

HTWK Leipzig  
Fachbereich IMN  
Sommersemester 2013

## Beleg im Fach Informationssysteme

Konzeption

Kurt Junghanns, B.Sc.  
Philipp-Rosenthal-Straße 32  
04103 Leipzig  
kurt.junghanns@stud.htwk-leipzig.de

Marcel Kirbst, B.Sc.  
Sieglitz 39  
06618 Molau  
marcel.kirbst@stud.htwk-leipzig.de

7. Mai 2013

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>4</b>
<b>2</b>	<b>Beschreibung der Datenquellen</b>	<b>5</b>
2.1	Facebook-API . . . . .	5
2.2	OpenStreetMap . . . . .	6
2.3	Weather-API . . . . .	7
2.4	Wetterstationen . . . . .	9
<b>3</b>	<b>Architektur des aufzubauenden Data Warehouse</b>	<b>9</b>
<b>4</b>	<b>Datenbankschemata</b>	<b>9</b>
<b>5</b>	<b>Beschreibung der anvisierten Analysen</b>	<b>9</b>
<b>6</b>	<b>Literatur- und Quellenverzeichnis</b>	<b>10</b>

# Abbildungsverzeichnis

1	Facebook Nutzerdaten . . . . .	5
2	OpenStreetMap Orte . . . . .	7
3	Wetterdaten . . . . .	8

# 1 Einleitung

Ziel dieses Belegs ist ein Data Warehouse zu erstellen und die Phasen des Data Warehousing zu durchlaufen. Als Datenquellen dienen dabei das soziale Netzwerk Facebook, Geodaten von OpenStreetMap sowie Wetterdaten.

Ziel ist die Gewinnung neuer Aussagen anhand der Korrelation dieser Daten.

## 2 Beschreibung der Datenquellen

### 2.1 Facebook-API

Facebook bietet für Entwickler APIs zur Abfrage öffentlicher Profildaten an. Mit einem eindeutigen Schlüssel ist es möglich diese Daten per HTTP abzufragen. Die Daten werden im JSON-Format ausgeliefert.

