# PREDICTIONS OF URBAN QUALITIES IN THE CITY OF ZURICH

*Jakub Lichman, Philippe Schlattner*

Department of Computer Science

*Florian Koch*

Department of Physics

ETH Zurich
Zurich, Switzerland
Email: {lichmanj, pschlatt, flkoch}@student.ethz.ch

## ABSTRACT

**Increasing importance of cities together with an increase of their population poses challenges that span beyond technological areas. Since increase of city inhabitants can worsen urban qualities in districts with huge population gain, the cities have to be prepared for monitoring various urban quality features like greenery, public lighting, air quality or pollution. Furthermore if urban qualities in such areas have to be maintained or even improved authorities have to be ready to initiate appropriate actions. The goal of our paper is to obtain, preprocess and visualise selected urban quality features on the map of Zurich. Each feature is visualised on its own visual layer and all layers are finally overlayed such that they show both attractive and unappealing places to live within Zrich. Afterwards we evaluated our results by gathering data about exceptional places by walking them.**

## 1. INTRODUCTION

Huge predicted increase in the number of world urban area residents from 54 to 66 percent in 2050 brings many challenges which every successful city has to overcome. Cities of the future have to ensure continuous improvement of their services despite growing population. Since it is impossible for a city to reach excellence in all services, authorities have to aim for compliance in a prioritised subset of its services. For instance, this year's Mercer's Quality of Living index[1] takes into account economic and political environment, infrastructure, public transport, health, recreation and housing to decide which cities are the most desirable to live in. Based on these factors, the list of most liveable cities was constructed. However, these types of rankings usually generalise the whole city and take average of all the districts which will most certainly overrate some and underrate other districts. In our paper we tried to distinguish urban qualities within a city, concretely in the city of Zurich. Its public data availability enabled us to make interesting predictions about urban qualities of the individual city districts which we lately verified by examining selected places in person.

Our experiment tries to exploit data that is publicly available for the city of Zurich as well as data that we tried to obtain ourselves. The goal was to create a map of the city with multiple layers where each layer will represent one factor of urban quality. By summing all factors and visualising them we can obtain an interactive map which visualises liveability of each part of the city. Later we tried to evaluate selected areas with the Smart Agora platform which enabled us to collect data for validation by walking in these areas.

The main contribution of this paper is to automate reasoning about urban qualities. Nowadays it is done by real estate agents which charge big amounts of money for such services. In the future their jobs can be completely automated and so computers will monitor urban qualities and suggest accommodation prices as well as provide sets of possible actions to improve quality of life in poorly rated locations.

**Related work.** This paper is mainly based on the previous work of Danielle Griego et al. [2] where an internet of things approach is used to obtain the citizen's perception of the selected areas of the city of Zurich. Paolo Neirotti et al. [3] in their paper provide comprehensive understanding of the notion of a smart city through the elaboration of natural resources and energy, transport and mobility, buildings, living, government, economy and people. Kunwar P. Singh et al. [4] used principal components analysis (PCA) to identify sources of pollution and tree based learning models to predict the urban air quality of Lucknow (India) using the air quality and meteorological databases.

This paper is organised as follows: Section 2 explains what a smart city is, its history, tools that are used in smart cities and its practical applications in the praxis. Section 3 describes the data that were used for our experiment and their sources. Section 4 defines the criteria by which we will rate the locations within Zurich. Section 5 explains greenery detection process in our approach and Section 6 describes method of creating the quality index map of the city. In Section 8 we comment on the evaluation of the predictions. Finally, Section 9 concludes this paper and outlines future work.

## 2. BACKGROUND

The term "Smart Cities" has recently attracted [5, 6, 7] some attention. It is usually referring to the real time analysis of whole cities and their population [8]. It can also imply economical innovation which is usually equated with entrepreneurship created by smart people.

In the process of establishing this expression, two main sources of data have been identified [8]. On the one hand this compromises the increasing number of stationary installed computer connected hardware, such as traffic cameras, telecommunication networks and building management systems. On the other hand the population carries an even faster increasing number of mobile computers like smart phones, navigation systems or even fully connected vehicles.

By combining these data sources and analysing them in real-time, one is able to not only steer the flow within a city but also predict the evolution. The Chinese central government has recently gained attraction in the media [9] for not only using the public surveillance system to track down criminals but also equip the police officers with glasses that have an integrated camera [10]. According to the same article, they plan to have a public credit system in place for all its 1.4 billion citizens by 2020, which will be crucial for obtaining jobs or getting permission to travel abroad.

The real time analysis of their currently over 170 million CCTV cameras [11] enabled them to arrest a suspect, who was identified by facial recognition amongst 60'000 visitors to a concert. The combination of various data sources enables the Chinese central government already to monitor public life to a large extent. However, it is only one example of how the consolidation can be used effectively.

