CityBench
ESPON CityBench for benchmarking European Urban Zones

Scientific Platform / Tools 2013/3/10

Final Report | Version 30 April 2014

EUROPEAN UNION
Part-financed by the European Regional Development Fund
INVESTING IN YOUR FUTURE

This report presents the final results of a "Scientific Platform and Tools" Project conducted within the framework of the ESPON 2013 Programme, partly financed by the European Regional Development Fund.

The partnership behind the ESPON Programme consists of the EU Commission and the Member States of the EU27, plus Iceland, Liechtenstein, Norway and Switzerland. Each partner is represented in the ESPON Monitoring Committee.

This report does not necessarily reflect the opinion of the members of the Monitoring Committee.

Information on the ESPON Programme and projects can be found on www.espon.eu

The web site provides the possibility to download and examine the most recent documents produced by finalized and ongoing ESPON projects.

This basic report exists only in an electronic version.

List of authors

Geodan:	Eduardo Dias Robert Oortwijn Elizabeth Kalinaki
UJI:	Ana Sanchis Luis Rodriguez Sven Casteleyn

Table of contents

List of Tables

List of Figures

A CityBench Executive Summary

This CityBench Final Report describes the (final version of the) CityBench Webtool, which is a tool aimed at benchmarking European Larger Urban Zones (LUZ) on a number of indicators through a web portal. This final version is the result of a development phase which started in January 2013.

ESPON CU provided the TPG with a Response document regarding CityBench Draft Final Report (DFR). This has been used as input for the development of the final version of the Report, the Webtool and the Admin Tool. The latter is a functionality enabling a stakeholder administrator to add, modify and/or delete themes and indicators as required. Since numerous data sources, including ESPON and Eurostat, are constantly creating and updating indicators, the Admin Tool guarantees the sustainability of the Webtool after completion of the CityBench project.

The feasibility of changes to the Webtool as requested in the ESPON CU Response document are listed in Annex V.

1. Main results and activities

Larger Urban Zones

The selection of Larger Urban Zones (LUZ) to be included in the CityBench Webtool was based on four criteria.

- Criterion 1: Include all European cities included in the OECD Metropolitan Areas database.
- Criterion 2: include all EU27+4 countries.
- Criterion 3: Include all LUZ with a population number of > 400,000.
- Criterion 4: Include additional LUZ in underrepresented countries.

Application of these criteria led to a set of 171 LUZ. At the request of ESPON CU, Liechtenstein/Vaduz was added, although it has no LUZ definition.

Indicators & themes

Following the project specifications, the CityBench project has prepopulated the Webtool with a limited set of indicators, belonging to a limited set of themes. Requests by the stakeholders regarding the themes and indicators to be included have been taken into account as much as possible. The theme division is in line with ESPON db themes, although the actual theme naming may differ.

- **Theme Economy and Population**: Population (density), GDP per head in PPS, % of persons unemployed, "Agree" with It is easy to find a job
- **Theme Connectivity**: Potential accessibility by rail / road / air, "Satisfied" with public transport
- **Theme Demography**: Ageing index, Old age dependency
- **Theme Social media**: Crisis awareness, Unemployment awareness
- **Theme Investment climate**: Ease of doing business, Gas / Electricity prices for industrial consumers
- **Theme Environment**: % of LUZ consisting of green urban areas, Maximum PM10 / CO2 concentration, Combined adaptive capacity to climate change, "Satisfied" with Green spaces such as parks and gardens
- **Theme Smartness**: High-Tech (total) patent applications to the EPO per million of inhabitants, IP Addresses, Share of renewable energy in gross final energy consumption
- **Theme Quality of life**: "Agree" with Easy to find a house, "Agree" with I'm satisfied to live in, "Satisfied" with Cleanliness, "Satisfied" with Cultural facilities such as concert halls, theatres, museums and libraries

Indicators were both obtained from official data sources (ESPON, Eurostat, EEA, DG Regio) and derived from social media harvesting (see below).

Social media harvesting

CityBench database includes also social media content. Predefined indicators are generated by real-time retrieval (harvesting) and subsequent transformation of data from social media and are stored as daily values. This data is automatically loaded into the CityBench database and, within the scope of this project, cannot be managed, extended or modified like other data sources indicators. Maintenance of this ETL (Extraction Transformation Loading) is well documented, to ensure its future performance.

Admin Tool

The database contains the above described preloaded indicators, but can be modified and extended, since the current indicator set is far from exhaustive: numerous national and European statistical institutes and enterprises have collected hundreds of indicators and are constantly in the process of updating them.

Thus, in order to warrant the sustainability of the Webtool after completion of the CityBench project, the TPG has created a functionality (the 'Admin Tool') which enables an administrator to add one or more new indicators, update already included ones with newly collected data, and deactivate or delete indicators no longer needed. In addition, themes may be renamed, added and deleted as well. The Admin Tool ensures that at any given time, the stakeholders are equipped with a tailor-made set of themes and indicators corresponding to their needs.

Webtool interface

Each time the CityBench Webtool is opened in a browser, the user is presented with a wizard. (For more advanced users, this wizard can be skipped in order to go directly to the Webtool.) The wizard, consisting of four steps, is meant to assist the user in selecting one or more cities and the indicator(s) and TNC and/or typology of his or her choice. After selecting, the user may choose between either compare the selected cities or view the similarity of the first selected (reference) city with the other(s).

When choosing compare, the user is presented with a graph for each selected indicator, showing the values per selected city and the country and European averages. Time series, if available, are included.

When choosing similarity, a map of Europe is shown, displaying the similarity of the selected cities to the reference city based on the selected indicator(s). The more similar a city is to the reference city, the larger the circle which represents it – and vice versa. Alternatively, similarity may be shown in a radial, with the reference city in the center and the other selected cities around it at approximately their geographic location. The more similar a city is to the reference city, the closer to the center it is located – and vice versa. In this "radial view", the circle size of each city represents the actual indicator value for one of the selected indicators. By selecting a different indicator for display, the circle sizes will change to reflect the values for that indicator.

The user may switch between map and radial view, similarity and compare, add/remove cities and indicators and change TNC and/or typology at any time.

Webtool backend

The CityBench application follows the service oriented architecture paradigm. This basically means that the provided data and functionality is served through web services. The unique entry point to these services is implemented in the form of a Web Geoportal.

The Data Layer contains the components managing the project data, such as the database, structured in a data model specifically designed and implemented for the CityBench project and containing data the from the external data sources (Eurostat, ESPON DB, Social Media and so on) and metadata about data and indicators available in the CityBench project.

The Service Layer contains the services that implementing well-known and standard-based interfaces allow to work with the data in the data layer. The definition of these interfaces together compound the CityBench API used by the CityBench tool, but also other applications could access the data in the CityBench data layer through this API.

The components of the service layer are the Rest Web Services, (REST/JSON) that supply functionality and data to the CityBench clients and the data and metadata documented API that makes it possible to consume CityBench information in an efficient way. In order to use the different CityBench data and functionality in an Internet browser, a JavaScript client API with JavaScript libraries could also be developed in the future with little effort, based on the documented API.

Communication activities

To ensure that the Webtool development fits the stakeholders' requirements, several actions for communication, interaction and dissemination have been deployed during the project timeframe. Meetings and workshops have been attended to present the project and obtain users' requirements and suggestions. Furthermore, an info website and a stakeholder's social network have been tested and several testing sessions and evaluations for the different versions of the Webtool have been carried out in order to obtain continuous feedback from stakeholders through discussions, social media and a questionnaire, the results of which are summarized below:

Questionnaire results

From 11 March 2014 to 22 April 2014, a web survey was conducted among stakeholders of the CityBench project. 18 reactions were received. The detailed response to each question is provided in Annex VI. Based on the 18 responses, we observed the following trends:

Usability: most respondents quickly learned - and remembered - how to work with the Webtool, enjoyed it, found it relatively easy to use and consistent and felt confident about using it. The majority thought it worked well in their browser, found the functions well-integrated and expects to use it frequently and recommend it to a colleague or friend. There were mixed feelings regarding the ease and speed to recover from mistakes and the degree to which the Webtool works as desired.

Usefulness: comparing one / one's own city to other cities based on similarity and viewing that in Map and radial view was considered very useful and interesting, as were the typology and region based filtering option and the graphs comparing up to 4 cities. The importance of accessing the different years for the indicators to see trends was confirmed. Generally the "similarity" concept, as used in the Webtool, was easily understood.

Needs or additional functionality: a majority stated that the Webtool is complete for a quick scan of cities, but would welcome the options to download data and images and to link to a map or radial view using email, Facebook or twitter. A question about showing real values rather than "compare" (i.e. relative) values of indicators yielded an equal amount in favor of either option.

Please see Annex VII for an overview of individual wishes regarding additional functionality and positive aspects / aspects to be improved or changed.

2. Need for further consideration / analysis / research

With the delivery of the final version of both the CityBench Report and Webtool, the CityBench project is completed. Nonetheless, there is considerable potential for expansions and continued developments.

Webtool interface

Explore the use of dynamic maps (enabling zoom, pan, and other navigation tools) and spatial selection tools which could improve the user experience of the Webtool.

Admin Tool

The tool could make available more complex processing and management of big data retrieved from social media, or customization of indicators and LUZ included in the CityBench database.

Future indicators to be included in Webtool

- It is foreseen that ever more Volunteered Geographic Information (VGI) will become available, as the usage of social media using location-aware devices keeps increasing. This source of potentially relevant data should be continuously monitored and explored, to possibly extract new, innovative indicators from it.
- The proposed list of indicators constitutes only a fraction of all data/indicators available (from ESPON, Eurostat, and other sources). ESPON admins may add indicators as required, as the Webtool is designed in such a way as to easily facilitate this.

Building time series

The geographic delimitations currently used by ESPON and Eurostat UA to collect data are different from the ones used previously. ESPON is moving from (the various levels of) the 2006 version of the NUTS nomenclature to the 2010 version. Since NUTS 3 units are the building blocks for the Metropolitan Regions (MR), the effect of using different NUTS versions will propagate to MR level as well. Similarly, the most recent Eurostat UA data collection is based upon the 2012 LUZ version, as opposed to previous Urban Audits that were based upon earlier LUZ delimitation versions. In both cases, acreage differences between the previous and current nomenclature versions may exist to a greater or lesser extent. This may (and probably will) lead to inconsistencies when building time series from data collected for different nomenclature versions. Therefore, if time series from ESPON and/or Eurostat data are to be built, it is important to ensure that data for different years was collected for the same nomenclature version or, if that is not the case, differences between the nomenclature delimitations are negligible.

Improving selected set of indicators

- Although the LUZ (version 2012) level was to be the backbone of the Webtool, it was found that the main provider of LUZ level indicators, the Eurostat Urban Audit dataset, still has many data gaps, especially when time series are considered. Therefore NUTS 3 / MR level indicators are used as an approximation of the LUZ level instead. Nonetheless, it is advisable to scan the Urban Audit dataset regularly for updates, so as to eventually replace NUTS 3 / MR with LUZ data.

- Several indicators currently selected for the Webtool have been collected at MR level or NUTS 3 level, the 2006 version. The rationale for using the 2006 version is that the ESPON database as yet contains very little data collected for the 2010 version of NUTS 3. It is expected that future data collections will be based upon the 2010 NUTS nomenclature. Once enough new NUTS 2010 based data, including time series, for the selected indicators will be available, this should replace the current data.

Enabling the integration of OGC / LOD

An important addition to the system would be to implement the capability to ingest data from distributed databases and sources using standard and well-known formats like OGC (Open Geospatial Consortium) or LOD (Linked Open Data) and/or to export visualized data in these formats.

In the past the usage of LOD has been suggested, both for current (e.g. CityBench Webtool) and future applications. The advantages to ESPON / Eurostat of using LOD could be summed up as follows.

- ESPON / Eurostat data can be found, accessed and interpreted much more easily, both by persons and by systems / applications capable of handling LOD.
- This may lead to increased usage of all the valuable data stored within ESPON / Eurostat db.
- Since data is linked rather than uploaded, LOD indicators included in the CityBench Webtool are updated automatically as the data source (ESPON / Eurostat db) is updated.
- If other ESPON project databases are LOD-enabled as well, each project will benefit from the fact that data needs to be updated at one location (the data source) only, instead of having to update each project database separately.
- Additionally, external data used by ESPON projects can remain at the source; there is no need to copy the data.
- Inclusion of new LOD indicator(s) in the Webtool using the ETL tool is much easier than uploading manually created sheets containing new indicator(s) with their metadata, because only a link to the new indicator(s) needs to be specified. Since the metadata format is standardized, there is no need to specify this manually; it is added automatically and error-free.
- Seamless combination of data with metadata and semantics.

Section 2.7 provides a detailed overview of the possibilities of adapting CityBench to using LOD.

B CityBench Final Report

1. Methodologies, models and concepts developed and used

Based upon the input from stakeholders' user experience, survey results and internal discussions and meetings by the TPG, development of the Webtool has continued and led to its final version. This Section describes the methodologies, models and concepts developed and used within the framework of the CityBench project.

1.1 European directive INSPIRE

The CityBench application follows the service oriented architecture paradigm. This basically means that the provided data and functionality will be served through web services. The unique entry point to these services is implemented in the form of a Web Geoportal.

Having this in mind, we follow two main principles:

- The application must be scalable and extensible. The CityBench application has to be adapted to new requirements and new data. This means that the application has to be designed in a modular way. Furthermore, these modules should have a well-known functionality and interface so the application is more easily maintain. Most important decisions in this line are: to design the architecture and implement the components according to international standards and initiatives. As we will describe further in the document, we rely on European directives to design the components serving required functionality and on international standards to implements their interfaces and serve the data.
- Underlying technological complexity is transparent for the user. The use of standards, and communication protocols of the CityBench system should be transparent for the user. The CityBench web portal provides user with a unique entry point to the system, offering an intuitive manner to handle data and providing advanced search and visualization methods to derive useful information out of the massive amount of existing data.

CityBench project architecture follows the European directive INSPIRE (INSPIRE, 2007) which defines a classical service-oriented three-layer architecture. CityBench application extends this architecture to include the required functionality in the form of Network services and the client applications in the form of a Web Portal.

As INSPIRE recommends, CityBench project implements its components according to the Open Geospatial Consortium (OGC), which defines open standards for the successful exploitation of geospatial resources on the Internet.

1.2 Agile software development

Agile software development is a group of software development methods based on iterative and incremental development, where requirements and solutions evolve through collaboration between self-organizing, cross-functional teams. It promotes adaptive planning, evolutionary development and delivery, a time-boxed iterative approach, and encourages rapid and flexible response to change. It is a conceptual framework that promotes foreseen tight iterations throughout the development cycle.

Following this paradigm, the system will not be implemented in a single so-called 'waterfall' but in an iterative and incremental approach. First the initial requirements are determined, and a first analysis and design is performed. Preferably in cooperation with the steering group the first functionalities and components to be implemented are selected and prioritized. Subsequently these are implemented in a first prototype. Based on the feed-back and evaluation of this prototype a next iteration is performed, resulting in modified and new functionalities and components. In this way the SG can control and steer the intermediate and final results.

Thus, the CityBench development will have a participatory strategy approach, an iterative, agile-based, collaborative methodology to design and develop the software platform itself. The focus will be on constant communication and collaboration not only among all consortium partners, but also with all stakeholders involved, such as coordinators of the case-studies, experts, and case study participants in order to, through various interactions, collect functional and user interface requirements to yield an effective solution to best suit end-users' needs.

Early prototype and agile-based methodology makes developers aware of any tendency of change in the technology field, thereby detecting potential problems early. Also it allows to incrementally and adaptively learn through the design and development of the software platform through the project lifecycle and the rigid deploy-and-test strategy within the agile development methodology enables to incrementally increase and control complexity.

1.3 System Usability Scale

Usability testing is a technique used in user-centered interaction design to evaluate a product by testing it on users. This can be seen as an irreplaceable usability practice, since it gives direct input on how real users use the system. Usability testing focuses on measuring a human-made product's capacity to meet its intended purpose.

The most used questionnaire for measuring perceptions of usability is the System Usability Scale (SUS), released into this word by John Brooke in 1986. SUS is technology independent and has since been tested on hardware, consumer software, websites, cell-phones, IVRs and even the yellow-pages. It has become an industry standard with references in over 600 publications.

The SUS is a 10 item questionnaire with 1 to 5 response options (from "strongly disagree" to "strongly agree"):

- I think that I would like to use this system frequently.
- I found the system unnecessarily complex.
- I thought the system was easy to use.
- I think that I would need the support of a technical person to be able to use this system.
- I found the various functions in this system were well integrated.
- I thought there was too much inconsistency in this system.
- I would imagine that most people would learn to use this system very quickly.
- I found the system very cumbersome to use.
- I felt very confident using the system.
- I needed to learn a lot of things before I could get going with this system.

The SUS, in a slightly rephrased version, was part of a web questionnaire prepared by the TPG in order to learn the stakeholders' opinion on the CityBench Webtool. Before filling in the questionnaire, the stakeholders were asked to carry out several tasks using the Webtool. The results of this questionnaire are summarized in and detailed in Annex VII.

1.4 Social media harvesting

1.4.1 Introduction

When studying the existing information to analyze and compare cities we cannot forget new trends in digital information and social media. Nowadays bottom-up initiatives complement the ecosystem of information available online. Regular citizens are sharing information about their surrounding and their cities through a large number of social networks. This increasing amount of information cannot be ignored, since it provides in most cases, real time information about places and events, which have impact on society. In this context we witness how location analytics industry is moving from a paradigm of lower volume, higher accuracy data to one of higher volume and lower accuracy. The CityBench project aims to generate indicators not only using official data but also indicators which offer a social view analyzing data shared by citizen through social networks.

Although georeferenced user-generated data still represents a small percentage, its growth is being greatly accelerated largely by the use of sensor-enabled devices. It is thus reasonable to foresee that huge amounts of georeferenced data will be available in an immediate future.

1.4.2 Indicators from social media data

It is of interest to the project to provide social media information about current circumstances regarding Crisis, Politics, Economy and unemployment. The interest of citizens of talking, asking or informing about these issues provides an overview about social concern of citizens and therefore it could influence in vulnerable fields such as financial markets. This also could provide useful information to support decision making to local government or a potential investor.

After studying first results obtained from social media sources, we defined a starting set of indicators related to a topic inside a LUZ delimitation, defined as ”topic items in a city per total items in that city”, intended to describe the social media interest of that city in the cited topic. For doing so, we generate lists of keywords (in every EU language) that illustrate the topic field to create a query to send to the Social Media Networks, then we create a batch process to send this query to Twitter and store the number of items posted in a city. This is done for every city, on a daily basis.

For the moment, indicators are shown in two different temporal resolutions: daily, as stored, and yearly aggregated, in order to enable comparability together with the rest of indicators, stored by year. Nevertheless, data is stored daily and can be presented on a daily/weekly/monthly/yearly basis or any other temporal interval that can be derived from the daily data..

First topic tested is Crisis awareness, presented as “inside a city LUZ perimeter in its social Media (Twitter) activity” and as “inside a city LUZ perimeter related to the total crisis awareness in social Media (Twitter) activity”. Being these definitions too close and leading to misunderstanding, we decided to keep only the first indicator, more related to the city than the second. Data for this topic is stored from November 2013, thus, year 2013 indicator takes into account data from November and December. Year 2014 indicators takes into account data from 1st January to the present day.

Second topic being implemented is Unemployment awareness inside a city LUZ perimeter in its Social Media (Twitter) activity, as suggested by ESPON CU. Data is stored from April 2014, thus, indicator for year 2014 start aggregating data from April.

1.4.3 Methodology

These predefined indicators about data extracted from the social networks are generated retrieving real-time data from social networks for each city and transforming them to derive indicator information, stored as daily values. This data is automatically loaded in the CityBench database and, within the scope of this project, cannot be managed, extended or modified like other data sources indicators.

Technical description and maintenance of this ETL (Extraction Transformation Loading) is well documented for ensuring its future performance and can be consulted on Annex III.

1.5 Similarity

The concept of “similarity” is a crucial component of the Webtool. By similarity we mean the degree to which indicator(s) values between LUZ are comparable.

Two processing steps are involved in deriving similarity values:

1. All indicators are normalized, i.e. for each indicator the lowest LUZ value is assigned a value of 0 and the highest a value of 1.
2. For one, two or three normalized indicators, the ‘Euclidean’ distance between the indicator(s) values for one LUZ and one or more other LUZ (or country / European average) is calculated. Euclidean distance (or Euclidean metric) is the distance between two points as measured with a ruler and using the Pythagorean formula to derive the (metric) distance. The more indicators are selected, the more dimensions are involved in the calculation (one dimension for each indicator added).

For example, if three (normalized) indicators (a, b and c) and two LUZ (A and B) are selected, the similarity between these LUZ is calculated as follows:

sqrt ((aA - aB)2 + (bA - bB)2 + (cA - cB)2)

If the calculated distance is 0, there is full similarity; if the distance is equal to the sqrt(number of dimensions), there is maximal dissimilarity. In other words, the closer the number is to 0, the higher the degree of similarity between the selected LUZ, based on the chosen indicators. Conversely, the closer the number is to 1 (or even exceeds 1, which occurs as more indicators are added), the more different the LUZ are.

1.6 Visualization of similarity

One of the key features of the Map view of the Webtool is the ability to observe regional patterns when showing similarity to a particular LUZ based on the values of one indicator or a combination of more indicators. However, it was found that in some cases 'outliers', i.e. a LUZ with either a very low or a very high value for a particular indicator, cause the regional patterns to be obscured while it is included in the LUZ selection (e.g. Luxembourg, which has a very high GDP per head). Therefore three classification methods (of the standardized values, using five classes) were explored to assess how they affect the map view, in an effort to minimize the effect of outliers on the overall appearance of the map – and preferably maximize visibility of regional patterns. In addition, the effect of using ranks rather than classified values was assessed.

Standardized GDP per head (2009, Metropolitan Region) was taken as sample indicator, with three LUZ for which similarity to other LUZ was shown. The three LUZ are Luxembourg (€ 75,191: highest GDP per head among the LUZ selected for the Webtool), Amsterdam (€ 40,568: intermediate GDP per head among the LUZ selected for the Webtool) and Plovdiv (€ 3,728: lowest GDP per head among the LUZ selected for the Webtool).

The results of different methods to display LUZ similarity information are as follows.

1. **Equal intervals**, where "*the attribute values are divided into n classes with each interval having the same width=Range/n*" (De Smith, Goodchild and Longley 2007). Classes used: 0 - 0.2, 0.2 - 0.4, 0.4 - 0.6, 0.6 - 0.8, 0.8 – 1.

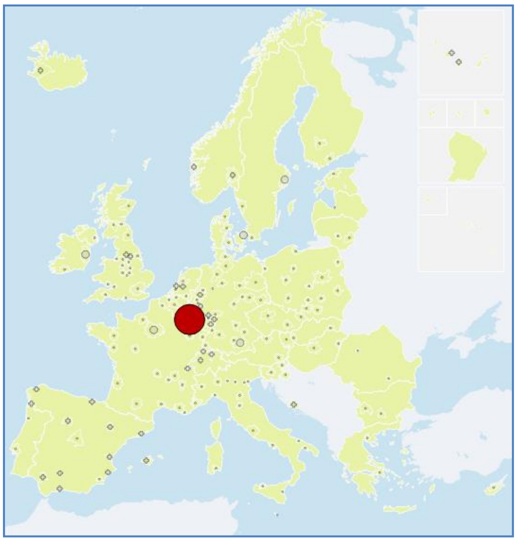


Figure 1-1. Equal interval: similarity to Luxembourg.

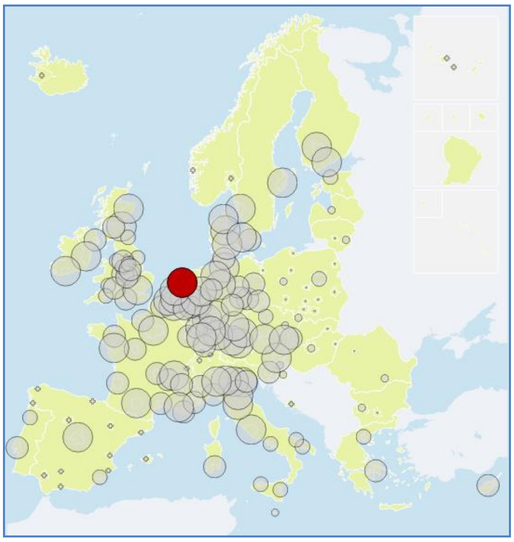


Figure 1-2. Equal interval: similarity to Amsterdam.

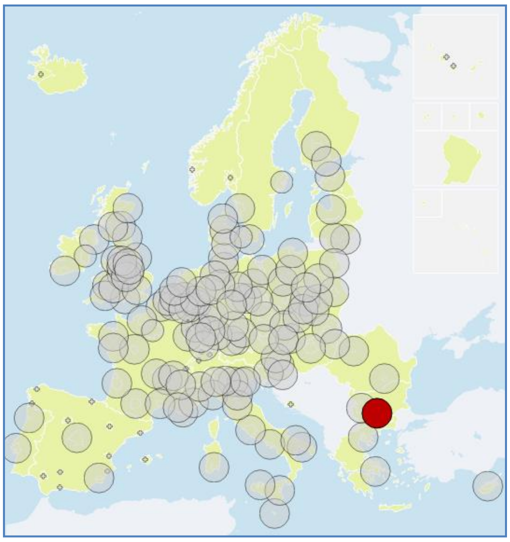


Figure 1-3. Equal interval: similarity to Plovdiv.

The figures above show that equal interval classification works best for LUZ with intermediate indicator values, since regional patterns are visible then. The high outlier (Luxembourg) does stand out, but the low outlier (Plovdiv) not at all; in the latter case any regional pattern is obscured by the high outlier.

2. **Jenks natural breaks**, which is a "*…variance-minimization classification. Breaks are typically uneven, and are selected to separate values where large changes in value occur*" (De Smith, Goodchild and Longley 2007). Classes used: 0 - 0.156, 0.156 - 0.293, 0.293 - 0.430, 0.430 - 0.617, 0.617 – 1.

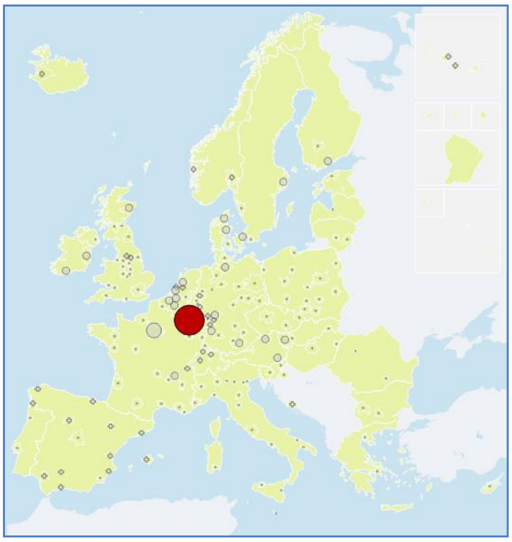


Figure 1-4. Jenks: similarity to Luxembourg.

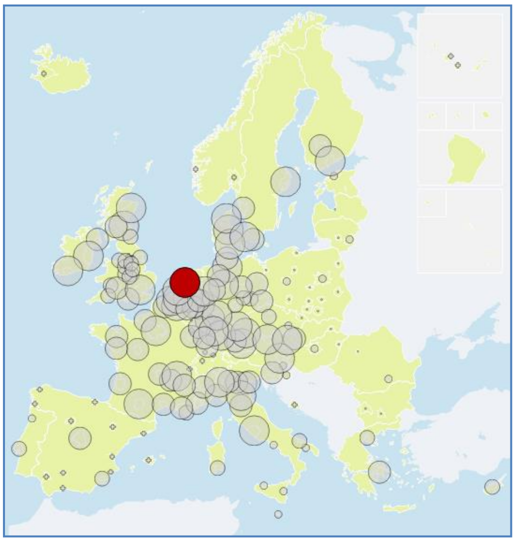


Figure 1-5. Jenks: similarity to Amsterdam.

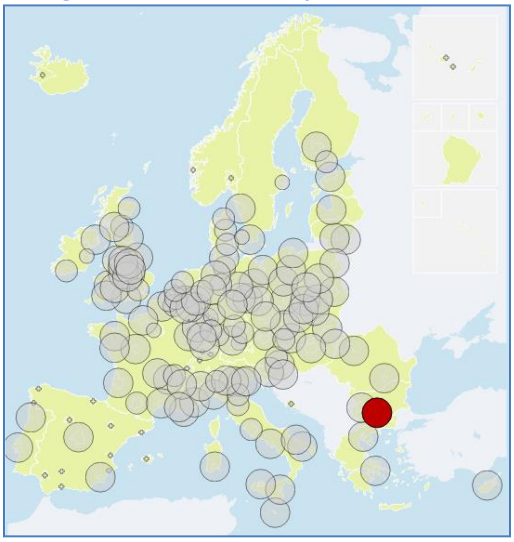


Figure 1-6. Jenks: similarity to Plovdiv.

In this case as well, the high outlier obscures any regional pattern when viewing similarity to a low GDP LUZ. Some pattern can be observed when viewing similarity to the high outlier; but again, patterns are best seen when viewing similarity to a LUZ with intermediate indicator values.

3. **Quintiles**, which means that "*intervals are selected so that the number of observations in each interval is the same*" (De Smith, Goodchild and Longley 2007, p.100). Classes used: 0.000 - 0.577, 0.577 - 0.654, 0.654 - 0.720, 0.720 - 0.838, 0.838 - 1.000.

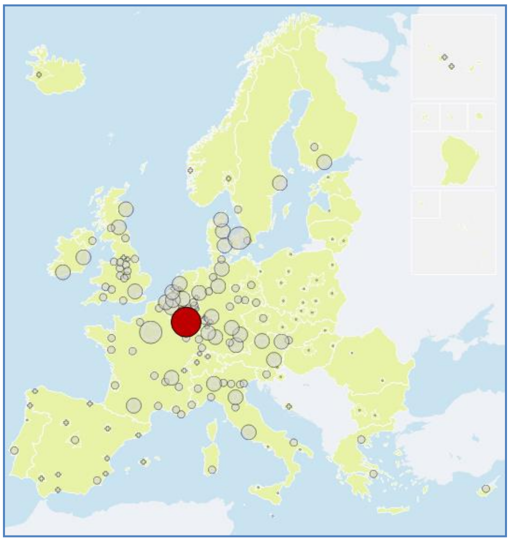


Figure 1-7. Quintiles: similarity to Luxembourg.

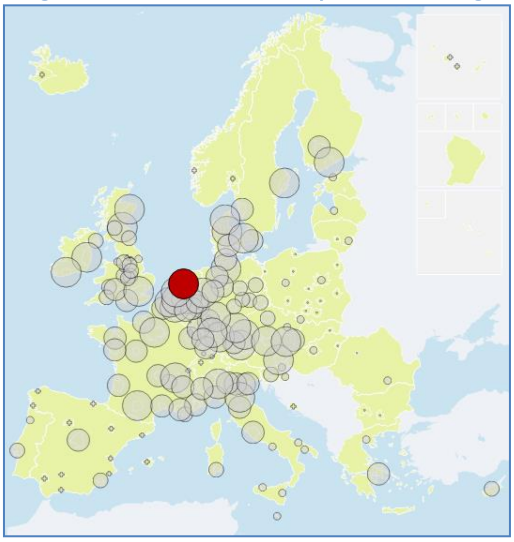


Figure 1-8. Quintiles: similarity to Amsterdam.

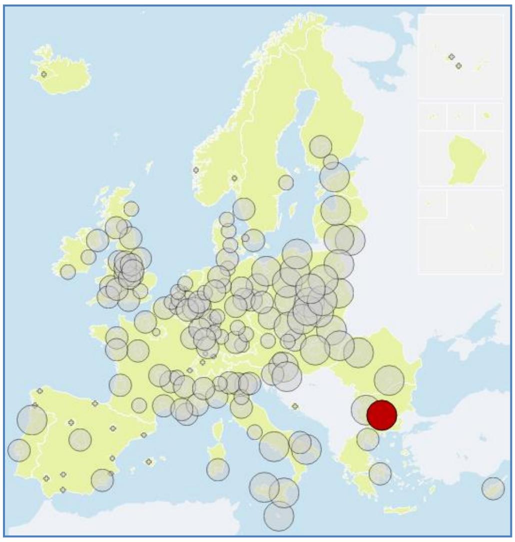


Figure 1-9. Quintiles: similarity to Plovdiv.

Compared to the previous classification methods, the regional pattern in case of the high outlier is clearer, whereas the low outlier shows some kind of regional similarity pattern, although hardly distinct. The intermediate indicator value yields a pattern similar to the ones generated by equal interval and Jenks.

4. **Ranking**: LUZ are ranked from low to high indicator values and assigned the ranking value. Ranking range: 1 – 171.

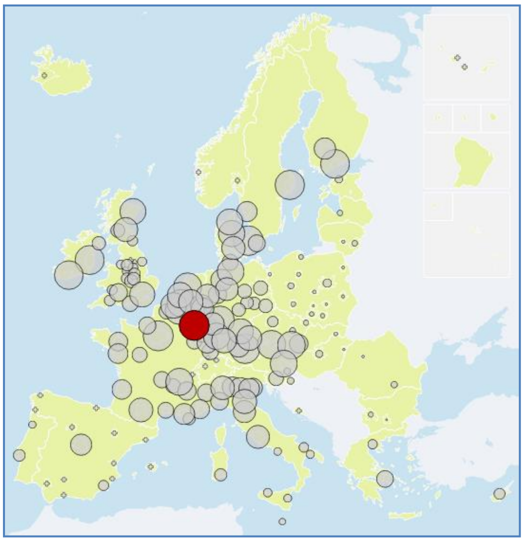


Figure 1-10. Ranking: similarity to Luxembourg.

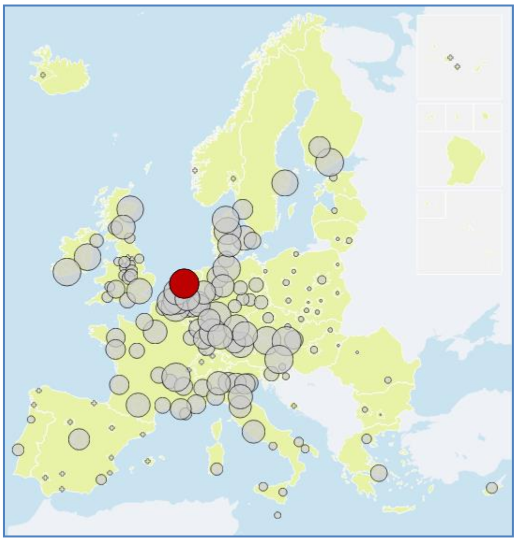


Figure 1-11. Ranking: similarity to Amsterdam.

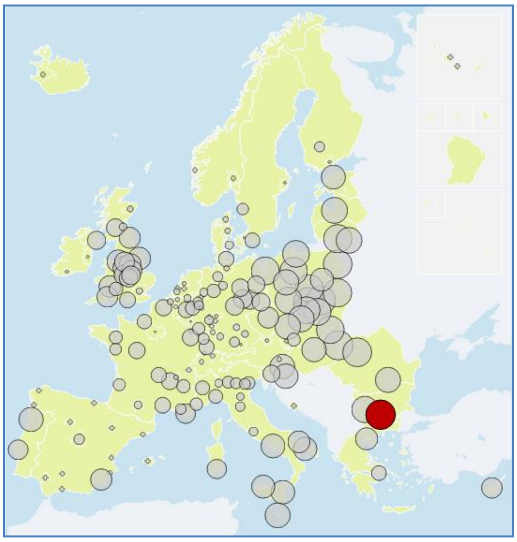


Figure 1-12. Ranking: similarity to Plovdiv.

Using the LUZ GDP ranking values yields rather different maps when showing similarity to high and low outliers: now Luxembourg is part of a pattern covering Central Europe, Ireland and Scandinavia, while Plovdiv shows a clear similarity to Eastern European LUZ. Also note that the maps for Luxembourg and Amsterdam are almost identical, indicating that GDP for Amsterdam reflects an order of magnitude common to Central Europe, Ireland and Scandinavia.

Based on the findings above, it was decided to use ranking values for the map view of the Webtool, since this method best reveals regional patterns. However, this does not apply to the 'radial view': because of its inherent non-geographic representation of similarity it does not show regional patterns. The radial view is therefore created from the true (standardized) values for an indicator. The calculation of similarity (i.e. the distance from a reference LUZ) in the radial view is derived from the values for all indicators selected, however the circle size corresponds to the standardized values of only one of the indicators selected, namely the top one in the list. If it is removed, the new upper one will be used to represent the circle size.

2. CityBench content

2.1 Data sources

Although many data sources were scanned for suitable indicators, ESPON CU requested that as the number of indicators to be initially included should be only 15-20, the ESPON projects and the Eurostat database should be the main sources of indicator data. Additionally, a limited number was to be derived from “other” sources, e.g. EEA and social media harvesting. A later request involved including DG Regio “Perception survey” results.

2.2 Geographic levels

LUZ level

ESPON CU has stated that the (current version of the) Larger Urban Zone geographic level was to be the backbone of the Webtool. Therefore, when searching for indicators attention was focused on indicators collected at LUZ level. The main provider of LUZ based indicators is the “Urban Audit”, an annual survey conducted by Eurostat. One of the issues encountered was that, until 2012, all LUZ indicators had been collected for a previous version of LUZ delineations. This version not only differs in size and shape of the LUZ, but also in number from the current LUZ version.

Eurostat published the first batch of indicator data collected for the current LUZ version at the beginning of 2014. However, this dataset contains as yet many data gaps. It was therefore decided to refrain from using Urban Audit data for now and replace it with comparable indicators from the “metropolitan regions” dataset, also from Eurostat (see below).

NUTS 3 level

A lot of relevant and recent indicator data is available at the NUTS 3 regional level. ESPON projects belong to the main suppliers of NUTS 3 indicator data. To be able to use NUTS 3 data in the Webtool, it has to be aggregated to the MR level, which is the NUTS 3 approximation of the major LUZ. This aggregation can be performed using a NUTS 3 - MR correspondence table.

Metropolitan region (MR) level

278 MR in 30 countries have been defined by Eurostat. These regions are aggregations of NUTS 3 areas and are considered by ESPON CU to be reasonable (and acceptable) approximations for the corresponding LUZ. This would mean that data collected at MR level could be included in the Webtool to constitute an additional set of indicators for selected LUZ. All but two included LUZ areas have a corresponding MR.

(Core) City level

The DG Regio “Perception surveys” have been collected at the (Core) City level, which, according to Eurostat, is “…the city as defined by its administrative and/or political boundaries”. As such, a Core City is always part of, and contained within, the LUZ to which it belongs. Thus (Core) City and LUZ are strictly speaking different geographic levels; moreover, there is no correspondence between (Core) City and NUTS 3 or MR. Perception survey results are nevertheless included as indicators in the Webtool.

