

WandelPAD VIS Projekt 2015

Philip Abernethy*
a1206672

Dominic Böhm †
a1263740

Andreas Fuchs‡
a1106307

Zusammenfassung

Das Projekt WandelPAD ist eine Visualisierung die zur Erforschung der Wanderung von und nach Wien dient. Obwohl Migration ein viel diskutiertes Thema ist und Daten dazu über Open Government für jeden zugänglich sind findet sich nur unzureichende Visualisierung zur Erkundung historischer Daten. Die Entwicklung geschah über mehrere Iterationen angefangen bei mehreren Low Fidelity Prototypen über mehr High Fidelity Prototypen zur Finalen Version. Zur Umsetzung war ursprünglich D3 vorgesehen, aufgrund von Schwierigkeiten wurde aber früh auf Tableau umgestiegen obwohl auch dieses nicht ohne Hürden ist wie in der Entwicklung und in der Benutzerstudie festgestellt wurde.

Keywords: Wanderung, Migration, Wien, Visualisierung

1 Motivation

Wien erfuhr in den letzten Jahren ein sehr starkes Bevölkerungswachstum und gehört zu den am schnellsten wachsenden Metropolen in Europa [1]. Dieser Trend wird laut Prognosen wie zum Beispiel [2] die kommenden Jahre anhalten und stellt die Stadt Wien sowohl vor große Herausforderungen als auch Chancen. Das Projekt WandelPAD soll die Möglichkeit bieten die Bevölkerungswanderung der letzten Jahre innerhalb Wiens zu erforschen. Die Visualisierung soll es ermöglichen Trends sichtbar zu machen und die Wirksamkeit vergangener Maßnahmen zu verifizieren. Die Software soll sowohl Stadtplaner die Bevölkerungsbewegungen beobachten und kontrollierend eingreifen wollen unterstützen als auch zur Erkundung für Journalisten, Soziologen oder jeden anderen mit Interesse dienen.

Die Datengrundlage hierfür liefert der Katalog *Wanderung nach Staatsangehörigkeit - Zählbezirke Wien - Zeitreihe* der von Open Government Wien bereitgestellt wird [3]. Dieser Katalog basiert auf den Meldedaten der Jahre 2007 bis 2013 und unterteilt die Daten nach Jahr, Zählbezirk, Staatsangehörigkeit der wandernden Person, Geschlecht und Ziel beziehungsweise Quelle der Wanderung in den Stufen anderer Zählbezirk, anderes Bundesland, und Ausland. Zählbezirke sind das kleinste von der Statistik Austria erfasste Gebiet und können mit [4] identifiziert werden.

2 Related Work

Angesichts des viel diskutierten Themas ist es wenig verwunderlich dass es schon eine Reihe von Visualisierungen gibt. Diese sind insbesondere in Zeitungen zu finden wie zum Beispiel [5] oder werden von der Stadt Wien selbst zur Verfügung gestellt [6]

*e-mail: a1206672@unet.univie.ac.at

†e-mail: a1263740@unet.univie.ac.at

‡e-mail: a1106307@unet.univie.ac.at

Allerdings sind diese Visualisierungen oft statisch oder bieten nur eine begrenzte Möglichkeit der Interaktion und Exploration. Diese Lücke hoffen wir mit unserem Projekt zu füllen.

3 Herangehensweise

Der natürlichste Weg Migrationsdaten darzustellen war für uns eine Karte. Zwar haben Karten eine vergleichsweise geringe Datendichte, jedoch sind sie die einfachste und lesbarste Art Ortsdaten zu visualisieren und örtlich bezogene Daten in Relation zu setzen.

Da das Datenset letztlich nur Summen über die Zu- und Abwanderung bietet erschien eine farbliche Kodierung anfangs sinnvoll. Es hat sich jedoch gezeigt, dass das allein nicht ausreicht um ein sinnvolles Bild der Daten zu vermitteln. Daher wurde die Darstellung um mehrere Balkendiagramme erweitert.

