

Association for Information Systems

AIS Electronic Library (AISeL)

Wirtschaftsinformatik 2021 Proceedings

Track 8: Smart City & E-Government

How Could Smart Cities Use Data? – Towards a Taxonomy of Data-Driven Smart City Projects

Babett Kühne
Universität Hamburg

Kai Heidel
Universität Hamburg

Follow this and additional works at: <https://aisel.aisnet.org/wi2021>

Kühne, Babett and Heidel, Kai, "How Could Smart Cities Use Data? – Towards a Taxonomy of Data-Driven Smart City Projects" (2021). *Wirtschaftsinformatik 2021 Proceedings*. 1.
<https://aisel.aisnet.org/wi2021/SSmartCity/Track08/1>

This material is brought to you by the Wirtschaftsinformatik at AIS Electronic Library (AISeL). It has been accepted for inclusion in Wirtschaftsinformatik 2021 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

How Could Smart Cities Use Data? – Towards a Taxonomy of Data-Driven Smart City Projects

Babett Kühne¹, Kai Heidel¹

¹ University of Hamburg, Department of Informatics, Hamburg, Germany
{babett.kuehne,kai.heidel}@uni-hamburg.de

Abstract. The process of urbanization has caused a huge growth in cities all over the world. This development makes the organization and infrastructure of an individual city increasingly important. In this context, the idea of a smart city is growing and smart city projects are beginning to appear. As the amount of data is growing with connected technologies, such projects rely on data as a key resource. However, current research does not provide an overview on these projects and which constructs are involved in data-driven smart city projects. Therefore, this research begins the building of a taxonomy on such projects through the establishment of a common language among researchers in this new field through eleven dimensions. Additionally, it develops a concrete conceptualization of data-driven smart city projects for practitioners as an initial guidance for the field of smart cities.

Keywords: Smart Cities, Taxonomy, Data-Driven Smart City, Smart City Projects

1 Introduction

A study from the United Nations Organization shows that 64,8 percent of the world population will live in urban regions by 2050 [1]. This lead to a 13.1 percent increase in the population of urban regions as compared to 2018 [1]. Most of these projected increases will take place on the African and Asian continents while a growth of 9.8 percent is expected in Europe [1]. Comparing the urbanization rates in Europe show huge growth in the coming years: 0.22 percent from 1990-2018, 0.33 percent from 2018-2030, and 0.38 percent from 2030-2050 [1]. It is clear that urbanization process will challenge cities all over the world.

This urbanization could be seen as the biggest change in the infrastructure and rebuilding processes of humanity [2]. Furthermore, the daily routine in an urban city becoming more digital and more increasingly inclusive of factors such as smartphones or sensors [3]. This digitalization could help to solve the challenges that are occurring due to urban growth. In doing so, information and communication technology (ICT) could be used to improve living conditions and quality of life as well as enable environmental protection [4]. Such ICT enables the city to become a smart city and improve the lives of citizens with the help of technology [5].

In order to use the data of ICT and inventory data in smart cities, data needs to be acquired, stored somewhere in some data structure, analyzed, and proceeded. Using these data analysis processes could generate new insights which could provide a new value proposition [6] for citizens. The usage of data in such a context is called data-driven [7]. Such data-driven smart city projects could be in the areas of industry, tourism, logistics, buildings, public transport and many others [8]. This research focuses on data-driven smart city projects that contain one use case in order to build a taxonomy.

As there are so many different use cases for data-driven smart city projects, current research does not provide an overview of the possibilities in this field. Perboli et al. provide three dimensions for a smart city: description, business model, and purpose [9]. These three dimensions are focused on a private versus public view. Furthermore, the taxonomy does not focus the data-driven aspect of a smart city and only differentiates between potential tools of a smart city. Niaros provide a taxonomy of smart cities in two dimensions: local versus global and capital versus commons [10]. This taxonomy is focusing on a more strategic view of a smart city project and does not integrate all possible smart city components. Additionally, the data-driven aspect is missing in this research. Thus, to the best of the current knowledge, there is no taxonomy of data-driven smart city projects available in research yet. However, the development of such data-driven smart city projects is growing but the theoretical understanding of such projects is lacking in knowledge, yet. In order to understand such projects more detailed and create methods and tools for the development of such initiatives, the knowledge of the constructs and characteristics of data-driven smart city projects needs to be increased. As a consequence, this research focuses on developing such a taxonomy by using the methodology of Nickerson et al. [11] with the goal of answering the following research question:

RQ: What are the empirically validated and conceptually grounded characteristics that describe data-driven smart city projects?

This taxonomy participates to the present body of knowledge in the field of smart cities by establishing a joint understanding of data-driven smart city projects. Such a shared knowledge contributes to the structuring of this field of research by supporting researchers as they position their work in this field. Furthermore, the shared understanding resulting from this taxonomy allows for the materialization of ideas and considerations that will lead to the development of design theories in the field of smart cities. However, the taxonomy could also support practitioners and offer initial guidance for assessing the chances and opportunities of a data-driven smart city project in order to analyze how this project could be implemented.

In order to do so, the paper is structured as follows. First, it provides an overview of the related research of smart cities. Second, the general approach of developing the taxonomy is described. Third, the development stages of the taxonomy for data-driven smart cities are presented, a detailed overview of the taxonomy itself is provided, and the evaluation is described. The paper closes by discussing the implications of the research, reflecting its limitations and describing possible next steps

2 Related Work

2.1 Smart Cities

The study of smart cities (SCs) is a new field in research. Several cities have started projects with the goal of improving their citizens' quality of life and rely on ICT as key drivers [12], and so data from this technology plays an important role in this context.

