Scientific article accompanying the MSc thesis in Complex Systems Engineering & Management

Deriving the Data Ecosystem to accelerate the heat transition in the Netherlands

Devin D.D. Diran[1]*

Abstract

The transition towards a sustainable energy system in the built environment is commonly referred to the incremental adoption of a variety of available technologies, practices and policies that may contribute to decrease the environmental impact, at reasonable costs and adequate quality standards. The complexity of the transition requires a sound information provision in order to make decisions which are future proof and optimal in the context of the system dependencies by both public and private parties. This information provision is lacking maturity, and in the era of Big and Open Data, it is believed that data has a significant role to play in the improvement of the information provision towards all stakeholders. This paper presents a study on the information provision in the Dutch energy transition with the focus on the thermal urban transition away from natural-gas. The study adopts the DE approach to answer the following research question: *What are the necessary elements, roles and context in an DE to enable the data-driven support of the local heat transition, by capturing citizen preferences, willingness to participate, and attitude?*

Number of words

11,132

Keywords

Data Ecosystem; Open Data; Sustainable Energy; Thermal Energy System; Urban

[1] *Faculty of Technology Policy & Management, Delft University of Technology, Delft, the Netherlands*

1. Background and problem definition

The challenge posed by climate change upon humanity can only be adequately addressed by drastically reducing human-induced greenhouse gas (GHG) Emissions to the point of zero-emissions [1]. The Netherlands is on the verge of implementing a new framework which should accelerate the energy transition, namely the Klimaatakkoord or Climate Agreement. The goal of the Climate Agreement is to propose and enable the realisation of measures to reduce GHG emission by 49% in 2030, relative to 1990, among the built environment, the electricity sector, industry, agriculture and land-use, and mobility [2].

For households in the Netherlands, the generation of heat for a variety of purposes such as space heating, warm water and cooking, contributes for over 82% in the final energy consumption [3]. This situation where the heating provision is dominated by natural-gas, over 87% in 2015 [4], is unique for the Netherlands since other countries have more diverse heating systems in terms of the energy carrier, technology for energy conversion and the location of heat generation [5]. In the Netherlands national policy is targeting the disconnection of buildings from the use of natural gas for heating purposes. To establish the disconnection from natural gas and the transition of the building stock, local governments will have a pivotal role together with the building owners and citizens. In the planning and realisation of the sustainable urban heating systems for all neighbourhoods, the technical aspects of the buildings, are not the only decisive factors. In addition, the inclusion of citizens and other stakeholders with regards to their interests and resources is decisive.

Because of this great share of energy consumption for heat in the urban setting, this study will focus on the provision of sustainable heat at low cost and in a reliable and natural-gas-free manner for households. Here, the increasing decentralisation in the generation of energy, and the increasing diversity in the energy system to fit the local technical and social conditions are taken into account [6].

To successfully fill in the novel responsibilities and tasks to kick-off the energy transition at the pace as desired in the Klimaatakkoord [2, 7], the local governments and stakeholders will need to innovate and be creative with the existing decision-making mechanisms. However, this decision-making requires a sound knowledge base to result in effective, supported and future-proof or robust decisions. Among the actors in the heat transition there is common consensus that the current information provision is lacking and that knowledge gaps are left unaddressed by the current information provision. However, there is little known about what these knowledge gaps are and how that differs over the various actors. A recent study on the knowledge gaps in the Netherlands for policy issues in the Energy transition, yielded knowledge gaps which are predominantly on the supply side of energy [8]. Subsequently an approach is proposed to establish a data strategy which targets the supply side.

This is acknowledged by the Climate Agreement, and in the program on the Improvement of the Information Provision in the Energy Transition, these knowledge gaps are mapped for the policy decisions and a plan is proposed to target these knowledge gaps with data of all kinds [2, 8]. However, the knowledge needs and subsequent knowledge provision and support to citizens, is not yet addressed.

However, a study by [9], on the local level in the Netherlands, surfaced that on the local level it is predominantly the demand side which prevails the most knowledge gaps. The public authorities require detailed, accurate and topical information on the citizens regarding their consumption behaviour, but also the attitude and willingness to participate in the heat transition, and the preferences towards technologies and solutions [10, 11, 12, 13]. This need for citizen information is met by a lack of knowledge on how to effectively, reliably and frequently capture this data and subsequently process the data to determine appropriate strategies of citizen engagement in the heat transition. To target this knowledge gap, this paper aims to answer the following research question:

What are the necessary elements, roles and context in an DE to enable the data-driven support of the local heat transition, by capturing citizen preferences, willingness to participate, and attitude?

Big Data and Open Data have been carrying great promise as technologies to enhance society, affecting all aspects of human activity, by incentivising citizen participation, transparency, economic growth and innovation [14]. The technological advances achieved in the last decades in particularly the ICT domain, such as the use of social media, smart phones and the Internet of Things is driving the exponential increase in the volume of data, leading to the so called Big Data. Accordingly to the immense growth in Big and Open Data, the infrastructure and technology to enable these technologies, such as Big data storage and processing facilities, Open data portals and platforms and tools and instruments, have been introduced and developed at an unprecedented pace to exploit the potential behind Big and Open Data. [14] and [15] argue that these facilities and infrastructure are utilised by many users, and also developed and maintained by many parties. Hence it is proposed to view them as part of a wider open Data Ecosystem (DE) in which each instrument or tool can add value as part of the puzzle. This is the main reason why in this research DE theory plays an important role to understand the involved stakeholders and the interaction between them in generating and utilising Open and Big Data for an inclusive heat transition. The Heat-transition will be addressed from the perspective of its current and future DE which should form the foundation for the provision of knowledge for decision making among all actors and stakeholders involved.

1.1 Structure of the paper

After introducing the problem and the intended approach in section 1, section 2 will proceed with the literature background on Big and Open data in energy systems and DE theory. Section 3 addresses the methods used briefly and the case of Utrecht, subsequently section 4 will address the results, and section 5 presents the proposed DE 2.0. Finally, section 6 discusses the findings, and the paper is concluded in section 7.

2. Literature background

2.1 The concepts of Big and Open data

Data is increasingly gaining value as an economic asset, however, this is not yet the case for policy making and public decision making in general. Data utilisation, among others, offers new insights into behaviours and patterns, with less dependence on surveying. However, it requires major efforts to bring together multiple actors from various disciplines and practices to study the under-explored relationship between data types [16].

2.1.1 Big Data

Big data, the first concept to be discussed in this section, and its role in both public and private decision-making has attracted significant research in the past decade [17, 18]. The technological advances achieved in the last decades, in particular the ICT domain, such as the use of social media, smart phones and Internet of Things, is driving the exponential increase in the volume of data leading to the so called Big Data. Hereby literature is quite diverse in the definition of Big Data, among others due to the strong developments. A commonly found definition is that big data means *data-sets that are too large for traditional processing systems and require new technologies* [19]. A more comprehensive definition, which also emphasises the increase in scale and scope of data is as follows: *"Big Data is a step change in the scale and scope of the sources of materials (and tools for manipulating these sources) available in relation to a given object of interest" [20, p. 349]*. This working definition will be applied for the remainder of the thesis, due to its comprehensiveness and it being more specific than the more common definition of Big Data in industry, where Big Data is defined by means of its 5 characteristics of Volume (e.g. a large number of objects or time series), Variety (Large variety among data types e.g. statistics, sensors, social media), Velocity (high speed of data generation, e.g. real-time data) and Veracity (the trustworthiness or validity of the data) [21]. With the increasing application of sensors, wireless network communication, advanced metering, and cloud computing technologies, large amounts of data are continuously being accumulated in the energy sector. These applications have predominantly been utilised in the context of Smart Grid research and power systems to enable generation side and demand side management, asset management and collaboration, and micro-grid and renewable energy generation [22].

2.1.2 Open Data

In addition to Big Data, this research addresses Open Data as a means or source to knowledge and decision support. The Open Knowledge Foundation, an international non-profit organisation, promotes Open Data as part of the their Open Knowledge campaign and define Open Data as *"a piece of data or content that is free to be used, reused, and redistributed - subject only, at most, to the requirement to attribute and/or share-alike" [23, p. 6].* Other characteristics, included by alternative definitions, include: standard formats, no restrictions from copy rights or patents, interoperable, digital and machine readable format, and Public Sector information [24].

For data to be classified as Open Data, the data should comply to the 10 criteria established in 2010 by the Sunlight Foundation, a NGO aiming to make governments and politics more accountable and transparent by means of among others open data, namely: complete, primary, timely, accessible, machine processable, non-discriminatory, non-proprietary, permanent, licence-free, and preferably free of charge [25].

Open Data is being increasingly considered as a major enabler of public service innovation [26]. The open access to data-sets, without any predefined financial and regulatory restrictions or conditions of use, and advanced analytical methods and tools have opened up endless possibilities to generate new knowledge to drive data-driven solutions. In the next subsection these applications of Big and Open data for energy and climate policy will the addressed.

2.1.3 Big and Open Data in Climate and Energy Policy

Big and Open Data can be sourced from both the energy infrastructure for generation [27] and distribution [24], as from the demand, influenced by end-user behaviour and building characteristics [28, 29]. In the heat transition, climate policy has a significant role to enable and guide the transition. For Evidence-based Policy-Making, the data-driven approach utilises data on energy, the environment and the economy, to establish a knowledge base and subsequently recognise problems, set priorities, establish and implement policy, and finally monitor and assess the efficacy of policy and the need for adaptive response [30, 31, 32]. Although this relevance and necessity for open data of adequate quality, integrity and detail is acknowledged by the sector and studied over a wide field of applications in the energy sector, energy policy research is lacking behind other sectors in promoting open and reproducible data and methods [33, 20].

2.1.4 Energy demand estimation and mapping

[34, 35, 36] study spatial mapping of energy potential and demand. [36] addresses how publicly available data-sets, or open data, on housing and energy can be used to map and plan mass retrofit and propose targeted low carbon measures across a city, in order to address the challenges of having: incomplete data and the inability to aggregate and consolidate private sector housing retrofit activities to minimise installation costs.

2.1.5 Dwelling retrofit or renovation

On the end-user side, dwelling retrofit is considered an effective mean to drastically reduce energy use and enable more innovative, sustainable and energy efficient heating and cooling solutions. [37] applies data-analytics and energy modelling on the building technical and consumption aspects to determine the effectiveness of non-structural retrofit strategies on energy efficiency in a residential building. A more data-driven approach is applied by [38], where smart meter and retrofit data from 1,600 non-residential buildings is collected and merged with data from previous energy saving retrofit measures to predict the success of energy saving measures based on the type of building and the industry type. No comparable research applied on residential buildings could be found.

2.1.6 Challenges and barriers for Big and Open Data

The work by [39] provides a overview of recurring challenges and barriers occurring in the application of open data and open data portals worldwide after a systemic review on papers discussing these challenges and barriers. The challenges. complemented with challenges found in other work, are as follows:

- Inconvenient data access, availability and findability
- Clarity of data purpose
- Inadequate or poor data collection
- Data license complications
- Difficult to understand the data
- Incomplete data
- Inadequate meta-data
- Poor data life-cycle management
- Poor data traceability
- High paced technological development [40].
- Inconsistent data [28].
- Lack of data standards [29, 41, 42]
- Communication of results [43].