Schließlich fand, um ein Bild der Trendentwicklung zu geben, auch ein Linienzug für die jeweiligen Nationalitäten seinen Weg in die Visualisierung.

3.1 Low Fidelity Prototyp

Zunächst wurden drei voneinander unabhängige Low Fidelity Prototypen entworfen, deren jeweiligen Vor- und Nachteile besprochen und aufgrund dessen für einen Prototypen entschieden der die Grundlage für die High Fidelity Implementierung bilden soll.

3.1.1 Low Fidelity Prototyp 1

Der erste Low-Fidelity Prototyp, zu sehen in Abbildung 1, besteht aus einer Karte und einem Diagramm wobei dieses abhängig von einer Auswahl gewählt werden kann.

Der erste Low Fidelity Prototype setzt sich aus einer interaktiven Karte und einem sekundären Teil zusammen, welcher durch eine Auswahl eine genauere Begutachtung der Daten ermöglicht. Um Trends leicht erkennen zu können, soll es die Möglichkeit von Small-Multiples geben, welche die Karte über die Jahre hinweg anzeigt.

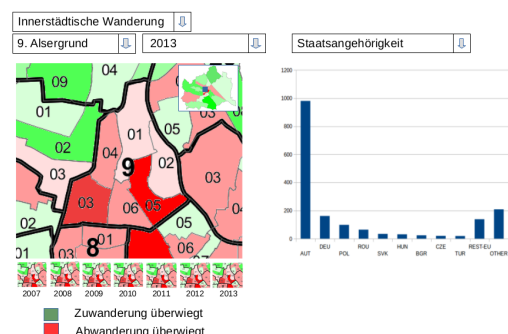


Abbildung 1: Erster Low Fidelity Prototyp

Die Vorteile dieses Prototypen sind dass hier die Trends der einzelnen Zählbezirke alle sofort ersichtlich sind, Einblicke in die historische Entwicklung durch Small-Multiples gegeben ist und dass zusätzliche Details angezeigt werden können. Die offensichtliche Schwäche ist dass die Detailanzeigen nicht miteinander verglichen werden können da eine gleichzeitige Anzeige nebeneinander nicht möglich ist.

3.1.2 Low Fidelity Prototyp 2

Der zweite Low Fidelity Prototyp legt den Fokus auf die Staatsangehörigkeit. Der Prototyp setzt sich aus einem Chord Diagramm der einzelnen Staaten, zwei Balkendiagrammen darunter die jeweils den Zu- und Abzug in die Zählbezirke zeigen sowie zwei Tortendiagramme die das Geschlechterverhältnis mitteilen zusammen. Die Vorteile die-

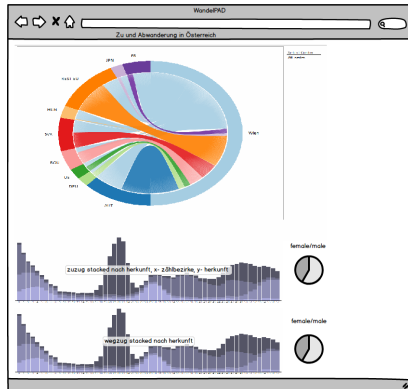


Abbildung 2: Zweiter Low Fidelity Prototyp

ses Prototypen sind dass die Bewegung nach Staatsangehörigkeit hier schnell ersichtlich ist, allerdings ist das Chorddiagramm schwierig zu verstehen und es fehlt an einer Karte um die 250 Zählbezirke auch geografisch zu verorten.

3.1.3 Low Fidelity Prototyp 3

Der dritte Low Fidelity Prototyp zeigt eine Karte Wiens mit ihren Bezirken. Innerhalb dieser Bezirke wird mittels zwei Stackbars jeweils die Zu- und Abwanderung getrennt nach der Herkunft beziehungsweise nach dem Ziel angezeigt. Das Anklicken eines Bezirks sollte diesen heranzoomen und entsprechend die Daten der Zählbezirke sichtbar machen. Die

