MONITORING MOVEMENT IN THE SMART CITY:

OPPORTUNITIES AND CHALLENGES OF MEASURING URBAN BUSTLE

N. Walravens[1*], B. Van de Vyvere [2], M. Van Compernolle[3], E. Vlassenroot[3], P. Colpaert [2]

[1] imec–SMIT, Vrije Universiteit Brussel, Brussels, Belgium – (nils.walravens@imec.be)
[2] imec–IDLab, Ghent University, Ghent, Belgium – (brecht.vandevyvere@ugent.be, pieter.colpaert@imec.be)
[3] imec–MICT, Ghent University, Ghent, Belgium – (mathias.vancompernolle@ugent.be, eveline.vlassenroot@ugent.be)

Commission VI

KEY WORDS: Open Data, Semantic Sensor Network, Smart Cities, Urban Bustle

ABSTRACT: One of the promises of the smart city concept is using real-time data to enhance policy making. In practice, such promises can turn out to be either very limited in what is actually possible or quickly trigger dystopian scenarios of tracking and monitoring. Today, many cities around the world already measure forms of urban bustle, i.e. how busy it is during specific periods of time. They do this for all kinds of purposes like optimising mobility flows, attracting tourism, monitoring safety during events or stimulating the local economy, and they employ divergent technologies: from analogue counting, over surveys, to more advanced near real-time tracking using mobile operator data. This fragmentation of approaches to measuring urban bustle creates some challenges for cities related to privacy, vendor lock-in, comparability of data, data quality and accuracy, historical and predictive analysis of data and so on. To tackle these challenges and formulate a standardised approach to measuring urban bustle, the thirteen largest cities in Flanders (Belgium), together with local technology vendors, co-created a "definition manual"; a document outlining indicators and relevant technologies for measuring urban bustle, as well as shared profile descriptions of residents and visitors of the city. This paper outlines the process and presents the results, an agreed-upon framework of standard profiles and indicators, which are useful to academics, public servants and technology companies involved in this topic.

1. INTRODUCTION

1.1 Context: Smart Flanders

For over a decade, city governments have been exploring what it can mean to be a "smart city". Turning the promises of the concept into practice, however, remains a challenge for cities today. Most agree that technology has *some* role to play in supporting or implementing policy, but how that role should be filled remains unclear and often the result of trial and error. The Smart City concept has also been criticized, *inter alia* for its self-congratulatory tendency, the commercial interests at play, as well as its push of ICT and the potential consequences towards reinforcing a digital divide (Graham, 2002; Hollands, 2008). Handing over too much control over the public domain to private companies raises concerns regarding democracy and the commodification of the public space (Greenfield, 2013; Townsend, 2013). Both are a far cry from what would be labelled as 'smart'.

At the same time, city governments are exploring what the concept can actually contribute to their daily practices and which role technology can play in providing better or 'smarter' services to citizens. Even the staunchest critics of the Smart City concept agree that *data* increasingly has a role to play in policy making (Hollands, 2008). While this of course has always been the case to more or lesser extent in policy making, the sheer amount of data that is becoming available today, as well as the combination of data from different sources and domains, can provide new types of tools and insights to policy makers. This can be data that comes from Internet of Things solutions (e.g. sensors in public parking garages), structured information in internal reporting systems, detailed data on the public domain (e.g. from satellite imaging) and so on.

In order to fully unlock the potential of this data however, it needs to be more easily available and accessible than today. In order to tackle some of these data-related challenges, the Smart Flanders program was initiated by the Flemish Government (Belgium) in early 2017. Smart Flanders was coordinated by IMEC, the largest non-profit technology research institute in Belgium, by an interdisciplinary team of researchers from communication sciences, organizational science and computer science. The goal of the 3-year program was to support the thirteen so-called centre cities in Flanders (approximately the largest cities in the region) and a representation of the Flemish Community in the Brussels Region (referred to as the 13+1), with defining and implementing a common open data policy. The program is followed up today by a steering group consisting of representatives of the cities, the cabinets of the Flemish ministers for Urban Policy and for Innovation, the Flemish agencies responsible for Interior Policy and Information, the Knowledge Centre Flemish Cities, and the Organization of Flemish Cities and Towns.