Country level

Some data is only collected at country level, often because it is related to country policies, regulations, economy, etc. An example of this is the set of 'Ease of Doing Business' indicators developed by IFC and World Bank. Using this kind of country level indicators implies that each LUZ within a specific country is assigned the same (country) value for a certain indicator. This is still interesting to the CityBench project, because it highlights differences between LUZ in different countries.

2.3 Included LUZ / MR areas

The selection of Larger Urban Zones (LUZ) to be included in the CityBench Webtool was based on four criteria.

- Criterion 1: Include all European cities included in the OECD Metropolitan Areas database.
- Criterion 2: include all EU27+4 countries.
- Criterion 3: Include all LUZ with a population number of > 400,000.
- Criterion 4: Include additional LUZ in underrepresented countries.

Application of these criteria led to a set of 171 LUZ. At the request of ESPON CU, Liechtenstein/Vaduz was added, although it has no LUZ definition. An overview of all LUZ, including their corresponding MR, is presented in Annex I.

Further development of the Admin Tool could include a LUZ delimitations uploading functionality; for now this has been adjourned to give priority to an easy to understand updating indicators functionality, well in line with ESPON CU requirements.

2.4 Themes and indicators

Following from ESPON CU feedback and an comparison of indicator value availability for similar indicators at LUZ and MR/NUTS 3 level , the draft final list of themes and indicators as proposed in the Draft Final Report was slightly modified. Annex II presents an overview of the final list of themes and indicators, including their operationalization.

However, it should be stressed that even the final list constitutes merely an, albeit well thought out, starting point for populating the Webtool with additional themes and/or indicators as needs for them arise. The 'Admin Tool', which is explained in Section 3.4, facilitates this. It also enables editing (e.g. adding newly collected values) and deactivating or even deleting existing indicators.

Annex III provides a detailed description on harvesting social media and deriving indicators from it.

Comparison of themes

An assessment has been performed regarding the degree to which the proposed CityBench themes correspond to those used in other projects. Table 2-1 presents an overview of ESPON db themes, current CityBench themes and themes useful to EIB (as listed in CityBench Project Specifications document and CU Response document).

Table 2-1. Comparison of ESPON, CityBench and EIB themes.

ESPON db themes	CityBench Webtool themes	Themes useful to EIB
Population and living conditions	Demography	Population age; Wealth
Labor market	Economy and Population	Education and Employment
Education		Education and Employment; Workforce Skills
Health and safety	Quality of life	Quality of life
Information society	Social media	
Agriculture and fisheries		
Transport and accessibility	Connectivity	Transport/accessibility
Environment and energy	Environment	Energy profile; Greenhouse gases
Science and Technology	Smartness	
Governance	Investment Climate	Business environment
Territorial structure	[Typology, TNC]	
	Investment Climate	Development demands

A comparison of ETMS dimensions and CityBench indicators is provided in Table 2-2.

Table 2-2. Comparison of ETMS dimensions with CityBench indicators.

ETMS dimension: Economic competitiveness	**ETMS dimension: Environmental qualities**	**ETMS dimension: Human capital**	**ETMS dimension: Social inclusion**	**ETMS dimension: Availability of Services and Functions**
GDP per inhabitant	% of LUZ consisting of green urban areas	population density	% of persons unemployed	Potential accessibility by rail
# of items being posted about 'Crisis'	Residential PM10	Ageing index	# of items being posted about 'Unemployment'	Potential accessibility by road
Ease of doing business	Combined adaptive capacity to climate change	Old age dependency		Potential accessibility by air
Gas / Electricity prices for industrial consumers	Share of renewable energy in gross final energy consumption			# of items being posted by tourists
High-Tech (total) patent applications to the EPO per million of inhabitants				IP Addresses

Table 2-1 shows that while there is a certain degree of correspondence between the theme lists, some differences exist as well. In particular, the number of Webtool themes is lower than that of either the ESPON db or the EIB wish list. The main reason for this is that the Webtool has been prepopulated with a limited set of indicators. To ensure that every theme contains two or three indicators, a maximum of seven themes was decided upon. Also, ESPON db theme 'Agriculture and fisheries' was not considered relevant for the Webtool's purpose of comparing cities.

Similar to the indicators, the Admin Tool facilitates adding, editing (i.e. renaming) and deleting of themes. In the latter case, all indicators belonging to that theme will move to the 'Other indicators' group. The Admin Tool functionality is explained in more detail in Section 3.4.

Urban Atlas data

ESPON CU requested the substitution of Corine Land Cover (CLC) 2006 data (used for calculating the amount of green urban areas within each LUZ) with Urban Atlas (UA) data. This data is only available in vector format from the European Environment Agency (EEA) website, with separate files for each LUZ. Processing these vector files (downloading, rasterizing, calculating green urban areas for each rasterized LUZ) would be a very time-consuming process. However, the Universitat Autònoma de Barcelona (Spain) has merged and rasterized the entire UA dataset and has made this rasterized data available to the TPG.

However, since it turned out that UA data uses LUZ 2004 instead of LUZ 2012 boundaries, it was agreed with ESPON CU to refrain from using UA data and to keep using CLC data.

DG Regio Perception Survey results

ESPON CU also requested the inclusion of DG Regio Perception Survey results. Although survey results are available for only part of the LUZ selected for the Webtool, the TPG has integrated some of the Perception Survey results into the final version of the Webtool.

NUTS 2 data

Several indicators requested by CityBench SG / ESPON CU have been collected at NUTS 2 level.

1. "Human Resources in Science and Technology": probably reference is made to an indicator with this name in Annex C to the ESPON SIESTA project. However, this data is available in Eurostat rather than ESPON db.
2. "Private R&D expenditure as % of GDP": probably 'Business Expenditure on R&D as % of GDP', also mentioned in Annex C to the ESPON SIESTA project, is meant. An exact match was found in neither Eurostat nor ESPON db, however Eurostat table 'Total intramural R&D expenditure (GERD) by sectors of performance and NUTS 2 regions' seems to correspond well.
3. "Regional quality of governance": this refers to the QoG EU Regional Dataset.

Integrating indicators collected at NUTS 2 level in a tool using the LUZ level as backbone poses several problems.

1. No correspondence table between NUTS 2 and LUZ (or MR) exists, therefore each LUZ has to be linked to a particular NUTS 2 region manually.
2. In some cases, NUTS 2 regions are much larger than the LUZ included and/or they contain more than one LUZ. Figure 2-1 presents some examples of this. The larger a NUTS 2 region is compared to a LUZ within its boundaries, the less realistic it will be to assign NUTS 2 derived indicator values to that LUZ. And if a NUTS 2 region contains more than one LUZ, there is the additional issue of how to allocate (a portion of) the NUTS 2 value to each LUZ.

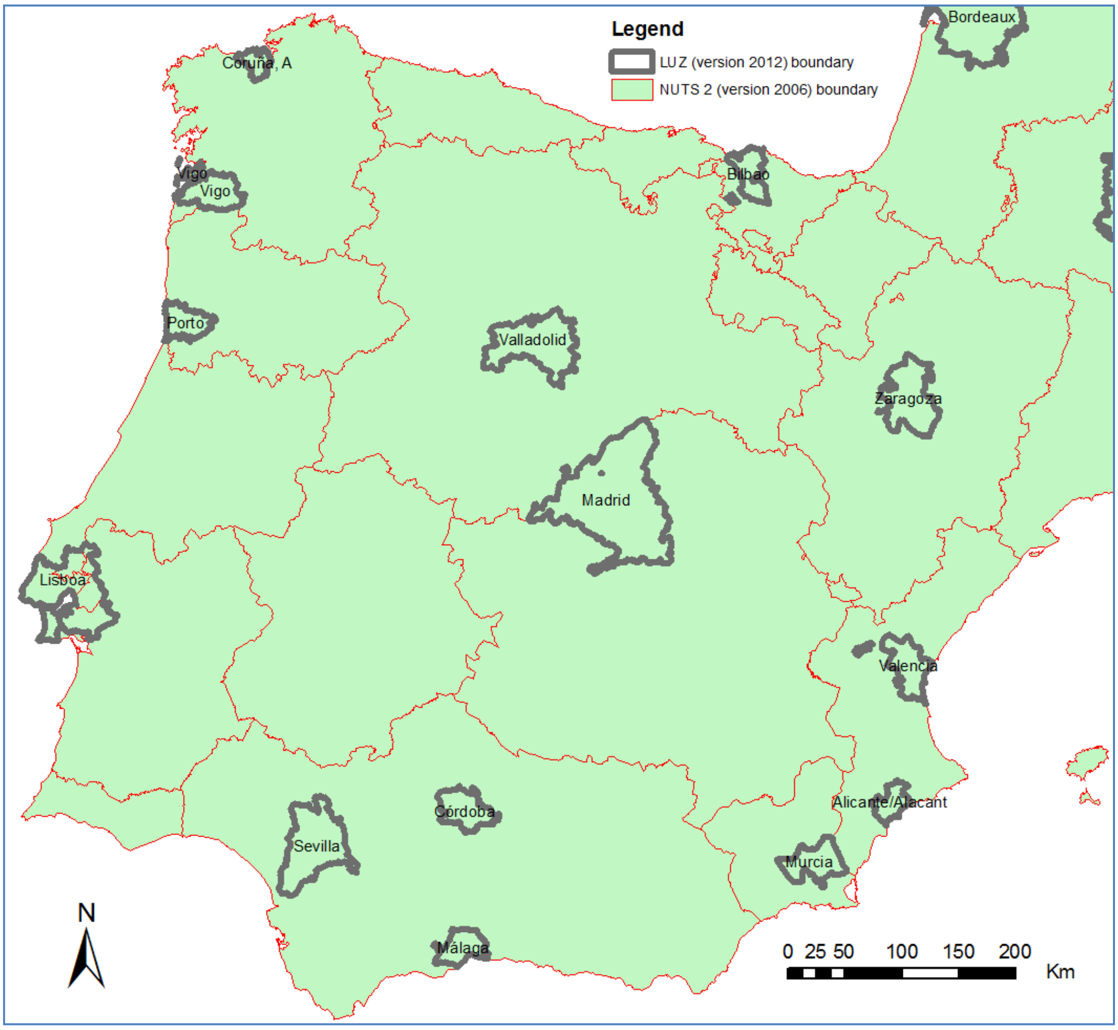


Figure 2-1. Examples of NUTS 2 regions much larger than LUZ or including more than one LUZ: Iberian Peninsula.

3. In other cases, one LUZ covers (parts of) more than one NUTS 2 region, which is exemplified in Figure 2-2. The problem here is deriving an indicator value at LUZ level from several NUTS 2 values, which necessitates taking into account e.g. NUTS 2 weighting factors (area? population?) and area of each NUTS 2 included in a LUZ.

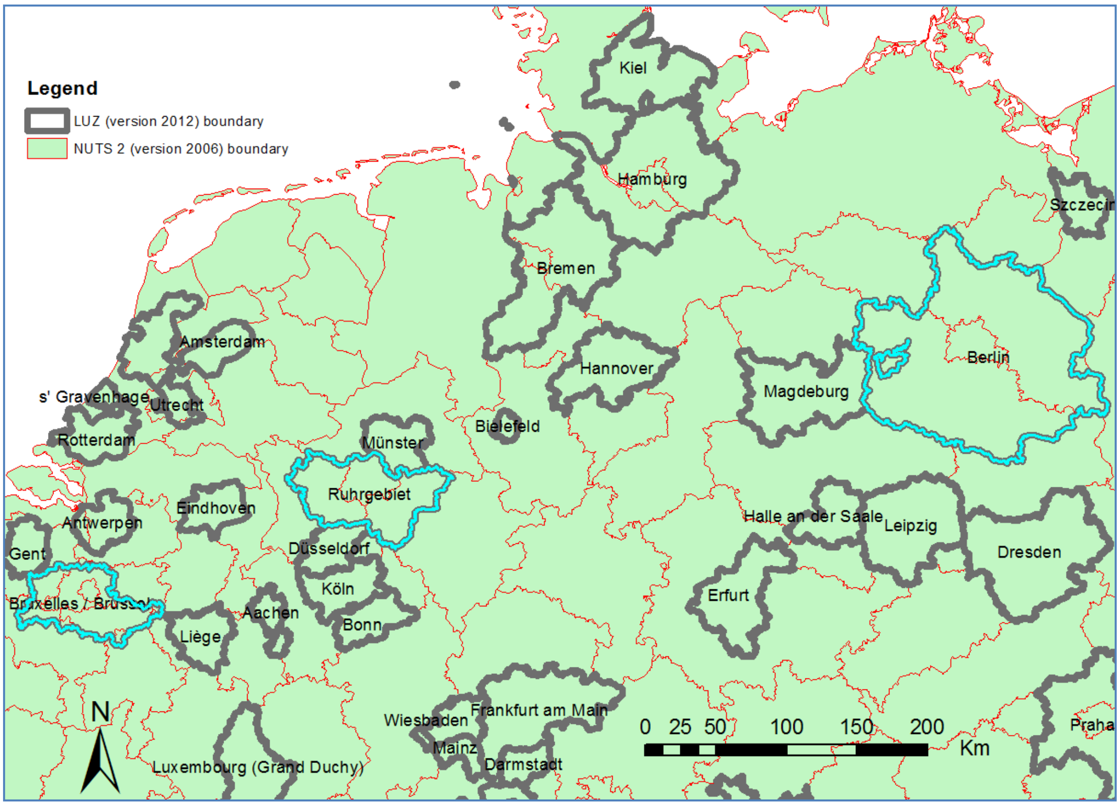


Figure 2-2. Examples of LUZ covering (parts of) more than one NUTS 2 region: Bruxelles, Ruhrgebiet and Berlin.

Whether these problems may be overcome depends on the topic covered by the indicator. NUTS 2 level indicators covering environmental variables, e.g. hours of sunshine, can be used as proxy for the LUZ level, since they are unlikely to differ much at these geographic levels.

The requested indicators, however, can be considered societal (demographic, economic, etc.) variables, which do tend to vary within and between adjacent regions. In this case the usage of indicators collected at NUTS 2 level is not recommended by the TPG. Therefore the requested NUTS 2 indicators have not been included in the Webtool.

2.5 Indicator visualization

The CityBench Webtool has two options for displaying indicator similarity values: the “Map” view and the “Radial” view.

Map view

The Map view of the Webtool shows how all included LUZ compare based on one, two or three indicators. The LUZ currently selected serves as the reference LUZ, to which the others are compared. The size of the circle representing a LUZ corresponds to the similarity of each LUZ to the reference LUZ: the larger the circle, the greater the similarity (and vice versa).

Radial view

The Radial view provides, for one, two or three indicators, an alternative way to show the similarity between the LUZ you selected and all other LUZ. The reference LUZ is located in the center of the graph, with the other LUZ placed around it, more or less at their actual geographic location relative to each other. The larger the distance between reference and other LUZ, the less similar they are, based on the chosen indicator(s). The LUZ circle size represents its actual value for the indicator selected for display.

2.6 Indicator time series

Whether time series can be included in CityBench depends on 1) data availability and 2) consistency of geographic units for data collection over time. 1) is self-explanatory: if for a particular indicator values for multiple years are (yet) lacking, no time series can be integrated into the Webtool for that indicator. 2) means that if for a particular indicator values from several years have been collected, comparable NUTS 3 / MR / LUZ boundaries should have been used for each collection year. If this is not the case, the time series may show fluctuations in indicator values which are not related to actual changes but to the varying size (or shape) of the geographic unit used for collecting from one year to the next. Generally the following observations apply.

1. Indicators covering several years created by ESPON projects (e.g. INTERCO, TRACC, TEL Update) are usually based on NUTS 3 version 2006 delineations, which can be aggregated to Metropolitan Region (MR) version 2006 level. Therefore time series for these indicators can be integrated in the Webtool.
2. Eurostat provides a Metropolitan regions database (*met*), which includes various indicators covering different years. Although metadata for the various tables is not available, it may be safely assumed that the same NUTS 3 version was used to calculate the yearly values. This data can therefore be used for time series.
3. Eurostat Urban Audit LUZ data should be approached with more care, as previous collection rounds used different LUZ delimitation versions. The latest Urban Audit data collection, released in October 2013, uses the LUZ version 2012 delineations. These may differ from previous delineations to a more or lesser extent (see also CityBench Indicators Report, Chapter 3). This is reflected by the fact that the Urban Audit database (*urb*) does not contain indicator values collected by previous Urban Audits for those LUZ that have been modified considerably, i.e. LUZ codes ending with *2*. Approximately 50% of the LUZ selected for inclusion in the Webtool have a code ending with 2, thus Urban Audit time series for these are not possible.
4. For indicators collected at country level inclusion of time series is straightforward, since country boundaries do no tend to change.
5. Gridded data covering different years, e.g. provided by the European Environment Agency, can be converted to indicator values for any geographic unit desired, such as LUZ 2012. If the same delineations are used for each year, the creation of a time series is straightforward.
6. Within the framework of the CityBench project, Social Media data is harvested real-time from various sources: Twitter, YouTube and Flickr. As the project is progressing, ever more data is collected. Provided that data collection will continue after its completion, this may serve as input for a time series of indicators derived from it.

For the indicators currently included in CityBench, time series have been included where possible. Users can explore the temporal data in the map and radial views by selecting the available years. And in the graph comparison tool, the indicators are presented as time series.

2.7 Enabling the integration of Linked Open Data

2.7.1 Introduction

Before elaborating on technical requirements and procedures for producing Linked Open Data from the CityBench data, it is important to understand what Linked Open Data is, and what its principles are. Based on this understanding, we then discuss different strategies for generating the Linked Open Data, leading to LOD datasets with varying embedded semantics, and of varying quality.

Linked Open Data refers to a way of structuring and making structured data available so that it is openly accessible, and linked with other data. Linked Open Data builds on Web technologies (most importantly, URI's- most often URL's - to represent resources), and uses Semantic Web technology (RDF, http://www.w3.org/RDF/) to represent data. A key issue of Linked Data is the fact that it should be linked, i.e., all data should be connected using semantic relationships.

Tim Berners-Lee outlined 4 principles for linked data:

- Use URIs to denote things.
- Use HTTP URIs so that these things can be referred to and looked up ("dereferenced") by people and user agents.
- Provide useful information about the thing when its URI is dereferenced, leveraging standards such as RDF, SPARQL.
- Include links to other related things (using their URIs) when publishing data on the Web.

2.7.2 Generating CityBench data as Linked Open Data

In the previous section we explained what Linked Open Data is, and to which principles it needs to adhere. It is hereby important to observe that these constitute technical requirements and design guidelines; they are not a definite guide for creating LOD. In fact, they leave open a broad range of possibilities when creating LOD sets, from basic raw data complying with the LOD principles, to sophisticated, well-described and well-linked LOD sets. Depending on the desired LOD set, the complexity to create it increases.

Specifically for CityBench, we discern three different cases which are increasingly rich, but also increasingly complex to create:

- Level 1: The CityBench data is dumped/exposed as a raw dataset, without further processing. This would make the raw dataset accessible as LOD. However, from a semantic and LOD point of view, the dataset would be relatively poor, with few semantic knowledge and few interlinking.
- Level 2: The CityBench data is expose as richer semantic information, including well-described domain information, re-use of existing vocabularies, and linking the CityBench data within the Linked Open Data cloud. As examples, consider re-using well known vocabularies (e.g. GeoNames, http://www.geonames.org/) and linking with well-known LOD sets (e.g., DBPedia, http://dbpedia.org/).

- Level 3: The CityBench data is exposed as a rich LOD set (see level 2), additionally including implicit domain knowledge and (semantic) relationships, which is implicitly present in CityBench, but not explicitly represented. As an example, consider the relationship between Metropolitan Regions and Large Urban Zones, which is only implicitly present.
- Level 4: The CityBench data is exposed as a rich LOD set which includes implicit domain knowledge and (semantic) relationships (see level 3), and the dataset is additionally enriched with derived information, not (explicitly) present in the original dataset, and currently typically programmatically calculated. As an example, consider how CityBench calculates indicators and performs certain operations (e.g. certain indicators exist only for some geographical levels (i.e. NUTS3) and the service converts indicators values and make it available for others (i.e. NUTS)). Other example of data created through the services is the calculation of certain indicators at the European level.

The decision of which type of LOD set needs to be generated depends primarily on its intended use, and will have implications for the technical realization, the required effort and financial support needed.

2.7.3 Technical considerations

We now elaborate on the technical aspects of generating a LOD set from CityBench data. We first state the current relevant technical CityBench specifications, then elaborate on how they influence LOD set generation.

The relevant technical CityBench specification are the following:

1. Data is stored in a relational database (PostgreSql, http://www.postgresql.org/).
2. The data is accessed through a REST (http://en.wikipedia.org/wiki/Representational_state_transfer) interface. Evidently, logical data relations and constraints are defined at database level, but progressive availability of datasets and evolution of the CityBench requirements led to the fact that some data logic is present in the REST services that access the data in the form of transformations or operations over data (i.e. aggregation, normalization).
3. The CityBench database is updated regularly, possibly using data obtained from external sources.
4. The CityBench database does not contain a big volume of data, but scalability and integration in with or into other (larger) systems needs to be taken into account.

Based on these technical observation related to the current CityBench Implementation, the following technical challenges for LOD creation arise:

A. How to expose existing relational data as RDF data?

There are two main possibilities to tackle the problem: 1/ conversion of data, and subsequent use of the converted dataset, 2/ using a layer/service on top of the existing DBMS that perform on-the-fly DBMS-RDF translation. Given the particularities of CityBench, in particular fact that data is regularly updated, and a considerable effort has already been put into importing/transforming/merging data from external datasets, the second solution seems preferable. On the other hand, the first solution presents a onetime effort. If the current data management code isn't hard to migrate, this might actually lead to a better performing solution. A hybrid solution, in which some data is made available on-the-fly, and other is transformed, may be the best solution if different parts of the dataset have highly different properties (e.g., importation, handling effort & cost, freshness requirements). It is thus clear that both solutions have their advantages and disadvantages, and a deeper study of these is needed to determine the best solution in the CityBench case.

For either alternative there are various possibilities for implementation. There is a range of tools available to assist in data conversion and for making data on-the-fly accessible as LOD. Some solutions offer both, and allow a hybrid approach. Just mentioning a few of these possible tools, APIs or services to use could be, for just mentioning few of them: Jena (https://jena.apache.org), Triplify (http://triplify.org/) or D2R Servers such as D2RQ (http://d2rq.org/) or OpenLink Virtuoso (http://virtuoso.openlinksw.com/wiki/main/Main/) for various data management tasks.

B. Which vocabularies to use?

The key to well-described and useful LOD is the re-use of existing vocabularies. This immediately raises the question: which vocabularies to use, and how to extend them to suit the particular needs of CityBench. The vocabularies to be used need to be carefully chosen for describing the data in the CityBench domain, whereby it is important to consider widely used vocabularies as starting point. Several LOD sets exists specifically in the geo-informatics field (e.g., GeoNames, http://www.geonames.org/ or LinkedGeoData, http://linkedgeodata.org/), possibly more specialized vocabularies exist, but also popular broader LOD sets need to be considered (e.g., DBPedia, http://dbpedia.org/). In any case, a thorough analysis needs to be performed and the alternatives need to studied.

C. How to integrate/link CityBench LOD with existing data?

The key of having real linked open data is linking it with other datasets, not only by re-using existing vocabularies, but also by re-using existing resources, and enriching the data with descriptions linking existing resources. Part of the data included in the CityBench database already exists, or is further described, in existing datasets, such as DBpedia, Geonames (for geographical features), or Eurostat linked data (http://eurostat.linked-statistics.org/). Integrating the CityBench LOD with these existing datasets is an essential feature of good linked data, and opens the door for data integration, enrichment, and ultimately, new applications incorporating data from the LOD cloud. Therefore, methods and techniques need to be studies to automate the search and selection of appropriate URI's to (re-)use and to further describe existing CityBench concepts. The ample research body that has been dedicated to this problem needs to be studied, and applied to the specific case of CityBench. Additionally, and unavoidably, a (additional) manual effort will be required.

2.7.4 Conclusion

From this document it is apparent that different parameters influence the technical choices on one hand, and the required effort and costs of the solution on the other hand. Of primary importance is the intended use of the LOD in the CityBench context, and in ESPON context in general. Choices regarding quality and extent of the Linked Open Data need to be made, and according to this, choices regarding technical solutions, and support software and hardware, need to be made. The chosen technical solution then determines the required manpower, and is impossible to estimate without a clear view on the requirements.

The chosen technical solution then also determines the required software that is best suited to realize it. However, it is safe to say that software costs is expected to be minimal, as open source tools exist, and commercial solutions are not overly expensive (cost estimate: 0 to 1000 EURO). Similarly, regarding required hardware, the equipment largely depend of the (extent of the) LOD set is of the elected, the services that needs to be supported, and the envisioned functionality on top of it. Also here, we expect a relatively small costs (cost estimate: 0 to 2000 EURO).

3. CityBench system

3.1 Database and Services

A detailed description of the CityBench technical architecture is provided in the CityBench Webtool Technical Report (Annex VI).

3.2 Usage scenarios: interrogation paths and application flow chart

The use cases delivered by ESPON CU and EIB served as input for the creation of ‘interrogation paths’, i.e. scenarios for approaches and interests corresponding to different user groups when accessing the Webtool. Four user groups were defined:

1. Private investor, looking for business opportunities.
2. Local official, comparing his/her own city to others or searching for cities similar to his/her own city.
3. European official, interested in regional / geographic performance (based on TNC and/or typology).
4. EIB investor, wishing to invest in under-achieving city, city type (typology) or region (TNC).

The objectives of the latter two user groups were perceived to overlap to a large extent; they were therefore combined into one interrogation path. The interrogation paths are shown in Figure 3-1, Figure 3-2 and Figure 3-3.

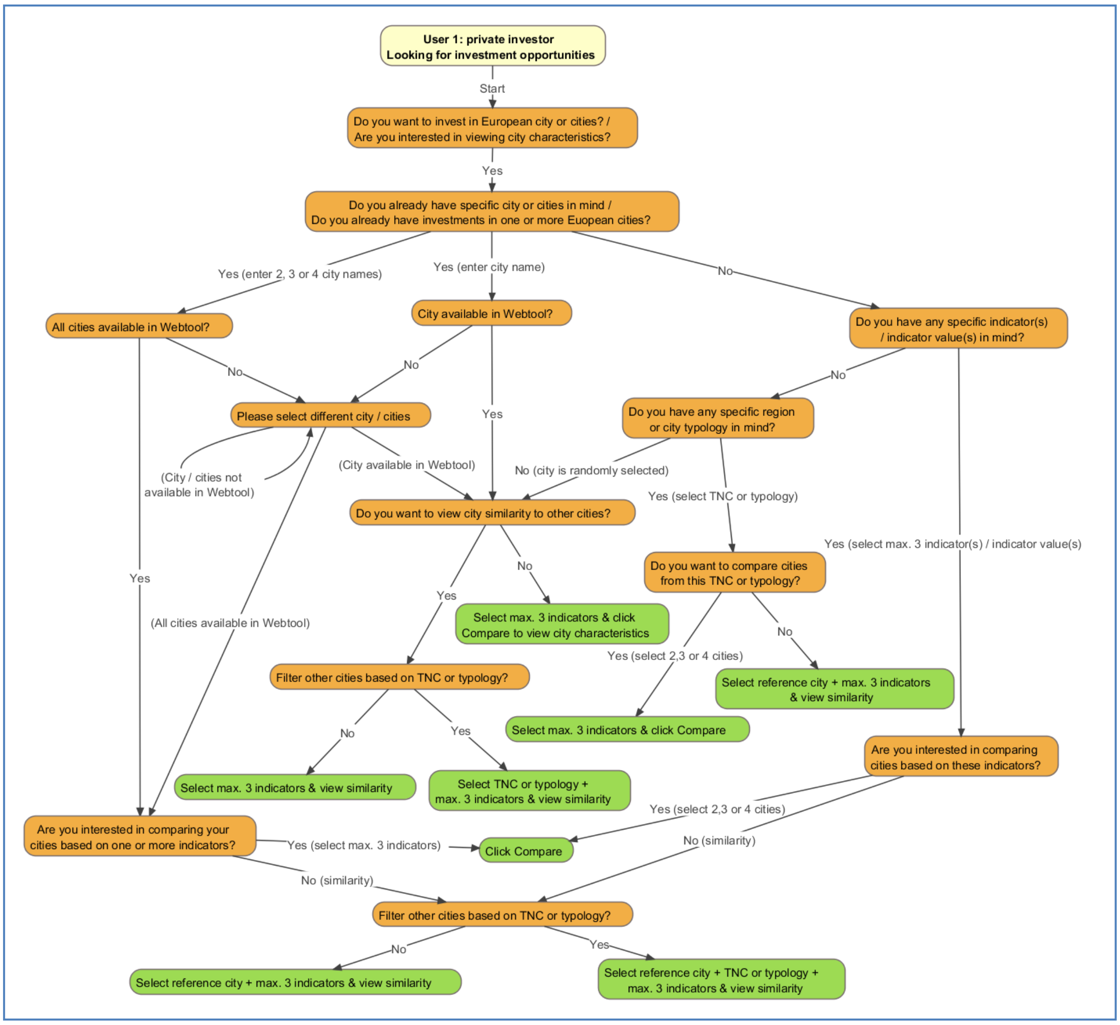


Figure 3-1. Interrogation path for private investor.

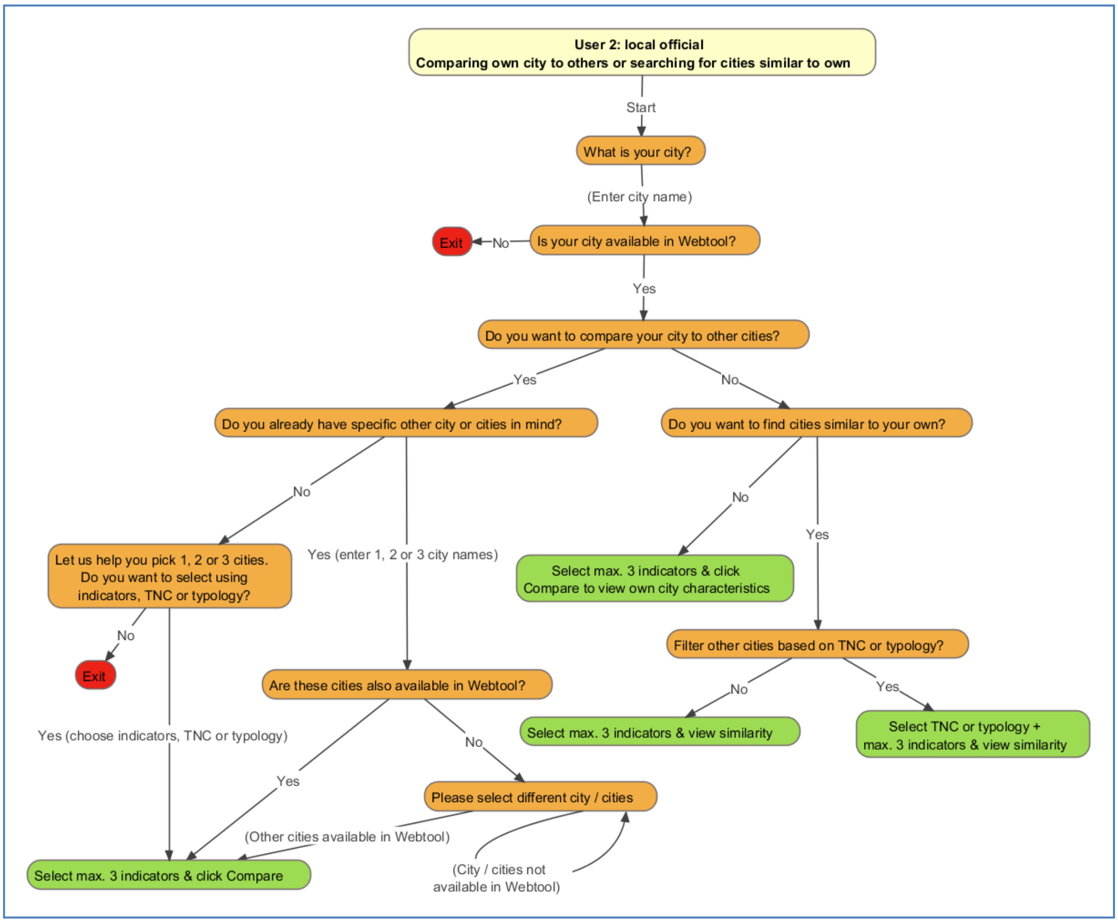


Figure 3-2. Interrogation path for local official.

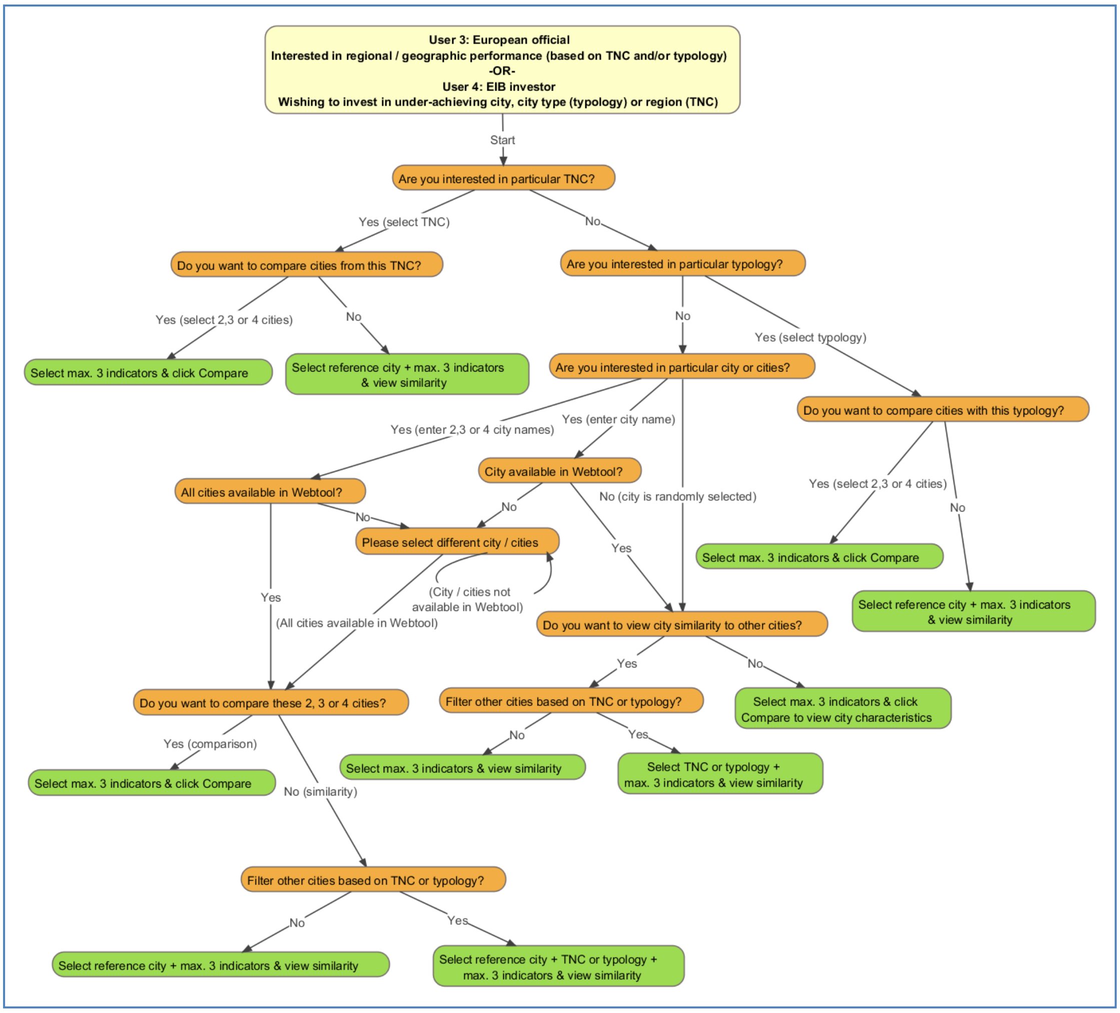


Figure 3-3. Interrogation path for European official or EIB investor.

Based on the interrogation paths in the previous Figures, a flowchart for the CityBench Webtool was designed, see Figure 3-4. The flowchart is an effort to combine the approaches and interests as reflected by the different user groups into a comprehensive application setup. This is especially useful for designing the Webtool Wizard, which is launched every time a user accesses the Webtool.

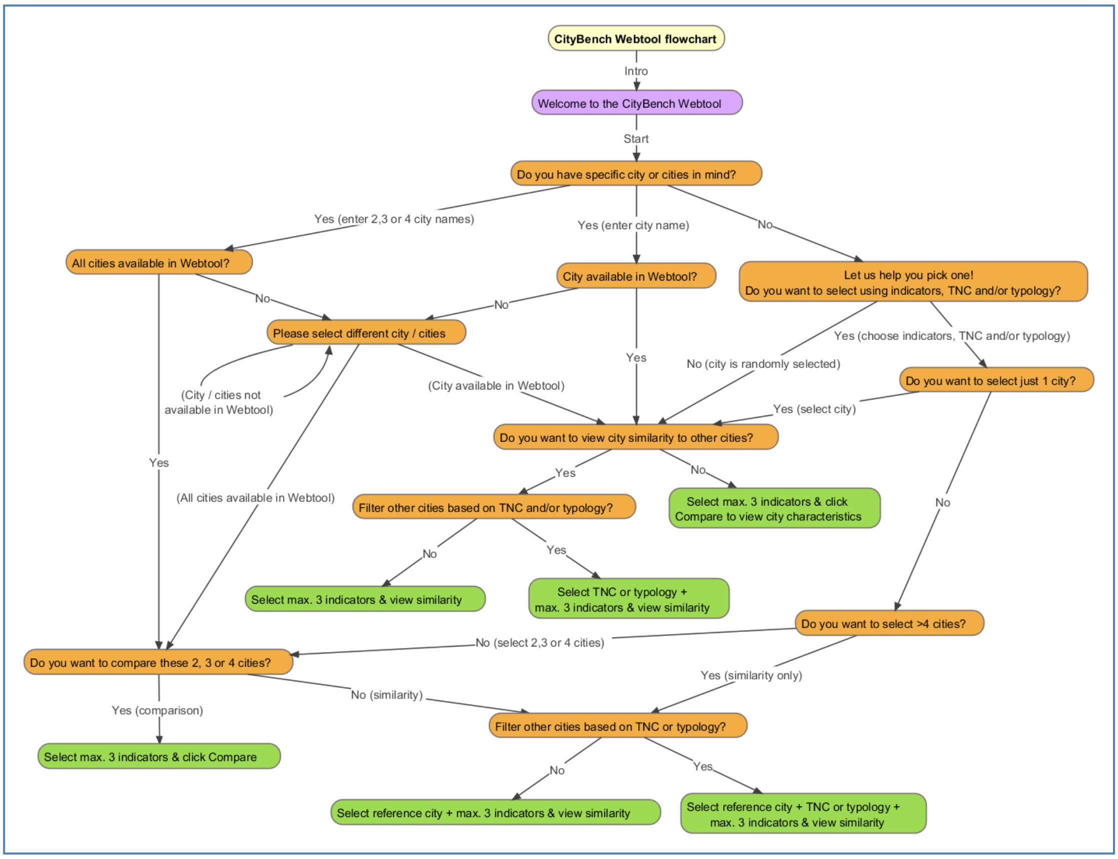


Figure 3-4. Application flowchart.