It also points out some of the threats. With the public credit and shame system [10] everyone has to constantly adhere to the accepted code of conduct in order to not loose points or be publicly denounced. This ubiquitous surveillance is feared by many privacy policy activists, slowing down similar efforts in western countries. In the end we will have to decide in which kind of world we all want to live in.

However the thus obtained data can also be used for the good. Responsive traffic management systems are a necessity in current megacities, where an increasing number of cars tries to reach their destination on roads that cannot be extended any further. And also the public transport systems reach their limits when an ever increasing number of people tries to get from point A to B all at the same time. Wisely employed camera systems not only allows to track the flow of people but the correct data analysis promises to resolve some of the problems.

In our project we tried to use data from different sources to rate the living quality for people within the districts of Zurich. As most of the real time data in Zurich is not publicly available, we set out to analyse less fast paced data first. The predictions should then be checked with data obtained from study participants, who would walk those areas to collect data with their smart phones. In the future expansion this could allow us to assess the living quality not only in general but also for different days of the week or even different times of the day.

## 3. DATA

To begin with, we evaluated data downloaded from the open data initiative of the city of Zurich [12]. Furthermore, we used maps from the Geographic Information System of the canton of Zurich GISZH [13] to get an idea of potentially interesting areas.

These two sources were used, since we believed that data published by institutions of the government is in general more reliable than that published by individuals. In addition, the data is considered to be of better quality and cover more of the area. This initial trust has been partially disappointed by finding quite a few data sets with only few points. As we wanted to predict the best living areas within Zurich, we also needed data that covers more or less the complete area of Zurich in order to avoid bias towards certain regions, which are better covered by the data sources. Therefore, we omitted data sets with subjectively too few entries, which we initially wanted to use due to their potential. These omitted data sets include the one on air quality, flat prices and traffic counts.

The data we finally used and analysed is composed of records for public street illumination, pedestrian zones, sighing points, zones where driving is prohibited as well as restaurants. These are all sets where the available data is spread across a large part of Zurich and that we figured might be indicative for how much we as students would appreciate living there. This is a highly subjective measure that took us some time to agree amongst ourselves on.

The analysis of the data is described in the Section 6. In addition we inferred the amount of greenery from satellite images as outlined in the Section 5 to evaluate how green the urban areas are. Finally we tried to evaluate the predictions by having people walk some of the areas and gather their impression by utilising Smart Agora. The details of the last bit are outlined in Section 8.

## 4. MEASURE OF ATTRACTIVENESS

As we are all students at ETH Zurich, we wanted to map the attractiveness to live in a certain area from the student point of view. Therefore, access to the public transport is crucial. However walking for up to 15 minutes does not bother us that much. Hence, we did not include stops of the public transportation in our analysis, since they are all distributed well enough and therefore we found a stop in a walking distance to each and every location within Zurich. On the other hand we did not have access to the detailed data on diverse measures. Accordingly we defined spheres of influence for all types of points we used for our analysis, that should represent how far we believe that item will influence how we perceive the area. This also means that accessibility is not that relevant, compared to the bars and restaurants. Another factor that is very important to us is the proximity to places we visit frequently. However, this is even more subjective and even in our small group of people we could not identify several points of interest, that are common to all of us except for study related locations.

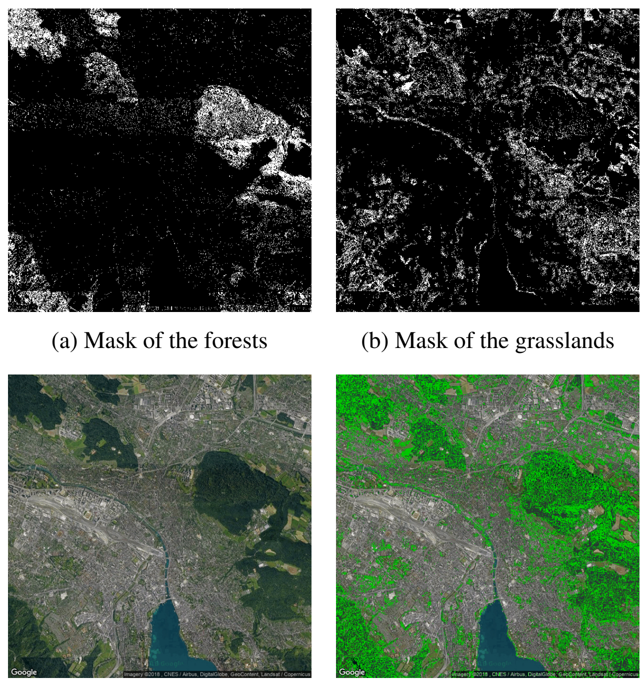
## 5. GREENERY DETECTION

Greenery has always played a crucial part in the construction of cities. The need for green spaces has been present since ancient times. City parks are traditional place of recreation and relax for all kinds of people in their spare time. Current trends in the architecture brought new ways of connecting buildings and greenery such as roofs with green surfaces or terraces with tree pots. These trends are a consequence of the natural human inclination towards nature. Therefore, we have decided to take greenery as a significant indicator of urban quality.

