

# Semantische Datenintegration

## Einleitung

Jakob Voß

Hochschule Hannover

2017-04-01

# Einleitung

# Semantische Datenintegration

- ▶ Problemstellung
  - ▶ Daten zusammenführen
  - ▶ Daten *sinnvoll* zusammenführen

# Beispielsweise Zusammenführung von

- ▶ Bildern/Audio/Video
- ▶ Dokumenten/Texten
- ▶ Datenbanken/Datensätzen

# Grundlagen

- ▶ Verständnis von Daten
  - ▶ Allgemein (was ist möglich)
  - ▶ Fachspezifisch (was ist praktisch)
- ▶ Verständnis von Datenverarbeitung (Sammeln, Konvertieren, Bereinigen . . .)
- ▶ Verständnis von Integration

# Organisatorisches

# Zeit

- ▶ Samstag, 1.4. 9:00 bis 13:00
- ▶ Freitag, 21.4. 14:00 bis 17:30
- ▶ Samstag, 10.6. 9:00 bis 16:30 (mit Mittagspause)

# Dokumente

- ▶ Vortragsfolien im Repository  
<https://github.com/hshdb/MWM-317-02>
- ▶ Interne Dokumente in Moodle  
<https://moodle.hs-hannover.de/course/view.php?id=5367>



# Leistungsnachweise

- ▶ Aufgabe
- ▶ Kurzvortrag (20 Minuten) und verschiedenartige Ausarbeitung
- ▶ Aufgabenbeschreibung und -Verteilung nach dem ersten Theorie-Block

# Personen

*Wer bin ich – und wenn ja, wie viele?*

(Precht 2007)

# Geplanter Ablauf

## 1. April

- ▶ Datenverarbeitung und -modellierung
- ▶ Aufgaben für Prüfungsleistungen
- ▶ Datenintegration
- ▶ Wikidata, eine integrierende Datenbank

## 21. April

- ▶ Identität und Identifizierung
- ▶ Integration von Dokumentformaten
- ▶ Mapping und Matching
- ▶ Übungen mit OpenRefine

## 10. Juni

- ▶ Methoden der Datenintegration
- ▶ Maschinelle Sprachverarbeitung
- ▶ Muster und Paradigma in Daten
- ▶ Kurzvorträge zu den Aufgaben
- ▶ ... *ihre Wünsche?*

# Literaturangaben

Naumann, Felix, und Ulf Leser. 2006. *Informationsintegration: Architekturen und Methoden zur Integration verteilter und heterogener Datenquellen*. Heidelberg: dpunkt-Verlag.

Precht, Richard David. 2007. *Wer bin ich – und wenn ja, wie viele?* Goldmann.