Due to contextual changes in the public sector, an unbundling of services from production processes appeared [13], cities are experiencing a shift in value creation from offering products to providing services to citizens [14]. SC services are therefore predominantly designed from the service-dominant logic's perspective [14, 15]. Another contextual change has occurred due to the increasing presence of digital technologies [13] and connection of cities [16]. An exponential growth in data, leveraged by connected technologies such as Internet of Things (IoT), currently characterizes cities. The data that these connected technologies produce, as well as that gathered from citizens' interaction with the connected technologies, can be used to stimulate innovation and to develop new projects aimed at contributing to the citizens quality of life [14, 17].

In general, the definition of a SC is inherently connected to the idea of a digital city (DC) [4]. These two are often used as synonyms although there is a difference between the two terms [5]. The idea of a DC developed during the 1900s. By contrast, SC started appearing regularly in research in the 2010s [5]. Both terms address the citizen and have the following goals regarding improving quality of life: improving electronic services, promoting social inclusion, supporting economic and political efficiency, and facilitating urban development [5]. However, there is a key difference between a SC and DC. A SC is limited by the city boundaries. By contrast, a DC is not limited by such boundaries and has only virtual boundaries [4]. In terms of infrastructure, a DC is only represented by its ICT while a SC includes all infrastructure, such as, streets, buildings, railways, and ICT. The idea of a citizen is also different in these two concepts. Each citizen of the city can profit from and enable services in a SC, even if they are not able to use ICT. In a DC, citizens are enablers and receivers that can profit from the DC only if they are able to use ICT. In summary, SC and DC differentiate from each other in some parts but share the goal of improving quality of life for citizens [5].

As the research field of SCs is relatively new, there is no definition that is well accepted in research yet [5]. All in all, a SC is characterized by its intercultural and social capital, the citizens' government, the smart strategic planning of the city, and ICT [5, 18]. The component 'smart' in this context can be defined in a city that is innovative, integrative, connected, efficient, effective, adaptive and attractive [19].

2.2 Dimensions of Smart Cities

Current research show dimensions of SCs from different viewpoints. As the research goal of this study is to develop a taxonomy, this section introduces all existing taxonomies and dimensions of SCs in detail.

In order to analyze which data is being used in a SC, we take a look onto the different applications in a SC. Lombardi et al. developed a schema to model the performance of a SC. After an extensive literature review, they identified the following dimensions of a SC: Smart Economy, Smart Mobility, Smart Environment, Smart People, Smart Living, and Smart Governance. The dimension Smart People is seen in a demographical context [20]. Other research takes a private economy perspective and developed the following dimensions: Smart Building, Smart Mobility, Smart Energy, Smart Infrastructure, Smart Technology, Smart Governance, Smart Citizen, and Smart Healthcare. The dimension Smart Citizen is seen in a demographical context, for example, usage of green mobility [21]. Lim and Maglio build the following dimensions based on 1234 news articles: Smart Device, Smart Environment, Smart Home, Smart Energy, Smart Building, Smart Transportation, Smart Logistics, Smart Farming, Smart Security, Smart Health, Smart Hospitality, and Smart Education [22]. They also introduce a hierarchical structure, meaning that the customer, provider or things could provide services like Smart Logistics through a linkage to Smart Devices and Smart Environments [22]. They provide five principles of smart services that a SC should have: (1) connection between humans and things, (2) processing of the data in the cloud, (3) wireless communication, (4) collection of data with context awareness, and (5) co-creation of value. If the first four principles are met, a co-creation of value between the customer and provider is possible [22]. All in all, these authors show different dimensions of SCs that somewhat overlap; Smart Mobility and Smart Transportation, for example, could be the same. However, these dimensions give the first indication of the context in which smart services in SCs appear.

In addition to these dimensions, Bischof et al. analyzed which data could be used in a SC. They structured the data according to the update frequency: static, semi-dynamic, and dynamic. Static data is leveraged at one time and could only be updated manually. Semi-periodic data is updated periodically. Dynamic data is updated all the time (every time new data appears) [23]. They developed the following data categories for SCs: transport, air quality, traffic, events of the city, services of the city, citizen data, and health data. The data is generated from different sources, which offers a challenge due to the homogenization of data. Such a homogenization is necessary in order to utilize the whole potential of a SC [23], but a centralization of the data could be helpful in order to make data usage easily assessable [24].

In order to identify all existing taxonomies in the field of SCs, a structured literature review is performed here [25] by searching for the combination of “taxonomy” and “Smart City” in different databases such as Science Direct, IEEE, ACM, and Google Scholar. After analyzing all titles, analyzing abstracts if necessary, and conducting a backward and forward search [25], two taxonomies were identified in the field of SCs. The selection criteria for identifying relevant papers were: (1) the paper offers dimensions for smart city projects and (2) these dimensions are specified by characteristics for smart city projects.