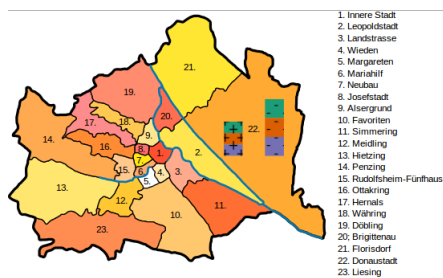


Abbildung 3: Dritter Low Fidelity Prototyp

Vorteile dieses Prototypen sind dass es eine kompakte Anzeige liefert jedoch werden die Zählbezirke und Details nur angezeigt wenn auf einen Bezirk fokussiert wird. Das macht vergleiche zwischen Bezirken schwierig. Ebenfalls unterscheiden sich die Größen der Bezirke erheblich und es ist schwierig

Stack-Bars zu erzeugen die in alle Bezirke passen und dennoch vergleichbar sind.

4 Implementierung

In diesem Abschnitt wird näher auf die Umsetzung der Low Fidelity Prototypen in Tableau [7] eingegangen. Es stellte sich durch Diskussionen und Feedback heraus, dass der Low Fidelity Prototyp 1 aus Abbildung 1, der am ansprechensten und auch am verständlichsten war.

4.1 Erster Prototyp

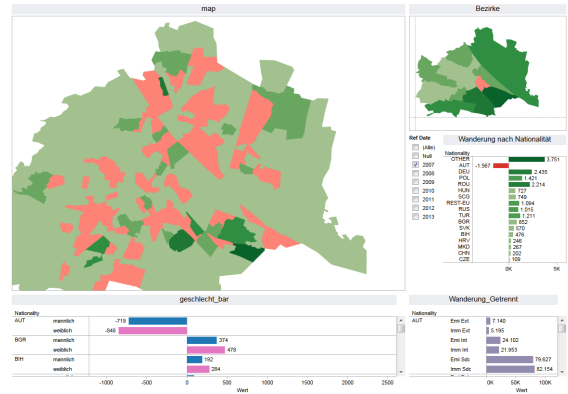


Abbildung 4: Erster High Fidelity Prototyp

Abbildung 4 zeigt den Prototypen verwirklicht mit Tableau, besonderer Wert wurde hier auf die Darstellung der Daten als Karte gelegt, sowie das Ausmerzen der Schwäche, des in Abschnitt 5.1.1 beschriebenen Prototypen, dass nur jeweils eine Detailanzeige zu jedem Zeitpunkt ersichtlich ist. Es ist möglich rechts, in der kleinen Karte von Wien, die einzelnen Bezirke zu sehen und in der großen Karte die Unterteilung nach Zählbezirken. Unter er Bezirkskarte befinden sich die Totalen Ein- bzw. Auswanderungssalden. Die Farbkodierung setzt sich aus grün für Zuzug und Rot für Wegzug zusammen. Im unteren Drittel spalten sich die totalen Salden noch in ihre einzelnen Komponenten auf und eine Aufteilung in männlich und weiblich ist möglich. Ersichtlich wird, dass die Implementation der Small-Multiples nicht in den Prototyp aufgenommen wurden. Dies hatte den Grund, dass Tableau (noch) keine Funktion bereitstellt unsere Zählbezirke als eben solche darzustellen. Die Möglichkeit bestünde wenn eine echte Karte Verwendung fände. Leider gab es zum Zeitpunkt der Implementierung keine geeignete Karte von Wiens Zählbezirken, somit musste der Input durch eine SVG-Konvertierung durchgeführt werden damit die Koordinaten als Polygone dargestellt werden können.

4.2 Zweiter Prototyp

Im zweiten Prototyp wurde auf das Feedback nach der Präsentation des ersten Prototypen eingegangen. Zu sehen ist in Abbildung 5, dass die Karte an Wichtigkeit verloren hat und nur noch eine kleinen Teil der Visualisierung darstellt. Auch wurde die Unterteilung in Bezirke weggelassen und stattdessen verblissen alle Zählbezirke welche sich nicht im selben Bezirk befinden wie die Auswahl. Ein weiterer Punkt war, dass die Grenzen nicht klar ersichtlich waren, deshalb wurden alle Zählbezirke noch umrandet. Über der Karte wurde eine vereinfachte Darstellung der Bezirkssalden mit der Idee der Small-Multiples implementiert, diese sollte über die Jahre im Datensatz die Wanderung nach Bezirken

zeigen. Der wesentliche Fokus wird aber auf die restlichen Daten gelegt, weshalb auch die zwei Diagramme im unteren Drittel an Fläche gewonnen haben.

