

Analisi di 130 Anni del Parlamento Svizzero*

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

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

1 Obiettivi del progetto

1. Creare un database relativo agli ultimi 130 anni del Parlamento

- Includere tutti i discorsi, le proposte di legge e le votazioni
- Monitorare le modifiche che sono state fatte alle camere
- Membri e attività della commissione

2. Analizzare l'evoluzione del parlamento svizzero

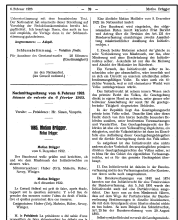
- In che modo i cambiamenti istituzionali influenzano la cooperazione parlamentare?
- Come e perché i membri del parlamento cambiano le loro posizioni?

2 Dati

- **Atti del parlamento svizzero** (entrambe le camere)
- **Documenti PDF**, cortesia dell'Archivio federale

Amtliches Bulletin

■ ≈ 50,000 documents



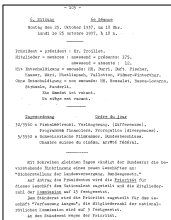
Summary of proceedings

■ ≈ 500 documents



Additional protocols

■ 1921-70, ≈ 30,000 doc.



- Per i dati relativi al periodo dal 1995 fino ad oggi consultiamo il database online del parlamento, collegandoli ai dati estratti dal 1891 al 1995

3 Fasi di pre-elaborazione

- Pre-elaborazione delle scansioni PDF
- Ripulire i file xml e arricchirli con informazioni strutturali

identifying margins



separating columns



ordering text

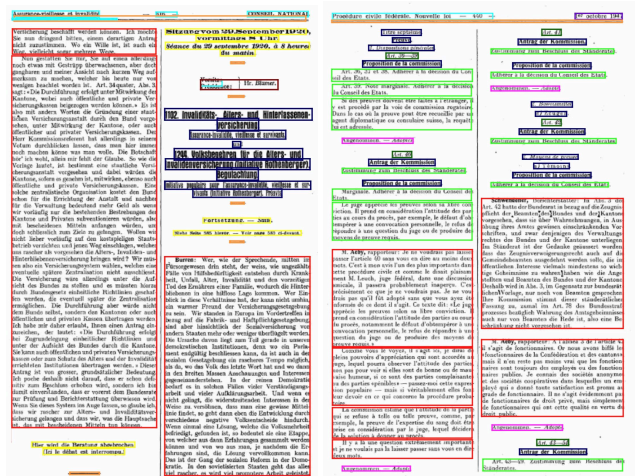


enriched XML file



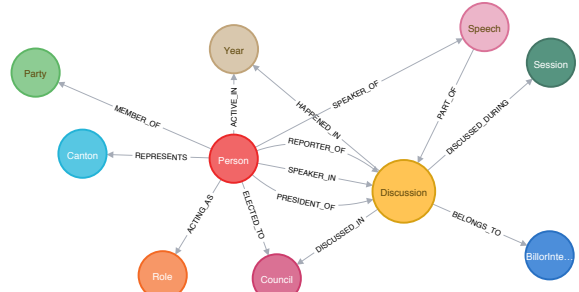
4 Strumento di estrazione PDF

- Abbiamo sviluppato un nuovo strumento per classificare i paragrafi di testo
- Approccio semi-supervisionato
- Distinguere titoli (blu scuro), discorsi (rosso), voti (rosa), proposte di legge (verde), ecc.



5 Costruire un database relazionale

- Collegiamo le proposte di legge ai relativi parlamentari, emendamenti proposti e discorsi
- Monitoriamo le **modifiche** alle proposte di legge e i risultati delle votazioni
- **Raggruppiamo** le proposte di legge per temi politici
- **Database grafico:** neo4j
- Contiene **entità, attributi, relazioni**



6 Contatti ed informazioni

Questo progetto è una collaborazione tra la Chair of Systems Design (ETHZ) e il Swiss Data Science Center (SDSC). Ulteriori informazioni e aggiornamenti: <http://u.ethz.ch/DD72W+>

