindestens ein Datenobjekt enthalten. Aus jedem dieser Cluster wird nun ein *Clusterobjekt* erstellt, indem die Werte der Attribute der einzelnen, enthaltenen Datenobjekte geeignet zusammengefasst werden. So wird zum Beispiel der geometrische Mittelpunkt aus allen, von den Objekten referenzierten Koordinaten berechnet. Außerdem werden alle referenzierten Quellen, Zeitspannen und religiösen Gruppen strukturiert in Listen gespeichert. Diese Clusterobjekte besitzen nun bereits die Struktur, die von der Visualisierungsprozedur verwendet wird, um die Visualisierung innerhalb der Kartenansicht zu erzeugen. Jedes dieser Objekte wird dabei mittels einer Baumstruktur visualisiert, wie sie in Abschnitt 5.3.1 beschrieben wird.

Da die Ergebnisse dieser ersten Aggregationsprozedur nicht ausreichend zufriedenstellend waren, wurde im Laufe der Entwicklung eine zweite Prozedur entwickelt, welche die Resultate der ersten verarbeitet. Diese Prozedur funktioniert ähnlich wie die erste, wobei hier die Clusterobjekte, und nicht die atomaren Datenobjekte aus dem Datensatz, geclustert und aggregiert werden. Dabei wird nun jedes Clusterobjekt mit jedem anderen verglichen basierend auf der Distanz der errechneten Mittelpunkte der Koordinaten und der geschätzten Größe der einzelnen Cluster. Die Größe eines Clusters, interpretiert als Radius der resultierenden Visualisierung, wird näherungsweise aus der Anzahl der aggregierten Datenobjekte berechnet. Ist die Distanz der Mittelpunkte nun geringer als die Summe der beiden Radien, so werden die beiden Cluster, in welchen die beiden Clusterobjekte enthalten sind, miteinander vereint. Nach dieser Prozedur liegen wieder mehrere Cluster vor, wobei aus jedem Cluster wiederum ein neues Clusterobjekt erstellt wird. Diese zweite Aggregationsprozedur wird wiederholt, bis sich das Resultat nicht mehr verändert.

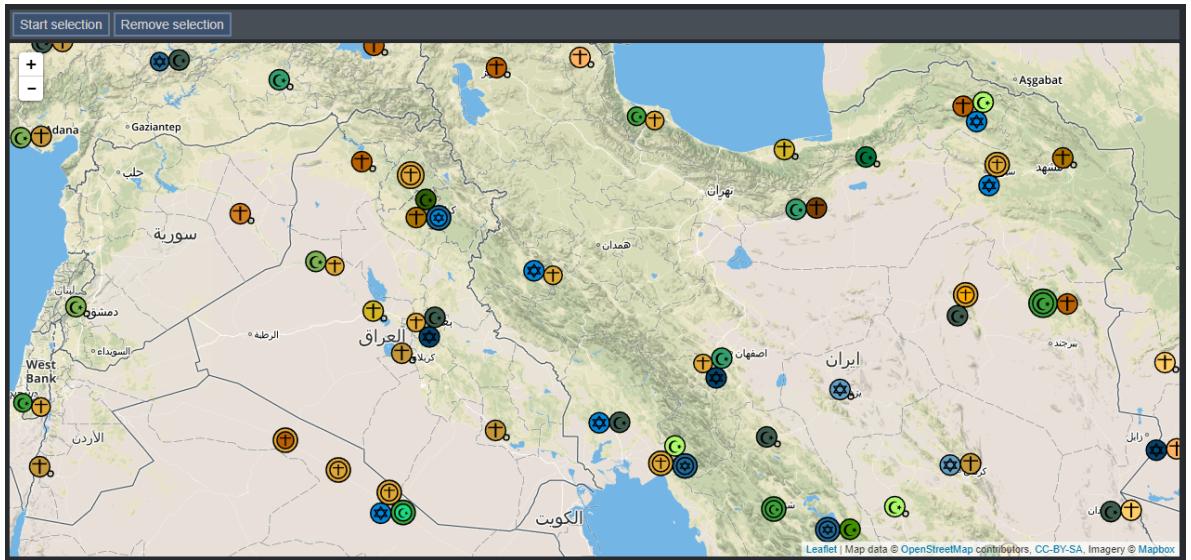
## 5.3 Ansichten

### 5.3.1 Kartenansicht

Die Kartenansicht, welche in Abbildung 5.2 dargestellt ist, dient der Visualisierung der geographischen Aspekte der Datenobjekte des ersten Datensatzes und stellt eine Funktion für die Selektion einer geographischen Region zur Verfügung. Jedes dieser Datenobjekte besitzt eine Koordinaten-Eigenschaft, welche die geographischen Koordinaten der entsprechenden Stadt oder des entsprechenden Ortes spezifiziert. Diese Spezifikation umfasst den Längengrad (Longitude) und den Breitengrad (Latitude). Die Visualisierung innerhalb dieser Ansicht stellt nun eine interaktive geographische Karte zur Verfügung, auf der die Datenobjekte geographisch und visuell repräsentiert werden. Der Aspekt, welcher in dieser Arbeit von besonderem Interesse ist, ist die Darstellung der Ansiedlung oder Aktivität verschiedener religiöser Gruppen im Raum des mittleren und nahen Ostens. Dementsprechend zielt diese Visualisierung darauf ab, die verschiedenen Gruppen zu visualisieren und diese Repräsentationen geographisch korrekt oder sinnvoll zu positionieren. Die Gruppen werden dabei durch graphische Baumknoten repräsentiert. Je nach Zoomstufe der Karte und je nach Bedarf werden die Baumknoten so aggregiert, dass unerwünschte, visuelle Überlappungen vermieden werden. Falls das Resultat der angewandten Aggregationsprozedur noch immer Knoten enthält, durch welche bei entsprechender

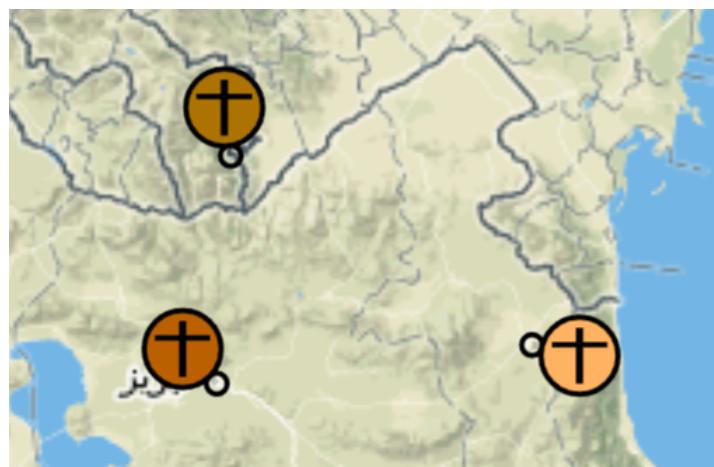
## 5 Umsetzung

---



**Abbildung 5.2:** Visualisierung der Aktivitäten verschiedener religiöser Gruppen innerhalb der Kartenansicht

Positionierung Kollisionen entstehen würden, sorgt ein kräftebasiertes Layout dafür, dass dies nicht geschieht.



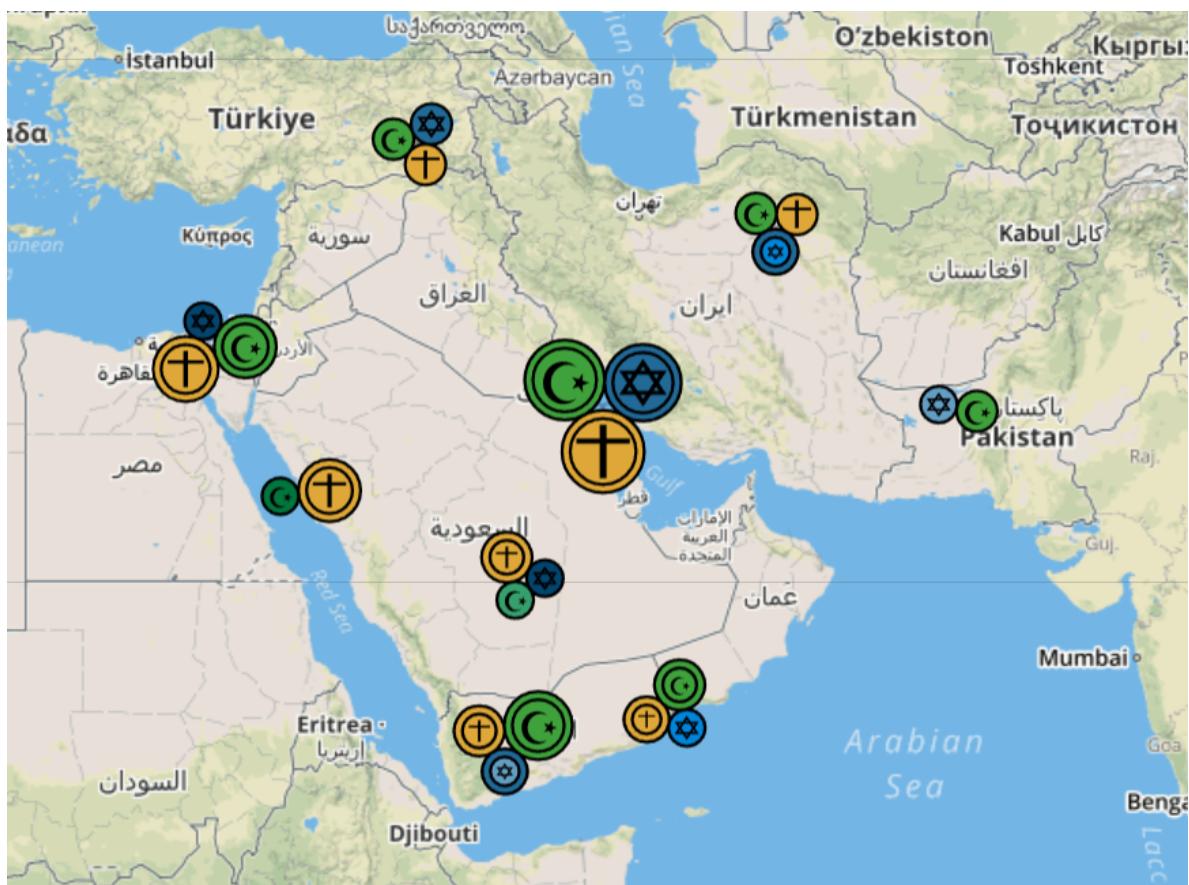
**Abbildung 5.3:** Graphische Baumstrukturen für die Repräsentation einzelner, nicht aggregierter Datenobjekte

Ein Baum, welcher ein nicht aggregiertes Datenobjekt repräsentiert, besitzt folgende Gestalt, welche auch in 5.3 anhand von drei Beispielen dargestellt ist:

- Ein kleiner Knoten (Wurzel) repräsentiert die Stadt, welche durch das entsprechende Datenobjekt und die darin spezifizierten Koordinaten referenziert ist. Dieser Knoten wird dargestellt

durch einen kleinen, schwarzen Kreis mit transparentem Inhalt. Dieser Kreis wird geographisch korrekt auf der Karte platziert.

- Ein größerer Kreis visualisiert die religiöse Gruppe, welche im entsprechenden, einzelnen Datenobjekt referenziert ist. Dieser Kreis ist entsprechend der referenzierten Gruppe eingefärbt und enthält ein Symbol, welches die Religion symbolisiert, der diese Gruppe angehört. Durch das kräftebasierte Layout orientiert sich dieser Knoten zu dem Knoten hin, welcher die entsprechend zugehörige Stadt repräsentiert. Somit hängen beide Knoten ohne Distanz direkt aneinander, wodurch die Zusammenghörigkeit deutlich wird.



**Abbildung 5.4:** Graphische Baumstrukturen für die Repräsentation aggregierter Datenobjekte

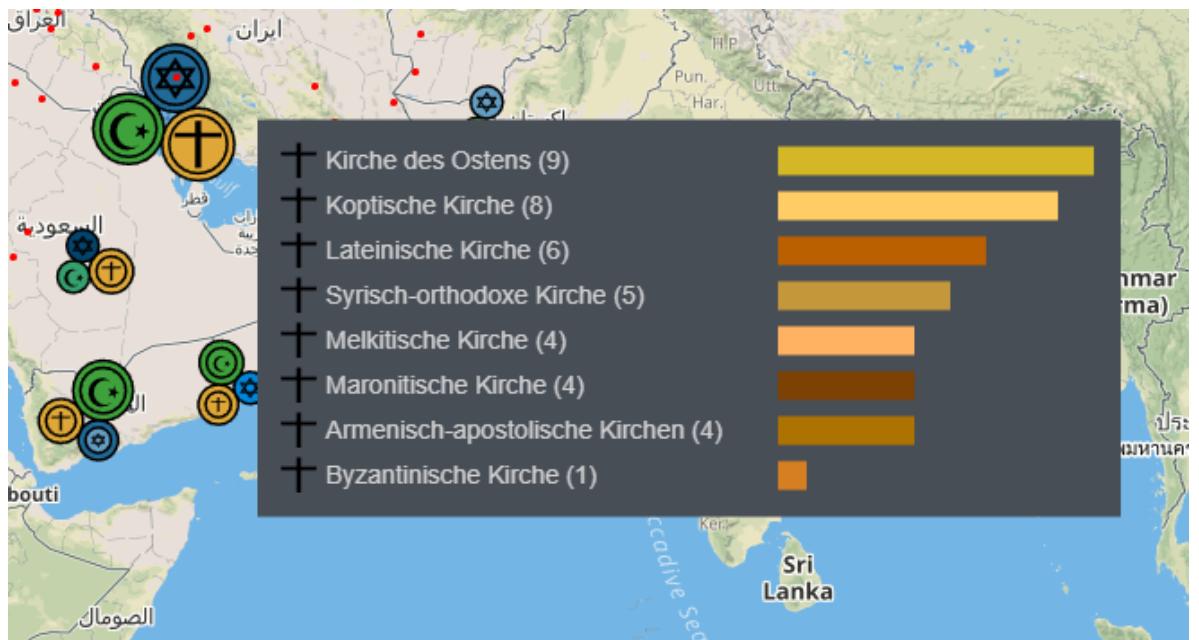
Die Darstellung von Bäumen, welche aggregierte Datenobjekte repräsentieren, ist in Abbildung 5.4 dargestellt und nachfolgend beschrieben:

- Aus den spezifizierten Koordinaten der einzelnen, aggregierten Datenobjekten wird der geometrische Mittelpunkt errechnet. Die Wurzel des Baumes wird an den entsprechenden Koordinaten dieses Mittelpunktes positioniert. Im Gegensatz zum Wurzelknoten eines Baumes, welcher nur ein einzelnes Datenobjekt visualisiert, ist der Knoten hier nicht sichtbar, da es sich dabei nicht um eine konkrete, real existierende Stadt handelt. Dieser Knoten dient lediglich der Ausrichtung und Positionierung der Knoten, welche die religiösen Gruppen repräsentieren.

## 5 Umsetzung

---

- Für jede übergeordnete Religion (hier: Christentum, Islam, Judentum), welche in den aggregierten Datenobjekten vorkommt, existiert ein von der Wurzel abgehender Ast. Sind beispielsweise mehrere Datenobjekte aggregiert, welche lediglich Gruppen einer einzigen übergeordneten Religion referenzieren, so geht vom entsprechenden Wurzelknoten nur ein Ast weg. Wurden jedoch Datenobjekte aggregiert, welche Gruppen von zwei oder drei übergeordneten Religionen referenzieren, so enthält der resultierende Baum entsprechend zwei oder drei Äste, welche von der Wurzel abgehen.
- Ein solcher Ast besitzt nun folgende Gestalt: An der Kante, welche vom Wurzelknoten abgeht, hängt am anderen Ende ein Knoten, welcher die übergeordnete Religion repräsentiert. Dieser Knoten ist entsprechend der Gruppe gefärbt, welche in der Hierarchie einen Vorgänger aller repräsentierten Gruppen darstellt, und beinhaltet das entsprechende Religionssymbol. Die Größe dieses Knotens ist abhängig von der Anzahl der repräsentierten, aggregierten Datenobjekte. Sofern ein solcher Knoten nur ein Datenobjekt repräsentiert, so besteht dieser Knoten lediglich aus einem entsprechend der Gruppe eingefärbtem Kreis, welcher das entsprechende Religionssymbol enthält. Repräsentiert ein solcher Knoten allerdings mehrere, aggregierte Datenobjekte, so besteht der Knoten aus einem Donutdiagramm, einem Kuchendiagramm darin und darin wiederum ist das entsprechende Religionssymbol enthalten. Das Kuchendiagramm und das Donutdiagramm ist jeweils für die Hervorhebung der Selektion von Bedeutung, was weiter unten erläutert wird.

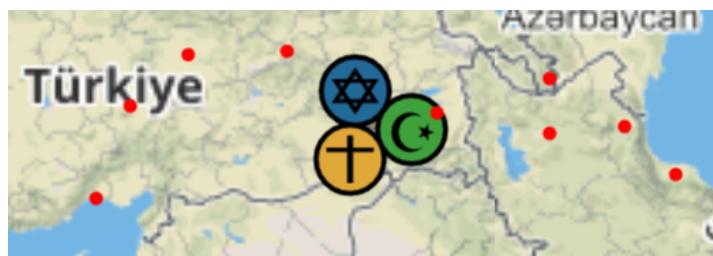


**Abbildung 5.5:** Ein Fenster veranschaulicht die Zusammensetzung eines Baumknotens, welcher mehrere, aggregierte Datenobjekte repräsentiert.

Durch die Aggregation von mehreren Datenobjekten zu einem Knoten gehen zunächst Informationen verloren, die nicht explizit visualisiert werden. Das ist allerdings notwendig, soll eine große Menge an Informationen auf engem Raum dargestellt werden. Um diesen Informationsverlust zu kompensieren,

können die zunächst nicht dargestellten Informationen durch Interaktion abgefragt werden. Bewegt der Anwender den Mauszeiger über einen Knoten, welcher mehrere aggregierte Datenobjekte repräsentiert, die wiederum verschiedene Gruppen referenzieren, so erscheint in der Nähe des Knotens ein Fenster, welches die Zusammensetzung des Knotens detailliert darstellt. Dieses Fenster ist anhand eines Beispiels in Abbildung 5.5 dargestellt. Diese Detailansicht enthält eine Liste, bestehend aus jeweils einem Listenelement für jede religiöse Gruppe, welche durch mindestens eines der aggregierten Datenobjekte referenziert wird. Jedes Listenelement besteht aus folgenden Bestandteilen:

- Symbol: Zunächst wird das Symbol der Gruppe beziehungsweise der übergeordneten Religion dargestellt.
- Bezeichner: Auf das Symbol folgt horizontal der Bezeichner der entsprechenden religiösen Gruppe. Zu diesem Bezeichner gehört ebenfalls in Klammern eine Zahl, welche angibt, wie viele Datenobjekte hierbei aggregiert wurden, welche diese Gruppe referenzieren.
- Balken: Ein Balken visualisiert durch seine horizontale Ausdehnung beziehungsweise durch seine Breite die oben erwähnte Anzahl der entsprechenden, aggregierten Objekte. Der Balken ist hier in der Farbe der religiösen Gruppe eingefärbt.