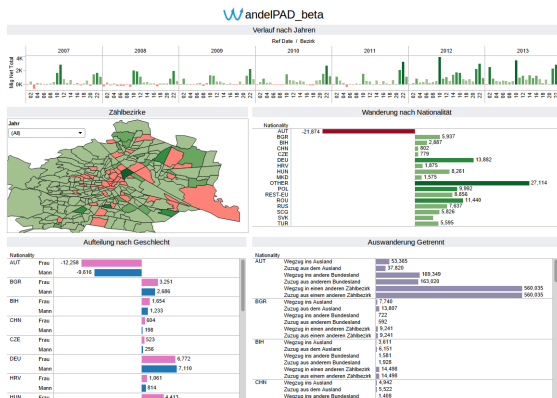


Abbildung 5: Zweiter High Fidelity Prototyp

4.3 Userstudie

Um die Implementierung weiter zu verbessern wurde der zweite High Fidelity Prototyp einer Benutzerstudie unterzogen. Diese wurde in drei Teile untergliedert. In der ersten Phase soll der Nutzer die Software frei erkunden um zu erkennen ob die Visualisierung verständlich ist. In der zweiten Phase soll der Benutzer mehrere Aufgaben lösen. So soll erkannt werden ob es möglich ist Wissen aus der Visualisierung zu extrahieren. In der dritten Phase hat der Benutzer die Möglichkeit seine Meinung über die Visualisierung mitzuteilen.

Die Auswertung der Studie hat mehrere Punkte identifiziert die den Benutzer verwirrt, nicht hilfreich oder gar hinderlich waren von denen die Wichtigsten hier kurz zusammengefasst werden.

Es wurden im Prototyp 2 noch mehrere Kürzel aus dem Katalog verwendet. So wurde etwa der Begriff *Mig Net Total* verwendet. Die User konnten nur selten herleiten das es sich hier um den Gesamtwanderungssaldo handelt. Dieses Problem war bekannt und war somit eine Bestätigung für die geplante Änderung.

Der *Verlauf nach Jahren* bereitete sehr viele Probleme. Das hier Bezirke dargestellt sind es sich sonst aber immer um Zählbezirke handelt hat häufig verwirrt. Viele User fanden es umständlich dass hier nur der Trend der Bezirke abgelesen werden kann, die Trendentwicklung eines Zählbezirks jedoch nur sehr umständlich erkundet werden kann und selbst für die Entwicklung der Bezirke wurde die Darstellung als zu klein empfunden. Ebenfalls kam es in Zusammenhang mit diesem Bereich zu häufigen Bugs. Dieser Bereich musste also nochmals komplett überarbeitet werden.

Auch blieben die Möglichkeiten der Karte komplett unentdeckt. Keiner der Studienteilnehmer fand selbstständig heraus dass die Karte vergrößert- und bewegbar ist oder dass eine Mehrfachauswahl möglich ist.

Aufgrund der Auswertung wurde eine Liste mit potentiellen Verbesserungen erstellt von denen die mit der höchsten Priorität in der finalen Implementierung umgesetzt wurden.

4.4 Finale Implementierung

Die Finale Implementierung zeigt das Ergebnis nachdem die Auswertung der Userstudie durchgeführt wurde und wichtige Änderungen vorgenommen wurden. In Abbildung 6 wird

dieser gezeigt. Wie oben genauer erläutert, erzielten die Small-Multiples nicht den gewünschten Effekt und wurden deshalb wieder verworfen, auf den Wunsch nach einer kontinuierlichen Darstellung der Entwicklung der Wanderung wurde eingegangen, sodass nun das mittlere Drittel eine Liniendiagramm mit dem historischen Trend einnimmt.