After introducing the concepts of Big and Open Data in this sub-section, and the encountered challenges and barriers, the following sub-section elaborates on the DE as presented in literature with the relevant theories

2.2 The DE Approach

After the discussion on the role of Big and Open Data in energy policy and in particular, urban thermal systems, in section 2, this section proceeds with the elaboration of the DE approach and the relevant theories. The section is subsequently concluded with a DE framework consisting of elements which encompass the roles, activities, and environment in DEs. This framework forms the means to recommend the suitable DE for the heat transition.

2.2.1 DE theory

The technological advances achieved in the last decades, in particular the ICT domain, such as the use of social media, smart phones and the Internet of Things is driving the exponential increase in the volume of data leading to the so called Big Data which is previously discussed in sub-section 2.1, together with Open Data. Accordingly to the immense growth in Big and Open Data, the infrastructure and technology to enable these technologies, such as Big data storage and processing facilities, Open data portals and platforms, and tools and instruments, have been introduced and developed at an unprecedented pace to exploit the potential behind Big and Open Data. [14] and [15] argue that these facilities and infrastructure are utilised by many users, and also developed and maintained by many parties. Hence it is proposed to view them as part of a wider open DE in which each instrument or tool can add value as part of the puzzle. Subsequently, in this research DE theory plays an important role to understand the involved stakeholders and the interaction between them in generating and utilising Open and Big Data for an inclusive heat transition. The Heat-transition will be addressed from the perspective of its current and future DE which should form the foundation for the provision of knowledge and support, to decision making among all actors and stakeholders involved.

Moreover, the literature review on DEs by [44] derived the following potential benefits of the DE approach: 1) improved political and social aspects, 2) improved economic aspects, 3) convenient data generation and utilisation, 4) improved communication and interaction between actors and 5) improved quality of data and services. These are thus benefits which are in line with the needs of modern day governance of complex challenges, such as the heat transition.

The term ecosystem is derived from its use in biology where an ecosystem represents a natural unit, functioning as a whole, consisting of plants, animals and microorganisms together with the non-living physical resources in the environment [44]. For Open Data, the Ecosystem metaphor is used to refer to the interdependent social system consisting of individuals, organisations, infrastructures, and resources that can be created in technology-enabled, information-intensive social systems [14, 44, 45]. In the study with the focus on the architectural elements of Big DEs by [15], the ecosystem metaphor is used to state that the Big Data challenge is not only that of its core technological components, but it is rather a complex whole of components to store, process, visualise, and deliver results. This whole of interrelated components is what is defined as the Big DE which deals with the evolving data, models, and supporting infrastructure over the entire Big Data Life-cycle.

2.2.2 DEs: Definitions

The term DEs is commonly used in studies, however there is little agreement on the definition of the term. The literature review by [44] surfaced 15 different definitions of DEs. In that literature review the work by [46, 47, 45] and [48] is presented as the most cited among the selected 29 primary studies published in the period between 2011 and 2016. The first work on DEs, according to the review by [44], is published in 2011 by [49] and [50].

[46] states that a DE consists of data cycles, and within these, intermediate data consumers, such as app developers and data wranglers process the data and share the cleaned, integrated and packaged data back into the ecosystem to be utilised by the end-consumers.

In the work by [51], DEs are generally defined as *"all activities for releasing and publishing data on the Internet, where data users can conduct activities such as searching, finding, evaluating, and viewing data and their related licenses, cleansing, analysing, enriching, combining, linking, and visualising data, and interpreting and discussing data and providing feedback to the data provider and other stakeholders"*[51, p. 3].

Alternatively, [49] look at a DE for linked data and provide the following definition: *"a data-based system where stakeholders of different sizes and roles find, manage, archive, publish, reuse, integrate, mash up, and consume data in connection with online tools, services, and societies."* [49, p. 326].

Although there is a large variety in definitions for DEs, certain characteristics of these ecosystems are common over the definitions. It can be observed that all mentioned definitions share the following components: 1) stakeholders and roles, 2) relationships and interactions, and 3) (digital) resources. In the proceeding sub-section these elements will be further elaborated on.

2.2.3 DEs: Characteristics and Elements

From the review study on DEs by [44], one of the main findings is that the various views on DEs, commonly identify these ecosystems as socio-technical systems with a large variety of technical and social elements involved and a high degree of inter-dependencies. Hence, the study and design of DEs requires thorough contextual understanding of human interactions, in relation to the technological, cultural, political, and economic context.

[52] argues that the context of DEs can be characterised in at least three different categories. First, the regulatory context entails laws, policies, standards, and agreements, and guides the structure of the DE and the relation between the actors. Second, the institutional context, where the actors operate under certain values, rules, and norms. This context is also called the environmental context by [51]. Third, [52] proposes the technological context, encompassing the IT resources and operators, but also other enabling technologies that contribute in connecting the DE elements.

Moreover, the review by [44] reports on characteristics of DEs, proposed by [48], that relate to the cyclical nature of processing resources, sustainability to maintain continuity of the ecosystem, user-centricity and demand-drivenness, and self-organisation.

Finally, [44] reports on the organisational structure of the DEs as relevant category of characteristics. This entails the actors and the type of role they take in the ecosystem, but also the interest and business models connecting these actors. Hereby the most common organisational structures found in the reviewed work are: 1) *Keystone-centric*, where stakeholders are organised around the Keystone actor responsible to directly or indirectly provide data, 2) *Intermediary-based*, where the data intermediaries are accountable to add value to the data, 3) *Platform-centric*, where a platform shapes the organisation of the DE, and 4) *Marketplace-based*, where the market provides the infrastructure, business-models, rules, and services for the exchange of data between actors.

After addressing the characteristics of DEs, the elements of DEs can be considered as the next operationalisation of the characteristics in the Ecosystem design. Several studies propose DE elements or components, in order to interact and achieve the functionality of the ecosystem. In table 1 it is presented which elements for both Open and big DEs are proposed by the work from [14, 53, 15] and [54].

In the first study to address these elements in this section, [14] propose a set of four core elements and three additional elements for DEs with the focus on Open Government Data (OGD). Next, three categories of ecosystem elements are proposed by [53]. Here it can be noticed that [53] is placing more emphasis on the social and regulatory aspects within the ecosystem. On the contrary, the set of elements proposed by [14] encompasses a broader set of aspects and particularly places the emphasis on the activities which are or can be carried out in DEs.

[15] and [54] study the Big DEs and propose different sets of elements. Compared to [15], [54] takes a broader approach in proposing the key elements of a Big DE where also the social aspects are taken into account, in addition to the technological aspects as previously presented by [15].

Finally, among these studies, it can be observed that none address both Big and Open DEs as the focus is on either Big or Open Data. Hence, this study will take an integrated approach for the DE, addressing both Big and Open Data.

Given the characteristics and elements of DEs, [50] proposes a set of steps to facilitate the development of an DE. However, with the focus on an Open Government Ecosystem. These steps are as follows:

1. identify the people and organisations that can take a role as essential component of the ecosystem
2. comprehend the nature of the transactions and interactions that occur between those entities
3. identify what resources are needed by each entity to engage with each other in transactions and interactions
4. observe the indicators to measure the health and performance of the ecosystem as a whole

These steps will be taken into account when the stakeholder map will be derived for the heat transition in the Netherlands. This, to comprehend the current DE and use it a foundation for the design of a proposed DE. The following sub-section elaborates further on the theoretical background behind the novel DEs.

Table 1. An overview of DE elements derived from literature

Ecosystem type	Element or Component	Source
Open Data Ecosystems	1) Releasing and publishing open data on the internet 2) Searching, finding, assessing and viewing data and the associated licenses 3) Cleansing, analysing, enriching, combining, linking and visualising data 4) Interpreting and discussing the data and providing feedback to stakeholders and the data providers 5) User pathways to inspire users on how data can be used 6) A quality management system 7) Meta-data to connect the elements	[14]
Open Data Ecosystems	1) Government policies and practices 2) Innovators as a combination of technology, business and the government 3) Users, civil society, and business	[53]
Big Data Ecosystems	1) Data models, Structures and Types, this concerns Data formats, file systems etc. 2) Big Data Management on e.g. Big Data Life-cycle, Big Data transformation and staging, and sourcing, curating, and archiving 3) Big Data analytics and tools, e.g. Big Data applications, the target use, presentation, and visualisation 4) Big Data Infrastructure for e.g. storage, computation, exchange, and Big Data operational support 5) Big Data Security, on in-rest and in-move data, and through trusted processing environments	[15]
Big Data Ecosystems	1) Infrastructure 2) Software and Technology 3) Service and Applications 4) Standards 5) Users 6) Social and Cultural factors 7) Government 8) Industry	[54]

2.3 DEs: theoretical foundation

According to [44] the theoretical basis for the DEs has been slow to develop as there is no common agreement on how theories should look like. When considering the work carried out in the field of DEs and as reviewed by [44], a final selection of 29 articles, a large variety of theories is applied to study and understand DEs. The main theoretical foundations adopted by DE studies are: Socio-technical systems theory, Resource dependency theory, Value chain theory, Business ecosystems, and Software ecosystems.

As was presented in sub-section 2.2.3, DEs, among others, include elements such as activities and processes, databases, work-flows, people, market parties, the government, and infrastructure. With these elements, DEs need to combine and integrate components from various ecosystems and domains [44]. It could be derived from the previous list of theoretical foundations, that alternative Ecosystem theories are commonly applied to study the novel DEs. [14] relates the DE to Business Ecosystems, Innovation Ecosystems, Digital Ecosystems, and Open Government Ecosystems.

The body of literature on DEs is thus growing, and applications are wide-spreading over various sectors and stakeholders. However, no application of DE research has been found on the energy sector. As other sectors, the energy sector is ever increasing in digitisation with the introduction of smart metering and control technology to promote energy efficiency and demand side management, but also to cope with the increasing diversity and share of (intermittent) energy sources which pose challenges to balance the grid. This is one reason why this study on the realisation of an inclusive and sustainable urban heat supply can build upon the DE approach. In this DE approach, the system as a whole needs to combine components from different domains in order to be a functioning whole, taking into account the dynamic interaction between the different factors and components.

With the theoretical background on the DE approach in the previous sub-sections, the following sub-section establishes an DE framework which will be utilised for the remainder of this research.

2.4 The DE framework: elements and roles

Building forward on the literature reviewed on DEs in the previous sub-sections, this section derives a DE framework to assess the case of the heat transition.

For DEs, in general, three types of actors or roles can be defined, according to [55, 56] these three roles are:

- *Data producers and providers*: the organisations or entities which produce and provide the data, whereby the data can be acquired over a myriad of technologies and the produced data can be categorised according to the various data types.
- *Data users or consumers*: the organisations or entities which can use the data and derive valuable insights. Categories of data users are: Academic, Commercial, Governmental or public, Non-profit and Citizens,
- *Data intermediaries*: Organisations or entities which organise and facilitate the exchange of data, and coordinate the participation of users and providers. These are the organisational functions of the intermediary. Additionally, the intermediaries may or may not have data-related functions, e.g. data pre-processing and the provision of technology for data sharing.

In figure 1, see the lower part referenced ”data-stakeholder context” in blue, it is depicted how these roles relate to each other and to the DE. For each role it is stated which challenges commonly occur in the DE.

