

Analyse du Parlement Suisse sur 130 Ans*

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

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

1 Objectifs du Projet

1. Créer une base de données consultable

- Comprend les discours, propositions de loi et votes
- Suivi des changements dans les chambres
- Analyse de la coopération parlementaire
- Analyse des activités des membres et des comités

2. Analyser l'évolution du parlement suisse

- Comment les changements institutionnels affectent-ils la coopération parlementaire?
- Comment et pourquoi les parlementaires changent-ils leurs positions?

2 Les Données

- **Actes** du parlement suisse (des deux chambres)
- **Documents PDF**, avec la permission des Archives fédérales

Amtliches Bulletin

- ≈ 50,000 documents



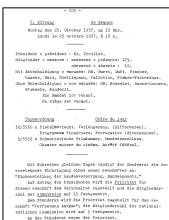
Summary of proceedings

- ≈ 500 documents



Additional protocols

- 1921-70, ≈ 30,000 doc.

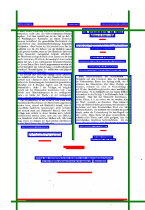


- Pour les données de 1995 à aujourd'hui, nous nous appuyons sur **la base de données en ligne du parlement** et la restructurons pour qu'elle corresponde aux données extraites de 1891 à 1995

3 Étapes Préliminaires

- Prétraitement des scans PDF
- Nettoyer les fichiers XML et les enrichir avec des informations structurales

identifying margins



separating columns



ordering text

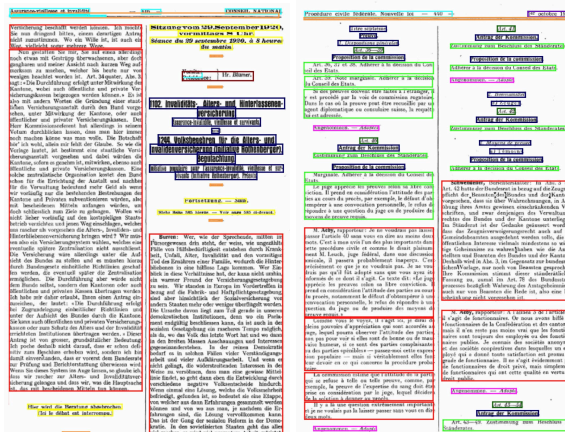


enriched XML file



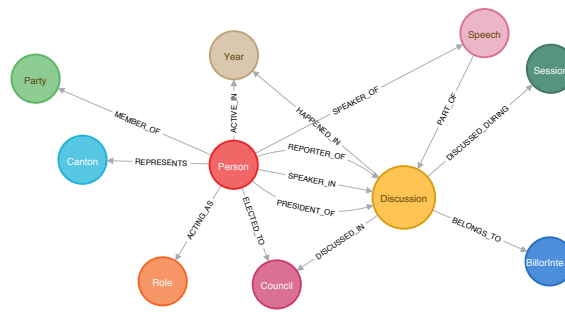
4 Outil d'Extraction

- Nous avons développé **un nouvel outil** pour classer les paragraphes de texte
- Distinguer entre les titres (bleu), les discours (rouge), les votes (rose), les propositions de loi (vert), etc.



5 Construire une Base de Données

- Nous lions les propositions de loi aux auteurs, aux propositions d'amendements et aux discours
- Nous **suivons les modifications** des propositions de loi et les résultats des votes
- Nous **regroupons** les propositions de loi par sujet
- Nous liions les mentions de propositions de loi et de personnes dans les discours
- **Base de données:** neo4j
- Contient des **entités**, des **attributs**, des **relations**



6 Contact et Informations Complémentaires

Ce projet est une collaboration entre la Chair of Systems Design (ETHZ) et le Swiss Data Science Center (SDSC). Informations complémentaires et mises à jour: <http://u.ethz.ch/DD72W>

