

Entwicklung einer Webanwendung zur Annotation spezifischer linguistischer Merkmale in Fließtexten

Masterarbeit Medieninformatik

Oliver Brehm

Eberhard Karls Universität Tübingen
Mathematisch-Naturwissenschaftliche Fakultät
Wilhelm-Schickard-Institut für Informatik

14. Februar 2018

Motivation

- **Lese- Rechtschreibschwäche:** Weit verbreitete Entwicklungsstörung
- Übungstexte mit Hervorhebung von Silben in verschiedenen Farben (Unterstützung bei der Wortdekodierung, Sprachrhythmus)
- **Zielgruppe:** LerntherapeutInnen, Linguisten, Eltern betroffener Kinder
- Webanwendung mit Fokus auf User Experience

Eine Webanwendung zur Textannotation

[Home](#) [Textanalyse](#) [Verifizierung](#)olibrehm@gmail.com

Einstellungen

Betonte Silbe:

Fett ☒

#E84F2F

Unbetonte Silbe:

Benutzen ☒

#2D2FFF

2. Unbetonte Silbe:

Benutzen ☒

#71A2E9

Wort Hintergrundfarbe:

Benutzen ☒

#FFF1DD

Silbe hervorheben:

☒ Vordergrundfarbe☐ Hintergrundfarbe

Schriftgröße:

16

Silbenabstand:

0.2

Wortabstand:

0.4

Zeilenabstand:

2.2

Vorschau

Einem reichen Manne, dem wurde seine Frau krank, und als sie fühlte, daß ihr Ende herankam, rief sie ihr einziges Töchterlein zu sich ans Bett und sprach: † Liebes Kind, Bleibe fromm und gut, so wird dir der liebe Gott immer beistehen, und ich will vom Himmel auf dich herablicken, und will um dich sein. † Darauf tat sie die Augen zu und verschied. Das Mädchen ging jeden Tag hinaus zu dem Grabe der Mutter und weinte, und blieb fromm und gut. Als der Winter kam, deckte der Schnee ein weißes Tüchlein auf das Grab, und als die Sonne im Frühjahr es wieder herabgezogen hatte, nahm sich der Mann eine andere Frau.

DRUCKEN

KOPIEREN (ZWISCHENABLAGE)

Text

Text:

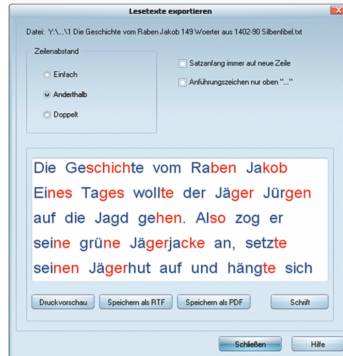
Einem reichen Manne, dem wurde seine Frau krank, und als sie fühlte, dass ihr Ende herankam, rief sie ihr einziges Töchterlein zu sich ans Bett und sprach: †Liebes Kind, bleibe fromm und gut, so wird dir der liebe Gott immer beistehen, und ich will vom Himmel auf dich herablicken, und will um dich sein. † Darauf tat sie die Augen zu und verschied. Das Mädchen ging jeden Tag hinaus zu dem Grabe der Mutter und weinte, und blieb fromm und gut. Als der Winter kam, deckte der Schnee ein weißes Tüchlein auf das Grab, und als die Sonne im Frühjahr es wieder herabgezogen hatte, nahm sich der Mann eine andere Frau.

