CITYkeys

Co-funded by the European Commission within the H2020 Programme.
Grant Agreement no: 646440.
2015-02-01…2017-01-31 (24 months).

Deliverable 2.1 Definition of data sets

Revision 0
Preparation date 2016-01-29 (m12)
Due date 2016-01-31 (m12)
Lead contractor............................. TNO

Authors:

Peter Bosch .... TNO
Sophie Jongeneel......................... TNO
Vera Rovers.................................. TNO
Hans-Martin Neumann.................. AIT
Stefan Vielguth .... AIT
Aapo Huovila .... VTT
Isabel Pinto Seppä........................ VTT
Topi Penttinen............................... VTT
Roland van der Heijden .... City of Rotterdam
Elli Kotakorpi................................. City of Tampere
Eva Pangerl .... City of Vienna
Sanja Malnar Neralić .... City of Zagreb
Daniel Sarasa Funes .... City of Zaragoza

Dissemination level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Deliverable administration										
No & name	**D2.1 Definition of data sets**									
Status	Final				**Due**	m12		**Date**	2016-01-31	
Author(s)	Peter Bosch, TNO; Sophie Jongeneel, TNO; Vera Rovers, TNO; Hans-Martin Neumann, AIT; Stefan Vielguth, AIT; Aapo Huovila, VTT; Isabel Pinto Seppä, VTT; Topi Penttinen, VTT; Roland van der Heijden, City of Rotterdam; Elli Kotakorpi, City of Tampere; Eva Pangerl, City of Vienna; Sanja Malnar Neralić, City of Zagreb; Daniel Sarasa Funes, City of Zaragoza									
Description of the related task and the deliverable in the DoW	Based on the KPI definitions, developed in WP1, this task will identify the common data sets that will be used in the calculation of different KPIs. The result will be the description of all data sets that will be needed as input for the Smart City data collection system and their aggregation levels. After the data set definition the task will focus on analysing, in the testing Cities, available data sources, their reliability, data access methods and existing data formats. In addition potential privacy and security issues will be screen. In WP4 collaboration activities and with the partner cities the potential reliability will be analysed.									
Planned resources PM	VTT	AIT	TNO	EUR	TAM	ROT	VIE	ZGZ	ZAG	Total
	2	2	4		1,6	0,5	0,5	0,5	1,5	12,6
Comments										

V	Date	Authors	Description
0	2015-11-26	TNO	1st draft contents
0	2015-12-14	VTT, TNO, AIT	Integrated version with inputs from all the partners
0	2015-12-18	VTT, TNO, AIT	1st draft ready for review
0	2016-01-18	VTT, TNO, AIT	2nd draft ready for review
1	2016-01-31	VTT, TNO, AIT	Final deliverable submitted

Table of contents

Executive summary 1

1. Introduction 3
1.1 Purpose and target group 3
1.2 Contributions of partners 4
1.3 Baseline 4
1.4 Relations to other activities 4

2. CITYkeys 5
2.1 Background 5
2.2 The CITYkeys selection of indicators 5

3. Concepts in city data 9
3.1 Data categorisation 9
3.2 Ownership of information 10
3.3 Open data initiatives and platforms 11

4. Data requirements 13
4.1 People 14
4.2 Planet 15
4.3 Prosperity 17
4.4 Governance 17
4.5 Propagation 18

5. Data set analysis 19
5.1 Analysis methodology 19
5.2 Data sets and their availability 20
5.3 Inventory of cities' open data 21
5.4 CITYkeys KPI sets and open data 27
5.5 Dispersion of data sources within the city 27
5.6 Data for project KPIs: cities initial testing plans 28

6. Data privacy 40
6.1 Open data, public data, internal data 40
6.2 Cities' data privacy policies 40

7. Conclusions 44
7.1 Summary of achievements 44
7.2 Relation to continued developments 46

8. Acronyms and terms 47

9. References 48

Appendix 1: SCIS and CITYkeys 49

Appendix 2: ITU testing plans 51

Appendix 3: List of city data sets and their availability 58

Appendix 4: List of project data sets and preliminary selection of relevance in testing 67

EXECUTIVE SUMMARY

This report describes the analysis of data requirements for CITYkeys selection of indicators, for assessing smart city projects and the corresponding indicators on city level. Next to that, the main results consist of the identification of data sets corresponding to the indicators and the analysis of available data sources, their reliability, formats, level of confidentiality and data access methods. In addition potential privacy issues are discussed.

Starting from the definition of a smart city and smart city projects, taking into account the wishes of cities and citizens as well as already existing frameworks, indicators have been selected that can function as Key Performance Indicators for tracking the progress towards smart city project and city objectives in terms of People, Planet, Prosperity, Governance and Propagation. Although the resulting selection of indicators includes a considerable body of indicators that can be used for evaluating smart city projects in various areas, not all indicators are equally suited for evaluating all types of smart city projects (i.e. some indicators are specifically suited for transport projects, etc.).

The contents in this report have been developed in parallel with the selection and definition of the smart city KPIs (project task 1.3). The results are structured in two public documents: 1) smart city KPIs and related methodology (project deliverable 1.4) which included the full description of all the CITYkeys framework indicators, including their data requirements, and 2) this report (project deliverable 2.1), which focuses on the analysis of the available data sources for the implementation of indicators. The testing of the indicators in 2016 is expected to lead to refinements in issues regarding the applicability and usability of the indicators and the corresponding data availability and quality.

There is a difference in the data collection process between the project and the city indicators. The project indicators are meant for assessing the success and potential for replication of smart city projects. For this a mixture of quantitative and semi-quantitative indicators is proposed. For the semi-quantitative indicators data needs to be collected using interviews or an analysis of policy documents. For the quantitative indicators data needs to be extracted from the project documentation. Cities might want to streamline the data collection for all their smart city projects, but a number of the project indicators will require qualitative information that can only be gathered by involving persons involved in the project (e.g. through interviews, questionnaires).

Data for the majority of the city indicators can be retrieved from statistical sources within the city administration or national or European statistics bureaus. One problem with that data is that it typically presents average annual figures for a whole city. For a city it may however be more interesting to analyse the differences between districts. Spatial data makes it possible to calculate indicators also for geographically restricted areas such as city districts. It is expected that CITYkeys testing phase will allow assessing, to some extent, the impact of a project on city level as the CITYkeys KPI framework contains project-to-city link for many indicators.

The data sets relevant for the CITYkeys KPIs were identified and analysed regarding their availability, sources, reliability, data access methods, existing data formats and level of confidentiality. Special attention was paid in the analysis of open data sets. The data availability rates for the data sets needed by smart city KPIs for the partner cities are in January 2016 as follows: Zaragoza 82%, Vienna 80%, Tampere 77%, Rotterdam 71% and Zagreb 52%. The average availability is 72%. The share of open data providing the needed data sets varies from 1% to 25%, and is 15% on average.

In total the smart city KPIs need 116 raw data sets. Depending on the city and calculation method less than 20 of open data sets were directly applicable for the calculation of the selected CITYkeys KPIs. This means that roughly 85% of needed data sets are either from city's internal systems; are not available and need to be generated; or they are public but do not qualify as open data. For an average of 13% of the data sets, their availability remains unknown at this stage of the project giving a possibility for slight changes in the availability rates. These data sets will need additional investigation during the testing phase but it is expected at this stage that those will not be available easily. External organisations and companies are the most common source for needed available data sets (32%) because some cities have outsourced their functions such as water management, electricity. Most of the available data sets coming from city departments (23%) originate from Mayor's office/Economy (including cities' statistics departments).

On the project level, most of the project KPIs are qualitative or semi-quantitative and need to be evaluated for example based on project documentation or interviews with project manager. Moreover, the boundaries of these data sets need to be further defined in each project separately because the coverage of relevant data is always project specific. Cities' preliminary plans for testing projects, relevant indicators and plans for data collection are presented.

It is important to note that data for all indicators will obviously not be available immediately in all cities. A city that engages in smart city indicators starts a process. The CITYkeys indicator framework is a methodology for such a process. The cities will need continuous development of the indicators to be used by the city and of the data collection mechanisms. Although not mandatory, it could be recommended as a CITYkeys goal that cities automate their data collection and framework feeding. These will not only benefit cities, but also will play a role towards the sustainability of CITYkeys results beyond the project's termination date.

Moreover, the definitions behind certain data sets and data quality obviously differ within countries, between cities and between city departments. The quality of the overall assessment depends on the quality of the indicators, which in turn depend on the underlying data. Managing data quality throughout the process is thus crucial for a good assessment and maximum comparability. When making comparisons a transparent communication of all meta data underlining the data sets is important, since it can explain how reliable the data is – and thereby the results of the corresponding indicator(s).

Some data can't be made open in its raw format due to privacy protection, including citizen privacy, and confidentiality issues. Cities still can have these data in their internal systems following for each data set the access rules and conditions specified in cities' privacy protocols based on national and European privacy regulations. All cities state to follow the law with regard to privacy.

1. INTRODUCTION

1.1 Purpose and target group

CITYkeys aims to speed up the transition to low carbon, resource-efficient cities by facilitating and enabling stakeholders in smart city projects and cities to learn from each other, create trust in solutions, and monitor progress. This is achieved by means of a common performance measurement framework.

The ultimate goal is to support the wide-scale deployment of smart city solutions and services in order to create impact on major societal challenges related to the cities' fast growth and to contribute to the Union's 20/20/20 energy and climate targets.

Cities will benefit from the CITYkeys results as these results support their strategic planning and allow measuring their progress towards smart city goals. In addition, benefits are created from the enhanced collaboration within and between cities, providing the possibility to compare solutions and find best practices. Solution providers will benefit from better insight into business opportunities for their products and services, and into the possibilities for replication in a different city or context. Industrial stakeholders will benefit from the recommendations for new business, e.g. based on open data. Accurate monitoring supporting the implementation of smart city projects should bring environmental benefits as reduction of CO_2 emissions, increased energy efficiency, increased share of renewables, as well as improving the quality of life through a better mobility, better communication between local authorities and their citizens, empowerment of citizens (i.e. smart citizens).

For the development of the performance measurement framework, CITYkeys is building on existing smart city and sustainable city indicator systems as well as filling existing gaps. Synergies with other initiatives such as the currently running H2020 project Smart City Information System (SCIS) are examined. For more information on the relation between SCIS and CITYkeys, see Appendix 1. The bases of the CITYkeys indicator framework are the traditional sustainability impact categories **People**[1]**, Prosperity**[2] **and Planet**[3], but the performance measurement framework includes specific smart city KPIs. These go beyond the traditional categories in showing not only the impact but also indices of the success factors for smart city endeavours (**Governance**) and the suitability for replication in other cities and circumstances (**Propagation**).

The work presented in this document includes:

- Identification of the common data sets that will be used in the calculation of the different KPIs;
- Analysis of the available data sources in collaboration with the partner cities: including reliability, data access methods and existing data formats, as well as potential privacy and security issues;
- Introduction to the city projects that will be used during the framework testing phase.

The transparent and flexible CITYkeys performance measurement framework will be able to handle different sizes of cities in different smart city development stages and thereby support different development strategies of smart cities and initiatives over a wide range of characteristics.

[1] Improving the quality of life of its inhabitants, commuting workers and students, and other visitors
[2] Building an innovation-driven and green economy
[3] Significantly improving its resource efficiency, decreasing its pressure on the environment and increasing resiliency

A list of the CITYkeys indicators' data requirements is presented in this report together with the analysis of the respective data sources, availability and accessibility of the needed data-sets for the implementation of the framework. Adjustments are to be expected after next stages of determining data collection procedures and testing.

The contents in this report are developed in parallel with the selection and definition of the smart city KPIs (project task 1.3). The results are structured in two public documents: 1) smart city KPIs and related methodology (project deliverable 1.4) which included the full description of all the CITYkeys framework indicators, including their data requirements, and 2) this report (project deliverable 2.1), which focuses on the analysis of the available data sources for the implementation of indicators. Future discussions on issues regarding data availability and data quality may bring further changes to the selection and definition of the indicators. Also the testing of the indicators in 2016 is expected to lead to refinements in these issues.

1.2 Contributions of partners

TNO as task leader and VTT as work package leader led most of the analysis presented in this report and edited most of the chapters of this deliverable. The indicator selection exercise, which included data requirements, was done on the basis of an intensive cooperative work between TNO, VTT and AIT in collaboration with all the partner cities. Following on the indicator selection in WP1, all the project partners have evaluated the existing indicators on the data availability for assessing smart city projects with the connected indicators on the city scale. The identification of data sets and analysis of their availability was led by VTT in collaboration with all the partners. The writing of this deliverable was distributed as follows: chapters 1, 2 and 4 were led by TNO, chapter 3 by AIT and chapters 5 and 6 by VTT. The conclusions were drawn together. The partner cities actively contributed with the localization and definition of the needed data sets and data sources in their own organization and contributed to specific sections of this report: Rotterdam, Tampere, Vienna, Zaragoza and Zagreb. The selection and description of initial testing plans, data collection methods and indicators relevant to those, presented in section 5.6, were led by the partner cities.

1.3 Baseline

The aim of CITYkeys is to develop an integrated indicator framework: a cross-sectoral, extended triple bottom line approach. Within this, T2.1 looks for the data requirements for the integrated performance measurement framework, building on existing common data sets whenever possible.

The baseline for this document is the extensive work developed in WP1: the requirements definition, which included the 1st inventory of available open data sets and the full selection and definition of all the indicators in the CITYkeys framework.

1.4 Relations to other activities

T2.1 relates to WP1 on the input side and to the other tasks of WP2 on the output side:

- T2.1 takes into account the results of WP1, building on the outcomes of the indicator selection process that took place in T1.3.
- T2.1 serves as input for further tasks in WP2, in which the indicators from T1.3 and data requirements from T2.1 will be further operationalised (data collection and calculation) and tested.

2. CITYKEYS

2.1 Background

The ultimate goal of CITYkeys is to support the speeding up of wide-scale deployment of smart city solutions and services in order to create impact on major societal challenges around the cities fast growth and the Union's 20/20/20 energy and climate targets. Therefore, CITYkeys aims to facilitate and enable stakeholders in projects or cities to learn from each other, create trust in solutions, and monitor progress, by means of a common integrated performance measurement framework.

2.2 The CITYkeys selection of indicators

A summary of the evaluation framework and of the selection methodology is extracted from WP1 reports and presented in this section in order to contextualize the data sources and data sets analysis. More details on the process leading up to that and the justifications of the choices made are presented in the report on smart city and project KPIs and related methodology (project deliverable D1.4).

2.2.1 The CITYkeys evaluation framework

The selection of indicators for the evaluation framework was based on the outcomes of a questionnaire regarding the needs of cities and citizens and internal discussions on the CITYkeys working definitions (Public available project report D1.1[4]). Figure 1 shows the resulting structure of the evaluation framework.

Figure 1. The CITYkeys indicator framework

The CITYkeys assessment method and the indicators are to be used to evaluate the success of smart city projects and the possibility to replicate the (successful) projects in other contexts. As follows from the smart city definition[5], success is determined by the transition across the

[4] http://www.citykeys-project.eu/citykeys/resources/general/download/CITYkeys-D1-1-Cities-and-citizens-needs-WSWE-9X4HNA

[5] A smart city is a city that efficiently mobilizes and uses available resources (including but not limited to social and cultural capital, financial capital, natural resources, information and technology) for efficiently improving the quality of life of its inhabitants, commuting workers and students, and other visitors [people]; significantly improving its resource efficiency,

entire ecological footprint of urban areas, simultaneously promoting economic prosperity, social aims and resilience to climate change and other external disturbances. Over the past decennia, the concept of sustainability - split up in the triple bottom line of social sustainability (People), environmental sustainability (Planet) and economic sustainability (Prosperity) - has become generally accepted in the development of indicator systems for national and regional urban development (SCOPE, 2007). The 3 Ps (people, planet, prosperity) have also gained considerable ground in company reporting (Kolk, 2004).

The extent to which smart city projects[6] are able to have an effect on social, ecological and economic indicators forms the core of the evaluation. However, this is not enough to determine the success of a smart city Project. Success is also determined by *how* projects have been - or will be - realised in various contexts. The **Governance** of developing and implementing urban smart city projects is a determining factor for high scores in People, Planet & Prosperity indicators (Fortune and White, 2006). Therefore we need to include a number of indicators to evaluate the importance of the city context (external factors) and quality of the development and implementation process (internal factors).

Finally, the ability of individual smart city projects to be adapted in other cities and contexts determines its ultimate effect in achieving European goals with regard to energy and CO_2 emissions. Under the **Propagation** category, smart city projects are evaluated to determine their potential for up-scaling and the possibilities for application in other contexts.