Smart Flanders was conceived as a participatory and an implementation-driven program, wanting to learn by doing and from cities' expertise and experience, identifying challenges and bottlenecks along the way. As a practical approach to this, a number of data pilots were defined in the program. The topics of the data pilots were chosen by the cities and the first theme to work on was real-time data on off-street parking availability. These parking garages are very often run by commercial companies and cities have limited to no insight into how and when these garages are used, as the data is owned by the

---

*Corresponding author

companies and nothing related to open data was included in the concession or contract. In this case, Smart Flanders negotiated with parking companies in an effort to publish the real-time and historical occupancy of the parking garages as linked open data. Trying to publish this data as open data yielded a lot of relevant lessons (Rojas Meléndez, 2018) and identified the need for more standardized clauses in contracts with third party suppliers.

The second data pilot focused on accessibility data; the physical accessibility of public buildings to be more precise. Currently, this data is fragmented and managed by different organizations (by the regional government, cities, non-profits and so on), so structuring, linking and publishing this information poses a significant challenge. After a number of capacity building and knowledge sharing sessions with stakeholders in the field, a trajectory was set up by the government of Flanders to completely redesign the vision and data collection methods for this type of information.

The final pilot, which ran during 2019, centred on a very broad topic: better understanding urban bustle through local data. Today, bustle in the city is measured in a lot of different domains, using different technologies and requiring divergent forms of collaboration with the private sector. The following section will explore some of these challenges.

1.2 Many Ways to Capture Urban Bustle

Aiming to better understand urban bustle at the local level is done for a wide variety of reasons and within different domains. The main domains mentioned by city administrations during the Smart Flanders program are the following:

- Local economy (retail, circulation in the city centre, vacant buildings etc.)
- Mobility (how and when to move from, to and in the city)
- Events (carnivals, fairs, Christmas markets, parades, accounting for marketing spend, etc.)
- Tourism (non-residents visiting the city)
- Culture (museums, concerts, events, etc.)
- Development of new urban areas and the public domain
- Real-time crowd management in the context of security (police)

To measure or assess urban bustle for these types of domains and activities, local governments work together with vendors who apply various - often technological - solutions. Various types of companies are active in this market and they distinguish themselves in size, experience, the technology used, whether or not to combine hardware and software in the offer, focus on a certain indicator (e.g. parking information) and so on. One of the most important factors in how insights can be built up based on measurements in the city is the technology used. Currently, the possibilities (and limitations) of a certain technology are decisive to what extent the captured data can be useful. These are popular as well as less frequently used measurement methods: WiFi sniffers; Infrared distance sensors; Data from mobile operators CRD; Magnetic loops in the ground; Floating car data; Manual surveys and tallying; Beacons; Cameras; Transactions in shops; Mobile applications; Social media data.A few initial elements emerge related to these technological choices. An important advantage is that the data collected by all these measurement methods can be exported as raw data to open formats, such as CSV. This makes them more widely usable within the city organisation, although they often have to be explicitly requested

from the supplier. A second advantage is that the accuracy of the data can be validated or placed in context by comparing the results of different measurement methods.

A number of disadvantages with the current state of affairs revolve around the way in which data is processed and the relationship with suppliers. Today almost every measurement method needs a certain calibration (and often also extrapolation) of the data, which makes it easier to contest the accuracy of the results. For example, there is the market share of a telecom supplier in an area that can play a role in counting visitors with mobile usage. A number of calibration methods can support this, but are themselves potentially prone to incomplete data collection, such as manual counting or a control camera (where human error can be made in both cases).

Furthermore, the reports that are delivered as a result of measurements roughly depend on three components: the supplier, the algorithm and the measurement technology. As a result, historical and baseline measurements are lost when one of these three things changes, which either makes it difficult to change supplier or the historical comparability (and therefore analysis potential) of the data is lost. Finally, new players are entering the market or existing companies are pivoting towards this market. This means that the above-mentioned problem of fragmentation of data models and datasets only increases.

Next to the questions of *why* (the domains) and *how* (the technologies used), *what* is being measured adds additional complexity to this field. A number of cities developed the indicators and profiles to be tracked in consultation with the vendor responsible for the measurements, but these are all tailored to the specific situation of the city. Furthermore, the question of which indicators should be measured informs the choice for a specific technology, as not all technological solutions can measure all indicators to sufficient detail and vice versa.

As a result of all this, performing comparable, trustworthy and privacy-protecting measurements of urban bustle becomes an incredible challenge. In order for cities to collaborate on this topic, it is first important to understand the characteristics of the challenges in detail. Building an inventory of current practices was a first important step to better understanding the status quo and to explore alignment between local governments. As cities shared and discussed these challenges (see the methodology section below), the most pressing issues were detected, which are outlined in the following section.