The first taxonomy shows three dimensions of a SC [9]. The dimensions are split into categories and these categories into major fields. According to Nickerson a taxonomy has only the elements, dimension, and characteristics [11]. The dimension ‘description’ in this taxonomy describes the main project features of the SC (context

and components) divided into objectives, tools, project imitators, and stakeholders. The dimension ‘business model’ shows the actions that need to be performed to introduce a new business model into an SC. This dimension has the following categories: management, infrastructure financing, and financial resources. The last dimension ‘purpose’ indicates the final goal of a SC and is characterized by the categories: client, product, and geographical target [9]. Besides the category ‘tools’, no data-driven specific characteristics were indicated in this taxonomy. The components are not very detailed, including things such as ‘data base’. Thus, the origin of the data is not clear; there were other data specific characteristics [23] that also could not be identified in this taxonomy. Thus, we argue that this taxonomy could give an overview of SC initiatives but does not meet data-driven goals that are the focus of this study.

The second taxonomy [10] involves differentiating the SC projects according to local versus global and capital versus commons projects. The projects could therefore be identified in four quadrants: corporate SC, commons-based SC, sponsored SC, and resilient SC [10]. As a consequence, the taxonomy is puts SC projects into these categories, which does not reflect the data-driven aspect of a SC. Thus, it does not answer the research question of this study.

3 Methodology

This study aimed to show the empirically validated and conceptually grounded characteristics that describe data-driven smart city projects and develop a taxonomy. In terms of taxonomy development, the methods of Nickerson et al. [11] were followed and their approach was adapted to this research context. This methodology seemed to be appropriate this study’s purpose as several information systems studies have successfully used this method in different study contexts [26-28], suggesting its robustness in developing taxonomies. The evaluation illustrates use cases, a common method for evaluating taxonomies [29].

The taxonomy development method suggested by Nickerson et al. [11] constitutes an iterative approach which allows researchers to build taxonomies conceptually, based on literature, and empirically.

In order to build a taxonomy using the method of Nickerson et al. [11], the following steps need to be performed: (1) Meta-characteristics that all dimensions and characteristics following in the methodology will be a logical consequence of are defined by the researcher. (2) The researcher defines ending conditions that need to be fulfilled entirely for the taxonomy development process to end. These conditions could be objective or subjective. An objective criterion is characterized by the condition that each dimension of the taxonomy contain characteristics that are exclusive and complete. The dimension also needs to be unique but in the last iteration it is not possible to split or summarize objects and characteristics of the taxonomy. Additionally, one object needs to be assigned to a dimension and the last iteration should not add any dimensions or characteristics. The subjective conditions to end the process are achieved if the taxonomy is succinct, robust, complete, expendable and explainable [11].

(3) The third step is differentiated between an empirical to conceptual, or conceptual to empirical approach. The first one is qualified if there is a lot of data available about the research object and not a lot of domain understanding. The second is qualified if there is a lot of domain understanding but not a lot of available data on the research object. If the domain understanding and the available data of the research object are at the same level, the researcher could decide which approach should be chosen [11]. Due to the huge data set and some extant literature in the field of data-driven SC projects, a mix was chosen here – the study starts with the conceptual to empirical approach in the first iteration and moves on to the empirical to conceptual approach for the following iterations. A detailed description of the development stages can be found in section 4.1.

In the conceptual to empirical approach, the researcher starts to build up knowledge about the dimensions and characteristics that are a logical consequence of the meta characteristics. Afterwards, the researcher examines the existence of each object that fulfills each characteristic of each dimension and decides which dimensions can be verified [11].

In the empirical to conceptual approach, a subset is first identified, which then needs to be classified. This subset is known to the researcher due to a literature review. Subsequently, common characteristics of these objects are identified as a logical consequence of the meta characteristics. If there is a characteristic that fits all objects it should be seen as useless as characteristics should be assigned to one dimension [11].

After performing one iteration, the researcher checks if the subjective and objective end conditions are fulfilled. If this is not the case, a subsequent iteration would be started until no new object could be identified and the subjective and objective end conditions are fulfilled [11].

4 Taxonomy

4.1 Development Stages

Step 1: Following the research design by Nickerson et al. [11], we start to define meta characteristics. The goal of the taxonomy is to show possible use cases and characteristics of data in a SC. As a consequence, the meta characteristics of the taxonomy are data and its usage in a SC.

Step 2: We adopted the subjective and objective criteria from Nickerson et al. [11]. However, the criterion of excluding mutual characteristics was not adopted because overlapping is possible in this context. A characteristic ‘mixed’ could avoid such overlapping but this would not be precise enough and would not meet the subjective criteria. All in all, we selected the following objective criteria: (1) each dimension contains characteristics which are complete, (2) each dimension is unique, (3) each characteristic of a dimension contains at least one object, (4) there are no added dimensions and characteristics in the last iteration, and (5) no dimension, characteristics or objects were summarized in the last iteration. The subjective criteria indicate that the taxonomy is: (1) precise, (2) robust, (3) complete, (4) expendable, and (5) explainable.

After the first two steps, the iterations to develop the taxonomy start. All dimensions and characteristics of the taxonomy will be described in detail in the next section.