In order to predict user preferences partially based on a greenery, we needed to acquire data about it. There are two basic ways of detecting trees, parks or grass areas on a map. First approach relies on a snapshot of a satellite view [2]. It is simple to implement and yields accurate results. Detection of green areas is via pixels with RGB values that lie within a specified interval. The only crucial requirement in order to have precise detection is to use high resolution satellite images of the selected area which in our case were taken from the Google maps satellite view.

Second approach is to detect greenery from the Google Street View (GSV) [14]. Xiaojiang Li et. al. detected greenery by examining street pictures taken from GSV. With this method they were able to explore greenery inside the cities in great detail. The idea behind the second method is in general more accurate since it can detect trees shadowed by taller buildings or shelters. However, its precision highly depends on the availability of street view and algorithms, which are in our view still not accurate enough. That is the reason why we have decided to use the first method.

We have developed the python script which does greenery detection similar to the first approach mentioned above. Firstly, we did the transformation of RGBA[1] pixel format into HSV[2] one where detection of colour spectrum is more accurate. However, during searching for the accurate green space range we encountered a few issues. Detection of grasslands together with forests was infeasible within a single interval. While grasslands are pixels of a bright green, forests are the ones with a very dark one together with a dark grey in the locations where one tree shadows the other. Therefore, we had to distinguish between them and thus do two separate detections. Combination of them then produced the desired result. Detection of forests, grasslands and their combination can be seen in the Figure 1. Detected locations are marked with the RGB value (0, 200, 0) which is solid green. Furthermore our method computes an intensity matrix which indicates influence of the detected greenery on every place within the city of Zurich.



(a) Mask of the forests (b) Mask of the grasslands

(c) Unmodified Zurich area (d) Greenery in the Zurich area

**Fig. 1**: Visualisation of the greenery detection phases. In the original image we firstly detected forlests then grasslands and finally we combined them and overlay above the original image.

[1]RGBA corresponds to Red, Green, Blue and Alpha, where the first three values lie in the range from 0 to 255 as in usual RGB and $\alpha$ or A gives the opacity in the range from 0=transparent to 1=opaque

[2]HSV corresponds to Hue, Saturation and Value, where Hue is an angle on the colour wheel, Saturation is the intensity of the colour and Value corresponds to the brightness

Our approach firstly produces a binary matrix $B$ of the image $P$ with greenery locations $G$ such that

$$B_{i,j} = \begin{cases} 1 & \text{if } P_{i,j} \in G \\ 0 & \text{if } P_{i,j} \notin G \end{cases}$$

In the second phase we perform so called ”smoothing” of the greenery edges. After the first phase where we have obtained matrix $B$, we now have to compute the influence of the greenery on the pixels without it i.e. the ones with the value 0. The reason why it is needed is because inhabitants do not live in the forests or on the meadows but rather profit from living close to them. Therefore, we have to ”influence” pixels that are in the close neighbourhood with the greenery ones. Our algorithm is iterative and so in every iteration visits every entry in the matrix and checks all neighbouring pixels, finds maxima of them and assigns its half to the current one. The method can be more formally defined as in Algorithm 1.

**Algorithm 1** Smoothing

**procedure** SMOOTHING(iterations, B)
  $S \leftarrow B$
  $i \leftarrow 0$
  **while** $i < 0$ :
    **forall** $p$ **in** $S$ :
      $p := max_{neighbours}(p)/2$
    $i \leftarrow i + 1$

The effect of the smoothing Algorithm 1 after 2 iterations on the Matrix 1 can be seen in the Matrix 2. Remaining 0 entries would be replaced by $0.125$ after the third iteration.

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{1}$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0.25 & 0.25 & 0.25 & 0.25 \\ 0 & 0.25 & 0.5 & 0.5 & 0.5 \\ 0 & 0.25 & 0.5 & 1 & 0.5 \\ 0 & 0.25 & 0.5 & 0.5 & 0.5 \end{pmatrix} \tag{2}$$

## 6. URBAN QUALITY PREDICTION

The core part of this work is the automated detection of living qualities through different parts in Zurich with help of publicly available data. It allows us to create an interactive map which visualizes liveability of each part of the city and lets us reason about the best/worst places to live at according to our perception of attractiveness.