Each of the major themes (people, planet, prosperity, governance and propagation) encompasses several specific policy goals. In many cases these are not all mentioned in a smart city strategy, but may be scattered over various policy documents in a city. For the design of the CITYkeys indicator framework we have arranged these policy goals under the major theme headings. For instance, under the theme People, subthemes conforming to policy ambitions are created (see Fig.1): increasing diversity and improving social cohesion, increasing safety, guaranteeing good education for every citizen, etc.

The reasons for doing so are:

- to underline the relation between policy ambitions and the key indicators that are to be used to measure progress towards these ambitions
- to provide the basis for comparing the indicators with each other, whereby users or user groups may attach weightings to policy goals (and thereby to the indicators belonging to a subtheme).
- to ease communication on the outcome of the indicators in terms which are familiar to the decision makers.

2.2.2 Indicators at project and city level

The CITYkeys evaluation framework will:

1. Evaluate the impact of a smart city project comparing before and after situations or comparing expected impact with a reference situation.

decreasing its pressure on the environment and increasing resiliency [planet]; building an innovation-driven and green economy [prosperity]; and fostering a well-developed local democracy [governance].

[6]A smart city project is a project that efficiently mobilizes and uses available resources (including but not limited to social and cultural capital, financial capital, natural resources, information and technology), and; has a significant impact in supporting a city to become a smart city along the four axis of sustainability mentioned above; actively engages citizens and other stakeholders; uses innovative approaches, and; is integrated, combining multiple sectors. A smart city project can be executed on the scale of: a single building, for instance improving the energy performance of a theatre; or a neighbourhood, for instance improving the waste collection; to the scale of a city or even a region, think of an improvement in the public transport system.

2. Show the progress of the city as a whole towards smart city goals, comparing the year under study with a reference year.
3. Assess how the project has contributed to the objectives at city level.

For the design of the indicator lists, we have started with creating a list of indicators that are useful and feasible to evaluate smart city projects (using the principles described in the next Section). With this list as a starting point we have scanned existing urban indicator sets for corresponding indicators for evaluating city policies. In a few cases it appeared possible to find a corresponding indicator, in which the impact of smart city projects can be immediately expressed (in other words: if one would add the results of all smart city projects in a city, this could immediately be translated in (or related to) the score of the city indicator). For instance, the reduction of CO_2 emissions by a smart city project can be loosely related to the city indicator 'yearly CO_2 emission', although of course the CO_2 emissions of a city are likely to be influenced by many other developments (other projects, economic developments, fuel price, climate etc.). In the majority of the cases it is not possible to add project indicator scores quantitatively, but an indicator on the city level can be found that expresses the same intentions, but using a metrics that cannot be applied on the project level. Appendix 3 of CITYkeys project deliverable 1.4 contains the overview of the link between the CITYkeys project and city indicators.

2.2.3 Target groups for the indicator system

Indicators serve decision making. Indicator outcomes, be it individual indicators or assessments based on multiple indicators should reach the relevant decision makers. The various parts of the CITYkeys indicators are aimed at decision makers on various levels.

The indicators on project level have two primary target groups:

1. those decision makers managing smart city projects (among others, project planners), who can use the indicators to learn about the relative success of smart city projects (how they have been performing, what have been factors determining performance) in order to improve in the next projects (more holistic, smarter), which requires integral in-depth knowledge of results and process of the project, and
2. decision makers in the city council, who need an insight in how the various projects they have decided upon, have been performing (also to be able to take better decision next time), for which a more aggregated overview may be more appropriate.

The project indicators can also be used in the design phase of a project: to give an impression on the expected performance based on design specifications, vis-à-vis already realized projects.

Because the European Commission is financing the so called lighthouse projects, they are (temporarily) in a similar position as a city council, needing insight in the performance of their investments.

The smart city indicators equally have two primary target groups:

- decision makers in the city council who need to follow to impacts of their smart city strategy over time essentially answering the questions has the city become smarter and what has been the final result, and
- national governments and European bodies, to follow if their smart city policies have resulted in more attention for the overall aims (of reducing energy use and greenhouse gas emissions, increasing citizens' participation, etc.). In addition national government and European institutions tend to use indicators to compare cities.

It is clear that for users of the city indicators progress over time is important. Thus, the city indicators should be formulated in such a way that they can easily be included in the city's programme for gathering regular statistics. The outcome of the indicator process, in turn, should get a regular place in the planning processes of the city.

Other groups that are using both project and city indicators include educational and knowledge institutes, and businesses. For citizens the indicators may help to get a better understanding of complex projects and their impacts.

3. CONCEPTS IN CITY DATA

This chapter is a general introduction to the main concepts, approaches and initiatives related to city data. It is introduced here as a supporting background to chapters 5 and 6. Chapter 5 will deepen the analysis on city data within CITYkeys. Chapter 6 assesses on data privacy issues. Finally the currently on-going EC project SCIS is shortly described and potential links to the CITYkeys project are evaluated.

3.1 Data categorisation

To describe data and data sets several categorisations can be used. The description below is adapted from the project Ready4SmartCities (R4SC)[7] which focused on the domains of building and urban data.

Categorization by domain

Usually data are described by domain (such as energy production and distribution, environment, housing, economic production). Data can be broken down into subcategories dependent on the specific use.
For instance data in the domain of energy production and distribution can be broken down into energy sources and forms:

- Primary sources e.g. renewables (wind, solar, geothermal) and conventional sources (coal, natural gas, oil)
- Transformed energy carriers, e.g. electricity, heat/cold, transport fuels.

Another possible breakdown is according to function in the distribution network and storage:

- Networks e.g. power grids, heat/cold supply, gas, district heating/cooling, gas
- Storage e.g. pumping plants and reservoirs, night-storage, batteries.

In the CITYkeys indicators, for example, there is one indicator on final energy consumption. The data set needed to calculate the indicator is broken down in energy consumption by energy carrier, as this is the usual way the data are collected. Data by energy carrier in different units such as tonnes of coal, litres of liquid fuel and kWh of electricity can be aggregated in one single indicator using conversion factors. To enhance the analysis, the indicator may be broken down by sector (transport, housing, street lighting, etc.)[8].

Resolution

Another attribute of data is their resolution of data and time reference, and aspects of ownership and standardisation of the data:

- Temporal framework (when does e.g. consumption occur?)
- Spatial framework
    - Geographical reference
    - Urban reference
        - Planning – residential, commercial, industry areas
        - Urban road systems – private and public transport
- Organizational framework
    - Functional units (ownership, who owns information and data?)

[7] See http://www.ready4smartcities.eu/project-details

[8] See CITYkeys public project deliverable D1.4 for more information. Available at www.citykeys-project.eu/

- Normative organizations
- Standards

The data descriptions presented in the annexes to the CITYkeys project deliverable 1.4 follow this format.

Influencing factors

Finally data can be accompanied by other data sets offering explanatory information, such as internal activity data about company and individual behaviour and expected/actual consumption, such as:

- Occupancy and user behaviour (user data to energy profiles for office work, commerce, residential use)
- Consumer device (use, consumption), or external influencing factors, such as
- weather data which may affect both the consumption and the production of energy (in case of solar or wind energy) or
- energy market data which may affect the prices of energy as well as the demand for it.

In CITYkeys, for instance, information on climate zones will be used to account for differences in building energy consumption in the final scoring of the indicators.

3.2 Ownership of information

Data are not only held by the city authorities. A multitude of stakeholders owns information on a city. According to the variety of domain data, the stakeholders possibly owning the information and required data sets can be described by a similar categorization.

The related institutions and bodies can be described by their activities:

- Municipality and governance bodies
  - local and city councils
  - municipal services
- Building and other asset managers
- Network managers, such as energy providers, traffic managers
- Other companies
- Scientific and other parties such as meteorological institutes
- Citizens

In general the problems with information sources and related data sets spread among city authorities and contributors can be described as follows:

- The existing data sets are often heavily fragmented, with differing granularity in terms of timely and spatial resolution
- Technically the formats are often proprietary and created with different syntax and semantics for similar and different contents
- Different countries may measure or collect what appears to be the same indicator or data using slightly different definitions.
- The owners of information do only have low motivation in publishing data without generating benefit out of it
- The needed data sets are therefore often dedicated to silos of protected information and could only be exploited with high effort or not all due to legal and privacy framework conditions

In CITYkeys we have tried to overcome the data definition issue by formulating the indicators as much as possible in line with common existing indicator schemes such as ISO 37120, to harmonize and facilitate data collection. The other issues arising from distributed ownership will be tackled during the testing phase in 2016 and addressed accordingly.

3.3 Open data initiatives and platforms

In this section an overview about current initiatives and platforms regarding publishing and revealing appropriate information and related data sets is given. As a first step some usual definitions regarding data and availability will be listed and shortly described.

Data and access

In general, information and related data sets can be described by different perspectives of access (see chapter 6.1.1 for details):

- Online access through networks is possible without authentication
- Data is available online but the access is restricted
- Information is not accessible online or data gathering requires manual work

Open Data

A Data set is defined as open, when it fulfils following three fundamentals:

1) data is available and is in readable form
2) data is published with a license which allows re-use and redistribution
3) data is published with equal terms for every user (The Open Knowledge Foundation, 2015)

For further details see http://5stardata.info/en/ and chapter 5.3.

Linked Data

The notion of Linked Data[9] can be described by the following properties:

- Related data published on the Web lowering the barriers for linking
- Recommended best practice for exposing, sharing and connecting pieces of data, information and technology using Semantic Web technologies e.g. URIs and RDF[10]

Linked Open Data

The notion of Linked Open Data can be described as extension of Linked Data by means of linking open content, i.e. data which is published on an open license basis.[11]

Ontology

In fact there exist a lot of similar definitions for this notion. There is unfortunately not one universally accepted definition available. The following definition should although be given:

- An ontology is a formal specification of a shared conceptualization
- An ontology describes
  - What kind of things exist or can exist
  - What manner of relations can those things have to each other[12]

In the following section a list of publicly accessible information resources on the Web is presented. The focus lies therefore on platforms using the principles of (Linked) Open Data

[9] http://linkeddata.org
[10] https://en.wikipedia.org/wiki/Linked_data
[11] http://linkeddata.org/faq
[12] http://semanticweb.org/wiki/Ontology

and Ontology look up services[13]. For further information about available formats see chapter 5.3.

Open data catalogues

- Open Government Data (http://opengovernmentdata.org/data/)
- Open Data Index (https://index.okfn.org/)
- Global Open Data Index (http://global.census.okfn.org/year/2015)
- European Union Open Data Portal (http://open-data.europa.eu/en/data/)
- Linking Open Government Data (http://logd.tw.rpi.edu/)
- The European Open Government Data Initiative (http://www.govdata.eu/en/europeanopen.aspx)
- Open Data initiatives from EU countries (e.g. https://www.data.gv.at/, https://open.wien.gv.at/site/open-data/)

Web portals, which contain data sets from a concrete organization or a domain

- UNdata http://data.un.org/
- The World Bank http://datacatalog.worldbank.org
- Engage project http://www.engagedata.eu
- DBPedia http://wiki.dbpedia.org/
- Eurostat Urban Audit http://ec.europa.eu/eurostat/web/cities/data/database
- The Episcope and Tabula Website http://episcope.eu/welcome/
- The Buildings Performance Institute Europe (BPIE) http://www.buildingsdata.eu/data-search
- Clean Web Initiative http://cleanweb.co
- Local statistics authorities e.g. http://www.statistik.at/web_en/statistics/index.html

Ontology search engines

- Watson (http://watson.kmi.open.ac.uk)
- Swoogle (http://swoogle.umbc.edu)
- Linked Open Vocabularies (LOV) (http://lov.okfn.org/data set/lov)

Linked Data data set catalogue

- Datahub (Data Management Systems) http://datahub.io
- Reegle (http://data.reegle.info/)
- Open Energy Information (OpenEI) (http://en.openei.org/data sets)

This enumeration of platforms and resources does not claim to be exhaustive as well. But it should give a rough overview about possible resource types with the characteristics of public accessibility which could be exploited for feeding the data sets needed for a comprehensive indicator system like CITYkeys.

[13] http://www.ready4smartcities.eu/project-details

4. DATA REQUIREMENTS

The CITYkeys set of indicators consists of a mixture of quantitative and semi-quantitative indicators[14]. The semi-quantitative indicators for example provide an assessment of the way smart city projects are executed and the way the city government stimulates smart city development and of their potential of smart city projects to be taken up by other cities. For the semi-quantitative indicators data need to be collected using interviews or an analysis of policy documents.

There is a difference in the data collection process between the project and the city indicators. The project indicators are meant for assessing the success of smart city projects. In individual project assessments data need to be collected from the project office, the project leader and/or others closely involved in the project. Cities might want to streamline the data collection for all their smart city projects, creating a reporting system and specific databases, but even then a number of the project indicators will require (qualitative) information that can only be gathered by involving persons involved in the project (e.g. through interviews, questionnaires).

Data for the majority of the city indicators can be retrieved from statistical sources within the city administration. Some have been made available in open source formats. However some of the governance indicators also require a person to gather the information. The share between these different data sources varies for different cities administration realities. Eventually it is recommended to automate the data collection and framework feeding for as far as possible.

The complete description of the data needed to compile the indicators is provided in Appendices 1 and 2 of the CITYkeys public report describing the complete indicators framework (D1.4). The template used for the indicator descriptions is presented in Table 1.

It is important to note that data for all indicators will obviously not be available immediately in all cities. A city that engages in smart city indicators starts a process. The CITYkeys indicator framework is a methodology for such a process. The cities will need continuous development of the indicators to be used by the city and of the data collection mechanisms.

Moreover, data quality obviously differs within countries, between cities and between city departments. The quality of the overall assessment depends on the quality of the indicators, which in turn depend on the underlying data. Managing data quality throughout the process is thus crucial.

[14] Quantitative indicators are based on quantities obtained using a quantifiable measurement process. Semi-quantitative indicators are based on qualitative information that is then assessed according to a Likert Scale (e.g. 1 = not at all; 2 = poor; 3 = fair; 4 = good; 5 = excellent). An example is found in section 4.1.

Table 1. Template used for indicator descriptions including data requirements

Name of the indicator		**Type of project for which the indicator is relevant (ICT, built environment, transport)**
Description incl. justification		
Definition		
Calculation		
Strengths and weaknesses		
Data requirements		
Expected data source		
Expected availability		
Collection interval		
Expected reliability		
Expected accessibility		
References		

4.1 People

4.1.1 Project

On the project level, the data for most indicators will have to come from the project documentation and/or interviews with the project leader and others involved in the project. By doing so, it is expected that it will be possible to assess most indicators relevant for the project. Due to the nature of the theme, many people-indicators are assessed on a semi-quantitative Likert scale, for example in box 1:

Box 1. Example of a semi-quantitative Likert scale.

The indicator "connection to the existing cultural heritage", defined as the extent to which making a connection to the existing cultural heritage was considered in the design of the project.

Not considered in the design – 1 — 2 — 3 — 4 — 5 — Very much considered in the design

Guideline for 'Connection to the existing cultural heritage on the premises by design'':

1. *Poor: no attention has been paid to existing cultural heritage.*
2. *Fair: heritage places have received some attention in the project, but not as an important element.*
3. *Average: some attention has been given to the conservation of heritage places.*
4. *Good: heritage places are reflected in the project design*
5. *Very good: heritage places are included in the project as clear and recognizable landmarks.*

This is typically an indicator that can be assessed on the basis of an interview or project documentation, if possible combined with a site visit. Coincidentally, this indicator is also an example of an input indicator. Most indicators in the people theme are outcome indicators[15]. Some indicators only need scoring after the project, others need a reference situation. The reference situation may be before implementation of the project, or a hypothetical reference (for example in the case of a newly built development in a greenfield setting). Most indicators can also be used ex-ante to evaluate the quality of plans.

4.1.2 City

On the city level, the city statistics bureau (e.g. population statistics), records of the executive departments/services (e.g. police records for crime data) and city open data (e.g. availability of amenities) are important sources of data for the People category. Most indicators rely on quantitative information. We will consider the availability and quality in detail during the testing, tuning and adjusting where necessary. For many of the more common indicators such as traffic accidents, it is expected that the data will be available, although the exact definitions and metrics used may differ throughout Europe. Some of the newer indicators, such as digital literacy may prove more difficult and it will have to be seen in the testing phase whether they can be operationalized using data. Moreover there are three qualitative indicators using Likert scales: although in our experience they are workable on the project level, it remains to be seen how they will work on the city level. In order to follow progress towards policy goals over time, data collection should take place annually.