3.3 Webtool Wizard

The flowchart presented in Figure 3-4 served as a starting point for the design of the Webtool Wizard, which guides the user through the process of selecting a city or cities. Please note that although the Wizard launches every time a user visits the Webtool homepage, it can be discarded at any time. This is especially useful for experienced Webtool users.

The Wizard consists of the following steps.

- Step 0: introductory screen, providing information about the Webtool.
- Step 1: screen for selecting city or cities (see Figure 3-5). When a city name is being typed, the Webtool will automatically complement it, provided that it includes that particular city. A circle on the map represents the location of the city. Once a city is specified, Step 2 will turn blue and become clickable. The user may specify up to four cities before continuing. The screen contains two additional buttons: ‘random city’ and ‘help me choose’. When the ‘random city’ button is clicked, the Webtool will randomly select one or more cities for the user. The ‘help me choose’ button will direct to a screen displaying the indicators and TNCs / Typologies. Here the user may set values for preferred indicators using sliders and/or filter based on TNC or Typology. The cities within the resulting selection will become visible in the Map view, from which the user may select up to four by clicking them.

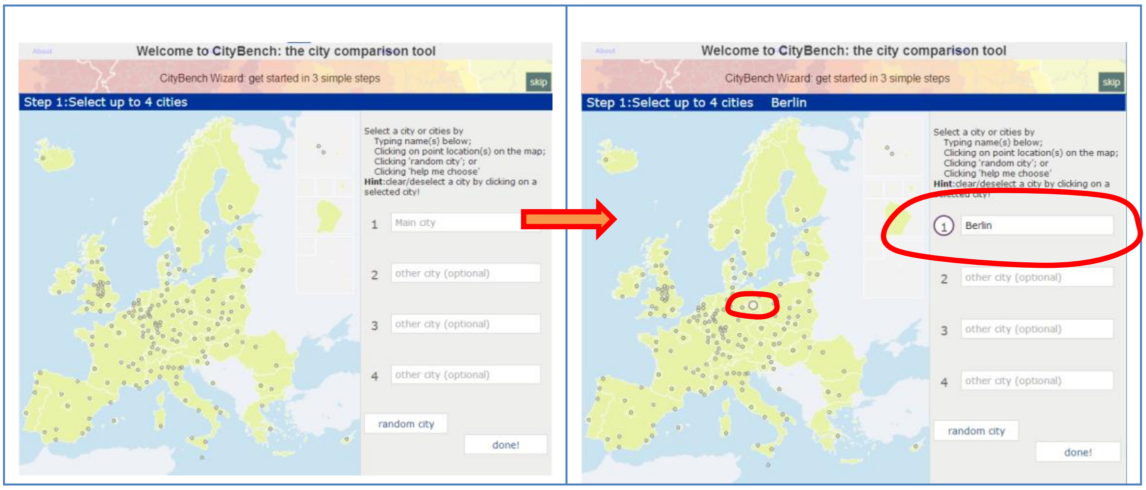


Figure 3-5. Wizard step 1: selecting a city or cities.

- Step 2: screen for selecting indicators and/or TNC / typology (see Figure 3-6). The indicators are available from a dropdown list, which becomes visible when a 'select an indicator' field is clicked. Similarly, clicking the 'select a region' or 'select a typology' field brings up dropdown lists of TNC and typologies, respectively. Once an indicator has been selected, Step 3 will turn blue and become clickable. Up to three indicators can be specified. The Webtool also enables a user to compare his/her city or cities to a sub selection, based on either region (TNC) or typology, of the included cities.

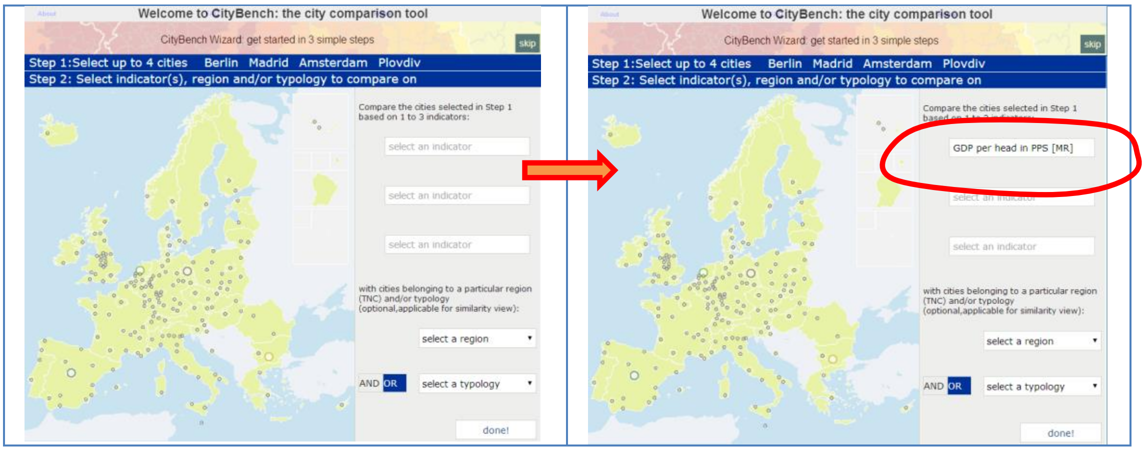


Figure 3-6. Wizard step 2: selecting indicators and/or TNC / typology.

- Step 3: screen providing – and explaining - the choice between comparing cities and looking for similar cities (see Figure 3-7). As the difference between 'comparing' and 'looking for similar' cities may not be clear to all, this step provides some text on both methods. Depending on the choice made here, the Wizard jumps to the Webtool presenting either a page with graphs and charts or a map and similarity graph.

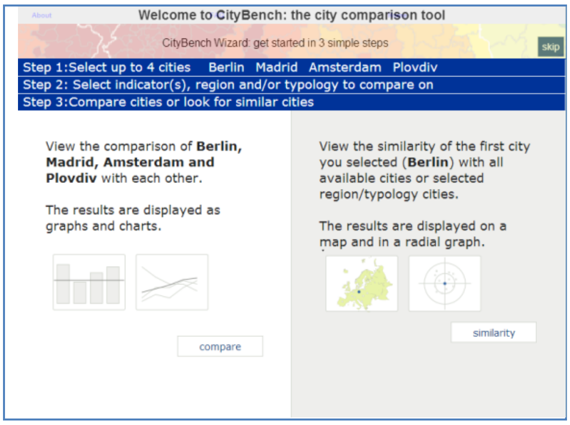


Figure 3-7. Wizard step 3: explaining and providing the choice between comparing cities and looking for similar cities.

3.4 CityBench Admin Tool

Following the project specifications, the CityBench project has pre-populated the Webtool with a limited set of indicators, belonging to a limited set of themes. Requests by the stakeholders regarding the themes and indicators to be included have been taken into account as much as possible. However, the current indicator set is far from exhaustive: numerous national and European statistical institutes and enterprises have collected hundreds of indicators and are constantly in the process of updating them.

Thus, in order to warrant the sustainability of the Webtool after completion of the CityBench project, the TPG has created a functionality (the 'Admin Tool') which enables an administrator to add one or more new indicators, update already included ones with newly collected data, and deactivate or delete indicators no longer needed. In addition, themes may be renamed, added and deleted as well. The Admin Tool ensures that at any given time, the stakeholders are equipped with a tailor-made set of themes and indicators corresponding to their needs.

3.4.1 Upload template

Using an automated upload tool entails that the data be presented in a standardized format. The template to be used for new / updated indicator values and accompanying metadata is provided as a separate deliverable and is also available for download from the Admin Tool interface.

At the request of ESPON CU, the TPG has performed an analysis of using the M4D template for uploading new or changed indicator values into CityBench. Although this is in principle possible, there are several reasons why the TPG has decided in favor of a custom CityBench template instead.

- Rather than being a Key Indicators / Case Study / Background Data dataset provider which uploads data to ESPON db, CityBench may be considered a dataset consumer which is populated with data downloaded from ESPON (and other) db.
- The M4D template is very detailed, meaning that it contains much more metadata fields than what is needed for uploading to CityBench.
- Not all fields in M4D template are in a fixed position, which makes it difficult to automate an upload functionality based on deriving info from specific cell

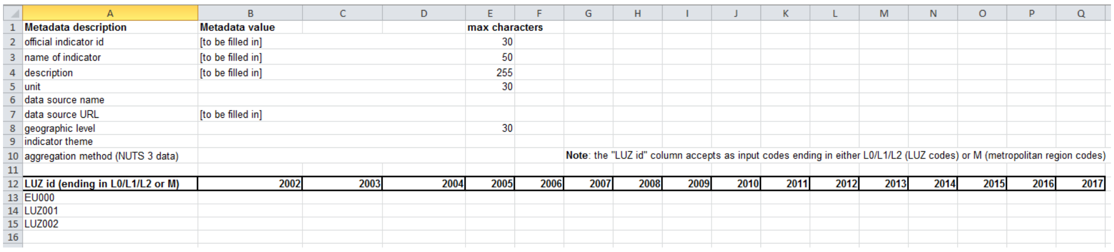
positions (example: Numerator / Denominator Name, which moves down as the number of Keywords and Temporal Extents increases).

- No-data in M4D formatted datasets downloaded from ESPON db is sometimes marked as ":" and sometimes as "N/A", which is difficult to account for.
- Actual indicator values are sometimes combined in one Data sheet and sometimes in several Data sheets, which makes it difficult to automate an upload functionality based on deriving indicator values from a specific sheet (and specific columns, from a specific row downward).
- Data sources other than ESPON, for example Eurostat (Urban Audit) and EEA (raster data), do not use M4D template format, which necessitates the manual formatting of data in these cases. Therefore, the simpler the template to use for upload to CityBench, the quicker this may be performed.
- The M4D template does not contain a field or values for "EU average", which according to ESPON CU should be included in CityBench.

The template contains several drop-down lists from which, for each new indicator, values have to be selected for:

- Data unit;
- Data source name;
- Geographic level of source data;
- Indicator theme;
- Aggregation method (applicable in case of NUTS 3 data).

The value(s) for EU average may optionally be added on a separate row; in that case LUZ_id should be specified as "EU000". Figure 3 8 shows a screenshot of the template.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Metadata description	Metadata value			max characters												
2	official indicator id	[to be filled in]			30												
3	name of indicator	[to be filled in]			50												
4	description	[to be filled in]			255												
5	unit				30												
6	data source name																
7	data source URL	[to be filled in]															
8	geographic level				30												
9	indicator theme																
10	aggregation method (NUTS 3 data)						**Note**: the "LUZ id" column accepts as input codes ending in either L0/L1/L2 (LUZ codes) or M (metropolitan region codes)										
11																	
12	LUZ id (ending in L0/L1/L2 or M)	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
13	EU000																
14	LUZ001																
15	LUZ002																
16																	

Figure 3-8. Screenshot of the data upload template.

The drop down lists will be prepopulated with various values. However, the lists may be modified by a Webtool administrator as required. If for example a particular data source not included in the data source drop down list is being used more often, the administrator can add that source as an entry to the data source list, after which it will be available from the drop down list.

3.4.2 Preparing indicator values for uploading to CityBench

Preparing non-LUZ/MR (NUTS 2 / 3, country) data for usage in the Webtool may involve manual processing of indicator values. In general the following processing steps apply (please note that other methods may be used to achieve the desired results):

1. Retrieve data from a data source (ESPON, Eurostat, EEA, etc.).
2. Process according to geographic level of data.
    a. Country level data: assign country value to each LUZ within its boundaries.
    b. NUTS 3 level data: using the NUTS 3 – MR correspondence table, aggregate NUTS 3 values to MR level by summing, (weighted) averaging, etc., depending on the topic covered by the indicator. Exceptions: Reykjavik and Vaduz, for which the NUTS 3 values are considered to represent the MR value.
    c. NUTS 2 level data: use NUTS 2 value as an approximation for LUZ value. Please note that NUTS 2 data should be approached carefully (see Section 0).
    d. Raster data: depending on the topic covered by the indicator, calculate the average, maximum, minimum, median, etc. for each LUZ.
    e. MR level data: no processing necessary.
3. If absolute indicator values are to be converted to relative values, choose the appropriate division unit, e.g. LUZ area or number of residents.

When uploading MR or aggregated (to MR) NUTS 3 level indicators the Admin Tool will automatically convert the MR codes in the data to the corresponding LUZ codes. For example, Berlin MR code (DE001M) is changed into Berlin LUZ code (DE001L1), which is recognized by the Webtool.

More information on indicator uploading using the Admin Tool is available in Annex IV.

3.5 Web text Sections

This Chapter presents the final version of the textual content of the Webtool. Attention will be directed primarily at key website Sections: Home, About, How to Use, FAQ, etc. The main prerequisite to be taken into account is that the text blocks should be in plain English, implying that the use of technical and/or scientific phrases should be restricted as much as possible. The text blocks should be considered mere suggestions, which are open to discussion with the stakeholders.

Home

Welcome to the CityBench Webtool. With this tool you can compare many European 'Larger Urban Zones' (LUZ) on the basis of one or more indicators. The indicators cover several themes: Economy and Population, Connectivity, Demography, Social Media, Investment Climate, Environment and 'Smartness'.

About

The CityBench Webtool has been developed within the framework of a project funded by the European Observation Network for Territorial Development and Cohesion (ESPON), an EU programme. The goal of the Webtool is to provide an easy-to-use interface that enables a quick benchmarking / comparison of two or more European Larger Urban Zones (LUZ). This information may be useful to a wide audience. Target groups for the Webtool include, but are not restricted to, policy makers, investors and companies in search of a new business location.

The indicators, or the data from which an indicator was derived, originate from various sources, most notably the ESPON and Eurostat databases. Other sources include the European Environment Agency (EEA), Doing Business and social media (Twitter, YouTube and Flickr).

More info on the CityBench project can be found here.

How to use

Video tutorials are available reflecting various use-cases and interrogation paths.

FAQ

<u>What is a Larger Urban Zone (LUZ)?</u>

The CityBench Webtool is aimed at benchmarking / comparing Larger Urban Zones (LUZ) rather than cities. A LUZ is "…an approximation of the functional urban area extending beyond the core city"; the core city being "…the city as defined by its administrative and/or political boundaries" according to Eurostat. In practice this means that a LUZ is virtually always larger than the city it contains and that values collected for LUZ level will, to a greater or lesser extent, differ from (core) city values.

What is NUTS 1/2/3?

According to Eurostat, “The NUTS classification (Nomenclature of territorial units for statistics) is a hierarchical system for dividing up the economic territory of the EU…”. The most commonly used levels are:

- NUTS1: major socio-economic regions (not used in Webtool)

- NUTS2: basic regions for the application of regional policies

- NUTS3: small regions for specific diagnoses

In case of the NUTS3 indicators in the Webtool, the NUTS3 values have been aggregated to metropolitan region level, as an approximation of the LUZ level.

In case of the NUTS2 indicator in the Webtool, the NUTS2 values have been assigned to each LUZ it contains.

Indicators which have been derived from NUTS 2/3 data are marked as such in the Webtool.

What is a metropolitan region (MR) and how does it compare to a LUZ?

A metropolitan region is an aggregation of one or more NUTS3 units and is considered the NUTS3 approximation of a larger urban zone. Please note that although many MR - LUZ pairs show a good correspondence, correspondence is poor in some cases (MR much larger than LUZ or vice versa).

Which LUZ are included and why?

171 LUZ are currently included in the Webtool. They were selected according to the following criteria:

Inclusion of all European cities that are part of another online city comparison tool, the OECD Metropolitan Explorer.

Inclusion of capitals of European countries not included in OECD to cover ESPON Space (i.e. EU27 plus Iceland, Liechtenstein, Norway and Switzerland).

Inclusion of all remaining LUZ with a population number of > 400,000.

Inclusion of additional LUZ in underrepresented countries, i.e. countries with only one or even zero LUZ with a population of > 400,000 by including a second city, provided that its population exceeds 200,000.

Why isn’t my city included in the Webtool?

If you are in a city which does not comply to any of the criteria specified above, it is currently not included. However, as indicator data availability increases, more and more (smaller) LUZ will be added to the Webtool, possibly including your city as well.

What do the various indicator themes represent?

The indicators are grouped according to 'themes', each covering a different topic.

- Theme Economy and Population: provides general information about a LUZ.
- Theme Connectivity: provides information on the degree to which a LUZ is connected to 'the rest of the world'.
- Theme Demography: provides information on the build-up of the LUZ population.
- Theme Social Media: includes indicators derived from data harvested from social media.
- Theme Investment Climate: provides indicators potentially of interest to investors.
- Theme Environment: provides an indication of the environmental / air quality of a LUZ.
- Theme 'Smartness': provides an indication of the degree to which a LUZ is prepared for future developments.

Why are some indicators not available for the city or cities I'm interested in?

Because of gaps in the source data, unfortunately completeness for most indicators is not 100%. This means that they do not cover all LUZ included in the Webtool. A small icon in the bar above an indicator displays its degree of completeness. It looks like this: . Dark blue represents the proportion of LUZ for which indicator values are available.

How can I compare two or more cities?

The CityBench Webtool allows the comparison of up to four LUZ, based on up to three indicators. To select LUZ for comparison, just click on the buttons below the map of Europe: 'Select the 1st (2nd/3rd/4th) city'. Also, select one, two or three indicators on which the comparison will be based by clicking 'remove' and/or 'Select a 1st/2nd/3rd indicator'. Once you have selected one or more LUZ and one or more indicators, click 'Compare'.

Rather than comparing between cities, can I compare a city to other geographic units?

Yes, you can. The CityBench Webtool has the option to compare LUZ indicator values to the average value of the country it belongs to or to the European average.

How is the similarity between cities calculated?

By similarity we mean the degree to which indicator(s) values are comparable.

Two processing steps are involved in deriving similarity values:

1) All indicators are normalized, i.e. for each indicator the lowest LUZ value is assigned a value of 0 and the highest a value of 1.
2) For one, two or three normalized indicators, the 'Euclidean' distance between the indicator(s) values for one LUZ and one or more other LUZ (or country / European average) is calculated. Euclidean distance (or Euclidean metric) is the distance

between two points as measured with a ruler and using the Pythagorean formula to derive the (metric) distance. The more indicators are selected, the more dimensions are involved in the calculation (one dimension for each indicator added).

For example, if you selected 3 (normalized) indicators (a, b and c) and two LUZ (A and B) then the similarity between these LUZ is calculated as follows: sqrt ((aA - aB)2 + (bA - bB)2 + (cA - cB)2).

If the calculated distance is 0, there is full similarity; if the distance is equal to the sqrt(number of dimensions), there is maximal dissimilarity. In other words, the closer the number is to 0, the higher the degree of similarity between the selected LUZ, based on the chosen indicators. Conversely, the closer the number is to 1 (or even exceeds 1, which occurs as more indicators are added), the more different the LUZ are.

What is shown on the different tabs: Map / Radial?

Map

Being the main tab, the Map view shows how all included LUZ compare based on one, two or three indicators. The LUZ currently selected serves as the reference LUZ, to which the others are compared. The size of the circle representing a LUZ corresponds to the similarity of each LUZ to the reference LUZ: the larger the circle, the greater the similarity (and vice versa). See for an explanation of the calculation of similarity: How is the similarity between cities calculated?

Radial

The Radial view provides, for one, two or three indicators, an alternative way to show the similarity between the LUZ you selected and all other LUZ. The reference LUZ is located in the center of the graph, with the other LUZ placed around it, more or less at their actual geographic location relative to each other. The larger the distance between reference and other LUZ, the less similar they are, based on the chosen indicator(s). The LUZ circle size represents its actual value for the indicator selected for display. See for an explanation of the calculation of similarity: How is the similarity between cities calculated?

4. European Urban Benchmarking Webtool Demonstration Report

Together with ESPON CU, the TPG developed a Webtool demonstration protocol targeting stakeholders and potential users of the tool (see Table 4-1). This demonstration protocol includes the necessary steps, to be followed by users, that demonstrates the capabilities of the tool. The demonstration protocol was used in the official testing of the Webtool.

Table 4-1. Demonstration protocol.

<table>
<tr><td>Brief description of the ESPON CityBench project</td><td>This ESPON project delivered the CityBench Webtool. This Webtool allows benchmarking cities against other similar cities, around issues such as demographic challenges, economic challenges, social disparities/polarization, urban sprawl and greenhouse gas emissions.

This benchmarking Webtool creates opportunities for the forming of innovative learning networks among cities. It will also provide a first indication of the suitability of financial engineering instruments in place-based policymaking for addressing risks and opportunities.

Eduardo Dias, representing Geodan Holding (NL) is the Lead Partner of this ESPON project.

Additional information about this ESPON project is available on the ESPON website: http://www.espon.eu/main/Menu_Projects/Menu_ScientificPlatform/citybench.html</td></tr>
<tr><td>Purpose of this testing</td><td>To elicit stakeholders' views on the practical potential of the CityBench Webtool to provide a first indication of the suitability of financial engineering instruments in place-based policymaking for addressing risks and development opportunities in cities. In particular it is meant to receive feedback on the possible ways to develop/improve the Webtool further in the perspective of creating opportunities for the forming of innovative learning networks among cities.</td></tr>
<tr><td>Target audience</td><td>Policymakers, practitioners and public and private investors that need to be informed about economic, social and environmental sustainability aspects of cities for decision making processes. Those who have been or may be involved in implementing or supporting the implementation of Financial Instruments for urban development including:

• Member States
• Other European Programmes promoting sustainable urban development
• Urban Development Fund Managers
• EC DG Regio officers
• Project promoters</td></tr>
<tr><td>The "quickscan" CityBench Webtool: expected role and feedback from the test drivers</td><td>Time expected to carry out the testing: 30 minutes, including the submission of the questionnaire.

1. Go to https://www.google.com/intl/en/chrome/browser/
2. Click "Download Chrome"
3. In the next screen, click "Accept and Install" (you might want to deselect "Set Google Chrome as my default browser" here)

Note: Please be aware that you do not need to have "administration rights" on your computer to perform step 2 and 3. In practice you will be able to use Chrome (without</td></tr>
</table>

<table>
<tr><td></td><td>installing it). During the testing phase the tool is running in Chrome but in the final version it will also work in other browsers.

4. Access the CityBench tool In Chrome through the following link: http://espon.geodan.nl/citybench3/
5. A wizard will guide you in using the tool through simple steps.
6. You may exit the wizard at any time by pressing 'skip', which will redirect you to the main page of the Webtool: a map of Europe with the point locations of the included cities. Then, you will be able to explore the tool according to your interest.
7. A FAQ section includes detailed information on a particular topic.
8. Explore the tool and try to use the various functionalities for 10-15 minutes. Some inspiration on 3 tasks scenarios related to different types of potentials users are described in the section below and might help you to use the tool and to assess its usefulness.
9. Fill in the questionnaire and submit your feedback by 22 April 2014 at the latest by using the following link: https://docs.google.com/forms/d/1nEGtRtYvUoVRSawRrfSL-kbzBcJsd4tsAtFZMMXz00Y/viewform</td></tr>
<tr><td>Roles that you might consider when testing the CityBench Webtool that could support you in the testing process</td><td><u>Scenario 1</u>
• You are:
A private investor assessing urban locations within the EU so as to help ensure sustainable investments for a new retail center.
You have available space in Barcelona, Münster, Athina and København. Which of these cities will you choose based on relevant indicators?
• You want to:
Compare and contrast multiple possible locations for your project in order to help assess the most viable location, based on the project premises: population, accessibility, purchasing power (or energy costs).
• Instructions:
1. Go to URL of Webtool (http://espon.geodan.nl/citybench3/#).
2. In Step 1 of the CityBench wizard, select Barcelona, Münster, Athina and København.
3. In Step 2 of the CityBench wizard, select for example three out of the following indicators: Resident population, Potential accessibility by road, GDP per inhabitant and/or Electricity prices for industrial consumers.
4. In Step 3 of the CityBench wizard, click Compare.

<u>Scenario 2</u>
• You are:
A practitioner from a public authority assessing urban locations within the EU so as to help ensure sustainable investments.
You have 10M € to fund projects to combat unemployment with a top grant of 200.000 €. What cut-off levels for unemployment rate and GDP per inhabitant would you specify for cities to be eligible for these grants?
• You want to:
Specify cut-off levels for certain indicators, as to target a call for funding to cities most in need, and exclude cities that are sufficiently developed, while taking into account the total budget you have to spend and the maximum amount to spend in one city. Specifically, you want to set cut-off levels for unemployment rate and GDP per inhabitant.
Instructions:</td></tr>
</table>

<table>
<tr><td></td><td>1. Go to URL of Webtool (http://espon.geodan.nl/citybench3/#).
2. In Step 1 of the CityBench wizard, click Skip to close the wizard and proceed to the Webtool.
3. Select one city to show the similarity to (the reference city).
4. From the available indicators, select: % of persons unemployed and GDP per inhabitant.

<u>Scenario 3</u>
• You are:
A regional/urban practitioner in Glasgow (which might be termed a second tier city – M1), based at the City Council in either an executive or political function, at the development and regeneration services.
Which cities are similar to Glasgow in terms of economic base, population size, other indicators or typology? Which are Glasgow's weaknesses and strengths compared to them?
• You want to:
Examine Glasgow's endogenous potentials and thus possibilities for new development opportunities and ideas.
More specifically, you want to explore and reflect on Glasgow's position relative to other similar cities, in terms of economic base, population size, other indicators or typology. Identify what advantages and disadvantages Glasgow has relative to these cities.
Instructions:
1. Go to URL of Webtool (http://espon.geodan.nl/citybench3/#).
2. In Step 1 of the CityBench wizard, select Glasgow.
3. In Step 2 of the CityBench wizard, select up to three indicators of your choice. Optionally, select as Typology: second tier metro region.
4. In Step 3 of the CityBench wizard, click Similarity.
5. In the Map view, besides Glasgow select up to three other cities you are interested in and click Compare.</td></tr>
<tr><td>Where and when to send your feedback</td><td>Please fill in the web questionnaire by 22 April 2014: https://docs.google.com/forms/d/1nEGtRtYvUoVRSawRrfSL-kbzBcJsd4tsAtFZMMXz00Y/viewform Additional comments should be sent by email to the Lead Partner of the ESPON CityBench project, Eduardo Dias, by the 22 April 2014.

Email: eduardo.dias@geodan.nl, T: +31 20 5711 323

Please put Sandra Di Biaggio (sandra.di.biaggio@espon.eu) of the ESPON Coordination Unit on copy.</td></tr>
<tr><td>Technical matters</td><td>Please raise these with Eduardo Dias (as above)</td></tr>
<tr><td>How your comments will be used</td><td>The comments will be collated by the project team and will be drawn into a series of recommended next steps for the consideration of the ESPON Coordination in further developing the tools and concept. In addition and if possible some adjustments will be implemented in the final version of the Webtool, which will be published on the ESPON website by May 2014.</td></tr>
</table>

The ESPON CU disseminated the protocol to stakeholders and other potential users, who carried out the steps outlined and, after that, provided feedback on their experience using the tool by completing an evaluation survey. Between 11 March 2014 to 22 April 2014, 18 reactions were received. The detailed response to each question is provided in Annex VII. We observed the following trends:

Usability: most respondents quickly learned - and remembered - how to work with the Webtool, enjoyed it, found it relatively easy to use and consistent and felt confident about using it. The majority thought it worked well in their browser, found the functions well-integrated and expects to use it frequently and recommend it to a colleague or friend. There were mixed feelings regarding the ease and speed to recover from mistakes and the degree to which the Webtool works as desired.

Usefulness: comparing one / one's own city to other cities based on similarity and viewing that in Map and radial view was considered very useful and interesting, as were the typology and region based filtering option and the graphs comparing up to 4 cities. The importance of accessing the different years for the indicators to see trends was confirmed. Generally the "similarity" concept, as used in the Webtool, was easily understood.

Needs or additional functionality: a majority stated that the Webtool is complete for a quick scan of cities, but would welcome the options to download data and images and to link to a map or radial view using email, Facebook or twitter. A question about showing real values rather than "compare" (i.e. relative) values of indicators yielded an equal amount in favor of either option.

Please see Annex VII for an overview of individual wishes regarding additional functionality and positive aspects / aspects to be improved or changed.

5. Dissemination & communication

During the project timeframe, effort has been reserved by the TPG to undertake disseminations works and these have been fruitful, since interest from public and private bodies have risen.

To ensure that the Webtool development fits the stakeholders' requirements, several actions for communication, interaction and dissemination have been deployed during the project timeframe. Meetings and workshops have been attended to present the project and get users' requirements and suggestions, an info website and a stakeholders social network have been tested and several testing sessions and evaluations for the different versions of the Webtool have been carried out in order to get continuous feedback from stakeholders through discussions, social media and questionnaires.

5.1 ESPON Meetings

ESPON activities concern different projects, organizations and partners. Several meetings took place in order to coordinate efforts between projects (and engage with stakeholders), the CityBench TPG participated in:

- ESPON info session - Brussels -19 November 2013
- USESPON BBSR-Workshop "The role of cities in the EU2020 strategy" – Berlin - 26 September 2013
- Territorial Co-operation for growth and jobs - Dublin - 13 and 14 June 2013
- Technical meeting on ESPON tools - Paris – 17 May 2013
- Several steering committee meetings – ESPON CU, Luxembourg.

5.2 Conferences and Workshops

During the timeframe of the project, activities and foreground have risen as a powerful and fruitful scientific source of knowledge. Partners have presented CityBench and related work in research conferences and workshops, such as:

- Citizen Science and Smart Cities. 6 February 2014 – IES – JRC, Ispra
- Geodesign Summit Europe. 18 – 19 September 2013 – Geofort, Herwijnen, The Netherlands
- Conferencia ESRI España. 2 -3 de October 2013 - Madrid, Spain
- The 16th AGILE International Conference on Geographic Information Science. 14-17 May 2013 – Leuven, Belgium

5.3 Research Papers

The TPG documented part of their research as papers and communications in scientific outlets, such as:

- Elizabeth Kalinaki, Robert Oortwijn, Ana Sanchis Huertas, Eduardo Dias, Laura Diaz, Steven Ottens, Anne Blankert, Michael Gould and Henk Scholten. 2014. "CityBench: A geospatial exploration of comparable cities" 17th AGILE Conference on Geographic Information Science. 2014, Castellón, Spain. 3-6 June

- Sven Casteleyn, Garrigós, I., Mazon, J.N. Ten years of Rich Internet Applications: a systematic mapping study, and beyond. ACM Transactions on the Web, Volume 8, Issue 3, 42 pages, 2014

- Ana Sanchis, Laura Díaz, Joaquin Huerta, Michael Gould. 2014. "Crowdsourced public participation of city building". In: Henk Scholten, Eduardo Dias, Danbi Lee (Eds) Geodesign by Integrating Design and Spatial science. Springer (Forthcoming)

5.4 Communication beyond project completion

Beyond the project lifetime, the TPG will carry out dissemination activities, such as:

- Present work on conferences and workshops (AGILE);
- Publish additional results in scientific outlets;
- Link to Webtool in partners websites (Geodan, Geotec, UJI);
- Press release after project finalization (through UJI communication Department).

C CityBench Scientific Report

The scientific report was converted to a paper format and submitted for review in a scientific conference. The conference was the AGILE conference, one of the leading European events in the Geographic Information Science field. The blind peer-review system of the conference selected this work for inclusion in the following scientific publication:

See: http://www.springer.com/earth+sciences+and+geography/geographical+information+systems/book/978-3-319-03610-6

Publisher: Springer

Title: Connecting a Digital Europe through Location and Place

Series: Lecture Notes in Geoinformation and Cartography

Editors: Huerta, Joaquín, **Schade**, Sven, **Granell**, Carlo

ISBN 978-3-319-03610-6

Due: June 30, 2014

CityBench: A geospatial exploration of comparable cities

Elizabeth Kalinaki[1], Robert Oortwijn[2], Ana Sanchis Huertas[3], Eduardo Dias[1, 2], Laura Diaz[3], Steven Ottens[2], Anne Blankert[2], Michael Gould[3] and Henk Scholten [2]

Abstract. In many city comparisons and benchmarking attempts, scores are purely one-dimensional with results split accordingly for each dimension. We propose a methodology for comparing cities on multiple dimensions implemented as a map-centric web tool based on a pairwise similarity measure. CityBench web tool provides a quick-scan geographical exploration of multidimensional similarity across European cities. With this dynamic method, the user may easily discover city peers that could face similar risks and opportunities and consequently develop knowledge networks and share best practices. This web tool is destined to provide economic/financing institutions', local governments' and policy makers' in Europe and beyond decision making support.

Keywords: city comparison, similarity measure, geo-visualization

1. Introduction

Why do similarly structured cities behave differently in socio-economic terms? And which cities across Europe might be similar to my city of interest if I have a number of different life aspects? This paper introduces the CityBench web tool an interactive, geospatial exploration of comparable (European) cities based on a pairwise similarity measure and a map-centric interface. The tool offers a quick and dynamic overview of the most similar cities for a user selected number of life aspects. The CityBench web tool can show cities apparently similar based on selected indicators and may help to identify geographic trends in the location of similar cities hence bringing GIS-like pattern analysis to an interactive web environment.

Often, economic/financial institutions such as the European Investment Bank Municipal and regional Unit (EIB-MRU) and EU policy makers wishing to obtain insight into regional trends in urban development or to explore effectiveness of regional urban cooperation programs, investors in search of a new business location as well as regional/urban practitioners searching for best practices or their own endogenous potentials and opportunities for co-operation, network or cluster forming carry out benchmarking studies. The studies in many cases like Helgason (1997), Lam et al. (2010) and Groenendijk (2004) involve identifying other 'peers' to compare a particular self with. The benchmarking and comparison process can benefit from technological developments offering new possibilities to create real-time, interactive map-centric visualization that give valuable insight.

For effective benchmarking, the regions or cities need to be clearly defined since comparing a capital city with a province or a municipality is not very useful due to the uniqueness of the different structures. Batty and Longley (1994) explain that definitions of cities rely upon definitions of boundaries or delineations although such definitions are never comprehensive. If cities are highly dynamic landscapes according to Ramalho and Hobbs (2012), their definitions need regular revisions to include functional urban regions- the large areas around the proper boundary of the

[1] Vrije University Amsterdam, Amsterdam, the Netherlands

[2] Geodan, Amsterdam, the Netherlands

[3] Jaume I University, Castellon, Spain

central city (Vasanen 2012), (Hall 2009) and (Parr 2005). This and other similar descriptions highlight delineations such as ‘inner city’, ‘metropolitan region’ and ‘central business district’. In this study, these delineations are normalized by aggregation or transformation into approximations of the recently proposed city delineation by Dijkstra and Poelman (2012) called a ‘larger urban zone’ (LUZ). In defining a LUZ, all grid cells covering the city with a density greater than 1500 inhabitants per square km are selected. From these, clusters with a minimum of 50,000 inhabitants are defined as ‘urban centers’. Then, all municipalities with at least half their population inside the urban center are selected as candidates to form the city. The LUZ is finally defined ensuring that there is a link to a political level with at least 50% of the population living in an urban center and at least 75% of the population of the urban center lives in a city (Dijkstra and Poelman 2012). Although different parts of the defined city boundaries can have very different patterns, European entities are proposing the LUZ as the most suitable city or region definition for comparability at European level and the backbone for city statistical data (EC, 2004).

The CityBench web tool was proposed to efficiently integrate existing and new urban indicator data at the LUZ level available from the ESPON 2013 programme[4] and Eurostat[5] for a geospatial city exploration and identification of geographic patterns and to provide stakeholders the necessary functionality to visualize city status and perform benchmarking. The tool’s purpose is to enable multi-scale, multidimensional representation of cities for finding similarity and geographical trends for economic/financing institutions’, local governments’ and policy makers’ decision making support.

2. Related work

Ammons (1999) emphasizes that choosing an appropriate benchmarking technique and carefully applying it are both essential for benchmarking success. City benchmarking as defined by Luque-Martínez and Muñoz-Leiva (2005) should allow cities to implement the most effective practices and capacities learned from other cities in order to improve their actions in what they offer. This statement is in agreement with Ammons (1999) who suggests that the idea behind benchmarking should not focus on how an organization stacks up but should be captured by what the benchmarked cities learned and how they better themselves from the lessons learned from the benchmarking.

Many benchmarking and city comparisons studies have been performed. Luque-Martínez and Muñoz-Leiva (2005) offer ground work on both benchmarking in general and city benchmarking while Amelang (2007) combined two popular approaches by historians (the thematic approach, in which a single issue or series of issues are examined in relation to more than one urban area and another that focused on the cities themselves with an aim to identify similarities and divergences in individual urban trajectories by directly contrasting one city with another) to understand the metamorphosis of Barcelona.

Holloway and Wajzer (2008) explain that there is no single methodology for conducting city benchmarking exercises. The authors identified ‘dimensions’ (-the particular facet of a city to be compared e.g. quality of life or economic competitiveness), ‘indicators’ (- the measure of performance for each dimension of study e.g. for quality of life, a stable political environment could be considered as one dimension of quality of life and an indicator for this may be the level of crime) and ‘scoring and ranking systems’ (the methods used to analyze, compare or benchmark dimensions, indicators and cities) as the three elements evident in most studies.

[4] http://www.espon.eu/main/

[5] http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/

Tuan Seik (2000) used many of these indicators such as health and family life, politics, religion and leisure to assess the quality of life in Singapore. In Luque-Martínez and Muñoz-Leiva (2005), approximately 180 indicators were gathered in different dimensions such as environmental management, housing, accessibility to evaluate the region of Andalusia, using Granada as the reference city. Baum (1997) performed a similar study with cities in Australia to test and explain the social polarization of Sydney. On a global scale, Lippman Abu-Lughod (1995) compared Chicago, New York and Los Angeles using a variety of indicators to specify a model for how forces generated at international level affect each city. Taylor and Walker (2001) used Friedmann's list of world cities (1986) and revised version Friedmann (1995) to produce a complex interweaving of hierarchical tendencies with distinct regional and interregional patterns from a multivariate analysis of their 'Service Complexes'.

The comparative study in Beaverstock et al. (2001) revealed a network model of inter-city relations which proved that there is more to benchmarking and city comparisons than which city comes first or last. It is noticeable in many studies that these comparisons are done to serve a specific purpose. European cases like Turok and Mykhnenko (2007) used city comparison and benchmarking to assess how 310 European cities had economically evolved over the period 1960-2005 revealing a general slowed growth over the last few decades at the time of their study. In Kasanko et al. (2006), a comparative analysis using 5 indicator sets of 15 European urban areas revealed that the structure of European cities had become less compact while Sager (2003) compared the performance of 'important' European cities in their function as tourist destinations to measure their gains and losses of the tourist market share in Europe.