2. URBAN BUSTLE

2.1 Challenges

As was illustrated above, measuring urban bustle using both low or high tech solutions, leads to a number of challenges. The following points of attention were formulated by the cities in the Smart Flanders programme and pertain to (1) the way measurements are currently taking place and how that impacts future endeavours, (2) the relationship with vendors offering very divergent solutions to some of the questions local governments have, and (3) how to make the data that is gathered useful for policy making. The cities presented their own experiences with measuring urban bustle thus far to each other and agreed on consolidating the challenges into the following three main ones.

2.1.1 Current approaches to measurements

A particular challenge today is how to deal with baseline measurements. When there is a change of vendor for example, the baseline measurement and/or history of the data is lost. This can even happen when the same vendor changes the measurement method. A related problem is the lack of common definitions and indicators as it is not always sufficiently clear what exactly is being measured. Uniform characteristics and definitions are important to achieve this. Cities often work out their own indicators individually: these are determined by what one wants to evaluate/measure and which technologies might be available to perform the measurements. Both policy and technology therefore determine what can be measured and how. Some cities carry out similar measurements, collaborating with the same vendor but on the basis of different indicators, leading to different results.

A second issue here is 'self-selection bias'. One must question whether the group of people detected is representative of the entire population. When a self-selection bias is present, individuals with certain characteristics are overrepresented or underrepresented in the data. This can make it difficult to compare or extrapolate numbers to the entire population. E.g. at an event for people over 65, fewer visitors will probably be traced using WiFi measurements compared to an event with a younger target audience. However, this does not mean fewer visitors were present at the event for 65+ people. The detection ratio for detecting the target group 65+ is likely lower than the one for detecting younger people. The risk of self-selection is strongly related to the technology used (WiFi/Bluetooth vs. cameras and magnetic loops) and contextual factors (e.g. events in an area with free WiFi, elimination of roaming charges in the EU...).

2.1.2 Relationship with vendors

When it comes to interactions with vendors and technology suppliers, cities have very mixed experiences. Promises that suppliers make in quotations in terms of deliverables cannot always be kept. One of the reasons for this is that suppliers do not always have off-the-shelf solutions on a technical level or that their offerings are still in early stages of development. Not all companies have the experience to meet the various expectations of every city. There is a clear difference in start-ups versus more established players.

Solutions are also often tailor-made, which requires an effort from the vendor that does not always outweigh the accuracy gain in the results or cannot meet the expectations of the city. Furthermore, vendors are not always aware of the possible consequences of technical interventions and/or communicate insufficiently or too late with their customers. Changes in the way certain calculations are made, or the addition of a sensor to an area can impact the results and need to be properly communicated with cities. Finally, the actual technical rollout of hardware is often underestimated in practice and takes longer than anticipated.

2.1.3 Making data useable and useful

A question raised by the cities is whether the data that are gathered should be open or shared and to what extent. Often, a political sensitivity is involved related to publishing certain types of data and these should be taken into account. Related to this are of course all matters related to privacy and compliance with the GDPR. WiFi and Bluetooth measurements are currently under pressure for example.

Another challenge comes from working with the data generated by solutions and their interoperability: data from different vendors can often not (easily) be combined. This can even be the case for data coming from the same type of solution, for example telecom providers can have a different topology for their cell towers which makes it difficult to combine mobile data from multiple telecom providers. As a result, potential insights from the data (e.g. causality between parameters) are lost. Often there is also a lack of capacity within the city administration to carry out its own analyses, and very often this is not worth the time or effort, for the reasons given earlier.

Finally, a lot of assumptions and/or extrapolations have to be made before a certain insight can be gleaned from the data. This makes it very difficult to draw concrete conclusions from the data. In general, there is little confidence in the figures, which makes their usability for policy problematic and quite expensive.

Based on these challenges formulated by the cities involved in the programme, it became clear there was a need for uniform indicators for measuring urban bustle, as well as shared profile descriptions of residents and visitors of the city. The following section outlines how these tools were developed.

2.2 Methodology

In order to come up with uniform profiles and indicators of urban bustle that could be used across cities and irrespective of technology, the cities and vendors needed to agree on a common language to describe them. To facilitate this, a total of 8 participatory workshops was organised during the course of 2019. The sessions were organised in the framework of the Smart Flanders programme and moderated by imec researchers. The attendees were representatives from the 13 centre cities with various backgrounds (e.g. from mobility, GIS, economy, tourism, IT departments and several others), but were all in some way involved in tracking and measuring urban bustle.