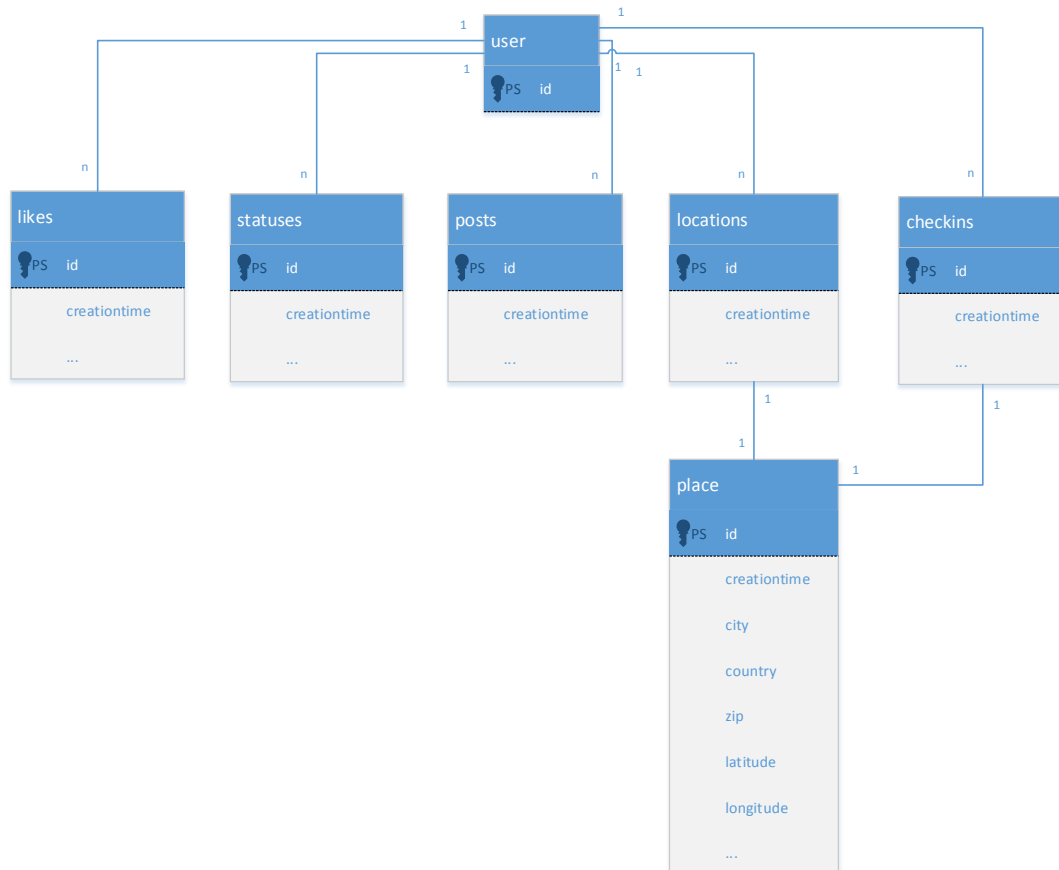


Abbildung 1: Facebook öffentliche Nutzerdaten

Abbildung 1 deutet den Charakter der zu erhaltenden Daten an.

Die aufgeführten Attribute und Entitäten sind für die Auswertung essentiell. Die Entitäten und deren Attribute sind jedoch optional.

Aus diesem Grund sollen im ETL-Prozess nur Datensätze mit vorhandenen Entitäten und Attributen ausgewertet werden.

Zur Abfrage der Nutzer muss eine Zeichenkette zur Filterung mitgegeben werden. Es findet eine Filterung des Nutzernamens statt.

Aus diesem Grund werden die weltweit häufigsten Namen als weitere Datenquelle herangezogen und damit die Nutzerdaten erhoben.

## **2.2 OpenStreetMap**

OpenStreetMap bietet neben umfangreichen Kartenmaterial auch den Service, entsprechend einer übergebenen Auswahl die darin befindeten Orte mit diversen Daten mit HTTP abzurufen.

Mit diesem Service sollen fehlende Informationen von Nutzerdaten von Facebook in Bezug auf Positionen und Orte geliefert werden.

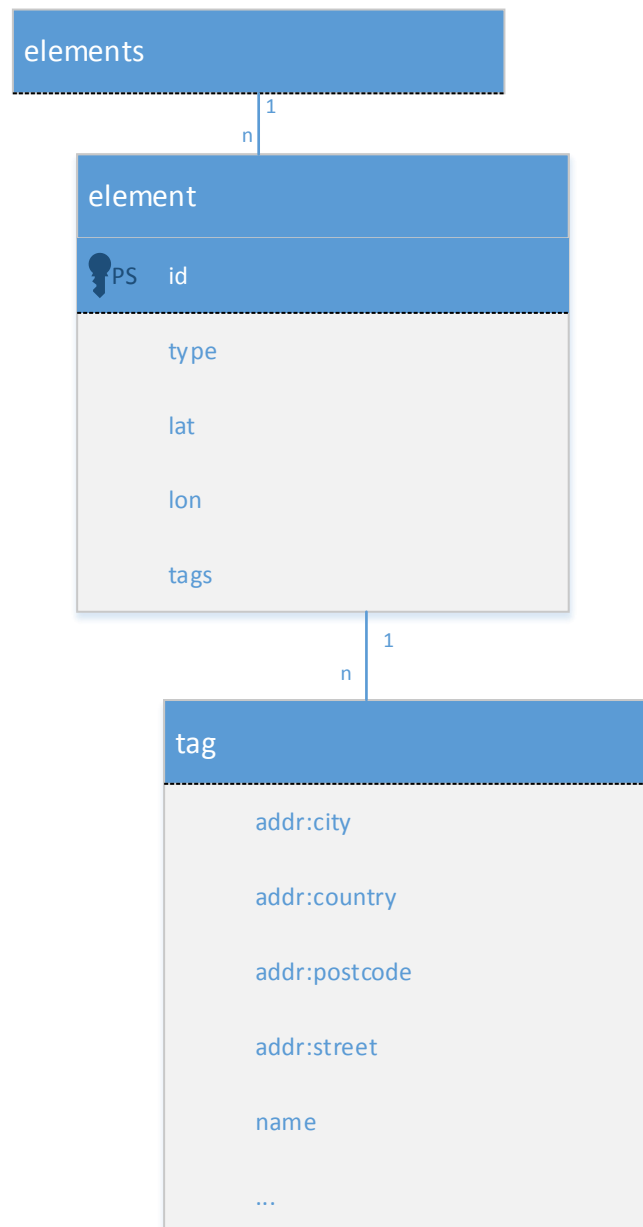


Abbildung 2: OpenStreetMap Orte

## 2.3 Weather-API

Weather Underground bietet angemeldeten Entwicklern per API Zugriff auf aktuelle und vergangene Wetter- und Klimadaten der ganzen Welt an. Auch hierbei werden die Daten mit HTTP abgefragt und im JSON-Format ausgeliefert.