Using indicators to benchmark cities or perform city comparisons has numerous advantages although as Holloway and Wajzer (2008) point out, city benchmarking has some limitations that undermine their validity for measuring and monitoring performance. These include the integrity and compatibility of data among cities, the overstatement of the cause and effect relationship between indicators and city performance and the subjectivity of the analysis and conclusions.

Most of the studies available use quantitative data from regional or governmental statistics offices and financial institutions and in most if not all the work involving indicators like in Luque-Martínez and Muñoz-Leiva (2005) and OECD (2012), the scores are purely one dimensional with results split accordingly for each dimension. If we want to determine the most similar cities to a reference city (LUZ), we need to supplement the above methodologies with a similarity measure. There are various similarity measures as described in Deza and Deza (2013). Preoţiuc-Pietro et al. (2013) performed a physical city similarity analysis in which the authors represented each city as a point of vector space to compute 'pairwise' similarities between the cities using the cosine similarity and later quantified their results using a Kendall Tau rank correlation coefficient to find for each city the most similar cities. The cosine similarity is one of the techniques used in Seth et al. (2011) to determine similar cities that are not necessarily geographically close. In ecology and biology science, use of similarity measures like the Cosine similarity and Euclidean distance similarity are used successfully to determine similar entities. For example in (Luo et al. 2001) the similarity between gene expression patterns is measured by computing the Euclidean distances for each pair of samples based on log-transformed ratios across all of the genes. A similar example by Jain et al. (2000) uses the Euclidean distance between two corresponding "FingerCodes" for fingerprint matching and in Mane et al. (2010) for face recognition.

The choice between a Cosine and a Euclidean distance similarity is dependent on the purpose of the similarity measure. While the two are closely related, using 'Euclidean distance' is most useful to us when determining the most similar cities to our given reference city.

3. Methodology

3.1. CityBench similarity measure

One of the oldest and most influential theoretical assumptions is that perceived similarity is inversely related to psychological distance (Ashby and Ennis 2007). The authors explain that our mental representations of objects, concepts, positions on issues typically vary on a variety of psychological dimensions. The numerical values of a particular ‘percept’ on each of these dimensions can be interpreted as the coordinates of this percept in a psychological feature space.

The percepts close together are perceived as similar and percepts far apart dissimilar. Following the above principals, the CityBench similarity measure defines the current cities’ multiple scores as geometric vectors and compares them across several indicator dimensions. It’s not enough to rank the cities and their respective scores in each indicator dimension as in Luque-Martínez and Muñoz-Leiva (2005), what is important for CityBench are the cities that come closest to a specific city that scores k in any number of indicator dimensions.

We define k as the combined score of the reference city on the one, two or three indicators selected by the user. The specific city with score value k is most similar to itself and then similar to cities that score k +/- 1. The similarity value (Scv) for any city against the reference city is given by equation [3.1] where c is the score of any city other than the reference city in any indicator dimension i and v is the reference city value.

$$S_{cv} = \sqrt{\sum_{Ii=1}^{n} (C_{Ii} - V_{Ii})^2} \quad [3.1]$$

For example to find the most similar cities to Luxembourg as the reference city using ‘GDP’ as A and ‘Ease of business’ as B indicator dimensions, we normalize the values for all cities for both A and B by subtracting the minimum value in that indicator dimension from the original city indicator value and dividing that by the difference of the maximum and the minimum values.

With these normalized values, we then compute $(CA\text{-}VA)^2$ and $(CB\text{-}VB)^2$ for each city which later provides each city’s Scv after computing the square root of their summation. If the list of cities and their respective Scv values provided by the equation above are displayed effectively, the values are capable of showing regional patterns based on the values of one indicator or a combination of indicators.

However, as with any statistical data, outliers cannot be avoided. Their sources can be traced to errors of measurement, faults in execution and or intrinsic variability (Woolley 2013). The presence of even just one outlier can offset the capacity to visualize variability of the remaining dataset therefore finding a best fit classification method can be used to observe the variability. Three data classification methods (Quintiles, Jenks natural breaks and Equal interval) and two alternative methods (ranking the values and using the original values) were explored to assess their effectiveness in displaying the similarity values.

In Table 1, samples of this exploration are shown for Luxemburg, Amsterdam and Plovdiv using original GDP (€) values (75,191, 40,568 and 3,728 respectively) as an indicator example. All the samples yield good visual regional patterns for cities with intermediate values (Amsterdam) while the high score outlier (Luxembourg) stands out in Quintiles and Jenks natural breaks but not as much in Equal intervals. For the low score outlier (Plovdiv), some regional patterns are only observable in the latter.

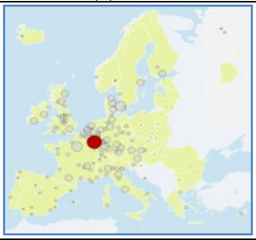
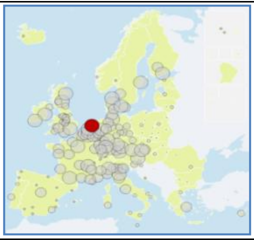
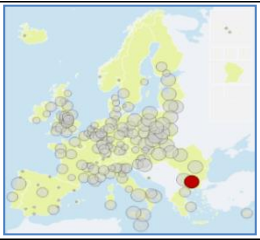
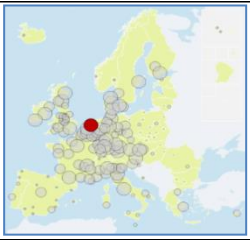
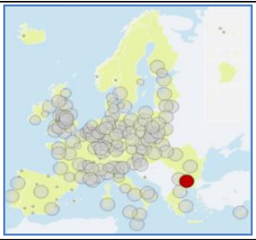
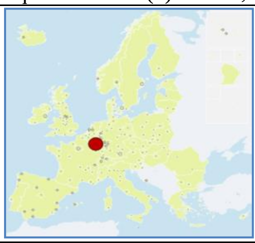
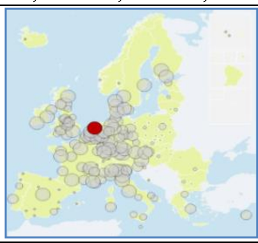
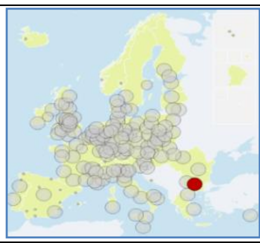
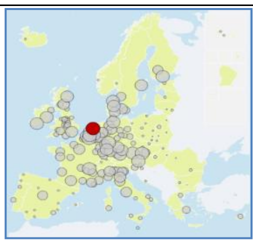
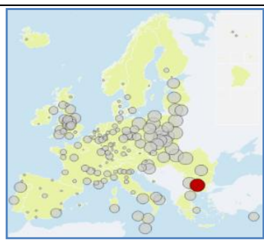
Luxembourg (highest GDP)	Amsterdam (in-between GDP)	Plovdiv (lowest GDP)
Quintiles (5): 0 - .57,0.57 - 0.65, 0.65 - 0.72, 0.72 - 0.84, 0.84 - 1.0		
		
Jenks Natural Breaks (5): 0 - 0.16, 0.16 - 0.30, 0.30 - 0.43, 0.43 - 0.61, 0.61 - 1		
		
Equal Intervals (5):0 – 0.2,0.2 – 0.4,0.4 – 0.6,0.6 – 0.8,0.8 – 1.0		
		
Based on ranking order (171)		
		

Table 0-1: Overview of similarity maps derived from different classification methods (using GDP (€) per head as indicator). The red dot represents the reference city against which similarity is measured and shown

The high score outlier (Luxembourg) in the ranking method becomes part of the Central Europe, Ireland and Scandinavia cluster while the low score outlier (Plovdiv) joins the Eastern Europe cluster.

This classification does not distinguish between the high score outlier and the middle score cities implying that Luxembourg with original value 75,191 and Amsterdam with original value 40,568 are almost identical. While this is good for reflecting an order of magnitude common to Central Europe, Ireland and Scandinavia, it is difficult to visually 'explain' the difference in the original values. Based on the above findings, to allow for visualizing regional patterns and to visualize effects of outliers, we opted to use both position ranking and the unique values (Table 2).

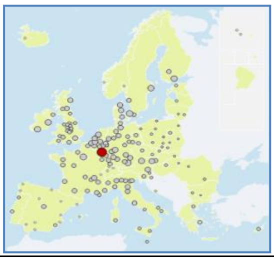
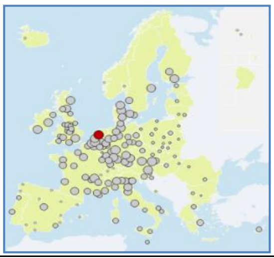
Luxembourg (highest GDP)	Amsterdam (in-between GDP)	Plovdiv (lowest GDP)
Unique values (Non-classified) 0.0-1.0		
		

Table 0-2: Similarity maps of the 3 cities using unclassified values from the pairwise measure

4. Implementation

4.1. Data

An extensive data audit was performed especially focusing on ESPON[6] and Eurostat[7] data. The audit assessed whether existing datasets, collected at various geographic levels (LUZ, NUTS 3, and Metropolitan Region) could be used for city benchmarking. Other data termed licensed / commercial and open / volunteered was considered as well. The data audit resulted in a provisional list of indicators which served as a basis input for CityBench.
In order to enable integration of these indicators into the tool, an extraction, transforming and loading (ETL) module was created for the purpose. For new indicators from other sources such as social media data, a module tailored to data of this nature was also developed. The ‘CityBench indicators ETL’ tool was developed to receive and consume data arranged in spreadsheets and load it into the CityBench database as new indicators, along with their ‘meta’ data. Fig. 1 below is a schematic representation of the CityBench ETL tool.

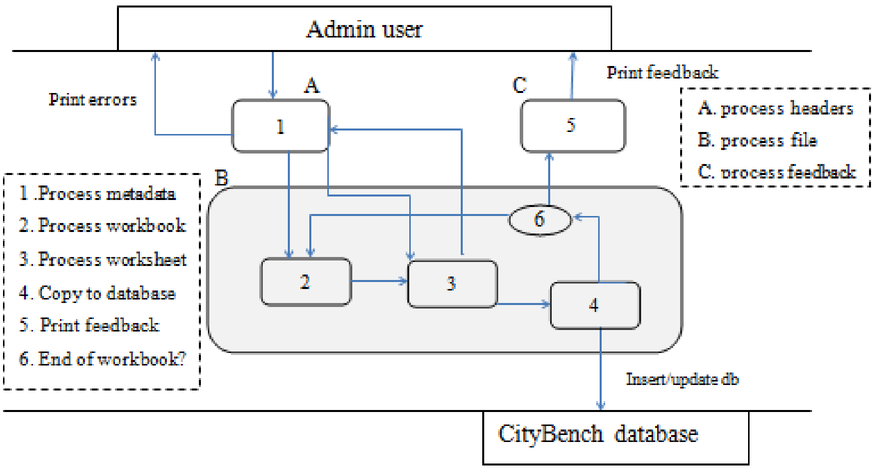


Fig. 9 CityBench ETL process flow

The CityBench ETL tool contains several subroutines for processing an uploaded file. The file is first processed by the “process headers which method ensures that

[6] http://www.espon.eu/main/

[7] http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/

the metadata for all indicators in the uploaded file is error free and will not add duplicates in the database. If the method finds no errors, the file is forwarded to the 'process file' method, which loops through the entire file reading data and creating new indicators for CityBench otherwise the user is asked to upload a new error free file.
Within 'process file', the 'process worksheet' method handles the individual processing of data for a single indicator. This process creates new unique indicator identification if the new indicator has no official id from the data provider and finally creates a new table in the CityBench database for the new indicator data while keeping track of the current process status. When all indicators in a file are loaded into the database, feedback is returned to the user with the results of the ETL process.
Using an ETL tool for data management contributes to the uniformity of the data to be used by the CityBench tool achieved by enforcing the usage of a data input template, which ensures that each indicator is uploaded in a standardized format and updates to indicators are uniformly performed and traceable.

4.2. Software frameworks and architecture

To achieve an interoperable, scalable and future extensible platform, CityBench employs the client-server architecture (Berson 1996). HTML5 and JavaScript at the client and Java on the server allow us to take advantage of the Rest model design.

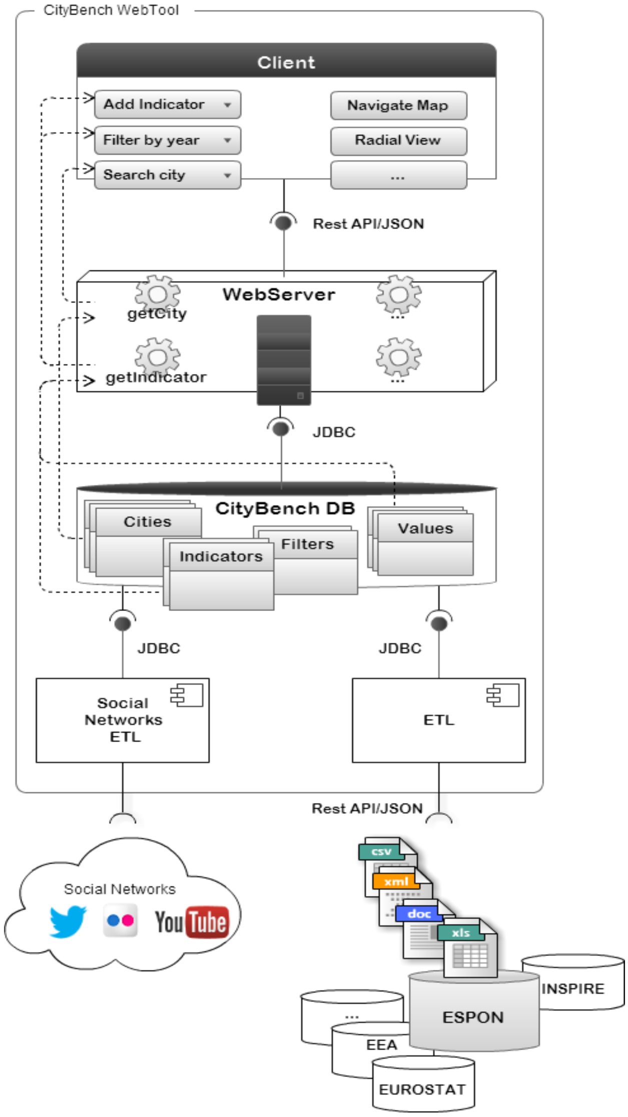


Fig. 10 technical architecture of CityBench

Employing jQuery.js[8] and d3.js[9] provide us a set of curated user interface interactions like the scaling bar and the later, creation and control over our dynamic and interactive graphical charts and diagrams (Murray 2013). Java allows implementation of restful web services (Martin 2009), (Burke 2009), (Sandoval 2009).These serve Asynchronous JavaScript and XML (AJAX) requests from the CityBench client allowing seamless data exchange between the client and server.
CityBench follows and extends the European directive INSPIRE (EC 2007) which defines a classical service-oriented three-layer architecture to include the required functionality in the form of network services and the client applications as a web portal (Tatnall 2005). In Fig. 2, this architecture is described from a technical perspective to consist of the web client at the application layer, the services responsible for data retrieval, the database for storage of all CityBench related data and the ETL modules.
From bottom up these layers are; A) the ‘Data and Services’ layer (external data fetching services and the ETL modules). This layer is configured to handle pre-processed data in Excel format from sources such as ESPON, EUROSTAT, European Environmental Agency (EEA), OpenFlights and social networks like Twitter (volunteered geographic information or VGI). B) The ‘Data’ layer whose

[8] http://jquery.com/

[9] http://d3js.org/

main component is a spatial database manages all web tool data needs. The database utilizes a custom data model well-adjusted to the needs of the CityBench tool and houses data from sources such as the ESPON database, metadata about the data and all indicators. C) The ‘Service’ layer services implement well-known and standard-based interfaces which provide data discovery, view, and upload functionality for the CityBench tool. Together they constitute the accessible CityBench API available to other applications with needs similar to the CityBench tool. D) The ‘App’ layer serves as the presentation layer comprising of a web portal which constitutes the visible CityBench tool and consists of a front-end user interface and a back-end login admin interface. Figure 5 and 6 are first CityBench prototype user interfaces with labeled functional areas.

5. Preliminary results

The CityBench web tool utilizes both a map centric and a ‘radial’ display for the results of the pairwise similarity measure. The sectioned interactive interface (Fig. 3) is meant to be easy to use, above and below the main view section are city selectors for choosing cities.

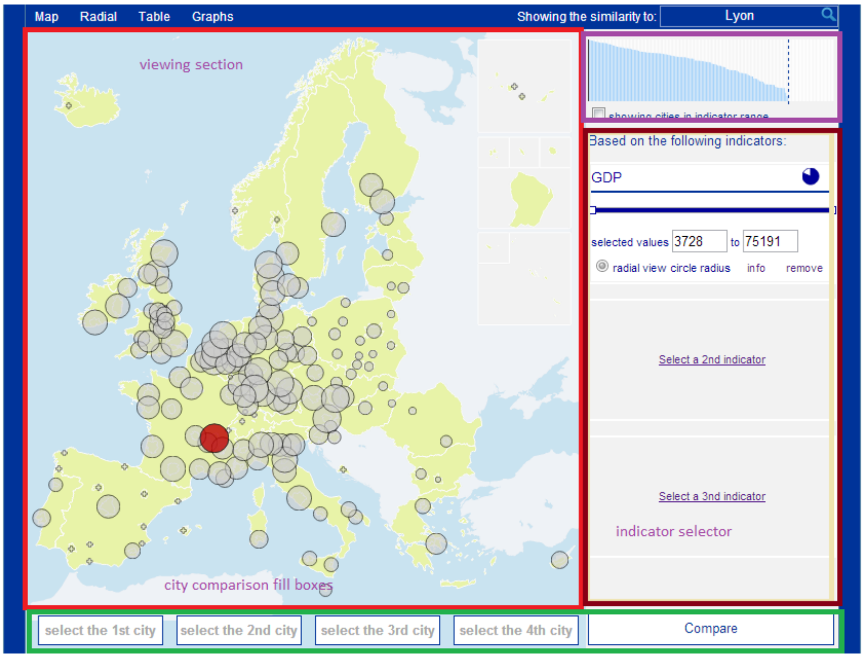


Fig. 11 CityBench interface

In both views (Figs. 3 and 4), there is a main view showing the results with each city represented as a circle whose radius is the pairwise measure score or a selected indicator value in Fig. 4. The reference city shown in red is naturally the most similar to itself and will always have the biggest radius in the main view and the highest bar in the list of cities to the right. Cities with no data in any of the selected indicators are shown with a ‘cross’ symbol in the map centric view (Fig. 3) and a white background in the bars widget located above the ‘indicator selector’ showing the distribution of values amongst the cities. In figure 3, the circle is placed in the middle of the geometry of the LUZ.

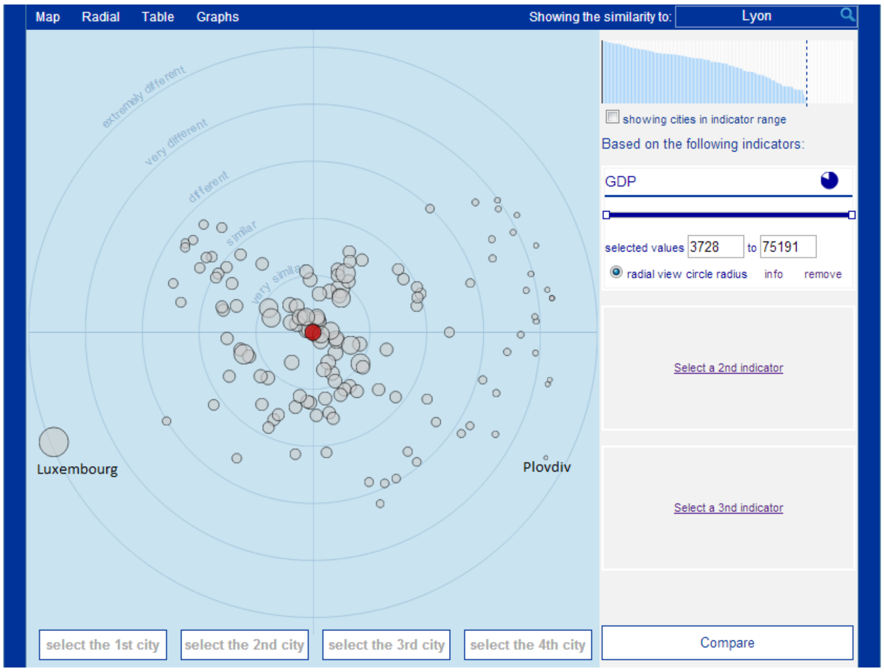


Fig. 12 CityBench radial view

The radial view in Fig. 4 assumes an arrangement where the cities seem to converge or radiate from the reference city in the centre. The closest to the centre, the more similar the city is to the reference city. Each city in this view is positioned approximately according to its actual geographic location, but also relative to the reference city. This view was inspired by a similar visualization by Tulp (2012). The circle represents a city's normalized value in a chosen indicator. In figure 4, Luxemburg has the biggest circle meaning that it has the biggest GDP value although it is very dissimilar to Lyon (in red). Plovdiv has the smallest circle and therefore the smallest GDP but also dissimilar to Lyon.

5.1. Available cities and indicators

A number of criteria were used in selecting the CityBench LUZ (Cities). The first criterion included all European LUZ with corresponding metropolitan regions in the OECD Metropolitan Areas database (MAdb). The MAdb has 268 metropolitan regions whose performances on a set of key indicators are visualized in the OECD Metropolitan Explorer[10] of which 114 were selected as starting point for populating CityBench. The second criterion included all LUZ from EU27+4 countries (i.e. EU27 plus Iceland, Liechtenstein, Norway and Switzerland). The third and fourth criteria supplemented the CityBench city list with LUZ from underrepresented countries. These countries are either small or relatively sparsely populated, or have a population that is distributed relatively equally over the country and may have only one or even no LUZ of which the population exceeds 400,000. Using a second city provided its population exceeds 200,000 ensured a proper representation of these countries in CityBench (criterion 4). The current list of cities may be available upon request.
Consideration for alternative data sources for the web tool resulted in use of social media for populating the CityBench web tool. This arose due to the need for relevant and up-to-date indicators as at the time of data auditing, ESPON and Eurostat the main sources of LUZ data and therefore indicators were in the process of reviewing and updating their databases. This meant that for several relevant indicators, recent data was not available. In addition, these databases do

[10] (http://www.oecd.org/statistics/datalab/metro-explorer.htm)

not cover all indicators potentially relevant to the CityBench mission. Geo referenced user-generated content is acquiring a fundamental role in a wide range of applications even though it currently represents only a small percentage. Tweets from Twitter for example may play a major role in response actions to emergencies (Lanfranchi and Ireson 2009) and (Núñez-Redó et al. 2011). Social media inclusion in CityBench utilized a custom implementation of social media extraction module out of scope of this research.
The search for data resulted in a set of indicators such as population and unemployment in an economy dimension, aging and old age dependency in a demography dimension, out flight route in a connectivity dimension. A complete list of these indicators and methodology to create them can be available on request.

5.2. City similarities

In both Sections 1 and 3, we mentioned that the CityBench tool would be helpful in the identification of geographic patterns using its similarity measure, the results indeed support this statement. We see a geographic pattern in figure 3 when viewing similarity to Lyon and notice a high concentration of similar cities to Lyon in Western Europe and a strip of dissimilar cities in Eastern Europe.

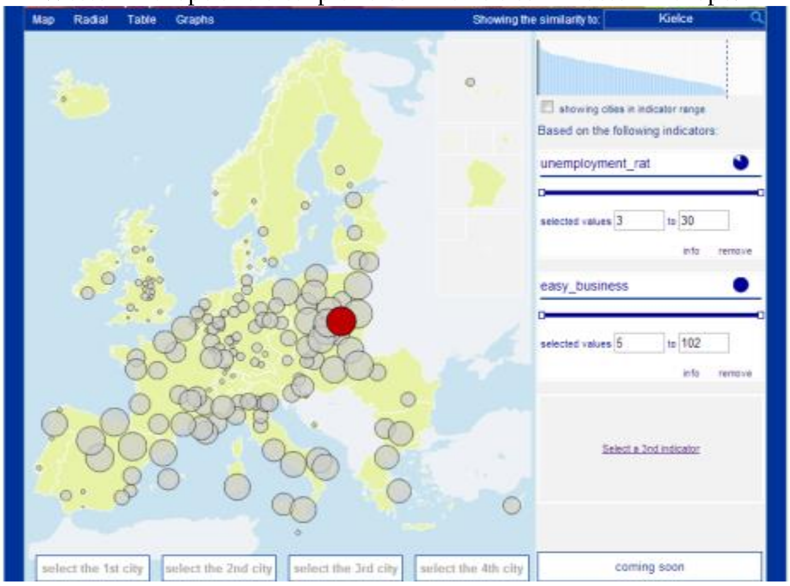


Fig. 13 similarity to Kielce

In Fig. 5, a regional belt covering most of Eastern Europe down to Spain and Italy is visible when viewing similarity to Kielce using unemployment ratio and ease of doing business indicators. Figs. 6 and 7 continue to stress similarities among western European cities in respect to western European reference cities with exceptions here and there such as in Fig. 8.
The CityBench web tool results are very data dependent. As such the tool is not meant to provide a static evidence of similarities or benchmarking, this means that certain users (researchers, investors, city officials, citizens) with specific questions can select the suitable indicators for their questions, and CityBench allows them to explore, show clusters or spatial patterns not earlier anticipated.

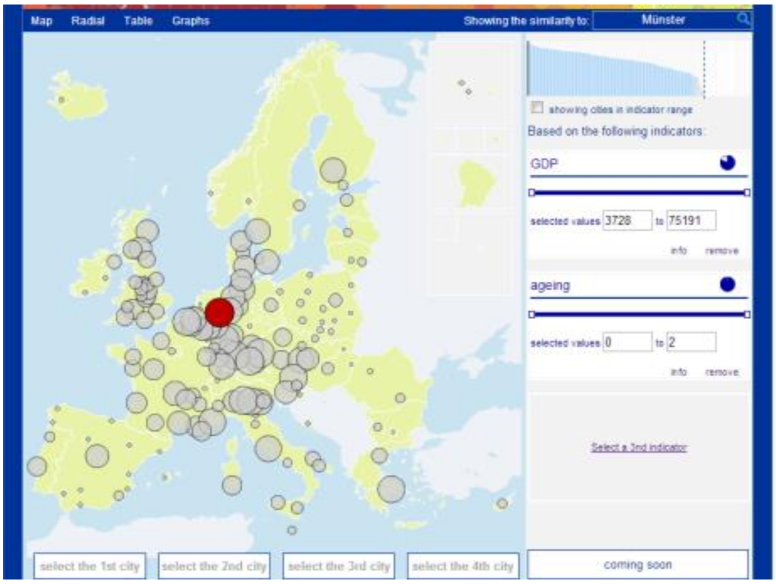


Fig. 14 similarity to Munster

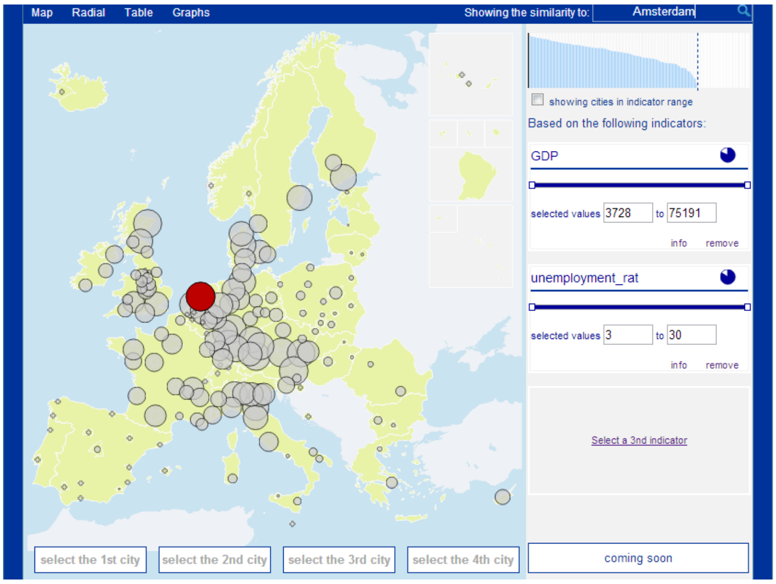


Fig. 15 similarity to Amsterdam

5.3. Which city is most similar to mine?

Using an example of 3 indicators (GDP, unemployment ratio, easy of doing business) with original values in table 3, below are the 10 cities most similar to each of the 6 cities Amsterdam, Cork, Helsinki, Luxemburg, Plovdiv and Madrid each used as a reference city as given by the CityBench similarity measure. We can already see relationships between cities in Table 4 whose top 4 are shown on the map centric in Fig. 8.

	city	gdp	unemployment ratio	easy business
max value		75191	30	102
	Amsterdam	40568	4.2	31
	Cork	44110	12.5	15

Helsinki	44074	6.4	11
Luxemburg	75191	4.4	56
Plovdiv	3728	8.5	66
Madrid	30040	16.1	44
min value	3728	3	5

Table 0-3 Original values of selected cities

The CityBench tool gives a quick scan into a city and cities that are most similar to it in respect to selected indicators. These results can be used by economists and policy makers to discover why similarly structured cities behave differently as it may be that cities that are similar may share similar risks and opportunities. These cities could learn from each other and also create a network of best practices.

Amsterdam		**Cork**		**Helsinki**	
Amsterdam	0.000	Cork	0.000	Helsinki	0.000
Linz	0.048	Dublin	0.029	Stockholm	0.037
Graz	0.055	Stockholm	0.203	København	0.088
Rotterdam	0.078	København	0.214	Frankfurt am Main	0.107
Wien	0.079	Tampere	0.224	Århus	0.125
Eindhoven	0.087	Helsinki	0.231	Aalborg	0.130
Karlsruhe	0.128	London	0.233	München	0.134
Frankfurt am Main	0.128	Düsseldorf	0.241	Hamburg	0.137
München	0.135	Aalborg	0.242	London	0.139
Regensburg	0.138	Paris	0.243	Karlsruhe	0.146

Luxemburg		**Plovdiv**		**Madrid**	
Luxembourg	0.000	Plovdiv	0.000	Madrid	0.000
Paris	0.476	Ostrava	0.121	Montpellier	0.121
Linz	0.523	Sofia	0.123	Lille	0.160
Amsterdam	0.549	Cluj-Napoca	0.126	Marseille	0.210
München	0.554	Brno	0.130	Rouen	0.240
Wien	0.575	Bialystok	0.132	Toulon	0.257
Lyon	0.584	Lódz	0.134	Nancy	0.280
Graz	0.592	Szczecin	0.134	Saint-Etienne	0.281
Bologna	0.603	Bydgoszcz	0.137	Nantes	0.284
Rotterdam	0.609	Kraków	0.137	Bordeaux	0.291

Table 0-4 similarity results

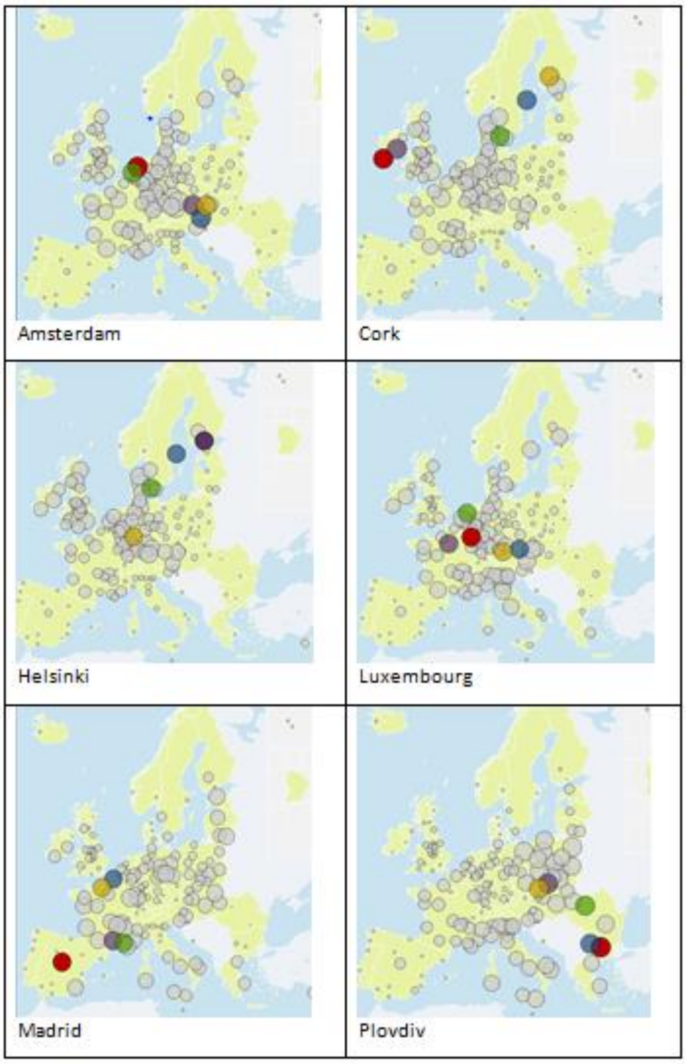


Fig 16 Plotting top 4 similar cities to each of the 6 cities

6. Limitations, future works and conclusion

6.1. Limitations

Some data from the vast collection of data from the main providers, ESPON and the EUROSTAT's Urban Audit (UA) collections is not yet available at LUZ geographical level revealing a shortage of directly comparable indicators across cities. Many indicators are collected at other regional scales other than the LUZ and conversion to LUZ level was not always possible.

In addition to the above, the latest edition of the UA uses different LUZ delineations resulting in the usage of different LUZ delineation version data when creating indicator time series from more than one UA edition. This may lead to errors in analysis due to the modifiable area unit problem (Openshaw 1984).

Even though we attempted to have the most complete indicator datasets, missing values for some cities in different indicators were unavoidable. Because of this, the presence of a city with no values in either one or more indicators means that this city is immediately dropped from the search or comparison operation resulting in an imperfect "image" of the current situation.

6.2. Future work

We have used a similarity measure to provide a quick scan of cities (European LUZ) revealing the most similar cities of a selected city in respect to selected indicators. A next step would be to test the robustness of this similarity measure by performing sensitivity analysis and testing different similarity methods. It would also be interesting to quantify the geographical trends of the distribution of similarity. Future research will attempt to use spatial auto-correlation and cluster analysis to quantify and further help identify spatial patterns.

6.3. Conclusions

Comparing and benchmarking cities is an important activity in determining profitable or investment ready cities, for the above reason, many tools exist that display data and rank cities based on pre-defined indicators such as the OECD metropolitan explorer.

CityBench offers to supplement such tools with a methodology that offers a multifaceted insight into similarity of cities for variable multi-dimensional indicators.

By highlighting the cities that are performing similar and different to a compared city, CityBench can be used to create opportunities for cooperation in cases where cities share similar challenges/opportunities or syndicates where low performing cities make alliances with high performers. Preliminary results show that CityBench can show regional patterns when meaningful indicators are used together.

Prototype feedback from stakeholders suggested that while the concept is worthwhile and provides a new insight into city comparisons, results should be carefully communicate in a clear and transparent manner.

Acknowledgments

We thankfully acknowledge the European Observatory for Territorial Development and Cohesion (ESPON) for funding this research and the team effort from both colleagues at Geodan Amsterdam, the Netherlands and Universitat Jaume I (UJI) Castellon, Spain.

Bibliography

Amelang J (2007) Comfigparing cities: a Barcelona Model? Urban History, 34, 173-189

Ammons DN (1999) A proper mentality for Benchmarking. Public Admistration Review Vol.59, No 2 , 105-109

Ashby F and Ennis D (2007) Similarity measures. Scholarpedia, volume 2, num 12, 4116

Batty M and Longley P (1994) Urban boundaries and Edges. In M. Batty and P. Longley, Fractal Cities: A Geometry of Form and Function (pp. 164-198). San Diego, CA and London: Academic Press.

Batty M and Longley PA (1994) Fractal Cities: A Geometry of Form and Function. San Diego, CA and London: Academic Press.

Baum S (1997) Sydney, Australia: A Global City? Testing the Social Polarisation Thesis. Urban Studies, 1881-1902.

Beaverstock JV, Hoyler M, Pain K et al (2001) Comparing London and Frankfurt as world cities: a relational study of contemporary urban change. Anglo-German Foundation for Study of Industrial Society, London.

Berson A (1996) Client/Server architecture. McGraw-Hill.

Burke B (2009) RESTful Java with JAX-RS. O'Reilly Media.Inc.

Deza MM and Deza E (2013) Distances and Similarities in Data Analysis. Encyclopedia of Distances, 291-305.

Dijkstra L and Poelman H (2012) Regional Focus (RF 01/2012). Retrieved from http://ec.europa.eu/regional_policy/sources/docgener/focus/2012_01_city.pdf.

EC (2004) Urban Audit: methodological handbook. Luxemburg: Office for Official Publications of the European Commission.

EC (2007) Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). Official Journal of the European Union, 50, 1-14.