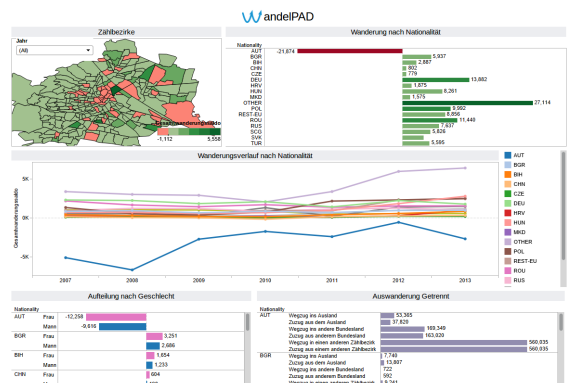


Abbildung 6: Finale Implementierung

5 Results

Hier soll das WandelPAD anhand eines kleinen Beispiels vorgestellt werden.

Bei Start des Programms sehen wir die Visualisierung mit der Auswahl auf ganz Wien, zu sehen in Abbildung 7. Links oben sehen wir den Bereich *Zählbezirk* bestehend aus einer Karte von der der zu betrachtende Zählbezirk gewählt werden kann und das Jahr, hier (*All*) für den gesamten Zeitraum. Rechts daneben sehen wir die *Wanderung nach Nationalität* die den Wanderungssaldo unterteilt nach Staatsangehörigkeit anzeigt. Darunter befindet sich der *Wanderungsverlauf nach Nationalität* der den Trend der gewählten Region in Betracht der Nationalität anzeigt. Links unten befindet sich die *Aufteilung nach Geschlecht* die die unterschiedliche Wanderung der einzelnen Nationalitäten nach Geschlecht anzeigt. Rechts unten befindet sich die *Auswanderung Getrennt*. Hier wird die Wanderung in absoluten Zahlen angezeigt im Gegensatz zu den anderen Bereichen bei denen immer nur die Differenz angezeigt wird. Darüber hinaus ist hier der Zu beziehungsweise Wegzug der Nationalitäten nach Ausgangspunkt beziehungsweise Ziel genauer aufgelistet.

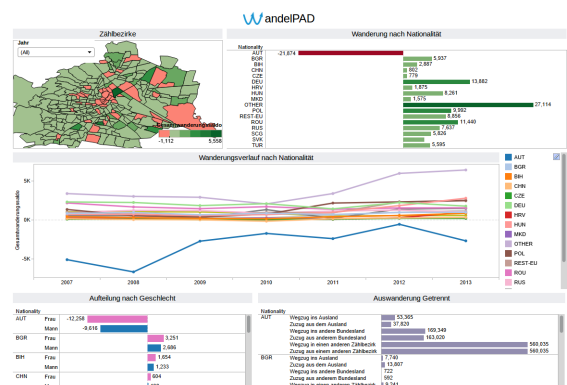


Abbildung 7: Startansicht

In unserem Beispiel interessieren wir uns im Besonde-

ren für den Zählbezirk 9090106 (Universitätsviertel) im 9ten Bezirk. Wir haben dazu etwas die Karte vergrößert und von dort diesen Zählbezirk gewählt, siehe Abbildung 8. Der gewählte Zählbezirk wird fett schwarz Umrandet hervorgehoben und alle Zählbezirke außerhalb des Bezirks werden ausgebleicht. Die Visualisierungen in den anderen Bereichen werden auf diesen Zählbezirk aktualisiert.