The first few sessions focused on knowledge sharing, as local representatives presented their current practices and experiences with technologies and suppliers to each other. This led to a first inventory, exposing the wide diversity of approaches, goals, indicators and technologies used to measure urban bustle. From this inventory, a first set of shared challenges was identified (as presented in the previous section), which were then discussed with technology vendors active in the field. During a half-day workshop, the cities presented these challenges to the vendors and discussed potential solutions and limitations in break-out sessions facilitated by the imec researchers. This led to two main conclusions: (1) the interest and need for vendors to align more amongst themselves when it comes to the technological solutions they offer and (2) the need for what was dubbed a "definition manual" that would allow cities to use uniform terms, profiles and indicators to describe the precise aspects of urban bustle they want to measure. The latter point was taken up to be developed further in the Smart Flanders network, with input from the vendors, to ensure the two parties would remain aligned, increasing the potential impact of the final product.

The remaining workshops then focused on the co-creative writing of this definition manual that outlines uniform profiles and indicators that cities would be able to use when asking the market for solutions related to measuring urban bustle. In order to ensure all parties involved spoke the same language, the choice was made to use an internationally standardised ontology to describe

the different aspects of the measurements. The choice quickly fell on using the Semantic Sensor Network ontology to structure the definition manual.

2.3 Semantic Sensor Network

The Semantic Sensor Network (SSN) ontology is an international standard endorsed by the World Wide Web Consortium (W3C) since 2017 and allows describing sensor data in all its facets[1]. Not only measurements of a sensor can be described (e.g. it is 30° in this building according to a thermometer), but it can also be used to describe software algorithms. The latter is used to answer more complex questions, such as "Are people going for a drink in the centre after the market took place?". These kinds of complex questions require a combination of sensor data. The cities from the Smart Flanders program already indicated that there is little confidence in the figures, because the way in which something is calculated is often not transparent (section 2.1.1). This is exactly what the SSN ontology also tackles: describing sensors (physical devices or software algorithms), what is observed, which measurement procedure is used, etc. In the next section we will discuss the concepts we reuse from SSN. Thereafter, we will describe urban bustle profiles, which are based on these concepts, that cities can implement to agree with vendors upon uniform descriptions of residents and visitors of the city.

An *Observation* is the central concept that bundles all aspects of a sensor measurement or software calculation. It is linked to a *Result* that contains the value (e.g. 5) and the time it holds true (e.g. between 1 pm and 2 pm). An observation is generated by a *Sensor*, which can be a device on the street or a software program. A *Procedure* allows to describe which steps were followed to obtain the observation. This can be used to communicate the extrapolation and calibration procedures (section 1.2) that were performed by the vendor and thus increase the confidence in numbers and figures. SSN also proposes referring to the data from which a new observation is made through the *Input* concept. This way, observations can be layered: raw observations from multiple sensors can be used as input for observations that indicate the general urban bustle on a city scale.

There are two concepts to specify the result of the phenomenon that has been observed: the *Feature Of Interest* (FoI) and the *Observable Property (OP)*. The FoI is the thing that is observed while the OP specifies the property of the FoI that is measured. For example, when the number of people is counted at the beginning of a shopping street, the FoI indicates the beginning of the shopping street's location and the OP refers to the property "number of people passing by". In the case of urban bustle, the FoI concerns almost exclusively a physical zone. In our profiles, Feature of Interest is therefore translated as Focus area.

Fig. 1 demonstrates an observation that states that 50.000 visitors are detected in a city (focus area). This phenomenon happened between 10am and 10.30am on 8th December 2019 and was made by applying a big data algorithm on the mobile phone location data of a network provider. The observation could link to an intermediate anonymized dataset of origin destination user flows that has been used but could also be a description without much detail. Cities can agree with a vendor which procedure will be performed to obtain observations. This offers a way to improve the relationship with suppliers how data is processed (Section 1.2).
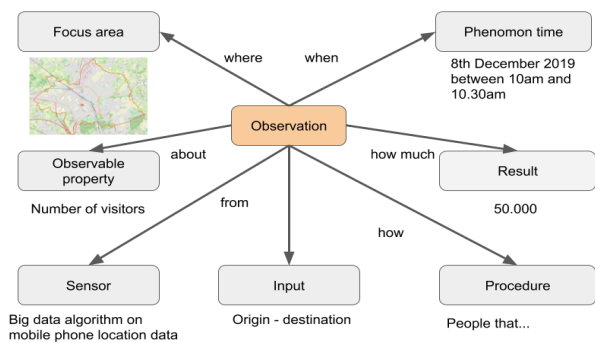
Fig. 1: Example observation related to urban bustle expressed with SSN concepts.