Der Zugriff ist in der kostenlosen Version auf 10 Anfragen pro Minute und 500 Anfragen pro Tag begrenzt.

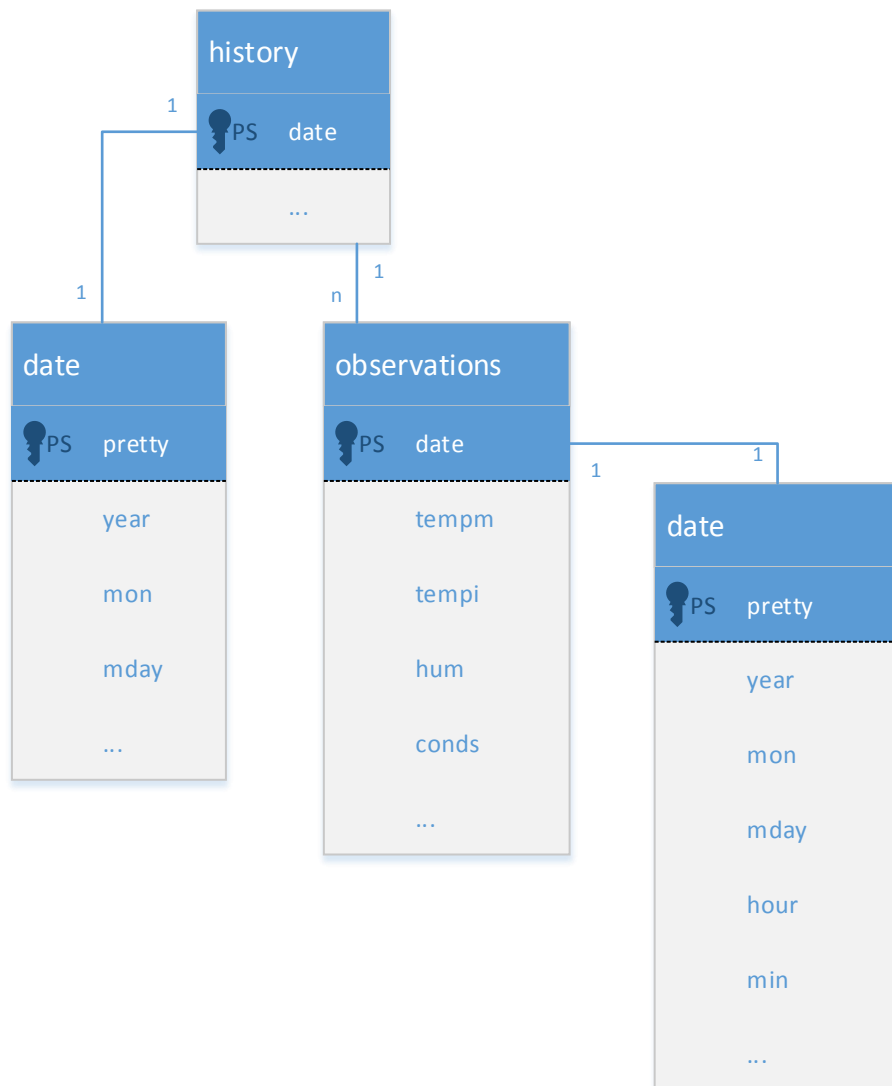


Abbildung 3: historische Wetterdaten von Weather Underground

Das obrige ERM stellt lediglich einen Teil der Granularität und Masse der Daten dar. Die bezeichneten Entitäten und Attribute sind für den ETL Prozess relevant.



2.4 Wetterstationen

3 Architektur des aufzubauenden Data Warehouse

4 Datenbankschemata

5 Beschreibung der anvisierten Analysen

## 6 Literatur- und Quellenverzeichnis

### Literaturverzeichnis

- [1] Kiumars Farkisch: *Data- Warehouse-Systeme kompakt*, Springer Verlag, 2011, ISBN: 978-3-642-21532-2

### Quellenverzeichnis

- [1] <http://www.wunderground.com/weather/api/d/docs>  
Abrufbar am 07.05.2013
- [2] <http://developers.facebook.com/docs/reference/api/>  
Abrufbar am 07.05.2013
- [3] <http://www.pentaho.de/>  
Abrufbar am 25.04.2013
- [4] [http://wiki.openstreetmap.org/wiki/Overpass\\_API/Language\\_Guide](http://wiki.openstreetmap.org/wiki/Overpass_API/Language_Guide)  
Abrufbar am 07.05.2013