# Nutzung von GeoDaten in den Sozialwissenschaften - Das Arbeiten mit OpenStreetMap Daten

Jan-Philipp Kolb

08 April 2016

#### Ausschnitte herunterladen

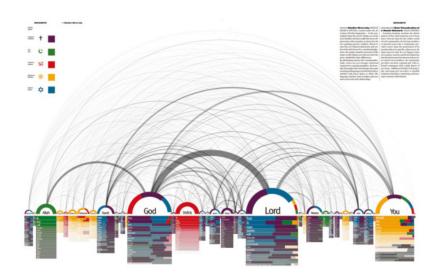
<www.openstreetmap.org/export>



Figure 1: osm export

### Das R-Paket XML - Gaston Sanchez

#### library("XML")



## Wichtiger für mich:

Das Arbeiten mit XML Daten http://gastonsanchez.com/work/webdata/getting\_web\_ data\_r4\_parsing\_xml\_html.pdf

# Getting Data from the Web with R Part 4: Parsing XML/HTML Content

Gaston Sanchez

April-May 2014

Content licensed under CC BY-NC-SA 4.0

Figure 3: Manual von Gaston Sanchez

## Funktionen im XML Paket

Function	Description
$\times$ mlName()	name of the node
xmlSize()	number of subnodes
xmlAttrs()	named character vector of all attributes
xmlGetAttr()	value of a single attribute
xmlValue()	contents of a leaf node
xmIParent()	name of parent node
xmlAncestors()	name of ancestor nodes
getSibling()	siblings to the right or to the left
xmINamespace()	the namespace (if there's one)

## Beispiel: administrative Grenzen Berlin

#### Administrative Grenzen für Deutschland

```
url <- "http://api.openstreetmap.org/api/0.6/
relation/62422"</pre>
```

```
BE <- xmlParse(url)
```

```
- Soun version="0.6" generator="CGinap 0.4.0 (19884 thorn-03.opentreetmap.org)" copyright="OpenStreetMap and contributors" attribution="http://www.openstreetmap.org/copyright="Itemse="http://opensdancommon.org/licenses/odb/1-0">
- Creation ind="0.4222" visible="itemse="36072269" timestamp="2015-12-20T19:49-522" user="thier" uid="278800">
- Smember type="node" ref="240109189" role="admin.centre/>
- Smember type="node" ref="240109189" role="admin.centre/>
- Smember type="node" ref="240109189" role="outer/>
- Smember type="node" ref="74195386" role="outer/>
- Smember type="node" ref="74185568" role="outer/>
- Smember type="node" ref="74185568" role="outer/>
- Smember type="node" ref="94055898" role="outer/>
- Smember type="node" role="94055898" role="0uter/>
- Smember type="node" role="94055898" role="0uter/
```

Figure 4: Administrative Grenzen Berlin

## Das XML analysieren

► Tobi Bosede - Working with XML Data in R

```
xmltop = xmlRoot(BE)
class(xmltop)
## [1] "XMLInternalElementNode" "XMLInternalNode"
## [3] "XMLAbstractNode"
xmlSize(xmltop)
## [1] 1
xmlSize(xmltop[[1]])
## [1] 322
Xpath, the XML Path Language, is a query language for selecting
```

xpathApply(BE,"//tag[@k = 'source:population']")

nodes from an XML document

## Mehr Beispiele, wie man mit XML Daten umgeht:

Daten aus XML extrahieren

http://www.stat.berkeley.edu/~statcur/Workshop2/ Presentations/XML.pdf

 Duncan Temple Lang - A Short Introduction to the XML package for R

http://www.omegahat.org/RSXML/shortIntro.pdf

#### Noch mehr Informationen

Web Daten manipulieren

```
http:
//www.di.fc.ul.pt/~jpn/r/web/index.html#parsing-xml
```

Information zu xquery

http://www.w3schools.com/xquery/

R und das Web (für Anfänger), Teil II: XML und R

```
http://giventhedata.blogspot.de/2012/06/r-and-web-for-beginners-part-ii-xml-in.html
```

String Manipulation

```
http://gastonsanchez.com/Handling_and_Processing_
Strings_in_R.pdf
```

#### Referenzen

#### citation("XML") ## ## To cite package 'XML' in publications use: ## Duncan Temple Lang and the CRAN Team (2016). XML: Too. ## ## Parsing and Generating XML Within R and S-Plus. R pack version 3.98-1.4. https://CRAN.R-project.org/package= ## ## ## A BibTeX entry for LaTeX users is ## ## @Manual{. ## title = {XML: Tools for Parsing and Generating XML \ author = {Duncan Temple Lang and the CRAN Team}, ## ## $vear = \{2016\}.$ note = {R package version 3.98-1.4}, ## url = {https://CRAN.R-project.org/package=XML}, ## ##