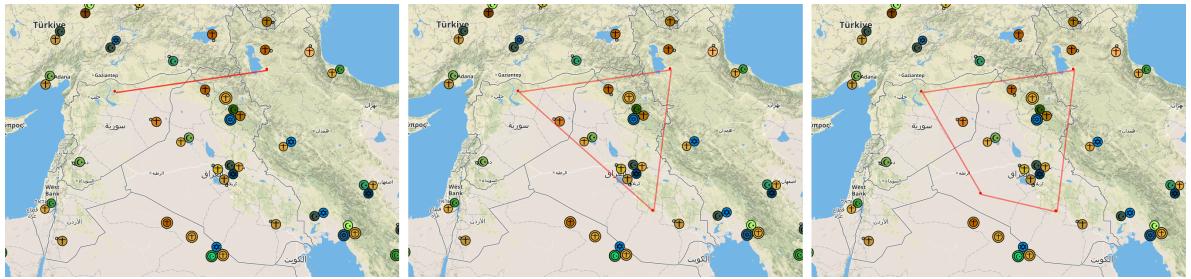
**Abbildung 5.6:** Rote Kreise visualisieren die Koordinaten der Städte, welche von den aggregierten Datenobjekten referenziert werden.

Werden mehrere Datenobjekte aggregiert, so gehen allerdings auch geographische Informationen verloren, da aus den Koordinaten dieser Objekte der Mittelpunkt errechnet und der resultierende Baum beziehungsweise die entsprechende Wurzel des Baumes dort positioniert wird. Zunächst ist also nicht mehr ersichtlich, welche Städte oder Orte hier eigentlich referenziert werden. Auch diese Informationen werden visualisiert, wenn der Anwender den Mauszeiger über einen Baumknoten bewegt. Dabei werden die Städte und Orte jeweils durch einen kleinen, roten Punkt dargestellt, welcher geographisch korrekt positioniert ist. Dies hilft dabei, zu verstehen, welcher geographische Bereich durch einen Baum abgedeckt wird. Ein Beispiel hierfür ist in Abbildung 5.6 dargestellt.

Mittels eines Buttons oberhalb der Karte kann der Selektionsmodus aktiviert werden. Ist dieser aktiv, so kann der Anwender durch punktuelles Klicken innerhalb der Karte sequentiell mehrere Punkte setzen. Jeder gesetzte Punkt wird automatisch mit dem vorherigen Punkt und dem ersten gesetzten Punkt verbunden, wobei automatisch die letzte Verbindungsline zwischen dem vorherigen Punkt und dem ersten gesetzten Punkt entfernt wird. Somit ist gewährleistet, dass die gesetzten Punkte und Linien immer einen geschlossenen Polygonzug ergeben. Dieser Sachverhalt ist in 5.7 dargestellt. Diese geographische Selektion kann nun entweder mittels des Buttons, durch welchen die Selektion gestartet wurde, gespeichert, oder mittels eines zweiten Buttons wieder entfernt und abgebrochen werden.

## 5 Umsetzung

---



**Abbildung 5.7:** Jedes Mal, wenn ein weiterer Punkt zu dem Polygonzug hinzugefügt wird, wird dieser neue Punkt automatisch mit dem ersten und dem letzten hinzugefügten Punkt verbunden, wodurch der Polygonzug zu jeder Zeit geschlossen ist.

Wird die Selektion gespeichert, so werden alle Datenobjekte, welche in diesem Polygon liegende Koordinaten referenzieren, selektiert und alle anderen Datenobjekte deseletktiert. Ein Beispiel für eine solche geographische Selektion ist in Abbildung 5.8 dargestellt. Wurde die Selektion gespeichert, so kann diese mittels des zweiten Buttons wieder entfernt werden.

Werden nun Datenobjekte durch die Interaktion mit einer Ansicht selektiert beziehungsweise gefiltert, so verändert sich die Gestalt der Baumknoten. Dabei muss zwischen Knoten, welche lediglich ein solches Datenobjekt repräsentieren, und Knoten, welche mehrere, aggregierte Datenobjekte repräsentieren, unterschieden werden:

- Ist das zugrundeliegende Datenobjekt selektiert, so wird der entsprechende Knoten, welcher dieses repräsentiert, dargestellt wie oben beschrieben. Wird das Objekt nun deseletktiert, so wird der farbige Kreis transparent und ausgegraut dargestellt. Diese Knoten kennen dementsprechend nur zwei Zustände. Beispiele hierfür sind in Abbildung 5.9 dargestellt.
- Sind alle Datenobjekte, welche durch einen Knoten repräsentiert werden, selektiert, so besitzt dieser die oben beschriebene Gestalt. Werden nun aber Datenobjekte deseletktiert, so verkleinern sich entsprechend jeweils das Kuchenstück und das Donutstück, welche prozentual den Anteil der selektierten Datenobjekte darstellen. Das Kuchenstück und das Donutstück, welche jeweils den nicht selektierten Anteil visualisiert, ist exakt so gestaltet, wie die nicht selektierten Knoten, welche nur ein Datenobjekt repräsentieren: Das Kuchenstück ist transparent und ausgegraut. In Abbildung 5.10 ist ein Beispiel hierfür dargestellt.

Für die Darstellung der prozentualen Anteile wurden einige Ideen gegeneinander abgewogen und die meiner Meinung nach sinnvollste und eindeutigste Variante ausgewählt. Alternativ bestand die Idee, das Symbol im Knoten vertikal abzuschneiden gemäß des prozentualen Anteils der selektierten Objekte. Wären also beispielsweise 50 Prozent der Objekte selektiert, so wäre das Bild vertikal bei der Hälfte abgeschnitten. Die Obere Hälfte wäre also nicht mehr sichtbar oder transparent. Die Interpretation dieser Darstellung ist allerdings sehr problematisch. Es ist nicht eindeutig und intuitiv, ob die Höhe, bei der das Bild abgeschnitten wird, oder der Flächeninhalt des verbleibenden Teiles des Symbols den prozentualen selektierten Anteil darstellt. Hinzukommend kann dieser Ansatz auch nur bei Symbolen einigermaßen funktionieren, welche vertikal mehr oder weniger symmetrisch sind.

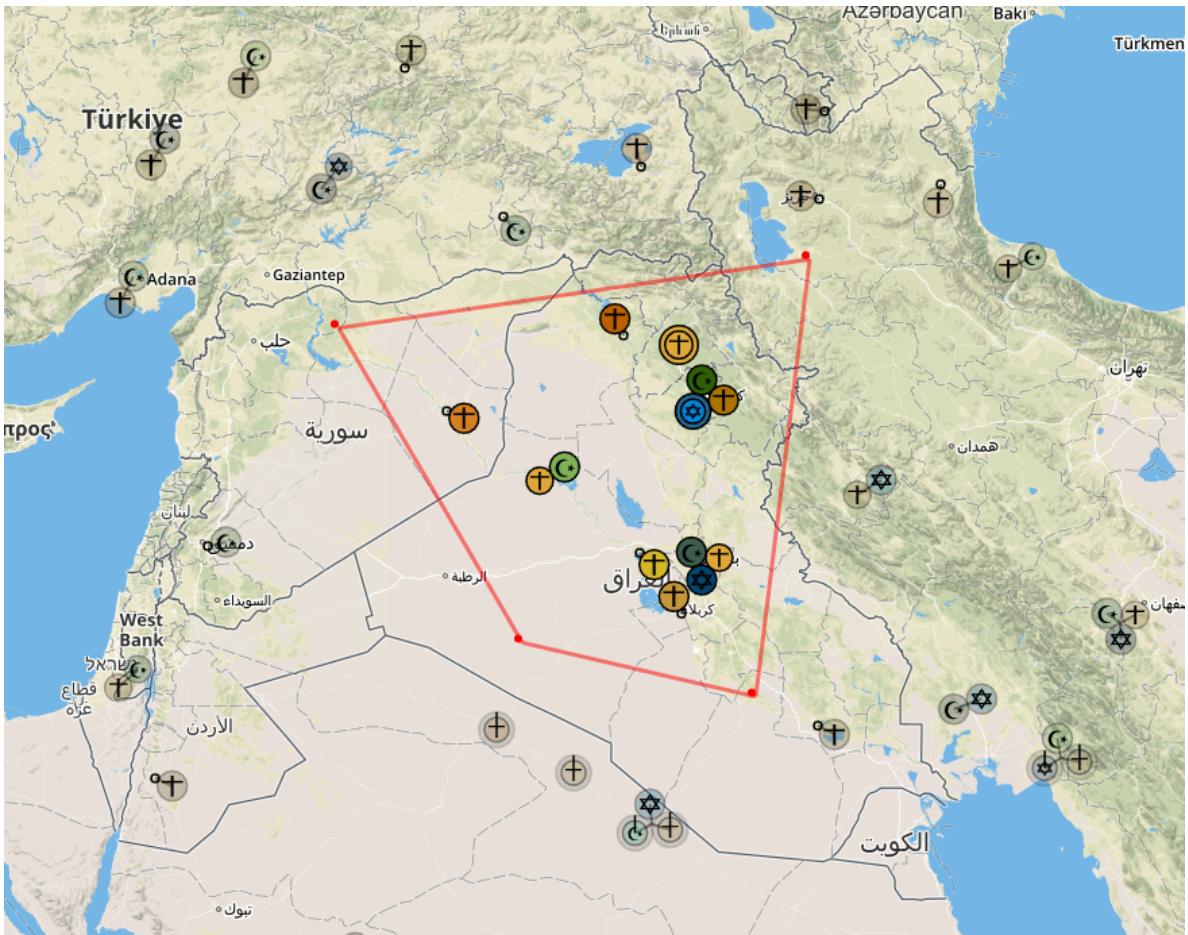


Abbildung 5.8: Ergebnis der geographischen Selektion innerhalb der Kartenansicht

Die selbe Idee kann auf einen farbigen Kreis im Hintergrund des Symbols übertragen werden. Aber auch hierbei besteht das Problem, dass nicht intuitiv und klar ist, wie der verbleibende Rest des Kreises zu interpretieren ist: Höhe des Schnittes oder Flächeninhalt?

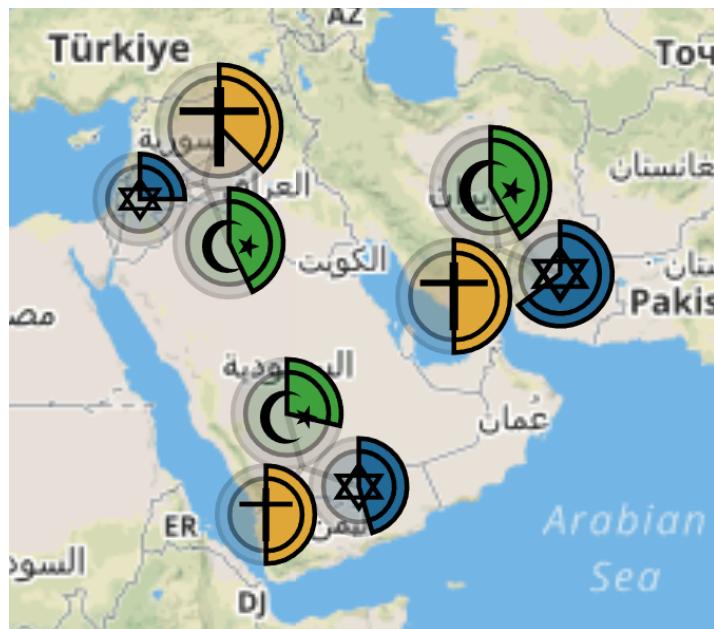
Wie auch innerhalb der Quellenansicht und innerhalb der Gruppenansicht könnte jeder Knoten zusätzlich mit einem Prozentbalken versehen werden, welcher die prozentualen Anteile gut interpretierbar darstellt. Allerdings ergäben sich hieraus mehrere Probleme. Die kollisionsfreie Positionierung der Knoten, welche nun jeweils aus einem Kreis und einem Rechteck bestehen würden, wäre aufwendiger und komplexer in der Umsetzung und in der Berechnung. Da die aktuelle Implementierung Kollisionen zwischen Knoten auf Basis der Radien dieser Knoten ermittelt, ist eine Anpassung nicht ohne Weiteres umsetzbar. Bezüglich der Positionierung der Knoten, also der Kreise und der Rechtecke, müsste außerdem gewährleistet werden, dass die Zugehörigkeit der Rechtecke zu den Kreisen jederzeit eindeutig erkennbar ist. Außerdem würde eine Hinzunahme von Rechtecken die Visualisierung innerhalb der Kartenansicht wesentlich komplexer und unharmonischer machen.

## 5 Umsetzung

---



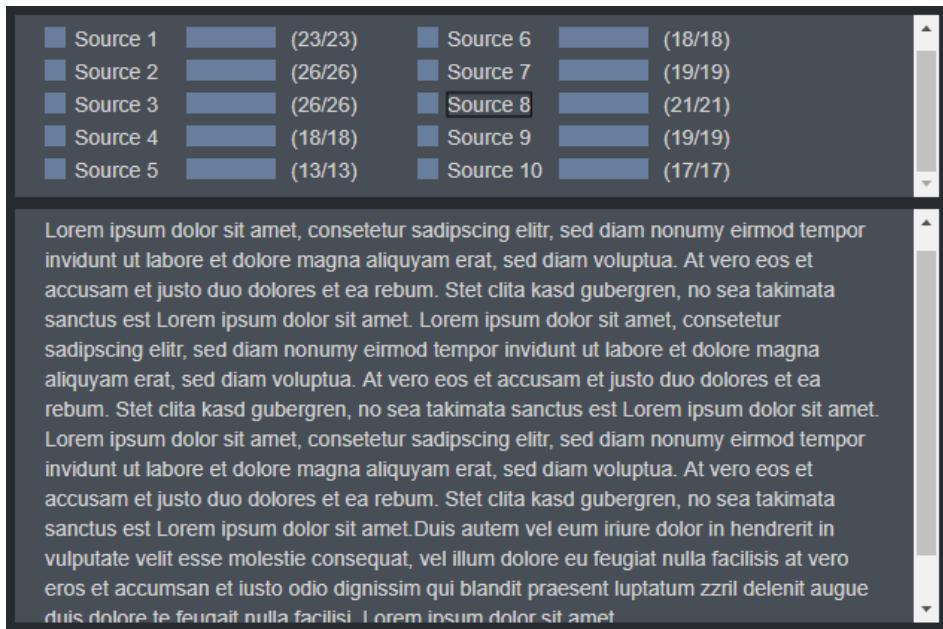
**Abbildung 5.9:** Knoten, welche einzelne Datenobjekte repräsentieren, im selektierten (links) und nicht selektierten (rechts) Zustand



**Abbildung 5.10:** Kuchen-und Donutstücke visualisieren selektierte und nicht selektierte Anteile von Knoten, welche mehrere, aggregierte Datenobjekte repräsentieren.

Eine weitere Idee wäre, den Anteil der selektierten Objekte durch die Opazität der Kreise im Hintergrund oder der Symbole selber darzustellen. Allerdings wäre es für den Anwender hierbei unmöglich, den selektierten Anteil grob abzuschätzen.

Alternativ zum kräftebasierten Layout könnten regelbasierte Ansätze benutzt werden, um die Knoten auf der Karte zu positionieren.



**Abbildung 5.11:** Quellenansicht, bestehend aus einer Liste aller Quellen und einem Abschnitt für die Anzeige des Textes einer ausgewählten Quelle

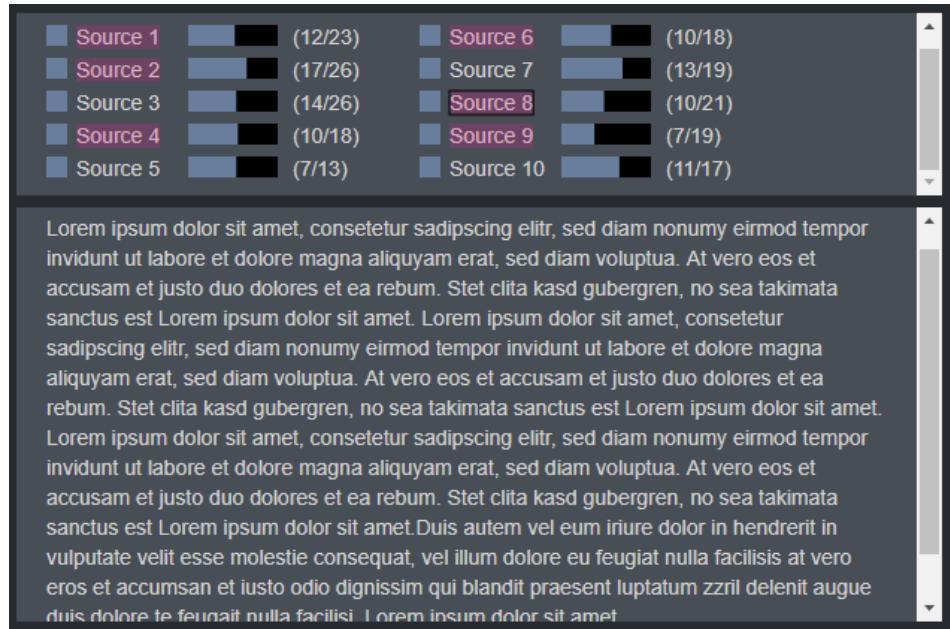
### 5.3.2 Quellenansicht

Die Quellenansicht besteht aus zwei voneinander abgetrennten Abschnitten: Der obere Abschnitt besteht aus Elementen für die Interaktion mit den einzelnen Quellen, während der untere Abschnitt lediglich der Anzeige des jeweilig ausgewählten Quellentextes dient. Die Quellenansicht ist in Abbildung 5.11 dargestellt. Der obere Abschnitt besteht konkret aus einer Liste von Elementen und dient der Visualisierung und der Selektion der verschiedenen Quellen. Jedes dieser Listenelemente repräsentiert eine Quelle, welche im ersten Datensatz referenziert wird. Ein Listenelement besteht aus folgenden Bestandteilen:

- Checkbox: Diese Checkbox dient der Selektion und Deselektion der entsprechenden Quelle. Wird eine Quelle selektiert, werden alle Datenobjekte aus dem ersten Datensatz in die Selektion aufgenommen, welche diese Quelle referenzieren. Die Deselektion einer Quelle nimmt alle entsprechenden Objekte aus der Selektion heraus.
- Bezeichner: Hier wird als Textelement der Bezeichner der entsprechenden Quelle angezeigt. Wird dieses Textelement angeklickt, erscheint der Textinhalt dieser Quelle im unteren Abschnitt der Quellenansicht. Wird ein Knoten innerhalb der Kartenansicht ausgewählt, welcher Datenobjekte des ersten Datensatzes repräsentiert, die unter anderem diese Quelle referenzieren, so wird der Quellenbezeichner durch ein pinkes Rechteck hervorgehoben. Diese Hervorhebung ist beispielhaft in Abbildung 5.12 dargestellt.
- Prozentbalken: Dieses Element visualisiert prozentual den selektierten und den nicht selektierten Anteil der Objekte aus dem ersten Datensatz, welche diese Quelle referenzieren. Das

## 5 Umsetzung

---



**Abbildung 5.12:** Für jede Quelle visualisiert ein Prozentbalken den selektierten Anteil. Wird ein Knoten innerhalb der Kartenansicht ausgewählt, so werden alle Quellen, welche in den durch den Knoten repräsentierten Datenobjekte referenziert werden, mittels pinker Rechtecke hervorgehoben.