In short, SSN makes it possible to distinguish the meaning of what, how and when is being observed. Moreover, this way of structuring makes it possible to compare different suppliers and cities. The city of Leeds (United Kingdom), for example, applies this for publishing the consumption (gas and electricity) of its public buildings (city hall, libraries, museums etc.) (Radulovic & Garcia-Castro). Or the city of Madrid (Spain) uses the same standard to model ticket validations on the metro (Rojas et al., 2018). In the next section, we will apply this set of SSN concepts for a profile-based approach to observe urban bustle in the city.

3. PROFILES AND INDICATORS

3.1 Profiles

Each city administration needs insights for a certain domain, such as the number of day visitors for tourism (section 1.2). A profile allows cities together with vendors to define how this need will be captured by using the SSN observation structure. The working group of Smart Flanders decided to create 6 *basic* profiles that are applicable over all domains and can be shared between cities. The first profile are residents of the city. Four profiles are linked to visitors and can be categorized in a matrix (Table 1). The columns indicate if the visitor is staying for one day or stays overnight and the rows indicate whether the visits happen on a regular basis or randomly. The sixth and final profile consists of observations of passersby that are made, but who do not linger or stay in the city: these are referred to as "in transit".

	Day visitor	**Staying visitor**
Structural	*Commuting students, commuting employees...*	*Students staying in dormitories, seasonal workers...*
Random	*Museum visitors, shoppers, concert goers...*	*Tourists staying multiple days...*

Table 1. Visitors are categorised based on how long they are 'seen' and how regularly this happens.

Determining whether a visitor is a day or stay-over visitor on a certain day can only be done after measurement data is available

from at least one day thereafter. In the elaboration of the basic profiles, we therefore assume that historical data will be used and analysed.

More specific questions can be answered by combining these basic profiles with activities, such as working, studying, shopping, tourism, transport, culture and events. For example, to determine the number of students that went shopping at the market, we also need to describe the procedure for determining shopping behaviour and look at both the profile of structural day visitors for commuting students and the residents profile for boarding students.

To apply SSN in a non-technical way, we use a scheme that fills in the different concepts from the previous section (cf. Semantic Sensor Network), allowing cities to use these basic profiles to formulate their policy questions in a more uniform way. We give an overview of these basic profiles in what follows.

3.1.1 Resident

To determine what a resident is, we use the model "Persoon basis", developed during the Open Standards for Linked Organisations program (OSLO) (Buyle, 2016). This states that a person can have a residency (permanent or temporary resident) up to a certain jurisdiction. It is crucial that a residence can be linked to this residency. In the case of a student in a student room, this person therefore has 2 residences. These are people who have a right of residence, which is in principle only granted for a very specific reason (study, work, etc.) and who therefore stay in the city for a longer period of time. This "fixed" residence is often calculated using a "most likely living place" algorithm (De Meersman et al.). A resident is described with the SSN structure in the table below.

Focus area	Geospatial area (e.g. the city boundaries)
Observable property	The total number of inhabitants. This can be either a permanent or temporary resident who is domiciled or not at his/her place of residence.
Procedure - description	- The number of permanent residents can be determined on the basis of Open Data (Statbel[2] / city monitor[3]). This outputs the number of domiciled persons in the city. - To determine the number of temporary residents, mobile data or Open Data (e.g. number of enrolled students) can be used. For this we refer to "structural visitors".
Procedure - extrapolation	N/A for permanent residents.
Input	Place of residence, pattern Datasets:

	- city monitor - number of structural visitors
Sensor	N/A
Result	Number (e.g. 300,000)
Phenomenon time	Per year (e.g. 2019)

Table 2. Description of the profile "Resident".