Friedmann J (1986) The world citu hypothesis. Development and Change, 69-83.
Friedmann J (1995) Where we stand: a decade of world city research. World Cities in a World-System, 21-47.
General Books LLC (2012) ETL Tools:extract, transform, load, Informatica, Data Transformation services, Ibm infosphere Datastage, Ab Initio, Clover.Etl. General Books LLC.
Ghoda A (2012) HTML5, JavaScript, and Windows 8 Applications. Windows 8 MVVM Patterns Revealed, 99-107.
Groenendijk N S (2004) The use of benchmarking in EU economic and social policies. Conference on The Future of Europe, (pp. 24-25). Odense .
Hall P (2009) Looking backward, looking forward:the city region of the mid-21st century. Regional Studies, 43, 803-817.
Helgason S (1997, Feb) International benchmarking:experiences from OECD countries. Procedings of the Danish Ministry of Finance Conference on International Benchmarking, (pp. 20-21). Copenhagen: The Stationery Office.
Holloway A and Wajzer C (2008) Improving City Performance through Benchmarking. International Cities Town Centres & Communities Society Conference. Sydney.
Jain A, Prabhakar S, Hong L et al (2000) Filterbank-based fingerprint matching. Image Processing, IEEE Transactions vol.9, no.5, 846-859.
Kasanko M, Barredo JI, Lavalle C et al (2006) Are European cities becoming dispersed? A comparative analysis of 15 European urban areas. Landscape and Urban Planning, Volume 77, Issues 1-2, 111-130.
Lam EW, Chan AP and Chan DW (2010) Benchmarking success of building maintenance projects. Facilities 28, no. 5/6, 290-305.
Lanfranchi V and Ireson N (2009) User requirements for a collective intelligence emergency response system. In Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology (pp. 198-203). British Computer Society.
Lippman Abu-Lughod J (1995) Comparing Chicago, New York and Los Angeles: testing some world city hypotheses. Lippman Abu-Lughod, J. (1995) Comparing Chicago, New York and Los Angeles: testing some world city hypotheses, in: P. KNOx and World Cities in a World System, Lippman Abu-Lughod, J. (1995) Comparing Chicago, New York and Los Angeles: testing some world city hypotheses, in: P. KNOx and P. 171-191.
Luo J, Duggan DJ, Chen Yidong SJ et al (2001) Advances in Brief:Human Prostate Cancer and Benign Prostatic Hyperplasia: Molecular Dissection by Gene Expression Profiling. Cancer Research, 4683-4688.
Luque-Martínez T and Muñoz-Leiva F (2005) City benchmarking: A methodological proposal referring specifically to Granada. Teodoro Luque-Martínez, Francisco Muñoz-Leiva, City benchmarking: A methodologCities, Volume 22, Issue 6, 411-423.
Mane AV, Manza RR and Kale KV (2010) The Role of Similarity measures in face recognition. International Journal of Computer Science and Application .
Martin K (2009) Java Web Services:Up and Running. O'Reilly Media, Inc.
Murray S (2013) Interactivee Data Visualization for the Web. O'Reilly Media, Inc.
Núñez-Redó M, Díaz L, Gil J et al(2011) Discovery and integration of Web 2.0 content into geospatial information infrastructures: a use case in wild fire monitoring. Availability, Reliability and Security for Business, Enterprise and Health Information Systems, 50-68.
OECD (2012) Redefining "Urban": A New Way to Measure Metropolitan Areas. OECD Publishing.
Openshaw S (1984) The Modifiable Areal Unit Problem. Elsevier Science Geo Abstracts.
Parr J (2005) Perspectives on the city-region. Regional Studies, 39, 555-566.
Preoţiuc-Pietro D, Cranshaw J and Yano T (2013) Exploring venue-based city-to-city similarity measures. Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing (pp. 1-4). Chicago, Illinois: ACM .
Ramalho C E and Hobbs RJ (2012) Time for a change: dynamic urban ecology. Trends in ecology & evolution, 27(3), 179-188.
Sager C (2003) European City Benchmark Study. Copenhagen: Wonderful Copenhagen.
Sandoval J (2009) Master Core REST Concepts and Create RESTful Web Services in Java. Packt Publishing Ltd.
Seth R, Covell M, Ravichandran D et al (2011) A Tale of Two (Similar) Cities-Inferring City Similarity through Geo-spatial Query Log Analysis. KDIR, 179-189.
Sikos L (2011) Web standards: mastering HTML5, CSS3, and XML. Apress.
Tatnall A (2005) Web Portals: The New Gateways to Internet Information and Services. Idea Group Inc. (IGI).
Taylor PJ and Walker D (2001) World Cities: A first Multivariate Analysis of their Service Complexes. Urban Studies, 23-47.
Tuan Seik F (2000) Subjective assessment of urban quality of life in Singapore (1997–1998). Habitat international, Volume 24, Issue 1, 31-49.
Tulp JW (2012, November) close votes. Retrieved November 15, 2013, from Tulp interactive: http://tulpinteractive.com
Turok I and Mykhnenko V (2007) The trajectories of European cities, 1960–2005. Cities, Volume 24, Issue 3 , 165-182.

Vasanen A (2012) Functional Polycentricity: Examining Metropolitan Spatial Structure through the Connectivity of Urban Sub-centres. Urban Studies, 49, 3627-3644.
Woolley T (2013) an investigation of the Effect of the Swamping Phenomenon on Several block Procedures for Multiple Outliers in Univariate Samples. Open Journal of Statistics, vol 3, no.5, 229-304.

Annex I. Final list of included Larger Urban Zones

Final list of LUZ and corresponding MR as included in the Webtool.

	LUZ 2012 code	LUZ 2012 name	Corresponding MR 2012 code (based on NUTS 3 V2006)	MR 2012 name
1	AT001L2	Wien	AT001M	Wien
2	AT002L2	Graz	AT002M	Graz
3	AT003L2	Linz	AT003M	Linz
4	BE001L2	Bruxelles / Brussel	BE001M	Bruxelles / Brussel
5	BE002L2	Antwerpen	BE002M	Antwerpen
6	BE003L2	Gent	BE003M	Gent
7	BE005L2	Liège	BE005M	Liège
8	BG001L2	Sofia	BG001M	Sofia
9	BG002L2	Plovdiv	BG002M	Plovdiv
10	CH001L1	Zürich	CH001M	Zürich
11	CH002L1	Genève	CH002M	Genève
12	CH003L1	Basel	CH003M	Basel
13	CH004L1	Bern	CH004M	Bern
14	CY001L1	Lefkosia	CY001M	Lefkosia
15	CZ001L1	Praha	CZ001M	Praha
16	CZ002L1	Brno	CZ002M	Brno
17	CZ003L1	Ostrava	CZ003M	Ostrava
18	DE001L1	Berlin	DE001M	Berlin
19	DE002L1	Hamburg	DE002M	Hamburg
20	DE003L2	München	DE003M	München
21	DE004L1	Köln	DE004M	Köln
22	DE005L1	Frankfurt am Main	DE005M	Frankfurt am Main
23	DE007L1	Stuttgart	DE007M	Stuttgart
24	DE008L2	Leipzig	DE008M	Leipzig
25	DE009L2	Dresden	DE009M	Dresden
26	DE011L1	Düsseldorf	DE011M	Düsseldorf
27	DE012L1	Bremen	DE012M	Bremen
28	DE013L1	Hannover	DE013M	Hannover
29	DE014L1	Nürnberg	DE014M	Nürnberg
30	DE017L0	Bielefeld	DE017M	Bielefeld
31	DE018L1	Halle an der Saale	DE018M	Halle an der Saale
32	DE019L2	Magdeburg	DE019M	Magdeburg
33	DE020L1	Wiesbaden	DE020M	Wiesbaden
34	DE025L1	Darmstadt	DE025M	Darmstadt
35	DE027L1	Freiburg im Breisgau	DE027M	Freiburg im Breisgau
36	DE028L1	Regensburg	DE028M	Regensburg
37	DE032L1	Erfurt	DE032M	Erfurt
38	DE033L2	Augsburg	DE033M	Augsburg
39	DE034L1	Bonn	DE034M	Bonn
40	DE035L1	Karlsruhe	DE035M	Karlsruhe

41	DE037L1	Mainz	DE037M	Mainz
42	DE038L1	Ruhrgebiet	DE038M	Ruhrgebiet
43	DE039L1	Kiel	DE039M	Kiel
44	DE040L1	Saarbrücken	DE040M	Saarbrücken
45	DE084L1	Mannheim-Ludwigshafen	DE084M	Mannheim-Ludwigshafen
46	DE504L1	Münster	DE504M	Münster
47	DE507L1	Aachen	DE507M	Aachen
48	DK001L2	Köbenhavn	DK001M	Köbenhavn
49	DK002L1	Århus	DK002M	Århus
50	DK003L1	Odense	DK003M	Odense
51	DK004L1	Aalborg	DK004M	Aalborg
52	EE001L1	Tallinn	EE001M	Tallinn
53	EL001L1	Athina	EL001M	Athina
54	EL002L1	Thessaloniki	EL002M	Thessaloniki
55	ES001L1	Madrid	ES001M	Madrid
56	ES002L1	Barcelona	ES002M	Barcelona
57	ES003L1	Valencia	ES003M	Valencia
58	ES004L1	Sevilla	ES004M	Sevilla
59	ES005L1	Zaragoza	ES005M	Zaragoza
60	ES006L1	Malaga	ES006M	Malaga
61	ES007L1	Murcia	ES007M	Murcia
62	ES008L1	Las Palmas	ES008M	Las Palmas
63	ES009L1	Valladolid	ES009M	Valladolid
64	ES010L1	Palma de Mallorca	ES010M	Palma de Mallorca
65	ES019L1	Bilbao	ES019M	Bilbao
66	ES020L1	Córdoba	ES020M	Córdoba
67	ES021L1	Alicante/Alacant	ES021M	Alicante/Alacant
68	ES022L1	Vigo	ES022M	Vigo
69	ES025L1	Santa Cruz de Tenerife	ES025M	Santa Cruz de Tenerife
70	ES026L1	Coruña (A)	ES026M	Coruña (A)
71	FI001L2	Helsinki	FI001M	Helsinki
72	FI002L2	Tampere	FI002M	Tampere
73	FR001L1	Paris	FR001M	Paris
74	FR003L2	Lyon	FR003M	Lyon
75	FR004L2	Toulouse	FR004M	Toulouse
76	FR006L2	Strasbourg	FR006M	Strasbourg
77	FR007L2	Bordeaux	FR007M	Bordeaux
78	FR008L2	Nantes	FR008M	Nantes
79	FR009L2	Lille	FR009M	Lille
80	FR010L2	Montpellier	FR010M	Montpellier
81	FR011L2	Saint-Etienne	FR011M	Saint-Etienne
82	FR013L2	Rennes	FR013M	Rennes
83	FR215L2	Rouen	FR015M	Rouen
84	FR016L2	Nancy	FR016M	Nancy

85	FR017L2	Metz	N/A (no corresponding MR)	
86	FR022L2	Clermont-Ferrand	FR022M	Clermont-Ferrand
87	FR026L2	Grenoble	FR026M	Grenoble
88	FR032L2	Toulon	FR032M	Toulon
89	FR035L2	Tours	FR035M	Tours
90	FR203L2	Marseille	FR203M	Marseille
91	FR205L2	Nice	FR205M	Nice
92	HR001L2	Grad Zagreb	HR001M	Grad Zagreb
93	HR005L2	Split	HR005M	Split
94	HU001L2	Budapest	HU001M	Budapest
95	HU005L2	Debrecen	HU005M	Debrecen
96	IE001L1	Dublin	IE001M	Dublin
97	IE002L1	Cork	IE002M	Cork
98	IS001L1	Reykjavík	IS001 **(NUTS 3 version 2006 code)**	Höfuðborgarsvæði **(NUTS 3 version 2006 name)**
99	IT001L2	Roma	IT001M	Roma
100	IT002L2	Milano	IT002M	Milano
101	IT003L2	Napoli	IT003M	Napoli
102	IT004L2	Torino	IT004M	Torino
103	IT005L2	Palermo	IT005M	Palermo
104	IT006L2	Genova	IT006M	Genova
105	IT007L2	Firenze	IT007M	Firenze
106	IT008L2	Bari	IT008M	Bari
107	IT009L1	Bologna	IT009M	Bologna
108	IT010L2	Catania	IT010M	Catania
109	IT011L2	Venezia	IT011M	Venezia
110	IT012L2	Verona	IT012M	Verona
111	IT022L2	Taranto	IT022M	Taranto
112	IT027L1	Cagliari	IT027M	Cagliari
113	IT028L2	Padova	IT028M	Padova
114	IT029L2	Brescia	IT029M	Brescia
115	LT001L1	Vilnius	LT001M	Vilnius
116	LT002L1	Kaunas	LT002M	Kaunas
117	LU001L1	Luxembourg	LU001M	Luxembourg
118	LV001L0	Riga	LV001M	Riga
119	MT001L1	Valletta	MT001M	Valletta
120	NL001L2	Den Haag	NL001M	Den Haag
121	NL002L2	Amsterdam	NL002M	Amsterdam
122	NL003L2	Rotterdam	NL003M	Rotterdam
123	NL004L2	Utrecht	NL004M	Utrecht
124	NL005L2	Eindhoven	NL005M	Eindhoven
125	NO001L2	Oslo	NO001M	Oslo
126	NO002L2	Bergen	NO002M	Bergen
127	PL001L2	Warszawa	PL001M	Warszawa
128	PL002L2	Łódź	PL002M	Łódź

129	PL003L2	Kraków	PL003M	Kraków
130	PL004L2	Wroclaw	PL004M	Wroclaw
131	PL005L2	Poznan	PL005M	Poznan
132	PL006L2	Gdansk	PL006M	Gdansk
133	PL007L2	Szczecin	PL007M	Szczecin
134	PL008L2	Bydgoszcz	PL008M	Bydgoszcz
135	PL009L2	Lublin	PL009M	Lublin
136	PL010L2	Katowice	PL010M	Katowice
137	PL011L2	Bialystok	PL011M	Bialystok
138	PL012L2	Kielce	PL012M	Kielce
139	PL024L2	Czestochowa	PL024M	Czestochowa
140	PT001L2	Lisboa	PT001M	Lisboa
141	PT002L2	Porto	PT002M	Porto
142	RO001L1	Bucuresti	RO001M	Bucuresti
143	RO002L1	Cluj-Napoca	RO002M	Cluj-Napoca
144	SE001L1	Stockholm	SE001M	Stockholm
145	SE002L1	Göteborg	SE002M	Göteborg
146	SE003L1	Malmö	SE003M	Malmö
147	SI001L1	Ljubljana	SI001M	Ljubljana
148	SI002L1	Maribor	SI002M	Maribor
149	SK001L1	Bratislava	SK001M	Bratislava
150	SK002L1	Košice	SK002M	Košice
151	UK001L2	London	UK001M	London
152	UK002L2	West Midlands urban area	UK002M	Birmingham
153	UK003L1	Leeds	UK003M	Leeds
154	UK004L1	Glasgow	UK004M	Glasgow
155	UK005L0	Bradford	UK005M	Bradford
156	UK006L2	Liverpool	UK006M	Liverpool
157	UK007L1	Edinburgh	UK007M	Edinburgh
158	UK008L2	Manchester	UK008M	Manchester
159	UK009L1	Cardiff	UK009M	Cardiff
160	UK010L2	Sheffield	UK010M	Sheffield
161	UK011L2	Bristol	UK011M	Bristol
162	UK012L1	Belfast	UK012M	Belfast
163	UK013L2	Newcastle upon Tyne	UK013M	Newcastle upon Tyne
164	UK014L1	Leicester	UK014M	Leicester
165	UK016L1	Aberdeen	UK016M	Aberdeen
166	UK018L2	Exeter	UK018M	Exeter
167	UK023L1	Portsmouth	UK023M	Portsmouth
168	UK025L2	Coventry	UK025M	Coventry
169	UK026L1	Kingston upon Hull	UK026M	Kingston upon Hull
170	UK027L1	Stoke-on-Trent	UK027M	Stoke-on-Trent
171	UK029L1	Nottingham	UK029M	Nottingham
172	LI000 **(NUTS 3 version 2006 code)**	Liechtenstein **(NUTS 3 version 2006 name)**	LI000 **(NUTS 3 version 2006 code)**	Liechtenstein **(NUTS 3 version 2006 name)**

Annex II. Themes and indicators

A. Final list of themes and indicators

	Indicator	Source (code)	Source (name)	Years	Geographic level (version)
Theme: Economy and Population					
	Population, total / Population density	3	ESPON M4D	2002-2011	NUTS 3 (2006)
	GDP per head in PPS	2	Eurostat met	2000-2010	MR (NUTS 3 2006?)
	% of persons unemployed	2	Eurostat met	2003-2012	MR (NUTS 3 2006?)
	"Agree" with It is easy to find a job	6	DG Regio	2006, 2009, 2012	Core city
Theme: Connectivity					
	Potential accessibility by rail, standardized ESPON	3	ESPON TRACC	2001, 2006	NUTS 3 (2006)
	Potential accessibility by road, standardized ESPON	3	ESPON TRACC	2001, 2006	NUTS 3 (2006)
	Potential accessibility by air, standardized ESPON	3	ESPON TRACC	2001, 2006	NUTS 3 (2006)
	"Satisfied" with public transport	6	DG Regio	2006, 2009, 2012	Core city
Theme: Demography					
	Ageing index	3	ESPON INTERCO	2002-2009	NUTS 3 (2006)
	Old age dependency	3	ESPON INTERCO	2002-2009	NUTS 3 (2006)
Theme: Social media					
	# of items being posted about 'Crisis'/·# items posted in the city	5	Twitter	Daily from November 2013	LUZ (2012)
	# of items being posted about 'Unemployment'/·# items posted in the city	5	Twitter	Daily from April 2014	LUZ (2012)
Theme: Investment climate					
	Ease of doing business	5	IFC / World Bank	2005-2013	Country
	Gas prices for industrial consumers	2	Eurostat	2008-2013	Country

			nrg		
	Electricity prices for industrial consumers	2	Eurostat nrg	2008-2013	Country
Theme: Environment					
	% of LUZ consisting of green urban areas	4	EEA	2006	LUZ (2012)
	Maximum PM10 concentration	4	EEA	2008	LUZ (2012)
	Maximum CO2 concentration	4	EEA	2008	LUZ (2012)
	Combined adaptive capacity to climate change	3	ESPON Climate	2011	NUTS 3 (2006)
	"Satisfied" with Green spaces such as parks and gardens	6	DG Regio	2006, 2009, 2012	Core city
Theme: Smartness					
	High-Tech (total) patent applications to the EPO per million of inhabitants	2	Eurostat	2001-2010	MR (NUTS 3 2006?)
	IP Addresses	3	ESPON TEL Update	2009	NUTS 3 (2006)
	Share of renewable energy in gross final energy consumption	2	Eurostat	2011	Country
Theme: Quality of life					
	"Agree" with Easy to find a house	6	DG Regio	2006, 2009, 2012	Core city
	"Agree" with I'm satisfied to live in	6	DG Regio	2006, 2009, 2012	Core city
	"Satisfied" with Cleanliness	6	DG Regio	2006, 2009, 2012	Core city
	"Satisfied" with Cultural facilities such as concert halls, theatres, museums and libraries	6	DG Regio	2006, 2009, 2012	Core city

B. Operationalization of indicators

<table>
<tr><th>Indicator</th><th>Source (name), Geographic level (version)</th><th>Calculation of indicator</th></tr>
<tr><td>Population, total / Population density</td><td>ESPON M4D, NUTS 3 (2006)</td><td>Resident population:
1. Download tabular data
2. Aggregate NUTS 3 POP_T values to MR level by using an MR – NUTS 3 (2006) correspondence table

Population density:
1. Calculate MR areas from MR outlines vector file (using Lambert azimuthal equal-area projection)
2. Divide aggregated POP_T values by MR areas</td></tr>
<tr><td>GDP per head in PPS</td><td>Eurostat met , MR (NUTS 3 2006?)</td><td>1. Download tabular data
2. Extract values for Gross domestic product (GDP) at current market prices - Purchasing Power Standard per inhabitant
3. Correct some MR codes (remove ‘C’ at end of code) to prevent inconsistencies with MR – LUZ correspondence table in CityBench Webtool</td></tr>
<tr><td>% of persons unemployed</td><td>Eurostat met , MR (NUTS 3 2006?)</td><td>1. Download tabular data
2. Extract values for Unemployment rates, Total, 15 years or over
3. Correct some MR codes (remove ‘C’ at end of code) to prevent inconsistencies with MR – LUZ correspondence table in CityBench Webtool</td></tr>
<tr><td>Potential accessibility by rail / road / air, standardized ESPON</td><td>ESPON TRACC , NUTS 3 (2006)</td><td>1. Download tabular data from project: ACC Update
2. Extract values for Potential accessibility, rail / road, ESPON average / Potential accessibility by air, standardized
3. For a particular year, determine value for each MR by assigning highest value from its included NUTS 3 regions to it, using an MR – NUTS 3 (2006) correspondence table</td></tr>
<tr><td>Ageing index</td><td>ESPON INTERCO , NUTS 3 (2006)</td><td rowspan="2">1. Download tabular data from project: INTERCO
2. Extract values for Ageing index / Old age dependency ratio
3. For a particular year, calculate average value for each MR from values for its included NUTS 3 regions, using an MR – NUTS 3 (2006) correspondence table (not available online) and assigning a weight to each NUTS 3 (2006) region proportional to its population number in that year (derived by extracting values for Population, total from ESPON db => Project: ESPON M4D)</td></tr>
<tr><td>Old age dependency</td><td>ESPON INTERCO , NUTS 3 (2006)</td></tr>
<tr><td># of items being posted about ‘Crisis’</td><td>Twitter, LUZ (2012)</td><td rowspan="2">4. Generate 2 lists with ‘Crisis’ resp. ‘Unemployment’ keywords, for each official language within selected LUZ
5. Create query and batch process to send this query to Twitter/Flickr/YouTube
6. Store number of items posted daily/weekly/monthly/yearly within each LUZ
7. For values per inhabitant: divide stored numbers by Resident population values</td></tr>
<tr><td># of items being posted about ‘Unemployment’</td><td>Twitter, LUZ (2012)</td></tr>
<tr><td># of items being posted by tourists</td><td>Twitter, LUZ (2012)</td><td>1. Create query and batch process to send this query to Twitter/Flickr/YouTube
2. Store number of items posted daily/weekly/monthly/yearly by foreigners within each LUZ</td></tr>
</table>

		3. For values per inhabitant: divide stored numbers by Resident population values
Ease of doing business	IFC / World Bank, Country	1. Download tabular data 2. Extract EU27+4 countries 3. Assign country rankings for Ease of Doing Business to each LUZ included in it
Gas / Electricity prices for industrial consumers	Eurostat nrg, Country	1. Download tabular data 2. Extract Gas / Electricity prices (Euro, Kilowatt/hour, excluding taxes and levies) for industrial consumers (Gas: Band I3; Electricity: Band IC) 3. Assign values for Gas / Electricity prices for industrial consumers per country to each LUZ included in it
% of LUZ consisting of green urban areas	EEA, LUZ (2012)	1. Download raster data 2. Perform Tabulate Area calculations on CLC 2006 classes for LUZ 2012 areas 3. Convert absolute values to % 4. Combine CLC classes 141 (Green urban areas) and 142 (Sport and leisure facilities) to obtain total % of green area inside LUZ
Maximum PM10 / CO2 concentration	EEA, LUZ (2012)	1. Download European Pollutant Release and Transfer Register (E-PRTR) diffuse air emission datasets 2. Calculate raster statistics 3. Reclassify raster into 7 classes using Jenks Natural Breaks 4. Perform Tabulate Area calculations on reclassified raster for LUZ 2012 areas 5. Manually derive maximum pollution class (1 to 7) for 5 x 5 km area within each LUZ
Combined adaptive capacity to climate change	ESPON Climate, NUTS 3 (2006)	1. Download tabular data from project: ESPON Climate (*no public access*) 2. Calculate average value for each MR from values for its included NUTS 3 regions, using MR – NUTS 3 (2006) correspondence table and assigning a weight to each NUTS 3 (2006) region proportional to its population number on 2009-01-01 (derived by extracting values for Population, total from ESPON db => Project: ESPON M4D)
High-Tech patent applications to the EPO	Eurostat, MR (NUTS 3 2006?)	1. Download tabular data 2. Extract values for High tech - total patent applications to the EPO per million of inhabitants 3. Correct some MR codes (remove 'C' at end of code) to prevent inconsistencies with MR – LUZ correspondence table in CityBench Webtool
Number of IP Addresses	ESPON TEL Update, NUTS 3 (2006)	1. Download tabular data from Project: TEL Update 2. Calculate value for each MR by aggregating values from its included NUTS 3 regions, using MR – NUTS 3 (2006) correspondence table
Share of renewable energy in gross final energy consumption	Eurostat, Country	1. Download tabular data 2. Assign values for Share of renewable energy in gross final energy consumption per country to each LUZ included in it
Satisfaction with: Public transport / Cultural facilities / Cleanliness / Green spaces; Satisfied to live in city; Ease of finding: good housing at reasonable price / job (Perception Survey results)	DG Regio, Core city	1. Download tabular data 2. Extract data for specific questions from different years (2012, 2009 and 2006) 3. Link city names in Perception survey tables to LUZ 2012 codes 4. Create table with LUZ 2012 codes and Perception survey results

Annex III. Social Media Indicators

When studying the existing information to analyze and compare cities we cannot forget new trends in digital information and social media. Nowadays bottom-up initiatives complement the ecosystem of information available online. Regular citizens are sharing information about their surrounding and their cities through a large number of social networks. This increasing amount of information cannot be ignored, since it provides in most cases, real time information about places and events, which have impact on society. In this context we witness how location analytics industry is moving from a paradigm of lower volume, higher accuracy data to one of higher volume and lower accuracy. CityBench project aims to generate indicators not only using official data but also indicators which offer a social view analyzing data shared by citizen through social networks.

Recent trends in information technology show that citizens are increasingly willing to share information using tools provided by crowdsourcing platforms to describe events with social impact. This is fuelled by the proliferation of location-aware devices such as smartphones and tablets, users are able to share information in these crowdsourcing platforms directly from the field at real time, augmenting this information with its location. However, there is still difficult to extract useful information from this big volume of raw data. It is necessary to generate indicators that resume in useful information the analysis of these data. As such content refers to phenomena that are bound to a location, georeferenced user-generated content is acquiring a fundamental role in a wide range of applications.

Simple georeferenced messages from social networks such as Twitter[1] may play a major role in response actions to emergencies (Schade et al., 2012)(Roche et al., 2012). Not only tweets but other types of data such as videos, audio files and pictures may also be related to a location, and being used in diverse situations such as volunteered–based map creation (Neis et al., 2012), collect in situ biodiversity data (Newell et al., 2012) and forestry data (Aragó et al., 2011). Although georeferenced user-generated data still represents a small percentage, its growth is being greatly accelerated largely by the use of sensor-enabled devices. It is thus reasonable to foresee that huge amounts of georeferenced data will be available in an immediate future.

It is of interest of the project to provide social media information about current circumstances regarding Crisis, Politics, Economy and unemployment. The interest of citizens of talking, asking or informing about these issues provides an overview about social concern of citizens and therefore it could influence in vulnerable fields such as financial markets. This also could provide useful information to support decision making to local government or a potential investor.

After studying first results obtained from social media sources, we have defined a starting set of indicators related to a topic inside a LUZ delimitation, defined as “Topic items in a city per total items in that city”, intended to describe the social media interest of that city in the cited topic.

Social Media data source used is Twitter. Topics chosen are “crisis” and “unemployment”.

Table AIII-1. Attitude/information about current circumstances indicators.

Source	Indicator	Spatial level (+ version)	Year	Unit / format	Remarks
Twitter	# items posted in a city containing “crisis” keywords per total items posted in that city	LUZ (2012)	2013-2014	%	Data is stored daily. Accumulated values are also stored to reach an annual indicator value

Twitter	# items posted in a city containing "unemployment" keywords per total items posted in that city	LUZ (2012)	2013-2014	%	Data is stored daily. Accumulated values are also stored to reach an annual indicator value.

Calculation procedure

- Generate lists of keywords (in every language) that illustrate the topic field to create a query to send to the Social Media Networks.
- Create a batch process to send this query to Twitter/Flickr/YouTube and store the number of items posted in a city. This is done for every city, on a daily basis.
- Indicator can be calculated in different temporal resolution: daily/ weekly/ monthly/ yearly

Crisis

- Query (UK001L2, en): service=Twitter&q= bailout, bank, business, bubble, capital, crash, corruption, credit, crisis, debt, deflation, bankruptcy, economy, euro, financial, investment, market, money, price, rate, recession, sme, stock market &bbox=&lat=3.212.622,546lon=3.635.630,281 &radius=70000&format=atom&end=26062013&start=26062013
- Projection: ETRS89_ETRS_LAEA
- Keywords (en,English): bailout, bank, business, bubble, capital, crash, corruption, credit, crisis, debt, deflation, bankruptcy, economy, euro, financial, investment, market, money, price, rate, recession, sme, stock market

Table AIII-2. Crisis keywords for the EU27+4 official languages.

language, LUZ	Keywords
bg,Bulgarian	bailout, банка, business, bubble, capital, crash, корупцията, кредит, криза, дълг, дефлация, несъстоятелност, икономиката, euro, финансови, инвестиционни, пазар, пари, цена, rate, рецесия, мсп, stock market
hr,Croatian	bailout, banka, business, bubble, kapital, crash, korupcija, kreditne, kriza, dug, ispuhavanja, stečaj, ekonomija, euro, financijska, investicijska, tržište, novac, cijena, stopa, recesija, malog i srednjeg poduzetništva, burza
cs,Czech	výpomoci, banka, business, bubble, kapitál, crash, korupce, úvěr, krize, zadlužení, deflace, konkurs, ekonomika, euro, finanční, investiční, trh, peníze, cena, rychlost, recese, malých a středních podniků, stock market
da,Danish	redningen, bank, business, bubble, capital, crash, korruption, credit, krise, gæld, deflation, konkurs, economy, euro, finansielle, investment, markedsaktører, money, pris, rate, recession, sme, stock market
nl,Dutch	bailout, bank, business, bubble, hoofdstad, crash, corruptie, krediet, crisis, schuld, deflatie, faillissement, economie, euro, financieel, investering, markt, geld, prijs, tarief, recessie, kmo, stock market
et,Estonian	kautsjoni, bank, business, bubble, capital, crash, korruptsiooni-, krediidi, kriisi, võlakirjad, deflatsioon, pankrot, majandus, euro, finants-, investeeringute, turg, raha, hind, rate, allakäik, vkede, stock market
fi,Finnish	bailout, bank, business, bubble, capital, crash, korruptio, luotto, kriisi, velka, deflaatio, konkurssi, talous, euro, rahoitus-, sijoitus, market, money, hinta, rate, lama, pk-yritysten, stock market
fr,French	renflouement, banques, business, bulle, capital, crash, la corruption, crédit, crise, dette, la déflation, la faillite, économie, euro, financiers, d'investissement, le marché, l'argent, le prix, taux, récession, pme, stock market
de,German	bailout, bank, business, blase, capital, unfall, korruption, kredit, krise, schulden, deflation,

	bankrott, wirtschaft, euro, finanzen, investitionen, markt, geld, preis, preise, rezession, kmu, stock market
el,Greek	διάσωσης, τράπεζα, επαγγελματικός, φυσαλίδας, capital, crash, διαφθορά, credit, κρίση, χρέος, αποπληθωρισμός, την πτώχευση, οικονομία, ευρώ, χρηματιστηριακές, επενδυτικές, αγορά, χρήματα, τιμή, τιμή, ύφεση, μμε, χρηματιστήριο
en,English	bailout, bank, business, bubble, capital, crash, corruption, credit, crisis, debt, deflation, bankruptcy, economy, euro, financial, investment, market, money, price, rate, recession, sme, stock market
hu,Hungarian	szanálását, bank, business, bubble, capital, crash, korrupció, hitel, válság, adósság, defláció, csőd, gazdaság, euro, pénzügyi, befektetési, piac, pénz, ár, árfolyam, recesszió, kkv-k, stock market
ga,Irish	fhóirithint, banc, business, bubble, caipitil, crash, éilliú a, creidmheasa, ghéarchéime, fiach, díbhoilsciú, féimheachta, geilleagar, an euro, airgeadais, infheistíochta, mhargaidh, airgead, praghas, ráta, chúlú eacnamaíochta, sme, stoc mhargaidh
it,Italian	bailout, banca, bussines, bolla, capitale, crash, la corruzione, credito, crisi, debito, deflazione, fallimento, economia, euro, finanziario, investimento, mercato, denaro, prezzo, tasso, recessione, pmi, azioni
lv,Latvian	bailout, banka, business, bubble, capital, crash, korupciju, kredīts, krīze, parādu, deflācija, bankrotu, ekonomika, eiro, finanšu, investīciju, tirgus, nauda, cena, rate, recesijas, mvu, akciju tirgus
lt,Lithuanian	finansinės pagalbos, bankas, bussines, burbulas, kapitalas, avarijos, korupcija, kredito, krizė, skolos, defliacija, bankrotas, ekonomika, euras, finansai, investicijos, rinkos, pinigai, kaina, reitingas, recesija, mvį, vertybinių popierių rinka
mt,Maltese	kawzjoni, bank, business, bubble, capital, crash, korruzzjoni, kreditu, kriżi, dejn, deflazzjoni, falliment, ekonomija, euro, finanzjarja, investiment, market, money, price, rata, reċessjoni, sme, stock market
pl,Polish	bailout, bank, business, bubble, capital, crash, korupcja, kredyt, kryzys, dług, deflacja, upadłość, gospodarka, euro, finanse, inwestycje, market, money, cena, cena, recesja, sme, stock market
pt,Portuguese	ajuda, banco, bussines, bolha, capital, crash, corrupção, crédito, crise, dívida, deflação, falência, economia, euro, financeira, investimento, mercado, dinheiro, preço, rate, recessão, sme, stock market
ro,Romanian	bailout, bank, business, bubble, capital, crash, corupția, de credit, criza, datoriilor, deflația, falimentul, economy, euro, financiar, de investiții, market, money, pret, rate, recesiunea, pentru imm-uri, market stock
sk,Slovak	výpomoci, banka, business, bubble, kapitál, crash, korupcia, úver, kríza, zadlženie, deflácia, konkurz, ekonomika, euro, finančné, investičné, trh, peniaze, cena, rýchlosť, recesia, malých a stredných podnikov, stock market
sl,Slovene	bailout, banka, poslovni, bubble, capital, crash, korupcije, credit, kriza, dolg, deflacija, ste?aju, gospodarstvo, evro, finan?ne, naložbe, trg, denar, cena, cena, recesija, msp, stock market
es,Spanish	rescate financiero, banco, negocios, burbuja, capital, crash, corrupción, crédito, crisis, deuda, deflación, bancarrota, economía, euro, financiera, inversiones, mercado, dinero, precio, velocidad, recesión, pyme, mercado de valores
sv,Swedish	bailout, bank, business, bubble, huvudstad, krasch, korruption, kredit, kris, skuld, deflation, konkurs, ekonomi, euro, finansiell, investering, marknadsaktörer, pengar, pris, rate, recession, smf, stock market
is,Icelandic	bailout, bank, business, bubble, capital, crash, spillingu, credit, crisis, skuld, verðhjöðnun, gjaldþrot, economy, euro, financial, investment, market, money, price, rate, samdráttur, sme, stock market
tu,Turkish	kurtarma, banka, business, kabarcık, sermaye, crash, yolsuzluk, kredi, kriz, borç, deflasyon, iflas, ekonomi, euro, finans, yatırım, market, para, fiyat, fiyat, durgunluk, kobi, borsa
lu,Luxembourgish	
no,Norwegian	bailout, bank, business, bubble, capital, crash, korrupsjon, credit, krise, gjeld, deflasjon, konkurs, economy, euro, financial, investment, market, money, prisen, rate, tilbakeslag, sme, stock market

Unemployment

- Query (UK001L2, en): service=Twitter&q= Demand, Employment, Increase, Job, Labour, Market, Productivity, Rate, Salary, Selfemployment, Underemployed, Unemployment, Wages, Work, Worker, Young &bbox=&lat=3.212.622,546lon=3.635.630,281 &radius=70000&format=atom&end=26062013&start=26062013
- Projection: ETRS89_ETRS_LAEA
- Keywords (en): Demand, Employment, Increase, Job, Labour, Market, Productivity, Rate, Salary, Selfemployment, Underemployed, Unemployment, Wages, Work, Worker, Young

Table AIII-3. Unemployment keywords for the EU27+4 official languages.

language, LUZ	Keywords
bg,Bulgarian	Търсенето, трудовата заетост, увеличаване, Job, труда, пазар, производителността, равнището на заплатата, Selfemployment, непълна заетост, безработица, заплати, работа, Работник, младежи
hr,Croatian	Potražnja, zapošljavanje, Povećanje, posla, rada, tržišta, produktivnost, stopa, plaća, Selfemployment, nedovoljno, Nezaposlenost, plaće, rad, radnik, mladih
cs,Czech	Poptávka, zaměstnání, zvýšení, práce, práce, trh, produktivita, rychlost, Plat, samostatná výdělečná činnost, Podzaměstnaní, nezaměstnanost, Mzdy, práce, dělník, mládeže
da,Danish	Demand, beskæftigelse, øge, Job, Labour, Marked, produktivitet, Rate, Løn, selvansættelse, underbeskæftigede, arbejdsløshed, Wages, Arbejde, Worker, Ungdom
nl,Dutch	Vraag, Werkgelegenheid, Verhoog, Job, Arbeid, Markt, Productiviteit, Rate, Salaris, Selfemployment, Underemployed, werkloosheid, lonen, Work, Worker, Jeugd
et,Estonian	Nõudlus, tööhõive suurendamine, töö, töö, tootlusele, Rate, Palk, Selfemployment, Vaeghõivatuid, töötus töötasu, töö-, töötaja-, noorsoo-
fi,Finnish	Kysyntä, Työllisyys, kasvu, Job, Labour, Market, tuottavuus, Rate, Palkka, Selfemployment, alityöllistettyjä, työttömyys, palkat, työ, työntekijä-, nuoriso-
fr,French	Demande, emploi, augmentation, travail, marché, productivité, taux, salaire, travail autonome, sous-emploi, chômage salaires, travail, travailleur, jeunesse
de,German	Die Nachfrage, Beschäftigung, Erhöhung, Job, Arbeit, Markt, Produktivität, Rate, Gehalt, Selbstständigkeit, unterbeschäftigt, Arbeitslosigkeit, Lohn, Arbeit, Arbeiter, Jugend
el,Greek	Ζήτηση, την απασχόληση, την αύξηση, Εργασίας, Εργασίας, Αγορά, παραγωγικότητα, το ποσοστό, Μισθός, αυταπασχόληση, υποαπασχολούμενοι, ανεργία, μισθοί, εργασία, εργάτης, Νεολαία
en,English	Demand, Employment, Increase, Job, Labour, Market, Productivity, Rate, Salary, Self-employment, Underemployed, Unemployment, Vacancy, Wages, Work, Worker, Young
hu,Hungarian	Kereslet, a foglalkoztatás, növekedés, munka, munka-, piac, termelékenység, Rate, Fizetés, Selfemployment, alulfoglalkoztatott, munkanélküliség, bérek, munka, munkás, ifjúsági
ga,Irish	Éileamh, Fostaíocht, Méadú, Jabanna, an Lucht Oibre, an Mhargaidh, Táirgiúlacht, Ráta, Tuarastal, Selfemployment, underemployed, Dífhostaíocht, Pá, Obair, Oibrí, Óige
it,Italian	La domanda, lavoro, aumentare, lavoro, lavoro, mercato, della produttività, di cambio, stipendio, lavoro autonomo, sottoccupati, Disoccupazione, salari, lavoro, lavoratore, la gioventù
lv,Latvian	Pieprasījums, darba tirgū, palielināt, Darba, Darba, tirgus, ražīgums, Rate, alga, Selfemployment, nepietiekami, bezdarbs, algas, darba, darbinieks, jaunatnes
lt,Lithuanian	Paklausa, Užimtumo, padidinimas, Darbas, Darbo, rinka, našumas, Reitingas, Atlyginimas Selfemployment, ne visu pajėgumu, Nedarbas, Darbo užmokestis, darbas, darbuotojas, jaunimo
mt,Maltese	Demand Impjiegi, Żieda, Job, Labour, Market, Produttività, Rata, Salarju, Selfemployment, sottoimpjegati, qgħad, Pagi, ix-Xogħol, Worker, Żgħażagħ
pl,Polish	Popyt, zatrudnienie, wzrost, praca, pracy, rynku, wydajność, szybkość, zalogi, samozatrudnienia, niepełne zatrudnienie, bezrobocie, Płace, Praca, Pracownik, młodzież
pt,Portuguese	Demand, Emprego, Aumento, Job, Trabalho, mercado, produtividade, Rate, Salário, Selfemployment, subempregados, Desemprego, Salário, Trabalho, Trabalhador, Juventude
ro,Romanian	Cerere, de muncă și creșterea, locuri de muncă, muncă, piață, a productivității, Rate, salariu, Selfemployment, sub-angajați, șomaj, salarii, munca, munca, tineret

sk,Slovak	Dopyt, zamestnanie, zvýšenie, práca, práca, trh, produktivita, rýchlosť, Plat, samostatná zárobková činnosť, Podzaměstnaní, nezamestnanosť, Mzdy, práca, robotník, mládeže
sl,Slovene	Povpraševanje, zaposlovanje, povečanje, Job, dela, trg, Produktivnost, Rate, plače, samozaposlitve, podzaposlenim, Brezposelnost, plače, delo, delavec, mladina
es,Spanish	Demanda, empleo, Aumento, trabajo, trabajo, mercado, productividad, velocidad, Salario, autoempleo, subempleados, Desempleo, salario, trabajo, trabajador, Juventud
sv,Swedish	Efterfrågan, arbetsmarknaden, öka, jobb, arbetsmarknad, marknad, produktivitet, Rate, Lön, egenföretagande, undersysselsatta Arbetslöshet, löner, arbete, arbetare, ungdom
is	Eftirspurnar, atvinnu, Aukning Job, Vinnumálastofnun, Market, framleiðni, Rate, Laun, Selfemployment, Vinnulítill, Atvinnuleysi, Laun, Vinna, Worker, Young
tu	Talep, İstihdam, Artış, İş, Çalışma, Market, Verimlilik, Hızı, Maaş, Selfemployment, Eksik İstihdam, İşsizlik, Ücret, İş, İşçi, Genç
lu	
no	Etterspørsel, sysselsetting, øke, Job, arbeids-marked, produktivitet, Rate, Lønn, Selfemployment, undersysselsatte, Arbeidsledighet, Lønn, Arbeid, Worker, Young

Other possible topics:

- Economy: Crisis, Bank, Money, Credit, SME, Financial, Euro, Debt, Investment, Unemployment, Employment, Job, Market, ECB, OECD, Labor Union, Trade Union, Stock Market
- Politics: Parliament, Europe, European Commission, Politics, Corruption, Right Wing, Left Wing, Union, Labor Union, Trade Union, Legislation, Law

Annex IV. Uploading new indicator data using Admin Tool

Below are three likely scenarios for populating CityBench with new indicator data. In a few steps, the procedure for uploading data using the custom CityBench upload sheet is explained.