Segment, welches den selektierten Anteil repräsentiert, ist farbig und das andere Segment für den nicht selektierten Anteil schwarz ausgefüllt. Die Aufteilung zwischen beiden Segmenten wird automatisch bei jeder Änderung einer Selektion in einer der vier Ansichten, also auch innerhalb der Quellenansicht, aktualisiert. Die Visualisierung der selektierten Anteile durch die Prozentbalken ist beispielhaft in Abbildung 5.12 dargestellt.

- **Zahlen:** Das letzte Element ist wieder ein Textelement, welches die Anzahl der selektierten und die Anzahl aller Objekte darstellt, welche die entsprechende Quelle referenzieren. Wie auch die Segmentierung des Prozentbalkens wird die erste Zahl bei jeder Änderung einer Selektion in einer der vier Ansichten aktualisiert.

Da für alle Ansichten in der linken Spalte des Layouts eine einheitliche Breite festgelegt ist, und die Gruppenansicht wesentlich mehr horizontalen Platz einnimmt, als für die Darstellung eines Listenelements benötigt wird, wurde die Quellenliste in zwei Spalten aufgeteilt. Dadurch wird die Liste kürzer und mehr Informationen können gleichzeitig angezeigt werden. Wird allerdings doch mehr vertikaler Platz durch eine höhere Anzahl an Quellen benötigt, so kann die Liste vertikal gescrollt werden.

Der untere Abschnitt der Quellenansicht dient zur Anzeige der Quellentexte. Diese Texte können gescrollt werden, falls sie mehr Platz benötigen, als die Ansicht zur Verfügung stellt.

Für die prozentuale Anzeige für jedes Listenelement wurde ein Prozentbalken gewählt, da dieser sehr einfach und korrekt interpretiert werden kann. Außerdem eignet sich ein Element mit horizontaler Ausdehnung gut für die Positionierung innerhalb horizontal ausgerichteter Listenelemente.

Der visuelle Eindruck durch ein Kuchendiagramm mit zwei Kuchenstücken (für den selektierten und den nicht selektierten Anteil) wäre weniger harmonisch und darüber hinaus auch schwieriger zu interpretieren als ein Prozentbalken.

Alternativ könnte man den oberen Teil der Ansicht auch komplett durch eine einzige Grafik ersetzen, beispielsweise durch ein Kuchendiagramm. Dieses Diagramm müsste dann aber jeweils zwei Kuchenstücke für jede Quelle darstellen. Damit gingen mehrere Probleme einher. Für jede Quelle müssten zwei Farben gewählt werden, die das Kuchenstück für den selektierten und das Stück für den nicht selektierten Anteil klar unterscheidbar machen. Bei einer hohen Anzahl an Quellen würde die Farbgestaltung hierbei schwierig werden. Außerdem wäre der visuelle Eindruck durch die Verwendung vieler, verschiedener Farben auf engem Raum eher unharmonisch und unangenehm. Diese Probleme ergeben sich für viele Ansätze, welche das Ziel verfolgen, alle Quellen, und damit auch beide Anteile für jede Quelle, in einer Grafik zu visualisieren. Diese Problematik hat ebenfalls das Design der Zeitleiste beeinflusst. Außerdem, wie bereits in Abschnitt 2.1 erwähnt, sind derartige Kuchendiagramme weitaus schwieriger zu interpretieren als einfache Prozentbalken.

#### 5.3.3 Gruppenansicht

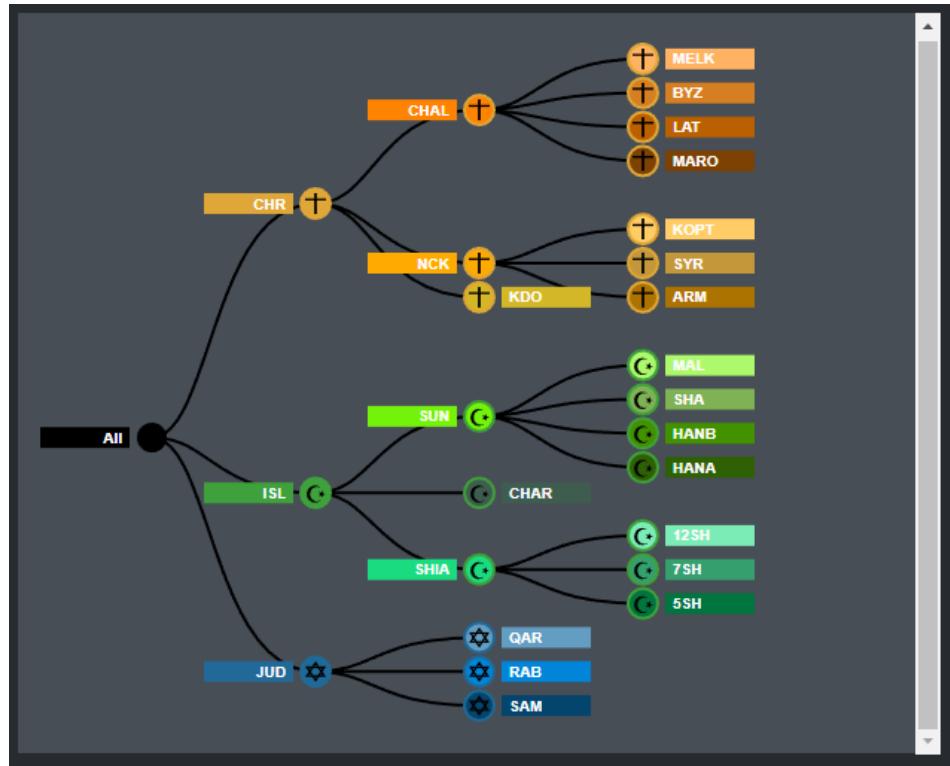
Wie bereits weiter oben angedeutet bestehen hierarchische Beziehungen zwischen den einzelnen religiösen Gruppen. Diese hierarchische Struktur wird aus dem zweiten Datensatz eingelesen. Dabei ist der Ansatz im aktuellen Zustand darauf ausgelegt, dass es sich um eine Struktur mit drei Objekten auf der obersten Hierarchieebene handelt. Diese drei Objekte repräsentieren die verschiedenen Religionen, welche in dieser Arbeit von Interesse sind: Das Christentum, der Islam und das Judentum. Alle anderen Gruppen sind in der Hierarchie unter diesen Religionen einzuordnen.

Da die Struktur der Religionen und Gruppen hier hierarchisch vorliegt, wurde eine Baumstruktur gewählt, um diesen Sachverhalt geeignet zu visualisieren. Diese ist in Abbildung 5.13 dargestellt. Die Erstellung des graphischen Baumes wird durch eine Prozedur realisiert, welche den zweiten Datensatz ausliest und für jedes Objekt daraus einen Baumknoten erstellt, welcher durch Kanten mit den entsprechenden Eltern-und Kindknoten verbunden ist. Die Wurzel des Baumes, welche die Gesamtheit aller Religionen beziehungsweise Gruppen repräsentiert, besitzt drei Kindknoten; jeder davon repräsentiert eine der oben genannten Religionen.

Die Ausrichtung des Baumes, beziehungsweise die Beziehungen von Elternknoten zu Kindknoten, ist horizontal, von links nach rechts, umgesetzt.

Die Knoten des Baumes bestehen jeweils aus folgenden visuellen Elementen:

- Kreis: Ein Knoten ist grundsätzlich durch einen Kreis repräsentiert, in dem und um welchen herum weitere Elemente positioniert sind. Der Kreis ist hierbei, sofern die Gruppe selektiert ist, gemäß eines Farbschemas eingefärbt, welches den verschiedenen Religionen und den darunter strukturierten Gruppen eine entsprechende Farbe zuweist. Dieses Farbschema wird auch in der



**Abbildung 5.13:** Gruppenansicht

Zeitleiste und in der Kartenansicht verwendet. Ist die entsprechende Gruppe nicht selektiert, so ist der Kreis schwarz gefärbt, wie in Abbildung 5.14 dargestellt.

Wird ein Knoten angeklickt, so wird der daran hängende Teilbaum ein-beziehungsweise ausgeklappt, je nach dem, in welchem Zustand sich dieser Teilbaum befindet. Ist der an einem Knoten hängende Teilbaum eingeklappt, so wird dies visuell durch einen kleinen, farbigen Punkt neben dem Kreis dargestellt, um zu verdeutlichen, dass es sich hierbei nicht um einen Blattknoten handelt.

Der restliche Baum dehnt sich gemäß des neu verfügbaren Platzes aus, wenn ein Teilbaum eingeklappt wird, und zieht sich wieder zusammen, sobald der Teilbaum wieder ausgeklappt wird. Diese Funktion bietet dem Anwender die Möglichkeit, den gegebenenfalls sehr komplexen Baum, für welchen nur eine beschränkte Ansichtsgröße zur Verfügung steht, nach seinen Bedürfnissen und Interessen zu konfigurieren, so das nur die für ihn relevanten Informationen dargestellt und fokussiert werden. Zwei Beispiele hierfür sind in Abbildung 5.15 dargestellt.

- Bild: Der Kreis beinhaltet ein Bild, welches ein Symbol vor transparentem Hintergrund darstellt. Jedem Knoten beziehungsweise jeder Gruppe ist damit das Symbol der entsprechenden, übergeordneten Religion zugeordnet. Denkbar wäre auch, jeder Gruppe ein einzigartiges Symbol zuzuweisen, allerdings liegen diese Informationen aktuell nicht vor. Wenn innerhalb der Kartenansicht ein Knoten ausgewählt wird, werden alle Gruppen, welche in den durch die

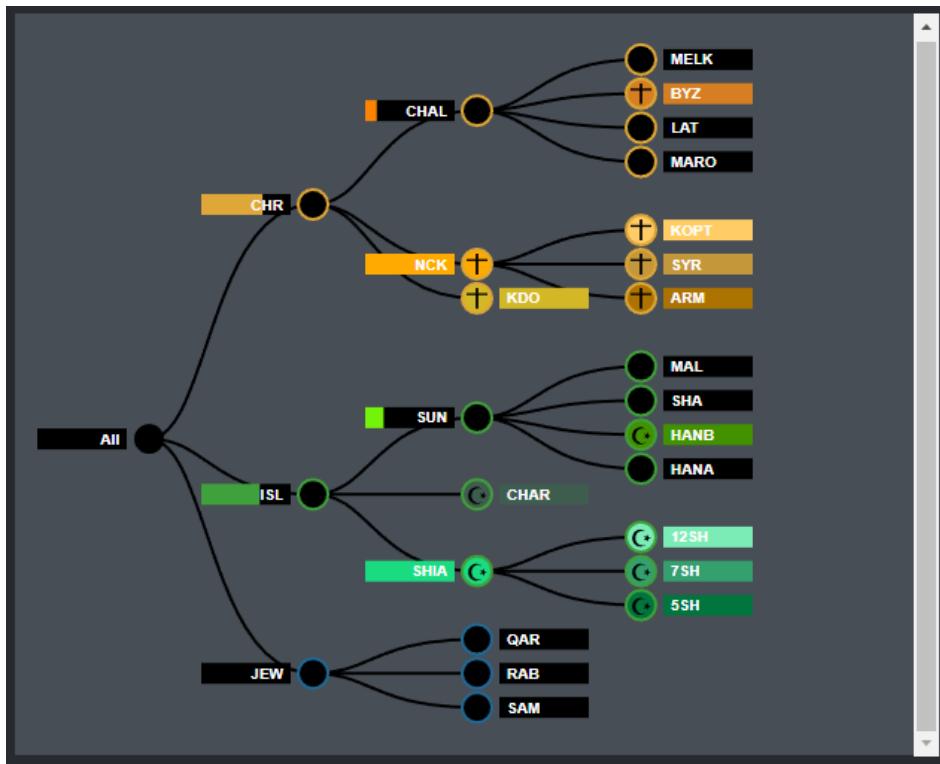
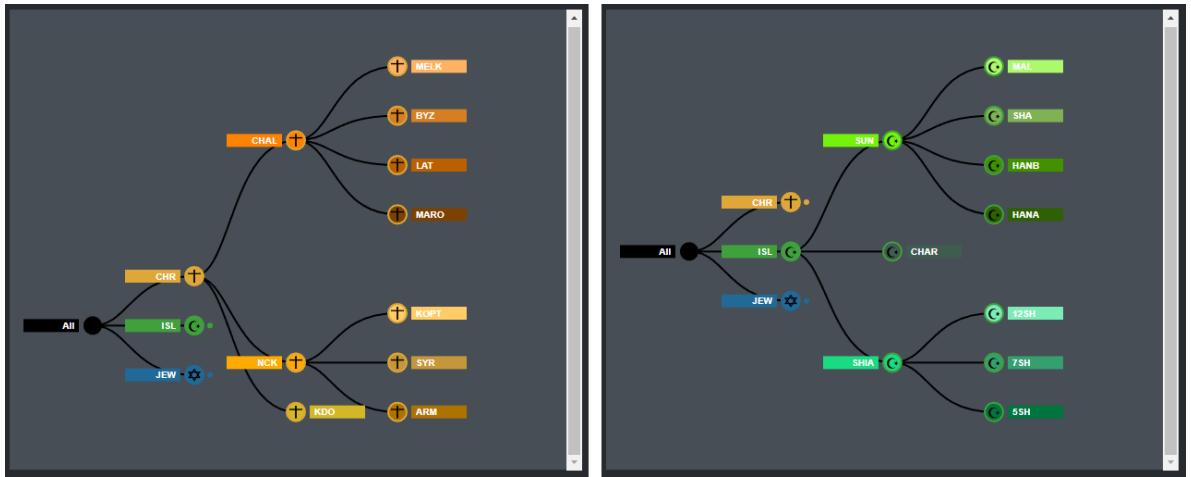


Abbildung 5.14: Der Inhalt deselektierter Knoten ist schwarz.

sen Knoten repräsentierten Datenobjekten referenziert werden, mittels pinker Umrandungen hervorgehoben. Ein Beispiel hierfür ist in Abbildung 5.16 dargestellt.

- Prozentbalken: Jedem Baumknoten ist außerdem ein Prozentbalken angeheftet, welcher prozentual den selektierten und den nicht selektierten Anteil aller Datenobjekte des ersten Datensatzes, welche die entsprechende Religion oder Gruppe referenzieren, visualisiert. Wird eine Änderung der Selektion innerhalb einer der vier Ansichten, also auch innerhalb dieser Ansicht, vorgenommen, so aktualisiert sich die Anzeige jedes Prozentbalkens automatisch gemäß der neuen Filterdaten, wie dies auch innerhalb der Quellenansicht geschieht. Der Prozentbalken ist rechts des Kreises angeordnet, falls es sich um einen Blattknoten handelt, andernfalls links davon. Würden die Prozentbalken generell rechts von den Kreisen platziert werden, so würden diese die Verzweigung der Kanten zu den entsprechenden Kindknoten verdecken. In Abbildung 5.16 wird ein Beispiel für die Visualisierung der selektierten Anteile mittels der Prozentbalken dargestellt.
- Label: Dieses Label ist durch einen einfachen Textknoten umgesetzt und enthält lediglich den abgekürzten Namen der entsprechenden Religion oder Gruppe. Die Platzierung des Labels ist identisch geregelt wie die Platzierung der Prozentbalken, also abhängig davon, ob es sich bei dem entsprechenden Knoten um einen Blattknoten handelt oder nicht. Damit liegen die Labels über den Prozentbalken. Beide Elemente sind jedoch durch Form, Größe und Farbe so gestaltet, dass die Interpretation der einen Darstellung nicht durch die jeweils andere gestört wird. Die

## 5 Umsetzung



**Abbildung 5.15:** Werden Teilbäume eingeklappt, so dehnt sich der restliche Baum gemäß des verfügbaren Platzes aus. Ein eingeklappter Teilbaum wird durch einen kleinen, farbigen Punkt rechts neben dem entsprechenden Knoten angedeutet.

Selektion eines Knotens, und damit auch des gesamten Teilbaumes, welcher an diesem Knoten hängt, wird durch Anklicken des entsprechenden Labels ausgelöst.

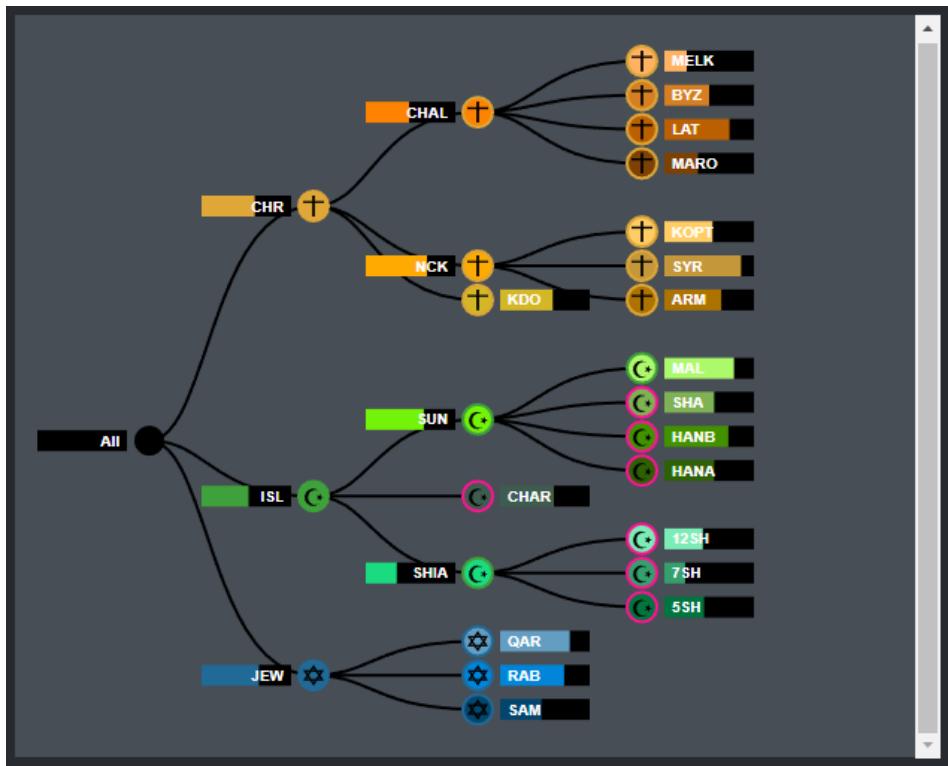
Die Selektionslogik der Baumknoten umfasst folgende zwei zur Verfügung stehende Operationen:

- Verfeinerung: Wird das Label eines Knotens, welcher Bestandteil eines bereits selektierten Teilbaumes ist, angeklickt, so wird die Selektion dahingehend verfeinert, dass der entsprechende, übergeordnete selektierte Teilbaum deselektiert, und lediglich der Teilbaum, welcher an dem ausgewählten Knoten hängt, selektiert wird.
- Generalisierung: Wird das Label eines Knotens angeklickt, welcher innerhalb des an diesem Knoten hängenden Teilbaumes einen selektierten Teilbaum enthält, so wird nun der gesamte an diesem Knoten hängende Teilbaum selektiert.