Our algorithm can be divided into the following parts. First of all, we create heat maps of interesting living quality factors for which we could find data in [12]. Interesting means, it matters to our perception of attractiveness, whereas reasonable means we got enough data to reason about the whole area of the city of Zurich. The intensities of the heatmap areas represent the proximity to certain quality factors. High values for close proximity, low values for almost no proximity. Therefore, the brighter an area is, the closer you would be to a quality factor you aim for (e.g. sighting points, pedestrian zones, ...).

In a second step, all these heatmaps are combined into a general heatmap by weighting them and adding them together. In this way we obtain a map, describing exactly where the areas are, where one can get as much quality factors as possible at the same time. At the very end by overlaying this general heat map on a Zurich map, we have a nice visualisation of our detection of liveability.

In the first part, we create heat maps for different living qualities. All the chosen quality fators depend on certain assets of the real world. First, we map the presence of those qualities to GPS positions. Whereas sighting points and pedestrian zones can be directly represented as GPS locations/paths, lightening of the city has been represented as the GPS positions of the lanterns in Zurich. Next, we used Algorithm 2 to determine the presence of a certain quality at a certain position.

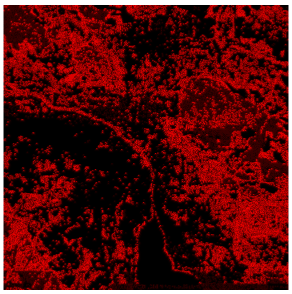
**Algorithm 2** Intensity Finder

**procedure** GET_INTENSITY(gpsPosition)
  *NN* $\leftarrow$ *getNeigherstNeighbor(gpsPosition)*
  *dist* $\leftarrow$ *getDistance(gpsPosition, NN)*
  **if** *dist* $> radius$ :
    **return** *0*
  *fMin* $\leftarrow$ *NN.intensity(atCenter, withRadius)*
  *fMax* $\leftarrow$ *NN.intenstiy(atRadius, withRadius)*
  *f* $\leftarrow$ *NN.intenstiy(distance, withRadius)*
  **return** *normalize(f, fMin, FMax)*

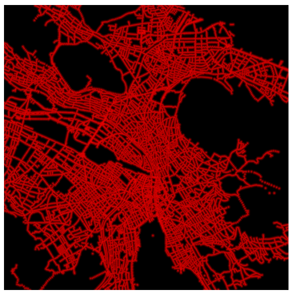
This algorithm can be used in different ways and with many different intensity functions and parameters. As presented in the example Algorithm 2 , it finds the closest point at which your quality is present. Then it determines the influence of this quality toward your present position. This influence is an intentsity value between 1 and 0 of a decreasing function that has a chosen radius and its maximum at its center and goes to zero at the radius border. In this work we use a decreasing exponential function as intensity function.

| Heatmap | Radius [m] |
|---|---:|
| Greenery | 0 |
| Lightning | 70 |
| Driving Prohibited | 70 |
| Pedestrian Zone | 70 |
| Sighting Points | 500 |

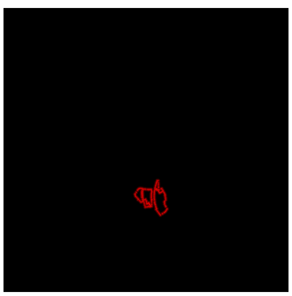
**Table 1**: Choosed radius for different qualities



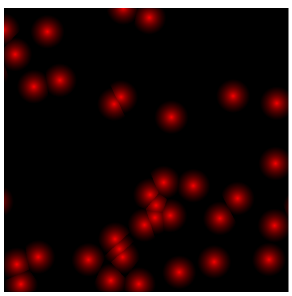
(a) Greenery heatmap, 80% green areas, 20% forests



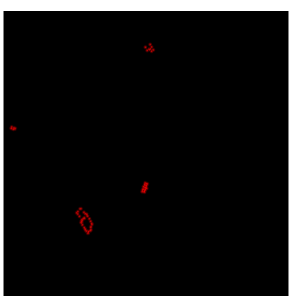
(b) Heatmap of the illumination in Zurich



(c) Heatmap of the pedestrian zones in Zurich



(d) Heatmap of the sighting points in Zurich



(e) Heatmap of driving prohibited zones in Zurich

**Fig. 2**: Visualisation of different intensity heatmaps for different qualities.

Next, the determination of a quality for the whole Zurich map was done in a discrete way. To determine how fine the spacing of the detection positions had to be, we divided the Zurich map by the number of pixels of the Zurich map image we were working with. Then, we mapped every pixel to a GPS position and calculated the influence intensity at this point until getting a heat map of this quality, shown in Figure 2. For having nice colors from black to white we multiplied the influence values going from 0 to 1 by 255 and saved them in the first component of the RGB tuple of each pixel of the image to have a nice red heatmap, shown in Figure 3.

