**Methods**

* Study region
  + To reduce the computation time needed for the analysis, this research will be focused on European cities with more than 500.000 inhabitants
  + This focus might be extended for future studies
  + ‘European cities’ in this study are based, partly on urban atlas city core layers and partly on ... (Manus layer specifications)
  + To account for the ability of city dwellers to leave their city if they live in proximity to a city border, GBS from a buffer of 1 km surrounding the city core will be included in the study
* Study design
  + Data
    - For the analysis of the walkable environment of European cities, we needed available and comparable data on public green spaces and residential buildings and their respective entry points.
    - Additionally, the analysis requires information on the population living in each residential building and a network that connects the buildings with the green spaces.

We aimed to incorporate publicly accessible data and open source software in order to allow i.) reproduction (e.g. with more recent data), ii.) assessments in data-scare regions, and iii.) comparative approaches covering a larger sample of cities.

* + - We agreed on using Urban Atlas (UA) and OpenStreetMap (OSM) as our main data sources.
    - All analysis was carried out and tested in R 4.1.3 using RStudio version 1.4.1717 and are made available on the GitHub repository www.github.com/blabohm/MA
    - To ensure comparability across all data sources, all data was acquired from the year 2018 where possible
    - Urban Atlas (UA)
      * UA is a land use / land cover (LULC) product from the Copernicus program of the European Union.
      * The 2018 UA version provides Europe-wide comparable data for 788 Functional Urban Areas with more than 50.000 inhabitants.
      * It represents 17 urban and 10 rural LULC classes with a MMU of 0.25 and 1 ha, respectively.
      * In addition to the spatial data, UA contains information on the population for the residential land use classes.
      * Since Copernicus does not offer an API, we downloaded the latest UA version available at the time of this study (v13) by hand.
    - OpenStreetMap (OSM)
      * OSM is a community-based project that provides free geospatial data.
      * The OSM community seeks to create a database of the entire planet that is free and editable.
      * For the analysis we downloaded the OSM data with the identifiers ‘building’ and ‘highway’.
      * We downloaded all OSM data within a 1 km buffer around the city core.
      * Since OSM is a community-based mapping service, the latest version of OSM data is expected to have the most information.
      * We acquired the OSM data via the OSM API, which we accessed via a custom-made tool that heavily relies on the R package ‘osmdata’.
      * For information on the OSM download tool see Appendix ...
  + Data pre-processing
    - Network
      * The network, in this analysis, represents the walkable environment of a city, which connects the entry points of the UGS with those of the residential buildings.
      * To acquire information on the street network of a city, we downloaded OSM data with the identifier ‘highway’, which represents all linestrings in the OSM database that are associated with streets and paths.
      * To secure comparability across European countries, we used all OSM highway classes, except for the class ‘highways’.
      * Linestrings identified with the string ‘highway’ represent motorways which are reserved for motorized use only.
      * To ensure network connectivity and reduce overlap, we cleaned the resulting network following the tutorial on network data pre-processing and cleaning by Lucas van der Meer.
      * Further information on the network pre-processing and cleaning steps can be found in Appendix ...
    - Buildings
      * The following analysis requires information on a cities’ residential buildings, their entry points and on how many persons inhabit each building.
      * We filtered the OSM ‘building’ polygons for residential buildings (include table with excluded words: ‘university’, ‘school’ etc.)
      * We only kept those OSM building polygons that were contained inside of urban atlas residential areas (UA class code starting with: 11).
    - Building entries
      * To detect building entries, we first calculated the centroids of each building.
      * Centroids had to satisfy the constraint, that the point has to lie inside the polygon (see st\_point\_on\_surface).
      * We snapped the centroids to the closest point on the cleaned street network and assumed the resulting points to be the building entries.
    - Population per building
      * The UA dataset provides information on population mostly on a city block level.
      * We disaggregated this data to the building level by distributing the population proportionally to a buildings base area.
      * This workflow follows the assumption, that the building structure, and thus the population per base area inside one city block is similar.
      * For buildings that were contained inside UA residential polygons that erroneously did not have population values, we used the mean population per square-meter of the corresponding UA residential class in the city.
    - Green spaces
      * The last data point required for the analysis is information on publicly accessible UGS and their entry points.
      * We filtered the UA data for the classes ‘green urban areas’ (code 14100) and ‘forests’ (code 31000) to ensure that all green spaces that are used in the analysis are publicly accessible.
      * All green spaces in the UA dataset come with information on area, which we double checked for consistency.
    - Green space entries
      * To reliably detect green space entries, we intersected the outline of the UA green spaces with the cleaned network.
      * We applied different buffer sizes until each green space in a given city had an entry point.
      * We used the resulting points as entry points of the green spaces for the further analysis.
      * Further information on the method and the validation / sensitivity analysis can be found in Appendix ...
    - Network blending
      * In the process of network blending, the entry points of the residential buildings and the UGS (now called ‘nodes’) are being ‘blended’ into the network.
      * During this process, the lines (now called ‘edges’) will be broken at every node location.
      * The node location now represents the new starting / ending points of the newly created edges.
      * For more detailed information on this process see Appendix ...
  + Analysis
    - Analysis to achieve Objective 1:
      * Short summary of Objective 1
      * Our first objective was to develop a consistent workflow that estimates the walkability between green space supply and demand
      * We intended to
      * Modeling service connecting areas / creating workflow for index creation
      * Index building
        + We assumed the distance people will usually walk to their nearest UGS to be 500 m.
        + By making this assumption, we also limited the computational requirements for the analysis.
        + We calculated the Detour Index (DI) and the Local Significance (LS) index for each UGS inside the city core plus a buffer of 1 km.
        + We accounted for the maximum walking distance by calculating both indices for the residential buildings within a network distance of 500 m between a building entry and the nearest entry point of a UGS.
      * Detour Index (DI)
        + The DI is an indicator of barriers in a network.
        + It accounts for the detours people have to take on their way to the nearest UGS.
        + It combines the Euclidean distance, i.e. the direct connection between two points, with the network distance:

