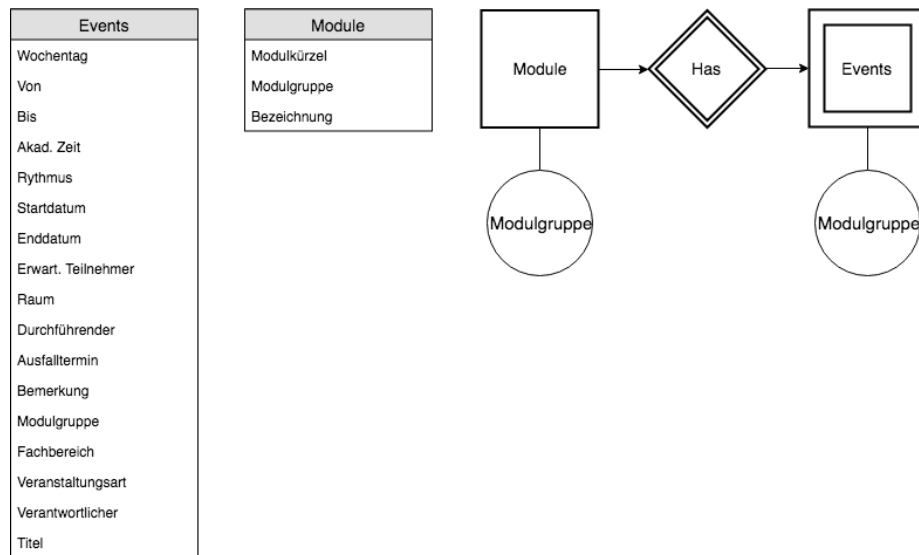


# Marvin Webcrawler

Jakob Eckstein, Markus Flicke

## Datenbankschema



## Programmierung

Wir sind von Python auf Java umgestiegen um das Projekt zuverlässiger warten zu können. Im Folgenden beschreiben wir welche Anforderungen der Webcrawler erfüllen muss.

### Navigation(n)

1. Hauptseite
2. Leere Suche
3. Einträge  $\rightarrow$  300
4. Go to Nth page
5. Event öffnen

6. Zur Suchseite zurückgehen

### **Auslesen(driver)**

1. Grunddaten  $\rightarrow [ < K, V > ]$
2. Eventtabelle  $\rightarrow [Headers], [[Values]]$
3. Module  $\rightarrow [Headers], [[Values]]$
4. Rückgabe als EventData

### **EventData**

Container für Ausgelesene Daten.

### **Hochladen(EventData)**

1. SQL Verbindung
2. EventData zum Hochladen vorbereiten
3. Hochladen