

Stolperwege

WiSe 2015 / 2016, Praktikumsbericht

Alexander Mehler, Giuseppe Abrami

Inhaltsverzeichnis

1	Praktikumsbeschreibung und Praktikumsziele	1
2	Informatrische Praktikumsziele	1
3	Zur Organisation des Praktikums	1
3.1	AG1 – Ontologische Modelle von Biographien	1
3.2	AG2 – Goldstandards für das maschinelle Lernen biographischer Daten	1
3.3	AG3 – Maschinelles Lernen biographischer Daten	2
3.4	AG4 – 3D-Modelle historischer Gebäude	2
3.5	AG5 – Visualisierung von Lebenswegen	2
3.6	AG6 – iOS-basierte Variante der Stolperwege-App	2
4	Zur Tätigkeit der im Rahmen des Praktikums finanzierten studentischen Hilfskraft	2
5	Ergebnisse und Nachhaltigkeit	2

1. Praktikumsbeschreibung und Praktikumsziele

In dem Praktikum *Stolperwege* haben sich 20 Studierende der Informatik der Goethe-Universität mit der Planung, Realisierung und Weiterentwicklung der gleichnamigen mobilen Applikation beschäftigt. Die *Stolperwege*-App zielt auf die Vermittlung der Geschichte des Holocaust im Kontext der neuen Medien, und zwar am Beispiel der Stadt Frankfurt. Die Zielgruppe der *Stolperwege*-App sind Schülerinnen und Schüler, Bewohnerinnen und Bewohner, Historikerinnen und Historiker und allgemein Besucherinnen und Besucher der Stadt Frankfurt, welche vor Ort den Spuren des Holocaust nachspüren wollen, um sich an der Schreibung einer entsprechenden *Public History of the Holocaust* zu beteiligen.

2. Informatrische Praktikumsziele

Aus technologischer Sicht zielte das Praktikum auf die Vermittlung der Programmiersprache Java, die Erstellung von Android- und iOS-Apps, die Verwendung entsprechender Frameworks, das objektorientierte Programmieren sowie auf die Kooperation innerhalb größerer Pro-

grammierprojekte. Aus der Sicht der Datenmodellierung ging es um die Vermittlung neuester Datenbanktechnologien (insbesondere am Beispiel so genannter Graphdatenbanken und Triple-Stores). Schließlich wurden aus texttechnologischer Sicht neueste Verfahren für die automatische Sprachverarbeitung und die so genannte *Relation Extraction* vermittelt.

3. Zur Organisation des Praktikums

Das Stolperwege-Praktikum wurde in Arbeitsgruppen (AG) unterteilt. Die Aufgabenbeschreibungen der AGs wurden im Rahmen einer anfänglichen Konzeptionierungsphase gemeinsam mit den Studierenden erarbeitet. Im Folgenden werden die Arbeitsgruppen im Einzelnen thematisch dokumentiert.

3.1 AG1 – Ontologische Modelle von Biographien

Die *Stolperwege*-App thematisiert biographische Daten im Kontext des Holocaust. Zur Modellierung solcher Daten sollten die Teilnehmer einschlägige soziologische Literatur (Sackmann 2013) wie auch die *Situationssemantik* (Barwise und Perry 1983) und bereits bestehende ontologische Entwürfe sondieren. Ferner ging es darum, die Datenmodellentwürfe der Gruppe anhand von beispielgebenden biographischen Artikeln aus der Wikipedia zu testen. Es sollte also geprüft werden, inwiefern diese als Folgen von Propositionen zu repräsentierenden sprachlichen Beschreibungen in den entsprechenden Modellentwurf integrierbar waren. Hierzu konnten die Studierenden auf die Arbeiten der AG *Texttechnologie* zurückgreifen und unter anderem das vorhandene, ontologiebasierte Annotationssystem namens *OWLnotator* (Abrami, Mehler und Pravida 2015) nutzen. Als Arbeitswerkzeuge dienten ferner der OWL-Editor namens *protégé* sowie eine MySQL-Datenbank.

3.2 AG2 – Goldstandards für das maschinelle Lernen biographischer Daten

Während AG1 auf die abstrakte Modellierung von Biographien fokussierte, zielte AG2 auf die Annotation von

textuellen Bausteinen solcher Biographien anhand eines eigens erstellten Annotationshandbuchs. Dies betraf in erster Linie die Annotation von Ortsnamen und Personennamen nach dem in (Benikova, Biemann und Reznicek 2014) beschriebenen Format, die Annotation von Zeiten nach dem Format der TimeML (<http://timeml.org/>) und die Annotation von Koreferenzrelationen zwischen Eigennamen, Paraphrasen und Proformen. Zur Bildung des entsprechenden Annotationskorpus wurden biographische Artikel der Wikipedia ausgewählt. Hierzu nutzte AG2 ein Verfahren zur Identifikation von Zielartikeln, das deren thematische Kategorisierung in Wikipedia auswertet. Die eigentliche Annotationsarbeit (im Sinne einer Nachkorrektur) erfolgte auf der Basis von automatisch erstellten Vorverarbeitungsergebnissen, welche die AG Texttechnologie mittels des TextImager bereitstellte. Als Annotationstool wurde WebAnno (Castilho u. a. 2014) eingesetzt und nach den Anforderungen von AG2 angepasst.

