

# Crawling von Datenschutzerklärung-Historien

## Zwischenpräsentation

Alexander Prull, Jörn-Henning Daug, Simon Kaleschke

Universität Leipzig

16. Dezember 2016

- 1 tmcrawl
  - Projektbeschreibung
  - Lösungsansatz
  - Softwarearchitektur

# Gliederung

## 1 tmcrawl

- Projektbeschreibung
- Lösungsansatz
- Softwarearchitektur

# Projektbeschreibung

## Motivation

- Aktuell geltende DSEs analysieren.
- Die Entwicklungsgeschichte von DSEs betrachten.
- Trends und Veränderungen beobachten.

## Aufgaben

- DSEs (täglich) extrahieren.
- Diese geeignet anzeigen.
- Unterschiede über Zeit darstellen.

DSE = Datenschutzerklärung

# Gliederung

## 1 tmcrawl

- Projektbeschreibung
- **Lösungsansatz**
- Softwarearchitektur

# Lösungsansatz

## Aufteilung

- Extraktion: Ruby, XPath, SQLite.
- Backend: Java, REST-Services.
- Frontend: HTML5, CSS3, JavaScript, AngularJS, Bootstrap.

## Arbeitspakete

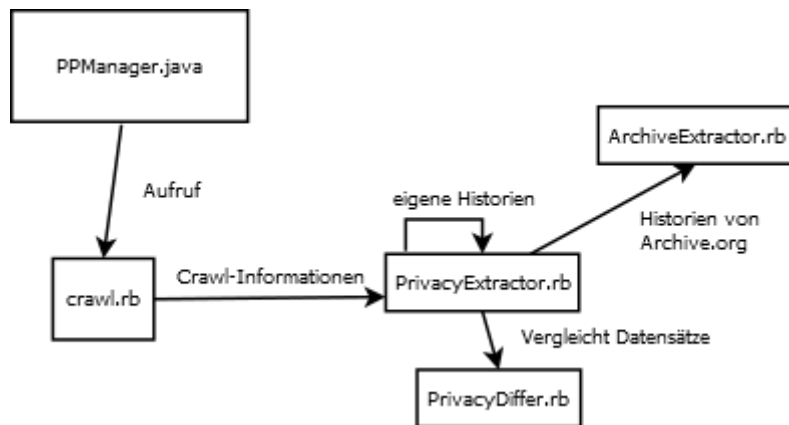
- Recherche (100 %)
- Grundgerüst (80 %)
- Feinschliff (0 %)

# Gliederung

## 1 tmcrawl

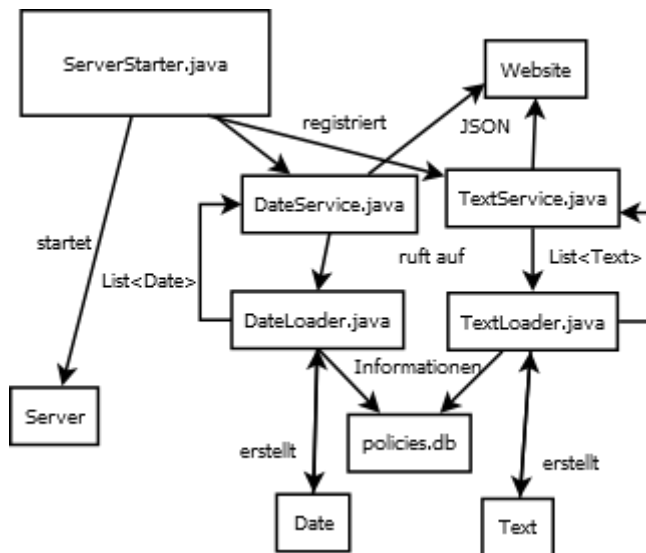
- Projektbeschreibung
- Lösungsansatz
- Softwarearchitektur

# Extraktion





# Backend



# Website