The scheme mentioned above only gives an initial step to observe residents. It's important that cities and vendors complete all the fields, except the Observable property field so cities can compare the same property. For example, the focus area can be defined more accurately using standard formats for describing geographical features (GeoJSON, WKT literal etc.). The Procedure of profiles is split into two parts: first, it has a description of the datasets that are used and secondly, the extrapolation measurements of the vendor can be clearly described. Input mentions indicators (e.g. place of residence, more about this in section 3.2) and datasets that can be used. The Sensor field allows to list the physical sensors or describe the used algorithms. If cities have a Internet-of-Things (IoT) registry available, then they can look up a certain sensor and mention its identifier in the Sensor field. At last, the Result and the Phenomenon time respectively indicate the expected output format and time span.

3.1.2 Structural day visitor

A structural day visitor are people who regularly visit the city, but do not stay for a long period of time.

Focus area	The city.
Observable property	Number of people who regularly visit the city during the day from a functional point of view (e.g. work (commuter), study (commuter), hobby, family visits)
Procedure - description	- They don't comply with the resident profile. - These visitors didn't spend the night in the city. - The visitor must have been detected during the day: between 7 a.m. and 11 p.m. - The duration of a visit must have lasted at least 1 hour, but not longer than 24 hours. - If the previous or subsequent day was also spent in the city (between 8 pm and midnight and 5 am and 10 am), this is considered as a residential visit. For example: ● Resident visitor (arrival on day 1)

---

[2] Statbel, the Belgian statistical office
https://statbel.fgov.be/nl/open-data

[3] https://gemeente-en-stadsmonitor.vlaanderen.be/naar-de-cijfers

	● Resident visitor (day 2) - The visits must comply with a certain pattern, for example at least once a week (frequency is weekly and number is at least 1).
Input	Duration, pattern, place of residence, origin/destination (mobile data)
Result	Number (e.g. 50,000)
Phenomenon time	Per day (e.g. 1/1/2019)

Table 3. Description of the profile "Structural day visitor".

3.1.3 Random day visitor

Random day visitors visit the city, but not on a regular basis, for example to sightsee, have a meeting, visit a museum and so on.

Focus area	The city.
Observable property	Number of people who visit the city, but do not do so regularly. This may be from a functional point of view (e.g. for a work meeting, sports competition), but also from a non-functional point of view (e.g. tourism, catering).

Procedure - description	- They are not residents. This means that these visitors do not have a fixed place of residence in the city.
	- The visitor must be detected during the day: between 7 a.m. and 11 p.m.
	- The duration of a visit must have lasted at least 1 hour, but also no longer than 14h.
	- If the previous or subsequent day was also spent in the city (between 8 pm and midnight and 5 am and 10 am), this is considered as a residential visit. For example:
	  ● Resident visitor (arrival on day 1)
	  ● Resident visitor (day 2)
	- The visits are irregular. In other words, there is no pattern.

Input	Duration, place of residence, origin/destination (mobile data)
Result	Number (e.g. 50,000)
Phenomenon time	Per day (e.g. 1/1/2019)

Table 4. Description of the profile "Random day visitor".

3.1.4 Structural staying visitor

Structural staying visitors regularly stay over in the city.

Focus area	The city.
Observable property	Number of people who regularly stay in the city at night. This can be functional (e.g. people doing a night shift) or non-functional (e.g. people with a second home).

Procedure - description	- They are not permanent residents. This means that these visitors stay in the city for a short or temporary period.
	- Please note that students who stay in a student accommodation are both temporary residents and structural visitors. If the number of students is known, it can be deducted from the number of structural visitors. For some services it is not relevant to count students.
	- The visitor must be detected at night (between 8 pm - midnight and 5 am - 10 am).
	- If the visitor is detected for several consecutive days, then the visitor is a resident over all these days. In that case, a visitor cannot be counted as both a day visitor and a day visitor. For example: someone stays in the city for 3 days:
	  ● Resident visitor (arrival on day 1)
	  ● Resident visitor (day 2)
	  ● Visitor (departure day 3)
	- The visits must comply with a certain pattern, for example at least once a week (frequency is weekly and number is at least 1).

Input	Duration, pattern, place of residence, origin/destination (mobile data)
Result	Number (e.g. 50,000)
Phenomenon time	Per day (e.g. 1/1/2019)

Table 5. Description of the profile "Structural staying visitor".

3.1.5 Random staying visitor

Random staying visitors stay in the city overnight, but not on a regular basis. They could for example be tourists on a multi-day city trip.