4.2 Planet

4.2.1 Project

On the project level, the data for most indicators will have to come from the project documentation and project monitoring (e.g. data extracted from sensors into a database). It is expected that, since reduction of energy use, CO_2-emissions, resource efficiency and improving environmental quality are an important drivers for smart city projects, data will be available to assess most of these indicators if the project addresses these topics. Due to the nature of the theme, many Planet indicators can be assessed quantitatively. In many cases, the rough data from the project documentation, or monitoring platforms, will need some conversion: for example, the rough data may be gas use in m^3/year, while the indicator requires % reduction in CO_2. The indicator descriptions provide guidelines for such conversion calculations.

Most indicators in the planet theme are impact indicators. Most indicators need to be scored in comparison to a reference situation. The reference situation may be before implementation of the project, or a hypothetical reference (for example in the case of a newly built development in a greenfield setting). Most indicators can also be used ex-ante to evaluate the quality of plans.

Note that within this theme, the indicators for (direct and indirect/embodied) energy use and CO_2 will be tested using the ITU L.1440 guidelines for assessing the impacts of smart city

[15] Input indicators refer to the resources needed for the implementation of an activity or intervention (e.g. number of staff involved in training). Process indicators refer to indicators that measure whether planned activities took place (e.g. number of trainings held). Output indicators add more details in relation to the product of the activity (e.g. people trained). Outcome indicators refer to the results in relations to objectives of an intervention (e.g. people trained as percentage of the target group). Impact indicators refer to the quality and quantity of long-term results generated by the project (e.g. the reduction in child mortality as a result of the training).

projects. A brief background is given in box 2. A simplified guideline for testing is provided in Appendix 2.

Box 2. Brief background of the ITU L.1440 guideline testing.

The International Telecom Union has developed guidelines for determining the environmental impacts, mainly first and second order energy use and GHG emissions, of ICT. This is known as the ITU L.1440 methodology for environmental impact assessment of information and communication technologies at city level.

The CITYkeys project was requested by the European Commission to test the ITU L.1440 methodology for environmental impact assessment of information and communication technologies at city level. This will be done in select cases as part of the testing.

Next to this, the ITU has developed a set of indicators for smart sustainable cities projects. The framework and indicators of the ITU smart sustainable cities initiative have been assessed.

The goal of the test is to:

1. Investigate whether it is possible to apply the methodology of ITU L.1440 to evaluate the environmental impacts of ICT in CITYkeys case studies;
2. Outline what data is needed, the quality, level of detail and the availability of such data in the cities;
3. Evaluate which capacity is roughly required to gather data and calculate results in terms of expertise, competences and working hour's costs.

It is expected that it is possible to calculate these indicators, but it may be costly in terms of data and capacity needed. It is also unclear to which extent the data on embodied energy/CO2 of ICT will be available. If the methodology is workable, it may become an integral part of the CITYkeys methodology.

It is suggested to test the methodology in the cities of Tampere and Rotterdam, since these cities are likely to have ongoing smart city projects, have a good data management system and a research institute with additional expertise (for e.g. LCA) is available to assist. The data are to be a consistent set from the most recent year available and/or for the duration of the project.

Deliverables include:

- A summarized manual on how to operationalize the indicators of ITU L.1440, to be integrally included in the CITYkeys selection of indicators, for which Appendix 2 provides a baseline.
- A full report on the case studies, going into the process of executing the methodology, data availability, the calculated results and following conclusions, and an evaluation on the workability of the methodology.

4.2.2 City

On the city level, the city statistics bureau (e.g. overall CO_2-emissions), records of the executive departments/services (e.g. air quality indicators) and city open data (e.g. information on green areas) are important sources of data for the Planet category. Almost all indicators rely on quantitative information. We will consider the availability and quality in detail during the testing, tuning and adjusting where necessary. For many of the more common indicators such as CO_2 and air quality, it is expected that the data will be available, although the exact definitions and metrics used may differ throughout Europe. Some of the newer indicators, such as Self-sufficiency for food may prove more difficult and it will have to be seen in the implementation and testing phase whether they can be operationalized using data. Moreover there is one qualitative indicator using a Likert scale: although in our experience Likert scales are workable on the project level, it remains to be seen how they will work on the city level. In order to follow progress towards policy goals over time, data collection should take place annually.

Note that within this theme, the indicators for (direct and indirect/embodied) energy use and CO_2 will be tested using the ITU L.1440 guidelines. A simplified guideline for testing is provided in Appendix 2.

4.3 Prosperity

4.3.1 Project

On the project level, the data for most indicators will have to come from the project documentation and/or interviews with the project leader and others involved in the project. Our experience is that collecting data for the Prosperity category can be difficult since this often includes sensitive information, for example about the business case of the project. Nevertheless, indicators such as Fuel Poverty can often be calculated from data that is also needed for calculating the Planet-indicators on energy, using some conversions. Many of the indicators in the subthemes green economy and innovation are very new. It remains to be seen whether these can be operationalized during the testing phase. Most indicators in the Prosperity theme are output indicators. Some indicators only need scoring after the project, others need a reference situation. The reference situation may be before implementation of the project, or a hypothetical reference (for example in the case of a newly built development in a greenfield setting). Most indicators can also be used ex-ante to evaluate the quality of plans. Indicators may not be applicable in some cases, for example if the project does not concern itself with open data many of the innovation-indicators would not apply.

4.3.2 City

On the city level, the city statistics bureau (e.g. GDP), records of the executive departments/services (e.g. public transport use) and city open data (e.g. open data indicators) are important sources of data for the Prosperity category. All indicators rely on quantitative information. We will consider the availability and quality in detail during the testing, tuning and adjusting where necessary. For many of the more common indicators such as unemployment and tourism intensity, it is expected that the data will be available, although the exact definitions and metrics used may differ throughout Europe. Some of the newer indicators, such as Green jobs may prove more difficult and it will have to be seen in the testing phase whether they can be operationalized using data. In order to follow progress towards policy goals over time, data collection should take place annually.

4.4 Governance

4.4.1 Project

On the project level, the data for most indicators will have to come from the project documentation and/or interviews with the project leader and others involved in the project. Our experience is that collecting data for the Process theme is most efficiently by interviewing three stakeholders (i.e. the project leader and two others, e.g. from the municipality or from the neighbourhood). Due to the nature of the theme, most indicators in the Governance theme are semi-quantitative process indicators; all are assessed by Likert scale. Most indicators only need scoring after the project, and can also be used ex-ante to evaluate the quality of plans. The indicators are relevant for any type of project.

4.4.2 City

On the city level, the city statistics bureau (e.g. voter participation), records of the executive departments/services (e.g. open public participation) and perhaps city open data (e.g. policy evaluations) are important sources of data for the Governance category. Approximately half of the indicators rely on qualitative information and half on quantitative information. We will consider the availability and quality in detail during the testing, tuning and adjusting where necessary. For many of the more common indicators such as voter participation, it is expected that the data will be available, although the exact definitions and metrics used may differ throughout Europe. Some of the newer indicators, such as Ease of access to information may prove more difficult and it will have to be seen in the testing phase whether they can be operationalized using data. Moreover the qualitative indicators need to be assessed using Likert scales: although in our experience they are workable on the project level, it remains to be seen how they will work on the city level. In order to follow progress towards policy goals over time, data collection should take place annually.

4.5 Propagation

4.5.1 Project

On the project level, the data for most indicators will have to come from the project documentation and/or interviews with the project leader and others involved in the project. In the Propagation theme, also the overall impression of the project on the assessor will play a role, for example with regard to the likelihood that the project will be replicated or scaled up in other areas or with other stakeholders. Due to the nature of the theme, most indicators in this theme are semi-quantitative outcome or impact indicators; all are assessed by Likert scale. Most indicators only need scoring after the project, and can also be used ex-ante to evaluate the quality of plans. The indicators are relevant for any type of project.

4.5.2 City

There are no indicators on the City level in this category.

5. Data set analysis

The full CITYkeys framework description (public report, project deliverable D1.4) contains for each indicator the definition, calculation method and data requirements including expected data sources, data availability and reliability. This section identifies the data sets relevant for the CITYkeys KPIs and analyses in the partner cities available data sources, their reliability, data access methods, existing data formats and level of confidentiality. Special attention is paid in the analysis of open data sets. A co-benefit of CITYkeys could be that data becomes more open and more frequently collected as cities start the process to evaluate their smart city (project) strategy. Within the project the status of these open data sets will be re-evaluated at the end of the testing phase (project task T2.4).

As reported in section 4, not all KPIs can be unpacked into data sets, as part of the KPIs are qualitative or semi-quantitative. Those are evaluated with a qualitative Likert-scale which is typically evaluated as expert assessment or through interviews or surveys. In addition, it is to be noted that the system boundaries for project indicators are project specific. Therefore the data set analysis of this chapter is mostly relevant for the quantitative city KPIs and is mostly restricted to those. Data sets for quantitative project KPIs are listed in Appendix 4 of this report and should be further defined in each project case. Those will be addressed more in detail in the testing phase in CITYkeys (project task T2.4).

Initial plans for the testing phase including case project descriptions, relevant KPIs and planned data collection methods are presented in section 5.6 and will be further refined later in the following phases of the framework implementation. The selection of relevant data sets for the testing projects is presented in Appendix 4. For a glossary of acronyms and terms, see chapter 8.

5.1 Analysis methodology

The first step was to unpack the KPI calculation formulas into data sets. Most of the KPIs are not available ready-calculated, and an indicator often consists of two or more single data sets. The same data set can be shared by multiple indicators, being referred here as “common data sets”. Therefore when common data sets were identified, and with a list of distinct data sets, it was possible to start collecting and analysing the respective data sources.

The following presents two examples of how an indicator is translated into needed data sets. First example is the indicator “Ratio of green and water spaces”, which is defined in the framework as follows “Share of green and water surface area of total land area”. In mathematical notation it would be:

$$\left(\frac{green \wedge water\ surface\ area}{Total\ land\ area}\right) * 100$$

From there we can see, that in order to calculate the indicator we need the data sets *Green surface area*, *Water surface area* and *Total land area*. When only a division of two numbers is done, there are not any more requirements for those data sets. They can either be plain values (e.g. *Total land area = 690 000 m*^{2}), or a spatial feature representing city borders in GIS-format, where the *Total land area* could then be calculated on the fly. Both data sets have their pros and cons, but in this case will lead into same result.

Another example of indicator is “Access to public transport”. Its definition is “% of inhabitants with a public transport stop/transportation connection (train, tram, subway) within reasonable (500m) distance”. The associated calculation formula is:

$$\left(\frac{population\ with\ a\ public\ transport\ stop/connection(train,tram,subway)<500m}{total\ population}\right)*100.$$

The needed data sets are *Total population*, *population data with coordinates* and *locations of public transport stops/connections*. *Total population* can be a plain value, but *locations of public transport stops/connections* and *population data with coordinates* both have to contain exact feature coordinates. Population data needed for this indicator could be for example centre coordinates of buildings, with an attribute telling how many persons live in that building. Then the indicator would be calculated with the following steps:

1. *Locate buildings, where the distance to the nearest bus stop is less than 500m*
2. *Calculate how many persons live in those buildings*
3. *Divide it by total amount of persons living in all buildings*
4. *Multiply it with 100% to get percentages.*

Total population can be retrieved from multiple data sources: from city statistics, from number of persons living in each building, etc. Depending on data quality, and initial data sources, different data sources can provide a different value for *total population*. Therefore it would be advised to use values extracted from the same data set both in numerator and denominator.

5.2 Data sets and their availability

The list of all data sets needed for the CITYkeys city KPIs and their availability in the partner cities is presented in Appendix 3 of this report. In this section we analyse, in the partner cities, the availability of the data sets for the calculation of the CITYkeys KPIs at city level.

The data availability rates for the data sets needed for the KPIs are in the partner cities as follows: Zaragoza 82%, Vienna 80%, Tampere 77%, Rotterdam 71% and Zagreb 52%. Average availability in cities is 74%. Share of open data ranges from 1% (Zagreb) to 25% (Tampere), and is 15% on average. For an average of 13% of the data sets, their availability remains unknown at this stage of the project giving a possibility that unavailability or availability rates might slightly change. These data sets will need additional investigation during the testing phase but it is expected at this stage that those will not be available easily. The most difficult data sets for cities in terms of availability are:

- *Expenditures by the municipality for a transition towards a smart city in €*
- *Total food demand (tonnes)*
- *# of green jobs (related to environmental service activities that contribute substantially to preserving or restoring environmental quality)*
- *# of houses or households with grey water reuse capability*

These data sets are unavailable in all cities (or their availability still remains unknown) and they have in common that most of them are not exactly countable, but require sophisticated calculations.

The availability of needed data sets in the five CITYkeys partner cities is presented in Table 2.

Table 2. Availability of needed data sets in the five cities

Data sets in total: 116	*Rotterdam*	*Tampere*	*Vienna*	*Zagreb*	*Zaragoza*
Available	*82*	*89*	*93*	*60*	*95*
	71 %	*77 %*	*80 %*	*52 %*	*82 %*
Public	*49*	*38*	*64*	*51*	*52*
	42 %	*33 %*	*55 %*	*44 %*	*45%*
Confidential	*6*	*9*	*9*	*7*	*0*
	5 %	*8 %*	*8 %*	*6 %*	*0 %*
Open data	*23*	*29*	*18*	*1*	*15*
	20 %	*25 %*	*16 %*	*1 %*	*13%*
Not available	*18*	*26*	*18*	*18*	*5*
	16 %	*22 %*	*16%*	*16 %*	*4%*
Unknown	*16*	*1*	*6*	*38*	*16*
	14 %	*1 %*	*5 %*	*33 %*	*14 %*

5.3 Inventory of cities' open data

This section focuses on the open data inventory provided by CITYkeys partner cities. The link to CITYkeys KPIs is made in the later sections.

A Data set is defined as open, when it fulfils following three fundamentals:

1) data is available and is in readable form;
2) data is published with a license which allows re-use and redistribution;
3) data is published with equal terms for every user (The Open Knowledge Foundation, 2015).

Therefore only the amount of data sets is not an indicator about how well a city has adapted open data policy, but the quality and technical properties are more important.

All five CITYkeys partner cities have an open data portal, where they publish their open data. Three portals, Rotterdam[16], Vienna[17] and Zagreb[18], are based on CKAN-platform, which is an open-source platform which was developed by The Open Knowledge Foundation (OKNF) and is managed by the CKAN Association (The Open Knowledge Foundation, 2015b). The cities of Tampere[19] and Zaragoza[20] have developed their own solutions.

Zaragoza's portal offers access to several services:

- a catalogue of 112 open data sets. The catalogue is defined using an internationally recognized vocabulary called DCAT ("data set catalogue").

[16] http://rotterdamopendata.nl/data set
[17] https://www.data.gv.at/suche/?publisherFilter_Stadt+Wien=on&connection=and&hideFilters
[18] http://data.zagreb.hr/
[19] http://www.tampere.fi/tampereen-kaupunki/tietoa-tampereesta/avoin-data.html
[20] https://www.zaragoza.es/ciudad/risp/

- SPARQL access point
- an API
- an app repository

The formats in which the open data from Zaragoza can be displayed are: (Geo-)JSON, RDF, CSV, etc. Zaragoza's open data is ranked with 5 stars (linked data) according to W3C consortium classification. Particularly innovative is the “Complaints and suggestions” data set, which is updated by both the city and the citizens. The latter become, through their contributors, “prosumers” of open data. Through this data set, everyone can track the city hall's service level. Additionally, the city offers access to other statistical information (mainly socio-economic information) through a human-oriented interface, that allows filtering and downloading in excel formal.

In total the 5 cities have 666 open data sets in their portals, making an average of 133 data sets per city. However, as said before, the quantity of open data provided tells only half of the story since the quality and usability of data are often more important. Currently the quality and update intervals on cities’ open data portals vary a lot for different data sets. Cities are following their own strategy to update certain information on their portal. Also data can be added during one project but not updated regularly. The use of some of the data sets seems minimal. Some potential reasons include the lack of security on data quality and reliability. To put it differently, data sets might be added to the portal from the “supplier’s point of view” instead of the “customers point of view”. On the other hand, another city estimates that around 90% of the city data (currently behind internal interfaces) could be made open. The problem is that it would need significant additional work since the level of publicity should be defined for all data and then should be processed accordingly. Due to the increasing interest and possible uses of open data, there is, however, a clear need to have reliable open data that can be easily used. An expected potential co-benefit of CITYkeys project could be that data becomes more open and more frequently collected as cities start the process to evaluate their smart city (project) strategy. Within the project the status of these open data sets will be re-evaluated at the end of the testing phase (project task T2.4). In CITYkeys WP3 new business opportunities will be identified and open data is a potential source of data for those.