The selection of the radius for each quality was made by ourselves as there is no right choice. We tried to agree on a value so it averages our opinions and gives us some more common and reasonable value. The values for some qualities are shown in Table 1. It is also interesting to notice, that our code also supports other influence functions (quadratic/linear) and another type of influence calculation, that not only considers the closest quality point but all the quality points. In this way a location which is close to a lot of those quality points might get a higher intensity than a point with a single closest quality point. Furthermore this might also seem more appropriate as a method of influence determination at a certain position, but due to very high runtimes of the algorithm, we had to make a compromise with the less accurate method.

**Algorithm 3** Heat Map Creator

```
procedure CREATEHEATMAP(gpsPosition)
    width ← image.width
    height ← image.height
    longitude ← longEast-longWest
    latitude ← latitudeNorth-latitudeSouth
    longStep ← longitude / widht
    latStep ← latitude / height
    for i in range(width)
        for j in range(height)
            x ← longWest + i * longStep
            y ← latNorth - j * latStep
            current ← 255 * get_intensity(x, y)
            result_pixel[i,j]  ← current
```

As all the heatmaps have the same dimensions as the map it is very easy to add them together to a combined heatmap, as shown in Figure 3. Here it was important to weigh them to remain in the 0 to 255 color intensity range to keep a representative information about the influence as we were working with images and not with matrices.

At the very end, the final heatmap was overlayed on the Zurich map to have a nice visualisation of the Zurich map in Figure 4.

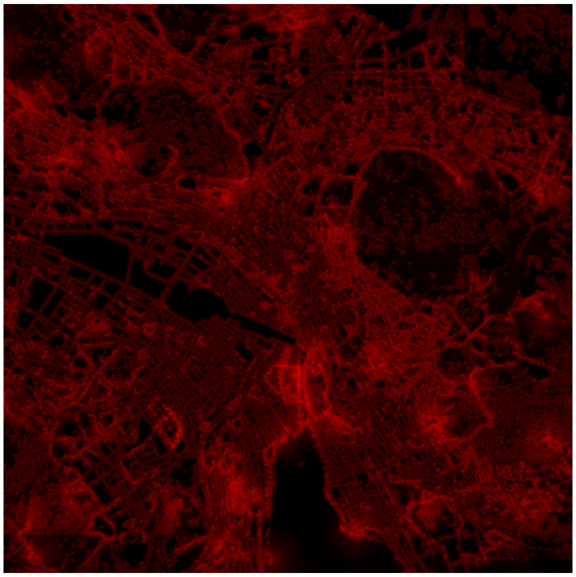


**Fig. 3**: Visualisation of the final combined heatmap.

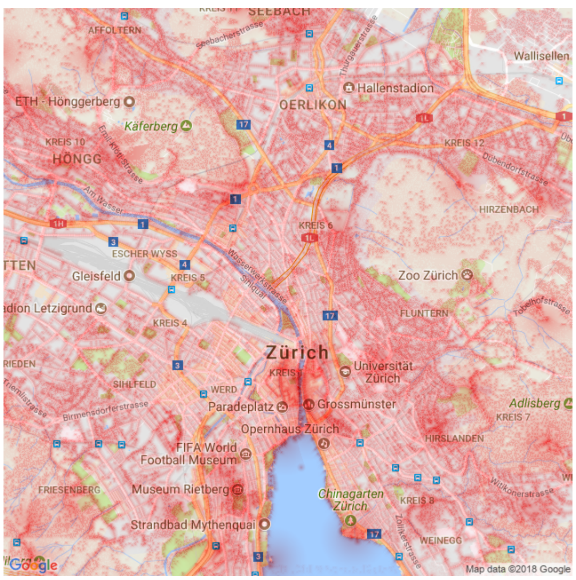



**Fig. 4**: Final visualisation of the heatmap overlayed on the Zurich map.

## 7. DISCUSSION AND VISUALIAZATION OF THE PREDICTIONS

It is interesting to notice, that the best places according to our classification (Figure 5) are also the places we would've considered to be appealing places before starting this project : Hoengg, Bucheggplatz, Central, Bellevue, Seebach, Fluntern and arround the lake (see Figure 6). All these central places offer greenery, a lot of lightening, nice views and close driving prohibited or pedestrian zones.