Focus area	A specific area (e.g. the city or the historical centre).
Observable property	Number of persons who stay overnight in the area, but do not do so regularly. This could be tourists, for example.
Procedure - description	Here, the same rules apply as the procedure of Structural Visitor, except

	that the visits are irregular (no pattern can be recognized).
Input	Duration, place of residence, origin/destination (mobile dates)
Result	Number (e.g. 50,000)
Phenomenon time	Per day (e.g. 1/1/2019)

Table 6. Description of the profile "Random staying visitor".

3.1.6 In transit

The sixth and final profile consists of observations of passersby that are made, but who do not linger or stay in the city: these are referred to as "in transit".

Focus area	The city and/or approach roads and ring roads
Observable property	Number of people passing through in the focus area. This could be car passengers, for example, but also train passengers.
Procedure - description	- They are not residents. - The person has been detected in the focus area for less than an hour. This person has not been detected long enough to be a visitor.
Input	Duration, origin/destination
Result	Number (e.g. 100)
Phenomenon time	Half-hourly. e.g. between 10:00 and 10:30 on 10/2/2019

Table 7. Description of the profile "In transit".

These six profiles can be used in a multitude of scenarios as a starting point: additional indicators can be added to further refine a profile, but these should always be matched with technological limitations, privacy considerations and other regulations. The definition manual also describes an example set of indicators that can be used to compose and further refine the profiles outlined above.

3.2 Indicators

An indicator represents a certain observation (e.g. there were 5 cars in the shopping street at 13h35) in an objective way. These are all things that can be measured without output from the cities (using a profile). An indicator therefore depends on the sensors and datasets that a supplier has at his disposal. It may be that a supplier has to do certain internal processing to get the indicator values from raw (sensor) data. For example: a bicycle-counting loop in the ground actually measures air pulses and still has to be processed (e.g. remove noise from playing children, reflections, simultaneity... eliminate) to calculate how many cyclists have passed by in total. In case a supplier does not have a complete picture in an area (e.g. due to market share) he has to extrapolate or in other words multiply by a certain factor when processing the raw data in order to approximate the complete picture.

Below is an overview of several measurable indicators. This list was drawn up in collaboration with the centre cities: an inventory was made of how measurements are made today and which indicators are used. The indicators that emerged were structured and adopted in the overview below. Subsequently, they were supplemented with indicators resulting from the questions present with the cities.

In the definition manual, an example is given for each of these indicators, which technology or measurement method can be used to measure them and what possible challenges there are in this way of measuring. This would lead us too far in the context of this paper, but we provide one as an example. The indicators, broken down in five main categories, are:

- **Time**: duration, interval, pattern
- **Location**: point, area, origin, destination
- **Demographics**: place of residence, age, social class
- **Means of transportation**: pedestrian, bus/tram, bicycle, car, train, number of road users
- **Context**: events, tickets sales, transaction data, interactions on digital media, speed, parking occupancy, traffic density…

Each of these 23 indicators is defined and described in the manual. By way of example, the indicator "duration" is provided below:

Definition: The duration of the period in which an observation is made. This can be expressed in minutes, hours or days.
Example: At least 60 min, at least 30 min and at most 120 min, present for 24 hours, present for more than 5 days.
Technology: Wi-Fi/Bluetooth tracking, mobile operator data, camera.

Generic indicators like the one referenced here can be complemented by more specific indicators, to build up a particular profile the city would like to identify - always respecting privacy and other existing legislation.

4. DISCUSSION AND OUTLOOK

One of the key concerns in engaging in these types of measuring activities is indeed privacy. Given the multitude of possible profiles, scenarios and purposes that may be involved in urban bustle measurements, it is not possible to make a comprehensive or conclusive assessment of the impact on privacy in general terms: after all, a privacy impact assessment can only be carried out for a specific case in which the actor processing the personal data can be identified. When this use case is known, it is strongly recommended to carry out a privacy impact assessment for every separate monitoring activity. As a general remark, it should not be possible to follow an individual, which means that unique identifiers such as e.g., smartphones, license plates, MAC addresses of digital devices and so on, cannot be used under the GDPR. It is strongly recommended to write a clear data processing agreement with suppliers who collect data on the territory of the local authority, establishing who is responsible, and where and for which purpose personal data is processed.

In the future, it will be necessary to investigate how measurements from heterogeneous sensors can be combined. First steps have already been undertaken[4], but more work needs to be done on standardizing common data models. This could not only lower the data heterogeneity between telecom providers (because of different cell tower topologies), but could also make local measurements (Wi-Fi, Bluetooth etc.) complementary with these more high-level measurements. For the time being, the basic profiles proposed here can be drawn up almost exclusively on the basis of mobile operator data, but there is a need for more local measurements in order to draw up more detailed profiles. The framework presented in this paper may provide part of the answer to pursue and stimulate cooperation between different suppliers.