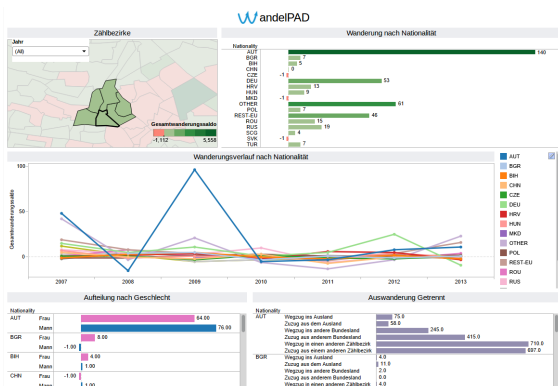


Abbildung 8: Zählbezirk

Wir bemerken im Jahr 2009 einen starke Spitze in der Zuwanderung und wollen dieses Jahr genauer betrachten. Dazu wählen wir im Drop Down Menü bei der Karte das Jahr 2009 worauf hin sich die anderen Bereiche wieder aktualisieren, siehe Abbildung 9. Wir stellen fest dass der positive Zuwanderungssaldo für viele der benachbarten Zählbezirke nicht gilt und sich überwiegend aus Österreichern zusammensetzt.

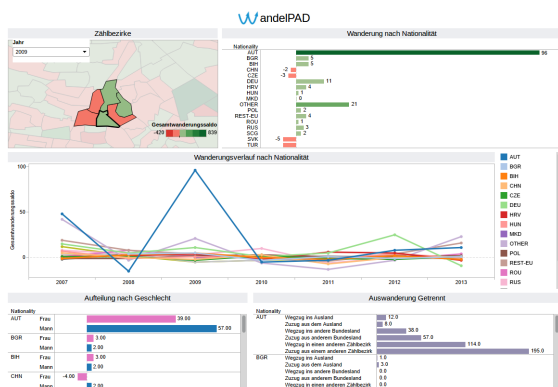


Abbildung 9: Jahr 2009

Wir wollen die Wanderung der Österreicher im Detail betrachten und wählen diese in der Wanderung nach Nationalität. Es werden in allen Bereichen ausgenommen der Karte die Ansicht aktualisiert so dass nur die Daten für Österreich angezeigt werden oder die anderen Daten ausgebleicht werden, siehe Abbildung 10. Wir stellen fest dass dieses Plus überwiegend aus Zuzug aus einem anderen Zählbezirk, also aus Österreichern die davor schon in Wien gelebt haben, entsteht und dass etwas mehr Männer als Frauen in diesen Zählbezirk gezogen sind.

Es kann unter [10] die Software auch selbständig probiert werden.

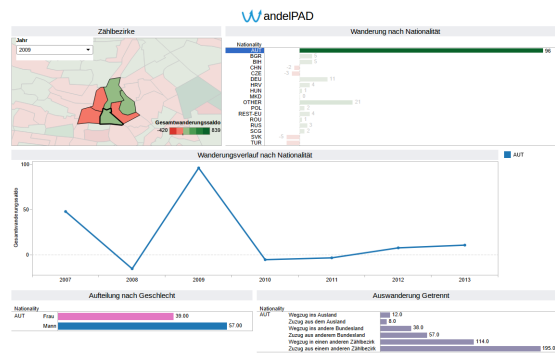


Abbildung 10: Österreich

6 Discussion

Das erste große Problem bestand darin eine geeignete Karte zu finden, welche auch die Zählbezirke beinhaltet. Leider konnte nur auf eine Vektorgrafik zurückgegriffen werden. Mittels python-Script wurden die Daten als Koordinaten abgespeichert und somit für einen einfachen Import vorbereitet. Tableau erwies sich hierbei als einfacher zu bedienen als eine Implementierung in D3 [8], weshalb auch Tableau im Endeffekt als Tool die bessere Wahl gewesen ist. Leider konnten damit aber nicht alle geplanten Interaktion und Diagramme (im Detail Small-Multiples) umgesetzt werden.

Ein weiteres Problem trat mit der Datenkonsistenz auf, da Tableau bei einem Import aus einem CSV-File keine komplexen Join-Operationen zulässt, kam es zu Beginn auch zu einer Immigration von circa 500 Millionen Menschen nach Wien, was in etwa für die Bevölkerung von ganz Europa sprechen würde. Auch für eine finale Veröffentlichung auf Tableau Public[9] musste der verwendete Datensatz kleiner als 10 Millionen Einträge sein. Durch das naive Kreuzprodukt wurden es aber von anfänglich circa 50 Tausend auf 38 Millionen Einträge vergrößert. Durch das Auflösen dieses Knotens mussten leider auch auf Funktionen in der Karte verzichtet werden. Es ist zwar möglich Zählbezirke auszuwählen und somit die ganze Applikation zu manipulieren, jedoch ist es nicht möglich ein Land im Graph auszuwählen und die Salden in der Karte so nach Nationalität zu filtern.