The quality of all available open data sets was rated according to the so called “5-Star Data” scale (Berners-Lee, 2015). In this 5-star scheme, points are given from 0 to 5 according to following scale:

1. Data is published in any format under an open license
2. Data is in structured format
3. Data is in open format
4. Data is in linked data format containing URIs
5. Data is linked to other data respectively

In practice getting points 4 and 5 requires that data is transformed into Linked Open Data (LOD). LOD is part of the W3C’s Semantic Web standard (W3C, 2015).

Semantic Web and Linked Data aim to provide machine understandable linking and binding between different data sets and their entities. The difference between structured and unstructured format can be described using the following example:

> *Let’s have a table with n rows and m columns. For human the same table is shown in three different formats, as image-file, PDF-file and an Excel-file. We only can see what’s rendered on the screen, so for us all three tables look equal. Computer itself, which is not capable to sense visual connections of things, is not easily able to read content of two first (image and PDF) files, but can read content of the Excel file. Therefore publishing data in unstructured format doesn’t yet make the data usable.*

Using as baseline the above concepts the quality of the open data sets screened was analysed. Average points for cities ranged from 2.17 to 2.96 with a standard deviation of 0.29. Being in the range between 2 and 3 points means, in general that cities have published most of their data in structured format, but not everything in non-proprietary format. For example if city publishes everything as Excel-spreadsheets it would result into rating 2.0 whereas publishing everything in CSV it would be then 3.0. CSV stands for comma-separated values, where each column is separated with a delimiter (for example comma character), and line breaks are done with a newline character.

Table 3 shows the cities' open data platforms, the number of available open data sets provided there as well as the average quality rating of the data sets according to the "5-Star Data" scale. The analysis was done by listing all open data sets from cities' open data portals into same table, and then giving stars based on the format for every single data set. It is common to publish one data set in multiple formats, which makes re-use easier by decreasing format-specific barriers, and also allows faster exploration of provided data.

Table 3. Number of open data sets and their average quality rating in CITYkeys partner cities

City	***Platform***	***Data sets***	***Rating***
Rotterdam	CKAN	103	2.17
Tampere	own	144	2.82
Vienna	CKAN	292	2.92
Zagreb	CKAN	16	2.63
Zaragoza	own	111	2.96
Total	**-**	**666**	**2.78**

5.3.1 Analysis of open data formats

As stated before in the open data definition, data is not seen re-usable open data until it's published in open structured format. Open format means that the creator of file format has published technical specification for the format, or the format is standardized by a standardization organization. For open data available we identified in which format the data is available, both to see if the data is usable, and to get a first insight, how easily it can be integrated into testing interface. In format analysis we preferred some formats over other, as explained below.

5.3.2 Spatial data

For data sets containing spatial data (coordinates, geometries) we first checked if the data is available as WFS (OGC, 2015). This is due to the fact that WFS (Web Feature Service) is a widely used open OGC-standard, and it gives developer more possibilities than a single file-based-format. But for spatial data not available as WFS, next formats to check were Shapefile and CSV. Shapefile itself is a binary format, but having an open specification it has become de facto standard file format in GIS. Spatial data makes it possible to calculate indicators also for geographically restricted areas such as city districts.

5.3.3 Application Programming Interfaces (APIs)

Some data sets were also usable as REST (Representational State Transfer) APIs, which itself is not a protocol, but an architectural guideline, how to implement it with standard HTTP methods. If a data set was available as REST API, we grouped all of them under the same API label, even though using them would require specific implementations.

5.3.4 Formats

For every data set we chose from provided formats the one, which was best in terms of usability, openness and data quality. Finally we ended up with a list of 31 different formats, where 16 formats were only available in four or less data sets. A big share of open data is spatial, almost half of the data sets, which makes that WFS was the most common format with the share of 32%. The wide adoption rate of WFS is possible partially thanks to Inspire directive, which defines requirements on what and how spatial data should be published (The European Parliament and the Council, 2007). After WFS follows CSV, with a 26% share. In practice CSV is also an output format of WFS-standard meaning that over 60% of all data sets are accessible in CSV format, making it at the moment the most commonly available open data format. But as JSON, WFS and other formats provide enhancements to CSV; therefore they are valued over it.

A big share of 86.5% of the open data sets is in easily machine readable formats. Five most common formats, which are not easily machine-readable, are: XLS, ZIP, ODS, PDF and HTML. Although all of them can be read programmatically, practical implementation of reader is complicated and in any changes of structure the read process will probably fail. One easily doable movement towards better rated open data is to publish all Excel and ODS based data sets also as CSV files. The division of different data formats in cities' open data portals is illustrated in Table 4.

Table 4. Data formats in cities' open data platforms

Format	***Data sets***	***%***	***Open structured format***
WFS	198	32,2 %	yes
CSV	173	28,1 %	yes
API	35	5,7 %	yes
WMS	27	4,4 %	yes
Excel	26	4,2 %	no
JSON	26	4,2 %	yes
SPARQL	23	3,7 %	yes
ZIP	16	2,6 %	no
ODS	11	1,8 %	no
PDF	11	1,8 %	no
SOLR	10	1,6 %	yes
HTML	9	1,5 %	no
ZML	8	1,3 %	yes
RSS	5	0,8 %	yes
WMTS	5	0,8 %	yes
Others	32	5,2 %	

5.3.5 Update interval

Open data which is kept up to date is more valuable for the user, than open data which is updated seldom or not at all. CKAN data set metadata holds information about when data set is created and when it's updated. However, it is to be noted that for real-time data sets, where only a reference URL is updated into CKAN, last update time in CKAN doesn't reflect the reality. Simply comparing timestamp from last update is ineffective because recently created and never updated data sets would qualify in that comparison unfairly well. This is why we compared total age of data set with time from last update. Formula is as follows:

$$\frac{Currentdate - Lastupdated}{Currentdate - Datasetcreated}$$

This equation would yield 1.0 when data set has been never updated and 0.0 if data set is updated at current date.

From the analysis of the partner cities we can see, that, for example, in Vienna the data sets are either updated frequently with short intervals, or not at all. These results might tell, that in Vienna data sets are automatically updated either to reflect on changes on underlying data stores, or periodically a new version is uploaded into CKAN, regardless of if the data set is already its newest version.

Rotterdam strives to keep the data relevant for city policies updated the 1st of January of each year, and also some data sets are updated real-time or continuously and reflect the actual situation.

The current version of the Tampere open data portal shows all updated data sets as new data sets, which makes the estimation of the update interval difficult. For statistical data sets, the update interval is most often once a year, but there are also real-time interfaces.

5.3.6 Open data from other sources

Open data relevant to CITYkeys and its partner cities is not limited into the open data portals cities have. In Zagreb, for example about 44% of data sets are available from public sources. In this case there is a web page "Zagreb annual" holding a lot of statistical and numerical information updated annually about the city[21]. Even though the data is available from public sources, it doesn't mean that it is more readable for computers or that the data collection would be easier than for example for data in printed format. In almost every case, when data is published on reports or on HTML-pages, it must be manually inserted into a KPI calculation system.

In the case of Tampere open data comes also from National Land Survey (NLS) of Finland, Finnish Transport Safety Agency (Trafi), Finnish Meteorological Institute (FMI) and Statistics Finland. For example FMI provides APIs for weather data, containing also air quality measurements, when the city itself doesn't have to own or open duplicate data set. All vehicles are registered at Trafi, when it also becomes natural source of the data. In Finland there exists at least two levels of topological databases, municipalities have their own GIS-systems, where the level of details is higher, and for example in some cases single trees are stored in their GIS-databases. NLS has their own topological database where the level of details is lower, but is consistent in the whole country.

In the case of Rotterdam, next to the open data and data from the city statistics office, CBS Statistics Netherlands has many data on the state of the country. For some data sets, more spatial detail (e.g. city level) is provided. For many data sets, more detail is available than displayed on the website, but can only be accessed for certain types of research, by certain people and often in exchange for a fee. The Royal Netherlands Meteorological Institute (KNMI) is the Dutch national weather service. Primary tasks of KNMI are weather forecasting, and monitoring of weather, climate, air quality and seismic activity as well as research on these topics. Rijkswaterstaat is responsible for the design, construction,

[21] http://www1.zagreb.hr/zgstat/ See "Grad Zagreb - osnovni statistički podaci (hr i en)" for data available in English.

management and maintenance of the main infrastructure facilities in the Netherlands. This includes the main road network, the main waterway network and water systems. Vehicle registration is with the RDW: tasks in the area of the licensing of vehicles and vehicle parts, supervision and enforcement, registration, information provision and issuing documents. Other agencies which may have relevant data include among others The Netherlands Enterprise Agency (RVO), which has made available many energy-related data. Also, the DSOs publish regular data on energy uses on the neighbourhood level, which has been made available to the municipalities. All these agencies are expected to work in accordance with Dutch privacy laws.

Concerning Zaragoza, on a national scale, the Spanish Government issued in 2009 the initiative "Aporta" to promote the development of the open data culture. This initiative has produced the country's open data portal found at datos.gob.es through which the national catalogue of open data is made available for re-use. The "Aporta" initiative also comprises of a vast array of actions covering transparency, economic development, legislation, public-private cooperation schemes, national and international coordination, as well as assessment and support.

In the case of Vienna a lot of information is meanwhile available by Open Data initiative[22]. It is a sub section of the Austrian Open Government Data initiative[23]. The following aspects are e.g. covered:

- Geo based data
- Mobility data
- Environmental data
- Budget data
- Statistical data

Geo-based data contains e.g. information about land usage, land models and land surveys, solar potentials and many more. Apart from available information on OGD portals, information about weather data can be retrieved by local meteorological institute Zentralanstalt für Meteorologie und Geodynamik (ZAMG). The data contains e.g. information about global radiation, air pressure, temperature, rainfall, wind, etc. and can be freely downloaded based on station, timeframe and time granularity. The Austrian statistics institute[24] offers as well a lot of publicly available information for free download covering the following aspects:

- Labour
- Population
- Education and culture
- Health
- Social
- Economy
- Energy, environment and mobility, and,
- Wealth and innovation.

At European level Eurostat[25] provides statistical data both for metropolitan areas and on country level. Even though for example the CITYkeys indicator “Gross domestic product” can be found from Eurostat, further analysis of metropolitan areas reveals that in that case of for example Tampere the aerial boundaries are the whole Tampere region, which consists of

[22] https://open.wien.gv.at/site/open-data
[23] https://www.data.gv.at/
[24] http://statistik.at/web_de/statistiken/index.html
[25] http://ec.europa.eu/eurostat

22 municipalities[26], and indicators should be calculated for only one of them. In the case of Rotterdam, the Eurostat regional data (NUTS2 level) encompasses the province of South-Holland (60 municipalities in total). In the case of Zagreb, it consists of 21 counties including the City of Zagreb. The City of Zagreb has a status as town and county in same time[27].

5.4 CITYkeys KPIs and open data

As explained in the beginning of this section, the calculation formulas of the quantitative city KPIs were transformed into needed data sets and a table of those needed data sets was created. Then this table was matched with the already existing inventory of cities' open data. In total the calculation of all city KPIs requires 116 raw data sets. Some KPIs already exists pre-calculated, and hence the amount of data sets slightly reduces. Depending on the city and calculation method less than 20 of open data sets were applicable for the selected CITYkeys KPIs. This means that roughly 85% of needed data sets are either from city's internal systems, are not available and need to be generates or they are public but do not qualify as open data.

After we combined the answers from cities into one table, it confirmed that about 52 data sets were such where at least one of the cities had answered that data is not available. And 21 were unavailable in at least two cities.

5.5 Dispersion of data sources within the city

The analysis of the available data sets also included their data sources location within the city administration. Different cities have different nominations and divisions for their internal departments, and also different outsourcing policies. In order to simplify the analysis the terminology was merged as follows (as seen in table 5):

- Mobility/Transport,
- Environment,
- Housing/urban planning,
- ICT,
- Social services,
- Open government/participation,
- Mayor's office/Economy (including statistics),
- Non-ICT Infrastructures, and,
- External organisations and companies.

With combined department names, the internal department where the most of the data originates is "Mayor's office/Economy (including statistics department)" with an average of 23% of the available data sets. Still external organisations and companies seem to be the most common source for needed data sets (32%), because some cities have outsourced their functions such as water management, electricity, and in addition some data sets are available nationwide, when the data comes for example from national statistical office.

[26] http://www.pirkanmaa.fi/en/tampere-region/municipalities-tampere-region
[27] For more information, see http://www.dzs.hr/Hrv/publication/2009/12-1-5_1h2009.htm

Table 5. Location of data within the different departments of city administrations

	Rotterdam	*Tampere*	*Vienna*	*Zaragoza*	*Zagreb*	*Average*
Mobility/Transport	3	5	4	7	2	4
Environment	10	6	15	17	3	10
Housing/urban planning	5	16	35	8	3	13
ICT	-		1	6	2	2
Social services	2		1	2	2	1
Open gov/participation	-		8	1	-	2
Mayor's office/ Economy (incl. statistics)	42	5	17	-	33	20
Non-ICT Infrastructures	7		-	-	-	1
External organisations and companies	13	57	12	37	15	27

5.6 Data for project KPIs: cities' initial testing plans

Most of the project KPIs are qualitative or semi-quantitative. They are defined so that they will be evaluated for example based on project documentation or an interview with the project manager. Some project KPIs also use quantitative data and the associated generic data sets are listed in Appendix 4 of this report. However, as explained before, the boundaries of these data sets need to be further defined in each project separately because the coverage of relevant data is always project specific. The project indicators will be addressed more in detail in CITYkeys testing phase (project Task T2.4). This section, however, presents cities' initial plans for testing projects, relevant indicators and plans for data collection using the following table (see Table 6).

The initial testing plans in each of CITYkeys five partner cities are presented in the following sub-sections. The initial selection of indicators relevant for those testing projects are presented in Appendix 4 of this report and will be further refined along the development work within project work package two towards the testing (project task 2.4).

Table 6. Cities initial plans for testing project and data collection

City	
Project name	
***Start and end** date of the project*	
***Project type** add relevant sector (e.g. building, energy, transport, ICT)*	
General description	
***Stakeholders** involved in the project including funding body*	
***Definition of the boundaries of the project** (geographical or other), please define the scope of the project (what is included and what is excluded)*	
***List of CITYkeys data sets relevant for the project**. You can make the selection with some colour in the excel list of project data sets*	
***Project data collection**. If the project has ended or is ongoing please describe your methodology for data collection, storage, etc. If the project is starting now please describe your planned methodologies, databases, etc. (see next row for non-quantitative data)*	
***Non-quantitative data:** please indicate your ideas how to evaluate the non-quantitative indicators for your project*	

5.6.1 Testing plans in Rotterdam

Table 7. Initial testing plans in Rotterdam

City	Rotterdam
Project name	E-mobility 3 cities NL - Boosting Electromobility Amsterdam - Rotterdam - Utrecht
***Start and end** date of the project*	Start: Q3 2012 End: Q1 2016
General description	Electric transportation is necessary for improving the air quality in cities and allows for the use of renewable energy in transport. The market for charging infrastructure is not fully developed. The business case for further investments in charging infrastructure can be bolstered by preparing the electrical grid to meet increasing demand. The Boosting Electromobility project stimulates the development of electric transport in the Randstad region by expanding the public charging

	infrastructure with strategically placed chargers and by promoting the use of electric vehicles. The four partners aim to increase the share of electric transport within the Randstad region. This region is an industrial and metropolitan area in west-central Netherlands including the cities of Amsterdam, Rotterdam, Utrecht and the Hague. The Randstad is one of the most important economic areas in Europe. The region's dense population and extensive economic activity make it suitable for electric transport. Actions include: 1. Prepare innovative and shared procurement of charging infrastructure 2. Install and manage chargers and fast-chargers 3. Initiate stakeholder-platforms 4. Promote electric transportation 5. Monitor and share knowledge
Stakeholders *involved in the project including funding body*	The project is a cooperation between the metropolitan area of Amsterdam, the cities of Rotterdam and Utrecht and the Royal Dutch Touring Club ANWB. The project is funded by the European Commission under the LIFE+ program.
Definition of the boundaries of the project *(geographical or other), please define the scope of the project (what is included and what is excluded)*	City level
List of CITYkeys data sets relevant for the project*. You can make the selection with some colour in the excel list of project data sets*	Selected in Appendix 4. ITU L.1440 methodology for assessing the impacts of smart city projects (see Appendix 2) will be tested within T2.4 in Tampere. Therefore, direct and indirect/embodied energy use and CO_2 will be tested. It is however still open whether a case project will be used for that purpose or if that evaluation will be made separately.
Project data collection*.* *If the project has ended or is ongoing please describe your methodology for data collection, storage, etc.* *If the project is starting now please describe your planned methodologies, databases, etc. (see next row for non-quantitative data)*	Project nearly finished. Relevant data will be available.
Non-quantitative data: *please indicate your ideas how to evaluate the non-quantitative indicators for your project*	Interview with project manager
Additional information*, e.g. link to*	http://ec.europa.eu/environment/life/project/Project

project web-page	s/index.cfm?fuseaction=search.dspPage&n_proj_id=4407 http://www.boostingelectromobility.eu/ http://www.rotterdam.nl/elektrischrijden http://www.rotterdam.nl/rotterdamelektrischineuropeseprojecten

5.6.2 Testing plans in Tampere

At the moment of writing this report (January 2016) Tampere has two relevant projects in mind for testing. Whether both of those or only one of them will be finally used in testing will be decided later.