NEU ANALYSIEREN

Grundlagen und Forschung

Vergleichbare Projekte

■ ABC der Tiere Silbengenerator



- celeco Druckstation
- an Verlage und Produkte gebunden

Linguistische Analyse

- beliebiger Eingabetext
- Parser/Tokenizer (Spacy mit Python)
- Tagging (Part-of-Speech, Lemma)

```
▼ [{annotation: {pos: "det", lemma: "der", hyphenation: "die", stress_pattern: "1", text: "die"},...},...]  
  ► 0: {annotation: {pos: "det", lemma: "der", hyphenation: "die", stress_pattern: "1", text: "die"},...}  
  ▼ 1: {,...}  
    ► annotation: {pos: "noun", lemma: "Wuerde", hyphenation: "Wür-de", stress_pattern: "10", text: "wuerde"}  
      lemma: "Würde"  
      pos: "NOUN"  
      text: "Würde"  
      type: "annotated_word"  
    ► 2: {annotation: {pos: "det", lemma: "der", hyphenation: "des", stress_pattern: "1", text: "des"},...}  
    ► 3: {,...}  
    ► 4: {annotation: {pos: "verb", lemma: "sein", hyphenation: "ist", stress_pattern: "1", text: "ist"},...}  
    ▼ 5: {,...}  
      ► annotation: {pos: "adj", lemma: "unantastbar", hyphenation: "un-an-tast-bar", stress_pattern: "0010",...}  
        lemma: "unantastbar"  
        pos: "ADJ"  
        text: "unantastbar"  
        type: "annotated_word"  
      ► 6: {pos: "PUNCT", lemma: ".", type: "ignored", text: "."}
```

Wortdatenbank

Silbentrennung und Wortbetonung gesucht

- Silbentrennung aus Hyphenator (Pyphen)
- **Wortbetonung:** Lexikon CELEX2 als Goldstandard
- Bestimmung der Betonung aus Phonologischer Darstellung (z.b. *antizipiert* / *&n-ti=i-'pirt*)
- Nicht alle Wörter (mit Flexionsformen) enthalten
- Generieren von Vorschlägen bei unbekannten Wörtern
- Text-to-Speech Systeme (z.B. MARY)

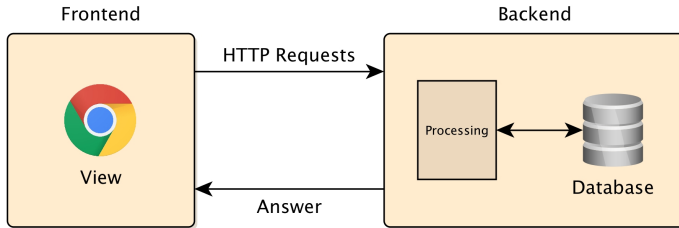
Entwicklung der Anwendung

Anforderungsanalyse

Erarbeitung der Anforderungen in Expertengesprächen:

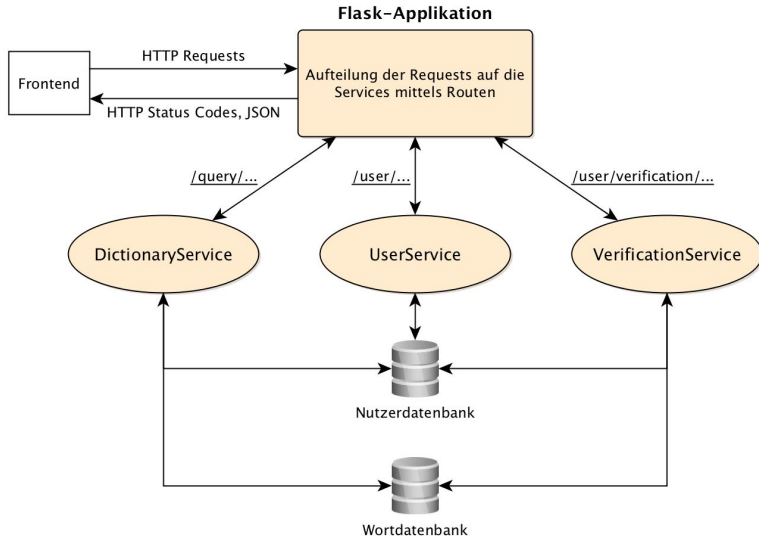
- Verwaltung der Wortdatenbank
- Nutzerverwaltung
- Textanalyse
- Nutzer: Registrierung, Anmeldung, Speicherung von Texten und Einstellungen
- Verifizierung neuer Datenbankeinträge durch NutzerInnen