Step 3.1: The first iteration starts with a conceptual to empirical approach. Thus, the literature review was performed to build up the findings (see section 2.2). The first dimensions identified and examined with additional literature were: user of the solution [22, 30, 31], connection to the user [22], domain of the application [20, 21, 32], and periodicity of the data [23]. If a dimension was identified as not exclusive and more than one characteristic can be identified for a data-driven SC project, it was marked with the additional note ‘NEX’ (not exclusive). After the first iteration, the following taxonomy was developed:

$$T = \{ \begin{array}{l} D_1 \text{ user of the solution} \mid D_1 = \{\text{things; customer; provider}\} \{\text{NEX}\} \\ D_2 \text{ connection to the user} \mid D_2 = \{\text{smart devices; smart environment}\} \{\text{NEX}\} \\ D_3 \text{ domain of the application} \mid D_3 = \{\text{governance; environment; mobility;} \\ \text{infrastructure; technology; citizen; services}\} \\ D_4 \text{ periodicity of the data} \mid D_4 = \{\text{static; semi-dynamic; dynamic}\} \end{array} \}$$

Step 3.2: This iteration switches to an empirical to conceptual approach. The database of the EU Smart Cities Information System was used in order to gather empirical information [33]. This platform opens information about SC projects in order to encourage the exchange know-how and collaboration between citizens, developers, cities, industry, experts, and research centers. The data quality of this platform is sometimes expandable and sometimes a clear definition of a SC is missing. As a consequence, a large number of the projects in the database are not SC projects (e.g. an energetic reconstruction of buildings). However, GrowSmarter, Triangulum, and Smarter Together were established as projects with detailed data. All projects have the goal to improve the quality of life of the citizens and the ecological footprint of the city. All the projects also contain different subprojects. Thus, these provide a good database to further develop the taxonomy. These projects were analyzed by two independent researchers with a structured content analysis using an open coding system [34]. These two independent coding systems were compared and synchronized afterwards. After analyzing the projects, dimensions and characteristics were added to the taxonomy which are listed as follows:

$$T = \{ \begin{array}{l} D_1 \text{ user of the solution} \mid D_1 = \{\text{things; customer; companies, cities}\} \{\text{NEX}\} \\ D_2 \text{ connection to the user} \mid D_2 = \{\text{smart devices; smart environment, smart data} \\ \text{platforms}\} \{\text{NEX}\} \\ D_3 \text{ domain of the application} \mid D_3 = \{\text{governance; environment; mobility;} \\ \text{infrastructure; technology; citizen; services}\} \\ D_4 \text{ periodicity of the data} \mid D_4 = \{\text{static; semi-dynamic; dynamic}\} \\ D_5 \text{ data ownership} \mid D_5 = \{\text{citizen; company; city}\} \{\text{NEX}\} \\ D_6 \text{ data storage location} \mid D_6 = \{\text{company server; cloud server; city server; open} \\ \text{data platform; end device}\} \{\text{NEX}\} \\ D_7 \text{ data processing} \mid D_7 = \{\text{manually; automatically; part-automatically}\} \\ D_8 \text{ data quality} \mid D_8 = \{\text{reviewed; review necessary; no review}\} \\ D_9 \text{ data type} \mid D_9 = \{\text{numeric measuring data; numeric data; geographic data;} \\ \text{textual data; machine recognizable data}\} \{\text{NEX}\} \end{array} \}$$

D₁₀ usage of the data | D₁₀ = {evaluation; analysis; monitoring; user application; open data portal} {NEX}

D₁₁ data user | D₁₁ = {citizen; company; city; journalist} }

Step 3.3: The empirical to conceptual approach is applied here as well. Empirical data was added from the database Nominet [35], which is mainly a provider for the registration of domains but also offers additional services. Nominet provided information about 150 projects. After filtering the projects according to SC characteristics and skipping projects analyzed in the second iteration, we analyzed 54 SC projects in total. These projects were analyzed by two independent researchers with a structured content analysis using an open coding system [34]. These two independent coding systems were compared and synchronized afterwards. This analysis led to the further development of the taxonomy as follows:

T = { D₁ user of the solution | D₁ = {things; citizen; companies; cities; journalists; researcher; developer} {NEX}

D₂ connection to the user | D₂ = {smart devices; smart environment; smart data platforms} {NEX}

D₃ domain of the application | D₃ = {governance; environment; economy; mobility; infrastructure; technology; citizen; services}

D₄ periodicity of the data | D₄ = {static; semi-dynamic; dynamic}

D₅ data ownership | D₅ = {citizen; company; city} {NEX}

D₆ data storage location | D₆ = {company server; cloud server; city server; open data platform; end device}

D₇ data processing | D₇ = {manually; automatically; part-automatically}

D₈ data quality | D₈ = {reviewed; review necessary; no review}

D₉ data type | D₉ = {numeric measuring data; numeric data; geographic data; textual data; video data; machine recognizable data} {NEX}

D₁₀ usage of the data | D₁₀ = {evaluation; analysis; monitoring; practical application; open data portal; atomization} {NEX}

D₁₁ interaction with the data | D₁₁ = {synchronic; asynchronous; no interaction} }

Step 3.4: The empirical to conceptual approach was applied again in the last iteration. In order to expand the number of SC projects, another database called Bable [8] was analyzed. This database is a spin-off of the Fraunhofer IAO. It provides an overview of realized SC projects and provides a platform to gather objects in this field. As many SC projects had already been analyzed, only ten more projects came from this database. Thus, a free search on Google was performed to find additional projects; eight additional ones were found [30, 32]. After analyzing these 18 SC projects, no new dimensions or characteristics could be identified. Additionally, the characteristics are complete, the dimensions are unique, and they each contain only one object. In terms of the guidelines from the methodology section, all objective end criteria were fulfilled at this point. Additionally, the taxonomy is succinct, robust, complete, expendable, and explainable. Thus, the subjective end criteria are also fulfilled. The taxonomy is presented in detail in the next section.