Table 8. Initial testing plans for the first possible testing project in Tampere

City	Tampere
Project name	"Solutions for electric mobility - Tampere leads the way"
***Start and end** date of the project*	11/2014-9/2016
General description	The project is about Electric mobility. It has 4 work packages: 1) Planning and implementing electric bus system procurement. The electric bus system will work as an innovation platform for ITS, 2) Innovation competition for electric transport, 3) Designing the user experience of electric transport in Tampere, 4) Communication and marketing. Project summary: Tampere aims to become a forerunner in electronic transport both nationwide and internationally by 2025. Through this project Tampere takes a significant step towards electric bus traffic and in promoting other innovations in electric transport. City of Tampere will be the first city in Finland to acquire electric buses as a public procurement for public transport services. One criterion in the procurement is that the electric bus system will be used as an innovation platform for intelligent transport systems. The procurement process and the lessons learned will be summarized in a procurement guide, which can be used by other Finnish cities. The project will create the foundation for significantly scaling up the electric bus system in the future. Furthermore, the project will determine criteria for designing the user experience of electric transport in Tampere as well as search for new electric transport solutions through an open innovation competition.

***Stakeholders** involved in the project including funding body*	Stakeholders: Funding by Tekes (Finnish innovation agency) TKL (Public transport operator, owned by the city) City of Tampere Robustco ltd (consulting company)
***Definition of the boundaries of the project** (geographical or other), please define the scope of the project (what is included and what is excluded)*	The activities mentioned in the project plan are included. The geographical boundary is Tampere.
***List of CITYkeys data sets relevant for the project**. You can make the selection with some colour in the excel list of project data sets*	Relevant project KPI's are shown in Appendix 4. Regarding the city indicators, Tampere will be focusing on mobility and innovation related indicators. Those are the ones that will be most relevant in the selected project as well, so there will be a connection between the two levels. ITU L.1440 methodology for assessing the impacts of smart city projects (see Appendix 2) will be tested within T2.4 in Tampere. Therefore, direct and indirect/embodied energy use and CO_2 will be tested. It is however still open whether a case project will be used for that purpose or if that evaluation will be made separately.
***Project data collection**.* *If the project has ended or is ongoing please describe your methodology for data collection, storage, etc.* *If the project is starting now please describe your planned methodologies, databases, etc. (see next row for non-quantitative data)*	The project doesn't have a methodology for data collection.
***Non-quantitative data:** please indicate your ideas how to evaluate the non-quantitative indicators for your project*	Interviewing project stakeholders.
***Additional information**, e.g. link to project web-page*	http://www.tampere.fi/tampereen-kaupunki/projektit/kaupunkikonsernin-hankkeet/eco2-hanke/hankkeet/sahkoisenliikenteenratkaisut_0.html

Table 9. Initial testing plans for the second possible testing project in Tampere

City	Tampere
Project name	Co-ZED, constructing close to zero energy district
***Start and end** date of the project*	Reported (2013-2015), construction is

	starting.
Project type *add relevant sector (e.g. building, energy, transport, ICT)*	New buildings and connection to energy system district heating and possible lake water heating. In addition peak power options were studied.
General description	
Stakeholders *involved in the project including funding body*	Tekes, Tampereen kaupunki, Tampereen sähkölaitos, Tampereen kaukolämpö, Verte, Fidelix, Skanska
Definition of the boundaries of the project *(geographical or other), please define the scope of the project (what is included and what is excluded)*	Härmälänranta in Tampere
List of CITYkeys data sets relevant for the project*. You can make the selection with some colour in the excel list of project data sets*	Energy consumption (heating, cooling, electricity), peak power, CO2 emission, also user feedback about preferred options were collected.
Project data collection*.* *If the project has ended or is ongoing please describe your methodology for data collection, storage, etc.* *If the project is starting now please describe your planned methodologies, databases, etc. (see next row for non-quantitative data)*	normal measurement
Non-quantitative data: *please indicate your ideas how to evaluate the non-quantitative indicators for your project*	Questionnaires might be possible.

5.6.3 Testing plans in Vienna

Table 10. Initial testing plans in Vienna

City	Vienna
Project name	SMARTER TOGETHER H2020-SCC1 lighthouse project
Start and end *date of the project*	01.02.2016 – 31.01.2019 (implementation) / 31.01.2021 (monitoring)
General description	Vienna would like to use the project area of the SMARTER TOGETHER Light House Project. Large social housing estates mainly built between 1945 and 1985 and owned by the City of Vienna – Wiener Wohnen or Non-Profit Housing Cooperatives (i.e. BWSG) need to be refurbished in the upcoming years. This refurbishment will have a big impact of local energy consumption unless measures at

	the supply side, especially at the district heating infrastructure, are taken.
Stakeholders *involved in the project including funding body*	European commission, City of Vienna, BWSG (housing company), Wiener Stadtwerke (Utility company), Kelag Wärme GmbH (district heating operator), Siemens Austria, Sycube, Austrian Post, AIT, University of St. Gallen, local citizens and SMEs
Definition of the boundaries of the project *(geographical or other), please define the scope of the project (what is included and what is excluded)*	Vienna's smart city lighthouse area is part of Simmering, the 11th and one of the outer districts in the South-East of Vienna. Simmering is a traditional workers' district. The area selected for SMARTER TOGETHER is located in its North-West. It is an area "in between" vast redevelopment sites (Vienna main station, Mautner-Markhof Areal), but not directly connected to them, and is as a whole a refurbishment area. The area covers about 1.5 km2 with some 21,300 inhabitants, hosts 12,000 jobs and is characterized by important social housing from between WW1 and WW2. With 14,200 inhabitants per km2 it is a rather dense area, way above the average in the district or the city as a whole.
List of CITYkeys data sets relevant for the project*. You can make the selection with some colour in the excel list of project data sets*	Selected in Appendix 4
Project data collection*.* *If the project has ended or is ongoing please describe your methodology for data collection, storage, etc.* *If the project is starting now please describe your planned methodologies, databases, etc. (see next row for non-quantitative data)*	Assess the status quo (t=0) to check the availability of data and the quality of the data sets. As second data point we can gather data on the planned end state of the project. Potentially, we could compare this desired end state with the "real" end state at the end of the Smarter Together project.
Non-quantitative data: *please indicate your ideas how to evaluate the non-quantitative indicators for your project*	Interviewing project stakeholders.

5.6.4 Testing plans in Zagreb

Table 11. Initial testing plans in Zagreb

City	City of Zagreb
Project name	ZagEE – Zagreb energy efficient city
Start and end *date of the project*	1 April 2013-31 March 2016

General description	The ZagEE - Zagreb energy efficient city represents an initiative for encouraging and realizing significant energy savings by implementing economically viable and energy efficient technologies and measures on buildings of different purposes owned by the City of Zagreb as on the public lighting system. The project is implemented as part of the IEE program for technical assistance 2012 - Mobilizing Local Energy Investment (MLEI) which is used to finance technical assistance and production of the necessary technical documentation for the application of measures of energy efficiency and renewable energy sources on objects included in the ZagEE project. The implementation of the project ZagEE began on April 1st, 2013 and it shall last for three years. The project value is 1.813.464 EUR. Energy refurbishment of public buildings includes the implementation of standard energy efficiency measures (restoration of facades, roofs, external joinery, internal lighting, change of energy sources…), as well as the application of renewable energy systems (solar collectors and photovoltaic systems) on the said buildings. The modernization of a part of public lighting will be the first project of such size in Croatia which will feature LED lamps with regulation during late night hours. The ambitious plan of renovating 87 public buildings and the replacement of a part of energy inefficient public lighting through the ZagEE project with an estimated investment of 29.379.114 EUR will result in high energy savings and a reduction of CO_2 emissions. The implementation of energy refurbishment investments, the local economy will gain a significant initial incentive through creating new business opportunities, new workplaces as well as contribute to positive economic shifts and boost economic development as a whole. The ZagEE project is the first project of this size and complexity in Croatia and wider

	region and the experience gained through its implementation can serve as an example and guidance to other public, local and regional self-governments that wish to implement energy refurbishment on their territory.
Stakeholders *involved in the project including funding body*	Co-funded by the IEE (Intelligent Energy Europe) programme of the EU (as part of the IEE program for technical assistance 2012 - Mobilizing Local Energy Investment (MLEI)) Stakeholders involved (in the attachment Project core team)
Definition of the boundaries of the project *(geographical or other), please define the scope of the project (what is included and what is excluded)*	The City of Zagreb.
List of CITYkeys data sets relevant for the project*. You can make the selection with some colour in the excel list of project data sets*	Selected in Appendix 4
Project data collection. *If the project has ended or is ongoing please describe your methodology for data collection, storage, etc.* *If the project is starting now please describe your planned methodologies, databases, etc. (see next row for non-quantitative data)*	Collection of the data of energy consumption by smart metering (electricity, heating, water, weather forecast) for technical analysis. Collection of bills of energy consumption for economic analysis. Investments initiated by energy refurbishment of the public buildings and modernisation of the public lighting. Capacity building in energy refurbishment of public buildings: trainings, number of workshop for the building managers and workshops for the city administration.
Non-quantitative data: *please indicate your ideas how to evaluate the non-quantitative indicators for your project*	Interviewing project stakeholders.
Additional information*, e.g. link to project web-page*	http://zagee.hr/?lang=en

The organization of the project team and data in the ZagEE project is presented in Figure 2.

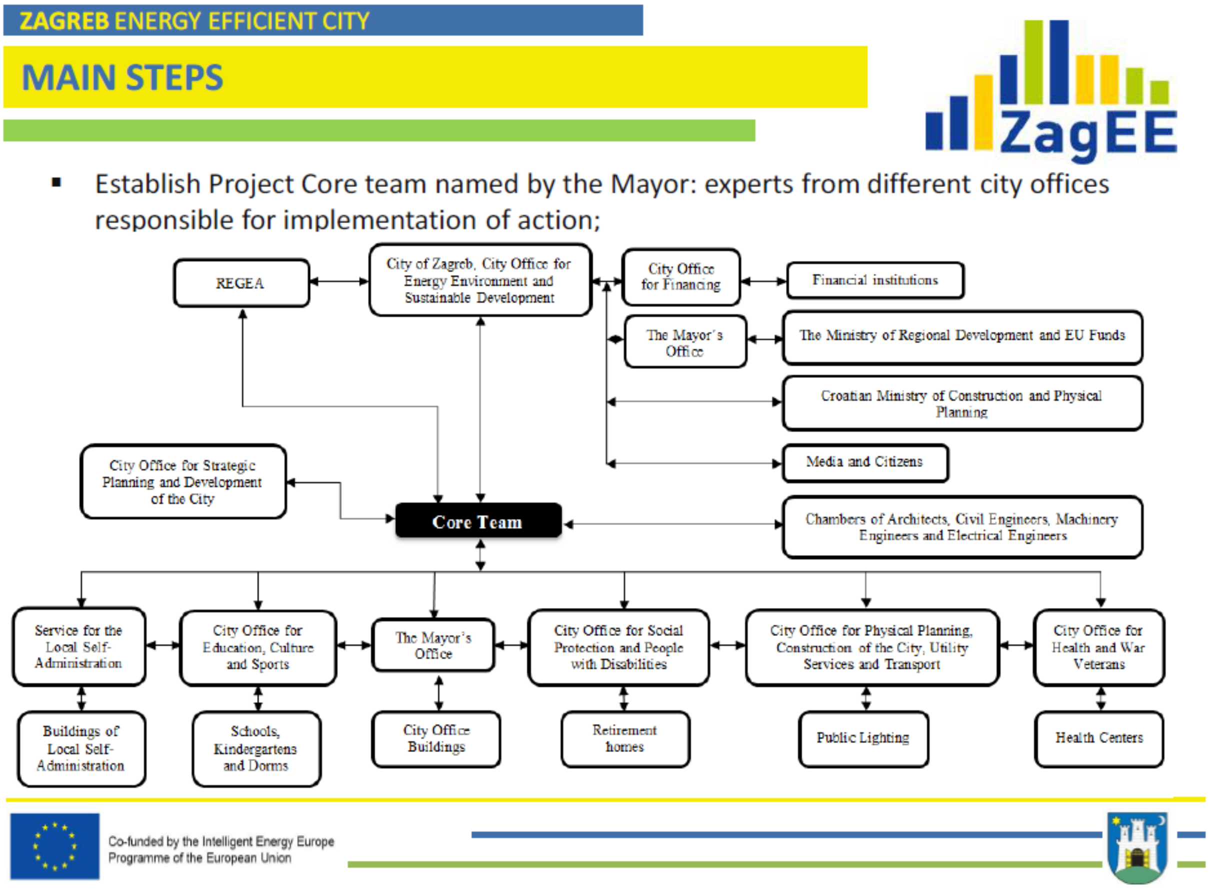

Figure 2. The organization of staff and data in the ZagEE project

5.6.5 Testing plans in Zaragoza

At the moment of writing this report (January 2016) Zaragoza has two relevant projects in mind for testing. Whether both of those or only one of them will be finally used in testing will be decided later.

Table 12. Initial testing plans for the first possible testing project in Zaragoza

City	Zaragoza
Project name	CIEM (Centro de Incubación Empresarial de la Milla Digital)
Start and end *date of the project*	Start construction: 2010, project on-going
General description	zero-emissions building holding a start-up incubator (so we can test both energy performance and innovation)
Stakeholders *involved in the project including funding body*	Zaragoza City Council, Init services (start-up incubation services), Zeroaplus (CIEM Data Lab project / energy data collection and exploitation)
Definition of the boundaries of the project *(geographical or other), please define the scope of the project (what is included and what is excluded)*	The project consisted on the construction of a Zero Emissions Building (CIEM) to serve as a start-up incubation facility. So both low energy construction and innovation are including in the project.

List of CITYkeys data sets relevant for the project*. You can make the selection with some colour in the excel list of project data sets*	Selected in Appendix 4
Project data collection*.* *If the project has ended or is ongoing please describe your methodology for data collection, storage, etc.* *If the project is starting now please describe your planned methodologies, databases, etc. (see next row for non-quantitative data)*	Data collection about the economic / innovation part (start-up incubation services) is done through the "Annual CIEM's Wealth generation report" http://www.ciemzaragoza.es/2014/04/el-informe-de-generacion-de-riqueza-%C2%A1ya-esta-en-ingles-wealth-generation-report/ Data collection about the energy consumption is on the process of being collected and published.
Non-quantitative data: *please indicate your ideas how to evaluate the non-quantitative indicators for your project*	We have internal surveys with non-quantitative data about the services provided to entrepreneurs and their overall satisfaction with the project.
Additional information*, e.g. link to project web-page*	www.ciemzaragoza.es

Table 13. Initial testing plans for the second possible testing project in Zaragoza

City	Zaragoza
Project name	"Caminos escolares"
Start and end *date of the project*	March 2013
General description	project "safe routes to schools" which includes mobility, environment and education
Stakeholders *involved in the project including funding body*	City Council and School Community
Definition of the boundaries of the project *(geographical or other), please define the scope of the project (what is included and what is excluded)*	Geographical boundaries is the city, and inside the city, those areas around the schools inside the project. Other boundaries: small civil works (bike routes, modification of barriers, signalization...), dissemination of good mobility practices, improving and using public space, participation, local shop engagement...
List of CITYkeys data sets relevant for the project*. You can make the selection with some colour in the excel list of project data sets*	Cycling roads, # of citizens reached (by the project), # of citizens considered stakeholders in the project, total transport energy consumption, CO2 emissions, emissions (all), project costs spent on local suppliers, contractors and service providers, number of

	green jobs created by project, Number of jobs created by project, Project's initial total investments, Project's annual total cash inflow, Project's annual total cash outflow, Average delay per vehicle kilometre (before and after project)
***Project data collection**.* *If the project has ended or is ongoing please describe your methodology for data collection, storage, etc.* *If the project is starting now please describe your planned methodologies, databases, etc. (see next row for non-quantitative data)*	Methodologies: – mobility surveys, activity surveys – meetings (minutes of meetings) – design of "mobility spiders" (pedestrian route optimization)
***Non-quantitative data:** please indicate your ideas how to evaluate the non-quantitative indicators for your project*	– activity surveys (to measure the level of satisfaction) – meeting minutes
***Additional information**, e.g. link to project web-page*	www.zaragoza.es/ciudad/caminoescolar/

6. DATA PRIVACY

6.1 Open data, public data, internal data

In chapter 5 we concluded that about 15% of data sets are available as open data, 44% as public and 5% as confidential. The difference between public and open data is that data can be publicly available, on internet, even if not in open data formats - for example, it is not provided with an open license or is not in structured format. This data is defined as public data, because it's publicly available, but does not qualify as open data. In Zagreb, for example about 44% of data sets are available from public sources. In that case there is a web page holding a lot of statistical and numerical information about the city[28]. Even when the data is available from public sources, it doesn't mean, that it is more readable for computers, or data collection would be easier, than for example data in printed format. In almost every case, when data is published on reports or on HTML-pages, it must be manually inserted into KPI-calculation system.