The elements of the DE are derived from the work by [14], focusing on Open Government Data, and [15], focusing on Big Data, whereby the elements proposed by these studies complement each other to cover the most important characteristics and interactions in DEs. The included elements come down to:

1. Data capturing and pre-processing to desired format
2. Data release or sharing
3. Searching, finding, viewing and assessing data and data licenses
4. Cleansing, linking, analysing and visualising data
5. Discussing data and providing feedback to providers
6. Meta-data to connect elements
7. Use-case promotion
8. Quality management

These elements capture the data life-cycle management or value chain where various approaches propose steps on how to gather data and add value to the data as it passes through the different steps. [57] elaborates on the following steps in the data value chain: 1) Data acquisition, 2) Data analysis, 3) Data curation, 4) Data storage, and 5) Data usage. In the DE for Big Data [15] propose a Big Data Life-cycle Management model consisting of the following steps: 1) Data Collection and Registration, 2) Data Filtering, 3) Enrichment and Classification, 4) Data Analytics, Modelling and Prediction, 5) Data delivery and Visualisation, and 6) a parallel step of 6a) Data re-purposing, 6b) Analytics re-factoring and 6c) Secondary processing. What these approaches commonly show is that the steps are organised in a linear way, whereby often feedback-loops occur towards earlier steps. The activities proposed as part of the steps in the Data Value Chain are thus taken into account in the DE elements depicted in figure 1 and surrounded by the Regulatory, Institutional or Environmental, and Technical context [52].

3. Research Methodology

3.1 Research design

This research uses a case study method, in particular an embedded single-case study design in applied according to the case-study design definition by [58]. The case study design encompasses the Netherlands as the main case, the stage where the heat transition has to be realised on a national level, and the municipality of Utrecht as the embedded case. The case study is executed as a qualitative case study, where the main mean of data collection are interviews. The benefit of the qualitative case study is that rich information can be derived from the interviews on the local issues regarding the heat transition, and the associated DE, via the critical interactions of complex social phenomena [59]. Hence, the case study meets in the necessity to derive the real-life context on the heat transition with regards to the actors involved, the current and future decision-making processes, the knowledge gaps based on the actual technological, social and economic context, and the data available. In addition, the case study enables the derivation and assessment of the Data Ecosystem as can be found in Utrecht.

3.2 Case description: the municipality of Utrecht

For this study, in order include as much as possible of the local heating system and to grasp the actual challenges regarding the heat transition and the associated knowledge gaps, the focus is placed on the Municipality of Utrecht in the case

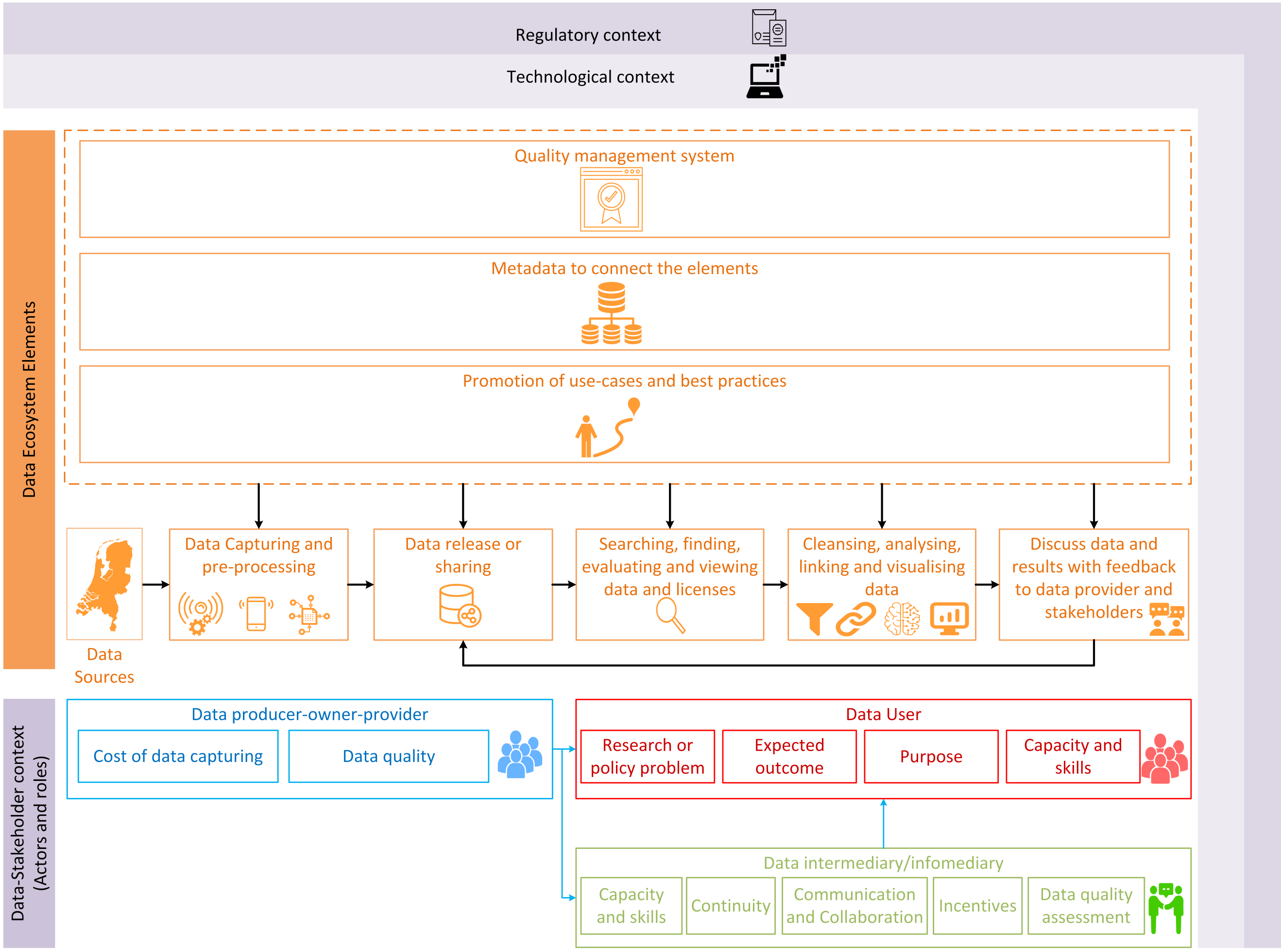

Figure 1. The DE according to [14], complemented with the social layer consisting of the corresponding roles

study. The municipality of Utrecht is located in the Utrecht province in the Netherlands and covers a surface area of 99.21 km^2. With the ambition to be climate neutral by 2030, Utrecht has its work cut out to make drastic changes to its energy landscape in the coming years [60]. In 2018, Utrecht counted 347,483 inhabitants, divided over 178,186 households, in 150,831 dwellings [61].

The building stock in Utrecht is dominated by the flat buildings and row houses with a total share of around 96%. When looking at the building period of the residential building stock in Utrecht, there is a relative over representation of the dwellings originating before 1945, also known as the pre-war dwellings.

In 2017, the residential dwellings in Utrecht consumed a total of 6.39 PJ, of which around 80% of the energy is used as heat for spatial heating, cooking and warm water. The remaining 20% of the energy use entails the electricity to power appliances and lighting. The heat supply is predominantly provided by natural gas, 3.6 PJ - 71%, while District heating accounts for the remaining 1.4 PJ - 29% [62]. Over 58% of the dwellings has energy label C or better.

In the Netherlands, after Rotterdam, Utrecht is the leader in terms of the installed district heating capacity. In 2015 this amounted to 52,800 dwellings connected to DHN developed and exploited by Eneco, and this is expected to increase to 58,000 connected dwellings in 2020 [4]. The most important source of heat for this DHN in Utrecht are the gas fired electricity plants of Lage Weide and Merwedekanaal [4].

Regarding the citizens of Utrecht and their motivation to have a positive contribution in the sustainability of the municipality, a survey in 2016 yielded that, when looking at the measured related to the heat demand, 85% of the respondents have some sort of window insulation method implemented such as double glass. Furthermore, 47% have their walls (partially) insulated, 43% have their roof (partially) insulated, 36% have their floor (partially) insulated, 33% have a water saving shower head and/or tap, 9% have a energy-manager e.g. TOON, 35% have a smart meter, 4% have a heat pump installed, and 1% have PVT panels installed for heating water [63].

3.3 Interview approach

For this study, both open and semi-structured interviews are utilised over two phases of the study as the main means of data collection. In the first phase of problem orientation, open interviews are conducted, to gather empirical information on the perception of the problem, the possible solutions, and the stakeholder field involved. In the phase of the problem analysis and the DE design, semi-structured interviews were conducted to derive the information necessary to answer the research questions. The set of stakeholders resulting from the first round of interviews represent the various elements in an urban thermal energy system and form the pool of entities to be interviewed in the second round. From this stakeholder set, a structured method is applied to select and approach the interviewees for the second round, these interviewees are presented in table A2, while table A1 presents the interviewees from the first round.

3.4 Method of coding and Data analysis

It should be noted that the processing and analysis is aided by the NVivo software package [64]. The data treatment consisted of two coding iterations where the simultaneous coding, or co-occurrence coding method is applied, initially on a predefined code-set based on literature, and second by a code-set enriched with the codes that can be derived from the empirical data. Each iteration is followed by a filtering round to clean up the codes used very rarely, and by merging codes which show significant similarity or a strong correlation.

After the date was coded, the analysis consisted mainly of a hierarchical cluster analysis, operationalised through a coding matrix, for the definition of themes and interrelations between themes [65, 66]. In addition to the definition of themes, the cluster analysis aids the classification of the data, by grouping data together in classes related to various characteristics of the respondents [66]. This enables to say something about how respondent characteristics relate to the data, for instance which challenge or knowledge gaps is commonly mentioned by a certain category of respondents. By doing so, the complexity and variation in the perceptions and attitudes of various actors and stakeholders can be better understood when studying the socio-technical system behind the heat transition.

4. Results

4.1 The socio-technical system

The socio-technical system around the heat transition can be described as very extensive, actors and stakeholders can be divided over roughly 7 categories: citizens, government, government authorities, market (construction, technical installation, energy utility and service, etc.), Intelligence (research and advisory), Real-Estate (developers, intermediates, and housing corporations), and Other (from network operators to financial institutions and local citizen initiatives). This extensive field of actors and stakeholders evolves around an equally extensive field of technologies at different states of maturity, sustainability and affordability, and with different characteristics on the energy source, temperature, and being an individual or collective system. The actor and stakeholder field now comes together with the technological field to make decisions, invest and adopt alternatives for natural gas in the heat supply. This is met with challenges on the uncertainty in the performance of technology and the allocation of cost among public and private parties.

In the actor and stakeholder field, one of the most interesting findings is that many of the actors and stakeholders are still in the process of comprehending the problem of natural-gas in the built environment and searching for their role to realise this. Given that these stakeholders are not yet aware of their role, it is also unclear what resources they have available for the heat transition, while their attitude towards the transition is broadly varying and unstable over the field. It is commonly mentioned that the government should be establishing the facilitating conditions, this will provide the stakeholders of clarity on their role, after which they can proceed to decision-making in the heat transition.

From the DE perspective some clear roles can be distinguished, for instance Stedin as network operator is taking the role as data supplier because their legal status forbids them to add value to the data. The municipality with the heat transition vision, and utility companies are taking the role as data user in supporting the planning and development activities in the heat transition. Kadaster and CBS are establishing themselves as true Data intermediates, whereby the role entails, on the one, hand DE organisational aspects regarding the establishment and maintenance of databases and the facilitation of the stakeholders involved in that process. On the other hand, there are the data aspects where they add value to data and release data for users to process. The role of standards organisations is an important one in DE to ensure data quality and interoperability, however such an authority only exists for government geo-data. For other data types this role is still vacant, yet badly needed.