Scenario 1: uploading ESPON M4D formatted data to CityBench

1. Download ESPON project data from http://database.espon.eu/db2/search, for example "Potential accessibility, rail, ESPON average"
2. Metadata: use following correspondence table between CityBench upload sheet and ESPON M4D template to populate fields:

Table AIV-1. Correspondence table between CityBench upload sheet and ESPON M4D template.

CityBench upload sheet	ESPON M4D template	M4D sheet	M4D cell
official indicator id	Code	Indicator	A3 (usually)
name of indicator	Name	Indicator	B3
description	Abstract	Indicator	C3 (usually)
unit	Numerator / Denominator Name	Indicator	[varies]
data source	Project	Dataset	B3
geographic level	Object type	Data	B4
indicator theme	Theme	Indicator	B7

3. Indicator data: copy the desired indicator columns and the Unit code column from row 4 downward (tab Data):

					RAIL_I_ESPON		RAIL_I_ESPON	
2					2001		2006	
3	Unit code	Object type	Version	Name	2001	Source	2006	Source
4	AT111	NUTS3	2006	Mittelburgenland	78.80000305	1	81.5	1
5	AT112	NUTS3	2006	Nordburgenland	82.90000153	1	85.40000153	1
6	AT113	NUTS3	2006	Südburgenland	63.20000076	1	65.19999695	1
7	AT121	NUTS3	2006	Mostviertel-Eisenwurzen	102.1999969	1	108.9000015	1
8	AT122	NUTS3	2006	Niederösterreich-Süd	93.40000153	1	96.80000305	1
9	AT123	NUTS3	2006	Sankt Pölten	105.9000015	1	115.8000031	1
10	AT124	NUTS3	2006	Waldviertel	74.30000305	1	77.69999695	1
11	AT125	NUTS3	2006	Weinviertel	72.80000305	1	77.40000153	1
12	AT126	NUTS3	2006	Wiener Umland/Nordteil	95	1	98.09999847	1
13	AT127	NUTS3	2006	Wiener Umland/Südteil	97	1	102.0999985	1
14	AT130	NUTS3	2006	Wien	112.5	1	119.0999985	1
15	AT211	NUTS3	2006	Klagenfurt-Villach	88.30000305	1	91.30000305	1
16	AT212	NUTS3	2006	Oberkärnten	87.90000153	1	93.30000305	1
17	AT213	NUTS3	2006	Unterkärnten	81.30000305	1	79.19999695	1
18	AT221	NUTS3	2006	Graz	84.80000305	1	84.69999695	1
19	AT222	NUTS3	2006	Liezen	84.59999847	1	85.19999695	1
20	AT223	NUTS3	2006	Östliche Obersteiermark	88.40000153	1	89.90000153	1
21	AT224	NUTS3	2006	Oststeiermark	59	1	58.70000076	1
22	AT225	NUTS3	2006	West- und Südsteiermark	70.40000153	1	68.90000153	1
23	AT226	NUTS3	2006	Westliche Obersteiermark	64.5	1	64.90000153	1
24	AT311	NUTS3	2006	Innviertel	106	1	114.5	1
25	AT312	NUTS3	2006	Linz-Wels	119	1	128.3000031	1
26	AT313	NUTS3	2006	Mühlviertel	88.59999847	1	95.19999695	1
27	AT314	NUTS3	2006	Steyr-Kirchdorf	95	1	101.6999969	1
28	AT315	NUTS3	2006	Traunviertel	108.1999969	1	114.8000031	1
29	AT321	NUTS3	2006	Lungau	70.19999695	1	72.80000305	1
30	AT322	NUTS3	2006	Pinzgau-Pongau	99.09999847	1	108.5999985	1
31	AT323	NUTS3	2006	Salzburg und Umgebung	123.8000031	1	134.3999939	1
32	AT331	NUTS3	2006	Außerfern	85.40000153	1	93.5	1
33	AT332	NUTS3	2006	Innsbruck	126.5999985	1	143.6000061	1
34	AT333	NUTS3	2006	Osttirol	80	1	83.90000153	1
35	AT334	NUTS3	2006	Tiroler Oberland	105.5	1	117.8000031	1
36	AT335	NUTS3	2006	Tiroler Unterland	125.9000015	1	140.5	1
37	AT341	NUTS3	2006	Bludenz-Bregenzer Wald	120.9000015	1	127.9000015	1
38	AT342	NUTS3	2006	Rheintal-Bodenseegebiet	129.8000031	1	136	1
39	BE100	NUTS3	2006	Arr. de Bruxelles-Capitale / Arr. van Brussel-Hoofdstad	251.6000061	1	255.5	1
40	BE211	NUTS3	2006	Arr. Antwerpen	220.3999939	1	220	1
41	BE212	NUTS3	2006	Arr. Mechelen	228.3999939	1	229.6999969	1
42	BE213	NUTS3	2006	Arr. Turnhout	187.3999939	1	188	1
43	BE221	NUTS3	2006	Arr. Hasselt	197.1000061	1	204.3999939	1
44	BE222	NUTS3	2006	Arr. Maaseik	179.1999969	1	181.6999969	1
45	BE223	NUTS3	2006	Arr. Tongeren	197.6000061	1	207.3000031	1
46	BE231	NUTS3	2006	Arr. Aalst	212.8999939	1	212.1999969	1
47	BE232	NUTS3	2006	Arr. Dendermonde	216.1999969	1	213.8999939	1
48	BE233	NUTS3	2006	Arr. Eeklo	176.1999969	1	175.5	1
49	BE234	NUTS3	2006	Arr. Gent	219.8999939	1	220.3999939	1
50	BE235	NUTS3	2006	Arr. Oudenaarde	213.1000061	1	210.8000031	1
51	BE236	NUTS3	2006	Arr. Sint-Niklaas	218.6999969	1	217.5	1
52	BE241	NUTS3	2006	Arr. Halle-Vilvoorde	217.1000061	1	221.3000031	1
53	BE242	NUTS3	2006	Arr. Leuven	233.3999939	1	241.3999939	1
54	BE251	NUTS3	2006	Arr. Brugge	201.6999969	1	200.6000061	1
55	BE252	NUTS3	2006	Arr. Diksmuide	189.6000061	1	186.8999939	1
56	BE253	NUTS3	2006	Arr. Ieper	199.8000031	1	196.6999969	1

Dataset / Indicators / Sources / Data

Figure AIV-1. Example of ESPON project data sheet.

4. Paste into CityBench upload sheet and change year(s) in column header(s) if necessary:

	A	B	C	D
1				
2	**official indicator id**	RAIL_I_ESPON		
3	**name of indicator**	Potential accessibility, rail, ESPON average		
4	**description**	Potential accessibility by rail, standardised (ESPON=100)		
5	**unit**	standardized value / 1		
6	**data source**	ACC Update		
7	**geographic level**	NUTS 3 version 2006		
8	**indicator theme**	Connectivity		
9	**Unit_code**	2001	2006	
10	AT111	78.80000305	81.5	
11	AT112	82.90000153	85.40000153	
12	AT113	63.20000076	65.19999695	
13	AT121	102.1999969	108.9000015	
14	AT122	93.40000153	96.80000305	
15	AT123	105.9000015	115.8000031	
16	AT124	74.30000305	77.69999695	
17	AT125	72.80000305	77.40000153	
18	AT126	95	98.09999847	
19	AT127	97	102.0999985	
20	AT130	112.5	119.0999985	
21	AT211	88.30000305	91.30000305	
22	AT212	87.90000153	93.30000305	
23	AT213	81.30000305	79.19999695	
24	AT221	84.80000305	84.69999695	
25	AT222	84.59999847	85.19999695	
26	AT223	88.40000153	89.90000153	
27	AT224	59	58.70000076	
28	AT225	70.40000153	68.90000153	
29	AT226	64.5	64.90000153	
30	AT311	106	114.5	
31	AT312	119	128.3000031	
32	AT313	88.59999847	95.19999695	
33	AT314	95	101.6999969	
34	AT315	108.1999969	114.8000031	
35	AT321	70.19999695	72.80000305	
36	AT322	99.09999847	108.5999985	
37	AT323	123.8000031	134.3999939	
38	AT331	85.40000153	93.5	
39	AT332	126.5999985	143.6000061	
40	AT333	80	83.90000153	
41	AT334	105.5	117.8000031	
42	AT335	125.9000015	140.5	
43	AT341	120.9000015	127.9000015	
44	AT342	129.8000031	136	
45	BE100	251.6000061	255.5	
46	BE211	220.3999939	220	
47	BE212	228.3999939	229.6999969	
48	BE213	187.3999939	188	
49	BE221	197.1000061	204.3999939	
50	BE222	179.1999969	181.6999969	
51	BE223	197.6000061	207.3000031	
52	BE231	212.8999939	212.1999969	
53	BE232	216.1999969	213.8999939	
54	BE233	176.1999969	175.5	
55	BE234	219.8999939	220.3999939	
56	BE235	213.1000061	210.8000031	

Figure AIV-2. CityBench upload sheet template with sample ESPON project NUTS 3 indicator values.

5. Save completed CityBench upload sheet in location and under name of choice.
6. Go to CityBench admin tool (http://espon.geodan.nl/CityBench3/admin.html) and click on button “Add new indicator”. The following popup screen is displayed:

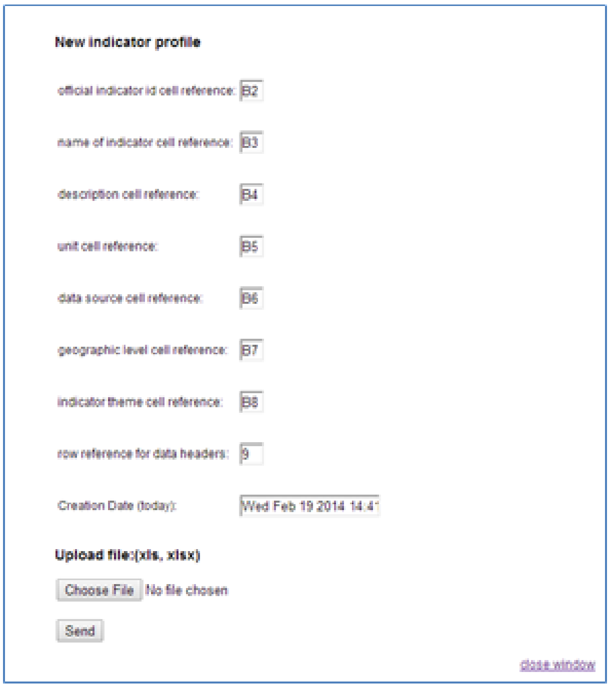


Figure AIV-3. CityBench Admin Tool interface: “Add new indicator” dialogue box.

7. If CityBench upload sheet is filled in as shown in step 4, the pre-filled cell references are correct. If necessary, cell references and/or Creation Date may be changed here.
8. Click button “Choose File” to browse to the location where saved CityBench upload sheet is stored, click “Open” and then “Send”; now the data is being uploaded to CityBench.

Scenario 2: uploading Eurostat data to CityBench

1. Download Eurostat data from http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database, for example “Population on 1 January by age groups and sex - larger urban zone (urb_lpop1)”.
2. Metadata in Eurostat tables is provided either at the top of the table (see example below) or at the bottom of it.
3. Indicator data: copy the desired indicator columns and the Cities column from below the column header row downward:

	A	B	C	D	E	F	G	H	I	J	K	L
1	Population on 1 January by age groups and sex - larger urban zone [urb_lpop1]											
2												
3	Last update	06.02.14										
4	Extracted o	11.02.14										
5	Source of d	Eurostat										
6	Short Descr	Short Description is not available										
7	INDIC_UR	DE1001V - Population on the 1st of January, total										
8												
9	CITIES	CITIES(L)/TI	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
10	BE001L2	Bruxelles / I	:	:	:	:	:	:	:	2468014	2531106	2565164
11	BE002L2	Antwerpen	:	:	:	:	:	:	:	1047451	1067584	1079819
12	BE003L2	Gent	:	:	:	:	:	:	:	573955	581784	584925
13	BE004L2	Charleroi	:	:	:	:	:	:	:	465277	465277	487217
14	BE005L2	Liège	:	:	:	:	:	:	:	720401	729977	738250
15	BE006L2	Brugge	:	:	:	:	:	:	:	222603	223862	224935
16	BE007L2	Namur	:	:	:	:	:	:	:	213514	216095	217841
17	BE008L1	Leuven	:	:	:	:	:	:	:	213816	216714	218458
18	BE009L1	Mons	:	:	:	:	:	:	:	192323	193956	195157
19	BE010L1	Kortrijk	:	:	:	:	:	:	:	169478	170288	171064
20	BE011L1	Oostende	:	:	:	:	:	:	:	124601	126160	127196
21	BG001L2	Sofia	:	:	:	:	:	:	:	:	1533269	1535819
22	BG002L2	Plovdiv	:	:	:	:	:	:	:	:	545934	544910
23	BG003L2	Varna	:	:	:	:	:	:	:	405118	417127	417293
24	BG004L2	Burgas	:	:	:	:	:	:	:	281027	279684	278508
25	BG005L1	Pleven	183683	181562	179098	178110	176620	175394	173837	173837	163076	161371
26	BG006L2	Ruse	:	:	:	:	:	:	:	197184	188004	187050
27	BG007L2	Vidin	:	:	:	:	:	:	:	69507	66316	65256
28	BG008L2	Stara Zagor	:	:	207506	206780	205884	205278	204201	:	163128	162113
29	BG009L1	Sliven	:	:	:	:	:	:	:	128249	125323	:
30	BG010L1	Plovdiv	:	:	:	:	:	:	:	116964	113195	112203
31	BG011L1	Shumen	:	:	:	:	:	:	:	108020	99990	99381
32	BG013L1	Yambol	:	:	:	:	:	:	:	103602	98386	97271
33	BG014L1	Haskovo	:	:	:	:	:	:	:	96499	94230	93305
34	BG015L1	Pazardzhik	:	:	:	:	:	:	:	121366	114938	:
35	BG016L1	Blagoevgrad	:	:	:	:	:	:	:	:	96975	96520
36	BG017L1	Veliko Tarnc	:	:	:	:	:	:	:	88724	88733	88417
37	BG018L1	Vratsa	:	:	:	:	:	:	:	75842	73956	72877
38	CZ001L1	Praha	:	1964750	:	:	2054032	2099282	2131451	2156097	:	:
39	CZ002L1	Brno	:	728101	:	:	757359	763276	767535	770568	:	:
40	CZ003L1	Ostrava	:	1153876	:	:	1151749	1152387	1149740	1145851	:	:
41	CZ004L1	Plzen	:	351701	:	:	360273	367325	369321	369217	:	:
42	CZ005L1	Ústí nad Lal	:	245356	:	:	249399	:	251443	251631	:	:
43	CZ006L1	Olomouc	:	:	:	:	230607	231339	231843	232226	:	:
44	CZ007L1	Liberec	:	246469	:	:	255997	258637	260185	260979	:	:
45	CZ008L1	České Bude	:	179369	:	:	184256	185584	186681	187799	:	:
46	CZ009L1	Hradec Král	:	159293	:	:	161349	162377	163011	163378	:	:
47	CZ010L1	Pardubice	:	159981	:	:	163926	166519	167481	168446	:	:
48	CZ011L1	Zlín	:	193068	:	:	192988	193154	192927	192639	:	:
49	CZ013L1	Karlovy Var	:	121430	:	:	119165	119923	119432	119289	:	:
50	CZ014L1	Jihlava	:	108292	:	:	111257	112031	112501	112707	:	:

Data / Data2 / Data3 / Data4 / Data5

Figure AIV-4. Example of Eurostat data sheet.

4. Paste into CityBench upload sheet and change year(s) in column header(s) if necessary:

	A	B	C	D
1				
2	**official indicator id**	DE1001V		
3	**name of indicator**	Population on the 1st of January, total		
4	**description**	Short Description is not available		
5	**unit**	Inhabitant		
6	**data source**	Eurostat		
7	**geographic level**	Larger Urban Zone (LUZ)		
8	**indicator theme**	Economy and Population		
9	**Unit_code**	**2010**	**2011**	**2012**
10	BE001L2	2468014	2531106	2565164
11	BE002L2	1047451	1067584	1079819
12	BE003L2	573965	581784	584925
13	BE004L2	465277	465277	487217
14	BE005L2	720401	729977	738250
15	BE006L2	222603	223862	224935
16	BE007L2	213514	216095	217841
17	BE008L1	213816	216714	218458
18	BE009L1	192323	193956	196157
19	BE010L1	169478	170288	171064
20	BE011L1	124601	126160	127196
21	BG001L2		1533269	1536819
22	BG002L2		545934	544910
23	BG003L2	405118	417127	417293
24	BG004L2	281027	279684	278508
25	BG005L1	173837	163076	161371
26	BG006L2	197184	188004	187050
27	BG007L2	69507	66316	65256
28	BG008L2		163128	162113
29	BG009L1	128249	125323	
30	BG010L1	116964	113195	112203
31	BG011L1	108020	99990	99381
32	BG013L1	103602	98386	97271
33	BG014L1	96499	94230	93305
34	BG015L1	121366	114938	
35	BG016L1		96975	96520
36	BG017L1	88724	88733	88417
37	BG018L1	75842	73956	72877
38	CZ001L1	2156097		
39	CZ002L1	770568		
40	CZ003L1	1145851		
41	CZ004L1	369217		
42	CZ005L1	251631		
43	CZ006L1	232226		
44	CZ007L1	260979		
45	CZ008L1	187799		
46	CZ009L1	163378		
47	CZ010L1	168446		
48	CZ011L1	192639		
49	CZ013L1	119289		
50	CZ014L1	112707		
51	CZ016L1	116797		
52	CZ018L1	127218		

Figure AIV-5. CityBench upload sheet template with sample Eurostat Urban Audit LUZ indicator values.

5. Next steps: same as in Scenario 1.

Scenario 3: uploading EEA data to CityBench

1. Download EEA data from http://www.eea.europa.eu/data-and-maps/data, for example Corine Land Cover 2006 raster data.
2. Metadata: derive from Metadata info provided for each dataset on the EEA website.
3. Indicator data, for example CLC: extract desired land use class(es) for LUZ 2012 delineations.
4. Copy and paste into CityBench upload sheet and change year(s) in column header(s) if necessary:

	A	B
1		
2	**official indicator id**	CLC141_CLC142
3	**name of indicator**	Total % urban green / sport / leisure areas
4	**description**	Total percentage of CLC classes Green urban areas and Sport and leisure facilities
5	**unit**	Ratio
6	**data source**	EEA
7	**geographic level**	Larger Urban Zone (LUZ)
8	**indicator theme**	Environment
9	**Unit_code**	**2006**
10	AT001L2	0.62
11	AT002L2	0.13
12	AT003L2	0.26
13	BE001L2	2.09
14	BE002L2	2.83
15	BE003L2	0.90
16	BE005L2	0.81
17	BG001L2	0.63
18	BG002L2	0.29
19	CH001L1	0.64
20	CH002L1	1.45
21	CH003L1	0.54
22	CH004L1	0.53
23	CY001L1	0.46
24	CZ001L1	0.91
25	CZ002L1	0.35
26	CZ003L1	0.50
27	DE001L1	0.76
28	DE002L1	1.03
29	DE003L2	1.10
30	DE004L1	3.04
31	DE005L1	1.17
32	DE007L1	0.69
33	DE008L2	1.27
34	DE009L2	0.26
35	DE011L1	2.56
36	DE012L1	0.57
37	DE013L1	1.60
38	DE014L1	0.55
39	DE017L0	1.54
40	DE018L1	0.66
41	DE019L2	0.44
42	DE020L1	0.94
43	DE025L1	0.94
44	DE027L1	0.26
45	DE028L1	0.18
46	DE032L1	0.55
47	DE033L2	0.39
48	DE034L1	0.75
49	DE035L1	1.01

Figure AIV-6. CityBench upload sheet template with sample EEA indicator values, calculated for LUZ 2012 areas.

1. Next steps: same as in Scenario 1.

Annex V. Feasibility of requested changes to Webtool

Requested change	Present in final delivery of CityBench Webtool?	Estimated labor costs (days)	Remarks
Supporting the deployment of the application to the ESPON server, in particular consider technical requirements from the ESPON's web service provider – Infeurope.	Yes	N/A	Special attention to social media ETL. Being discussed with Infeurope.
Make sure that the tool is working in most common browsers (Mozilla, Explorer, Chrome).	Yes	N/A	The Webtool is compatible with IE, Firefox and Chrome (recent versions).
Make clearer the meaning of the size of the circles in the map and in the radial map, make more explicit the central city (place the city search/holder inside the map?). Explain (add a legend) what the circles mean. In general there is a need to add an explanation on similarity and comparison concepts.	Yes	N/A	A legend at the top of the map was added that includes the circle sizes and values (e.g. 0, 0.5, 1). Also the central city is part of the legend which makes it more explicit.
Check the possibility to maximize the size of the screen for displaying info, in particular check the possibility to open the map in a new window that can go full screen.	No	12 days	As this will have impact on the complete layout of the tool, extensive adjustments to the User interface will have to be performed.
Add hover tool-tip when hovering the completeness pie chart (say what it is/means)	Yes	N/A	Done.
Graphs need a facelift (do not use grey background with black, some labels overlap values).	Yes	N/A	The appearance of the graphs was changed in line with the overall layout.
When starting a "new" round of a search/comparison the users need all the time to click the browser button "back", so perhaps a new functionality on the top of the interface, called " start over" or "new search" introduced, therefore user wouldn't have to use "back" button on the browser.	Yes	N/A	A link "Back to wizard" is added.
In the radial view at first it is not easy to understand that a user can switch between different indicators, because the button is somehow disappearing in the interface, perhaps it could be made bigger or different somehow.	Yes	N/A	Visualization of the radial view has been optimized by adding a legend. Also the switch is now in the legend.
In the description of the indicators, if possible include a weblink to the source.	Yes	N/A	Done.
Social media: "job searches" as a social media indicator.	No	15 days	The ability to make custom social media indicators is not available at this moment.
Social Media: Only one indicator should be used the 3 indicators feel roughly similar "crisis", use one.	Yes	N/A	"Unemployment awareness" is being stored daily, together with the "crisis awareness". So, just two indicators for social media (duplicated for daily and yearly data).
Try to use the daily to demonstrate the power of real time information.	No	5 days	A possible solution would be: "Last: day / week / month / year" as options for showing daily indicator values.

Rename "crisis" to "crisis perception" or something more descriptive.	Yes	N/A	“Crisis” has been renamed to “Crisis perception”.
Provide a short note giving instructions to the CU on how to collect the social media indicator and keep it updated in the tool.	Yes	N/A	The instructions for maintaining the deployed system have been prepared.
Add the possibility to export maps/graphs/etc...	No	2 - 3 days	Print Screen can still be used as a workaround, the dedicated export takes approximately 2 to 3 days to create.
Check the possibility to provide a simple file (.xls) including the list of cities included in the tool, plus the indicators (geographical level and year of reference available). / Add a new service in the home screen that compiles an overview of what can be found in the tool in an “.xls” a list of all cities and all indicators available. So no data, just a sort of blink into the database. Ana, can you create such a service that downloads a link from the interface?	Yes	N/A	A download button in the upper toolbar of the Webtool (next to FAQ) enables the download of the requested information in CSV format, which is both compatible with Excel and readable for non-MS Excel users (cross platform compatibility).
Do not include the download data option. Users should be guided to the original data sources for additional info/analysis.	Yes	N/A	Weblinks to the original data will be provided under the “info” button.
On the graph show e.g. “Ireland” instead of “average for dublin’s country average”	Yes	N/A	
Implement layout adjustments on the draft version of the European Urban Benchmarking Tool aimed at streamlining the ESPON tool websites to be delivered by the ESPON CU to the Lead Partner after November 2013 (based on the feedback from the ESPON’s layout service provider)	Yes	N/A	Done.

Annex VI. CityBench Webtool Technical Report

The Webtool has been designed and implemented through an agile approach, where continuous implementation-testing-feedback loops have been developed. The last prototype delivered, result from all the discussions and testing is a map-centered with text boxes and dropdown lists options for visualization of cities and indicators, fed from its own database through a set of get-services. It is complemented with an excel based Admin tool, and can be described as follows.

VI.1 Specific functionality

The idea is to implement the ESPON required scenarios, and allow all the individual use cases. From the initially web tool analysis and together with the use cases, the project arose some natural language questions to run the use cases and fit users' requirements, such as:

- I want to find my city. Where can I specify its name? Can I click on the map?
- I'm not sure about the correct spelling of a certain city. Can I try with some letters? Does the web tool provide auto-complete functionality [Can I click on the map?]?
- I would like to compare (select) cities sharing some characteristics: size (area), administrative level (capital, second tier city…), zone (Eastern Europe…), population (inhabitants, density...), unemployment, and other indicators (to study)
- I would like to compare (select) cities within a distance/ time from my city.
- To which cities does it make sense to compare a given city?
- Which are the data available for a (list of) city/ies?
- Can I download these data?
- Can I see only a set of indicators/themes (being the themes a pre-selected set of indicators)?
- What is the performance [position/interval in the list] of a given (list of) city/ies for each indicator? This is, can I see the position in the list instead of the value for the given indicator? Can I rank indicators by best performance [position in the list] (if one city)?
- Can I select number and type of intervals to divide an indicator to allocate my city in?
- When comparing several cities, can I choose a way of ordering indicators? [Best to worse performance, most to least populated data, grouped by themes… for each city]
- What is the quality/information we have about these data?
- What is the value of a certain indicator for a year (timeframe) in all/a set of cities? Can I have the top X cities? Can I see the temporal evolution?
- Can I see the position in the list instead of the value for the given indicator?

From this starting list, and as a result from the different implementation- testing-feedback- implementation loops, the final Webtool provides the functionality listed below:

- Choose (type) city to be compared from the 172 regular-covered ESPON space cities with relevant data (after entering a/a set of character/s it shows the available list of cities starting with this/those character/s)

- Show all the available indicators (official and unofficial sources) and their metadata
- Select a certain indicator
- Show cities based on given indicator value-frames
- Show indicator interval of selected cities
- Show indicator value of a selected city
- Show the values/intervals of selected indicators from different years
- Filter cities by Transnational Programme or Typology
- Find similar cities: the radial view where the first selected city is in the middle of the graph and the other cities are distribute around it, where closeness represents similarity
- Compare different cities (LUZ) using intuitive visualizations tools such as graphs
- Self-define composite indicators (up to three indicators together to create a similarity ranking)
- Upload a new indicator and delete or update an existing indicator

Which respond to the discussions with the steering committee and questionnaires feedback, whose concerns on functionality are included in the following list:

- Pre-selection of cities –allow users to select cities by territorial co-operation programme areas (transnational and cross-border) or using typologies (such as mountain cities, coastal, etc.). These thematic and geographic pre-selection would allow urban practitioners the opportunity to see benchmarking and networking potential.
- Extended benchmarking – to allow benchmark not just between cities, but also to compare performances in relation to the national capital city, as well as national and EU (averages). Still to be decided, if it should be used a derived average from the additional cities (and national capital) available in the CityBench database, or using a different (national) indicator.
- Scale classification – having 7-9 points is seen as too many and may be reduce understandability. User research should be carried to decide on the appropriate number of classes (3, 4, 5, or 7), being i.e. 5 the default number of classes presented to the user, or if it is possible, to allow user created classes.
- Composite indicators –composite indicators should be approached with great care, as they present both methodological and presentation challenges.
- A two-tier approach will be used, based on the availability of data. The TPG should note that coverage across the ESPON space is required for this limited number of cities. Using a first subset of cities, while considering the future use of all LUZ is considered a good compromise approach between full usability and longer-term coverage aspirations.
- Economic and smartness indicators are considered crucial; the development of these indicators is on-going. Indicators on e.g. skills, innovation, enterprise births and deaths, competition and R&D expenditure might prove very useful in determining the economic and "smart" potential of an area, in addition, Development/Investment demands, Energy and internal accessibility are also being considered.
- To respond to the previous point, the use of alternative, multilevel and innovative data sources is encouraged. Which includes: Unofficial indicators available at country-level data (esp. regulatory/ease of doing business: www.doingbusiness.org); Unofficial city data (TomTom, hotels databases,

citations indexes, most livable cities, UN habitat ranking, OpenStreetMap). The providence of this data should be honestly and transparently presented.

- Exclude indicators with less than 50% coverage while considering those above this threshold.
- Easy to use data loading methodologies and tools to warranty the future sustainability and updates of the tool by the ESPON CU (as administrators of the tool)

VI.2 Other requirements

On the usability context, the Webtool must be easy to learn and understand and it must be intuitive and effective to use (able to answer the questions of the user), for example, which other cities are similar to LUZ X, for a set of indicators?

On the data context, the system must use data from the ESPON Database and explore data from external sources to include indicators not available or not fully covered in the ESPON database and try to scale data to the LUZ regions.

The CityBench project will preferably use Open Source Software components to implement the required functionality. Source code and configuration file developed for the CityBench project specifically will be made available as Open Source.

The project should use Open Standards to give access to the CityBench data and must be able to use relevant Open Standards to use data from other sources (e.g. INSPIRE data from NMCA's, data from EIB and Eurostat).

The system must be deployable at the ESPON CU and it must be maintainable by employees or contractors of ESPON CU.

VI.3 Prototype Architecture

As described previously, CityBench application architecture follows the service oriented architecture paradigm. This basically means that the provided data and functionality will be served through web services. The unique entry point to these services is implemented in the form of a Web Geoportal.

Having this in mind, we follow two main principles:

1. The application must be scalable and extensible. The CityBench application has to be adapted to new requirements and new data. This means that the application has to be designed in a modular way. Furthermore, these modules should have a well-known functionality and interface so the application is more easily maintain. Most important decisions in this line are: to design the architecture and implement the components according to international standards and initiatives. As we will describe further in the document, we rely on European directives to design the components serving required functionality and on international standards to implements their interfaces and serve the data.

2. Underlying technological complexity is transparent for the user. The use of standards, and communication protocols of the CityBench system should be transparent for the user. The CityBench web portal provides user with a unique entry point to the system, offering an intuitive manner to handle data and providing advanced search and visualization methods to derive useful information out of the massive amount of existing data.

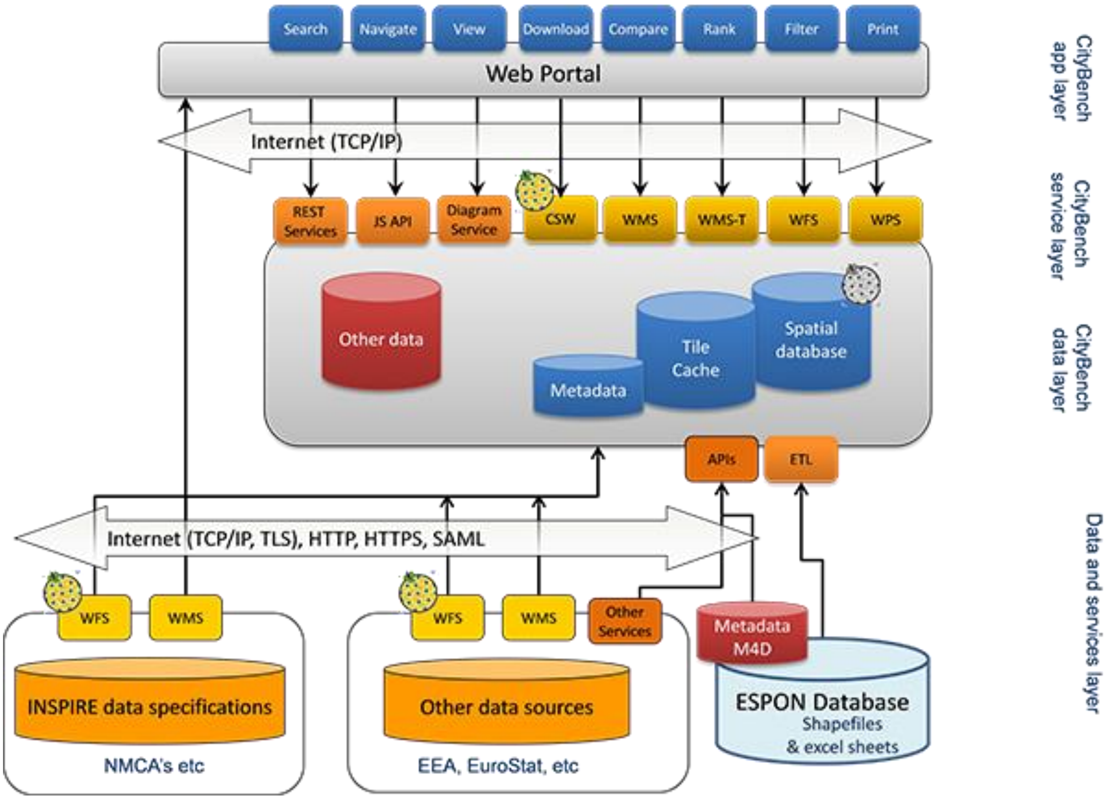


Figure 1. - Conceptual Architecture CityBench

CityBench project architecture follows the European directive INSPIRE (INSPIRE, 2007) which defines a classical service-oriented three-layer architecture. CityBench application extends this architecture to include the required functionality in the form of Network services and the client applications in the form of a Web Portal.

As INSPIRE recommends, CityBench project implements its components according to the Open Geospatial Consortium (OGC), which defines open standards for the successful exploitation of geospatial resources on the Internet.

Figure 1 shows the CityBench project conceptual architecture. On top of the figure we can observe the three main layers, which are application, service, and data layer. At the bottom, the data and service layer represents how the project architecture access and harvest data coming from external sources such as ESPON database. We detail in the next sections how this is implemented in the prototype.

