

# 130 Jahre Schweizer Parlament: Eine Analyse\*

Dr. Laurence Brandenberger, Chair of Systems Design, ETH Zürich

Kollaborierende: Prof. Frank Schweitzer, Sophia Schlosser; SDSC†: Luis Salamanca, Fernando Perez Cruz

## 1 Projektziel

### 1. Aufbau einer flexiblen Datenbank aus 130 Jahren Schweizer Parlament

- ▶ Inkl. aller Reden, Geschäfte und Abstimmungen
- ▶ Nachverfolgung der Änderungen an Gesetzesentwürfen
- ▶ Kommissionsmitgliedschaften und Aktivitäten

### 2. Analyse der Entwicklung des Parlamentes

- ▶ Wie beeinflussen institutionelle Änderungen die Zusammenarbeit im Parlament?
- ▶ Wie und wann ändern ParlamentarierInnen ihre Position?
- ▶ Welche Themen werden wie behandelt?

## 2 Daten

- ▶ **Beratungen** des Schweizer Parlamentes (beide Kammern)
- ▶ **PDF Dokumente** vom Bundesarchiv

#### Amtliches Bulletin

• ≈ 50,000 documents



#### Summary of proceedings

• ≈ 500 documents



#### Additional protocols

• 1921-70, ≈ 30,000 doc.

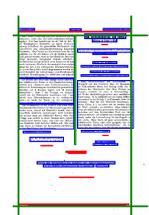


- ▶ Für die Jahre 1995-heute greifen wir auf die **Online Datenbank** des Parlamentes zurück und verbinden sie mit den Daten von 1891-1995

## 3 Vorbearbeitung

- ▶ Vorbearbeitungsschritte der PDF Scans
- ▶ Anpassung der xml-Dokumente

identifying margins



separating columns



ordering text

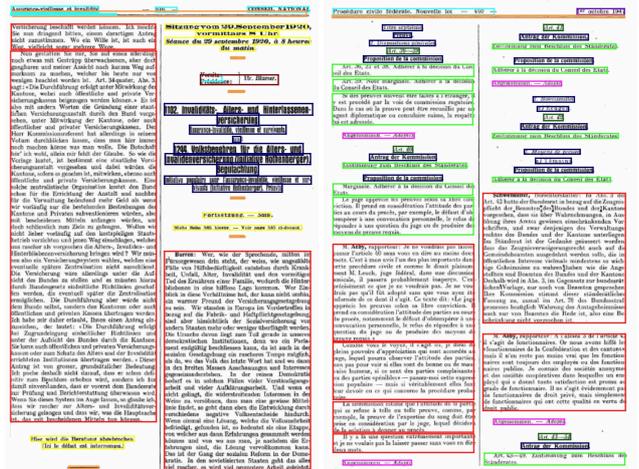


enriched XML file



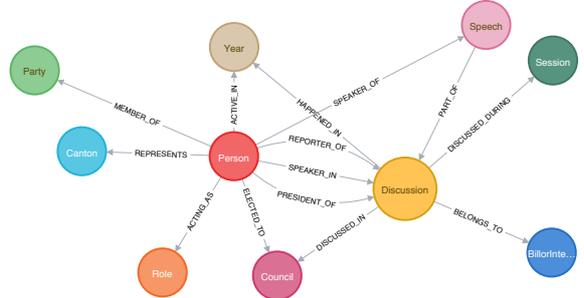
## 4 Extrahierungs-Software

- ▶ Wir entwickeln eine **neue Software** um Textparagrafen in PDFs zu klassifizieren
- ▶ Auseinanderhalten von Titeln (dunkelblau), Reden (rot), Voten (pink) und Anträgen (grün), etc.



## 5 Aufbau einer Relationalen Datenbank

- ▶ Wir verlinken ParlamentarierInnen mit ihren Anträgen und Reden
- ▶ Wir verfolgen **Änderungen an Gesetzesentwürfen**
- ▶ Wir gruppieren Geschäfte nach Themengebiete
- ▶ Wir verlinken Nennungen von ParlamentarierInnen oder Geschäften in Reden
- ▶ **Datenbank:** neo4j
- ▶ Enthält **Entitäten, Attribute, Beziehungen**



## 6 Kontakt und Weitere Informationen

Dieses Projekt besteht aus einer Zusammenarbeit der Professur für Systemgestaltung (ETHZ) mit dem Swiss Data Science Center  
 Weitere Informationen und Updates: <http://u.ethz.ch/DD72W+>