4.2 The knowledge gaps and data needs

From the empirical data it can be concluded that the stakeholders often address challenges in co-occurrence with knowledge gaps. In other words, many of the challenges in the heat transition are a consequence of lacking or sub-optimal knowledge, entailing inadequate information provision to the decision-makers responsible and accountable for decision-making in the heat transition, and stakeholders affected by these decisions. The identified knowledge gaps can be defined over five main themes: 1) the energy system and environment, 2) the data ecosystem, 3) dwellings and end-users, 4) market and economic aspects, and 5) the decision-making process.

Of these categories, the most knowledge gaps, measured by the variety that can be derived from the empirical data, fall under the themes of 1) the energy system and environment, and 2) the dwellings and end-users. In particular, very little is known about 1) the detailed characteristics of dwellings that impact the potential and costs of retrofit and thermal installation upgrades, 2) the perceptions and attitude, leading to the willingness to participate by the citizens and building owners, and 3) the preferences of stakeholders for natural-gas alternatives and their role in the heat transition.

4.3 Data-base and Data platform inventory

This study produced an extensive, but not exhaustive, inventory of data-bases currently available, with relevance for heat transition decision making. A total of 24 data-bases are presented in this inventory, categorised over: supply side data, demand side data, building stock data and energy statistics data, see figure 3 for a schematic presentation of these data-bases. It can be concluded that the majority of the data openly available nowadays is on the supply side. The infrastructure data for distribution and storage is also openly available. However, on the demand side, including the dwelling characteristics and the end-user characteristics and behaviour, there is little data captured and released as open data. A significant share of the demand data thus remains to be released, while there also is data with great potential which remains to be captured, namely: 1) citizen preferences and attitudes towards the alternatives for natural-gas and retrofit measures, and 2) data on dwelling envelope retrofit and thermal installation measures already conducted in dwellings.

Besides the data-bases, an inventory is produced on the data platforms or portals targeting energy, and in particular heat. This inventory yielded 9 platforms with varying functionality and data feed. See table A3 and A4 for a description on these platforms. It could be concluded that when placing these platforms and data-bases in the context of an ecosystem, there already is a very rich ecosystem. However there is very little known on the links and interaction between the platforms and data-bases. This means that several platforms have redundant functionality, while other platforms complement each other with their functionality. However, because the overview of the ecosystem is poor, opportunities to jointly utilise platforms are left unexploited.

Data-driven strategies have the potential to address these knowledge gaps, however significant barriers are experienced by the stakeholders in the development and execution of effective data-driven strategies in the current DE. The design of the DE2.0, needs to take into account barriers regarding: 1) restricting (privacy) legislation, 2) data ecosystem barriers e.g. difficult and cumbersome data search and acquisition and poor data quality and detail level, 3) stakeholder barriers, e.g. lacking willingness to share data, and 4) high perceived costs.

4.4 Barriers to data-driven approaches in the heat transition

There is a large amount of data currently available or data that could become available in the foreseeable future through, among others, the data platforms [9]. However, at the current stage of data and platform applications for the heat transition, reoccurring challenges and barriers are encountered. In the interviews with stakeholders in the Dutch heat transition, questions addressed the extent to which these stakeholders work with data and data portals or platforms and which challenges and barriers they encounter. The following challenges are encountered in practice:

The interviews with stakeholders in the Dutch heat transition, included questions which address how familiar these stakeholders are with data and data portals or platforms and which challenges and barriers they encounter during the development and execution of the data-driven strategy. From this overview it can be observed that many encountered barriers, overlap with the barriers mentioned in literature as was presented in sub-section 2.1.6. The following barriers are encountered in practice:

- Privacy legislation, namely the "Algemene Verordening Gegevensbeheer" (AVG) as the Dutch implementation of the EU General Data Protection Regulation (GDPR), and the unfamiliarity on how to deal with this AVG and privacy sensitive data of consumers is the most recurring barrier stated by the interviewees, namely 11 out of the 18 interviewees. Privacy legislation is not only impacting the release of data, but also the analysis of data and technologies such as Big and Open Linked Data (BOLD). The application of BOLD, by linking general data to derive individual and richer information is perceived as illegal and thus limits these activities (Business Developer, Eneco, 2018).
- Difficult and lengthy processes to find and access the necessary data via the appropriate platforms or portals is the second most mentioned barrier. There is a lot of time and expertise required to get access to the data and to utilise the data effectively, among others, this is caused by the dispersed distribution of data and inconvenient interfaces. This process is experienced as inconvenient, devious, and complex by several interviewees and for instance HoogravenDuurzaam questions whether the added value of the data and platforms outweighs the hassle (Chairman, HoogravenDuurzaam, 2018). In addition, the data may be subject to restrictions and costs and this is experienced as a barrier (Strategic Adviser, Kadaster, 2018).
- On the data ecosystem, other barriers include the immature state of organisations with regards to data-driven strategies. Organisations such as the municipality of Utrecht, Heijmans and the citizen initiatives are struggling to comprehend the technological complexity and lack the facilitating data ecosystem to support them in developing and executing the data-driven strategies. In the commercial sector, businesses lack the skills and do not have a clear business case with the added value of utilising and sharing data, hence this is a barrier towards the participation of the commercial sector in the Data Ecosystem (Project Director, Heijmans, 2019).
- Poor data quality and incomplete data, e.g. missing entries in the BAG, and poor data-base compatibility when combining data are a significant barrier (Business Developer, Eneco, 2018).
- The most detailed data available as Open Data on plat-

forms and portals are on the regional and municipality level, there is little outreach to the neighbourhood level or lower. The lack of detailed data is thus a barrier.
- Distrust under citizens leads to hesitation and a low willingness to share data, for instance citizens opt to turn off the smart meter or not register PV panels. This is enforced by the fact that it is not always clear what the purpose is of the data, and if citizens are not convinced by the purpose, the willingness to share data is low (Business Developer, Stedin, 2018; Statistical Officer, CBS, 2018). The poor inclusion of the citizens as data source in the data ecosystem is considered a barrier towards fostering the willingness and awareness.
- Legislation vs. policy instruments, there is a mismatch between, on the one hand, what policy instruments are imposing on actors in terms of the data and services they need to deliver, and on the other hand, the restriction posed by legislation such as privacy legislation. Due to the legal position of the network operators they are not allowed to add value to the data they own, e.g. visualising the data is not allowed. However, adding value to data is expected from the network operators in for instance the Regional Energy Strategies. The network operators are thus limited to release only raw open data (Business Developer, Stedin, 2018).
- For organisations looking to organise the Data Ecosytem and improve the release of data, a barrier mentioned is the lack of cooperation by data-owners such as the energy providers (Statistical Officer, CBS, 2018).
- Finally, the high perceived costs of the data-driven strategies necessary for the training and acquisition of skilled labour and the acquisition of the IT infrastructure and software, is commonly mentioned as a barrier withholding the large scale roll-out of data-driven strategies.

With the insights gained in this section on the socio-technical system, the knowledge gaps and the barriers, the following section presents the DE proposed for the heat transition.

5. The proposed DE: elements and roles

In this section the DE is proposed which aims to improve on the current DE, by taking into account the three building blocks: 1) the information needs in the heat transition, 2) the existing and potential data-bases and platforms with the encountered challenges and barriers, and 3) potential technologies to improve on the data capturing, exchange and utilisation. The proposed DE is exhibited in figure 2.

5.1 The DE Elements 2.0

The following paragraphs each elaborate on the DE elements which are added to or adjusted from the DE sketched for Utrecht and to meet in the needs for the heat transition in the Netherlands. This proposal is made subject to an expert validation in which Dr. Anneke M.G. Zuiderwijk-van Eijk, Assistant Professor at the TU Delft and a leading researcher in the field of Open Data and DEs, participated. This validation session yielded issues on the initial design regarding: overlap between elements, lacking visualisations and the ambiguity among terms in the design.

5.1.1 Quality management system

Poor data quality and incomplete data is explicitly stated by several data-users interviewed, as challenge when utilising open data. Furthermore, the quality of insights gained from the current data platforms, is also stated as a challenge, in particular by the Citizen Initiatives. The existence of a quality management system is thus critical for stakeholders to have confidence in the decision support provided by open data driven applications and platforms. In the current Ecosystem, geo-data standards are established by Geonovum, while other data-types lack standards when compared to geo-data. For an overall effective DE, reaching further than geo-data, this quality system should be expanded to all other types of data. Relevant for the heat transition are thus energy potential data, data on technology performance and costs, and data on the consumer's demand and behaviour. A fair share of this data is not geo-data, leaving a significant share of the data without quality standards. Moreover, the data-management system should not be limited to the development of standards, but also to communicate these standards towards the DE participants and, moreover, to monitor and enforce the compliance to these quality standards. For the provision of data to the energy transition the following criteria should be included in the quality management system: reliability, completeness, topicality, continuity, independence, veracity (source integrity and quality), and interoperable data formats [8]. By ensuring these quality factors for the DE, the barriers and challenges encountered in the current ecosystem, as presented in section 4.4, can be taken away.

5.1.2 Meta-data

Meta-data is important for users to gain a thorough understanding on the data and assess whether the data suits their needs. In addition, meta-data may benefit the DE in improved storing, preservation, accessibility, visualisation and interoperability of open data [14]. The three type of meta-data necessary are [67]: 1) descriptive meta-data on the characteristics of the data, enabling convenient data identification and discovery, 2) structural meta-data on the composition of the data-base, enabling transparency and architecture improvement, and 3) administrative meta-data on the intellectual property of data and data archiving, enabling clarity on the conditions for use. In addition to the above-mentioned meta-data, it is proposed to establish meta-data standards on the landscape of data-bases and platforms, and the associated providers and users of those. The term proposed is Network meta-data and this Network meta-data should provide clarity to providers, users and intermediates on the links between data-bases and platforms. Network meta-data may contribute in the interoperability between