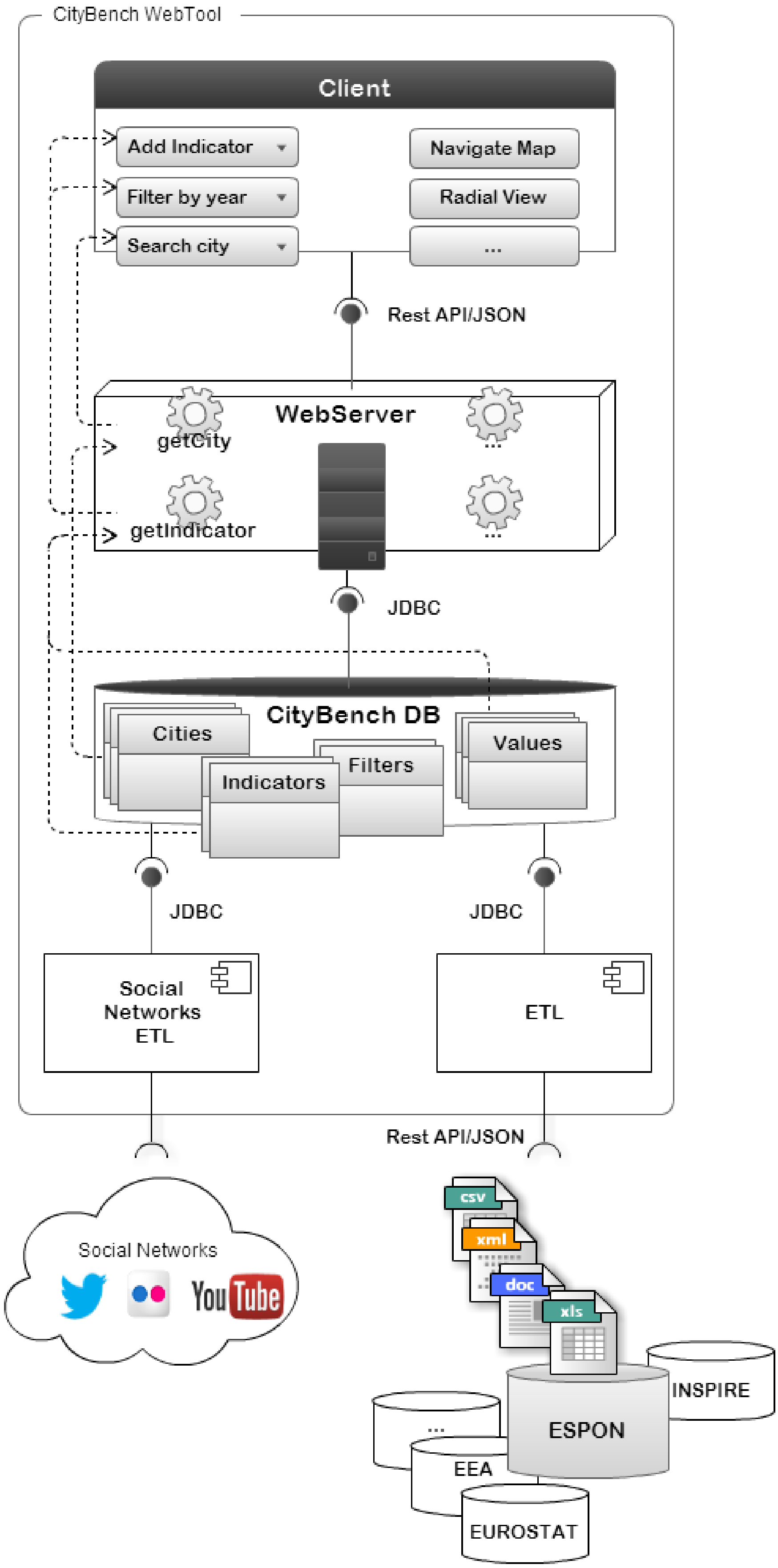


Figure 2. Prototype architecture

The Figure 2 shows the prototype technical architecture. It presents a more detailed view of the architecture highlighting the main technological components. From the general perspective the system is divided in 4 main parts (indicated layers in Figure 1) it includes: the web client at the application layer, where the different functionality and data about cities and indicators are presented to the users, the services, implemented in Java language, provide the functionality to access and retrieve the different data about cities and indicators through a well-defined Rest API, the communication with the database uses JDBC protocol. The database implementing the CityBench data model containing the project data to be delivered by the services, and the ETLs, in charge of populating the database with the data obtained from different sources at the data and service layer. Most of functionality of the client is related with the presentation of the indicators in a per city basis. The data of the indicators and the cities is provided by the services hosted in the WebServer presented in the figure. The services are used as a bridge for providing access to the data previously loaded through the ETLs and stored in the CityBench database.

VI.4 Design

The CityBench architecture is a service-oriented architecture composed of three layers. Each of these layers has distinct characteristics and its components implement a distinct role in the CityBench system.

VI.4.1 External Data Services Layer

This layer contains the data sources used by the CityBench system. Also the external services, which implement interfaces to external data sources, are part of this layer. Its components are:

ESPON Database

The ESPON database is available as a vast collection of Excel sheets. In order to be used by the CityBench system, relevant sheets will be loaded into a spatial database through an ETL process (see ETL).

Eurostat database

Similar to the ESPON database, the Eurostat database is mainly available as a vast collection of Excel sheets. In order to be used by the CityBench system, relevant sheets will be loaded into a spatial database through an ETL process (see ETL).

European Environmental Agency (EEA) database

The EEA provides land cover data (Corine Land Cover) in both vector (not used) and raster format and data on air pollution in raster format. The raster layers are processed in ArcGIS to calculate land cover and air pollution values per LUZ. These values are subsequently loaded into the CityBench system.

OpenFlights data

OpenFlights data is available as simple csv files (without column headers). Before the indicator values derived from this data can be loaded into CityBench, the data has to be processed in Excel and ArcGIS (a.o. calculating values per LUZ).

IFC / World Bank data

IFC / World Bank data is available as an Excel sheet. In order to be used by the CityBench system, relevant values (i.e. Ease of Doing Business in EU27+4 countries) will be loaded into a spatial database through an ETL process (see ETL). Country values will be assigned to each LUZ they contain.

Data derived from Social Networks (VGI)

The CityBench project complements official data and indicators with VGI. We have defined a set of indicators reflecting the social impact of current topics such as unemployment and the economic crisis. Also, short-term personal mobility (tourism) may be derived from VGI. This way, social networks such as Twitter, Flickr and YouTube become external data sources used by the CityBench services to populate the CityBench database.

VI.4.2 Data Layer

This layer contains the components managing the project data, these are:

Spatial database

The project database containing structured data obtained from the external data sources such as the ESPON database. This database is structured in a data model specifically designed and implemented for the CityBench project.

Metadata

Metadata about data and indicators available in the CityBench project will also be stored in the database.

Official data ETL (Admin Tool)

The Extract Transform and Load (ETL) component will extract the data from the ESPON database, transform it into CityBench data model (implemented by the spatial database), and load them into the spatial database.

Social data ETL

CityBench project includes also an ETL for social content. Indicators about data extracted from the social networks are generated retrieving data from social networks for each city and transforming them to derive indicator information. The data is finally loaded in the CityBench database.

VI.4.3 Services Layer

The CityBench Service Layer contains the services that implementing well-known and standard-based interfaces allow to work with the data in the data layer. The definition of these interfaces together compound the CityBench API used by the CityBench tool, but also other applications could access the data in the CityBench data layer through this API. The components of the service layer are:

Rest Web services

REST/JSON services will supply functionality and data to the CityBench clients. The interface and supported functionality will be clearly documented. This services main functionality is to deal with the discovery and retrieval of data available in the database. They will allow the CityBench web tool to retrieve all data available about cities and indicators in the way requested by the users. These services will be integrated as part of the Data and Metadata API.

Data and Metadata Services API

The data and metadata API will make it possible to consume CityBench information in an efficient way. This API will be documented so developers (from the CityBench project, but also external developers) can implement this interface integrate CityBench data and functionality.

JavaScript client API

In order to use the different CityBench data and functionality in an Internet browser, JavaScript libraries could be easily developed with further effort. These libraries would use other Open Source JavaScript libraries such as JQuery and D3. The JavaScript CityBench libraries could be stored on the CityBench webserver to be deployed at ESPON CU.

VI.5 Implementation

This section describes the implementation process of the current prototype: technical decisions on programming languages, data formats, communication protocols and APIs, and platforms of deployment.

VI.5.1 Architecture and deployment

The prototype is deployed according to the architecture described in the design section. On one hand, database and data are prepared and deployed in a server. On a second phase, web services will be implemented to give the web tool access to the data. The web tool will prepare and advanced and intuitive user interface for the users to interact with the data and functionality.

As a platform for deployment, CityBench project uses different test servers; both partners Geodan and UJI perform local tests in local servers. For more reliable tests on this first prototype and its exposition to the ESPON community, CityBench application has been replicated and deployed in two servers, one managed by UJI and other Geodan.

Database instance:

- http://geo4.dlsi.uji.es/
- database: CityBench

Services available at

- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/

Client application available at:

- http://espon.geodan.nl/citybench3/#

Following sections will describe other implementation details of the external data and services, data, service and application layer.

VI.5.2 External Data and Services layer

As the management of these data and services is out of scope for the CityBench project, no implementations will have to be carried out in this layer.

VI.5.3 Data Layer

Data

From the initial analysis we have finally identified a set of derived functionality that requires, at least, the availability of following data:

- Cities (like ID, Name, Alternative Name or Geographical definition). They allow locating cities and searching for them, and relating them to other information in the database.
- City characteristics (like, Country, Languages, Typology, Transnational Regions or Indicators having information about the city). They allow us filtering and selecting groups of cities, which might have same interests.
- Indicators (like ID, Name, Description or Units). They allow us searching for indicators, and relating them to other information in the database.
- Indicators characteristics (like Data Source, Theme or Geographical level to which the data are referenced). They allow us filtering and selecting groups of indicators, which might define a coherent comparison scenario.
- Values (like Indicator, City or Year). They are the core of this tool and allow us to compare and evaluate relations between cities, like similarity or 'distance', graphically represented in the web tool.

Database

The Spatial Database has been implemented using an Open Source product, namely PostgreSQL with PostGIS extension for spatial support. This database stores at the moment data and metadata about cities and indicators.

Data Model

We have identified relevant entities and their relations; this has resulted in a logical data model that has been implemented using Postgres database. Figure 3 shows a diagram of the CityBench data model:

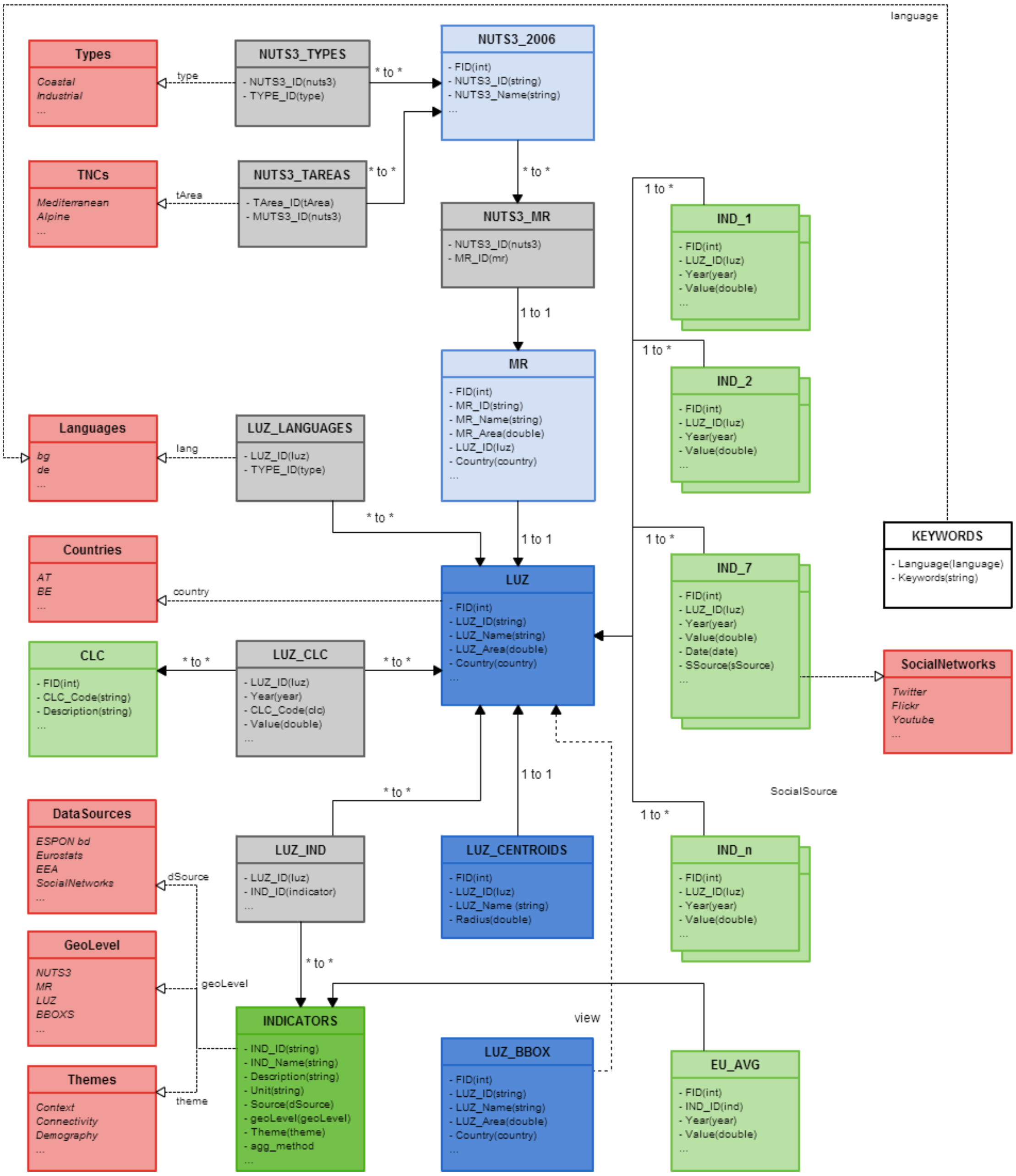


Figure 3. CityBench Data model

There we can see, in the middle column, the cities data and their relation with the different geographical levels (blue tables). In the left part of the model we can see the filtering data (red tables), for both cities and indicators, and at its bottom, the definition of indicators (dark green tables) and their relation with cities. The right part of the model represents the indicators values (light green tables), each row representing an indicator theme, and a special one for storing European averages for each indicator, each year.

Grey tables are connecting tables. White tables represent data modelled but not yet loaded into the database. Some views are created from other tables through a query, and are not stored and wasting storage space.

ETLs

Both the official data and the social data ETLs are tools implemented using Java programing language and SQL to support modifications of the CityBench database.

These tools work on the first load and the futures updates after delivering the tool to ESPON.

Admin Tool

The official ETL takes into account the necessary data; the original format and data source and the CityBench data model to define data transformations from standard ESPON excel sheets to the CityBench data model. It also delivers a document describing how new data from other ESPON projects should be submitted. See more info on Annex IV.

Social Media ETL

The social data ETL collects data obtained through the APIs offered by different social networks. The data collected is used for gathering information from European cities regarding different subjects of interest. For this is utilized different BigData technologies used for the collection, storage and processing of the data. The social network selected for extracting data is Twitter. The information extracted is loaded into the CityBench database.

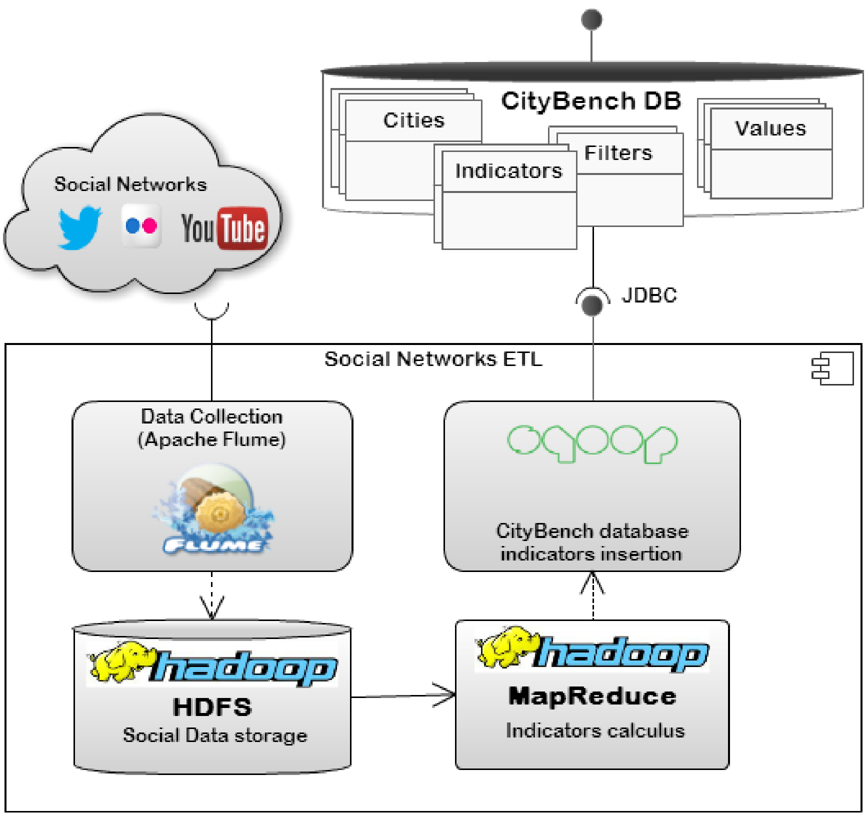


Figure 4 Social Data ETL

The social data ETLs is a tool implemented using Java programing language and SQL to support modifications of the CityBench database. This tool works on the first load and continuously (automatically) updates data after delivering the tool to ESPON.

For this is utilized different BigData technologies used for the collection, storage and processing of the data. The social network selected for extracting data is Twitter. The information extracted is loaded into the CityBench database.

The indicators are generated through the use of different software used for collecting; storing and aggregating social networks data. The general workflow is presented in Figure 4. The whole process is implemented using technologies for handling big data volumes, as is the case data coming from social networks. The data is collected by a service called Apache Flume, which allows reliably obtaining data from the different social networks and storing it in HDFS. HDFS (Hadoop Distributed File System) is the storage component of the social ETL. This file system is in charge of keeping the data obtained from the social networks for later analysis and indicators value extraction.

The data analysis is performed later by algorithms implemented using Hadoop MapReduce framework, which allows running algorithms appropriate for extracting the indicators values from potentially large data volumes.

Later, the indicators values are loaded into the CityBench database by a Sqoop service which takes them form the output of the MapReduce algorithms implemented.

Besides these tools, first load of the database has been supported using the Open Source Software PgAdmin.

The social networks data collection includes three main task, collection and storage, processing of the data and output of the results to an external database. Although the initial implementation only includes Twitter data (tweets) analysis the aim is to extend an initial core for processing data from other social networks. Due to the potential volume of data that we are aiming to collect and process we considered convenient to use big data tools.

The **data collection** at this point of the project is performed on Twitter services using the Twitter streaming API. The collection is performed using Apache Flume, a distributed and reliable service for collecting, aggregating and moving large amounts of data. A general depiction of how Apache Flume works is in the following picture.

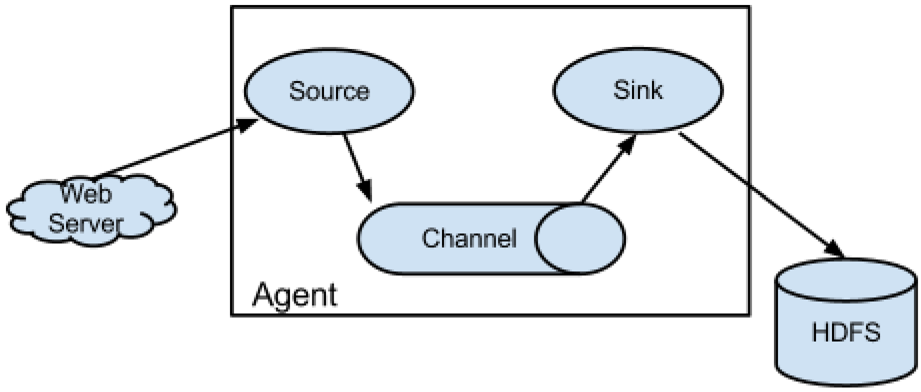


Figure 5. Data Collection

For this we implemented a source (as shown in the picture) that consumes the data streamed by the Twitter services and allows to group and store the data in HDFS, the Hadoop distributed file system. This enables the distributed processing of the data by using Hadoop library and tools. The source is based on an implementation explained in the article http://blog.cloudera.com/blog/2012/09/analyzing-twitter-data-with-hadoop/ and adapted for the purpose of the project.

Apache Flume services were configured to store tweets in a file path structure that enables the temporal classification of them and at the same time allows processing of the data based on temporal ranges. At this point of the project the folder structure is this:

/user/flume/tweets/seqs/%{query}/%{city}/%Y/%m/%d/%H

Where query is an id that refers to the %{query} set to the Twitter services, %{city} refers to a classification performed by the flume source implemented and %Y, %m, %d, %H refers to temporal properties of the data (in this order year, month, day, and hour).

For extending this setup in order to collect data from other social networks we plan to implement sources capable to request data from such networks, possibly by resting them through their APIs.

Data storage relies in HDFS. This file system allows to reliably storing the data collected while enabling distributed data processing and several optimizations for working with large amounts of data. The one of the most important features of HDFS are its replication capability. The files system guarantees that the data stored is going to have the certain number of replicas (configured). A depiction of the architecture of HDFS is shown in the following picture.

HDFS is designed to be used with a write once, read multiple times philosophy. It supports a group of roles that can be implemented by different machines in a cluster. A Namenode machine would be in charge of keeping a catalog (Metadata) of all the files stored in blocks of the file system and being the access point for reading the data stored. Another role is the data node, which is in charge of actually storing the data.

The previous picture shows how a client (an application operating on the file system) can write a piece of data (in a DataNode) and this will be replicated among other datanodes in the cluster. Later, another client could request access to a file (to the NameNode) and access the file in the closest datanode location. Data locality is supported by defining the rack a given datanode belongs to, which serves for achieving fastest access at processing time.

In our project, we are storing the data in sequence files of 25 Mb (might be increased in the future) for avoiding problems such as “the small file problem”. The sequence files contain the tweets delivered by Twitter streaming API for an hour in a certain city. Although there is not guarantee that the files will have this size due to the file classification most of the data is stored in files of this file size.

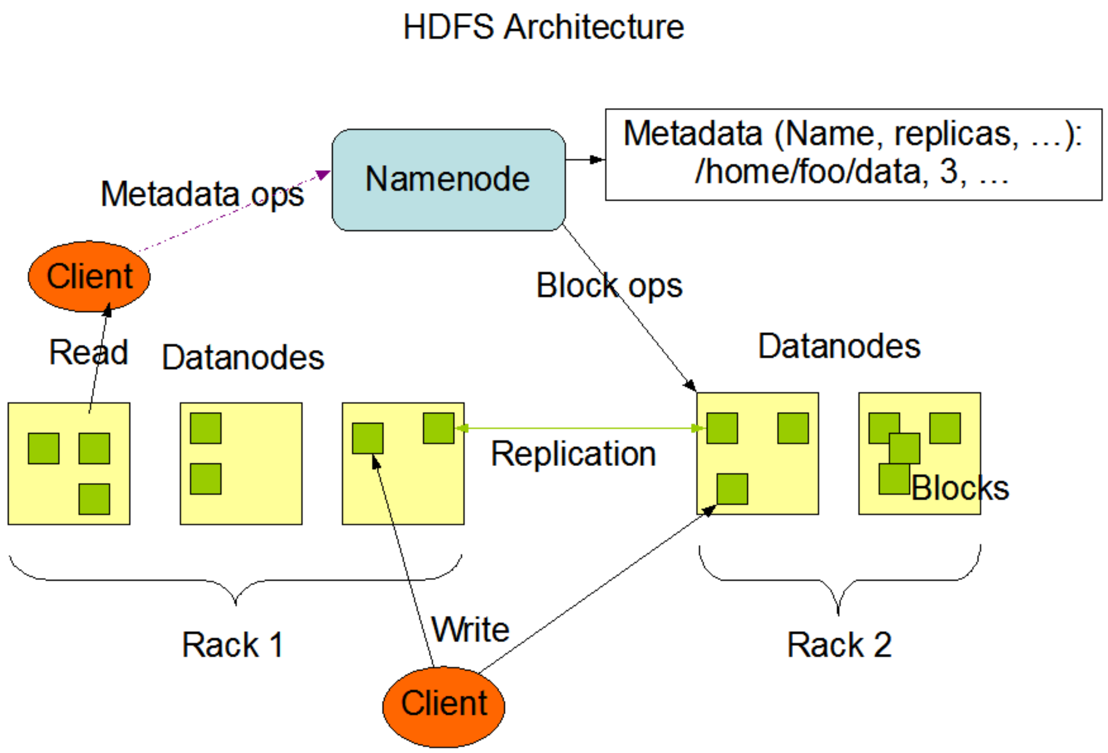


Figure 6. Data storage

For **processing** the data we are using Hadoop MapReduce, a framework for writing applications for processing large amount of data. The processing performed at this point of the project is very basic, mainly including summarization and simple classification algorithms. The input of the algorithm is the sequence files mentioned before. The Hadoop framework divides the sequence file in "splits" to be processed independently in a distributed environment, so that the algorithm processes one data item (tweet) at the time.

The result of processing the data is a set of text files containing (key, value) pairs containing the information summarized. Some of the output of the processing over a dataset of more than 3 million is shown in the next table:

Key	**Value**
job_201309251303_0141:-ECONOMY:TWEE:COUNT_DATA_ITEMS	3700455
job_201309251303_0141:-ECONOMY:TWEE:NO_TERM_IN_DATA_ITEM	1440643
job_201309251303_0141:-ECONOMY:TWEE:UNK:credit	546361
job_201309251303_0141:-ECONOMY:TWEE:UNK:crisis	423628
job_201309251303_0141:-ECONOMY:TWEE:UNK:debt	287649
job_201309251303_0141:-ECONOMY:TWEE:UNK:economic_crisis	4364
job_201309251303_0141:-ECONOMY:TWEE:UNK:euro	201119
job_201309251303_0141:-ECONOMY:TWEE:UNK:financial	216896

job_201309251303_0141:-ECONOMY:TWEE:UNK:investment	106929
job_201309251303_0141:-END_TIME	2013-10-04T11:38:27.000 +02:00
job_201309251303_0141:-BEGIN_TIME	2013-09-25T17:16:33.000+02:00

For gathering the **output** of the previous processing step and writing it to a database we are using Apache Sqoop, a tool for transferring data between Hadoop and relational databases. By using Sqoop we transferred the output of the MapReduce. Sqoop allows calling a stored procedure that handles the data insertion in the database. In our project we use the stored procedure for decoding the data produced by the MapReduce process (as shown in the previous table) and inserting the data in an intermediate table to be processed further.

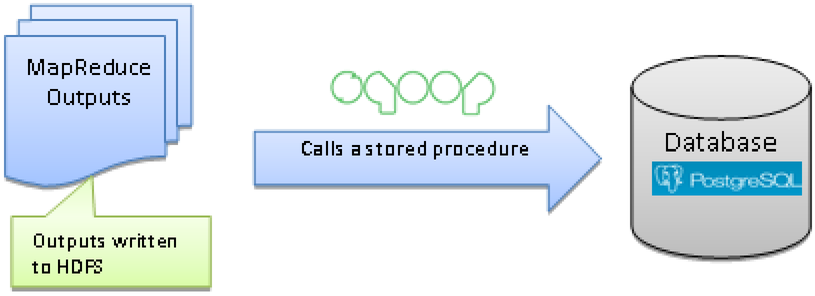


Figure 7. Data output to external databases

One of the purposes of our project is to have data collected and inserted in our database in a daily basis. In order to achieve this is necessary run the process described previously can be used Apache Oozie. Apache Oozie provides services **for executing and scheduling a data workflow** specified through an XML document. The workflow we described in our project is very simple and includes the steps mentioned before: data processing and output to external databases. The workflows contain a set of actions that can be assigned certain precedence; Oozie allows specifying (among others) MapReduce and Sqoop actions which fit the requirements of our project. Oozie also provide the concept of coordinator, a service that allows specifying a schedule for a workflow. Using this feature we have implemented a coordinator that runs our workflow at the beginning of each day for processing the data of the previous day.

VI.5.4 Services & API

According to INSPIRE, our system should provide a set of predefined functionality through standard-based services. These services (in yellow in Figure 1) offer the discover, view, and download functionality. This GIS services to manage the map has been skipped from the prototype. The functionality has been moved to the web client and future modifications on this will depend new requirements analyzed from the feedback received from the prototype users.

Other more specific functionality is offered by the rest of services in the CityBench Services Layer depicted in orange in Figure 1. They are components specifically developed for the CityBench system. They have been implemented in Java (server side) and JavaScript (client side). These web services expose a Restful interface and work with well known formats such as JSON, so the API can be easily described and they can be reused for future uses. For extending client compatibility, services are exposed. Next we describe the API to connect to services, available for getting **cities** locations with the supported filters, getting **indicators** metadata with their supported filters, getting indicators **values** for all (a set of) cities and years, getting social media indicators values for all (a set of) cities and days and getting equivalent LUZ entities for a given set of MRs or NUTS3 (**transformers**).

Values services support 1 to 3 indicators, none to n cities and none to n years. When a query parameter is left blank, service will return all items for that query parameter. This means, if you ask for values for one indicator one year, and not specific city, service will return values for that indicator, that year, in all cities.

Cities

getAllCities Service

This service will return the list of all LUZ from the database.

- Takes: -
- Returns: [LUZ_ID, LUZ_Name, long, lat]
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/objects/cities

getAllCities Country filter Service

This service is a query parameter from the previous service and will return the list of all LUZ from the database belonging to a given country

- Takes: countryID
- Returns: [LUZ_ID, LUZ_Name, long, lat]
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/objects/cities?country=AT

getAllCities Typology filter Service

This service is a query parameter from the previous service and will return the list of all LUZ from the database belonging to a given typology.

- Takes: typeID
- Returns: [LUZ_ID, LUZ_Name, long, lat]
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/objects/cities?typology=CAP

getAllCities Trans National Cooperation Programs filter Service

This service is a query parameter from the previous service and will return the list of all LUZ from the database belonging to a given transborder/crossborder area.

- Takes: tnc_code
- Returns: [LUZ_ID, LUZ_Name, long, lat]
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/objects/cities?tnc_code=7

getAllCities Language filter Service

This service is a query parameter from the previous service and will return the list of all LUZ from the database having as official language a given language.

- Takes: languageID
- Returns: [LUZ_ID, LUZ_Name, long, lat]
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/objects/cities?city_lang=nl

getAllCitiesWithTNC Service

This service will return the list of all indicators from the database classified by theme.

- Takes: -
- Returns: {TNCs: [{id:1,name:p1,cities: [{id:x1,name:q1},{id:xn,name:qn}]}] ,…, [{id:n,name:pn,cities: [{id:y1,name:r1},{id:yn,name:rn}]}]}
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/objects/citieswithtnc

getAllCitiesWithTypology service

This service will return the list of all indicators from the database classified by theme.

- Takes: -
- Returns: {Typologies: [{id:1,name:p1,cities: [{id:x1,name:q1},{id:xn,name:qn}]}] ,…, [{id:n,name:pn,cities: [{id:y1,name:r1},{id:yn,name:rn}]}]}
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/objects/citieswithtypologies

Indicators

getAllIndicators Service

This service will return the list of all indicators from the database.

- Takes: -
- Returns: [all Indicators' data]
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/manager/indicators

getAllIndicators Theme filter Service

This service is a query parameter from the previous service and will return the list of all indicators from the database belonging to a given theme.

- Takes: themeID
- Returns: [all Indicators' data]
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/manager/indicators?theme_id=1

getAllIndicators Data Source filter Service

This service is a query parameter from the previous service and will return the list of all indicators from the database whose data proceed from a given data source.

- Takes: dSourceID
- Returns: [all Indicators' data]
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/manager/indicators?source_id=2

getAllIndicators Geographical Level of data capture filter Service

This service is a query parameter from the previous service and will return the list of all indicators from the database whose data has been captured on a given geographical level basis.

- Takes: geolevelID
- Returns: [all Indicators' data]
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/manager/indicators?geolevel_id=C

getAllIndicatorsWithThemes Service

This service will return the list of all indicators from the database classified by theme.

- Takes: -
- Returns: {themes: [{id:1,name:p1,indicators: [{id:x1,name:q1},{id:xn,name:qn}]}] ,..., [{id:n,name:pn,indicators: [{id:y1,name:r1},{id:yn,name:rn}]}]}
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/manager/indicatorsandthemes

Completeness Service

This service will return the number of cities having data for a given indicator or for a given indicator and year.

- Takes: indicatorID or (indicatorID, year)
- Returns: count
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/manager/completeness?name=CB3201I
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/manager/completeness?name=CB3201I&year=2001

IndicatorTheme Service

This service will return the theme from a given indicator.

- Takes: indicatorID
- Returns: themeID
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/manager/indicatortheme?indicator_id=EC2031V

IndicatorGeolevel Service

This service will return the geographical level at whch the data was captured from a given indicator.

- Takes: indicatorID
- Returns: geolevelID
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/manager/indicatorlevel?indicator_id=EC2031V

Values

getIndicator Service

This service will return a list of all cities with the correspondent value for a given indicator or for a given indicator in a given year.

- Takes: indicatorID or (indicatorID, year)
- indicatorID is not NULL
- NULL values for year takes all years
- Returns: {"results ":[{"year":"y1 ","luz_id ":"c1","value":"n1"},{}]}
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/manager/indicatordata?name=CB3201I
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/manager/indicatordata?name=CB3201I&year=2001

getIndicatorPerMRsPerYears Service

This service will return the list of LUZ together with the value and year of the given indicator, when the indicator table is defined at MR geographic level, for a given set of (all) MRs and years.

- Takes: (indicatorID, [MR_ID], [year])
- indicatorID is not NULL
- NULL values for MR and year takes all MR and years
- Returns: {"results":[{"year":"y1","luz_id":"c1","value":"n1"},{}]}
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/manager/indicatorsmrsyears?indicator_id=XX0000X&mrs=DE003M,DE002M,DE001M&years=2001,2006

getIndicatorPerNUTS3PerYears AVG Service

This service will return the list of LUZ together with the aggregated AVERAGE value and LATEST year of the given indicator, when the indicator table is defined at NUTS3 geographic level, for a given set of (all) NUTS3s and years.

- Takes: (indicatorID, [NUTS3_ID], [year])
- indicatorID is not NULL
- NULL values for NUTS3 and year takes all NUTS3 and years
- Returns: {"results":[{"year":"y1","luz_id":"c1","value":"n1"},{}]}
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/extras/ind_luz_from_nuts32_summary?ind_name=XX0001X&ids=AT112,AT125,AT126,AT127,AT130&agg_fn=av_my&years=2001,2006

getIndicatorPerNUTS3PerYears SUM Service

This service will return the list of LUZ together with the aggregated ADDED value and LATEST year of the given indicator, when the indicator table is defined at NUTS3 geographic level, for a given set of (all) NUTS3s and years.

- Takes: (indicatorID, [NUTS3_ID], [year])
- indicatorID is not NULL
- NULL values for NUTS3 and year takes all NUTS3 and years
- Returns: {"results":[{"year":"y1","luz_id":"c1","value":"n1"},{}]}
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/extras/ind_luz_from_nuts32_summary?ind_name=XX0001X&ids=AT112,AT125,AT126,AT127,AT130&agg_fn=sv_my&years=2001,2006

getIndicatorPerNUTS3PerYears MAX Service

This service will return the list of LUZ together with the aggregated MAXIMUM value and LATEST year of the given indicator, when the indicator table is defined at NUTS3 geographic level, for a given set of (all) NUTS3s and years.

- Takes: (indicatorID, [NUTS3_ID], [year])
- indicatorID is not NULL
- NULL values for NUTS3 and year takes all NUTS3 and years
- Returns: {"results":[{"year":"y1","luz_id":"c1","value":"n1"},{}]}
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/extras/ind_luz_from_nuts32_summary?ind_name=XX0001X&ids=AT112,AT125,AT126,AT127,AT130&agg_fn=mv_my&years=2001,2006

getIndicatorPerCountriesPerYears Service

This service will return the list of LUZ together with the value and year of the given indicator, when the indicator table is defined at Country geographic level, for a given set of (all) Countries and years.

- Takes: (indicatorID, [Country_ID], [year])
- indicatorID is not NULL
- NULL values for Country and year takes all Countries and years
- Returns: {"results":[{"year":"y1","luz_id":"c1","value":"n1"},{}]}
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/manager/indicatorscountriesyears?indicator_id=XX0002X&years=2001,2006&countries=AT,BE

getIndicatorsPerCitiesPerYears Service

This service will return an indicators matrix with their units and the available years matrixes with the list of cities, values and country averages, and EU average value for each year and indicator from a given set of (for all) indicators (stored at LUZ geographical level), cities and years.

- Takes: ([indicatorID], [LUZ_ID], [year])
- indicatorID is not NULL (1 to 3)
- indicatorID.geolevel = LUZ
- NULL values for LUZ and year takes all LUZ and years
- Returns: {"indicators":[{"ind_id ":"i1","unit":"u","years":[{"year":"y1","values":[{"luz_id":"c1","value":"n1","country":"m1"},{},{}],"EU":"p1 "},{}]}
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/manager/indicatorscities?names=CB3201I,CB3202I&cities=AT001L2,AT002L2&years=2001,2006

getIndicatorsAllGeo PerCitiesPerYears Service

This service will return an indicators matrix with their units and the available years matrixes with the list of cities, values and country averages, and EU average value for each year and indicator from a given set of (for all) indicators, cities and years.

- Takes: ([indicatorID], [LUZ_ID], [year])
- indicatorID is not NULL (1 to 3)
- NULL values for LUZ and year takes all LUZ and years
- Returns: {"indicators":[{"ind_id ":"i1","unit":"u","years":[{"year":"y1","values":[{"luz_id":"c1","value":"n1","country":"m1"},{},{}],"EU":"p1 "},{}]}
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/manager/indicatorscitiesagg?names=CB3701V,XX0002X,XX0001X&cities=AT001L2,BE001L2&years=2001,2006,2009

getSMIndicatorsPerCitiesPerDays Service

This service will return an indicators matrix of available days matrixes with the list of cities and values for each day and indicator from a given set of (for all) indicators, cities and dates. Date scan be expressed as single days, sets of dates or date ranges.

- Takes: ([indicatorID], [LUZ_ID], [date])
- indicatorID is not NULL (1 to 3 IDs)
- NULL values for LUZ takes all LUZ
- date is not NULL (date/ date,…,date/ date::date)
- Returns: {"indicators":[{"ind_id":"i1","unit":"u","dates":[{"date":"d1","values":[{"luz_id":"c1","value":"n1"},{}]},{}]},{}]}

- geo4.dlsi.uji.es/citybenchservices/webresources/v1/manager/smindicators?indicator_ids=CB5011I,CB5012I&luz_ids=AT001L2,AT002L2&dates=2013-12-20
- geo4.dlsi.uji.es/citybenchservices/webresources/v1/manager/smindicators?indicator_ids=CB5011I,CB5012I&luz_ids=AT001L2,AT002L2&dates=2013-12-20,2014-01-01
- geo4.dlsi.uji.es/citybenchservices/webresources/v1/manager/smindicators?indicator_ids=CB5011I,CB5012I&luz_ids=AT001L2,AT002L2&dates=2013-12-20::2014-01-01

Transformers

MR2LUZ Service

This service will return the LUZ list from a given set of MRs.

- Takes: [MR_ID]
- Returns: [LUZ_ID]
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/extras/mrtoluz?ids=DE003M,DE002M,DE001M

NUTS32LUZ Service

This service will return the LUZ list from a given set of NUTS3s.

- Takes: [NUTS3_ID]
- Returns: [LUZ_ID]
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/extras/nuts3toluz?nuts_ids=AT112,AT125,AT126,AT127,AT130,AT221,AT225,AT312,AT313

C2LUZ Service

This service will return the LUZ list from a given set of Countries.

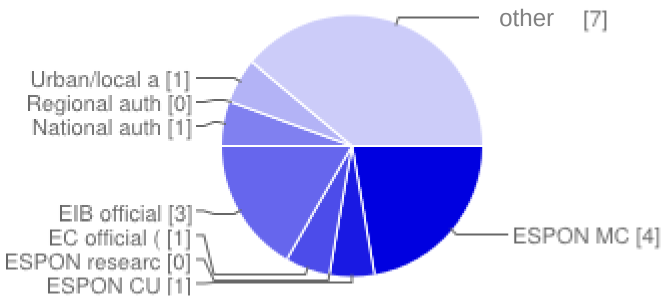
- Takes: [country_ID]
- Returns: [LUZ_ID]
- http://geo4.dlsi.uji.es/citybenchservices/webresources/v1/extras/ctoluz?country_ids=AT,BE

Annex VII. Detailed survey results

From 11 March 2014 to 22 April 2014, a web survey was conducted among stakeholders of the CityBench project. 18 reactions were received. The detailed response to each question is provided in this Annex.