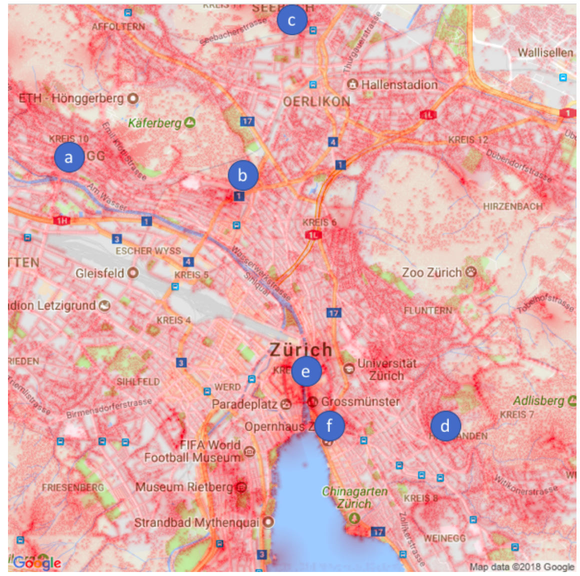



**Fig. 5**: Most appealing places according to our prediction.

(a) Hoengg Center

(b) Near Bucheggplatz

(c) Seebach

(d) Hirslanden

(e) Near Central

(f) Bellevue

**Fig. 6**: Visualisation of different places which were classified as very appealing (source : Google Street View)

Places that weren't classified as very appealing (Figure 7) don't have to be necessarily not appealing in general. They just don't fit our targeted qualities as much as the better classified zones. The lack of datasets of other qualities we would've liked to include (restaurants, bars, detailed public transportation, ...) leads us to results where some worse classified zones might still offer a lot of other qualities we would appreciate. Nevertheless they resepct our quality criterias defined in Section 6, as they often don't include greenery, pedestrian zones, driving prohibited zones nor sighting points (see Figure 8).

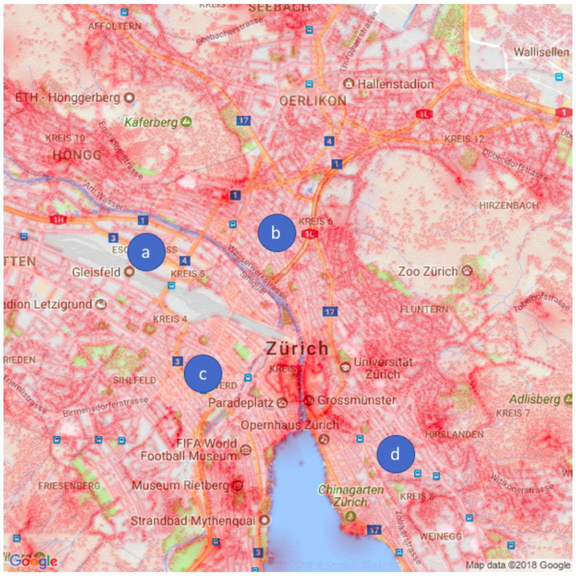



**Fig. 7**: Least appealing places according to our prediction.

(a) Escher Wyss

(b) Unterstrass

(c) Werd

(d) Hegibachplatz

**Fig. 8**: Visualisation of different places which were classified as not very appealing (source : Google Street View)

## 8. PREDICTIONS EVALUATION

For evaluating the predictions we derived in the previous section, we chose to let people rate that areas based on their visual impression after walking through the respective part of the city. For this we initially wanted to use Smart Agora, as it allows to pose the questions at various locations.

Instead we generated a map with markers in various positions and asked people to answer the same questions at those locations. Whereas this does not allow to gather additional information, it is more comfortable to use and easier to collect data from different people. All participants were asked to answer the five questions, where the first two are single choice and the last three are multiple choice questions. The order of the questions was the same for each path but randomly varied between different locations. The questions and answer options, Section 12.1, locations of the questions, Figure 11 and greenery along the walking areas, Figure 10 can be found in the appendix.

Due to the so far very limited amount of data gathered, the evaluation has to be postponed, as conclusions drawn would at most reflect the opinion of individuals. The two areas selected for the detailed analysis are once in the center of Zurich, where we found a very high density of quality factors and the other in Oerlikon, where the quality factors are significantly less dense. However, from personal inspections at those locations, we would attribute similar quality factors for both regions. The perceived differences for various streets in that region show a greater fluctuation, than does the mean between the areas. A more detailed analysis taking into account further areas would be needed.

## 9. DISCUSSION

In this chapter we would like to discuss problems we encountered during our project and outline possible improvements.

**Smart Agora.** In order to use Smart Agora effectively, we would appreciate the following additions to the platform. Firstly, it would be necessary to save parts of a task in case there is a problem with the internet connection so we can resume at a later point. Correcting typos would add the app more professional look. For this a preview of the saved tasks is crucial, preferably in a more user friendly form than a json file. Also the ability to delete erratic questions after the task was saved might be useful, combined with the graciously implemented ability to upload json files directly. Unfortunately the format of the data was changed in between, which required us to find the non documented changes and adopt them inside our json files.

