

# Scraper Documentation

Web Scraper to request data that is not provided through an API

## Table of content

1. [Scraper Doku](#)
2. [Table Of Content](#)
3. [Getting Started](#)
  - a. [Prerequisites](#)
  - b. [Installing](#)
4. [Overview](#)
5. [Structure](#)
6. [Functionalities](#)
  - a. [Study Rooms](#)
7. [Further Development](#)
8. [Built With](#)
9. [Versioning](#)
10. [Authors](#)

## Getting Started

These instructions will get you a copy of the project up and running on your local machine for development and testing purposes.

### Prerequisites

You will need a current version of [node & npm](#).



### Installing

After cloning the repository, install the dependencies. You can then run the project.

```
# install dependencies
npm install

# serve at localhost:8000
npm start
```

### Overview

The bot is basically a *Node-Express-Backend*. Incoming requests are checked and specifically handled.

## Structure

The bot is separated into two files. `index.js` contains the fundamental logic. At the moment there is only one resource but we expect to expand this with more resources. A resource is represented by a *route*. If the user request the lists of study rooms at our university, the script notices the request at the specified route and prepares a *JSON* response.

The second script, `scrape.js`, takes care of the actual web scraping. The given URL is requested with *axios* and then parsed with *cheerio*.

## Functionalities

### Study Rooms

When the resource is requested we scrape the [ASTA Website](#) and try to return a list of available rooms, that are provided for students, from our university.

## Further Development

Add a new route for every resource in `index.js` and prepare functions in `scrape.js` to scrape the requested data from given Websites.

## Built With

- [Node.js](#)
- [Express.js](#)
- [Axios](#)
- [Cheerio](#)

## Versioning

We use [SemVer](#) for versioning. For the versions available, see the [tags on this repository](#).

## Authors

- **Kai Nessig** - *Initial work* - [GitHub](#)

See also the list of <https://github.com/beuthbot/deconcentrator/graphs/contributors> who participated in this project.

Nutzungshinweis: Auf dieses vorliegende Schulungs- oder Beratungsdokument (ggf.) erlangt der Mandant vertragsgemäß ein nicht ausschließliches, dauerhaftes, unbeschränktes, unwiderrufliches und nicht übertragbares Nutzungsrecht. Eine hierüber hinausgehende, nicht zuvor durch *datenschutz-maximum* bewilligte Nutzung ist verboten und wird urheberrechtlich verfolgt.