4.2 Taxonomy for Data-Driven Smart City Projects

Our final taxonomy for data-driven SC projects can be found in table 1. The dimensions and characteristics of the taxonomy are as follows: (1) The dimension '*domain of application*' summarizes all areas of use in a SC project [20-22]. This dimension summarizes similar domains such as demographic, health, and education. The characteristics of these dimension are governance, environment, economy, mobility, infrastructure, technology, citizen, and services [20-22]. Measurement of the air quality can be seen as an environment characteristic [32]. (2) The second dimension is '*usage of the data*' dimension. This dimension contains qualitative and quantitative evaluation and analyzes of the data [36]. Monitoring of the data and its practical application as well as open data portals were identified [37]. An automatic control with the data was also identified as one characteristic of this dimension [38]. The characteristics of this dimension are not exclusive because the data could be used in different use cases at the same time, such as analysis and practical application [36]. (3) The third dimension is about the '*user of the solution*' in a SC [39]. Users could be citizens, companies, cities, or journalists [36, 39]. Another potential user of the solution could be a thing, such as a streetlight that could use data to control themselves automatically [38]. Researchers and developers were also identified as data users of the solution [40]. This dimension is not exclusive because users could use the solution simultaneously; the city and companies could use the solution at the same time, for example [36]. (4) A fourth dimension was identified as the '*connection of the user*' [22]. This dimension could include smart devices and smart environments [22]; smart data platforms were also identified as a connector in this dimension [41]. The characteristics of this dimension are not exclusive because different users could participate a smart system or solution at the same time and through different connection possibilities [22]. (5) '*Data ownership*' was identified as a fifth dimension [36]. This could be citizens, companies, public authorities respectively public facilities – summarized as cities – or mixed forms [36]. As mixed forms are possible, this dimension is not exclusive. (6) The next identified dimension was the '*data storage location*'. Data is mostly stored on a company's own server, in cloud solutions, in open data platforms, and on servers belonging to the city [36, 42]. These data storage locations are mainly used by the creators of the data, while end users are more likely to save the data on their end device or to open the data via app or website [24]. As the data needs to be saved on one platform in order to analyze it but it is possible to transfer the data from different platforms, this dimension is not exclusive. However, it is possible to gather the data from many different storage locations. (7) Another identified dimension was '*data processing*' which includes the data creation and processing. The creation and processing can be done manually, automatically, or partly automatically, meaning that some parts are processed automatically and some input needs to be processed manually [36]. As the last characteristic is a combination of both first ones, this dimension is exclusive.

(8) The next dimension of the taxonomy is ‘*data quality*’. This can vary depending on whether the data could be pre-checked by the data supplier or not. A review of the data quality is necessary in the latter case [36]. Thus, the data could be reviewed, not reviewed or partly reviewed. As the last characteristic is a combination of both first ones, this dimension is exclusive. (9) As data can be structured differently, the dimension ‘*data type*’ was identified. This could be numeric measuring data, numeric other data, geographical data, and textual data [36]. Numeric measuring data is an extra characteristic although this data type is a subtype of numerical data as the empirical study identified an accumulation of numeric measuring data. Further, we could identify the characteristic video data which is used in smart car parking systems for instance [30]. If the data is meant to be processed automatically, it needs to be in a structure of machine recognizable data [24]. This characteristic was therefore added to the dimension. All in all, it is possible for data to be available in different data types at the same time; data can be both, numerical data and machine recognizable data, for example. As a consequence, this dimension is not exclusive in the taxonomy. (10) As a subsequent dimension, ‘*interaction with the data*’ was identified [31]. For example, a citizen could use a smart traffic light system that enables them to receive a green light at every intersection on their way [31]. In this case, the characteristic is synchronic. More examples show that there could be also no communication with the user resulting in everything being automatically controlled, e.g. smart lighting in the city [38]. Additionally, an asynchronous interaction is also possible, such as if the user is receiving messages under certain data circumstances [43]. This dimension is exclusive because we do not indicate a combination of these three characteristics. (11) The last dimension of the taxonomy is ‘*periodicity of the data*’ which could be static, dynamic, or semi-dynamic [23]. As the last characteristic is a combination of the two previous ones, this dimension is exclusive.

4.3 Evaluation

As mentioned by Szopinski et al., researchers use often use cases to evaluate their taxonomy [29]. Thus, we decided to use this method in order to evaluate our data-driven SC projects taxonomy.

Table 2. Use Cases for Evaluation

Category	Use Cases	Examples
Internet of Things	18	Toyotas Woven City, Smart Road in Hamburg, Train Station Berlin Südkreuz, VTG Connect
Online Services	6	Smart Networks for Citizens Participation, Windcloud 4.0, ELEVATE Delta
Smart Grids	5	Grow Smarter Cologne, Sync Fuel
Robotics	7	Robotic Vessels as a Service, Powder Buddy
Big Data	5	Port Monitor Harbor Hamburg, NUNAV

Category	Use Cases	Examples
Network technologies	1	My SMARTLife electricity
Artificial Intelligence	2	Project HEAT, Forum 4.0
Augmented Reality	1	Speicherstadt Digital

As many different projects from many databases had already been analyzed, other databases that could offer use cases that reflect data-driven SC projects were sought out. A SC project was found in the database ‘Smart City Compass’ [44]. This database contains 45 different smart city projects (excluding projects which are doubled in the database) in Germany. The different categories used by the database can be found in Table 2. The evaluation of the taxonomy occurs through its implementation across all 45 cases in the database by two independent researchers. If there was not enough data about the project available in the database, more available information was sought out on the internet. Neither of the researchers could identify a case that did not fit into the taxonomy or had characteristics that needed to be added into the taxonomy of data-driven smart city projects. How often each characteristic was selected during the evaluation of the 45 use cases can be seen in Figure 1.