However using the app was even more troublesome. The generated checkbox questions with seven answers would freeze the app on trying to save the selected answers. That is given that the positioning worked correctly [3] and the pop-up

[3] cf. Figure 9 for an erratic localisation

to answer the question actually appeared once one arrived at the designated location. Moreover, we would expect to be able to answer the questions in a random order, if they were set-up in ”Simple” and not in ”Sequence” mode. However, a warning appeared after the third question that some had been skipped and visiting those locations did not allow us to answer these questions.

Occasional crashes of the entire app might be expected in the development phase. However, if non of the entered data is stored locally on the phone, this resets every input, including the researcher ID and potentially answered questions. This complete data loss, which occurs on every quit of the app, renders any tasks with more than a couple of questions extremely dangerous. Even with our short tasks of five questions each, we were not able to complete any of them at the first attempt. This user experience will hinder wide spread participation, which would be required for obtaining the collection of larger data sets.

## 10. CONCLUSION

This paper suggests possible usage of publicly available data for measuring urban qualities within a city. As the results of the evaluation show our predictions were quite accurate. The model and algorithms we developed are able to handle arbitrary number of features as well as records. Therefore further improvements in accuracy could be achieved just with providing more data with better accuracy.

## 11. REFERENCES


[1] “Mercer 2018 Quality of Living Rankings,” online: https://mobilityexchange.mercer.com/Insights/quality-of-living-rankings, 2018, Accessed: 12/May/2018.

[2] Danielle Griego, Varin Buff, Eric Hayoz, Izabela Moise, and Evangelos Pournaras, “Sensing and Mining Urban Qualities in Smart Cities,” in *2017 IEEE 31st International Conference on Advanced Information Networking and Applications (AINA)*. mar 2017, pp. 1004–1011, IEEE.

[3] Paolo Neirotti, Alberto De Marco, Anna Corinna Cagliano, Giulio Mangano, and Francesco Scorrano, “Current trends in Smart City initiatives: Some stylised facts,” *Cities*, vol. 38, pp. 25–36, jun 2014.

[4] Kunwar P. Singh, Shikha Gupta, and Premanjali Rai, “Identifying pollution sources and predicting urban air quality using ensemble learning methods,” *Atmospheric Environment*, vol. 80, pp. 426–437, dec 2013.

[5] Aidin Ferdowsi, Ursula Challita, Walid Saad, and Narayan B. Mandayam, “Robust Deep Reinforcement Learning for Security and Safety in Autonomous Vehicle Systems,” may 2018.

[6] Mohammad Abu-Matar and John Davies, “Data Driven Reference Architecture for Smart City Ecosystems,” may 2018.

[7] Muhammad Usman, Muhammad Rizwan Asghar, Imran Shafique Ansari, Fabrizio Granelli, and Khalid Qaraqe, “Technologies and Solutions for Location-Based Services in Smart Cities: Past, Present, and Future,” *IEEE Access*, vol. 3536, no. c, pp. 1–1, 2018.

[8] Rob Kitchin, “The real-time city? Big data and smart urbanism,” *GeoJournal*, vol. 79, no. 1, pp. 1–14, Feb 2014.

[9] “China’s ever-expanding survaillance state,” online: https://thediplomat.com/2018/04/chinas-ever-expanding-surveillance-state/, 2018, Accessed: 12/May/2018.

[10] “China turns to tech to monitor, shame and rate citizens,” online: https://www.cnet.com/news/china-turns-to-tech-to-monitor-shame-and-rate-citizens/, 2018, Accessed: 12/May/2018.

[11] “Chinese man caught by facial recognition at pop concert,” online: http://www.bbc.com/news/world-asia-china-43751276, 2018, Accessed: 12/May/2018.

[12] “Zurich open data,” online: https://data.stadt-zuerich.ch/, Accessed: 08/May/2018.

[13] “Geographic information system of the canton of Zurich,” online: http://maps.zh.ch/, Accessed: 08/May/2018.

[14] Xiaojiang Li, Chuanrong Zhang, Weidong Li, Robert Ricard, Qingyan Meng, and Weixing Zhang, “Assessing street-level urban greenery using Google Street View and a modified green view index,” *Urban Forestry and Urban Greening*, vol. 14, no. 3, pp. 675–685, 2015.

## 12. APPENDIX

### 12.1. Questions for survey

The following questions were asked to evaluate our predictions. After each question one finds the possible answers. The first two questions are single choice, whereas the other three are multiple choice. The order of questions was the same for each participant, but varied for the different locations.