Diese Selektionslogik ist in den Abbildungen 5.17 und 5.18 dargestellt.

Für die Visualisierung der prozentualen Anteile wurden Prozentbalken gewählt, da diese einfach und gut interpretierbar sind. Alternativ könnten die einfachen Kreise, welche jeweils das entsprechende Religionssymbol beinhalten, durch Kuchendiagramme ausgetauscht werden, wie dies auch in der Visualisierung der Kartenansicht getan wurde. Allerdings könnte die Kombination von komplexen Grafiken im Hintergrund und Symbolen im Vordergrund unerwünschte Eindrücke und Verwirrungen hervorrufen, welche durch die geringe Größe der Kreise noch verstärkt werden würden. Die geringe Größe der Knoten ist durch den Mangel an Platz, welcher zwangsläufig in MCV-Ansätzen zu Kompromissen führen muss, begründet.

Eine Möglichkeit um den durch den Baum benötigten Platz zu verringern besteht darin, lediglich jedem Blattknoten einen Prozentbalken zuzuordnen. Diese Designentscheidung würde allerdings



**Abbildung 5.16:** Für jede Gruppe visualisiert ein Prozentbalken den selektierten Anteil. Wird ein Knoten innerhalb der Kartenansicht ausgewählt, so werden die Gruppen, welche in den durch diesen Knoten repräsentierten Datenobjekten referenziert werden, durch pinke Umrandungen hervorgehoben.

die vorliegende hierarchische Struktur der Religionen und Gruppen ignorieren und dem Anwender Informationen vorenthalten, die dieser kognitiv selber aus den Informationen der Blattknoten zusammensetzen müsste.

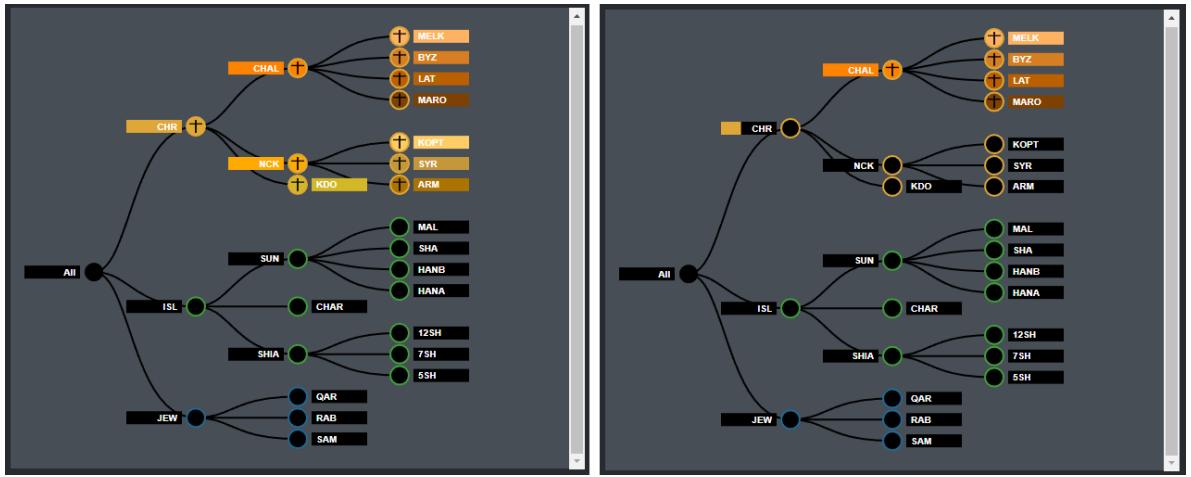
Ein weiteres Argument gegen die Verwendung solcher Kuchendiagramme in diesem Kontext ist die schwierigere Interpretation gegenüber einfacher Prozentbalken, wie bereits in Abschnitt 2.1 erwähnt.

Für den Baum wurde eine horizontale Ausrichtung gewählt, weil jeder Knoten des Baumes jeweils einen Prozentbalken und ein Label besitzt, welche horizontal ausgerichtet werden sollten, um die Lesbarkeit und die Interpretierbarkeit zu unterstützen. Würde der Baum vertikal und die Prozentbalken und Labels dagegen horizontal ausgerichtet werden, so würde ein weitaus höherer horizontaler Platzbedarf daraus resultieren.

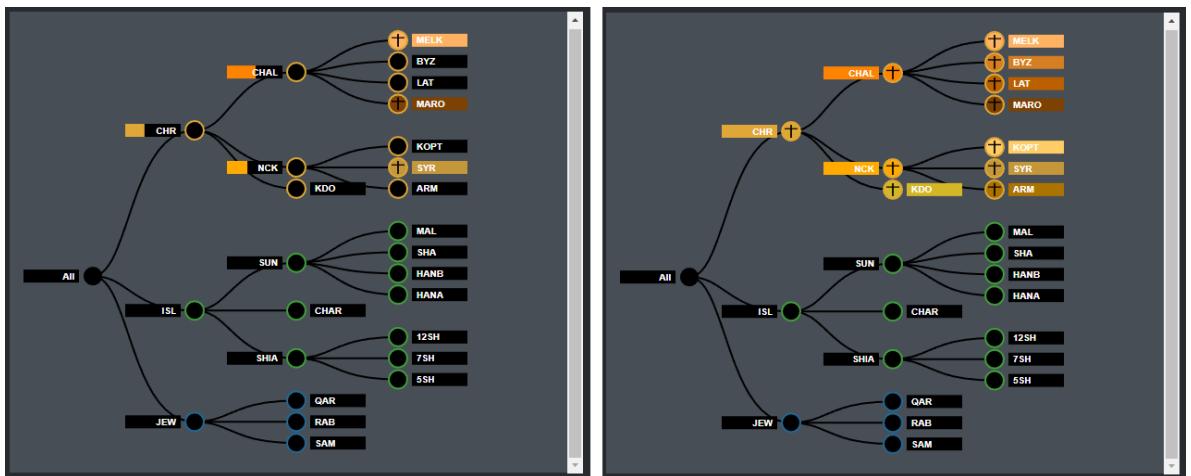
#### 5.3.4 Zeitleistenansicht

Jedes Datenobjekt aus dem ersten Datensatz beschreibt eine Zeitspanne, innerhalb welcher die entsprechende religiöse Gruppe an dem entsprechenden Ort angesiedelt war. Diese Zeitspanne ist

## 5 Umsetzung



**Abbildung 5.17:** Wird innerhalb eines selektierten Teilbaumes ein Teilbaum ausgewählt, so wird die Selektion verfeinert. Links ist der an dem Knoten 'CHR' hängende Teilbaum selektiert. Durch die Auswahl des Knotens 'CHAL' wird die Selektion dahingehend verfeinert, dass nicht mehr der gesamte an 'CHR' hängende Teilbaum selektiert ist, sondern nur noch der an 'CHAL' hängende Teilbaum.



**Abbildung 5.18:** Wird ein Teilbaum ausgewählt, welcher bereits teilweise selektiert ist, so wird die Selektion auf den gesamten Teilbaum ausgeweitet. Links sind nur die Knoten 'MELK', 'MARO' und 'SYR' selektiert. Wird nun der Knoten 'CHR' ausgewählt, wird der gesamte daran hängende Teilbaum selektiert (rechts).

durch zwei Jahreszahlen spezifiziert. Für die Visualisierung dieser Informationen und letztendlich auch für die Selektion einer Zeitspanne durch den Benutzer wurde ein gestapeltes Säulendiagramm implementiert, welches in Abbildung 5.19 dargestellt ist. Nachfolgend werden die Dimensionen dieses Diagramms kurz erklärt:

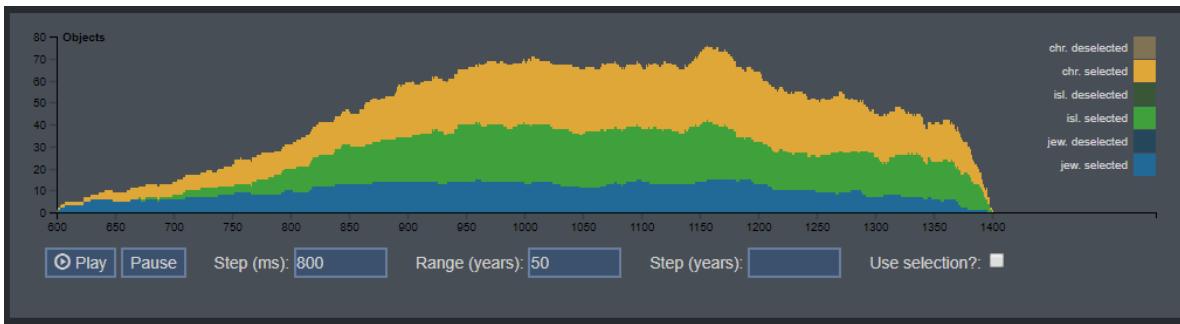


Abbildung 5.19: Zeitleistenansicht

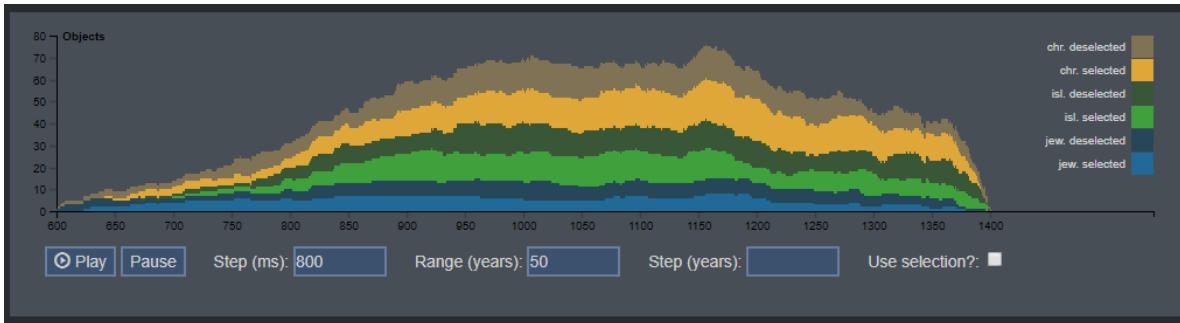


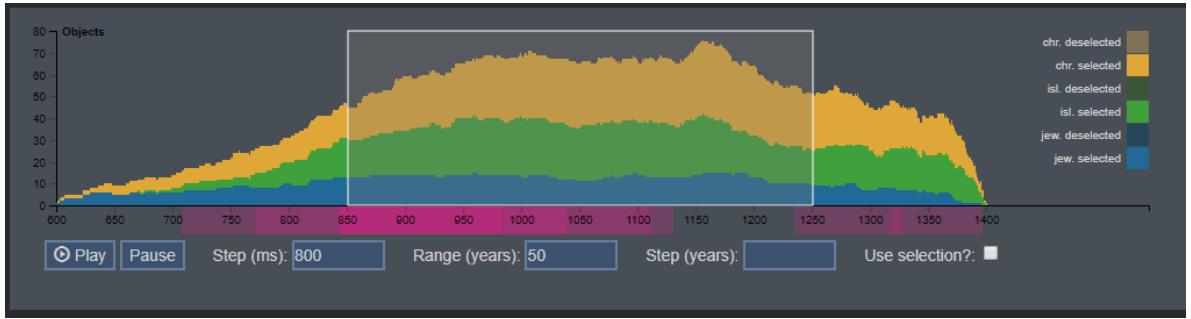
Abbildung 5.20: Zeitleistenansicht mit selektierten und nicht selektierten Abschnitten

- x-Dimension: Die x-Achse repräsentiert den zeitlichen Verlauf, beginnend mit dem frühesten und endend mit den spätesten Jahr, welche jeweils im ersten Datensatz referenziert sind. Jedem Jahr in dieser Zeitspanne wurde hierbei ein diskreter Wert auf der x-Achse zugewiesen. Innerhalb dieser Dimension kann der Anwender eine Zeitspanne selektieren.
- y-Dimension: Die y-Achse beschreibt die Anzahl der Objekte, welche für den entsprechenden Wert der x-Achse relevant sind. Relevant sind hierbei alle Datenobjekte aus dem ersten Datensatz, welche eine Zeitspanne referenzieren, in der das Jahr des entsprechenden x-Wertes enthalten ist. Die Stapelung der Säulen visualisiert anteilig, wie viele Datenobjekte, welche dem entsprechenden x-Wert zugeordnet sind, welche übergeordnete Religion (hier: Christentum, Islam, Judentum) referenzieren. Dementsprechend sind die Säulen in drei Abschnitte, also einen Abschnitt für jede übergeordnete Religion, eingeteilt. Werden allerdings durch die Selektion innerhalb anderer Ansichten Datenobjekte deseletktiert, so wird jeder der drei Abschnitte nochmal aufgeteilt in einen Abschnitt, welcher den selektierten Anteil repräsentiert, und einen Abschnitt, welcher den entsprechend nicht selektierten Anteil repräsentiert. Die Farben der einzelnen Abschnitte wurden so gewählt, dass die Zusammengehörigkeit von selektiertem zu nicht selektiertem Anteil eines Religionsabschnittes deutlich und intuitiv erkennbar ist, die Unterscheidung zwischen diesen beiden Abschnitten und die Unterscheidung zwischen mehreren Religionsabschnitten aber dennoch gewährleistet ist. Ein Beispiel für die Aufteilung in die drei beziehungsweise sechs Abschnitte ist in Abbildung 5.20 dargestellt.

## 5 Umsetzung

---

Wird nun innerhalb einer anderen Ansicht eine Selektion verändert, so werden lediglich die Säulen und deren Aufteilungen in die verschiedenen Abschnitte gemäß der neuen Filterdaten aktualisiert, nicht aber die Achsen und die Legende. Falls die linke Spalte des Layouts eingeklappt wird, dehnt sich das Säulendiagramm aus, um den nun verfügbaren Platz auszunutzen.

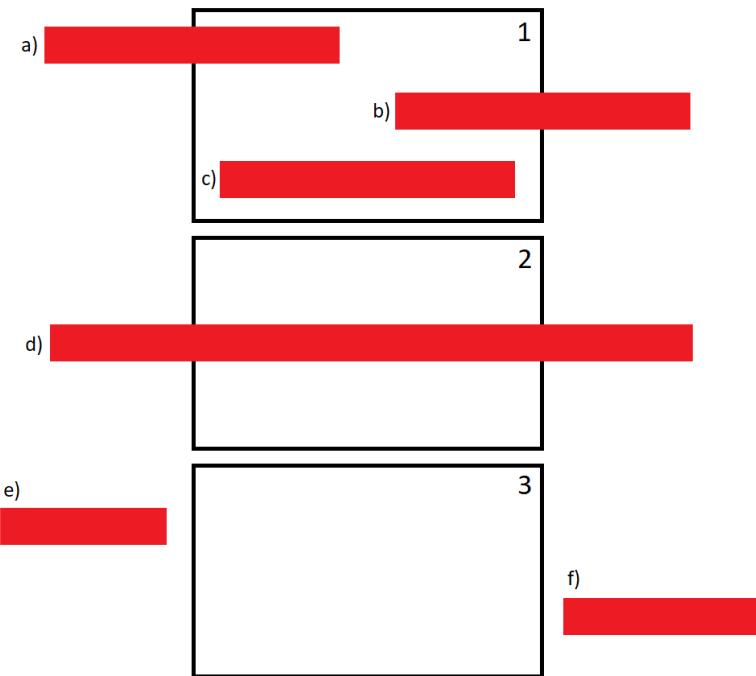


**Abbildung 5.21:** Die Selektion wird durch ein transparentes Rechteck über den gestapelten Säulen visualisiert. Wird ein Knoten innerhalb der Kartenansicht ausgewählt, so werden die in den durch den Knoten repräsentierten Datenobjekten referenzierten Zeitspannen mittels transparenter, pinker Rechtecke über der Beschriftung der x-Achse hervorgehoben.

Für die Selektion einer Zeitspanne innerhalb des Diagramms stehen folgende Operationen zur Verfügung:

- Durch das Klicken und Gedrückthalten mit dem Mauszeiger in den Bereich der Säulen innerhalb des Diagrammes wird der Anfang oder das Ende der selektierten Zeitspanne ausgewählt. zieht der Anwender nun den Mauszeiger nach links oder rechts erscheint ein transparentes Rechteck im Vordergrund der Säulen, welches sich von dem ausgewählten Anfang beziehungsweise dem Ende der Zeitspanne bis zur aktuellen Position des Mauszeigers erstreckt. Lässt der Anwender nun die Maustaste los, bleibt das Rechteck in seiner aktuellen Gestalt und zeigt damit die selektierte Zeitspanne an. Während der Anwender das zweite Ende des Rechteckes an die gewünschte Position zieht, werden zu jedem Zeitpunkt, in dem das Rechteck einen weiteren Wert der x-Achse mit einschließt, die Visualisierungen in den anderen Ansichten automatisch aktualisiert.
- Wurde bereits eine Selektion getätigt, so kann der Anwender diese Selektion manipulieren. Durch Klicken auf das Rechteck und Gedrückthalten und Bewegung der Maus in horizontaler Dimension kann das Rechteck, und damit automatisch auch die Selektion, nach links oder rechts verschoben werden. Auf die selbe Art und Weise kann der linke Anfang oder das rechte Ende des Rechtecks geklickt und nach links oder rechts gezogen werden, um den selektierten Bereich zu verkleinern oder zu vergrößern.
- Möchte der Anwender die Selektion nun aufheben, um wieder die gesamte Zeitspanne zu selektieren, so funktioniert dies, indem er innerhalb des Bereiches der Säulen im Diagramm auf einen Abschnitt klickt, der nicht vom Selektionsrechteck verdeckt ist.

Ein Beispiel für eine Selektion innerhalb der Zeitleiste ist in Abbildung 5.21 dargestellt.



**Abbildung 5.22:** Überschneidet sich die durch ein Datenobjekt referenzierte Zeitspanne (a, b, c, d in 1, 2) mit der selektierten Zeitspanne (schwarzes Rechteck), so wird das Datenobjekt in die Selektion aufgenommen. Gibt es keine Überschneidung (e, f in 3), so wird das Datenobjekt nicht aufgenommen.