In der Benutzerstudie wurde festgestellt dass Tableau dem Nutzer ein hinderliches oder zumindest nicht optimales User-interface bereit stellt. Die angesprochene unerkannte Interaktionsmöglichkeit mit der Karte stammt daher dass Tableau keine Anhaltspunkte für den User liefert dass die Karte interagirbar ist. Es gibt auch kein Hinweise wie ein gewählter Bereich wieder abgewählt werden kann und je nach Bereich unterscheidet sich dieses Verhalten auch. Ebenfalls fand kein User die hilfreiche Sortierfunktion die nur beim darüber Fahren mit der Maus erscheint. Weiters gibt es bei mancher Auswahl einen Pop Up mit der Möglichkeit die Auswahl aus dem Datenset zu Entfernen (*exclude*) oder das Datenset nur auf diese einzuschränken (*keep only*). Bei Auswahl, die noch dazu manchmal unfreiwillig geschieht da das Pop Up etwas zeitverzögert auftaucht, fanden die Benutzer keine Möglichkeit diesen Schritt wieder rückgängig zu machen. Es scheint dass Tableau eindeutig auf Benutzer ausgelegt ist, die schon mit dieser Software erstellten Visualisierungen vertraut sind.

Eine Software ist erfahrungsgemäß nie komplett und auch WandelPAD könnte in mehreren Bereichen noch weiterentwickelt und verbessert werden. So sind einige Verbesserungsvorschläge aus der Auswertung der Benutzerstudie noch

nicht umgesetzt da sich bisher nur auf jene mit der höchsten Priorität fokussiert wurde. Auch könnte die Visualisierung vermutlich enorm von einer Portierung auf D3 oder ähnliche Frameworks profitieren, da diese deutlich mehr Freiheiten im Bezug auf das Interface und den Datenzugriff bieten.

7 Schlussfolgerung

WandelPAD ist eine Visualisierung die das interaktive Erkunden der Wanderung von, nach und innerhalb Wiens in großer Fülle erlaubt. Es ist dabei anderen Visualisierungen zu diesem Thema weit voraus und hat sich über mehrere Iterationen weiterentwickelt. Die Visualisierung wird unter [10] bereitgestellt.

Literatur

- [1] <http://diepresse.com/home/panorama/wien/3861043/Wien-waechst-schneller-als-erwartet> (Letzter Zugriff: 27. Juni 2015)
- [2] <http://www.wien.gv.at/statistik/bevoelkerung/prognose/> (Letzter Zugriff 28. Juni 2015)
- [3] <https://open.wien.gv.at/site/datensatz/?id=47bdd615-28b4-4c35-9b7c-dba823475552> (Letzter Zugriff: 27. Juni 2015)
- [4] http://www.statistik.at/web_de/services/publikationen/751/index.html?includePage=detailedView&pubId=19§ionName=Regionale+Gliederungen (Letzter Zugriff: 28. Juni 2015)
- [5] <http://kurier.at/chronik/wien/1-741-246-einwohner-wien-nun-zweitgroesste-deutschsprachige-stadt/19.895.417> (Letzter Zugriff: 27. Juni 2015)
- [6] <https://wien2025.wien.gv.at/site/wachstumsprozesse-verandern-die-wiener-gesellschaft/>
- [7] <http://www.tableau.com> (Letzter Zugriff: 27. Juni 2015)
- [8] <http://d3js.org/> (Letzter Zugriff: 27. Juni 2015)
- [9] <https://public.tableau.com/s/> (Letzter Zugriff: 27. Juni 2015)
- [10] <https://public.tableau.com/profile/dominic.b.hm#!/vizhome/WandelPAD/andelPAD> (Letzter Zugriff: 27. Juni 2015)