For the on average 5% of data sets which are labelled as confidential the reason can be, for example, that either exposing the data would raise privacy issues or the data is in such a form that making it public is complicated. For example one city estimates that around 90% of the city data (currently behind internal interfaces) could be made open. The challenge is that it would need significant additional work since the level of publicity should be defined for all data and then should be processed accordingly.

Some data is not made open in its raw format due to privacy protection, including citizen privacy, and confidentiality issues. Examples of these kind of data are un-aggregated population data, private buildings energy consumption etc. CITYkeys project and indicators calculation will not access this raw data or any confidential data. The indicator calculation, during the testing phase, will be based on aggregated and anonymized data.

6.1.1 Data access methods

Data accessibility can be divided into three categories: 1) Data is available over common networking protocols without access constraints, 2) data is available online, but requires authentication, and, 3) data is not accessible online, or requires manual work to get data out from internal systems. Open data mostly satisfies the first category requirement, as it is also part of the open data definition that data should be accessible in machine readable format, at machine findable location.

Data sets which contain confidential information are either not exposed into public internet or are available after authentication.

A third category is then data, which is either extracted from documents, or from internal systems and then a file or resulting plain number is transferred as is, and later access to that data requires same process to be repeated again.

6.2 Cities' data privacy policies

Table 14 presents the data privacy policies and procedures to handle data collected from public sources in the five partner cities.

[28] http://www1.zagreb.hr/zgstat/ See "Grad Zagreb - osnovni statistički podaci (hr i en)" for data available in English.

Table 14. The procedures in the five cities to handle data collected from public sources

City	Installed protocols, national and/or EU level regulations
Tampere (partner 5)	The city of Tampere follows the Finnish Personal Data Act, which complies with the EU Directive on the Protection of Personal Data. If personal data is being collected, the purpose of data collection should be known and the collected data shouldn't be used for any other purpose. The owner of the data i.e. the administrator of the register should also be defined. A register description should be made when a new person register is established. City of Tampere has specific instructions on what aspects should be considered when personal data is collected. Un-aggregated population data within Tampere's internal system (Oracle Spatial) is a weekly copy of information stored in national Population Information System (original registry owner is Finland's Population Registry Centre). For this data Tampere is actually following the same regulations as the previously mentioned national registry centre. In addition, Tampere has many other internal systems of its own that contain personal data of its citizens. In these cases City of Tampere is the registry owner. All registries have to have their own data files that are describing in what conditions data may be used. Tampere has three different GeoServer[29] installations for different purposes. One GeoServer is only for providing open data, which holds 67 data sets (23.01.2015). Another GeoServer is primarily for serving data for Tampere's own public map service, and access to that server with 207 data sets is restricted with authentication. The third is a public facing GeoServer with 335 data sets is for internal use, with authentication enabled, even though access can be granted to third parties as well for research purposes. In Tampere, access to data sets is controlled with authentication and also with role-based access control, giving possibility to only give access to the data sets that the user really needs. Access constraints are also defined in the metadata according to Inspire metadata schema. The confidential data sets are not exposed to GeoServer and access to them is only possible for persons with access to city's primary geodatabase.
Rotterdam (partner 6)	The city of Rotterdam follows the Dutch 'Protection Personal Data Act' (Wet bescherming persoonsgegevens Wbp), which complies with the EU Directive on the Protection of Personal Data (95/46/EG). If personal data is being collected, the purpose of data collection should be known and the collected data shouldn't be used for any other purpose. The owner of the data i.e. the administrator of the register should also be defined. A register description should be made when a new person register is established. Data can be anonymized or aggregated to a different level so Personal Data are still protected.
Vienna (partner 7)	The city of Vienna follows the following Directives and Federal Laws and regulations in their data collection:

[29] http://geoserver.org/

	•Directive 95/46 / EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, Official Journal no. L 281, 23/11/1995, p 0031-0050 •Federal Law on the Protection of Personal Data (Data Protection Act 2000 - Data Protection Act 2000), Federal Law Gazette I No. 165/1999, as amended. •Federal Law on regulations to facilitate electronic transactions with public entities (E-Government Act - E-GovG), Federal Law Gazette I No. 10/2004, as amended... •In addition, there are numerous data protection regulations in the respective sectoral laws (such as: § 152 Commercial Code 1994, § 18 Registration Act 1991).
Zaragoza (partner 8)	The Municipality of Zaragoza informs users that the collection and processing of personal data that is made through its online site is subject to the provisions of the current legislation on data protection, the Organic Law 15 / 1999, Protection of Personal Data, Royal Decree 1720/2007 of 21 December, approving the Regulations implementing Law 15/1999 of 13 December on the Protection of Data approved personal and other applicable regulations, in order to ensure at all times the privacy of users, and the secrecy and security of their personal data. Access to and the use of the online site constitutes acceptance of this privacy policy is detailed below. The collection and processing of personal data that is made through the electronic town hall of Zaragoza has the sole purpose of providing the services provided therein, as well as those required by the users it is accompanied by the mandatory reporting obligations set out in Article 5 of Law 15/1999, of December 13, Protection of Personal Data. Users are also informed that their data will be treated with appropriate security measures in accordance with the provisions of Article 9 of Law 15/1999, of December 13, Protection of Personal Data and Title VIII Royal Decree 1720/2007 of 21 December, approving the Regulations implementing Law 15/1999 of 13 December on the Protection of Personal data. Also, data confidentiality is guaranteed and will not be communicated to third parties except in cases necessary to manage the services provided through this online site as well as under the current rules. Users may exercise their rights of access, rectification, cancellation and opposition of their data by contacting, in the required legal form to the City of Zaragoza - (Department of Science and Technology) located at Via Hispanidad No. 20, 50071 Zaragoza. Publication of lists containing personal data that could be carried out in this online site complies with the current Data Protection legislation. These lists are not a source of public access and cannot be reproduced, transmitted and recorded by any retrieving information system without the consent of those concerned and the approval of the City of Zaragoza.

Zagreb (partner 9)	The city of Zagreb follows the national legislation, which complies with the EU Directive on the Protection of Personal Data.

7. CONCLUSIONS

Based on the KPI definitions developed in WP1, the data sets to be used in the calculation of different KPIs have been identified. After the data set definition, the available data sources, their reliability, data access methods, existing data formats and level of confidentiality in the five partner cities were analysed. In addition, potential privacy issues were screened.

7.1 Summary of achievements

The CITYkeys set of indicators consists of a mixture of quantitative and semi-quantitative indicators. The semi-quantitative indicators for example provide an assessment of the way smart city projects are executed, the way the city government stimulates smart city development and of the potential of smart city projects to be taken up by other cities. For the semi-quantitative indicators data needs to be collected using interviews or an analysis of project or policy documents. There is a difference in the data collection process between the project and the city indicators. The project indicators are meant for assessing the success and potential for replication of smart city projects. In individual project assessments data needs to be collected from the project office, the project leader and/or others closely involved in the project. Cities might want to streamline the data collection for all their smart city projects, creating a reporting system and specific databases, but even then a number of the project indicators will require (qualitative) information that can only be gathered by involving persons involved in the project (e.g. through interviews, questionnaires).

Data for the majority of the city indicators can be retrieved from statistical sources within the city administration or national or European level. The problem of such data is that it typically presents average (annual) figures for the whole city. For a city it may however be more interesting to analyse the differences between districts. Spatial data makes it possible to calculate indicators also for geographically restricted areas such as city districts. It is expected that CITYkeys testing phase will allow, to some extent, the evaluation of a project's impact on city level as CITYkeys KPIs contain project-to-city link for many indicators. Additional sources of open data and public data can be found in national or European institutions (e.g. statistics bureaus). In any case, national data sources provide excellent support for data availability in all the five partner cities. EUROSTAT is potentially also of support for many data, but the problem is that information is given for regions and not for cities or municipalities as defined nationally.

It is important to note that data for all indicators will obviously not be available immediately. A city that engages in smart city indicators starts a process. The CITYkeys indicator framework is a methodology for such a process. The city will need to continuously develop the indicators to be used by the city and the data collection mechanisms. Moreover, the definitions behind certain data sets and data quality obviously differ between countries, between cities and between city departments. The quality of the overall assessment depends on the quality of the indicators, which in turn depend on the underlying data. Managing data quality throughout the process is thus crucial. When making comparisons a transparent communication of all meta data underlining the data sets is important, since it can explain how reliable the data is – and thereby the results of the corresponding indicator(s).

The data sets relevant for the CITYkeys KPIs were identified and the required data sources were analysed regarding their availability, reliability, data access methods, existing data formats and the level of confidentiality in the five partner cities. Special attention was paid in the analysis of the open data sets. It is to be noted that the system boundaries for project indicators are project specific and that the semi-quantitative indicators require an interview-based approach. Therefore the data set analysis was mostly relevant for the quantitative city KPIs. Most of the KPIs are not available ready-calculated, and an indicator often combines information from two or more single data sets, while some (common) data sets were necessary to calculate a range of indicators. Next, an inventory of the availability of the needed data sets in the cities was made. In total the smart city KPIs need 116 raw data sets. Some KPIs already exist pre-calculated, and hence the amount of data sets slightly reduces. The data availability rates for the data sets needed for the KPIs in the partner cities are as follows: Zaragoza 82%, Vienna 80%, Tampere 77%, Rotterdam 71% and Zagreb 52%. Average availability in cities is 72%.

Depending on the city and the calculation method less than 20 of open data sets were directly applicable for the selected CITYkeys KPIs. This means that roughly 85% of needed data sets are either from city's internal systems; are not available and need to be generated; or they are public but do not qualify as open data. For an average of 13% of the data sets, their availability remains unknown at this stage of the project giving a possibility for slight changes in the availability rates. These data sets will need additional investigation during the testing phase but it is expected at this stage that those will not be available easily. The most difficult data sets for cities in terms of availability are:

- Expenditures by the municipality for a transition towards a smart city in €
- Total food demand (tonnes)
- # of green jobs (related to environmental service activities that contribute substantially to preserving or restoring environmental quality), and,
- # of houses or households with grey water reuse capability.

These data sets are unavailable (or the availability remains unknown) in all cities and they have in common that most of them are not exactly countable, but require sophisticated calculations.

External organisations and companies are the most common source for needed available data sets (32%), because some cities have outsourced their functions such as water management and electricity. In a few exceptional cases data may come from a national statistical office or other nationwide sources. Most of the available data sets coming from city departments (23%) originate from Mayor's office/Economy (including statistics departments).

The share of the data sets available as open data in the five partner cities varies from 1% to 25%, and is 15% on average. In total the five partner cities have 666 open data sets in their portals, making an average of 133 data sets per city. However, the quantity of open data provided tells only half of the story since the quality and usability of data are often more important. Currently the quality and update intervals on cities' open data portals vary a lot for different data sets. Cities are following their own strategy to update certain information on their portal. Also data can be added during one project but not be updated after the project ends. The use of some of the data sets seems minimal. Some potential reasons include the lack of security on data quality and reliability. To put it differently, data sets might be added to the portal from the "supplier's point of view" instead of the "customers point of view". On the other hand, another city estimates that around 90% of the city data (currently behind internal interfaces) could be made open. The problem is that it would need significant additional work since the level of publicity should be defined for all data and then should be processed accordingly. Due to the increasing interest and possible uses of open data, there is, however, a clear need to have reliable open data that can be easily used. An expected potential co-benefit of CITYkeys project could be that data becomes more open and more frequently collected as cities start the process to evaluate their smart city (project) strategy. On the other hand, it is eventually recommended to automate the data collection as far as possible.

Within the project the status of these open data sets and rates of automated data collection will be re-evaluated at the end of the testing phase (project task T2.4). In CITYkeys WP3 new business opportunities will be identified and open data is a potential source of data for those.

Three portals, Rotterdam[30], Vienna[31] and Zagreb[32], are based on CKAN-platform, which is an open-source platform which was developed by The Open Knowledge Foundation (OKNF) and is managed by the CKAN Association (The Open Knowledge Foundation, 2015b). The cities of Tampere[33] and Zaragoza[34] have developed their own solutions.

A big share of 86.5% of open data sets is in easily machine readable formats. Five most common formats, which are not easily machine-readable, are: XLS, ZIP, ODS, PDF and HTML. Although all of them can be read programmatically, practical implementation of a reader is complicated and in any changes of structure the read process will probably fail. One easily doable movement towards better rated open data is to publish all Excel and ODS based data sets also as CSV files. According to the five star open data quality rating (Berners-Lee, 2015), the average quality of the open data in the five partner cities is in the range between 2 and 3 points, meaning that cities have published most of their data in structured format, but not everything in non-proprietary format. For example if a city publishes everything as Excel-spreadsheets it would result into rating 2.0 whereas publishing everything in CSV it would be then 3.0. The dominant formats are WFS with 32% share and CSV with 28% share. A big share of open data is spatial, almost half of the data sets, which makes WFS the most common format with its 32% share. It is common to publish one data set in multiple formats, which makes re-use easier by decreasing format-specific barriers, and also allows faster exploration of provided data.

Some data cannot be made open in its raw format due to privacy protection, including citizen privacy, and confidentiality issues. Examples of these kind of data are un-aggregated population data, private buildings energy consumption etc. Cities still can have these data in their internal systems and access to it can be given for other parties as well under certain conditions: no constraints; available online after authentication; and not accessible online, or manual work to get data from internal systems is required. All cities state to follow the law with regard to privacy.

7.2 Relation to continued developments

With the identification and analysis of the common data sets, a new phase in the operationalisation of the CITYkeys indicator assessment framework has started. Cities' preliminary plans for testing phase - including relevant indicators and plans for data collection - were identified in this report and will be refined later. Part of the cities prefer to focus on testing indicators on project level while other are more interested in city level. Apart of testing indicators in testing project, the next steps in city level will be to combine the information on calculation of the indicators, and the common data sets, to build algorithms for KPI calculation in harmony with existing (open data) systems. Then, the conceptual first draft of the framework in an overarching structure (the smart city performance measurement system) will be outlined. And finally, it will be evaluated to which extent it is possible to build a connection between project and city level.

[30] http://rotterdamopendata.nl/data set
[31] https://www.data.gv.at/suche/?publisherFilter_Stadt+Wien=on&connection=and&hideFilters
[32] http://data.zagreb.hr/
[33] http://www.tampere.fi/tampereen-kaupunki/tietoa-tampereesta/avoin-data.html
[34] https://www.zaragoza.es/ciudad/risp/

8. ACRONYMS AND TERMS

CSV .... Comma-separated values
DCAT .... Data Catalogue Vocabulary
GIS .... Geographic information system
HTTP .... Hypertext Transfer Protocol
KPI .... Key Performance Indicator
LOD............................. Linked Open Data
OGC .... Open Geospatial Consortium
REST .... Representational State Transfer
SPARQL...................... SPARQL Protocol and RDF Query Language
URI .... Uniform Resource Identifier
WFS............................. Web Feature Service
W3C .... World Wide Web Consortium

9. REFERENCES

Berners-Lee, Tim. 2015. Linked Data – Design Issues. Online. Available at http://www.w3.org/DesignIssues/LinkedData.html Accessed 08.12.2015, 27.07.2006

The European Parliament and the Council. 2007. Directive 2007/2/EC. Available at http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32007L0002

Fortune, Joyce and Diana White, 2006. Framing of project critical success factors by a systems model. International Journal of Project Management 24 (2006) 53–65.