Finally, the value of a participatory approach to this type of trajectory is perhaps the most important take-away. When cities feel urgency around a common challenge, there is extremely high value in brining their experience together, discussing it in detail, consolidating on the main challenges and presenting these to the market. Such a participatory approach leads to solutions that are not only valuable, but also more sustainable.

5. CONCLUSION

This paper describes the creation of a definition manual for cities, a document with the aim of creating a common framework of agreements on the construction of profiles that can be used in the context of measuring urban bustle. The framework and the basic profiles that are provided here can be used by (local) authorities when tendering certain solutions to third parties.

Six basic profiles are defined, to which activities or certain characteristics can then be assigned. By using the same basic profiles across different authorities and linking them to existing standards such as OSLO and SSN, market parties can be approached in a more uniform way to achieve more comparable and reliable results.

Going forward, and as the amount of data gathered on urban bustle and life in the city, privacy will become an ever-increasing point of attention. Local authorities and the suppliers they contract will need to take privacy impact assessments into account in order to leverage the insights provided by technological solutions, while ensuring that citizens' information is secure.

Currently, the definition manual is being used by a number of cities in Flanders, when they negotiate with technology suppliers about these types of solutions. More time is needed to evaluate the success of the use of these preliminary standard formulations, but initial signs of this manual being a practical tool have been illustrated by its use in a few contract negotiations and as an inspiration in other, international projects.

ACKNOWLEDGEMENTS

---

4 https://www.imeccityofthings.be/nl/projecten/cityflows

REFERENCES

Radulovic, F. & Garcia-Castro, R. Leeds city council energy consumption ontology: Available online: http://smartcity.linkeddata.es/lcc/ontology/EnergyConsumption-content/index.html

Buyle, R., De Vocht, L., Van Compernolle, M., De Paepe, D., Verborgh, R., Vanlishout, Z., De Vidts, B., Mechant, P. & Mannens, E. (2016). OSLO: Open Standards for Linked Organisations. *Association for Computing Machinery (ACM)*, pp. 126-134. Available online: https://biblio.ugent.be/publication/8504205/file/8504206

Graham, S., 2002. Bridging Urban Digital Divides: Urban Polarisation and Information and Communication Technologies. *Urban Studies, 39*(1), pp. 33–56.

Greenfield, A., 2013. The City is Here for You to Use. Wired, 5 February.

Hollands, R., 2008. Will the Real Smart City Please Stand Up?. *City, 12*(3), pp. 303-320.

Townsend, A., 2013. Smart Cities. New York: Norton & Company.

Van Compernolle, M., Waeben, J. & N. Walravens, 2018. *Eindrapport Smart Portrait*. Public Report for Kenniscentrum Vlaamse Steden and Agentschap Binnenlands Bestuur. Available online:
http://www.kenniscentrumvlaamsesteden.be/overhetkenniscentrum/Documents/Eindrapport%20Smart%20Portrait_PUBLIEK%20-%20definitief.pdf

Walravens, N., Van Compernolle, M., Colpaert, P., Ballon, P., & De Marez, L. (2019). Open Data and the Core Competences of Government: Lessons from Flanders, Belgium. In *Proceedings of the Research Conference on Communications, Information and Internet Policy (TPRC)*.

Rojas Meléndez, J., Van de Vyvere, B., Gevaert, A., Taelman, R., Colpaert, P., & Verborgh, R. (2018). A preliminary Open Data publishing strategy for real-time data in Flanders. In *Proceedings of the Web Stream Processing Workshop (pp. 1847–1853)*.

De Meersman, F., Seynaeve, G., Debusschere, M., Lusyne, P., Dewitte, P., Baeyens, Y., Wirthmann, A., Demunter C., Reis, F. and Reuter, H.I. (2016). Assessing the Quality of Mobile Phone Data as a Source of Statistics. Paper presented at the European Conference on Quality in Official Statistics, Madrid, Spain, May 31 -- June 3. Available online: https://ec.europa.eu/eurostat/cros/system/files/assessing_the_quality_of_mobile_phone_data_as_a_source_of_statistics_q2016.pdf. Accessed 13 March 2017.

Revised July 2020