Um zu erklären, welche Daten bei welcher Selektion selektiert oder eben nicht selektiert werden, muss zwischen folgenden Fällen unterschieden werden:

- Falls der Anfang und/oder das Ende der Zeitspanne eines Datenobjektes aus dem ersten Datensatz innerhalb des selektierten Bereiches liegt, so wird das entsprechende Objekt in die Selektion aufgenommen.
- Falls der Anfang der Zeitspanne eines Datenobjektes vor dem Anfang der selektierten Zeitspanne und das Ende entsprechend nach dem Ende der selektierten Zeitspanne liegen, so wird das entsprechende Objekt ebenfalls in die Selektion aufgenommen.
- Andernfalls, also falls sowohl Anfang als auch Ende der Zeitspanne eines Datenobjektes entweder vor oder nach der selektierten Zeitspanne liegen, es also keine Überschneidung mit dieser gibt, so wird das entsprechende Objekt nicht in die Selektion aufgenommen.

Diese Fälle werden in Abbildung 5.22 schematisch dargestellt.

Die Zeitleiste kann animiert werden, so dass nacheinander aufeinanderfolgende Zeitspannen automatisch selektiert werden. Für die Konfiguration der Animation stehen folgende Eingabemöglichkeiten zur Verfügung:

## 5 Umsetzung

---

- Millisekunden: Mittels eines Eingabefeldes kann spezifiziert werden, wie lange ein Selektionszustand beibehalten werden soll, bevor die nächste Zeitspanne selektiert wird. Falls keine Eingabe getätigt wird, wird der Standardwert von 800 Millisekunden verwendet.
- Zeitspanne: Über das nächste Eingabefeld kann spezifiziert werden, wie viele Jahre eine Selektion umfassen soll. Der Standardwert beträgt 50 Jahre.
- Zeitschritte: Dieses Eingabefeld dient der Spezifikation der Anzahl der Jahre, um welche die Selektion beziehungsweise die Zeitspanne beim Wechsel zum nächsten Selektionszustand verschoben werden soll. Wird hier kein Wert explizit angegeben, so wird die Zeitspanne immer um die Anzahl der darin enthaltenen Jahre verschoben, so dass keine Überschneidungen zwischen den Selektionszuständen entstehen.
- Kontext: Mittels einer Checkbox kann bestimmt werden, ob die gesamte Zeitleiste oder nur die manuell selektierte Zeitspanne animiert werden soll.

Wenn der Anwender durch das Bewegen des Mauszeigers über einen Knoten innerhalb der Kartenansicht entsprechend repräsentierte Datenobjekte auswählt, so werden die davon referenzierten Zeitspannen durch transparente, pinke Rechtecke über der Beschriftung der x-Achse hervorgehoben. Durch die Transparenz werden die Überschneidungen mehrerer Zeitspannen ersichtlich. Ein Beispiel hierfür wird in Abbildung 5.21 dargestellt.

Alternativ oder auch zusätzlich könnte die Zeitleiste als Stacked Line Graph umgesetzt werden. Denkbar wäre eine Option, mittels welcher der Anwender die aktuell angezeigte durch die jeweils andere Visualisierungsform austauschen könnte.

Um zu vermeiden, dass das gestapelte Säulendiagramm zu viele Säulenabschnitte enthält, könnte auch für jede Religion ein separates Diagramm erstellt werden, welches jeweils nur zwei Abschnitte enthält - einen, für den selektierten und einen für den nicht selektierten Anteil. Dieser Ansatz ging allerdings mit einem wesentlich höherem Platzbedarf einher. Außerdem ist der Vergleich von visualisierten Daten innerhalb einer Grafik effektiver als der Vergleich zwischen mehreren voneinander getrennten Grafiken.

Um die Abstraktion der Daten zu vermeiden, könnten alle referenzierten Gruppen durch einen eigenen Abschnitt innerhalb des Diagramms dargestellt werden. Hierbei wäre allerdings ab einer gewissen Anzahl von Gruppen die Wahl der Farben für die vielen gestapelten Säulen besonders schwierig, sollen doch die Abschnitte voneinander gut unterscheidbar und jeweils wiederum in zwei zueinander gehörende Abschnitte unterteilt sein. Aus diesem Grund wurden innerhalb dieser Ansicht alle Gruppen einer Religion zu einem Abschnitt für diese übergeordnete Religion zusammengefasst beziehungsweise abstrahiert.

Falls die exakte Repräsentation der einzelnen Daten nicht von großer Bedeutung ist, könnte die x-Achse so spezifiziert werden, dass aufeinanderfolgende x-Werte gleichförmige Sprünge von mehreren Jahren repräsentieren.

Für das Hervorheben der Zeitspannen bei der Auswahl repräsentierter Datenobjekte durch einen Knoten in der Kartenansicht wurde eine überlappende Darstellung transparenter Rechtecke gewählt,

da eine Darstellung nicht kollidierender Objekte mit einem enormen Platzbedarf einherginge, besonders bei Knoten, welche sehr viele Datenobjekte repräsentieren, wie dies bei Knoten der niedrigeren Zoomstufen meist der Fall ist.

### 5.3.5 Verbindungen

In diesem Abschnitt sollen die Verbindungen zwischen den verschiedenen Ansichten, welche in den vorangegangenen Abschnitten bereits angerissen wurden, kurz rekapituliert werden. Der beschriebene Ansatz enthält vier Ansichten, welche für die Selektion und Filterung der Daten im Hintergrund verwendet werden.

Die Kartenansicht bildet das Herzstück der Anwendung und präsentiert die selektierten und gefilterten Daten auf einer geographischen Karte. Jede Ansicht ist jeweils mit jeder anderen Ansicht verbunden, da die Visualisierungen in jeder Ansicht zu jeder Zeit von den selektierten und gefilterten Daten abhängen. Wird also die Selektion innerhalb einer Ansicht geändert, so aktualisieren sich die Visualisierungen in allen anderen Ansichten.

In der Kartenansicht können nicht nur Daten selektiert, sondern auch ausgewählt werden, was dazu führt, dass entsprechende Daten in den anderen Ansichten hervorgehoben werden. Somit bestehen jeweils zwischen der Kartenansicht und einer der anderen Ansichten zwei Beziehungen in verschiedene Richtungen.

Diese Beziehungen sollen nachfolgend nochmal kurz zusammengefasst werden:

- Innerhalb der Quellenansicht ist jede einzelne Quelle mit einem Prozentbalken versehen, welcher visualisiert, wie viel Prozent der relevanten Datenobjekte aktuell selektiert sind. Ändert sich die Selektion innerhalb dieser oder einer anderen Ansicht, so werden diese Prozentbalken entsprechend aktualisiert. Wird ein Knoten innerhalb der Kartenansicht durch das Bewegen des Mauszeigers über diesen ausgewählt, so werden die Bezeichner aller Quellen, welche in den von diesem Knoten repräsentierten Datenobjekten referenziert werden, innerhalb der Quellenansicht visuell hervorgehoben.
- Innerhalb der Gruppenansicht besitzt jeder Baumknoten, und damit jede religiöse Gruppe, einen Prozentbalken. Der Prozentbalken visualisiert, wie viel Prozent der Datenobjekte aus dem ersten Datensatz, welche die entsprechende Gruppe referenzieren, aktuell selektiert sind. Der Balken wird aktualisiert, sobald sich die Selektion innerhalb dieser oder einer anderen Ansicht ändert. Wird ein Knoten innerhalb der Kartenansicht ausgewählt, so wird in der Gruppenansicht hervorgehoben, welche Gruppen in den entsprechend repräsentierten Datenobjekten referenziert werden.
- Die gestapelten Säulen innerhalb der Zeitleiste geben jeweils an, wie viele Datenobjekte welche übergeordnete Religion referenzieren. Innerhalb jedes Religionsabschnittes wird nochmal zwischen zwei Abschnitten unterschieden, welche den jeweils selektierten und nicht selektierten Anteil der Datenobjekte visualisieren. Ändert sich nun die Selektion innerhalb einer anderen Ansicht, so ändert sich hier die Aufteilung der beiden Abschnitte innerhalb jedes von der Änderung betroffenen Religionsabschnittes. Wird innerhalb der Kartenansicht ein Knoten

## 5 Umsetzung

---

ausgewählt, so wird an der Beschriftung der x-Achse der Zeitleiste hervorgehoben, welche Zeitspannen durch die durch diesen Knoten repräsentierten Datenobjekte referenziert werden.

- Wird innerhalb einer der vier Ansichten die Selektion verändert, so aktualisiert sich die Darstellung jedes Knotens innerhalb der Kartenansicht. Knoten, welche lediglich einzelne Datenobjekte repräsentieren, wechseln visuell zwischen dem selektierten und dem nicht selektierten Zustand, während sich bei Knoten, welche mehrere aggregierte Datenobjekte repräsentieren, die Kuchen- und Donutdiagramme aktualisieren, welche angeben, wie viel Prozent der repräsentierten Datenobjekte aktuell selektiert sind.

Detailliertere Beschreibungen dieser Beziehungen sind in den entsprechenden Abschnitten dieses Kapitels zu finden.

# 6 Ergebnisse

In diesem Kapitel werden zunächst einige, ausgewählte Anwendungsszenarien präsentiert und erläutert. Anschließend werden die Ergebnisse einer Expertenbefragung dargestellt. Zusammenfassend werden dann die Resultate dieser beiden Abschnitte diskutiert.

## 6.1 Anwendungsszenarien

Die nachfolgenden Abschnitte beschreiben, wie der entwickelte Ansatz verwendet werden kann, um verschiedene Fragestellungen zu bearbeiten. Wichtig ist hierbei anzumerken, dass die verwendeten Daten zufällig generiert wurden und daher keinesfalls historisch belegte Sachverhalte beschreiben.

### 6.1.1 Untersuchung einer Quelle

Mittels dieses Anwendungsszenarios soll beispielhaft angedeutet werden, wie der entwickelte Ansatz dabei helfen kann, einzelne Quellen und deren Zusammensetzung zu analysieren. Konkret soll dabei die Quelle 5 untersucht werden. Da standardmäßig beim Start des Programms alle Quellen selektiert sind, deseletieren wir nun alle Quellen, bis auf Quelle 5, mittels entsprechender Checkboxen, wie in Abbildung 6.1 dargestellt.



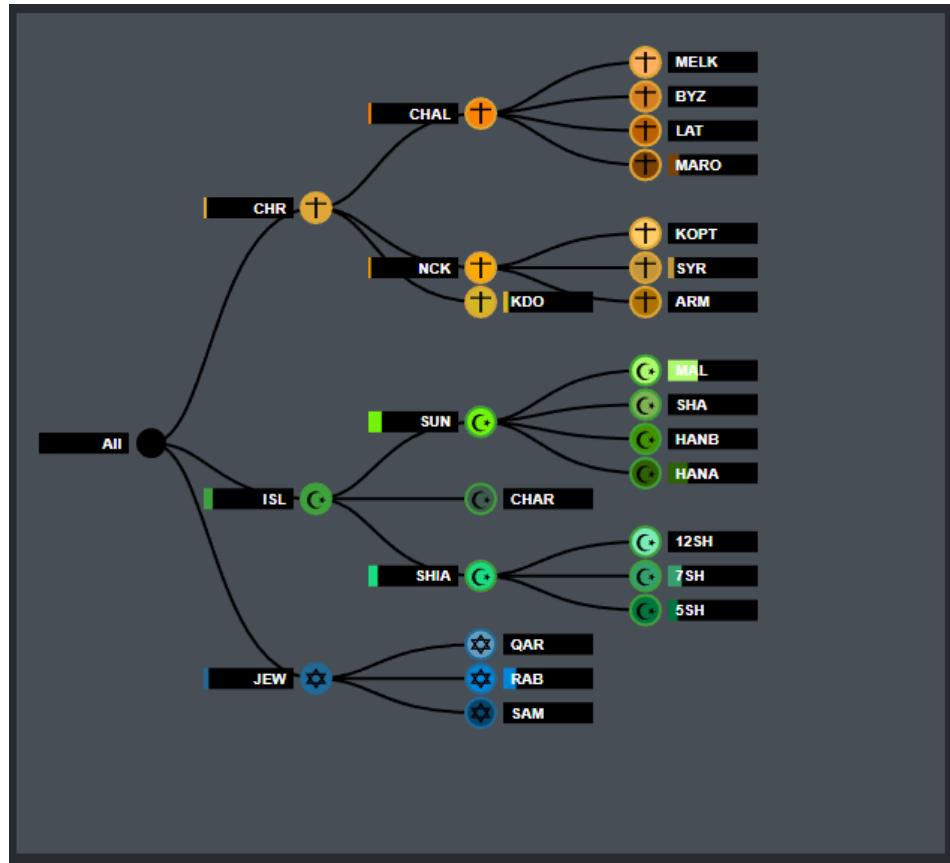
Abbildung 6.1: Alle Quellen, bis auf Quelle 5, wurden deseletiert.

Durch die Betrachtung der aktualisierten, anderen Ansichten können wir nun Informationen über diese Quelle erarbeiten. In der Gruppenansicht, dargestellt in Abbildung 6.2, können wir erkennen, dass die Datenobjekte, welche Quelle 5 referenzieren, überwiegend islamische Gruppen beschreiben.

Selektieren wir nacheinander die drei Hauptäste des Gruppenbaumes, so können wir innerhalb der Quellenansicht ablesen, wie viele Objekte jeweils eine Gruppe der entsprechenden Religion referenzieren. So erfahren wir, dass drei Objekte christliche Gruppen, acht Objekte islamische Gruppen und lediglich zwei Objekte jüdische Gruppen beschreiben. Die drei Selektionszustände sind in Abbildung 6.3 dargestellt.

## 6 Ergebnisse

---

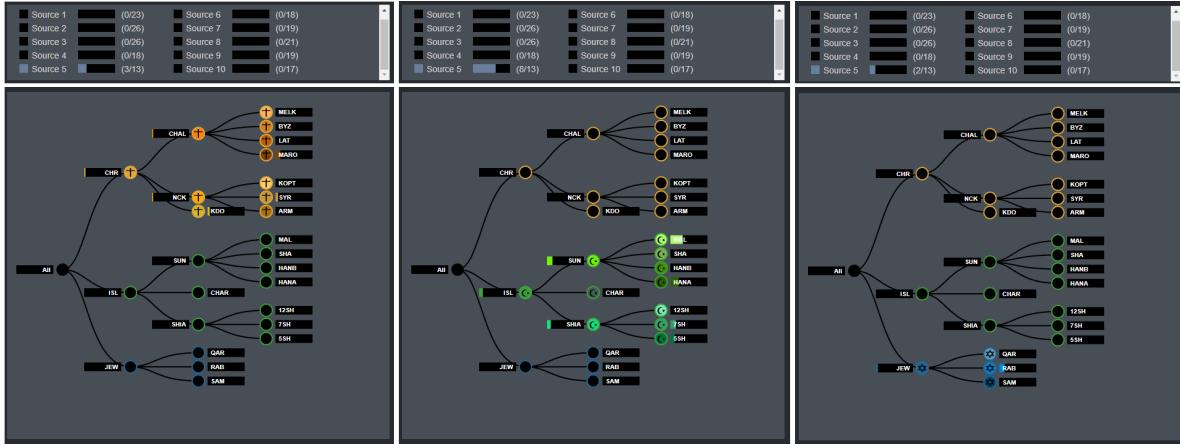


**Abbildung 6.2:** Islamische Gruppen sind stärker vertreten als christliche oder jüdische Gruppen.

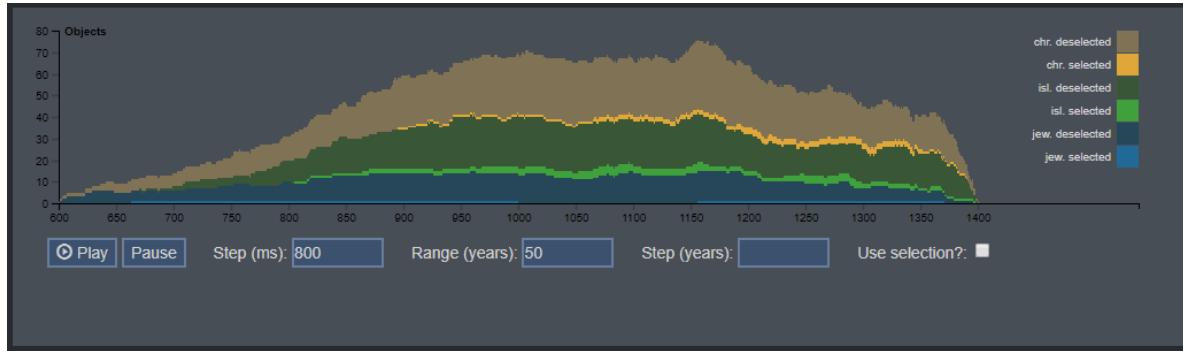
Nun selektieren wir wieder die gesamte Hierarchie der religiösen Gruppen durch einen Klick auf das Label der Wurzel des Baumes. Durch Betrachtung der Zeitleiste, dargestellt in Abbildung 6.4, können wir ablesen, in welchen Zeiträumen religiöse Gruppen, welche durch Quelle 5 beschrieben werden, aktiv waren.

So erkennen wir, dass jüdische Gruppen zwischen 650 und 1000 n. Chr. und zwischen 1150 und 1400 n. Chr. aktiv waren. Islamische Gruppen waren durchgängig zwischen 800 und 1400 n. Chr., christliche Gruppen zwischen 850 und 1400 n. Chr. aktiv. Nun wollen wir bestimmen, in welchen Regionen des ausgewählten geographischen Bereichs die entsprechenden Gruppen angesiedelt waren. Dafür klappen wir die linke Spalte und die Zeitleiste weg und vergrößern damit die Kartenansicht. Sowohl auf niedriger als auch auf höherer Zoomstufe wird deutlich, dass die in Quelle 5 beschriebenen Gruppen über den gesamten Bereich verteilt waren. Dies kann anhand der Abbildungen 6.5 und 6.6 nachvollzogen werden.

## 6.1 Anwendungsszenarien



**Abbildung 6.3:** Drei Selektionszustände: Erst sind alle christlichen, dann alle islamischen und schließlich alle jüdischen Gruppen selektiert.



**Abbildung 6.4:** Religionen, beschrieben durch Quelle 5, im zeitlichen Verlauf

### 6.1.2 Untersuchung einer Religion

In diesem Abschnitt soll mittels des entwickelten Ansatzes eine Religion untersucht werden. Für dieses Beispiel wählen wir die jüdische Religion. Dafür selektieren wir in der Gruppenansicht über den entsprechenden Knoten den gesamten Teilbaum, welcher diese Religion und deren Gruppen beschreibt. Mittels der Quellenansicht, dargestellt in Abbildung 6.7, können wir nun analysieren, welche Quellen für die Betrachtung dieser Religion bedeutend sind.

Auf den ersten Blick erkennen wir anhand der Prozentbalken, dass die jüdische Religion in allen Quellen nur geringfügig vertreten ist. Die Quellen mit den meisten Objekten, welche Gruppen der jüdischen Religion beschreiben, sind die Quellen 2, 9 und 10. Das sind außerdem die Quellen, innerhalb welcher die jüdische Religion prozentual größer vertreten ist, als in den anderen Quellen. Durch die Betrachtung der Zeitleiste ist zu erkennen, dass die jüdische Religion durchgehend von 600 bis 1400 n. Chr. aktiv war, allerdings ist auch hier zu erkennen, dass sie schwächer vertreten war, als die beiden anderen Religionen. Dies ist in Abbildung 6.8 zu sehen.