Kolk, A., 2004. “A Decade of Sustainability Reporting: Developments and Significance.” International Journal of Environment and Sustainable Development 3, no. 1 (2004): 51-64

OGC. 2015. Web Feature Service. Online. Available at http://www.opengeospatial.org/standards/wfs Accessed 9.12.2015

The Open Knowledge Foundation. 2015. What is Open Data?. Online. Available at http://opendatahandbook.org/guide/en/what-is-open-data/ Accessed 08.12.2015

The Open Knowledge Foundation. About | ckan - The open source data portal software. 2015b. Available at http://ckan.org/about/ Accessed 14.12.2015

SCOPE, 2007. Sustainability Indicators: A Scientific Assessment. Edited by T. Hák, B. Moldan and A.L. Dahl. Washington: Island Press. 2

W3C. Semantic Web – W3C. 2015. Semantic Web. Online. Available at http://www.w3.org/standards/semanticweb/ Accessed 08.12.2015

APPENDIX 1: SCIS AND CITYKEYS

A short description of the currently running H2020 project Smart City Information System (SCIS)[35] is presented and preliminary, possible linking to the CITYkeys project is evaluated. The SCIS project brings together information for project developers, cities, institutions, industry and experts to collaborate on the creation of smart cities and an energy-efficient urban environment. SCIS focuses on energy efficiency (energy demand reduction, CO_2 emission, renewable energy). SCIS is oriented towards the neighbourhood/building level, so synergies with the project level of CITYkeys may be possible for projects that are not heavily ICT-oriented. Information about economic monitoring as well as social and policy monitoring is also collected by SCIS.

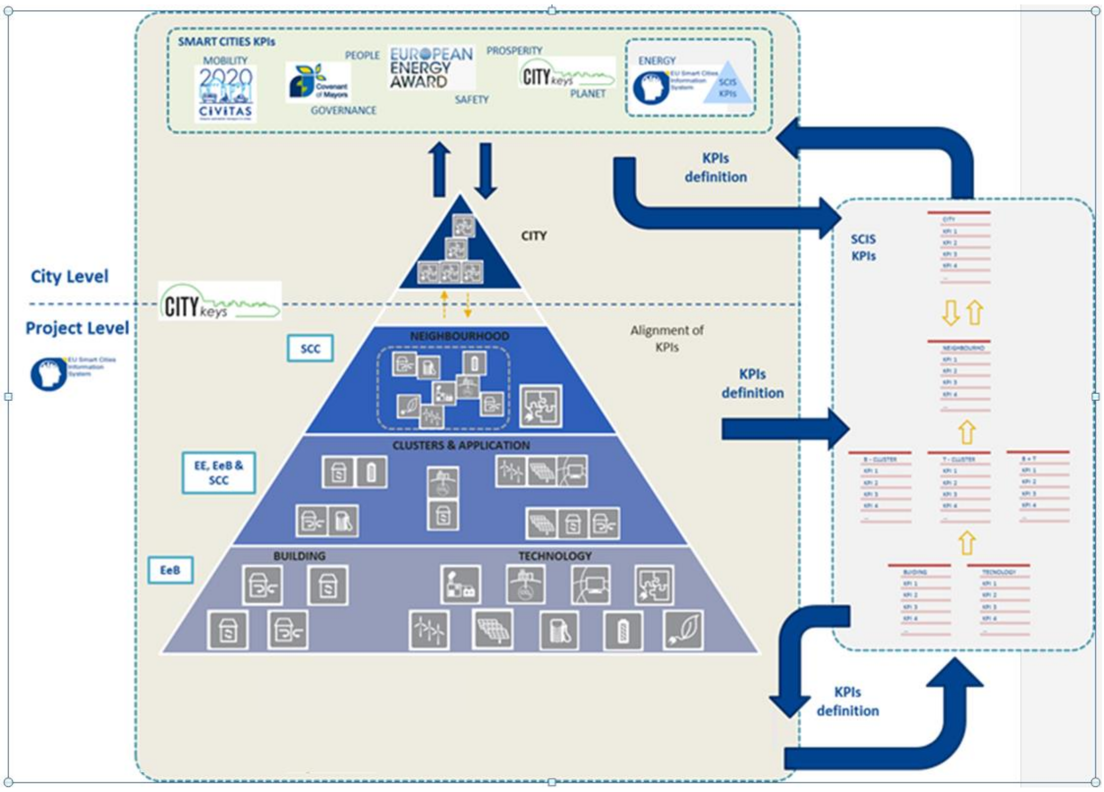

Figure 3. CITYkeys and possible linking to SCIS

Similar to CITYkeys, SCIS also deals with a city and a project level. As shown in Figure 3 a project (= demonstration site) in SCIS scope (and in some cases also for CITYkeys) can consist of one building or a number of buildings (same for ICT, Mobility or Energy Supply).

Within CITYkeys framework, the assessment takes place on project level and on city level. For SCIS the assessment of the impact is done by looking at each entity and trying to aggregate the overall impact (in some cases e.g. also building block level instead of building level might be the lowest granularity).

[35] http://smartcities-infosystem.eu/

Also similar to CITYkeys, for an existing area and project type the monitoring of the performance need to be done before and after the implementation, establishing a baseline for new sites.

To ensure that a link from the SCIS to CITYkeys project by means of transferring/exchanging indicator data is feasible, a clear strategy with regard of aggregating KPIs is needed. A possible way forward is as follows:

The SCIS project can propose a structure e.g. to separate the different achieved reductions by sector. SCIS can establish those KPIs with the existing KPI framework and provide them to CITYkeys project. CITYkeys ensures due to the simultaneous development of project and city level indicators that the results also can be used by the cities for replication and scalability of the ‘smart solution’ once the KPI framework is implemented by the city.

From the current perspective it is planned that SCIS will adapt a certain set of CITYkeys city level indicators to extend the SCIS framework. Additional alignment is planned for the indicators describing themes of qualitative, life cycle assessment and social aspects. The reason for that is that the focus of the SCIS project is more on the assessment of technological aspects of projects. Therefore from the current SCIS perspective it is useful to use the output from CITYkeys to extend its own indicator set by these categories.

Taking these specialties into account means that an exchange on indicator level between SCIS and CITYkeys in both directions makes sense mostly on the project level in particular for the indicators referring to the CITYkeys theme of energy and climate mitigation, because it addresses more the technical implementation aspects of realized projects. For this theme, an alignment of concerned indicators by means of formulating appropriate calculation rules (breakdown, aggregation, normalization) can be done, so that an exchange of information is possible without the risk to loose information due to different meanings and or uncertainties in indicator definition.

The current work plan regarding SCIS shows that the actions of data collection, reworking the concept and alignment are mostly done. Now the next phase is to do the alignment with new lighthouse projects to incorporate their specific issues into SCIS. Therefore the gaps of the current systems with regards to requirements of the new projects are identified. If there will be gaps by means of indicator information which is missing in SCIS but addressed in CITYkeys the attempt will be to grasp and incorporate this information from the CITYkeys project.

Therefore the effort to put an emphasis on the integration and linking of both projects is still pursued to make the linking integration and linking of both projects possible at the best.

APPENDIX 2: ITU TESTING PLANS

Introduction

Motivation

The International Telecom Union has developed guidelines for determining the environmental impacts, mainly first and second order energy use and GHG emissions, of ICT. This is known as the ITU L.1440 methodology for environmental impact assessment of information and communication technologies at city level.

The CITYkeys project was requested by the European Commission to test the ITU L.1440 methodology for environmental impact assessment of information and communication technologies at city level. This will be done in select cases as part of T2.4. Next to this, the ITU has developed a set of indicators for smart sustainable cities projects. The framework and indicators of the ITU smart sustainable cities initiative have been assessed within T1.2 and T1.3.

Goal

The goal of the test is to:

1. Investigate whether it is possible to apply the methodology of ITU L.1440 to evaluate the environmental impacts of ICT in CITYkeys case studies;
2. Outline what data is needed, the quality, level of detail and the availability of such data in the cities;
3. Evaluate which capacity is roughly required to gather data and calculate results in terms of expertise, competences and working hour's costs.

Scope

It is suggested to test the methodology in the cities of Tampere and Rotterdam, since these cities are likely to have ongoing smart city projects, have a good data management system and a research institute with additional expertise (for e.g. LCA) is available to assist. The data are to be a consistent set from the most recent year available and/or for the duration of the project.

Hypothesis

It is expected that it is possible to calculate these indicators, but it may be costly in terms of data and capacity needed. It is also unclear to which extent the data on embodied energy/CO2 of ICT will be available. If the methodology is workable, it may become an integral part of the CITYkeys methodology.

About this document

This document is a summarized manual on how to operationalize the ITU L1440 guideline. The matching CITYkeys indicators are: reduction in annual final energy consumption (by ICT), embodied energy of materials – quantitative (of ICT goods and networks), carbon dioxide emission reduction, reduction in life cycle CO2 emissions. See also the matching table below:

	ITU		CITYkeys indicators	
Energy	Tier 1	Direct energy use	Reduction in annual final energy consumption (by ICT)	Project and city level
	Tier 2	Indirect energy use	Embodied energy of materials – quantitative	Project level only
CO2 emissions	Tier 1	Direct CO2 emissions	Carbon dioxide emission reduction	Project and city level
	Tier 2	Indirect CO2 emissions	Reduction in life cycle CO2 emissions	Project level only

The document basically consists of three parts:

1. Simply calculating the ICT footprint of households and organisations: mainly relevant for CITYkeys on the city level
2. Calculating the footprints of projects and services: mainly relevant for CITYkeys on the project level
3. How to calculate expected footprint of pilots in different cities, for example to assess plans or to fill up data gaps: relevance for CITYkeys to be seen

Calculation steps ITU for households and organizations

1. Decide on project boundaries
    - Geographical: district, city, country
    - Sectoral: Inhabitants/households, organizations (large, small, # ICT-sector or not, # public administration)
    - Time horizon
2. Decide on goal and scope
    - What do you want to investigate and why?
    - What is included and what is not (cut-offs)? (more detail on the project boundaries)
    - What is your reference situation?
3. Decide on type of assessment
    - Tier 1: use stage (only direct CO2 emissions)
    - Tier 2: full life cycle (includes indirect CO2 emissions)
4. Estimate your inventory
    - Number of ICT items (appliances i.e. PCs, laptops, tablets, printers, mobile phones, servers, server rooms, base stations for cellular networks, ADSL/Wi-Fi etc.)
    - Use pattern (educated estimation of time spent ON and time spent STANDBY)
    - Power needed for each item and change in power for each use pattern
5. Calculate results
    - Multiply with emission factor from relevant electricity mixes (i.e. the country of use, country of production, country of decommissioning)
    - Show different intersections in results:

- Share of different ICT goods of total ICT footprint
- Impact of the different life cycle phases
- Share of different users
- Share per person/employee
- Share of use patterns
- Etc.

6. Sensitivity analysis
    - Include full life cycle (in case of a tier 1 assessment → expand to a tier 2 assessment)
    - Choose extremes from a bandwidth in literature
    - Double and halve important parameters and evaluate their impact on the overall results
7. Conclusions
8. References (very important!!!)
    - Statistics
    - Other LCA studies
    - Manufacturer

Note 1: Transport as a life cycle step was neglected (cut off).

Note 2: For data centres, the assessing city shall take into account all of the impacts of a data centre located within the boundaries of the city and used by households and organizations inside and outside city boundaries. From an environmental perspective it is better to have one large facility than several small scattered facilities in the region that are less energy efficient. Additional separate reporting is allowed.

Calculation steps ITU for ICT projects

Example: the real-time traffic monitoring of a bus line. The extra electricity used in servers may or may not be offset by the increased use of public transport, displacing private transport and corresponding CO_2 emissions.

1. Decide on project boundaries
    - City, country
    - Sector
    - Time horizon
      I.e. project is initiated in year 0 and runs from year 1-10.
2. Decide on goal and scope
    - What do you want to investigate and why?
      To monitor traffic conditions in a bus line by installing performance and consumption meters that communicate with other buses, the bus driver, bus stops and apps. May lead to increased use of public transport displacing private vehicles.
    - What type of project is it?
      See ITU L.1430 for a categorization of projects
    - What is included and what is not (cut-offs)?
      The displacement of private transport is out of scope (?)

- What is your reference situation?
  The buses ride without the software
- Prior conditions that may affect the project

3. Decide on type of assessment
   - *Tier 1: use stage (only direct CO2 emissions)*
4. Inventory stakeholders/project participants + contact details
   Municipal administration; City transports company; ICT GNS developers; Telecommunications company; Project proponent
5. Description of the project
   - *Reduce energy use/GHG emissions by optimization of time tables*
   - *Raise awareness on fuel use, traffic congestions*
   - *Better service to citizens, possibly leading to more users and less private transport (out of scope!)*
   - *Activities include: driving bus, operating the bus meters, data analysis, network communication to users and buses*
6. Estimate your inventory
   - Number of ICT items: hardware, software, networks (i.e. PCs, laptops, tablets, printers, mobile phones, servers, server rooms, base stations for cellular networks, ADSL/Wi-Fi etc.)
7. Calculate results
   - Multiply with emission factor from relevant electricity mixes (i.e. the country of use); *toe (diesel)*
   - Compare to baseline: benefit due to less fuel use; cost due to meter usage, data centre, network communications
8. Risk analysis
   - What risks can be identified and how can they be mitigated?
9. Conclusions
10. References (very important!!!)
    - Statistics
    - Other LCA studies
    - Manufacturer

Calculation steps ITU for ICT services

Example: Smart meter services for 30.000 users in a city area in country X. To understand the reduction in energy use needed to compensate for first order life cycle impacts of the ICT service.

1. Decide on project boundaries
   - City, country
   - Sector
   - Time horizon
     I.e. project is initiated in year 0 and runs from year 1-10.
2. Decide on goal and scope

- What do you want to investigate and why?
  To understand the reduction in energy use needed to compensate for first order life cycle impacts of smart meters.
- What is included and what is not (cut-offs)?
  Include transportation for installation and maintenance; exclude end-of-life treatment of old/new meters (since the GHG emissions of this phase are marginal)
- What is your reference situation?
  Regular meters
- Prior conditions that may affect the project

3. Decide on type of assessment
   - *Tier 2: use stage and lifecycle effects*
4. Description of the project
   - *Smart meters enable users to better manage their energy use through continuous information about the amount of energy being used and the associated costs, thus enabling a reduction in carbon emissions of energy users based on scenarios or actual data (in relation to a ref situation). Moreover, meters no longer need to be read manually (less vehicles/vehicle use: xx km/yr/smart meter avoided through automated meter reading, xx vehicles less needed)*
5. Estimate your inventory
   - Number of ICT items: hardware, software, networks (i.e. PCs, laptops, tablets, printers, mobile phones, servers, server rooms, base stations for cellular networks, ADSL/Wi-Fi etc.)
     This is where it gets technical, but these are just example terms to describe the equipment.
     Long Term Evolution (LTE) Radio Access Network with related site infrastructure (dedicated or shared?!): radio base station site; radio units; digital units; antennas; peripherals; cabinet; new antenna towers. LTE core network: servers and mechanics for the Evolved Packet Core to handle the LTE air interface. Dedicated data centre: servers; mechanics; peripherals. Meter: meter itself; modem; remotely operated switch.
   - Align with 1410
   - Make assumption about lifetimes
   - Make a system flow chart
6. Calculate results
   - Multiply with emission factor from relevant electricity mixes (i.e. the country of use; country of production); *toe (diesel)*
   - Compare to baseline: benefit due to less fuel use; cost due to meter usage, data centre, network communications
7. Conclusions
8. References (very important!!!)
   - Statistics
   - Other LCA studies

- Manufacturer

Calculation steps ITU for estimations of products or services based on pilots/experiences in other cities

Example: Smart work for a large company; use of ICT services to reduce need for business travelling and commuting; extrapolate results of a case study to a city

1. Decide on project boundaries
    - City, country
    - Time horizon
2. Decide on goal and scope
    - What do you want to investigate and why?
      Smart work for a large company; use of ICT services to reduce need for business travelling and commuting; extrapolate results of a case study to a city
    - What is included and what is not (cut-offs)?
    - What is your reference situation?
      Situation without the project
    - Prior conditions that may affect the project
      Comparison between contexts: differences in profile data (demographic, ICT maturity factors), parameters that influence the applicability of the case study
3. Decide on type of assessment
    - *Tier 1: use stage (only direct CO_2 emissions)*
4. Description of the project
    - *Reduce energy use/GHG emissions by reducing the need for business travelling and commuting*
5. *Activities include: travelling and commuting, ICT services from home (hardware and software), heating on at home during the day*
6. Estimate your inventory
    - Number of ICT items: hardware, software, networks (i.e. PCs, laptops, tablets, printers, mobile phones, servers, server rooms, base stations for cellular networks, ADSL/Wi-Fi etc.)
      Only existing ICT goods, services and networks were used.
7. Calculate results
    - Multiply with emission factor from relevant electricity mixes (i.e. the country of use); *toe (diesel, gasoline, natural gas)*
    - *Compare to baseline: benefit due to less fuel use; cost due to server usage, data centre, network communications*
      CO_2 reduction from reduction in car commuting per office employee per year: X office workers, of which Y commute by car (note: tendency to e-commute may be higher among those who live far from work) Include in comparison: average time of journey to work; average distance to work; percentage of commuting performed by car; average co2 emission from car (g/km); existence of smart work program y/n; does it mean that the heating is on during the day; does the network need to be expanded due to increase of data traffic; how does that compare to the footprint of transportation? Qualitative analysis of other parameters (importance of parameters e.g. lack of incentives)

8. Conclusions
9. References (very important!!!)
    - Statistics
    - Other LCA studies
    - Manufacturer

Reporting

Deliverables

1. A summarized manual on how to operationalize the indicators of ITU L.1440, to be integrally included in the CITYkeys selection of indicators, for which this document provides a baseline.
2. A full report on the case studies, going into the process of executing the methodology, data availability, the calculated results and following conclusions, and an evaluation on the workability of the methodology.

