

Datenprojekt: eulaw

F. Preu, J. Körtner, F. Gantenbein

18. September 2017

Zielsetzung

- Datenbank des EU-Primärrechts
 - ▶ Vollständige Erfassung des EU-Primärrechts zu jedem möglichen Zeitpunkt
 - ▶ Der Text soll anschließend *maschinenlesbar* sein
 - ▶ Die Erstellung der Datenbank muss *reproduzierbar* sein

Weshalb hilft uns Jan Biesenbenders Arbeit nicht?

1. Einheit in Jan Biesenbenders Arbeit sind „Veränderungen“
2. Verträge können nicht wieder auf Artikelebene gebracht werden
3. Nur Veränderungen zu dem Vertrag von Lissabon werden erfasst
4. Fehlende Reproduzierbarkeit

Grundidee

- Wir speichern den Text aller Verträge
- Wir halten die in den Verträgen festgehaltenen Änderungen fest und führen sie genau wie vorgeschrieben um.
- Somit:
 1. jede Veränderung nachvollziehbar
 2. zu jedem Zeitpunkt gültiges Primärrecht verfügbar
- Voraussetzungen:
 1. Vertragstexte mit Referenznummer die Textstruktur abbildet
 2. Für jeden Änderungsvertrag einen Veränderungskatalog

Texte & Texquelle

- Alle Verträge von `wikisource.org`
 1. Quelle hält (stichprobenartigen) Vergleich mit deutschen, englischen Originaltexten stand
 2. Einfach in eine maschinell lesbare Form zu bringen
- Konkret: scrapen aller Verträge von `wikisource.org` und Nachbearbeiten
- Abbildung der Textstruktur in der Referenznummer
- Beispiel: Euratom-Vertrag

Kodieren der Veränderungen

- Schritte:
 1. Sichtung der Verträge um die unterschiedlichen Anweisungen zu identifizieren
 2. Die Veränderungskataloge erstellen: Beispiel Merger-Vertrag
 3. Veränderungen ausführen

Text ersetzen

- `replace`: einen Artikel mit Referenznummer `id` ersetzen.
- `replace_txt`: in einem Artikel mit Referenznummer `id`, Text finden und ersetzen.
- `replace_txt_globally`: in einem Vertrag mit Referenznummer `id` einen Ausdruck ersetzen.

Text einfügen

- `insert`: neuer Artikel mit Referenznummer `id` zu einem Vertrag hinzufügen
- `insert_txt`: einen bestehenden Artikel mit Text ergänzen
- `add`: neuer, unabhängiger Gesetzestext mit Referenznummer `id` hinzufügen

Weitere

- `repeal`: Artikel aufheben
- `repeal_txt`: Artikel teilweise aufheben
- `renumber`: Artikel neu nummerieren z.B. infolge eines neuen Kapitel

Vorführung der Datenbank und Visualisierung

- Beispiel: Datenbank als Tabelle
- Beispiel: Visualisierung der ECSC Artikel

Was noch gemacht wird

- Alles wird als git zur Verfügung gestellt
- Anleitung zum Arbeiten mit der Datenbank
- Visualisierung