Dimension	Characteristics								exclusive
<i>Domain of application</i>	Governance (3)	Environment (4)	Economy (5)	Mobility (9)	Infrastructure (4)	Technology (3)	Citizen (9)	Services (8)	yes
<i>Usage of the data</i>	Evaluation (12)		Analysis (34)	Monitoring (10)	Practical application (15)	Open data portal (21)	Automatization (32)		no
<i>User of the solution</i>	Things (18)	Companies (22)	Cities (37)	Citizen (35)	Journalists (6)	Researcher (11)	Developer (13)		no
<i>Connection to the user</i>	Smart devices (34)			Smart environments (16)		Smart Data platform (17)			no
<i>Data ownership</i>	Citizen (8)			Company (32)		City (36)			no
<i>Data storage location</i>	Company server (17)		Cloud server (32)	City server (13)	Open data platform (18)		End device (7)		no
<i>Data processing</i>	Manually (8)			Automatically (15)		Part-automatically (22)			yes
<i>Data quality</i>	Reviewed (14)			review necessary (23)		No review (8)			yes
<i>Data type</i>	Numeric measuring data (23)		Numeric data (18)	Geographical data (17)	Textual data (5)	Video data (7)	Machine-recognizable data (45)		no
<i>Interaction with the data</i>	Synchronic (18)			Asynchronous (22)		No interaction (5)			yes
<i>Periodicity of the data</i>	Static (9)			semi-dynamic (21)		Dynamic (15)			yes

Figure 1. Evaluation of the taxonomy

In order to illustrate the evaluation, one use case out of the 45 use cases was selected to be described and illustrated in the next paragraphs. The selected use case is called ELEVATE Delta. This is a SC initiative which aims to improve the lives of wheelchair users. ELEVATE Delta is an app which shows all elevators in the city with real time data of their functionality. Therefore, a wheelchair user can plan their trip from A to B in the city with this app with the advantage of knowing where elevators

are located, and real time information regarding whether the elevator is disabled and cannot be used [44]. This case is a data-driven SC project because (1) it improves the lives of wheelchair users and (2) the data of the elevators is an essential key resource of the app. Thus, this project is a suitable use case to evaluate this study's data-driven SC project taxonomy.

Dimension	Characteristics								exclusive
<i>Domain of application</i>	Governance	Environment	Economy	Mobility	Infrastructure	Technology	Citizen	Services	yes
<i>Usage of the data</i>	Evaluation		Analysis	Monitoring	Practical application	Open data portal	Automatization		no
<i>User of the solution</i>	Things	Companies	Cities	Citizen	Journalists	Researcher	Developer		no
<i>Connection to the user</i>	Smart devices			Smart environments		Smart Data platform			no
<i>Data ownership</i>	Citizen			Company		City			no
<i>Data storage location</i>	Company server	Cloud server		City server	Open data platform		End device		no
<i>Data processing</i>	manually			automatically		Part-automatically			yes
<i>Data quality</i>	reviewed			review necessary		No review			yes
<i>Data type</i>	Numeric measuring data	Numeric data	Geographical data	Textual data	Video data	Machine-recognizable data			no
<i>Interaction with the data</i>	synchronic			asynchronous		No interaction			yes
<i>Periodicity of the data</i>	static			semi-dynamic		dynamic			yes

Figure 2. Evaluation Example of the taxonomy

As seen in Figure 2, the taxonomy could be used to identify the characteristics of the use case. All exclusive dimensions reflect only one characteristic of this dimension in the given use case. As the project improves the mobility of wheelchair users, the characteristic of the dimension '*domain of application*' is mobility. The data is used in different ways in this case. The first way is in the monitoring of the elevators. Secondly, the data is available in an open data portal. Thirdly, the data has a practical application because it supports the wheelchair users. Fourthly, the data evaluates the functionality of the elevators. These are the four characteristics of the dimension '*usage of the data*'. As the '*users of the solution*' are citizens with wheelchairs, the characteristic in this dimension is citizen. The '*connection to the user*' takes place via smartphone, smart devices, and the app - a smart data platform. As the project was initiated by a company, the '*data ownership*' lies with them. Due to the small size of the company, they use a cloud server for the '*data storage location*'. As the data on the elevators is partly processed automatically and partly manually, the dimension '*data processing*' is reflected by a part-automatically processing. The '*data quality*' dimension is reviewed because by the quality department of the company. The dimension '*data type*' is reflected by the three data types: geographical data (location of the elevator), textual data (information about the elevator), and machine recognizable data (automatic processing of the data). The '*interaction with the data*' is synchronic because it is real

time data. Lastly, the '*periodicity of the data*' is semi-dynamic because of the mix of automatically and manually processed data.

5 Discussion and Conclusion

Constructing on a well-established methodology from information systems literature introduced by Nickerson et al. [11], a taxonomy of data-driven SC projects was developed in this study. Overall, four iterations were conducted, one being conceptually based on a literature review of current SC literature and three being empirically grounded in a heterogeneous set of SC projects. The taxonomy of data-driven SC projects consists of eleven dimensions, each represented through a distinct set of characteristics and all providing a means to conceptualize data-driven SC projects as a phenomenon. The evaluation of the taxonomy indicates its reliability in terms of classifying and distinguishing cases of data-driven SC projects different studies. Thus, the taxonomy was proven to be useful and it meets its purpose by providing empirically validated and conceptually grounded characteristics of data-driven SC projects.