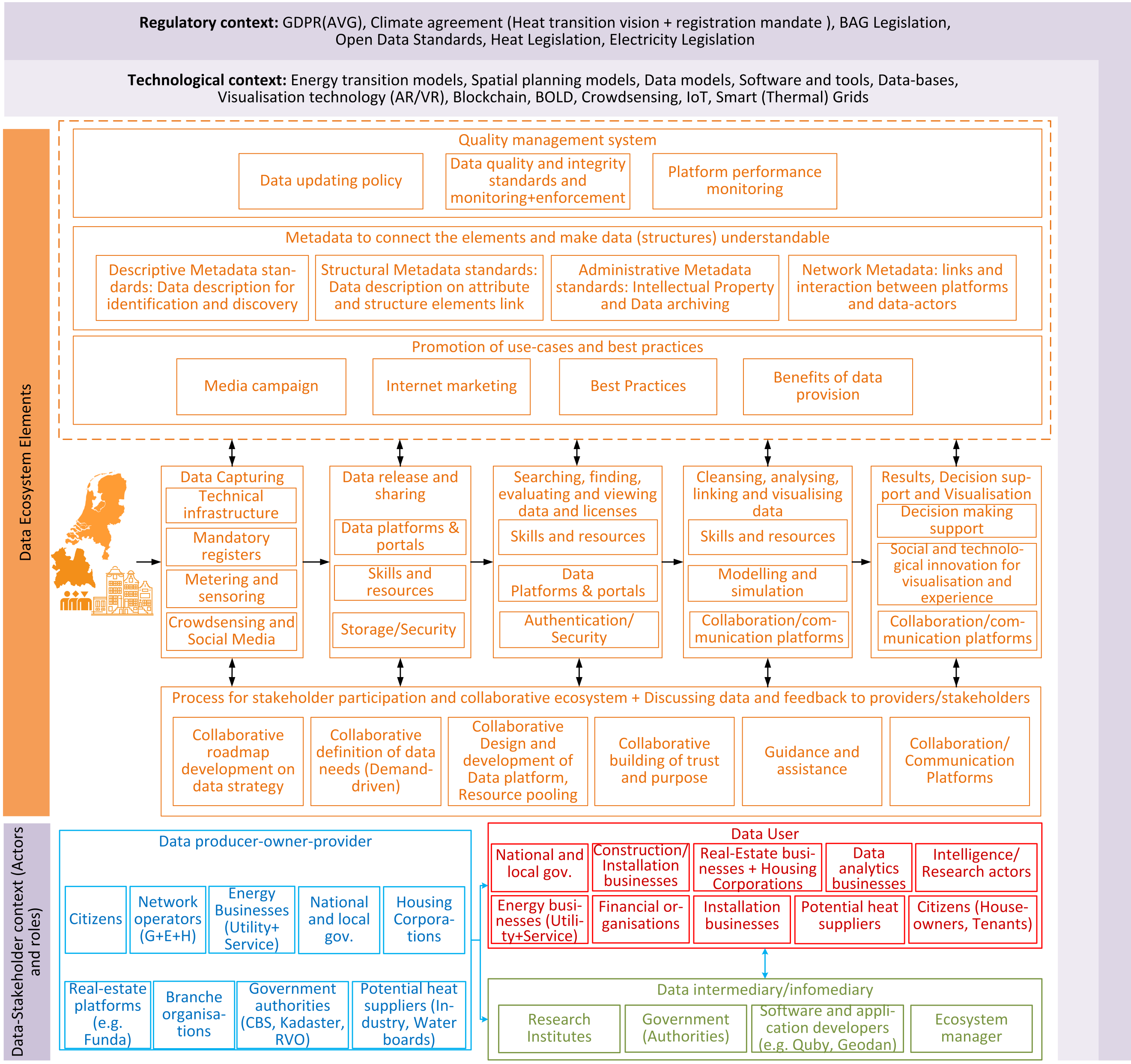

Figure 2. The DE proposed for the heat transition in the Netherlands, adapted from [9]

data-bases and the extent to which platforms can complement each other in terms of functionality and released data. Hence, the challenge of dispersed data and difficulties with finding the appropriate data or platform is directly targeted by this element.

5.1.3 Use-case promotion to improve data and platform use

For open data systems to be effective, a large enough field of users and data providers is necessary. Among the interviews for this study the utilisation of platforms and portals currently available is very low. In order to improve the use of the data and get most out of the potential, Use-case promotion informs and inspires users about data applications. It is proposed to promote best practices and innovative use-cases in which the added value and purpose of data utilisation becomes apparent to the users. Here, users are both parties developing software and tools to add value to data, but also the end-users in the context of the heat transition, the citizens, which do not necessarily directly utilise data, but are a rich source of data. Being aware of the added-value of the data, motivates citizens, and other decision-makers such as municipalities, to utilise data-driven tools and methods to support their decision making. In addition, by clearly promoting the use-cases, the purpose of data utilisation becomes better known by the data-providers, incentivising data sharing.

5.1.4 Process for stakeholder participation and DE collaboration

In the heat transition there is a strong demand for process support to the actors which have to make decisions and investments such as the municipalities, housing corporations and the citizens (Business Developer, Stedin, 2018; Consultant, Overmorgen, 2018; Chairman, HoogravenDuurzaam, 2018; Energy Ambassador, EnergieU, 2018; Strategic Advises, Municipality of Utrecht, 2018). This process support ranges from the technical knowledge, to financial aspects, but also process support regarding the identification and utilisation of data-driven decision support methods and tools. In the proposed DE there is the addition of the stakeholder participation and collaboration element, which is merged with the element mentioned by [14] on the discussion of the data and the provision of feedback to the data providers and stakeholders. This element should account for the inclusion of the data providers and users early in the DE where platforms are designed and developed, all the way to the utilisation of the platforms and the complete life-cycle of the data. On the demand side, i.e. for the data users, the aim is that this inclusion creates 1) platform designs with functionality in which the needs of the users are taken into account, a demand-driven approach and 2) increased familiarity to the platform from the early start. This is expected to result in an increased utilisation of data platforms and portals.

On the supply side, as presented in section 4.4, there is a lack of trust by citizens in authorities and businesses which aim to capture or acquire, and utilise citizen data. This distrust is caused by privacy concerns, but moreover so by the the citizens not being aware of what the purpose is of the data utilisation. Without knowing what the data will be used for, citizens are very hesitant to share data (Product Developer, Stedin, 2018; Product Owner, Geodan, 2018; Statistics Officer, CBS, 2018; Geo-Architect, RVO and Geonovum, 2018). The proposed DE element of ”Process for stakeholder participation and DE collaboration”, aims to built trust and awareness on the supply side, by including (potential) data suppliers in the identification and definition of the added value of data utilisation and the functionality of platforms and portals. Ultimately, the enhanced trust can contribute in increasing the willingness of citizens and other data holders to share data.

Hence, by adopting this inclusive and demand-driven approach, and first determine what the needs are of the stakeholders in the heat transition, and what data is necessary to address those needs, the purpose of the data-driven approach is clear and both data suppliers and users are made aware of the added value of data. This is expected to boost trust, confidence, and the sense of relevance, and ultimately result in an increased willingness to share data by the data suppliers and utilisation of platforms and portals by data user.

Finally, in the current DE there is no presence of data assessment and feedback towards the data providers in order to improve the data quality where needed. The two-way feedback between users and providers over the development and life-cycle of platforms, as included in this element, stimulates the continuous improvement of the data quality, but also platform quality.

5.1.5 Result and data communication and visualisation

It is mentioned by [43] that it is still a challenge to communicate immense amounts of data to stakeholders as comprehensible and actionable information. This aspect was first embedded in the element of cleaning, linking, analysing and visualising data, in the DE framework derived from literature. However, the visualisation or presentation of results is taken separately in proposed DE, because it is believed to be of a different nature and depend on different theories and processes as compared to the more technical aspects of data cleaning, linking and analysis. For instance socio-psychological research is relevant to determine what and how to visualise insights from data to citizens, but less relevant regarding data cleaning, linking and analysis (Product Owner, Geodan, 2018). Do note that these two elements are strongly intertwined, and need to align the activities to each others objectives.

5.2 The DE context

5.2.1 Data-Stakeholder Context

In figure 3 the data-stakeholder context as it currently is, is depicted. In the proposed DE it is plead to enable the role of the DE Manager. The responsibility on the earlier mentioned Data Quality System, can be taken up by the DE Manager which is currently lacking in the DE for the energy transition. Several organisations pick up tasks on the organisation and management of the Ecosystem, but there is no coherence in the DE management, leading to duplicate functionality of data platforms, but also poor interoperability. DE wide meta-data management and coordination of the stakeholder participation process, are also tasks which may be picked up by the DE Manager.

In the national program to improve the information provision to the energy transition, these tasks are proposed to be carried out by a Data Commission, consisting of experts from key stakeholders namely: Ministry of Economic Affairs and Climate, Ministry of Internal Affairs and the Kingdom Relations, CBS, Kadaster, PBL, Rijkswaterstaat and RVO [8]. In this proposal by [8], the sole participants in the Data Commission are public parties and Data Intermediates according to the role definition in this research. However, from the DE approach it is recommended to expand the composition of this DE Manager with representatives from: 1) the data suppliers, e.g. the network operators, 2) the data users, e.g. the energy utility and service businesses, and 3) the standards organisations, i.e. Geonovum. With this composition, the relevant interest are represented in the DE Manager, benefiting comprehensive DE organisation and management.

5.2.2 Regulatory Context

In the regulatory context, the Climate Agreement has an important role. In [8] it is stated that this agreement may provide the legislatory basis to impose certain mandates on the citizens

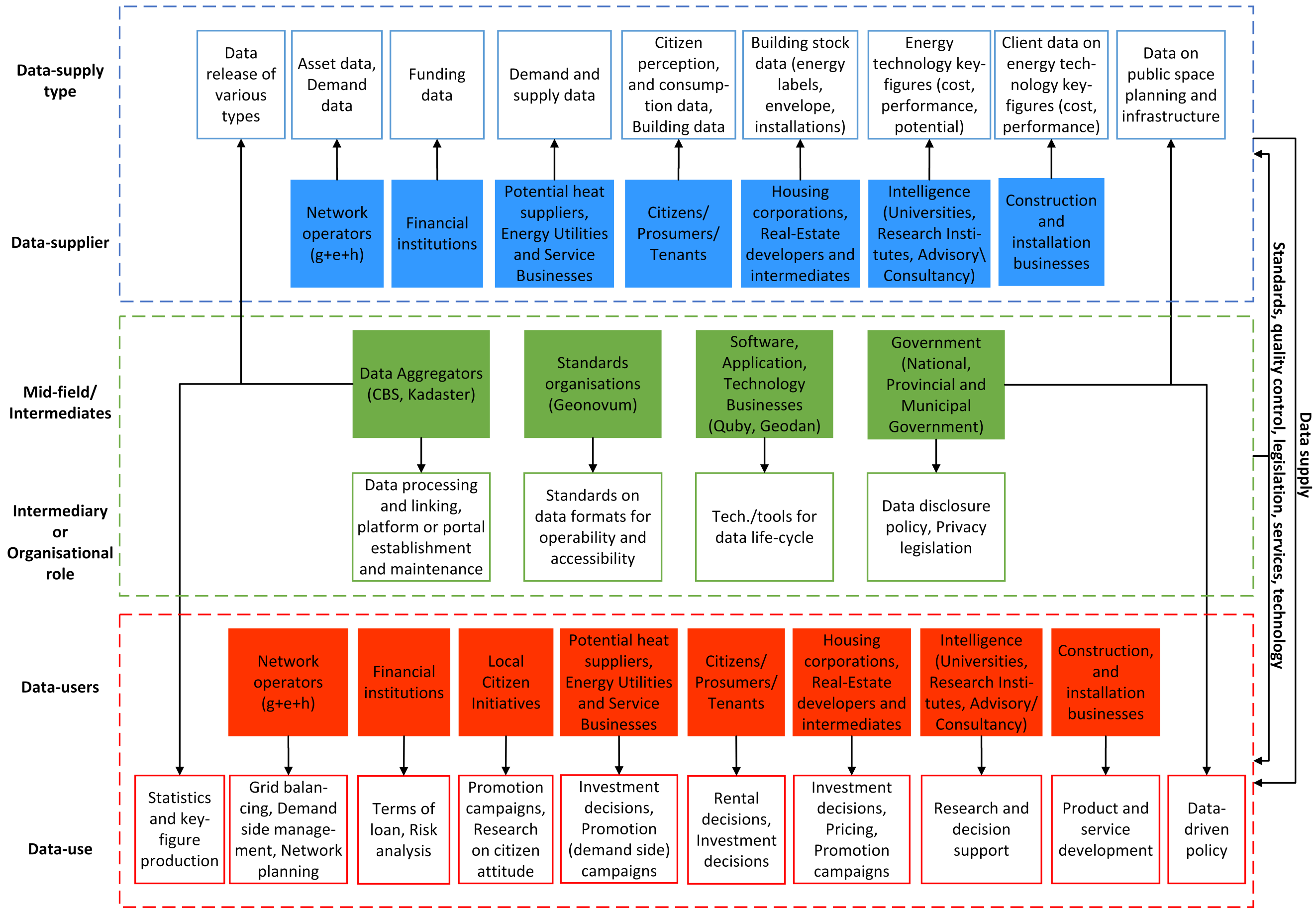

Figure 3. The detailed depiction of the data-stakeholder landscape, adapted from [9]

to register for instance the PV panels on their roof and the insulation measures taken. In the DE this could thus lead to an increase in the data availability by addressing the currently present blind spots or missing data.