3.3 AG3 – Maschinelles Lernen biographischer Daten

Die Daten der AG2 sollten von AG3 dazu herangezogen werden, Verfahren des maschinellen Lernens zu trainieren, welche in der Lage sind, die entsprechenden Daten automatisch zu annotieren. Hierzu wurde eine Bandbreite von Werkzeugen für das Taggen von Eigennamen, Zeitausdrücken und Relationen sondiert, getestet und einer Fehleranalyse unterzogen. Zur Verbesserung der Trainingsergebnisse wurden überdies die Eigennamenliste der Deutschen Nationalbibliothek zu Rate gezogen und insbesondere dem Namenstagger verfügbar gemacht. Die Lernergebnisse fallen je Zielklasse und Umfang entsprechend verfügbarer Trainingsdaten sehr unterschiedlich aus, machen in der Tendenz jedoch die zu 80% korrekte Annotation etwa von Ortsnamen und Ortsnamenbestandteilen erwartbar.

3.4 AG4 – 3D-Modelle historischer Gebäude

Der Wandel unserer Städte macht es notwendig, die teils nicht mehr vorhandenen städtischen Umwelten von Lebenswegen digital zu rekonstruieren, etwas im Hinblick auf zerstörte Gebäude oder nicht mehr vorhandene Wege und Straßen. AG4 hatte entsprechend die Aufgabe, sich mit dem historischen Stadtbild der Stadt Frankfurt und speziell mit zerstörten Gebäuden zu beschäftigen. Hierzu war eine umfangreiche Recherche- und Digitalisierungsarbeit in Zusammenarbeit mit dem Frankfurter Stadtarchiv notwendig. Ziel war es, ein einzelnes Gebäude in Form eines digitalen 3D-Modells (mittels Unity) zu rekonstruieren, dieses in der Stolperwege-App anzuzeigen und entsprechend traversierbar zu machen. Als Beispiel wurde das Palais Rothschild in der Zeil 92-98 ausgewählt.

3.5 AG5 – Visualisierung von Lebenswegen

Um es Nutzerinnen und Nutzern der Stolperwege-App zu ermöglichen, den Lebenswegen von Menschen vor Ort nachzuspüren, bedarf es einer geeigneten Visualisierung, und zwar jenseits der Möglichkeiten, welche etwa Google Maps oder OpenStreetMap bereits bieten. AG5 hatte entsprechend die Aufgabe einen Baukasten für die Vi-

sualisierung von visuellen Komponenten für die Darstellung von Lebenswegen auf der Basis von Google Maps zu entwickeln. Hierzu sollten Bausteine für die Bezeichnung und Unterteilung von Lebenswegen in Abschnitte sowie für die rekursive Eingrenzung von Orten und deren Vernetzung geschaffen werden. Auf diese Weise wird es möglich, Lebensabschnitte zu typisieren, zu beschreiben und graphisch sichtbar zu machen, um ihnen schließlich vor Ort nachzuspüren.

3.6 AG6 – iOS-basierte Variante der Stolperwege-App

AG6 beschäftigte sich mit der Teilnachbildung der Stolperwege-App auf der Basis des Apple iOS-Betriebssystems.

4. Zur Tätigkeit der im Rahmen des Praktikums finanzierten studentischen Hilfskraft

Die im Rahmen des Praktikums finanzierte Hilfskraft, welche mit 40 Stunden im Monat beschäftigt wurde, unterstützte die Teilnehmerinnen und Teilnehmer des Praktikums im Hinblick auf folgende Aufgabenstellungen:

1. Bereitstellung einer RESTful-Basis für die Webservices der Stolperwege-App.
2. Android-Basisprogrammierung als Grundlage für die programmierenden Arbeitsgruppen.
3. Erstellung von Testdaten für die einzelnen Arbeitsgruppen.
4. Technische Unterstützung der Annotationsgruppe in Bezug auf die Nutzung von WebAnno.
5. Implementierung der Klassenrepräsentationen der ontologischen Modelle in Abstimmung mit AG1.

5. Ergebnisse und Nachhaltigkeit

Die Ergebnisse des Praktikums werden über die Website¹ der Studiengruppe *Public History* nachnutzbar gemacht. Zudem wird das Praktikum im Sommersemester 2016 fortgeführt, um die Entwicklungsergebnisse aus dem Wintersemester in einen ersten lauffähigen Prototypen zu integrieren.

Literatur

- Abrami, Giuseppe, Alexander Mehler und Dietmar Pravida (2015). „Fusing Text and Image Data with the Help of the OWLnotator“. In: *Human Interface and the Management of Information. Information and Knowledge Design*. Hrsg. von Sakae Yamamoto. Bd. 9172. LNCS. Springer, S. 261–272. DOI: 10.1007/978-3-319-20612-7_25.
- Barwise, Jon und John Perry (1983). *Situations and Attitudes*. Cambridge: MIT Press.
- Benikova, Darina, Chris Biemann und Marc Reznicek (2014). „NoSta-D Named Entity Annotation for German: Guidelines and Dataset“. In: *Proceedings of LREC*. Reykjavik, Iceland.

¹<http://www.stolperwege.hucompute.org/>

- Castilho, Richard Eckart de u. a. (2014). „WebAnno: a Flexible, Web-based Annotation Tool for CLARIN“. In: *Proceedings of the CLARIN Annual Conference 2014. CAC'14*. Soesterberg, Netherlands.
- Sackmann, Reinhold (2013). *Lebenslaufanalyse und Biografieforschung. Eine Einführung*. Wiesbaden: Springer VS.