Introducing a thoroughly developed, reliable taxonomy of data-driven SC projects offers immediate implications for research and practice. The taxonomy offers an effectual way to indicate the characteristics of data-driven SC projects – effectual because the taxonomy allows researchers to describe data-driven SC projects in a consistent approach and to distinguish them from each other. Thus, the taxonomy contributes to the existing body of knowledge on data-driven SC projects, resulting in a shared language that has been missing up to this point. In particular, the added descriptive knowledge helps to build a better understanding of the key dimensions and the characteristics they entail. A common knowledge based on this taxonomy highlights the materialization of SC project ideas and understanding among scholars that will lead towards the development of a deeper theorizing process on data-driven SC projects. All of this leads to the creation of new ideas according to data-driven smart city projects and researchers can better position their work in the SC field based on this taxonomy. A common understanding of data-driven SC projects also gives rise to implications for organizations. Despite the taxonomy's simplicity, it may prove to be highly effective in identifying the chances and opportunities for data-driven SC projects. This could be due to the different characteristics in the taxonomy, such as the opportunity of cloud computing, was not taken into account at the beginning of the data-driven SC project initiative. Therefore, the application of the taxonomy allows for strategic differentiation of SC initiatives. Practitioners may also benefit from the taxonomy as it could provide them with initial guidance in terms of the materialization of ideas and considerations regarding new SC projects.

Nevertheless, this research has some boundaries and limitations that open up the potential for future research. The methodology of the taxonomy follows a design science approach building an effective solution in order to solve identified problem [45]. However, it is not given that it provides an optimal solution and other independent studies may generate different results. The results from our evaluation are also limited in their generalization because all the use cases which are taken into consideration were

from Germany. Applying the taxonomy to a set of use cases with global origins would help better evaluate the concept; future research should re-evaluate and adjust the taxonomy accordingly. Common databases in this field were used to collect the dataset of SC projects. It appeared that this study captured a sufficient set of diverse SC projects from different contexts and regions. Yet, there are still more use cases that have not been realized which may lead to increased diversity in the dataset.

There is also potential for this taxonomy to serve future research as it lays the foundation for a deeper process to theorize the nature of data-driven SC projects. With a reliable taxonomy at hand future research may focus on collecting more use cases of SC initiatives, leading to high-order constructs such as data-driven SC projects archetypes. Such archetypes could help to further theorize how data can be further used in the SC context. Overall, this research lays the foundation for future developments in the field of data-driven SC projects and initiatives.

References

1. United Nations, <https://www.un.org/development/desa/publications/2018-revision-of-world-urbanization-prospects.html> (Accessed: 10.07.2019)
2. Sankhe, S., Vittal, I., Dobbs, R., Mohan, A., Gulati, A., Ablett, J., Gupta, S., Kim, A., Paul, S., Sanghvi, A.: India's urban awakening: Building inclusive cities, sustaining economic growth. McKinsey Global Institute (2010)
3. Barreto, L., Amaral, A., Baltazar, S.: Urban mobility digitalization: towards mobility as a service (MaaS). In: 2018 International Conference on Intelligent Systems (IS), pp. 850-855. IEEE (2018)
4. Dameri, R.P.: Smart city implementation. Progress in IS; Springer: Genoa, Italy (2017)
5. Dameri, R.P., Cocchia, A.: Smart city and digital city: twenty years of terminology evolution. In: X Conference of the Italian Chapter of AIS, ITAIS, pp. 1-8. (2011)
6. Hartmann, P.M., Zaki, M., Feldmann, N., Neely, A.: Big data for big business? A taxonomy of data-driven business models used by start-up firms. In: University of Cambridge, Cambridge Service Alliance, pp. 1-30 (2014)
7. Kühne, B., Böhmman, T.: Data-Driven Business Models - Building the Bridge Between Data and Value. 27th European Conference on Information Systems Stockholm, Sweden, pp. 1-16 (2019)
8. Bable, <https://www.bable-smartcities.eu/explore/use-cases/?L=0> (Accessed: 24.04.2020)
9. Perboli, G., De Marco, A., Perfetti, F., Marone, M.: A new taxonomy of smart city projects. Transportation Research Procedia 3, pp. 470-478 (2014)
10. Niaros, V.: Introducing a taxonomy of the "smart city": Towards a commons-oriented approach? tripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society 14, pp. 51-61 (2016)
11. Nickerson, R.C., Varshney, U., Muntermann, J.: A method for taxonomy development and its application in information systems. European Journal of Information Systems 22, pp. 336-359 (2013)
12. Chourabi, H., Nam, T., Walker, S., Gil-Garcia, J.R., Mellouli, S., Nahon, K., Pardo, T.A., Scholl, H.J.: Understanding smart cities: An integrative framework. In:

- 45th Hawaii international conference on system sciences, pp. 2289-2297 (2012)
13. Anttiroiko, A.-V., Valkama, P., Bailey, S.J.: Smart cities in the new service economy: building platforms for smart services. In: *AI & society* 29, pp. 323-334 (2014)
 14. Giourka, P., Sanders, M.W., Angelakoglou, K., Pramangioulis, D., Nikolopoulos, N., Rakopoulos, D., Tryferidis, A., Tzovaras, D.: The smart city business model canvas—A smart city business modeling framework and practical tool. In: *Energies* 12, pp. 4798 (2019)
 15. Vargo, S.L., Lusch, R.F.: Evolving to a new dominant logic for marketing. In: *Journal of marketing* 68, pp. 1-17 (2004)
 16. Cheng, B., Longo, S., Cirillo, F., Bauer, M., Kovacs, E.: Building a big data platform for smart cities: Experience and lessons from santander. In: *2015 IEEE International Congress on Big Data*, pp. 592-599. IEEE (2015)
 17. Bifulco, F., Tregua, M.: *Service Innovation and Smart Cities: Linking the Perspectives. Innovating in Practice*, pp. 261-287. Springer (2017)
 18. Albino, V., Berardi, U., Dangelico, R.M.: Smart cities: Definitions, dimensions, performance, and initiatives. In: *Journal of urban technology* 22, pp. 3-21 (2015)
 19. Su, K., Li, J., Fu, H.: Smart city and the applications. In: *2011 International Conference On Electronics, Communications and Control (ICECC)*, pp. 1028-1031. IEEE (2011)
 20. Lombardi, P., Giordano, S., Farouh, H., Yousef, W.: Modelling the smart city performance. *Innovation*. In: *The European Journal of Social Science Research* 25, pp. 137-149 (2012)
 21. Sullivan, F., <https://ww2.frost.com/wp-content/uploads/2019/01/SmartCities.pdf> (Accessed: 18.01.2019)
 22. Lim, C., Maglio, P.P.: Data-driven understanding of smart service systems through text mining. In: *Service Science* 10, pp. 154-180 (2018)
 23. Bischof, S., Karapantelakis, A., Nechifor, C.-S., Sheth, A.P., Mileo, A., Barnaghi, P.: Semantic modelling of smart city data. In: *Wright State University CORE Scholar Conference*, pp. 1-5 (2014)
 24. Tcholtchev, N., Farid, L., Marienfeld, F., Schieferdecker, I., Dittwald, B., Lapi, E.: On the interplay of open data, cloud services and network providers towards electric mobility in smart cities. In: *37th Annual IEEE Conference on Local Computer Networks-Workshops*, pp. 860-867. IEEE (2012)
 25. Webster, J., Watson, R.T.: Analyzing the past to prepare for the future: Writing a literature review. In: *MIS quarterly* 26, pp. xiii-xxiii (2002)
 26. Paukstadt, U., Gollhardt, T., Blarr, M., Chasin, F., Becker, J.: A taxonomy of consumer-oriented smart energy business models. In: *Proceedings of the 27th European Conference on Information Systems*, pp. 1-17 (2019)
 27. Nagel, E., Kranz, J., Sandner, P., Hopf, S.: How Blockchain Facilitates Smart City Applications—Development of A Multi-Layer Taxonomy. In: *Proceedings of the 27th European Conference on Information Systems*, pp. 1-17 (2019)
 28. Heinrich, K., Roth, A., Zschech, P.: Everything Counts: A Taxonomy of Deep Learning Approaches for Object Counting. In: *Proceedings of the 27th European Conference on Information Systems*, pp. 1-16 (2019)
 29. Szopinski, D., Schoormann, T., Kundisch, D.: Because your taxonomy is worth it: Towards a framework for taxonomy evaluation. In: *Proceedings of the 27th European Conference on Information Systems*, pp. 1-19 (2019)

30. Smart Parking, <https://www.smartparking.com/smartpark-system/anpr> (Accessed: 21.04.2020)
31. Cardwell, D.: Copenhagen lighting the way to greener, more efficient cities. *New York Times* 8 (2014)
32. iCity Project, <http://icityproject.eu/information-systems-map> (Accessed: 13.04.2020)
33. EU SCIS, <https://smartcities-infosystem.eu/content/about-smart-cities-information-system-scis> (Accessed: 10.04.2020)
34. Mayring, P.: On generalization in qualitatively oriented research. In: *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research* (2007)
35. Nominet, <https://www.nominet.uk/about> (Accessed: 19.04.2020)
36. Enarson, L.: *Data management plan for GrowSmarter*. Stockholm: City of Stockholm (2017)
37. Stoppel, V., Klassen, B.: *Smart City Munich Lighthouse Project - Documentation of activities and achievements*. München: Stadt München D4.1.3 (2019)
38. Tviligh, <https://www.tvilight.com/case-studies> (Accessed: 21.04.2020)
39. Schieferdecker, I., Tcholtchev, N., Lämmel, P., Scholz, R., Lapi, E.: Towards an Open Data Based ICT Reference Architecture for Smart Cities. In: *2017 Conference for E-Democracy and Open Government (CeDEM)*, pp. 184-193. IEEE (2017)
40. Greater London Authority, <https://data.london.gov.uk> (Accessed: 25.04.2020)
41. VMZ Berlin, <https://www.smarter-together.eu/file-download/download/public/1257> (Accessed: 07.05.2019)
42. Rolando, D., Palm, B., Claesson, J., Nilsson, A., Robert, M., Shahrokni, H., https://grow-smarter.eu/fileadmin/editor-upload/reports/grow-smarter_validation.pdf (Accessed: 08.05.2019)
43. Bigbelly, <https://bigbelly.com/platform> (Accessed: 24.04.2020)
44. ThiS, <https://www.smartcity-kompass.de/smartcity/elevate-delta/> (Accessed: 15.07.2020)
45. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. In: *MIS quarterly*, pp. 75-105 (2004)