**Abbildung 6.5:** Geographische Verteilung religiöser Gruppen, welche durch Quelle 5 Beschrieben werden, auf niedriger Zoomstufe

Selektieren wir nacheinander die drei Gruppen dieser Religion innerhalb der Gruppenansicht, so können wir anhand der Zeitleiste erkennen, dass die Gruppe der Qaraiten erst ab 750 n. Chr. vertreten war, während die anderen beiden Gruppen nahezu durchgängig zwischen 600 und 1400 n. Chr. vertreten waren. Dieser Vergleich ist in Abbildung 6.9 dargestellt.

Nun selektieren wir wieder die gesamte Religion und vergrößern die Kartenansicht durch Einklappen der linken Spalte und der Zeitleiste. So erkennen wir auch hier wieder, dass die jüdische Religion über den gesamten Bereich verteilt aktiv war.

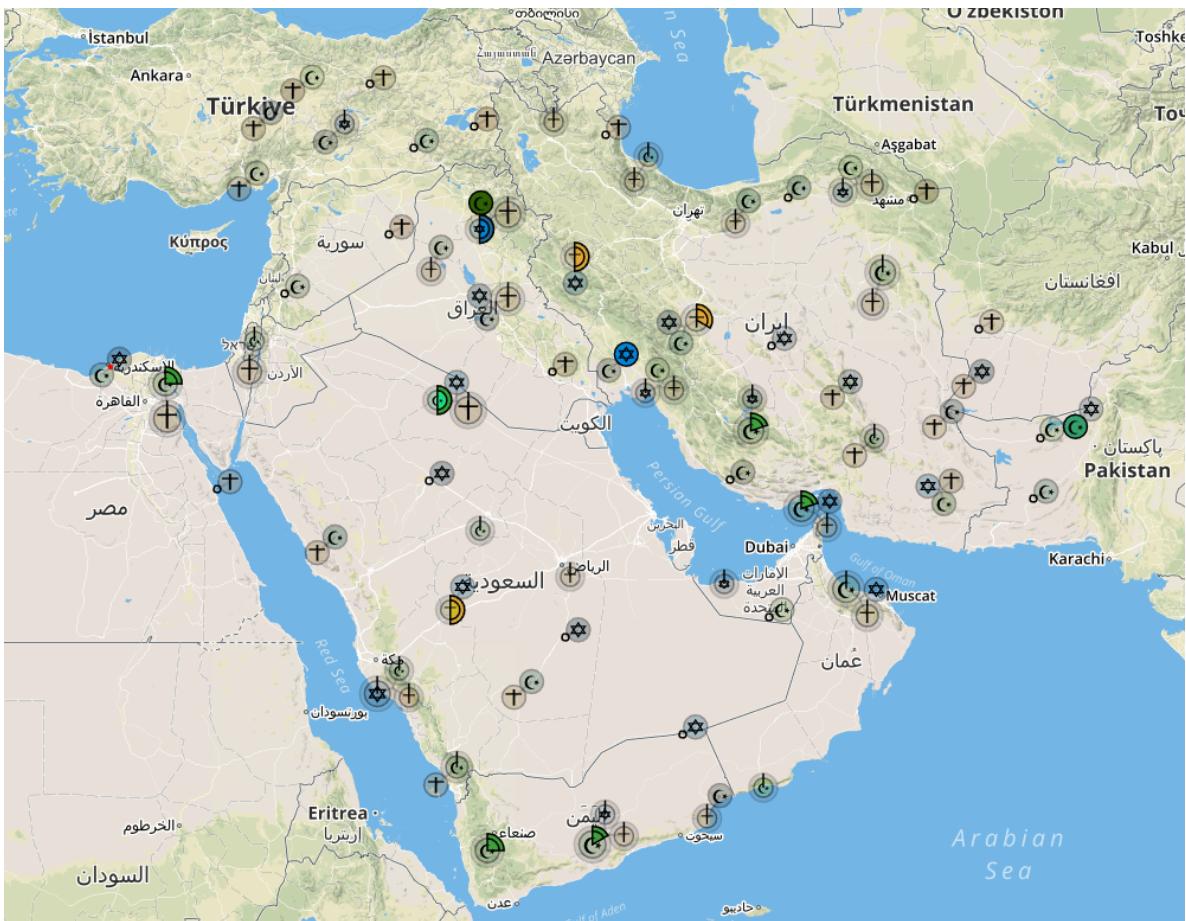
### 6.1.3 Untersuchung einer Zeitspanne

In diesem Abschnitt soll angedeutet werden, wie mittels des vorgestellten Ansatzes eine Zeitspanne analysiert werden kann. Dafür selektieren wir innerhalb der Zeitleistenansicht die Zeitspanne zwischen 600 und 700 n. Chr., wie in Abbildung 6.10 abgebildet.

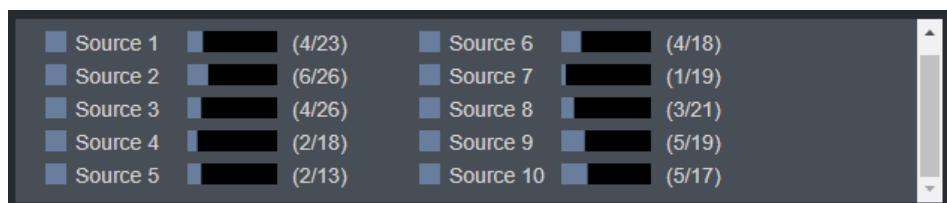
Bereits in dieser Ansicht ist zu erkennen, dass nur sehr wenige Datenobjekte diese Zeitspanne referenzieren. Außerdem kann abgelesen werden, dass hier hauptsächlich die jüdische und die christliche Religion aktiv waren, und nur geringfügig die islamische. Innerhalb der Quellenansicht fällt auf, dass die Quellen 4 und 10 prozentual am meisten Objekte referenzieren, welche diese Zeitspanne beschreiben, wie in Abbildung 6.11 dargestellt.

Klappen wir nun innerhalb der Gruppenansicht alle drei Äste, welche die drei Religionen beschreiben, ein und vergleichen die Prozentbalken, so fällt sofort auf, dass in dieser Zeitspanne hauptsächlich die jüdische Religion aktiv war, gefolgt von der christlichen, während die islamische Religion nur sehr geringfügig vertreten war. Dieser Vergleich ist in Abbildung 6.12 nachvollziehbar.

## 6.1 Anwendungsszenarien



**Abbildung 6.6:** Geographische Verteilung religiöser Gruppen, welche durch Quelle 5 Beschrieben werden, auf höherer Zoomstufe



**Abbildung 6.7:** Quellenansicht, während lediglich der jüdische Teilbaum in der Gruppenansicht selektiert ist

Wenn wir nun wieder den gesamten Gruppenbaum betrachten, können wir anhand der Prozentbalken sofort erkennen, dass in dieser Zeitspanne vor allem die Gruppen der Rabbaniten (RAB) und der Byzantinischen Kirche (BYZ) vertreten waren, wie in Abbildung 6.13 zu sehen.

## 6 Ergebnisse

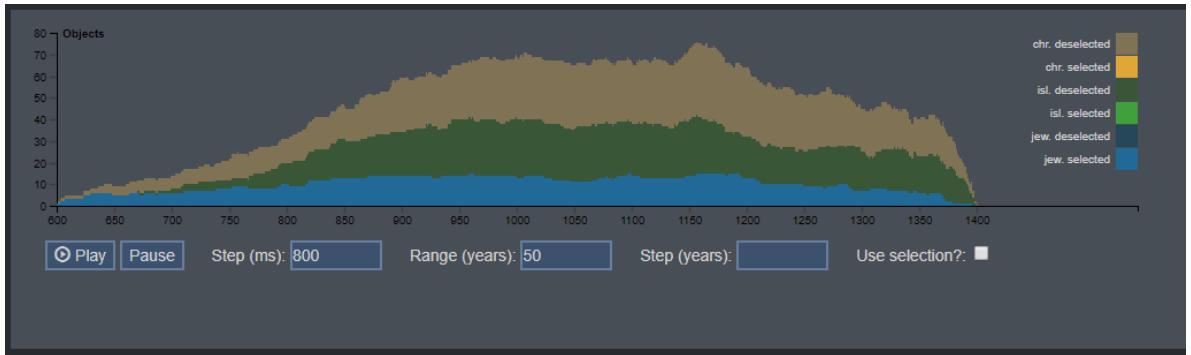


Abbildung 6.8: Die jüdische Religion im zeitlichen Verlauf

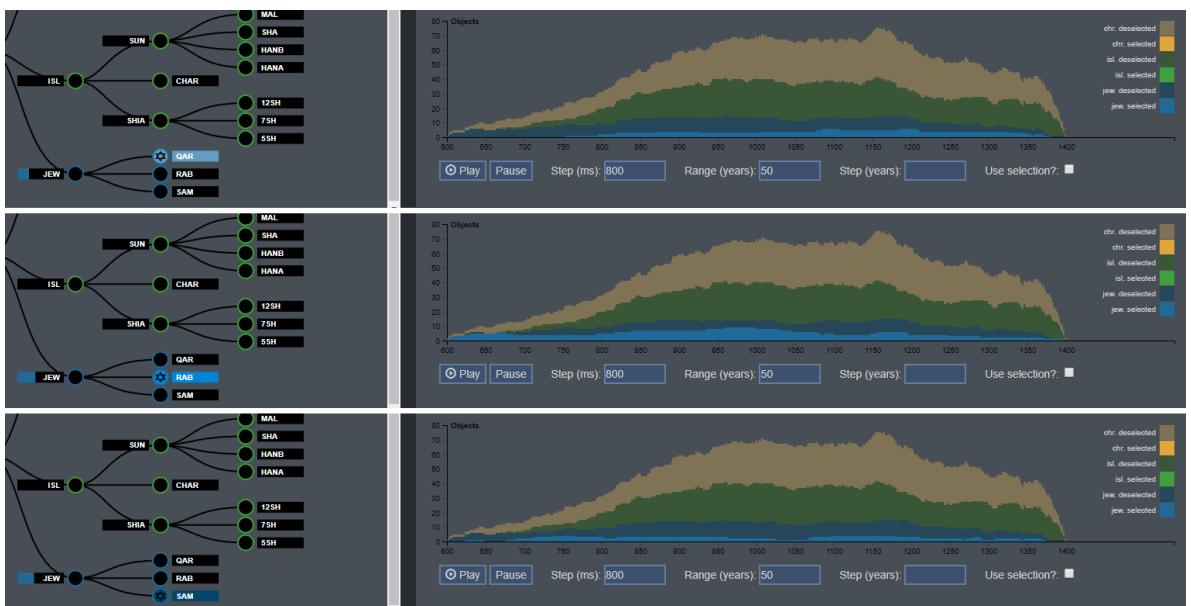


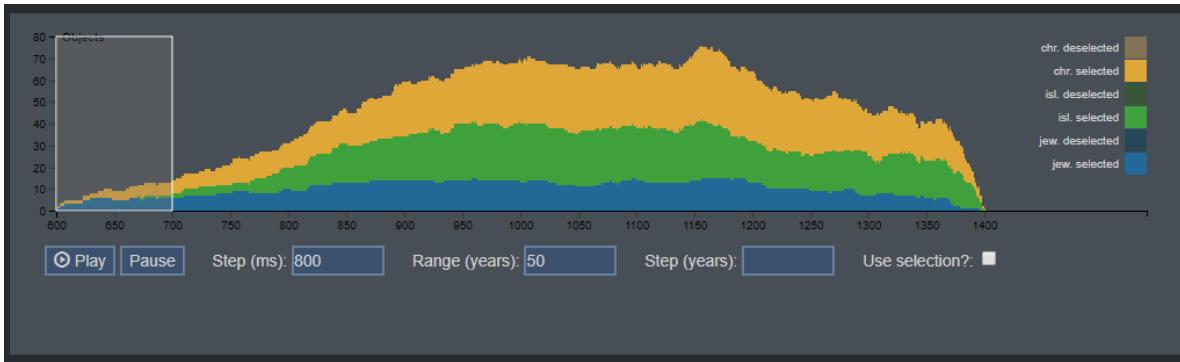
Abbildung 6.9: Einzelne, jüdische Gruppen im zeitlichen Verlauf

Betrachten wir nun die Kartenansicht, dargestellt in Abbildung 6.14, so stellen wir fest, dass während dieser Zeitspanne wenige Gruppen im Norden und Osten des geographischen Bereichs angesiedelt waren.

### 6.1.4 Untersuchung einer Region

Durch diesen Abschnitt soll angedeutet werden, welche Möglichkeiten der entwickelte Ansatz bietet, um geographische Regionen zu analysieren. Dafür klappen wir zunächst die linke Spalte und die Zeitleiste weg, um die Karte vergrößert betrachten zu können, und selektieren innerhalb der Kartenansicht durch das Zeichnen eines Polygons eine geographische Region, wie in Abbildung 6.15 dargestellt.

## 6.2 Expertenbefragung



**Abbildung 6.10:** Die Zeitspanne zwischen 600 und 700 n. Chr. ist in der Zeitleistenansicht selektiert.



**Abbildung 6.11:** Quellenansicht, während die Zeitspanne zwischen 600 und 700 n. Chr. in der Zeitleistenansicht selektiert ist

Nun lassen wir uns wieder die anderen Ansichten anzeigen, um die ausgewählte Region zu untersuchen. Betrachten wir die Quellenansicht, so sehen wir auf den ersten Blick anhand der Prozentbalken, dass die selektierte Region ungefähr ein Viertel der meisten, aufgelisteten Quellen ausmacht. Die Quellen 4 und 8 scheinen für diese Region von besonderer Bedeutung zu sein, da hier durch die Selektion innerhalb der Karte jeweils etwa die Hälfte der entsprechenden Datenobjekte selektiert wurden. Diese Erkenntnisse können anhand von Abbildung 6.16 nachvollzogen werden.

Durch die Betrachtung der Gruppenansicht, dargestellt in Abbildung 6.17, stellen wir fest, dass Gruppen der islamischen und der jüdischen Religion in dieser Region etwa gleich stark vertreten sind, während die Gruppen der christlichen Religion hier schwächer repräsentiert sind.

Selektieren wir nun lediglich die Quellen 4 und 8, so kann in der Gruppenansicht abgelesen werden, dass dadurch etwa die Hälfte aller Objekte, welche die Gruppe der Hanbaliten referenzieren, selektiert wurden. Diese Gruppe scheint also eng mit dieser geographischen Region verbunden zu sein.

## 6.2 Expertenbefragung

Da diese Arbeit als Auftakt für ein Kooperationsprojekt zwischen dem Institut für Visualisierung und Interaktive Systeme der Universität Stuttgart und dem Lehrstuhl für Mittelalterliche Geschichte II der Goethe-Universität in Frankfurt dient, bestand die Möglichkeit, einen ausgewiesenen Experten zu dem entwickelten Ansatz bezüglich Nutzen, Stärken, Schwächen und zukünftigen Arbeiten zu befragen. Die Ergebnisse dieser Befragung werden in diesem Abschnitt strukturiert zusammengefasst.

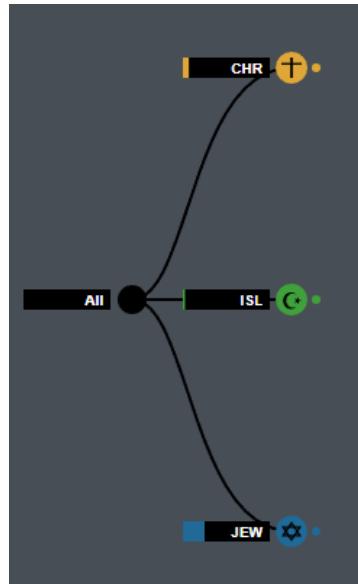


Abbildung 6.12: In der Gruppenansicht wurden alle drei Religionsäste eingeklappt.

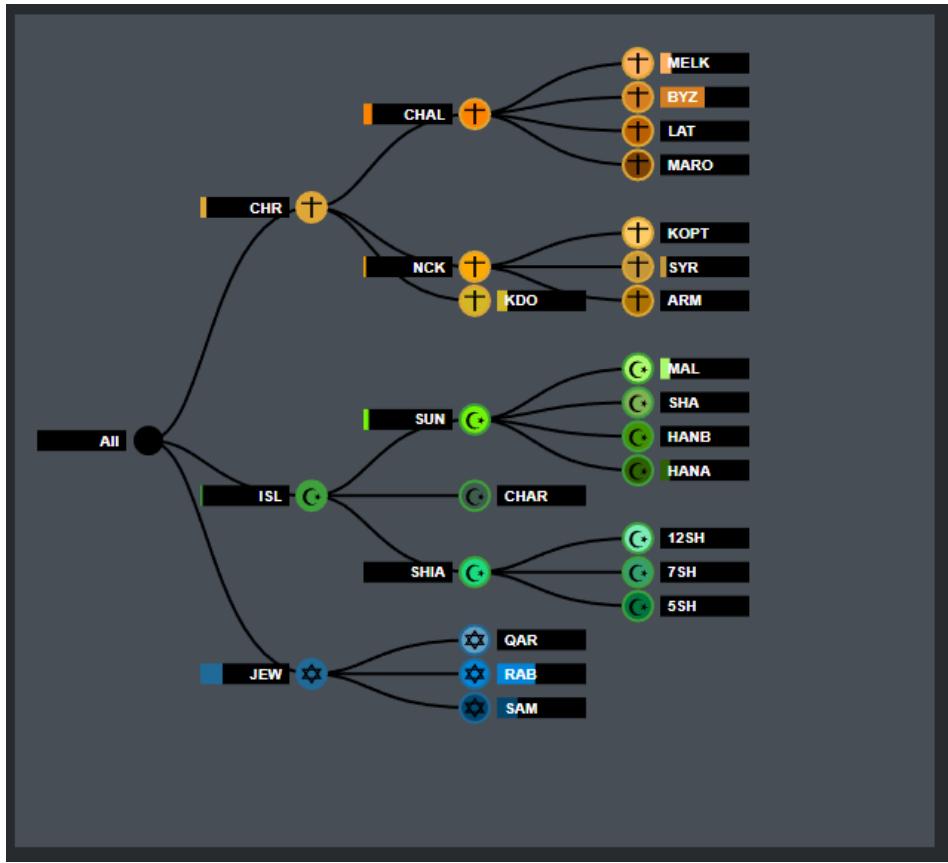
### 6.2.1 Allgemeines

Zunächst ist zu erwähnen, dass der entwickelte Ansatz sehr positiv bewertet und als nützlich angesehen wurde. Den Nutzen sieht der Experte in der Lösung komplexer Fragestellungen, welche mittels statischer, analoger Karten nur schwer zu beantworten wären, und in dem Entdecken und Erforschen neuer, bisher unbekannter Muster und Eigenschaften des visualisierten Datensatzes. Die Möglichkeit, eine größere Fülle an komplexen Daten auf einer Karte zu visualisieren, die manuelle oder automatisierte Selektion des zeitlichen Aspektes durch die Zeitleiste und die Möglichkeit, die Karte nach Bedarf zu zoomen, wurden als hauptsächliche Vorteile gegenüber statischen Karten genannt.