1. Would you like to live hear if you had the opportunity?
    - certainly yes
    - rather yes
    - rather not
    - certainly not
2. Is it likely, you could currently afford a flat in this area?
    - certainly yes
    - rather yes
    - rather not
    - certainly not
3. What do you like about this place?
    - It's close to transport.
    - It's close to shopping.
    - It's close to other places I often visit.
    - It's calm with lots of green.
    - It's visually appealing.
    - It's bright and open.
    - It's modern and well maintained.
4. What do you dislike about this place?
    - It's far from transport.
    - There are no/insufficient shopping possibilities nearby.
    - It's too far to the places I usually visit.
    - It's very busy and there are no tree nearby.
    - It looks ugly.
    - It's very confined and poorly illuminated.
    - It's old and shabby.
5. What can you see from here?
    - A stop for public transport.
    - Parking facilities.
    - Store for buying food.
    - A restaurant or a bar.
    - Construction site.
    - A park.
    - A nice flat that I could live in.

### 12.2. Code

All data used, the code developed for the analysis and further resources can be found in the following GitHub repository: `https://github.com/limo1996/ETH-DataScience`

### 12.3. Images

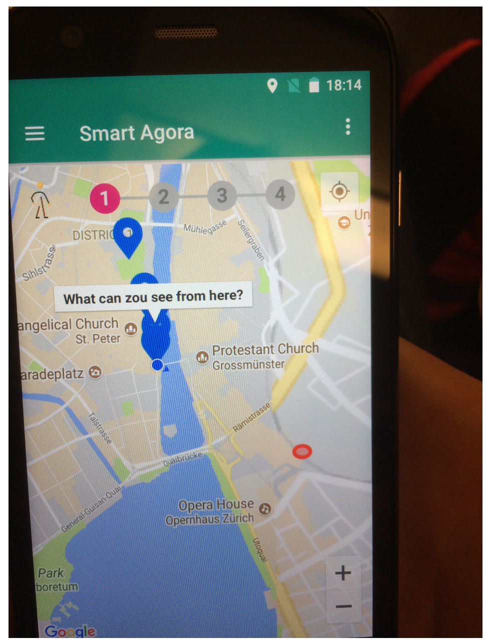



**Fig. 9**: Smart Agora app with location indicated by the blue dot with arrow and the sphere for answering questions shown in red.

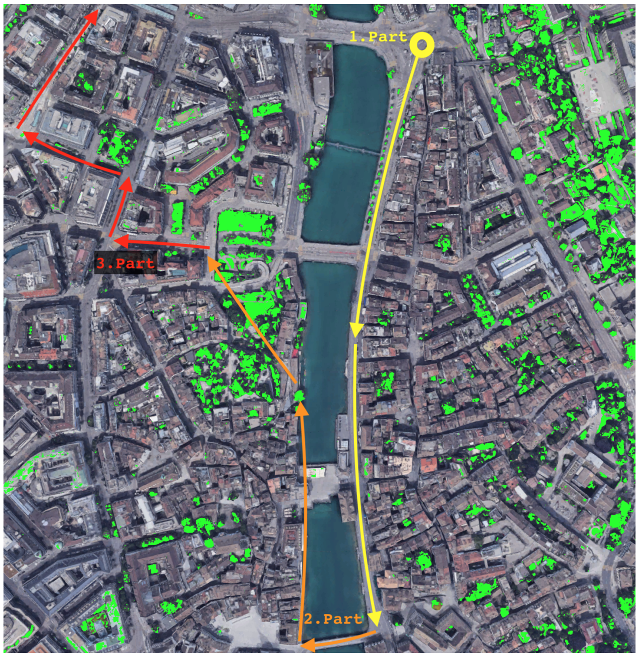



**Fig. 10**: Greenery detection in the area of our path.

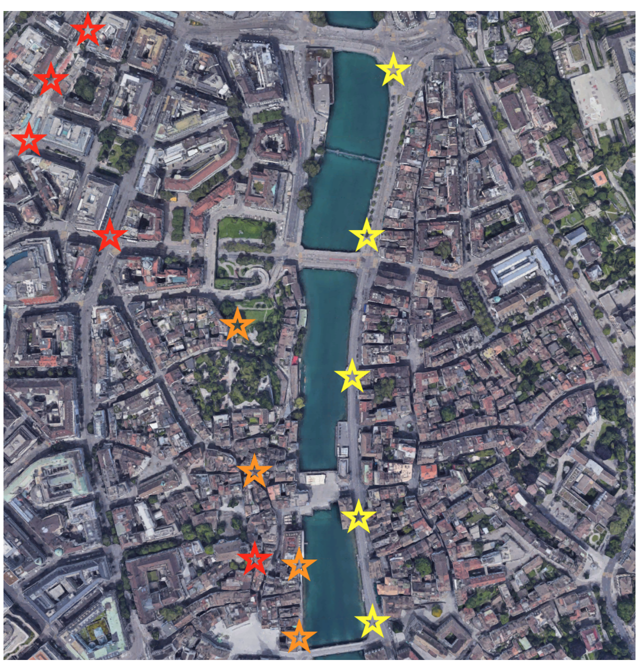


**Fig. 11**: Locations of question points along the path