

Warning. The length `\marginparwidth` is less than 2cm and will most likely cause issues with the appearance of inserted todonotes. The issue can be solved by adding a line like `\setlength{\marginparwidth}{2cm}` prior to loading the todonotes package.

# Bachelor-Forschungsprojekt Informatik: Relevante OSM-Tags vorschlagen

Marco Hildebrand, XXXX, stXXXX@stud.uni-stuttgart.de

Lukas Baur, 3131138, st141998@stud.uni-stuttgart.de

Felix Bühler, 2973410, st117123@stud.uni-stuttgart.de

July 12, 2018

## **1 Einleitung**

### **1.1 Projektbeschreibung**

## 2 Vorgehensweise

Anschauen von wiki xml dump

unbrauchbare daten, da viel untereinander verlinkt ist.

herunterladen der tags: <https://taginfo.openstreetmap.org/>

## 3 Gettings started

### 3.1 languages

einfach eine liste aller sprachen bekommen mithilfe `taginfo-wiki.db`.

Die kann man von <https://taginfo.openstreetmap.org/download> herunterladen.

### 3.2 export-links

herunterladen der osm-wiki sitemap <https://wiki.openstreetmap.org/sitemap-index-wiki.xml>

davon interessiert uns nur `sitemap-wiki-NS_0-0.xml` der rest enthält daten zu den nutzern, diskussionen und historie

## 4 crawl

alle gesammelten link in die `links.txt` legen

```
scrapy crawl osmWiki -t json -o keys.json
```

### 4.1 pretty json

```
python -m json.tool keys.json > keys-pretty.json
```

## 5 Anhang