Die Anwendung eines MCV-Ansatzes, wobei die Selektionen in den einzelnen Ansichten koordiniert sind, sodass sich jede Selektion in den anderen Ansichten widerspiegelt und mittels Selektionen in mehreren Ansichten gleichzeitig komplexe Anfragen formuliert werden können, wurde als äußerst hilfreich wahrgenommen.

Obwohl die Kombination von Selektion und Hervorhebung von Informationen als sinnvoll und wichtig beschrieben wurde, sorgten diese beiden parallelen und ähnlichen Funktionen zunächst für Verwirrung. Aus der Sicht des Experten könnte eine Verbesserung dieses Aspektes erreicht werden, wenn die Hervorhebung nur sichtbar wäre, solange sich der Mauszeiger über dem entsprechenden Element befindet, oder falls die Hervorhebung nicht durch das bloße Bewegen des Mauszeigers über ein solches Element, sondern durch explizites Anklicken dieses erfolgen würde.

Bezüglich der Datenobjekte im Hintergrund wurde angemerkt, dass es äußerst hilfreich und wichtig wäre, für ein einzelnes Objekt mehrere Quellen angeben zu können. Außerdem ist eine Erweiterung des Ansatzes denkbar, sodass dieser ebenfalls anwendbar wäre, auch wenn bestimmte Daten fehlen. So kann in realen Szenarien oft nicht präzise bestimmt werden, um welche religiöse Gruppierung es sich bei einem Vorkommen genau handelt, was aktuell ein Problem darstellt, da in den Datenobjekten



**Abbildung 6.13:** Gruppenansicht, während die Zeitspanne zwischen 600 und 700 n. Chr. in der Zeitleistenansicht selektiert ist

bisher explizit eine Gruppe aus der Ebene der Blattknoten referenziert werden muss. Darüber hinaus wäre es gut, könnte der Ansatz auch Daten, welche nicht eindeutig belegt sind und auf Vermutungen basieren, entsprechend verarbeiten und visualisieren.

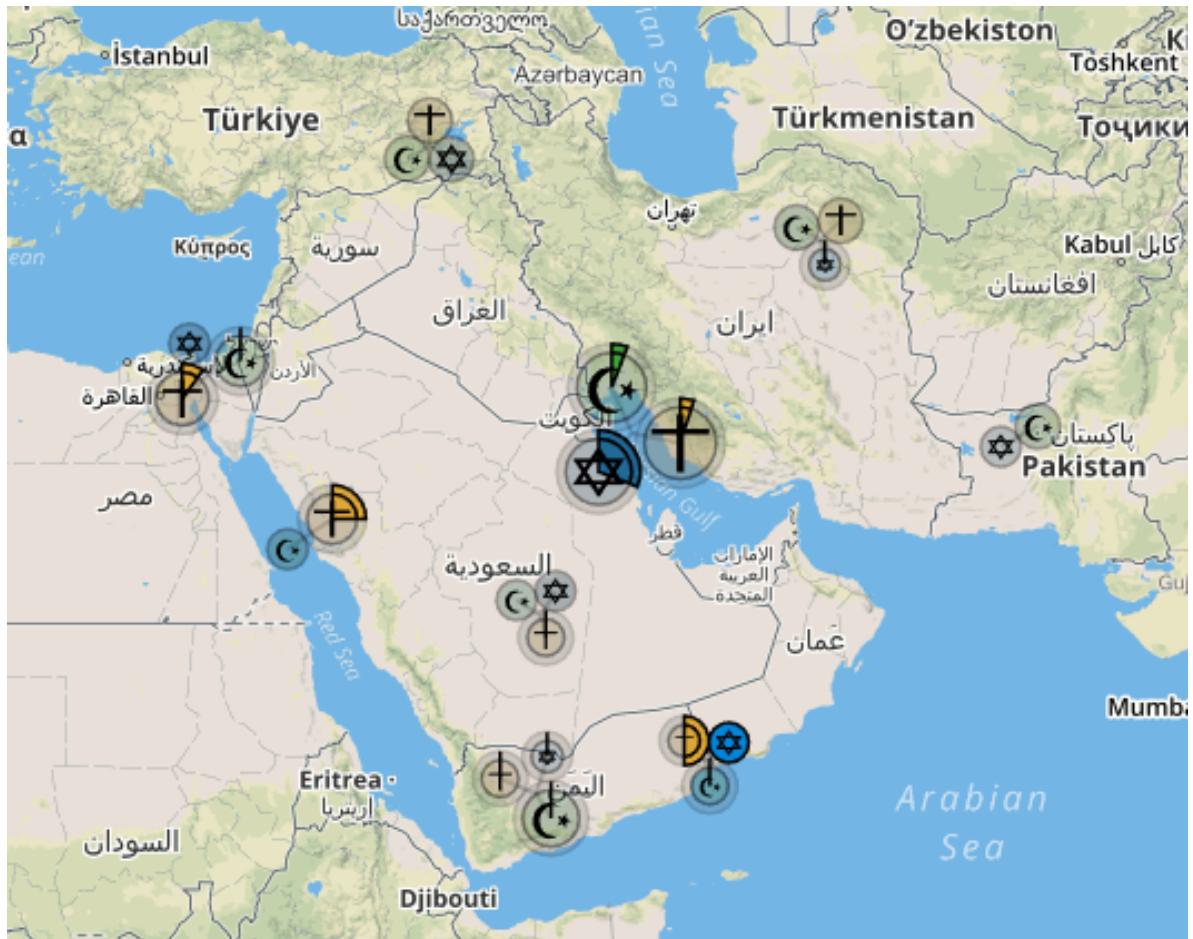
### 6.2.2 Kartenansicht

Allgemein wurde die Visualisierung innerhalb der Kartenansicht positiv bewertet, wobei dennoch einige Merkmale diskutiert und Verbesserungsvorschläge formuliert wurden. So wurde angemerkt, dass es aktuell nicht intuitiv und eindeutig sei, welche visuellen Elemente einzelne und welche mehrere Städte repräsentieren.

Die Auswahl von runden Grafiken, also von Kreisen, Kuchendiagrammen und Donutdiagrammen, für die Visualisierung auf der Karte wurde als sinnvoll und gut geeignet beschrieben. Eine alternative Umsetzung mittels Prozentbalken wurde abgelehnt, da vermutet wurde, dass dadurch der Gesamteinindruck unordentlicher und unübersichtlicher werden würde. Die Kombination aus Kuchendiagrammen

## 6 Ergebnisse

---



**Abbildung 6.14:** Kartenansicht auf niedriger Zoomstufe, während die Zeitspanne zwischen 600 und 700 n. Chr. in der Zeitleistenansicht selektiert ist

und Donutdiagrammen dagegen wurde hinterfragt, sodass wohl die Verwendung von ersteren alleine einen besseren Eindruck ergeben würde.

Das Fenster, welches bei der Auswahl bestimmter Knoten erscheint und deren Zusammensetzung visualisiert, wurde allgemein als ausgesprochen hilfreich bewertet. Für Irritationen und Verwirrung sorgte dabei, dass dieses Fenster nicht bei jedem Knoten erscheint, sondern nur bei Knoten, welche aggregierte Objekte repräsentieren, die wiederum verschiedene religiöse Gruppen referenzieren. Außerdem wurde angemerkt, dass die Balken innerhalb dieses Fensters ebenfalls den aktuellen Selektionszustand widerspiegeln könnten.

Bezüglich der Selektion innerhalb der Kartenansicht durch das Zeichnen eines Polygons wurde angemerkt, dass es ausgesprochen hilfreich wäre, außerdem vordefinierte Polygone für die Selektion zur Verfügung zu stellen, welche geographische Regionen beschreiben. Dass aktuell lediglich ein Polygon gleichzeitig gezeichnet werden kann wurde nicht als problematische Einschränkung wahrgenommen. Problematisch dagegen sei die Tatsache, dass mittels des gezeichneten Polygons die Daten im Hinter-

## 6.2 Expertenbefragung

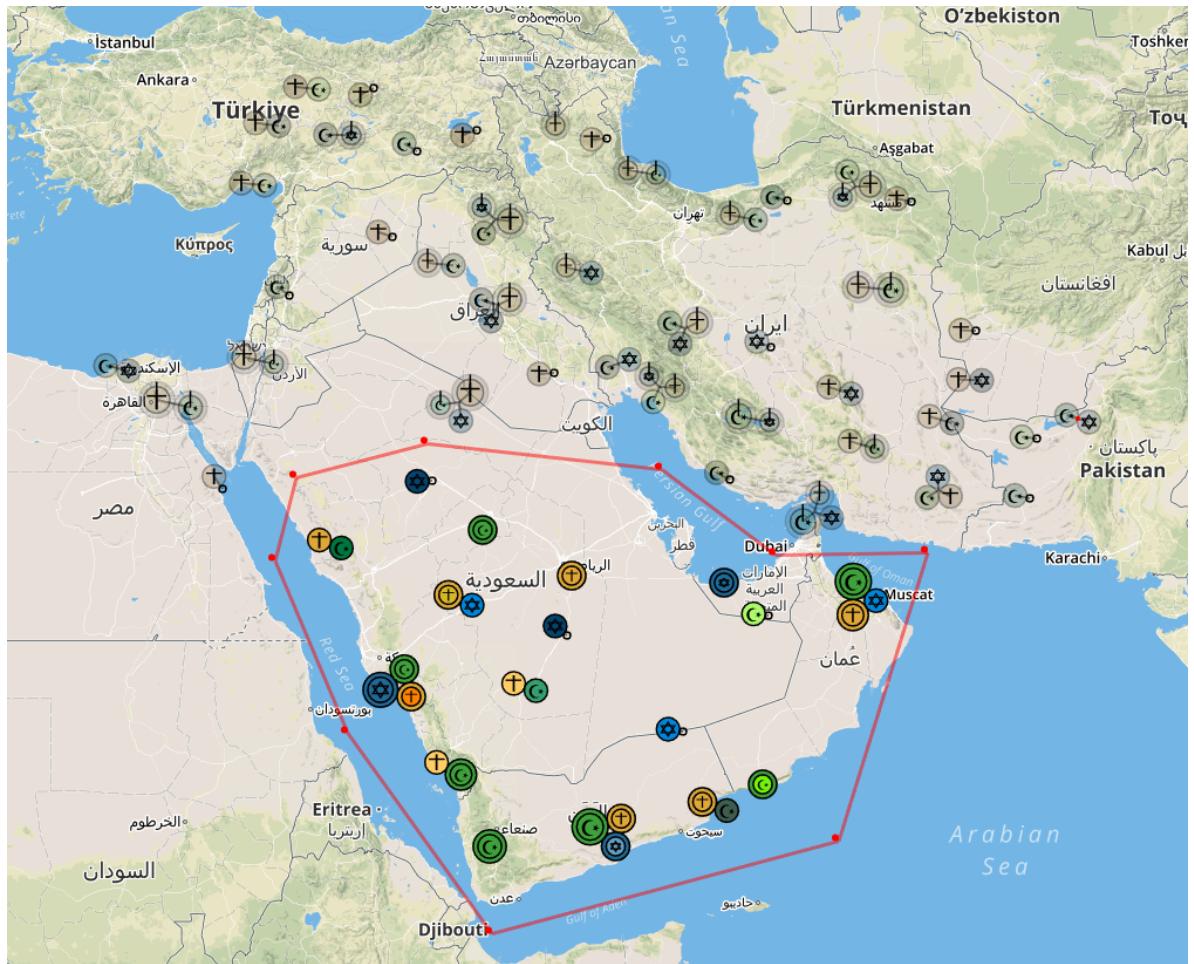


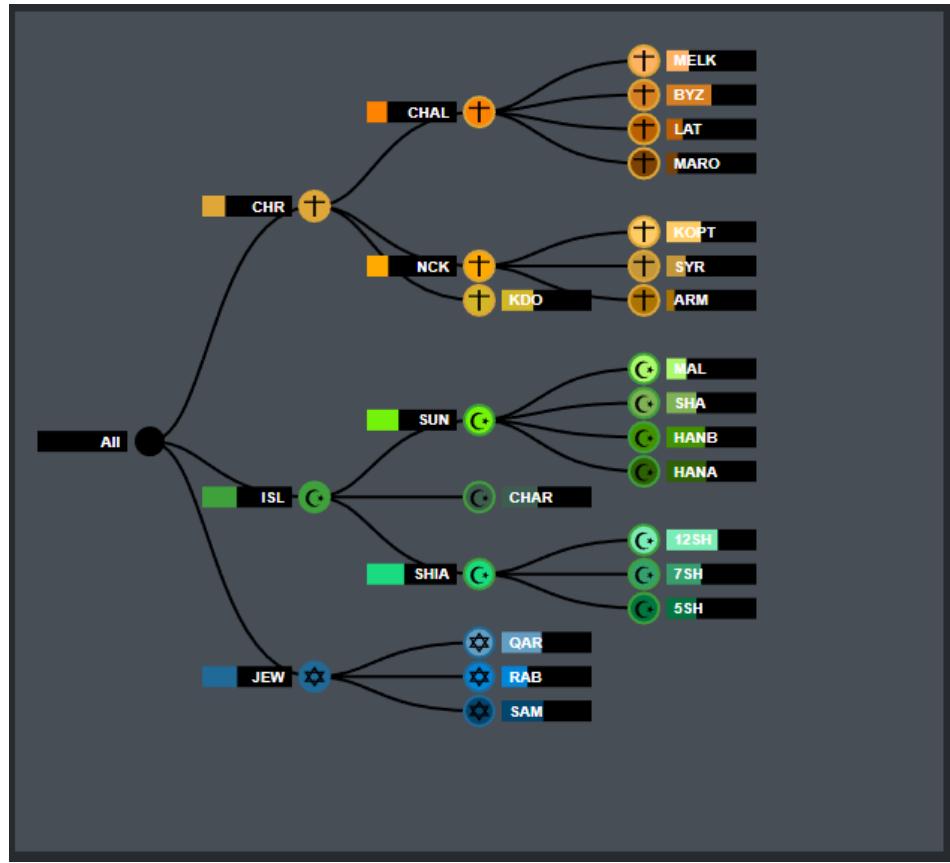
Abbildung 6.15: Eine geographische Region wurde durch das Zeichnen eines Polygons selektiert.



Abbildung 6.16: Quellenansicht, während eine geographische Region in der Kartenansicht selektiert ist

grund, und nicht die visuellen Elemente selektiert werden, wodurch irreführende visuelle Eindrücke entstehen können.

Dass visuelle Elemente, welche Objekte repräsentieren, die aktuell deselektiert sind, nicht einfach verschwinden, sondern lediglich ausgegraunt werden und somit in den Hintergrund rücken, war für den Experten zunächst ungewohnt und verwirrend, wobei sich dieser laut eigener Aussage



**Abbildung 6.17:** Gruppenansicht, während eine geographische Region selektiert ist

schnell daran gewöhnte. Als Vorteil dieser Umsetzung wurde erkannt, dass der Anwender auch mit deselektierten Elementen interagieren und durch die visuellen Hervorhebungen in den anderen Ansichten herausfinden kann, warum diese aktuell deselektiert sind.

Überraschenderweise wurde die Auswahl der Karte im Hintergrund der Visualisierung als nebensächlich und unbedeutend bezeichnet. Lediglich die Darstellung von Gelände, wie zum Beispiel von Seen, Flüssen und Gebirgen, sei dabei relevant. Außerdem sollte der Detailgrad nicht zu hoch sein. Ob die Karte nun farbig oder in Graustufen dargestellt wird, spielt für den Experten keine Rolle. Der Ansatz könnte aber dahingehend erweitert werden, so dass der Anwender zwischen verschiedenen Karten oder verschiedenen Darstellungen einer Karte wechseln könnte.

Kontrovers wurde diskutiert, ob es hilfreich wäre, würde beim Zoomen eine Animation darstellen, wie sich Knoten zu mehreren, kleinen Knoten aufteilen oder zu größeren Knoten zusammenfügen. Neben den offensichtlichen Vorteilen wurde als Nachteil angeführt, dass solch eine Animation auch verwirrend und überfordernd wirken kann, falls viele visuelle Objekte dargestellt sind.

### 6.2.3 Quellenansicht

Bezüglich der Quellenansicht wurden nicht viele Merkmale diskutiert. Die Kombination aus visuellen Prozentbalken und der Angabe expliziter Zahlen wurde positiv bewertet. Falls der Datensatz allerdings einige Quellen beschreibt, sei eine schlichte Auflistung dieser gegebenenfalls nicht mehr praktikabel. Daher müsste eine geeigneter Repräsentationsform hierfür entwickelt oder gefunden werden.

### 6.2.4 Gruppenansicht

Die Entscheidung, jede einzelne religiöse Gruppe eindeutig mittels einer Farbe zu assoziieren, wobei sich die Farben von Gruppen innerhalb eines Astes ähneln, so dass diese Zusammengehörigkeit erkenntlich wird, wurde als sinnvoll und praktikabel bezeichnet. Problematisch sei dabei, dass die Unterscheidbarkeit bei einer hohen Anzahl von Gruppen schwierig ist, allerdings wurde auch angemerkt, dass es aus der Sicht des Experten wohl keine bessere und gleichzeitig intuitive Möglichkeit gibt, um dies besser zu lösen.

Die Selektionslogik innerhalb der Gruppenansicht wurde als zu komplex und kompliziert bewertet, wobei auch problematisch ist, dass einzelne Knoten oder Teilbäume aktuell nicht explizit deseletiert werden können.

Eine sinnvolle Erweiterung wäre hier, wenn der Anwender die Möglichkeit hätte, mehrere Gruppen dahingehend auszuwählen, so dass lediglich solche Elemente innerhalb der Kartenansicht hervorgehoben werden, welche alle selektierten Gruppen repräsentieren.

Dass einzelne Teilbäume bei Bedarf eingeklappt werden können, so dass sich der restliche Baum gemäß des verfügbaren Platzes ausdehnt, wurde zwar positiv bewertet, allerdings würde diese Funktion nicht benutzt werden, weil damit relevante Informationen versteckt werden würden.

### 6.2.5 Zeitleistenansicht

Die Entscheidung, die Daten innerhalb der Zeitleiste auf der Abstraktionsebene der Religionen zu visualisieren, und nicht auf der Ebene der einzelnen religiösen Gruppen, wurde als sinnvoll und praktikabel erachtet.

Auch die Umsetzung der Zeitleiste mittels gestapelter Säulen sei ein guter Kompromiss, um platzsparend alle drei Religionen innerhalb eines Diagramms vergleichend darzustellen. Da hierbei diese und andere, alternative Repräsentationsformen jeweils gewisse Vor- und Nachteile bieten, wäre es ideal, wenn der Anwender hier zwischen verschiedenen Darstellungen wechseln könnte.