In addition, the Environmental Code or ”Omgevingswet” will turn in a significant presence in the DE for the energy transition when enacted by 2021. By means of the anticipated Construction file or ”Bouwdossier”, encompassing a digital archive on each dwelling containing relevant building envelope data, under the Digital Scheme of the Environmental Code or ”Digitaal stelsel Omgevingswet”, novel instruments are provided to enrich the data availability on dwellings. In order for these instruments to be effective, and ensure that the data needs for the heat transition are adequately met by these instruments, it is beneficial that these instruments are developed and implemented within the DE. The development of these instruments in the DE, by involving the DE actors, opens up the knowledge on the data needs in the heat transition encountered and identified by these actors and enables the Construction file to be equipped with the relevant and necessary data variables.

5.2.3 Technological Context

The technological context of the current DE for the heat transition can be described sparsely populated and in its infancy. Currently, energy transition models such as Vesta MAIS utilise available data-bases and the available platforms are casually used by stakeholders in an attempt to visualise and comprehend the challenges in the heat transition. To this point, actual decision making has not been supported adequately in a data-driven way from the technological context. This can be assigned to the technology being in its infancy, governance processes not completely embracing data-driven support, but also the lack of critical data mainly on the demand side. Where the technology is steadily improving, these technologies in the field of data analysis and visualisation are not relevant without the adequate data supply. Hence, the technological context is recommended to be expanded with innovations in the line of Blockchain for secure and reliable data storage and exchange, CrowdSensing for advanced citizen data acquisition, BOLD for enhanced insights from distributed and diverse data, and Artificial and Virtual Reality (AR/VR) to provide end-users with an engaging and familiar experience, rather than conventional data visualisation.

6. Discussion

Previous studies on the knowledge gaps and information provision in the energy transition, see for instance [8], present knowledge gaps which predominantly address the supply side where potential data sources are mapped with regards to the potential and cost factors of sustainable heat sources such as geothermal and solar energy. Accordingly, [8] propose infrastructure and resources developed on this supply side. However, empirical data collected by this study indicates that although the data on the supply side is necessary, it is the demand side which is poorly mapped at this point. This forms an issue as the bottom-up and highly diverse nature of the heat transition requires ample knowledge on the preferences and attitude of end-users. This is in particular the case in a transition, where not every end-user is equally enthusiastic and aware of the needs and urgency behind the transition. Local approaches need to do justice to the willingness of each citizen and stakeholder to participate and have an active role. To this end it is necessary to have up to date information of the preferences of these stakeholders regarding their role and participation.

A fair share of the commonly mentioned challenges and barriers in literature on Big and Open Data in energy systems, are also found to be present in the derived challenges and barriers encountered in the current DE in the Dutch thermal system. For instance recurring barriers such as Inconvenient data access, availability and findability, Inadequate or poor data collection, and Lacking data standards, as presented in leading work for Open and Big Data in energy systems by [40, 29, 41, 42] and [43], are confirmed by the stakeholders. However, new barriers surfaced from the empirical data on the heat transition in the Netherlands. Barriers such as: distrust among citizens leading to poor willingness to share data, contradicting policy requirements and legislation, and the disperse character of data over various platforms and owners, are not found to be reported in literature before.

The reported lack of case studies in the field of DE, is addressed by this research. Here it could be shown that DE elements presented by DE literature are indeed considered as relevant in the case for the heat transition. Moreover, it could be exhibited that although the importance of these elements are acknowledged, real life DEs are lacking in the organisational completeness whereby for the Urban Thermal Energy System DE, it could be found that the element of a quality management system is a critical absentee, resulting in recurring challenges with respect to data quality and integrity.

Finally, among studies proposing elements on effective DE, such as [14, 53, 15] and [54], it can be observed that none address both Big and Open DEs as the focus is on either Big or Open Data. The study presented in this thesis took an integrated approach, looking into both Big and Open Data, due to the nature and scale of data in the Urban Thermal Energy System. As a result the derived DE analysis framework and the proposed DE design form a more comprehensive picture, where both Big and Open data are considered, and where the scope reaches further than technology, towards the Data-Stakeholder context and the Regulatory context.

7. Conclusion and future research

7.1 Conclusion on the main findings

The socio-technical system For the climate goals to be realised, the energy landscape in the built environment will have to drastically change. In this process, challenged by technological novelty, high costs and social complexity, it is desired for the local governments along with the citizens and stakeholders to jointly work towards prioritising, aligning and executing decisions on how to shape the heat transition from the household level, up to the municipal level [2]. However, municipalities, citizens and stakeholders are lacking the specific, accurate and objective knowledge and capacities to effectively execute the complex decision-making on policy and investments in an effective and inclusive heat transition [68, 8].

The socio-technical system around the heat transition can be described as very extensive, actors and stakeholders can be divided over roughly seven categories: citizens, government, government authorities, market (construction, technical installation, energy utility and service, etc.), Intelligence (research and advisory), Real-Estate (developers, intermediates, and housing corporations), and Other (from network operators to financial institutions and local citizen initiatives). This extensive field of actors and stakeholders evolves around an equally extensive field of technologies at different states of maturity, sustainability and affordability, and with different characteristics on the energy source, temperature, and being an individual or collective system. The actor and stakeholder field converges with the technological field to make decisions, invest and adopt alternatives for natural gas in the heat supply. This is met with challenges on the uncertainty in the performance of technology and the allocation of cost among public and private parties.

In the actor and stakeholder field, many of the actors and stakeholders are still in the process of comprehending the problem of natural-gas in the built environment and searching for their role in the envisioned transition. Given that these stakeholders are not yet aware of their role, it is also unclear what resources they have available for the heat transition, while their attitude towards the transition is widely varying and unstable over the field. It is commonly mentioned that the government should be establishing the facilitating conditions, this should provide the stakeholders of clarity on their role, after which they can proceed to decision-making in the heat transition.

Knowledge gaps among actors and stakeholders From the empirical data it can be concluded that the stakeholders often address challenges in co-occurrence with knowledge gaps. In other words, many of the challenges in the heat transition are a consequence of lacking or sub-optimal knowledge, entailing inadequate information provision to the decision-makers responsible and accountable for decision-making in the heat transition, and stakeholders affected by these decisions. The identified knowledge gaps can be defined over five main themes: 1) the energy system and environment, 2) the data ecosystem, 3) dwellings and end-users, 4) market and economic aspects, and 5) the decision-making process.

Of these categories, the most knowledge gaps, measured by the variety that can be derived from the empirical data, fall under the themes of 1) the energy system and environment, and 2) the dwellings and end-users. In particular, very little is known about 1) the detailed characteristics of dwellings that impact the potential and costs of retrofit and thermal installation upgrades, 2) the perceptions and attitude, leading to the willingness to participate by the citizens and building owners, and 3) the preferences of stakeholders for natural-gas alternatives and their role in the heat transition.

the proposed DE 2.0 The main research question of this research targeted *the necessary elements, roles and context in an DE to enable the data-driven support of the local heat transition, by capturing citizen preferences, willingness to participate, and attitude*. The DE 2.0 is the proposed DE which answers this research question. It takes into account the shortcomings of the current DE, the data needs derived from the knowledge gaps identified among the stakeholders in the heat transition, the inventory of available data-bases and platforms, potential technologies, but also the challenges and barriers experienced.

The DE for the heat transition is in its infancy phase at this point with challenges, barriers and shortcomings in all elements as proposed by the DE framework derived from literature and presented in sub-section 2.4. Of the eight elements, the current DE is particularly falling short in 1) the Data discussion and feedback, 2) Meta-data, 3) Use-case promotion and 4) Quality management system. In addition, the data feed is sub-optimal whereby, in particular, little to no data is included for the demand side on the citizens. Currently, energy transition models such as Vesta MAIS utilise available data-bases and the available platforms are casually used by stakeholders in an attempt to visualise and comprehend the challenges in the heat transition. To this point, actual decision-making has not been supported adequately in a data-driven way from the technological context. This can be assigned to the technology being in its infancy, governance and decision-making processes not completely embracing data-driven support, but also the lack of critical data, mainly on the demand side. Where the technology is steadily improving, these technologies in the field of data analysis and visualisation are not relevant without that adequate data supply. In order to establish a DE which effectively supports the decision-making for the heat transition, the lacking data supply on heat demand, dwelling characteristics, and citizen attitude and perceptions requires the main focus. This will require, not only stable and adequate data infrastructure and technologies such as BC, CS and BOLD, but also an adequate quality management system, new roles such as the DE Manager, and a process for stakeholder participation and DE collaboration.

In the DE 2.0 a new element is proposed, namely the process for *stakeholder participation and DE collaboration with data discussion and feedback*. This process aims to contribute in increased involvement of the stakeholders, leading to a better familiarity with the ecosystem, but also fosters trust among data suppliers for an improved willingness to share data. Moreover, this element also inherits data discussion and feedback. By doing this along the DE, first, the quality of the data and infrastructure is continuously assessed and ideally improved, and second, continuous interaction enables a DE which is aware of the specific data needs, and targets those data needs effectively and efficiently.

Finally, new roles are proposed, among others, that of the Data Ecosystem Manager. Not only is that entity responsible for data standards and quality assurance, but also to coordinate the previously mentioned process of stakeholder participation and DE collaboration.

7.2 Limitations and future research

The Data Ecosystem approach is a fairly novel approach which takes a very broad view on objects, actors and interactions to include for and around data-driven strategies. This is a strength of the DE approach, enabling a comprehensive understanding. However, a limitation is that this approach stays on the meta-level. Hence, further research is necessary to operationalise findings from the DE approach on the data infrastructure and activity level.

A limitation of qualitative case studies may be related to the interpretivism, where [59, p. 281] talk about ”person-specific, artistic, private/interpretive acts which cannot be viably verified or replicated by others”. This limitation can be related to the data treatment in this research and the lack of coding verification, due to the fact that the interview data was only coded by the author. Discussions and verification did take place on the definition of codes and the derivation of themes, however the actual assignment of the codes to the interview data was subject to the author's interpretation of the interview data.

The DE framework derived from literature and the application of this framework to design the DE 2.0 did not undergo the thorough validation as would be desired. The validation was twofold. First, both the DE framework and DE 2.0 designs were made subject to expert validation where the completeness and topicality were assessed. Second, the DE framework underwent empirical validation by applying it on the Utrecht case and by assessing in the interviews if the DE elements are acknowledged by the interviewees. This validation can be improved by increasing the participants in the expert validation, and applying the framework to different municipalities to compare the generisability of the Utrecht case.