Where *DI* is the Detour Index, *d(i, j)* is the Euclidean distance between points *i* and *j*, and *nd(i, j)* is the network distance between the points *i* and *j*.

* + - * + In the case of this analysis, the two points are the entry points of a residential building and the nearest entry point of a UGS.
        + The DI can assume values between 0 and 1.
        + A DI value of 1 represents a straight line between building entry and UGS entry, while a DI value closer to 0 means that the inhabitants of the building have to take a sub-optimal route to the nearest UGS.
        + If one building has access to several UGS within a network distance of 500 m, we decided to use the mean DI value.
      * Local Significance (LS)
        + The LS is an indicator of how many people have access to a UGS.
        + The LS also considers the size of an UGS as well as the distance between people’s homes and the UGS entries:
        + Where *LS* is the Local Significance, *p* is the population of building *i*, *A* is the area of UGS *j* and *nd(i, j)* is the network distance between the entry points of building *i* and UGS *j.*
        + This indicator can assume infinite values.
        + A higher population and area, as well as a lower network distance lead to higher LS values.
        + We attached the LS values to each segment (edge) of the path between building and UGS entries.
        + We summed the LS values of overlapping paths from multiple buildings, leading to higher values on higher frequented edges.
        + A more detailed summary on index building can be found in Appendix …
      * Validation steps
        + Check for percentage / area of UA residential polygons covered by OSM buildings
        + Do the same for network (?)
    - Analysis to achieve Objective 2:
      * Short summary of Objective 2
      * Comparing cities / detecting mismatch
      * European level
        + Comparing distribution of DI and LS values in different cities with histograms / line plots (?)
        + Using LS / m (?)
      * City scenarios
        + To demonstrate application possibilities of the indices, we visualized the DI and LS values for area surrounding the Lene-Voigt-Park (LVP) in Leipzig.

Introduction of the LVP situation

Location: East of Leipzig, close to city center

History: Former train station area, unused since 1942. Since then brownfield. Conversion to park starting in 2000, partly open to public since 2001, fully open since 2004.

Surrounding city structure:

Population structure:

Green spaces in proximity to LVP:

* + - * + To further describe the impact of changes in different model parameters, we tested three different scenarios and calculated the change of the index values to the base model.
        + Scenario 1 – Unlimited access

In the first scenario we demonstrated how the LS and DI indicators change if all barriers obstructing access to the LVP were removed.

To model unlimited access, we distributed hypothetical entry points every 5 meters on the network surrounding the LVP.

* + - * + Scenario 2 – Green space development

In the second scenario we investigated the impact of a development of the green spaces surrounding the LVP to residential buildings.

We assumed the following green spaces in the north of the LVP to be developed to high density residential buildings: Reudnitzer Park, Staphaniplatz, and the green space between Täubchenweg, Perthesstraße and Gerichtsweg.

To implement this scenario, we converted the former green space entry points to building entries.

We multiplied the size of the parks by the 95th percentile of the population per square meter value derived from the urban atlas high density residential class in the surrounding two kilometers.

We distributed the outcome uniformly across the former green space entries.

* + - * + Scenario 3 – Population increase

In the third scenario we modeled a population increase in the residential areas surrounding the LVP.

For each residential building in a distance of 2 km to the LVP, we increased the population value to the 95th percentile of the respective urban atlas residential class.

* + - * + Scenario 4 – Ensemble scenario

In the final scenario we applied the changes from the unlimited access, the green space development and the population increase scenarios and gathered them in one ensemble model.