Die Möglichkeit der Animation des zeitlichen Verlaufs samt aller Konfigurationsmöglichkeiten wurde äußerst positiv bewertet. Diese Funktionalität sei nicht nur hilfreich, sondern sogar wichtig und notwendig für die effiziente und effektive Lösung der Problemstellungen des Anwenders.

Darüber hinaus wurde angemerkt, dass es von Vorteil wäre, könnte der Anwender mehrere Zeitspannen gleichzeitig selektieren, so dass die verschiedenen Resultate ebenfalls in den anderen Ansichten unterschieden und miteinander verglichen werden könnten.

### 6.3 Diskussion

Mittels der weiter oben beschriebenen Anwendungsszenarien wurde angedeutet, wie der entwickelte Ansatz verwendet werden kann, um verschiedenste Problemstellungen zu lösen, den Datensatz zu analysieren und um neue Erkenntnisse zu gewinnen. Durch die Expertenbefragung wurde deutlich, dass der Ansatz einige, bedeutende Vorteile gegenüber statischen, analogen Karten bietet und gut geeignet ist, um den Anwender bei der Lösung komplexer Fragestellungen und dem Auffinden und Analysieren neuer, unbekannter Muster und Eigenschaften der Daten zu unterstützen.

Die Anwendung eines Multiple Coordinated Views-Ansatzes stellte sich als wichtige und richtige Entscheidung heraus. Die Repräsentation verschiedener Aspekte des selben Datensatzes in verschiedenen Ansichten hilft dem Anwender bei der Analyse und unterstützt ihn dabei, ein besseres und tiefes Verständnis für diesen zu entwickeln. Durch die Koordination zwischen allen Ansichten können komplexe Anfragen formuliert und anhand der Ergebnisse neue Erkenntnisse gewonnen oder Vermutungen bestätigt oder widerlegt werden. Die Auswahl der Dimensionen beziehungsweise der verschiedenen Ansichten eignet sich gut, um die adressierten Probleme und Fragestellungen zu bewältigen. Allerdings wurde besonders durch die Expertenbefragung deutlich, dass die Benutzung des Ansatzes zwar schnell erlernbar ist, die Komplexität der verschiedenen Ansichten, der unterschiedlichen Visualisierungsformen und der reichhaltigen Interaktionsmöglichkeiten doch auch überfordernd sein und zu Verwirrung führen kann. Dementsprechend muss der Ansatz in der Zukunft dahingehend weiterentwickelt werden, so dass die Fülle an dargestellten Informationen und die Interaktivität gewahrt bleiben, die hohen kognitiven Anforderungen an den Anwender aber verringert werden.

Die Aggregation der Datenobjekte, die Darstellung der religiösen Gruppen mittels Knoten- und Baumstrukturen und die Anordnung dieser mittels eines kräftebasierten Layouts erzeugen gute, übersichtliche und strukturierte visuelle Ergebnisse. Auch die Verwendung von umstrittenen Kuchen-Diagrammen erweist sich in diesem Ansatz als nützlich und hilfreich. Die Entscheidung, Elemente, welche nicht selektierte Daten repräsentieren, nicht einfach zu verstecken, sondern visuell in den Hintergrund zu rücken, stellt sich ebenfalls im Nachhinein als richtig und gut heraus, da damit essentielle Vorteile einhergehen, welche dominanter sind als die Nachteile. Ein Problem der aktuellen Umsetzung ist allerdings die angewandte, primitive Aggregationsprozedur, welche teilweise nicht intuitive Ergebnisse liefert. Diese ist in der Zukunft durch ein klügeres Verfahren zu ersetzen, welches gegebenenfalls nicht nur auf den Daten, sondern auch auf den visuellen Elementen basiert. Die Hervorhebung von visuellen Elementen in anderen Ansichten bei der Interaktion mit Elementen in der Kartenansicht stellt sich als wichtige und unabdingbare Funktion heraus, wobei die Kombination aus Selektion und Hervorhebung aktuell noch verwirrend und irritierend wirken kann. Der Ansatz ist dementsprechend zu modifizieren, sodass diese beiden Funktionen klar und intuitiv voneinander unterschieden werden können.

Die aktuelle Umsetzung der Quellenansicht erfüllt ihren Zweck und visualisiert essentielle Informationen. Diese ist gut geeignet für die übersichtliche Darstellung weniger Quellen. Problematisch wird der angewandte Listenansatz dann, wenn viele Quellen enthalten sind, so dass nicht alle gleichzeitig betrachtet werden können. Hierbei ist ein geeigneter Ansatz zu entwickeln oder zu wählen, um diesen Sachverhalt besser darstellen zu können.

### 6.3 Diskussion

---

Die Gruppenansicht visualisiert intuitiv und übersichtlich hierarchische Daten, welche im Kontext dieser Arbeit von essentieller Bedeutung sind. Die eindeutige Farbzuweisung ist eine gut geeignete Möglichkeit, um die Unterscheidbarkeit der einzelnen Gruppen zu unterstützen. Da die Auswahl von unterscheidbaren Farben mit der wachsenden Anzahl der Gruppen problematisch und schwierig wird, ist der entwickelte Ansatz um weitere Unterscheidungsmerkmale zu erweitern. Dabei könnten zum Beispiel verschiedene Formen oder Symbole verwendet werden. Die Möglichkeit, Religionen, Äste und Gruppen individuell zu selektieren, ist hierbei besonders relevant. Daher sollte die Selektionslogik erweitert werden, so dass diese Elemente in der Hierarchie explizit deseletiert werden können.

Die Verwendung eines gestapelten Säulendiagramms für die Umsetzung der Zeitleiste stellte sich als durchaus geeignet heraus, wobei jede Form der Darstellung zeitlicher Verläufe mit gewissen Vor-und Nachteilen einhergeht. Daher wäre es sinnvoll, dem Anwender die Möglichkeit zu geben, zwischen verschiedenen Darstellungsformen der Zeitleiste wechseln zu können. Eine Verbesserung der gewählten Umsetzung kann erreicht werden, indem Animationen die Veränderungen der Säulen visualisieren, sobald Selektionen in anderen Ansichten geändert werden, anstatt die Säulen einfach vollständig neu zu zeichnen. Besonders durch die Expertenbefragung wurde deutlich, dass die Animation der Zeitleiste von herausragender Bedeutung für den gesamten Ansatz und für die Arbeit damit ist.



# **7 Zusammenfassung und Ausblick**

In diesem Kapitel wird zunächst eine kurze Zusammenfassung der Arbeit präsentiert. Anschließend werden im Rahmen eines Ausblicks einige Punkte angesprochen, die beschreiben, wie der präsentierte Ansatz in der Zukunft erweitert oder modifiziert werden kann.

## **7.1 Zusammenfassung**

Im Rahmen dieser Arbeit sollte ein Ansatz für die interaktive Visualisierung von georeferenzierten, historischen Daten entwickelt werden. Durch die Verwendung digitaler, interaktiver Karten und geeigneter Visualisierungstechniken und durch die Anwendung eines Multiple Coordinated Views-Ansatzes sollte der Anwender bei der Erforschung und Analyse entsprechender Datensätze unterstützt werden.

Dafür wurden zunächst einige, verwandte Arbeiten detailliert betrachtet und deren Stärken sowie Schwächen analysiert. Anschließend wurden wichtige Grundlagen aus den Bereichen der Visualisierung, der Geovisualisierung und aus dem Bereich der Multiple Coordinated Views erarbeitet und zusammengefasst.

Aus den resultierenden Erkenntnissen wurde ein Konzept für den zu entwickelnden Ansatz erstellt und detailliert beschrieben. Der entwickelte Ansatz besteht aus verschiedenen Ansichten, welche jeweils einen Aspekt des zugrundeliegenden Datensatzes visualisieren und Interaktionsmöglichkeiten für die Selektion und Filterung der Daten zur Verfügung stellen. Durch die Verknüpfung der verschiedenen Ansichten können komplexe Anfragen formuliert und Selektionen getätigten werden, wodurch verborgene Eigenschaften des Datensatzes aufgedeckt und der Fokus auf die wichtigen Aspekte der Daten gelegt werden können. Um die übersichtliche Darstellung der Daten zu gewährleisten, werden geeignete Aggregationsverfahren für die Strukturierung und Zusammenfassung der Daten verwendet.

Dieses Konzept wurde anhand einer prototypischen Implementierung umgesetzt und ausführlich beschrieben. Dabei wurde detailliert auf die verschiedenen Ansichten und die Verknüpfung zwischen diesen eingegangen.

Anhand von kurzen, beispielhaften Anwendungsszenarien wurde ange deutet, wie und für welche Aufgaben der entwickelte Ansatz verwendet werden kann und für welche Arten von Problemstellungen er Lösungen bietet. Im Rahmen einer Expertenbefragung wurden die Stärken, Schwächen und mögliche, zukünftige Erweiterungen und Verbesserungen des Ansatzes ermittelt und beschrieben. Die Ergebnisse wurden in einer kurzen Diskussion strukturiert zusammengefasst.

## 7.2 Ausblick

Basierend auf den Einsichten und Erkenntnissen aus der Expertenbefragung und der Diskussion in Kapitel 6 sollen hier mögliche Erweiterungen und Verbesserungen des präsentierten, entwickelten Ansatzes beschrieben werden.

Zunächst ist hierbei zu erwähnen, dass der entwickelte Ansatz bisher lediglich auf zufällig generierte Daten angewandt und mittels dieser evaluiert wurde. Dies muss erneut durchgeführt werden, sobald konkrete, historische Daten vorliegen, um zu ermitteln, wie gut der Ansatz für reale Szenarien funktioniert.

Das Datenmodell im Hintergrund muss erweitert, verändert und ergänzt werden, um auf reale Problemstellungen angewandt werden zu können. So muss es beispielsweise möglich sein, für einzelne Datenobjekte mehrere Quellen anzugeben und zu spezifizieren, wie viele Menschen durch dieses Objekt in Bezug auf Zeitspanne, Stadt und Gruppe beschrieben werden. Außerdem sollen Ansätze entwickelt und eingebaut werden, mittels welcher unsichere Daten angegeben und entsprechend visualisiert werden können.

Aktuell werden visuelle Informationen in allen Ansichten hervorgehoben, wenn innerhalb der Kartenansicht ein Knoten ausgewählt wird. Dieser Ansatz ist zu erweitern, so dass die Hervorhebung auch geschieht, wenn visuelle Elemente in der Quellenansicht, Gruppenansicht oder in der Zeitleistenansicht ausgewählt werden.

Der entwickelte Ansatz funktioniert aktuell lediglich mit den angegebenen religiösen Gruppen und Religionen. Dieser Sachverhalt sollte geändert werden, so dass die hierarchische Struktur der Gruppen und Religionen durch den Benutzer angegeben und modifiziert werden kann.

Die Quellenansicht stellt wichtige Informationen zur Verfügung und funktioniert sehr gut für die Darstellung weniger Quellen. Der angewandte Ansatz ist zu überdenken und gegebenenfalls zu ersetzen, so dass auch viele Quellen übersichtlich und informativ visualisiert werden können.

Für die effiziente Arbeit mit dem Ansatz wären Funktionen hilfreich, mittels welcher innerhalb einzelner Ansichten mehrere Selektionen getätigt werden können, so dass diese in den anderen Ansichten verschieden dargestellt werden. Damit ließen sich mehrere Selektionen direkt miteinander vergleichen, während der Vergleich aktuell lediglich sequentiell möglich ist.

Um den Anwender dabei zu unterstützen, gewonnene Erkenntnisse zu dokumentieren, können Funktionalitäten für die Annotation visueller Elemente in allen Ansichten entwickelt und integriert werden.

Bezüglich der Selektion innerhalb der Kartenansicht ist zu überdenken, ob es weiterhin sinnvoll ist, mittels der gezeichneten Polygone die Daten im Hintergrund, und nicht die visuellen Elemente zu selektieren. Außerdem könnte der Ansatz so erweitert werden, dass vordefinierte Polygone, welche verschiedene geographische Regionen beschreiben, für die Selektion zur Verfügung gestellt werden.

Um die Vor- und Nachteile verschiedener Repräsentationsformen der Zeitleiste zu kompensieren, könnten verschiedene Darstellungen angeboten werden, wobei der Anwender zwischen diesen auswählen und diese jederzeit wechseln können sollte.

## 7.2 Ausblick

---

Die angewandte Aggregationsprozedur funktioniert gut, um übersichtliche visuelle Ergebnisse zu erzielen. Allerdings entstehen dabei Fälle, bei denen nicht intuitiv verständlich ist, warum bestimmte, weit auseinanderliegende Städte zusammengefasst wurden. Daher sollte diese Prozedur verbessert oder komplett ersetzt werden.



# Literaturverzeichnis

- [BC87] R. A. Becker, W. S. Cleveland. Brushing Scatterplots. *Technometrics*, 29(2):127–142, 1987. URL <http://www.jstor.org/stable/1269768>. (Zitiert auf Seite 19)
- [Ber83] J. Bertin. Semiology of graphics: diagrams, networks, maps. 1983. (Zitiert auf den Seiten 7, 14, 15 und 16)
- [BG03] D. Brodbeck, L. Girardin. Design study: using multiple coordinated views to analyze geo-referenced high-dimensional datasets. In *Proceedings International Conference on Coordinated and Multiple Views in Exploratory Visualization - CMV 2003* -, S. 104–111. 2003. doi:10.1109/CMV.2003.1215008. (Zitiert auf den Seiten 7 und 23)
- [BMMS91] A. Buja, J. A. McDonald, J. Michalak, W. Stuetzle. Interactive data visualization using focusing and linking. In *Visualization, 1991. Visualization '91, Proceedings., IEEE Conference on*, S. 156–163, 419. 1991. doi:10.1109/VISUAL.1991.175794. (Zitiert auf Seite 19)
- [CDW<sup>+</sup>16] I. Cho, W. Dou, D. X. Wang, E. Sauda, W. Ribarsky. VAiRoma: A Visual Analytics System for Making Sense of Places, Times, and Events in Roman History. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):210–219, 2016. doi:10.1109/TVCG.2015.2467971. (Zitiert auf den Seiten 7 und 25)
- [CR98] E. H.-H. Chi, J. T. Riedl. An operator interaction framework for visualization systems. In *Proceedings IEEE Symposium on Information Visualization (Cat. No.98TB100258)*, S. 63–70. 1998. doi:10.1109/INFVIS.1998.729560. (Zitiert auf den Seiten 7, 14 und 16)
- [DCCW08] M. Dörk, S. Carpendale, C. Collins, C. Williamson. VisGets: Coordinated Visualizations for Web-based Information Exploration and Discovery. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1205–1212, 2008. doi:10.1109/TVCG.2008.175. (Zitiert auf den Seiten 7 und 27)
- [FD01] M. Friendly, D. J. Denis. Milestones in the history of thematic cartography, statistical graphics, and data visualization. URL <http://www.datavis.ca/milestones>, 32, 2001. (Zitiert auf Seite 13)
- [HM90] R. B. Haber, D. A. McNabb. Visualization idioms: A conceptual model for scientific visualization systems. *Visualization in scientific computing*, 74:93, 1990. (Zitiert auf den Seiten 7 und 18)
- [JHS13] S. Jänicke, C. Heine, G. Scheuermann. GeoTemCo: Comparative visualization of geospatial-temporal data with clutter removal based on dynamic Delaunay triangulations. In *Computer Vision, Imaging and Computer Graphics. Theory and Application*, S. 160–175. Springer, 2013. (Zitiert auf den Seiten 7 und 33)

- [LBWR94] G. L. Lohse, K. Biolsi, N. Walker, H. H. Rueter. A classification of visual representations. *Communications of the ACM*, 37(12):36–50, 1994. (Zitiert auf Seite 19)
- [Mac04] A. MacEachren. *How Maps Work: Representation, Visualization, and Design*. The Guilford Press, 1 Auflage, 2004. URL <http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20&path=ASIN/157230040X>. (Zitiert auf Seite 14)
- [Mun14] T. Munzner. *Visualization analysis and design*. CRC press, 2014. (Zitiert auf Seite 13)
- [Rob00] J. C. Roberts. Multiple view and multiform visualization, 2000. doi:10.1117/12.378894. URL <http://dx.doi.org/10.1117/12.378894>. (Zitiert auf den Seiten 19 und 20)
- [Rob04] J. C. Roberts. Exploratory visualization with multiple linked views. 2004. (Zitiert auf Seite 20)
- [Rob07] J. C. Roberts. State of the Art: Coordinated Multiple Views in Exploratory Visualization. In *Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization (CMV 2007)*, S. 61–71. 2007. doi:10.1109/CMV.2007.20. (Zitiert auf den Seiten 18, 19 und 20)
- [RR09] R. Roth, K. Ross. Extending the Google Maps API for Event Animation Mashups. *Cartographic Perspectives*, 0(64), 2009. URL <http://www.cartographicperspectives.org/index.php/journal/article/view/cp64-roth-ross>. (Zitiert auf den Seiten 7, 29, 30, 31 und 32)
- [S<sup>+</sup>46] S. S. Stevens, et al. On the theory of scales of measurement. 1946. (Zitiert auf Seite 14)
- [SB94] M. Sarkar, M. H. Brown. Graphical Fisheye Views. *Commun. ACM*, 37(12):73–83, 1994. doi:10.1145/198366.198384. URL <http://doi.acm.org/10.1145/198366.198384>. (Zitiert auf Seite 24)
- [Sch08] M. Scherr. Multiple and coordinated views in information visualization. *Trends in Information Visualization*, 38, 2008. (Zitiert auf den Seiten 18 und 19)
- [UFK<sup>+</sup>89] C. Upson, T. A. Faulhaber, D. Kamins, D. Laidlaw, D. Schlegel, J. Vroom, R. Gurwitz, A. van Dam. The application visualization system: a computational environment for scientific visualization. *IEEE Computer Graphics and Applications*, 9(4):30–42, 1989. doi:10.1109/38.31462. (Zitiert auf den Seiten 7 und 18)
- [WBWK00] M. Q. Wang Baldonado, A. Woodruff, A. Kuchinsky. Guidelines for Using Multiple Views in Information Visualization. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, AVI ’00, S. 110–119. ACM, New York, NY, USA, 2000. doi:10.1145/345513.345271. URL <http://doi.acm.org/10.1145/345513.345271>. (Zitiert auf den Seiten 18 und 21)

Alle URLs wurden zuletzt am 26. 9. 2017 geprüft.

### **Erklärung**

Ich versichere, diese Arbeit selbstständig verfasst zu haben. Ich habe keine anderen als die angegebenen Quellen benutzt und alle wörtlich oder sinngemäß aus anderen Werken übernommene Aussagen als solche gekennzeichnet. Weder diese Arbeit noch wesentliche Teile daraus waren bisher Gegenstand eines anderen Prüfungsverfahrens. Ich habe diese Arbeit bisher weder teilweise noch vollständig veröffentlicht. Das elektronische Exemplar stimmt mit allen eingereichten Exemplaren überein.

---

Ort, Datum, Unterschrift