This study targets the data-driven support in the heat transition and aims to provide an as complete as possible picture on the relevant data types and the technical aspects around the data life-cycle in Urban Thermal Energy Systems. However, it should be mentioned that due to the meta-level of the analysis, the total of data types included in the study is non-exhaustive and subject to fast development.

7.2.1 Recommendations for future research

Given the main findings of this research and its limitations, the following presents recommendations towards future research:

- This research was kept on an exploratory level, due to the lack of quantitative data, no statistical statements can be made regarding the relations between the factors and the relative importance and impact of each factor. It is recommended to extend the research on the willingness to share energy and dwelling data to the statistical level and quantify the importance of the factors.
- The technologies explored to have potential in closing the gaps between the current and desired DE, e.g. BC, BOLD, and CS, are found to still pose major barriers in their current state of maturity. Hence, it is recommended to initiate research on the implementation of these technologies, for the socio-technical needs and characteristics of the heat transition, regarding both the technological and operational aspects, as well as the societal embeddedness.
- Businesses encounter significant challenges in the energy sector to quantify the value of releasing their data, hence cannot develop profitable businesses propositions to release relevant data as a commercial asset. Subsequently, it is recommended to study possible strategies and business cases for commercial parties in the energy sector on how data can become a business asset, in order to incentivise the commercial parties to release this data, enrich the DE, and improve the information provision.
- Finally, another field of research can be on how to embed the DE findings in policy making, in other words, research on policy measures which enable the realisation of the proposed DE 2.0 in the operational spheres of the Dutch heat transition. Relevant aspects here are: models for subsidies or alternative incentives, organisational models for the DE manager, and governance models for common data infrastructure.

References

[1] United Nations. *Paris Agreement*. United Nations, 2015.

[2] Klimaatberaad. *Voorstel voor hoofdlijnen van het Klimaatakkoord*. Sociaal-Economische Raad, 2018.

[3] Energie Beheer Nederland. Energie in nederland: Huishoudens, 2016.

[4] Marijke Mankveld, Robin Matton, Reinoud Segers, Jurrien Vroom, and Annemiek Kremer. *Monitoring warmte 2015*. ECN, 2017.

[5] Robin Niessink and Hilke Rösler. *Developments of Heat Distribution Networks in the Netherlands*. ECN, 2015.

[6] Jasminka Young and Marleen Brans. Analysis of factors affecting a shift in a local energy system towards 100% renewable energy community. *Journal of Cleaner Production*, 169:117–124, 2017.

[7] Vanesa Castan Broto. Urban governance and the politics of climate change. *World development*, 93:1–15, 2017.

[8] G Dekker, K Keller, O Swertz, J Vroom, M Mink, A van den Hoek, L Noordegraaf, N Hoogervorst, J Matthijsen, J Baltussen, L Dijkshoorn, and G Nijsink. Vivet: Voorstellen om de informatievoorziening energietransitie te verbeteren. Technical report, 2019.

[9] Devin Diran. Data-driven decision making towards inclusive and sustainable urban heating in the netherlands: the data ecosystem approach, 2019.

[10] Stelios Grafakos, Alexandros Flamos, and Elena Enseñado. Preferences matter: A constructive approach to incorporating local stakeholders' preferences in the sustainability evaluation of energy technologies. *Sustainability*, 7(8):10922–10960, 2015.

[11] Binod Prasad Koirala, Yashar Araghi, Maarten Kroesen, Amineh Ghorbani, Rudi A Hakvoort, and Paulien M Herder. Trust, awareness, and independence: Insights from a socio-psychological factor analysis of citizen knowledge and participation in community energy systems. *Energy research & social science*, 38:33–40, 2018.

[12] Bernhard J Kalkbrenner and Jutta Roosen. Citizens' willingness to participate in local renewable energy projects: The role of community and trust in germany. *Energy Research & Social Science*, 13:60–70, 2016.

[13] Stefanie Hatzl, Thomas Brudermann, Kathrin Reinsberger, and Alfred Posch. Do public programs in 'energy regions' affect citizen attitudes and behavior? *Energy Policy*, 69:425–429, 2014.

[14] Anneke Zuiderwijk, Marijn Janssen, and Chris Davis. Innovation with open data: Essential elements of open data ecosystems. *Information Polity*, 19(1, 2):17–33, 2014.

[15] Yuri Demchenko, Cees De Laat, and Peter Membrey. Defining architecture components of the big data ecosystem. In *2014 International Conference on Collaboration Technologies and Systems (CTS)*, pages 104–112. IEEE, 2014.

[16] Marijn Janssen and George Kuk. Big and open linked data (bold) in research, policy, and practice. *Journal of Organizational Computing and Electronic Commerce*, 26(1-2):3–13, 2016.

[17] Aggeliki Androutsopoulou and Yannis Charalabidis. A framework for evidence based policy making combining big data, dynamic modelling and machine intelligence. In *Proceedings of the 11th International Conference on Theory and Practice of Electronic Governance*, pages 575–583. ACM, 2018.

[18] Amanda Clarke and Helen Margetts. Governments and citizens getting to know each other? open, closed, and big data in public management reform. *Policy & Internet*, 6(4):393–417, 2014.

[19] Foster Provost and Tom Fawcett. Data science and its relationship to big data and data-driven decision making. *Big data*, 1(1):51–59, 2013.

[20] Martijn Poel, Eric T Meyer, and Ralph Schroeder. Big data for policymaking: Great expectations, but with limited progress? *Policy & Internet*, 10(3):347–367, 2018.

[21] Martijn Poel, R Schroeder, J Treperman, M Rubinstein, E Meyer, B Mahieu, and M Svetachova. Data for policy: A study of big data and other innovative data-driven approaches for evidence-informed policymaking. *Report about the State-of-the-Art. Amsterdam: technopolis, Oxford Internet Institute, Center for European Policy Studies*, 2015.

[22] Kaile Zhou, Chao Fu, and Shanlin Yang. Big data driven smart energy management: From big data to big insights. *Renewable and Sustainable Energy Reviews*, 56:215–225, 2016.

[23] Daniel Dietrich, J Gray, T McNamara, A Poikola, R Pollock, J Tait, and T Zijlstra. Open data handbook documentation. *Open Knowledge Foundation, Cambridge, UK*, 2012.

[24] J Visser. Barriers in open energy data: An exploratory study into open energy data barriers and their mitigation strategies. Master's thesis, Delft University of Technology, 2015.

[25] Frederika Welle Donker and Bastiaan van Loenen. How to assess the success of the open data ecosystem? *International Journal of Digital Earth*, 10(3):284–306, 2017.

[26] Maarja Toots, Keegan McBride, Tarmo Kalvet, and Robert Krimmer. Open data as enabler of public service co-creation: exploring the drivers and barriers. In *E-Democracy and Open Government (CeDEM), 2017 Conference for*, pages 102–112. IEEE, 2017.

[27] Michel Noussan, Matteo Jarre, and Alberto Poggio. Real operation data analysis on district heating load patterns. *Energy*, 129:70–78, 2017.

[28] Evelina Di Corso, Tania Cerquitelli, Marco Savino Piscitelli, and Alfonso Capozzoli. Exploring energy certificates of buildings through unsupervised data mining techniques. In *2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, pages 991–998. IEEE, 2017.

[29] Paul A Mathew, Laurel N Dunn, Michael D Sohn, Andrea Mercado, Claudine Custudio, and Travis Walter. Big-data for building energy performance: Lessons from assembling a very large national database of building energy use. *Applied Energy*, 140:85–93, 2015.

[30] Anne Fleur van Veenstra and Bas Kotterink. Data-driven policy making: The policy lab approach. In *International Conference on Electronic Participation*, pages 100–111. Springer, 2017.

[31] European Commission. Quality of public administration: A toolbox for practitioners, 2017.

[32] Danny YC Wang, Amy JC Trappey, Charles V Trappey, SJ Li, et al. Intelligent and concurrent analytic platform for renewable energy policy assessment using open data resources. In *MOVING INTEGRATED PRODUCT DEVELOPMENT TO SERVICE CLOUDS IN THE GLOBAL ECONOMY*, pages 781–789, 2014.

[33] Stefan Pfenninger, Joseph DeCarolis, Lion Hirth, Sylvain Quoilin, and Iain Staffell. The importance of open data and software: Is energy research lagging behind? *Energy Policy*, 101:211–215, 2017.

[34] TV Ramachandra and BV Shruthi. Spatial mapping of renewable energy potential. *Renewable and sustainable energy reviews*, 11(7):1460–1480, 2007.

[35] Andy van den Dobbelsteen, Rob Roggema, Nico Tillie, Siebe Broersma, Michiel Fremouw, and Craig Lee Martin. Urban energy masterplanning—approaches, strategies, and methods for the energy transition in cities. In *Urban Energy Transition*, pages 635–660. Elsevier, 2018.

[36] Rajat Gupta and Matt Gregg. Local energy mapping using publicly available data for urban energy retrofit. In *Building Information Modelling, Building Performance, Design and Smart Construction*, pages 207–219. Springer, 2017.

[37] Anuj Karkare, Abhimanyu Dhariwal, Sumedh Puradbhat, and Monika Jain. Evaluating retrofit strategies for greening existing buildings by energy modelling & data analytics. In *Intelligent Green Building and Smart Grid (IGBSG), 2014 International Conference on*, pages 1–4. IEEE, 2014.

[38] Clayton Miller. Predicting success of energy savings interventions and industry type using smart meter and retrofit data from thousands of non-residential buildings. In *Proceedings of the 4th ACM International Conference on Systems for Energy-Efficient Built Environments*, page 17. ACM, 2017.

[39] Juan Ribeiro Reis, José Viterbo, and Flavia Bernardini. A rationale for data governance as an approach to tackle recurrent drawbacks in open data portals. In *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age*, page 73. ACM, 2018.

[40] Lucy Linder, Damien Vionnet, Jean-Philippe Bacher, and Jean Hennebert. Big building data-a big data platform for smart buildings. *Energy Procedia*, 122:589–594, 2017.

[41] Özgür Göçer, Ying Hua, and Kenan Göçer. A bim-gis integrated pre-retrofit model for building data mapping. In *Building Simulation*, volume 9, pages 513–527. Springer, 2016.

[42] Yixing Chen, Tianzhen Hong, and Mary Ann Piette. Automatic generation and simulation of urban building energy models based on city datasets for city-scale building retrofit analysis. *Applied Energy*, 205:323–335, 2017.

[43] Christoph F Reinhart and Carlos Cerezo Davila. Urban building energy modeling–a review of a nascent field. *Building and Environment*, 97:196–202, 2016.

[44] Marcelo Iury S Oliveira, Glória de Fátima Barros Lima, and Bernadette Farias Lóscio. Investigations into data ecosystems: a systematic mapping study. *Knowledge and Information Systems*, pages 1–42, 2019.

[45] Teresa M Harrison, Theresa A Pardo, and Meghan Cook. Creating open government ecosystems: A research and development agenda. *Future Internet*, 4(4):900–928, 2012.

[46] Rufus Pollock. Building the (open) data ecosystem. *Open knowledge foundation Blog*, 31, 2011.

[47] Barbara Ubaldi. Open government data. 2013.