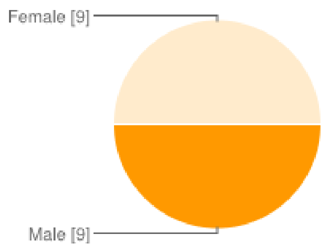
1. Your background

1.1. What is your function?



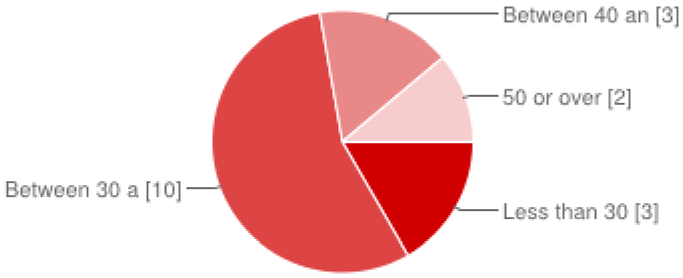
ESPON MC	**4**	22%
ESPON CU	**1**	6%
ESPON researcher	**0**	0%
EC official (practitioner, researcher)	**1**	6%
EIB official	**3**	17%
National authority	**1**	6%
Regional authority	**0**	0%
Urban/local authority	**1**	6%
Overig	**7**	39%

1.2. Gender:



Male	**9**	50%
Female	**9**	50%

1.3. What is your age?



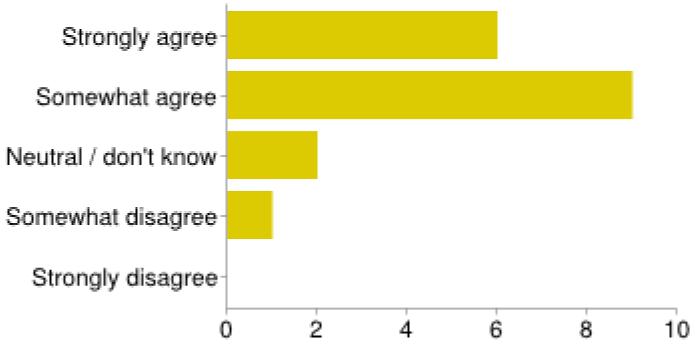
Less than 30	**3**	17%
Between 30 and 39	**10**	56%
Between 40 and 49	**3**	17%
50 or over	**2**	11%

1.4. Country where you work:

Portugal
Luxembourg
France
UK
Czech Republic
Romania
luxembourg
Ireland
Greece
Cyprus
United Kingdom
Belgium

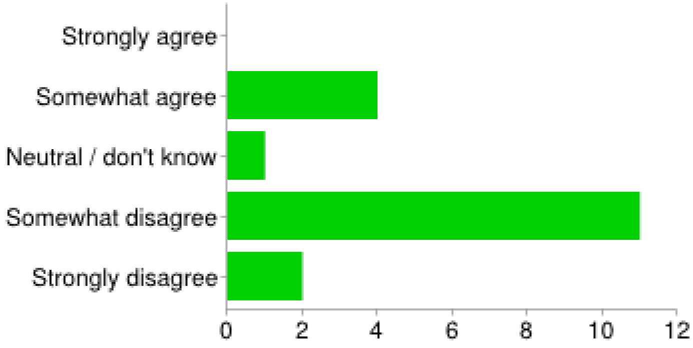
2. Usability

2.1. I enjoyed using the CityBench Webtool



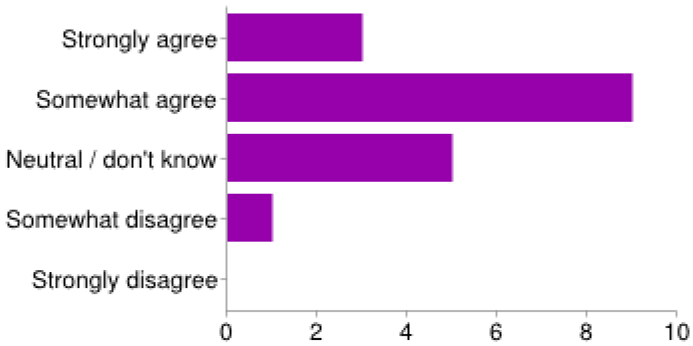
Strongly agree	**6**	33%
Somewhat agree	**9**	50%
Neutral / don't know	**2**	11%
Somewhat disagree	**1**	6%
Strongly disagree	**0**	0%

2.2. I found the Webtool unnecessarily complex



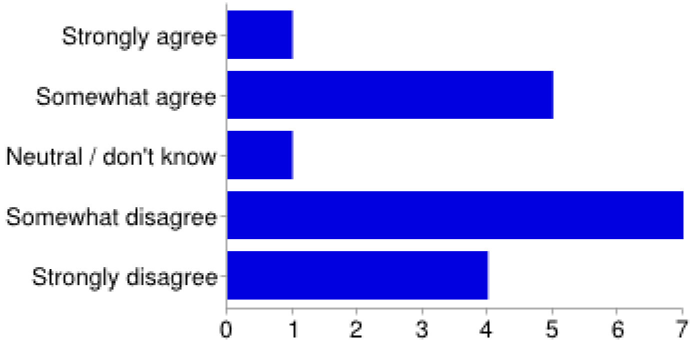
Strongly agree	**0**	0%
Somewhat agree	**4**	22%
Neutral / don't know	**1**	6%
Somewhat disagree	**11**	61%
Strongly disagree	**2**	11%

2.3. I think that I would like to use this Webtool frequently



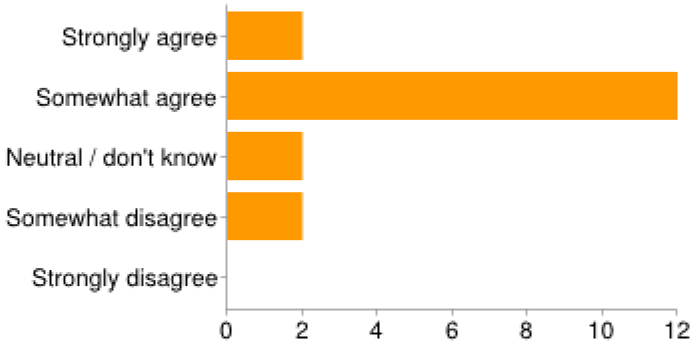
Strongly agree	**3**	17%
Somewhat agree	**9**	50%
Neutral / don't know	**5**	28%
Somewhat disagree	**1**	6%
Strongly disagree	**0**	0%

2.4. I think that I would need the support of a technical person to be able to use this Webtool



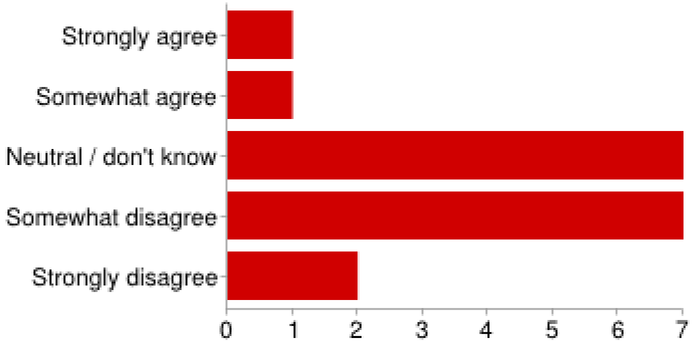
Strongly agree	**1**	6%
Somewhat agree	**5**	28%
Neutral / don't know	**1**	6%
Somewhat disagree	**7**	39%
Strongly disagree	**4**	22%

2.5. I thought the Webtool was easy to use



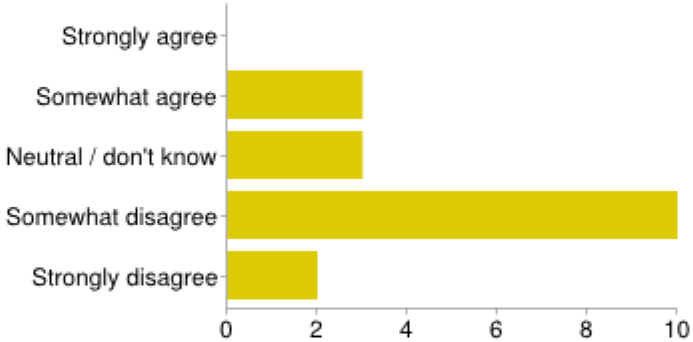
Strongly agree	**2**	11%
Somewhat agree	**12**	67%
Neutral / don't know	**2**	11%
Somewhat disagree	**2**	11%
Strongly disagree	**0**	0%

2.6. I thought there was too much inconsistency in this Webtool



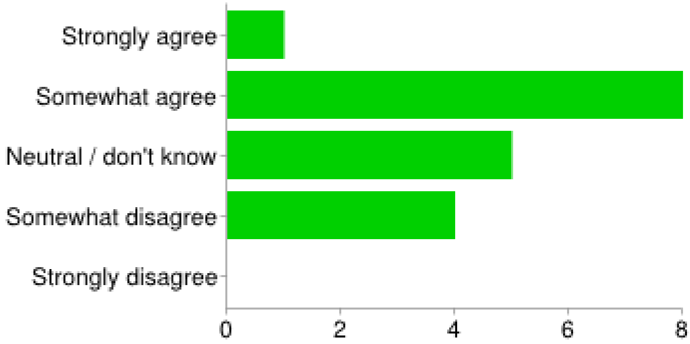
Strongly agree	**1**	6%
Somewhat agree	**1**	6%
Neutral / don't know	**7**	39%
Somewhat disagree	**7**	39%
Strongly disagree	**2**	11%

2.7. I found the Webtool very cumbersome to use



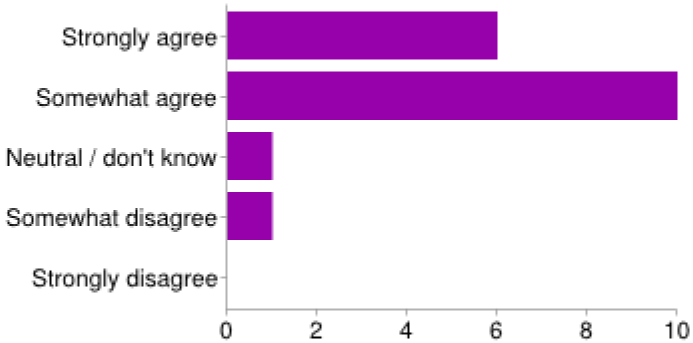
Strongly agree	**0**	0%
Somewhat agree	**3**	17%
Neutral / don't know	**3**	17%
Somewhat disagree	**10**	56%
Strongly disagree	**2**	11%

2.8. I found the various functions in this Webtool were well integrated



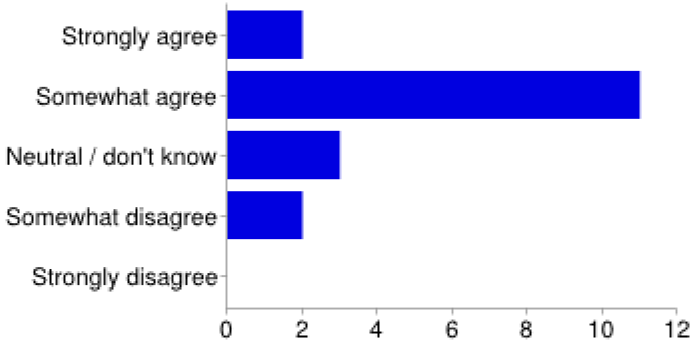
Strongly agree	**1**	6%
Somewhat agree	**8**	44%
Neutral / don't know	**5**	28%
Somewhat disagree	**4**	22%
Strongly disagree	**0**	0%

2.9. I learned to use the Webtool quickly



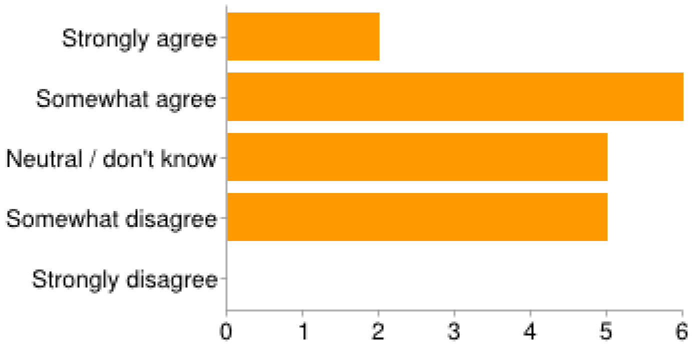
Strongly agree	**6**	33%
Somewhat agree	**10**	56%
Neutral / don't know	**1**	6%
Somewhat disagree	**1**	6%
Strongly disagree	**0**	0%

2.10. I felt very confident using the Webtool



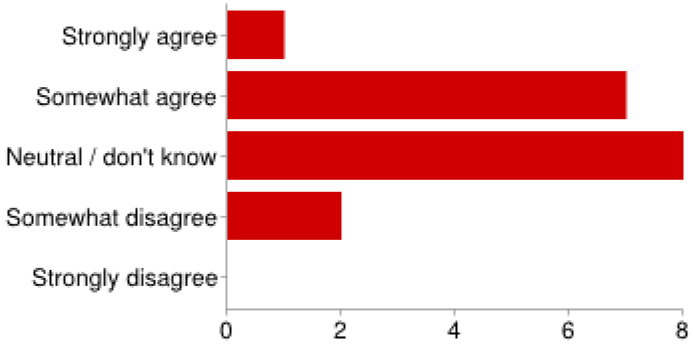
Strongly agree	**2**	11%
Somewhat agree	**11**	61%
Neutral / don't know	**3**	17%
Somewhat disagree	**2**	11%
Strongly disagree	**0**	0%

2.11. I found it easy and quick to recover from mistakes



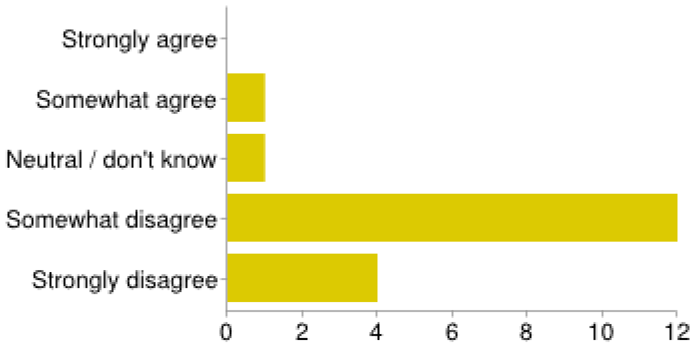
Strongly agree	**2**	11%
Somewhat agree	**6**	33%
Neutral / don't know	**5**	28%
Somewhat disagree	**5**	28%
Strongly disagree	**0**	0%

2.12. I would imagine that most people would learn to use the Webtool very quickly



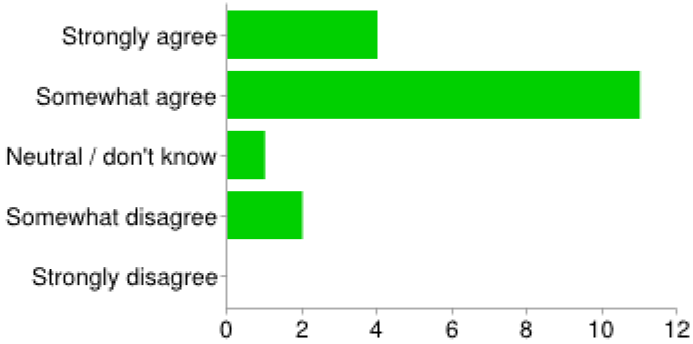
Strongly agree	**1**	6%
Somewhat agree	**7**	39%
Neutral / don't know	**8**	44%
Somewhat disagree	**2**	11%
Strongly disagree	**0**	0%

2.13. I needed to learn a lot of things before I could get going with this Webtool



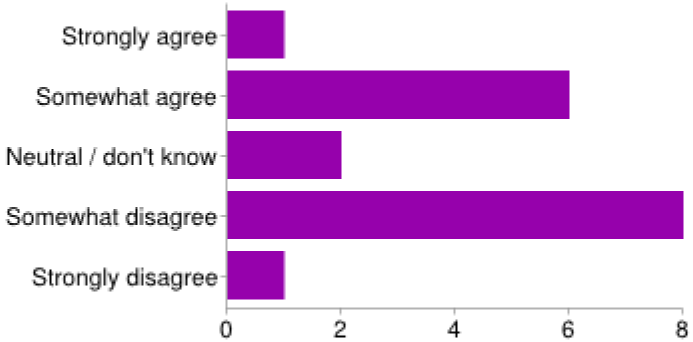
Strongly agree	**0**	0%
Somewhat agree	**1**	6%
Neutral / don't know	**1**	6%
Somewhat disagree	**12**	67%
Strongly disagree	**4**	22%

2.14. I found it easy to remember how to use the Webtool and to understand it



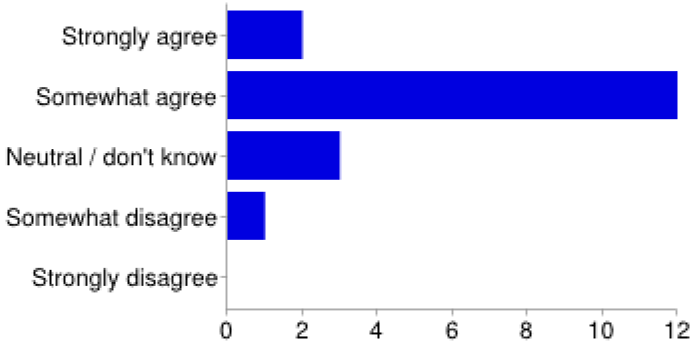
Strongly agree	4	22%
Somewhat agree	11	61%
Neutral / don't know	1	6%
Somewhat disagree	2	11%
Strongly disagree	0	0%

2.15. The Webtool works the way I want it to work



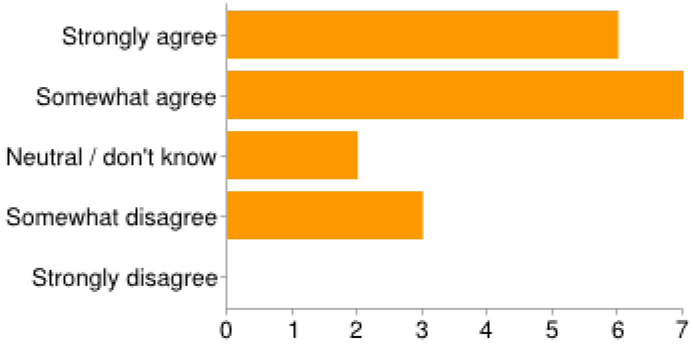
Strongly agree	1	6%
Somewhat agree	6	33%
Neutral / don't know	2	11%
Somewhat disagree	8	44%
Strongly disagree	1	6%

2.17. I would recommend the Webtool to a colleague / friend



Strongly agree	**2**	11%
Somewhat agree	**12**	67%
Neutral / don't know	**3**	17%
Somewhat disagree	**1**	6%
Strongly disagree	**0**	0%

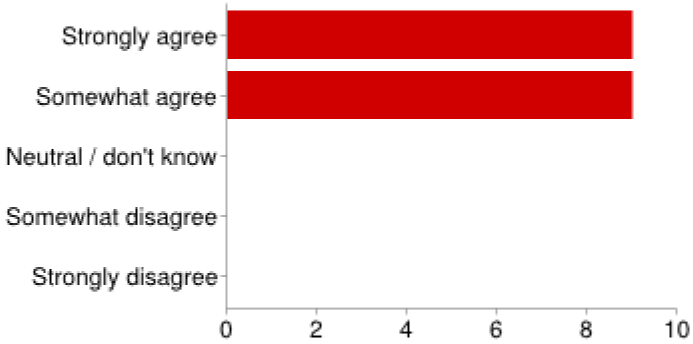
2.18. The Webtool works well in my browser



Strongly agree	**6**	33%
Somewhat agree	**7**	39%
Neutral / don't know	**2**	11%
Somewhat disagree	**3**	17%
Strongly disagree	**0**	0%

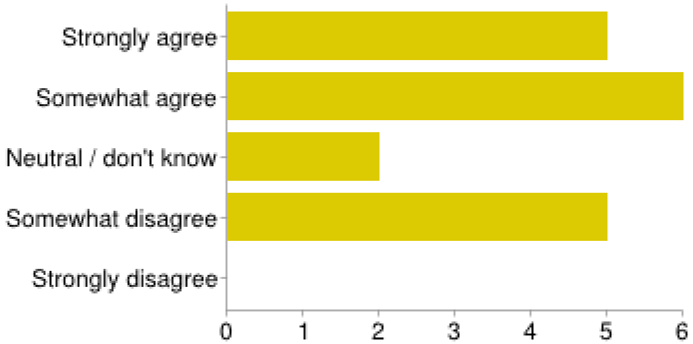
3. Usefulness

3.1. It is useful to compare ONE city with all the other cities based on similarity



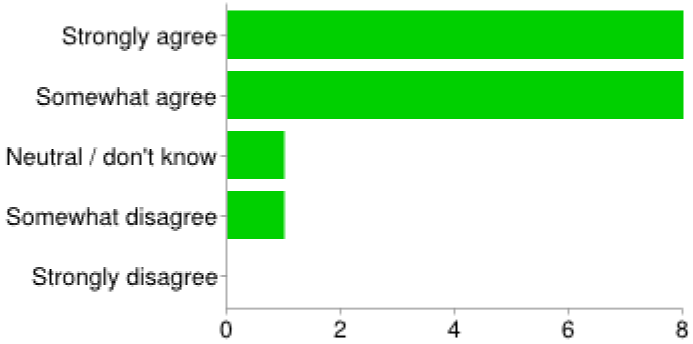
Strongly agree	**9**	50%
Somewhat agree	**9**	50%
Neutral / don't know	**0**	0%
Somewhat disagree	**0**	0%
Strongly disagree	**0**	0%

3.2. Similarity is an easy concept to understand in the Webtool



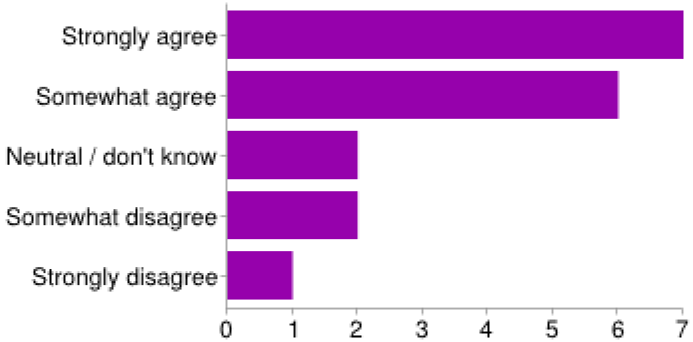
Strongly agree	**5**	28%
Somewhat agree	**6**	33%
Neutral / don't know	**2**	11%
Somewhat disagree	**5**	28%
Strongly disagree	**0**	0%

3.3. Similarity between your city and all other cities is interesting to see in the MAP view



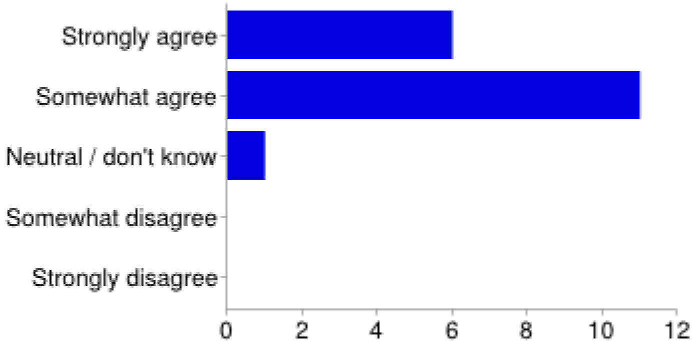
Strongly agree	**8**	44%
Somewhat agree	**8**	44%
Neutral / don't know	**1**	6%
Somewhat disagree	**1**	6%
Strongly disagree	**0**	0%

3.4. Similarity between your city and all other cities is interesting to see in the RADIAL view



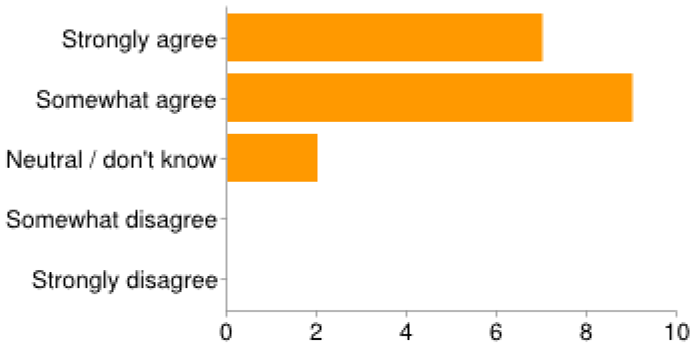
Strongly agree	**7**	39%
Somewhat agree	**6**	33%
Neutral / don't know	**2**	11%
Somewhat disagree	**2**	11%
Strongly disagree	**1**	6%

3.5. It is useful to filter the cities based on typology



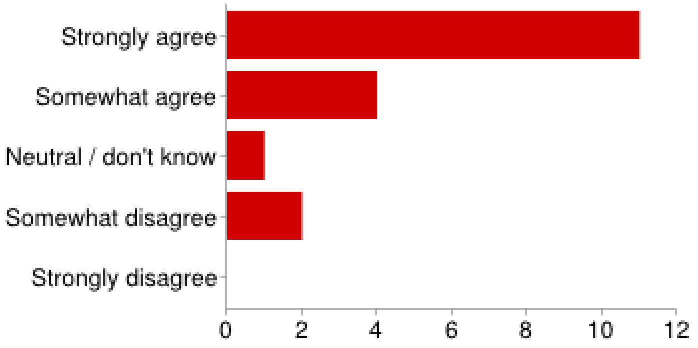
Strongly agree	**6**	33%
Somewhat agree	**11**	61%
Neutral / don't know	**1**	6%
Somewhat disagree	**0**	0%
Strongly disagree	**0**	0%

3.6. It is useful to filter the cities based on a region



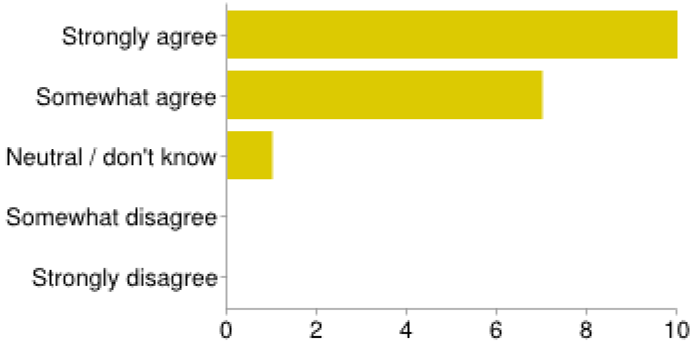
Strongly agree	**7**	39%
Somewhat agree	**9**	50%
Neutral / don't know	**2**	11%
Somewhat disagree	**0**	0%
Strongly disagree	**0**	0%

3.7. The graphs comparing just up to 4 cities are useful



Strongly agree	**11**	61%
Somewhat agree	**4**	22%
Neutral / don't know	**1**	6%
Somewhat disagree	**2**	11%
Strongly disagree	**0**	0%

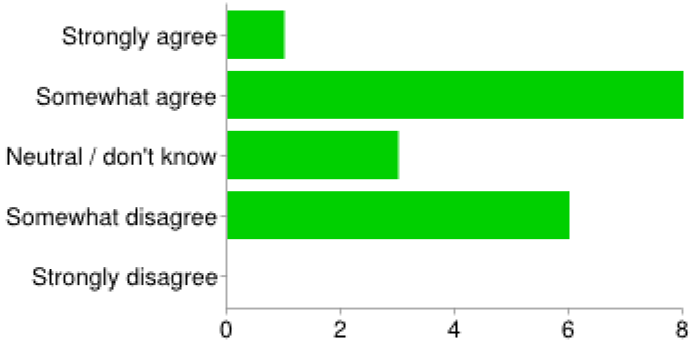
3.8. It is important to access the different years for the indicators to see trends



Strongly agree	**10**	56%
Somewhat agree	**7**	39%
Neutral / don't know	**1**	6%
Somewhat disagree	**0**	0%
Strongly disagree	**0**	0%

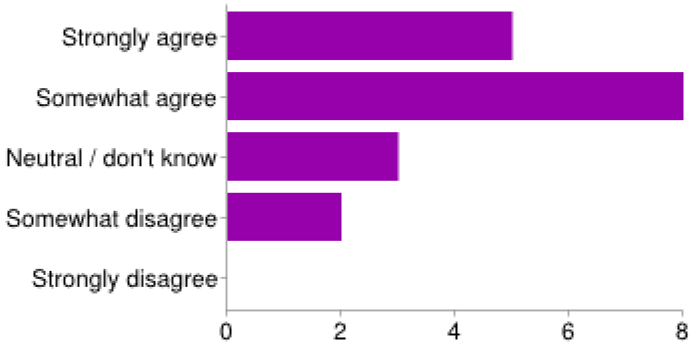
4. Needs or additional functionality

4.1. The system is complete for a quick scan of cities



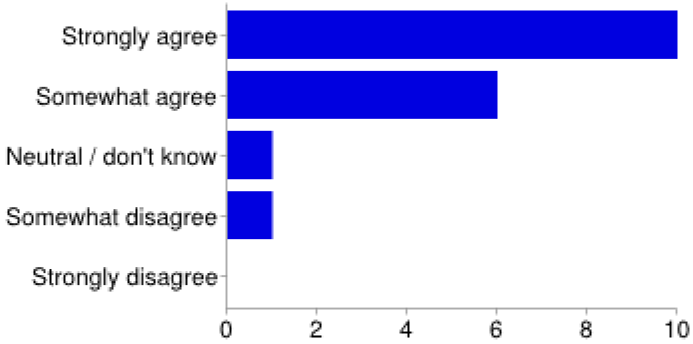
Strongly agree	**1**	6%
Somewhat agree	**8**	44%
Neutral / don't know	**3**	17%
Somewhat disagree	**6**	33%
Strongly disagree	**0**	0%

4.2. I need to download the data for further analysis



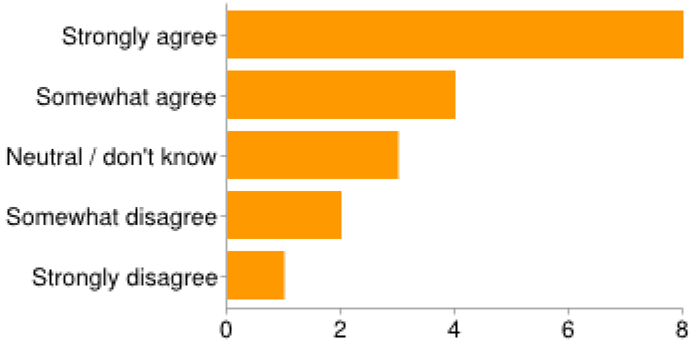
Strongly agree	**5**	28%
Somewhat agree	**8**	44%
Neutral / don't know	**3**	17%
Somewhat disagree	**2**	11%
Strongly disagree	**0**	0%

4.3. I would like to download images with the map, radial or graphs comparison



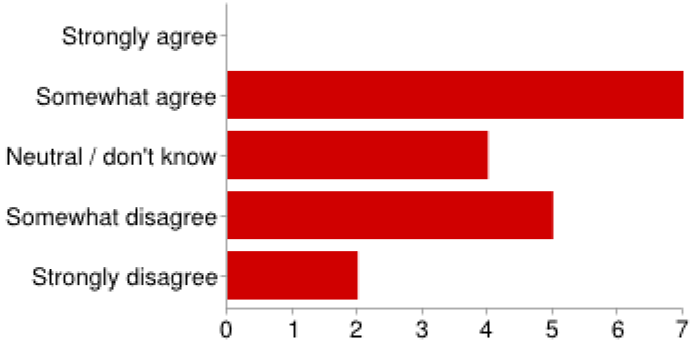
Strongly agree	**10**	56%
Somewhat agree	**6**	33%
Neutral / don't know	**1**	6%
Somewhat disagree	**1**	6%
Strongly disagree	**0**	0%

4.4. I would like to link to a map or radial view using email, facebook or twitter



Strongly agree	**8**	44%
Somewhat agree	**4**	22%
Neutral / don't know	**3**	17%
Somewhat disagree	**2**	11%
Strongly disagree	**1**	6%

4.5. I don’t need to compare to one city, I want to see the real values of indicators



Strongly agree	**0**	0%
Somewhat agree	**7**	39%
Neutral / don't know	**4**	22%
Somewhat disagree	**5**	28%
Strongly disagree	**2**	11%

4.6. Other (Additional functionality not mentioned above you would like to see added)

Very intutive system to use -- interesting to compare regions/cities by typology.
Not possible now to zoom in, in the map view to select a specific city to compare. Would also be good to be able to zoom in to read the map better. Not all indicators are populated for all the cities: this gives a strange result in the map (all circles are equally big), maybe it is useful to have a pop-up indicating for which indicator no data is available, now you have to search yourself by adding and deleting indicators.
The menu for selecting an indicator is not always active.
- would be useful to have a dynamic range of values for "similarity" and keep those values (or custom added ones) when the year of data is changed (currently jumps automatically to min. and max.) - the axes in the radial are not named and not used in a consistent way (this would be an added value in case 1 or 2 dimensions are assessed)... cities are now randomly scattered around. - the colors of the legend in graphs doesn't always correspond with the data in the graph itself
The radial view is not clear and easy to understand
more indicators more cities option to zoom in certain regions (e.g. Eastern Europe)
One disadvantage was the inability to return from the compare/similarity screen in step 3 to steps 1 or 2 in order to adjust my selection, so I would like to see this added.
export graphs, images. better explanation on the meaning of the size of the circles both in radial and map views. plus the meaning of similarity values in the radial view: 1 means equal to and 0 means completely different, i suppose.

5. General
5.1. Please specify positive aspects of the system (max. 3)

Quick comparison of the cities. Quick characteristics along the chosen indicators.
I definitely appreciated the aim of the tool, and I think it could be a very useful value added to MA. I have found the graphs very useful; same comment applies to the map and city selection.
overall good graphic interface map and "radial" a good idea - although radial IS misleading... potential for complementary use with other tools is there (but...-)
- Comparability of cities' indicators - View of regional systems of cities - Analysis of different indicators
Layout overall is generally nice
Very intutive system to use interesting to compare regions/cities by typology Much needed system for learning and knowledge-sharing
-quick and easy to use - pleasant lay-out
The city size on the map provides a great indication of the degree of similarity between cities! The fact that up to 3 indicators can be grouped is also a very powerful tool for comparison.
Quick way to find similarities between cities.
It is attractive. The idea on the similarities is very interesting.
immediate picture of comparison easy to identify the trend and picture of the situation direct access to points of information needed
- easy to use - good quick scan - it sort of seems to operationalise the urban audit (?)
Quick comparison of cities Quick overview of cities ranking for certain indicators
Little input is required from the user in order to get results. The tool is generally intuitive. Makes the data accessible.
innovative, easy to use, fast
Both functions (Compare and Similarity) can be useful for research purposes as well as for professionals working in consulting or technical advice. i'm not sure about the added value for civil servants in local authorities a part from starting basic benchmarking and networking analysis (before applying for URBACT type of networks). In URBACT partners of transnational networks could use the tool to easily visualize data about partnership's differences and complementarities. I wouldn't present CityBench as a tool to support financial engineering instruments in place-based policymaking as far as data are at city level scale and not at a lower scale (neighborhood)

5.2. Please specify aspects of the CityBench Webtool that should be improved or changed and explain why (max. 3)

High degree of similarity of the European towns resulting in too many similar cities calls for more specific indicators, which would allow less cities and better focussing.
It might be useful to add the possibility of extracting data using the tool. It is understood that this is developing prototype, as several things probably should be improved (first of all visualization, in my browser it did not perfectly work). It is not super easy to use, it takes a bit before you get used to it, and this might annoy users (how do you go back, after you perform a comparison?).
navigation not always easy some visual glitches should be sorted out richer range of examples to be provided in diocumantation
- A tour through the tool would be helpful, like a video guiding the user through a practical example - Possibility of downloading data in excel should be useful - Longer time series for the indicators should be useful as well

Several layout glitches (unreadable tags) and some problems with moving between views
The sub-indicators were a bit difficult to understand ie) Under "Economy and Population" - "Agree" it is easy to find a job -- not sure where the data is extract. Under "Social Media" -- what is "crisis". Also useful to have a "back button" and remove symbols like "ø" or ä, ö, ü, (München or København) b/c the database doesn't recognise these symbols.
- lack of indicator availability
The fact that some cities are shown with a "+" sign, which makes it unable to select another city after I finish benchmarking one. For instance, I select "Luxembourg", I benchmark it according to the 1st criterion under Economy and Population. That is fine. I then remove this indicator and I try to select Lublin, which, like other cities, appears with a "+", however I cannot select it.
It's not clear to me what the radial view means. Having more indicators would be interesting.
Sometimes the tool is doing strange things - not active menus, the names of the regions appear several times. It would be better if the definitions for the different types of regions are available.
the radial view is not visible and easy to understand -- when comparing the cities and you also provide the country average, the lines of the cities are not visible or easy to identify -- some more words explaining the indicators are needed (in a subsection or a separate window), in order to understand what exactly these indicators represent
- the tool is only as good as the data in it. To be useful, it needs to be based on up to data information: some of the data used are from 2009... - cities might find it useful to be able to compare on more detailed issues as well: e.g. levels of air pollution, the size of the creative industry. But lacking data is likely to be an obstable here.
There is a bug for the chart legend (colour of legend does not match)
Step 1: I did not know that I needed to click complete after selecting cities in step 1 so I progress to step 2 which would not work for me. Step 2: The drop-drown menu with radio buttons inside it for selecting indicators is a little cumbersome. 3 indicators to compare is a low number, I would like to see more.
export graphs, maps, figures better explanation of size of the circles both in radial and map view. add explanation of similarity values ensure that indicators in particular the ones from social media are updated regularly
1) Explore complementarities with RSFC (www.rsfc.eu). In particular it should be possible to align indicators that in any case should be sensibly increased to make the tool attractive not only to single visitors 2) Link the tool to websites presenting case studies or good practices of sustainable urban development (URBACT, DG REGIO, EUKN, EnergyCities, etc) to create connections between data visualization and concrete ideas and solutions to urban challenges. (e.g. comparing % of urban green area and being able to see how "similar" cities are dealing with public green spaces)

www.espon.eu

The ESPON 2013 Programme is part-financed by the European Regional Development Fund, the EU Member States and the Partner States Iceland, Liechtenstein, Norway and Switzerland. It shall support policy development in relation to the aim of territorial cohesion and a harmonious development of the European territory.

ISBN 978-2-919777-68-6