Aufbau der Anwendung



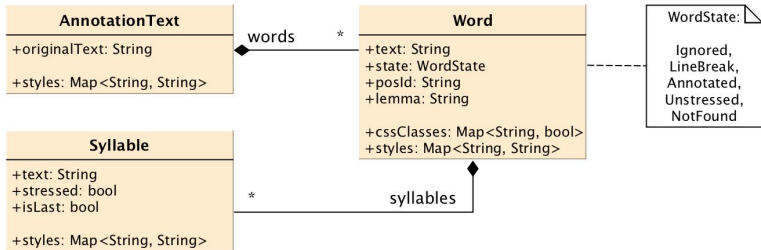
- RESTful API mit Python und Flask
- Wort- und User-Datenbanken (SQLite)
- Alternativen recherchiert (NoSQL, Backend-as-a-Service)

Backend



Frontend

- Single-Page-Application
- AngularDart Framework
- Datenmodell in Dart Klassen
- HTML/CSS Templates mit Data Binding



Evaluation

Ergebnisse der Textannotation

Einem reichen Manne , dem wurde seine Frau krank , und als sie fühlte , dass ihr Ende herankam , rief sie ihr einziges Töchterlein zu sich ans Bett und sprach : " Liebes Kind , bleibe fromm und gut , so wird dir der liebe Gott immer beistehen , und Ich will vom Himmel auf dich herabblicken , und will um dich sein . " Darauf Tat sie die Augen zu und verschied . Das Mädchen ging jeden Tag hinaus zu dem Grabe der Mutter und weinte , und blieb fromm und gut . Als der Winter kam , deckte der Schnee ein weißes Tüchlein auf das Grab , und als die Sonne im Frühjahr es wieder herabgezogen hatte , nahm sich der Mann eine andere Frau .

Ei-nem rei-chen Man-ne , dem wur-de sei-ne Frau krank ,
und als sie fühl-te , dass ihr En-de her-an-kam , rief sie ihr
ein-zi-ges Töch-ter-lein zu sich ans Bett und sprach : "
Lie-bes Kind , blei-be fromm und gut , so wird dir der lie-be
Gott im-mer bei-ste-he-n , und ich will vom Him-mel auf dich
her-ab-blic-ken , und will um dich sein . " Dar-auf tat sie
die Au-gen zu und ver-schied . Das Mäd-chen ging je-den
Tag hin-aus zu dem Gra-be der Mut-ter und wein-te , und
blieb fromm und gut . Als der Win-ter kam , deck-te der
Schnee ein wei-ßes Tüch-lein auf das Grab , und als die
Son-ne im Früh-jahr es wie-der her-ab-ge-zo-gen hat-te ,
nahm sich der Mann ei-ne an-de-re Frau .

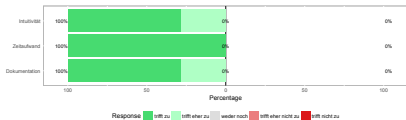
Einfache, abwechselnde
Darstellung

Andere Farben, manuelle
Aktivierung der Betonung

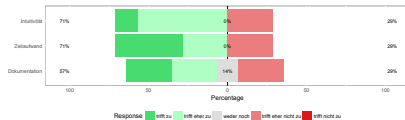
Nutzertest

- Interview mit *Thinking Aloud*
- Zusätzlich *After Scenario Questionnaire* für quantitative Daten
- Pilottest
- Sieben ProbandInnen aus zwei Gruppen

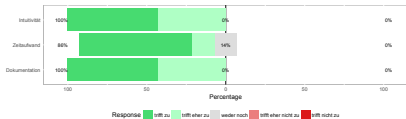
Ergebnisse der Szenarien



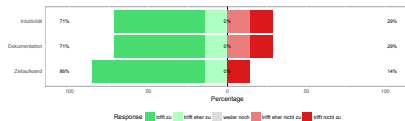
1: Nutzerkonto



3: Annotationsvorlagen



2: Textanalyse



4: Texte wiederverwenden

Fazit

- Automatische Erstellung von Übungstexten, Arbeitserleichterung
- Erweiterbare Wortdatenbank durch Crowdsourcing
- Intuitive Bedienbarkeit
- Erweiterung möglich (Modul für SchülerInnen, mobile App...)

Danke für die Aufmerksamkeit!

→ Fragen und Demo