APPENDIX 3: LIST OF CITY DATA SETS AND THEIR AVAILABILITY

Data set	Rotterdam	Tampere	Vienna	Zagreb	Zaragoza
population data, coordinates	confidential	confidential	public	public	public
basic health care services, coordinates	open data	open data	open data	public	public
number of inhabitants	public	open data	public	public	public
number of traffic accidents	public	open data	public	public	open data
number of violence, annoyance, crimes	open data	public	public	public	public
public transport stops, coordinates	open data	open data	open data	public	open data
Number of vehicles for sharing (cars and bicycles)	no	public	unknown	public	open data
bicycle paths (km)	public	open data	confidential	open data	open data
street network (km)	public	open data	public	public	yes
motorways (km)	public	open data	public	unknown	yes
basic public infrastructure (i.e. services/facilities provided by the town/city council for the general public to use, with or without charge (e.g. community centres, sports grounds, restrooms, drinking fountains), coordinates	open data	confidential	confidential	public	open data
commercial amenities providing goods for daily use, coordinates	open data	confidential	unknown	unknown	yes
Number of fixed broadband subscriptions	no	no	public	public	public
Area covered by public	no	no	public	public	yes

Wi-Fi (km2)					
Total city area (km2)	public	open data	public	public	yes
# of schools with environmental education program	unknown	yes	public	unknown	public
Total # of schools	public	open data	open data	public	yes
# of people reached by digital literacy programs	no	no	unknown	public	unknown
Total target group for digital literacy programs	unknown	no	unknown	public	unknown
# of dwellings per category (categories listed below). Answers on the respective rows below.					
* Detached residential, large (>116 m2)	public	public	public	public	public
* Detached residential, small (≤116 m2)	public	public	public	public	public
* Duplex or townhouse, large (>116 m2)	public	public	public	no	public
* Duplex or townhouse, small (≤116 m2)	public	public	public	no	public
* Dwelling unit in multiunit building with no elevator, large (>116 m2)	public	public	no	no	public
* Dwelling unit in multiunit building with no elevator, medium (>70 to ≤116 m2)	public	public	public	no	public
* Dwelling unit in multiunit building with no elevator, small (≤70 m2)	public	public	no	no	public
* Dwelling unit in multiunit building with elevator, 4 stories or fewer, large (>116 m2)	public	public	no	no	public
* Dwelling unit in multiunit building with elevator, 4 stories or fewer, medium (>70 to ≤116 m2)	public	public	no	no	public
* Dwelling unit in	public	public	no	no	public

multiunit building with elevator, 4 stories or fewer, small (≤70 m2)					
* Dwelling unit in multiunit building with elevator, 5 to 8 stories, large (>116 m2)	public	public	no	no	public
* Dwelling unit in multiunit building with elevator, 5 to 8 stories, medium (>70 to ≤116 m2)	public	public	no	no	public
* Dwelling unit in multiunit building with elevator, 5 to 8 stories, small (≤70 m2)	public	public	no	no	public
* Dwelling unit in multiunit building with elevator, 9 stories or more, large (>116 m2)	public	public	no	no	public
* Dwelling unit in multiunit building with elevator, 9 stories or more, medium (>70 to ≤7116 m2)	public	public	no	no	public
* Dwelling unit in multiunit building with elevator, 9 stories or more, small (≤70 m2)	public	public	no	no	public
* Live-work space, large (>116 m2)	public	no	no	no	public
* Live-work space, small (≤116 m2)	public	no	no	no	public
* Accessory dwelling unit, large (>116 m2)	unknown	no	public	no	public
* Accessory dwelling unit, small (≤116 m2)	unknown	no	public	no	public
Total number of dwellings	public	confidential	open data	public	public
total ground floor space in city (m2)	public	confidential	public	unknown	yes
ground floor space used for commercial or public	public	confidential	confidential	unknown	yes

purposes (m2)					
Public outdoor recreational spaces (m2)	public	open data	public	public	yes
green spaces (m2)	public	open data	public	public	open data
Total final energy consumption in the city per year	public	public	public	confidential	open data
total energy consumption of public buildings (total + electricity, heating, hot water and cooling separately)	no	confidential	public	confidential	public
total energy consumption of all buildings	public	yes	public	public	public
total electricity consumption of buildings	public	yes	public	public	yes
total heating consumption of buildings	no	yes	public	public	yes
total hot water consumption of buildings	no	yes	public	unknown	yes
total cooling consumption of buildings	no	no	public	unknown	yes
total energy consumption for transport	unknown	no	public	confidential	public
total energy consumption by all street lightning	open data	yes	public	confidential	public
total energy used by ICT	unknown	no	public	unknown	no
total consumption of energy generated from renewable sources (total + electricity and heat separately)	public	no	public	confidential	public

CO2 emissions (tonnes/yr)	public	yes	public	unknown	public
Total material consumption(tonnes/yr)	no	no	confidential	unknown	yes
Total water consumption	yes	yes	confidential	public	public
# of houses or households with grey water reuse capability	no	unknown	no	unknown	no
Total water abstraction (m3/yr) within relevant area	unknown	yes	public	unknown	public
Total long-term freshwater resources (m3) within relevant area	unknown	open data	public	unknown	yes
water delivered to billed customers during one year	yes	confidential	public	unknown	public
Food produced in 100 km radius (tonnes)	no	no	public	public	public
Total food demand (tonnes)	no	no	no	unknown	no
Brownfield area redeveloped (m2/yr)	public	no	confidential	confidential	unknown
Total brownfield area in the city (m2)	public	no	confidential	confidential	unknown
Summer temperatures within the city (Celsius degrees)	open data	open data	public	public	no
Summer temperatures outside the city (Celsius degrees)	open data	open data	public	public	public
NO2 emissions (g/yr)	open data	open data	open data	public	open data
PM2.5 emissions (g/yr)	open data	open data	open data	public	open data
Annual average concentrations of data sets listed below. Answers on the respective rows below.					
* NO2	open data	open data	open data	public	open data

* PM10	open data	open data	open data	public	open data
* Daily PM10	open data	open data	open data	public	open data
* Ozone (8h average)	open data	open data	open data	unknown	unknown
* SO2	open data	open data	open data	public	open data
* Benzene	unknown	open data	open data	unknown	unknown
Number of days exceeding average 50 µg/m3. Answers on the respective rows below.					
* NO2	open data	open data	open data	public	open data
* Daily PM10	open data	open data	open data	public	open data
* Ozone (8h average, exceeding 120 µg/m3)	open data	open data	open data	unknown	unknown
Municipal waste recycled (t/yr)	open data	public	public	public	public
Municipal waste produced (t/yr)	open data	public	public	public	public
Green and water surface (m2)	open data	open data	public	public	public
# of endemic species present in the city	unknown	no	public	unknown	public
# of people in labour force not in paid employment or self-employment, but available and seeking for work OR unemployment rate	public	open data	public	public	public
Total # of people in labour force OR unemployment rate	public	open data	public	public	public
Total number of unemployed youth OR youth unemployment rate	public	public	public	public	public
Total youth labour force OR youth unemployment rate	public	no	public	public	public
Modelled fuel costs	confidential	no	public	public	yes

(consumption * price) OR # of citizens in fuel poverty					
Household incomes	public	public	public	yes	public
# of citizens living in affordable housing	public	confidential	confidential	unknown	public
# of companies with ISO 140001 certificate	unknown	no	unknown	unknown	public
# of companies	confidential	public	open data	unknown	yes
Annual public procurement using environmental criteria (expressed both as total number and M EUR)	no	yes	public	unknown	yes
Total annual public procurement (M EUR)	no	yes	confidential	unknown	yes
# of green jobs (related to environmental service activities that contribute substantially to preserving or restoring environmental quality)	no	no	no	unknown	no
Total # of jobs	confidential	public	public	public	yes
Gross domestic product on the level of the city	public	public	public	public	yes
Number of new companies registered per year	confidential	open data	public	unknown	yes
Median disposable annual household income	public	public	public	unknown	yes
# of workers in creative industries	confidential	no	public	unknown	yes
# of people in labour force in paid employment or self-employment	public	public	public	public	yes

# of open information facilities/innovation hubs in the city, whether private or public	unknown	public	no	unknown	unknown
R&D expenditure in the city (€/yr)	no	public	public	unknown	yes
# of open government data sets in the city	no	public	public	unknown	unknown
Travel times by person at different times	unknown	no	yes	unknown	unknown
Travel distances made by vehicle (km)	yes	no	yes	unknown	unknown
Number of trips made with public transport	open data	yes	public	public	public
# of people moving into the city and moving out of the city per year	yes	public	public	public	unknown
# of people in the city under 14 and over 65	public	public	public	public	unknown
# of people in the city between the ages of 15 and 65	public	public	public	public	unknown
# of international events held	unknown	public	public	unknown	unknown
# of tourist nights in the city per year	no	yes	public	public	public
Total amount of open public participation processes	unknown	public	public	unknown	yes
# people who voted in last municipal election	open data	public	open data	unknown	public
# of people entitled to vote	public	public	open data	unknown	yes
Expenditures by the municipality for a transition towards a smart city in €/yr	unknown	no	no	unknown	unknown
	Rotterdam	Tampere	Vienna	Zagreb	Zaragoza

Available	82	89	93	60	95
	71 %	77 %	80 %	52 %	82 %
Public	49	38	64	51	52
	42 %	33 %	55 %	44 %	45 %
Confidential	6	9	9	7	0
	5 %	8 %	8 %	6 %	0 %
Open data	23	29	18	1	15
	20 %	25 %	16 %	1 %	13 %
Not available	18	26	18	18	5
	16 %	22 %	16 %	16 %	4 %
Unknown	16	1	6	38	16
	14 %	1 %	5 %	33 %	14 %

APPENDIX 4: LIST OF PROJECT DATA SETS AND PRELIMINARY SELECTION OF RELEVANCE IN TESTING

Data set	Rotterdam	Tampere	Vienna	Zagreb	Zaragoza
waiting times (to health care), before and after project				x	
# traffic accidents, before and after project				x	
# crime incidents, before and after project				x	
cycling roads (metres), before and after project			x	x	x
# of citizens reached by the project	x	x	x		
# of citizens considered stakeholders in the project	x	x	x		
# of dwellings of a housing category in the project:					
Detached residential, large (>116 m2)					
Detached residential, small (≤116 m2)					
Duplex or townhouse, large (>116 m2)					
Duplex or townhouse, small (≤116 m2)					
Dwelling unit in multiunit building with no elevator, large (>116 m2)					
Dwelling unit in multiunit building with no elevator, medium (>70 to ≤116					

m2)					
Dwelling unit in multiunit building with no elevator, small (≤70 m2)					
Dwelling unit in multiunit building with elevator, 4 stories or fewer, large (>116 m2)					
Dwelling unit in multiunit building with elevator, 4 stories or fewer, medium (>70 to ≤116 m2)					
Dwelling unit in multiunit building with elevator, 4 stories or fewer, small (≤70 m2)					
Dwelling unit in multiunit building with elevator, 5 to 8 stories, large (>116 m2)					
Dwelling unit in multiunit building with elevator, 5 to 8 stories, medium (>70 to ≤116 m2)					
Dwelling unit in multiunit building with elevator, 5 to 8 stories, medium (>70 to ≤116 m2)					
Dwelling unit in multiunit building with elevator, 5 to 8 stories, small (≤70 m2)					
Dwelling unit in multiunit building with elevator, 9 stories or more, large (>116 m2)					
Dwelling unit in multiunit building					

with elevator, 9 stories or more, medium (>70 to ≤7116 m2)					
Dwelling unit in multiunit building with elevator, 9 stories or more, small (≤70 m2)					
Live-work space, large (>116 m2)					
Live-work space, small (≤116 m2)					
Accessory dwelling unit, large (>116 m2)					
Accessory dwelling unit, small (≤116 m2)					
Total # of dwellings in the project			x	x	
# of social dwellings in the project					
Commercially or publicly used ground floor space (m2), before and after the project					
Public outdoor recreational space within 500m (m2), before and after the project					
Green space within 500m (m2), before and after the project					
Total final energy consumption of the project, consisting of:	x				
electricity consumption (and specifically ICT)	x				x
consumption of energy for heating and cooling					
consumption of energy for hot water					
consumption of	x	x	x		x

energy for transport					
floor area buildings (m2)			x		x
Embodied energy of materials in project (J)	x				
Embodied energy of materials in reference case (J)	x				
Total annual renewable energy generation of the project	x		x		x
Direct CO2 emissions, before and after project. See the indicator “Reduction in direct (operational) CO2 emissions achieved by the project” for further guidance on emission factors.	x	x	x		x
Indirect (life cycle) CO2 emissions, before and after project	x				x
Local hourly energy demand, consisting of:			x		
local hourly electricity demand			x		
local hourly heating and cooling demand			x		
local hourly hot water demand					
local hourly transport demand					
Local hourly renewable energy generation capacity, consisting of:					
local hourly renewable heating and cooling generation (for example by means of geothermal			x		

energy or ATES)					
local hourly renewable electricity generation (for example by means of PV and wind turbines)			x		
local hourly storage capacity			x		
Total material use (kg) due to project, actual and reference situation		x			
Recycled and re-used materials used in project (kg)		x			
Rapid-growth renewable materials (e.g. bamboo, cork, straw, cotton insulation, agrifiber, natural linoleum (Marmoleum), wool, wheat board and strawboard) used in project (kg). See the indicator "Share of renewable materials" for further guidance.		x			
Recyclable materials used in project (kg)		x			
Total volume of water consumed (m3), before and after the project			x		
Volume of water consumed (m3) at the city level					
Volume of water consumed that is re-used grey and rain water (m3), before and after the project					
Volume of water consumed from local resources (m3), before and after the project					

Inhabitants or workplaces (#/ha), before and after the project					
Food production (kg) within 100 km radius, before and after the project					
Total food demand (kg) within the project					
NO2 emissions (kg/yr), before and after the project		x	x		x
PM2.5 emissions (kg/yr), before and after the project		x	x		x
Noise levels (dB), before and after the project	x	x	x		
Solid waste collected (kg/yr), before and after the project			x		
Blue and green space (m2), before and after project			x		x
Project costs spent on local suppliers, contractors and service providers (€)	x	x	x		x
Total project costs spent on suppliers, contractors and service providers (€)	x				
Jobs created by the project (#)	x	x	x		
Energy costs (€/yr), before and after the project					
Gross household income (€/yr)					
Fixed housing costs after the project (€/yr)					
Companies involved in the project that have ISO-14001					

certification (#)					
Total companies involved in the project					
Costs for the end user, before and after the project (€/yr)		x			
Project's initial total investments (€)		x	x		
Project's annual total cash inflow (€/yr)		x	x		
Project's annual total cash outflow (€/yr)		x	x		
Payback period of project investment (yr)		x	x		
Subsidies received (€)					
Travel time (h) in peak hours, before and after the project					
Users actively involved in project (#)					
Total # of inhabitants					
# of visitors (offline and online) to project (web) site					

Note: if there is no reference available in terms of a before situation, then provide a business as usual reference.