[48] Maximilian Heimstädt, Fredric Saunderson, and Tom Heath. Conceptualizing open data ecosystems: A timeline analysis of open data development in the uk. In *Conference for E-Democracy and Open Governement*, page 245. sn, 2014.

[49] Li Ding, Timothy Lebo, John S Erickson, Dominic DiFranzo, Gregory Todd Williams, Xian Li, James Michaelis, Alvaro Graves, Jin Guang Zheng, Zhenning Shangguan, et al. Twc logd: A portal for linked open government data ecosystems. *Web Semantics: Science, Services and Agents on the World Wide Web*, 9(3):325–333, 2011.

[50] Tim Davies. Open data: infrastructures and ecosystems. *Open Data Research*, 2011.

[51] Anneke Zuiderwijk, Marijn Janssen, Geerten van de Kaa, and Kostas Poulis. The wicked problem of commercial value creation in open data ecosystems: Policy guidelines for governments. *Information polity*, 21(3):223–236, 2016.

[52] Francois Van Schalkwyk, Michelle Willmers, and Maurice McNaughton. Viscous open data: The roles of intermediaries in an open data ecosystem. *Information Technology for Development*, 22(sup1):68–83, 2016.

[53] Eunice Mercado-Lara and J Ramon Gil-Garcia. Open government and data intermediaries: the case of aiddata. In *Proceedings of the 15th Annual International Conference on Digital Government Research*, pages 335–336. ACM, 2014.

[54] Dong-Hee Shin and Min Jae Choi. Ecological views of big data: Perspectives and issues. *Telematics and Informatics*, 32(2):311–320, 2015.

[55] Marc van den Homberg and Iryna Susha. Characterizing data ecosystems to support official statistics with open mapping data for reporting on sustainable development goals. *ISPRS International Journal of Geo-Information*, 7(12):456, 2018.

[56] Iryna Susha, Marijn Janssen, and Stefaan Verhulst. Data collaboratives as a new frontier of cross-sector partnerships in the age of open data: taxonomy development. 2017.

[57] Edward Curry. The big data value chain: definitions, concepts, and theoretical approaches. In *New horizons for a data-driven economy*, pages 29–37. Springer, Cham, 2016.

[58] Robert K Yin. *Case study research and applications: Design and methods*. Sage publications, 2017.

[59] Matthew B Miles and A Michael Huberman. *Qualitative data analysis: An expanded sourcebook*. sage, 1994.

[60] Gemeente Utrecht. De energie van de stad 2017, 2018.

[61] Waarstaatjegemeente.nl. Bouwen en wonen: Utrecht, 2019.

[62] Klimaatmonitor. Totaal energiegebruik woningen, n.d.

[63] Gemeente Utrecht. Utrecht monitor: Energie, 2018.

[64] QSR International. Nvivo, 2019.

[65] David Henry, Allison B Dymnicki, Nathaniel Mohatt, James Allen, and James G Kelly. Clustering methods with qualitative data: a mixed-methods approach for prevention research with small samples. *Prevention science*, 16(7):1007–1016, 2015.

[66] Laura Macia. Using clustering as a tool: Mixed methods in qualitative data analysis. *The Qualitative Report*, 20(7):1083–1094, 2015.

[67] Jenn Riley. *Understanding Metadata: What is Metadata, and what is it For?* NISO Press, 2017.

[68] Cameron Roberts and Frank W Geels. Conditions and intervention strategies for the deliberate acceleration of socio-technical transitions: lessons from a comparative multi-level analysis of two historical case studies in dutch and danish heating. *Technology Analysis & Strategic Management*, pages 1–23, 2019.

Appendices

Table A1. Interviews in the context of the pre-problem analysis

#	Role and Subject	Organisation	Stakeholder type	Interview type
1	Researcher - The Policy Lab at TNO	TNO	Research Institute	face to face
2	Professor and Researcher - Regulating Energy Markets	TNO and University of Amsterdam	Research Institute	face to face
3	Policy Officer - The energy transition in The Hague	Municipality of The Hague	Municipality	face to face

Table A2. Interviews for the problem analysis and ecosystem definition

#	Role and subject	Organisation	Stakeholder type	Interview type
1	Chairman - Coordinating and stimulating sustainable initiatives in the Hoograven district	HoogravenDuurzaam	Local citizen initiatives	Face to face
2	Initiator and Chairman - Coordinating and stimulating sustainable initiatives in the Scheveningen district	Gasvrij Scheveningen and VNG	Local citizen initiatives	Face to face
3	Energy Ambassador - Gathering and answering questions from citizens and providing technical support	EnergieU	Energy corporation	(Video)call
4	Business Developer - Developing Heating solutions for residential purposes such as District heating networks and Heat pumps	Eneco	Energy utility company	Face to face
5	Director - Leading the development of integrated sustainable concepts for e.g the residential sector	HOMIJ DEC	Installation company	(Video)call
6	Head of Energy and Project Director - Establishing the new Energy department and leading the Hart van Zuid project	Heijmans	Construction company	Face to face
7	Senior Advisor and Coordinator Technology - Organisation of the sustainability of the building stock	Mitros	Housing corporation	Face to face
8	Product Developer - Developing data-driven products and services within the network company and towards clients	Stedin	Network company	Face to face
9	Consultant and Product Owner - Developing the Geomagine product for VR and AR for urban spatial applications	Geodan	Data product and service provider	(Video)call
10	Innovation-manager Sustainability - Developing new financial products and services around sustainability	Volksbank	Financial sector	(Video)call
11	Strategic Advisor - Developing new and data-driven approaches for the heat transition in Utrecht	Municipality of Utrecht	Municipality	(Video)call
12	Geo-architect RVO and adviser PDOK - Data management at RVO and technical advise on the PDOK development	RVO and Geonovum	Government Authority	Face to face
13	Statistical Officer - Producing statistics on the energy system	CBS	Government Authority	Face to face
14	Product Developer - Innovative data applications for data-driven statistics	CBS (Center for Big Data Statistics)	Government Authority	Face to face
15	Strategic Adviser - Development of the public data release by Kadaster and advising on the role of energy	Kadaster	Government Authority	(Video)call
16	Product owner PDOK - Development of the public data release by Kadaster	Kadaster	Government Authority	(Video)call
17	Consultant - Advising public and private parties on the natural-gas-free heat transition in a data-driven style	Overmorgen	Consultancy or Advisory firm	(Video)call
18	Researcher - Research on deriving and coping with the preferences of citizens in the heat transition	TU Delft	Research & Education Institute	Face to face

Figure A1. The overview of data sources (current and potential) with relevance for the heat transition, own figure

Table A3. An overview of data-platforms or portals which can be utilised for urban thermal energy system decision making in the Netherlands, clicking on the platform name directs the reader to the online resource

Data-platform/portal	Theme	Description
PICO - Geodan, TNO, Alliander, Ecofys, ESRI Nederland and NRG031	Energy use, Buildings, Spatial area, Energy potential, Sustainable generation, Energy efficiency, Infrastructure, Spatial planning	PICO provides information on the energy use up to the local level and identifies where opportunities and potential lies to best save energy or generate locally
Warmteatlas - RVO	Infrastructure, Emissions, Energy supply, Energy potential, Spatial areas	The Warmteatlas from RVO presents heat demand and supply related information on geographic maps. On the supply side this is: locations potentially suitable for heating- and cooling storage, deep geothermal, biomass and waste heat. The demand side presents e.g, gas consumption
Nationale EnergieAtlas - National Institute for Public Health and the Environment	Energy use, Sustainable generation, Infrastructure, Energy potential, Spatial area mapping, Spatial planning	The National EnergieAtlas is the information portal from the national government which maps current non-renewable and renewable energy generation. In addition, insights are provided on the potential of an area to become sustainable. Kadaster data on property ownership, potential NOM dwellings and governmental buildings, is included.
Klimaatmonitor- Rijkswaterstaat	Emissions, Energy use, Renewable energy, Labour and investments, Residential buildings, Service and utility buildings, Mobility, Industry and agriculture, Infrastructure, Social characteristics	The Klimaatmonitor by Rijkswaterstaat is an extensive platform with dashboards on mainly energy related aspects, but in addition it is enriched with a variety of underlying data on the environmental, societal and economic aspects of areas
CBS in uw buurt - CBS	Fossil energy (natural gas, coal and oil) delivery, Electricity and Heat use, Renewable energy generation	CBS in uw buurt is the digital portal which maps CBS Statline data geographically on the neighbourhood level.
PDOK Platform and Viewer - Kadaster	Energy use, Sustainable generation, Energy potential, Spatial area mapping, Spatial planning, (Subsurface) infrastructure, Hydrolaugical system	PDOK, or Public-service on the map, is a national geographical data portal or platform which combines, releases and visualises the geo-data-bases from the geo-register (Kadaster), BAG, AHN, Ministry of internal affairs and kingdom relations, Ministry of economic affairs, CBS, National Hydrological Instrumentarium and ”Het Waterschaps Huis”.
Energy atlas or platforms of provinces and municipalities e.g. Warmte transitie Atlas Zuid Holland and Lokale Energie Etalage	Energy use, Sustainable generation, Infrastructure, Energy potential, Spatial area mapping, Spatial planning	These portals have comparable functionality as the above-mentioned platforms. However, the focus is on the specific area (province or municipality) for which the platform is built and maintained. Often these local platforms are enriched with more detailed and accurate local data relative to e.g. the platforms from the national government.
Open Data Portals or Platforms of national and local governments e.g. Utrecht Open Data - Municipality of Utrecht' Dutch National Data portal - the National Government and Waarstaatjegemeente.nl - VNG	Ranging over various variables on society, economy and the environment	These portals focus on the release of the open data and less on the visualisation and analysis of that data. The data utilised for the Energy platforms and the visualisation on those platforms is often also released over the open data portals in raw format. The data portal of the national government released 12,397 data-sets up to now, while the local portal in Utrecht has released 605 data-sets
BAG viewer - Kadaster	Address location, Building function, Building surface, Building contour, Built year	The BAG Viewer presents BAG data online, both graphically and on a map. Different layers can be selected depending on the zoom level. The BAG viewer is not meant to extract large portions of BAG data, for this more suitable API's, such as BAG Extract, are developed

Table A4. An overview of the data-platform or portal functionality

Data-platform/portal	Level	Geo-visualisation	Monitoring	Bench-marking	Potential	Download
PICO	national, regional, local, neighbourhood	yes	yes	no	yes	maps of visualised area with data layers
Warmteatlas	national, regional, local, neighbourhood	yes	no	no	yes	maps of visualised area with data layers
Nationale EnergieAtlas	national, regional, local, neighbourhood	yes	no	no	yes	no
Klimaatmonitor	national, regional, local, neighbourhood	yes (limited)	yes	yes	no	yes (CSV, PDF, PPT, GIF, Open Office etc.)
CBS in uw buurt	national, regional, local, neighbourhood	yes	yes	no	no	maps of visualised area with data layers
PDOK Platform and Viewer	national, regional, local, neighbourhood	yes	yes	no	yes	data-sets + maps of visualised area with data layers
Energy atlas or platforms of provinces and municipalities	national, regional, local, neighbourhood	yes	yes	no	no	varies per platform
Open Data Portals or Platforms of national and local governments	national, regional, local, neighbourhood	yes	yes	no	no	download data-sets
BAG viewer	national, regional, local, neighbourhood, individual	yes	no	no	no	maps of visualised area with data layers

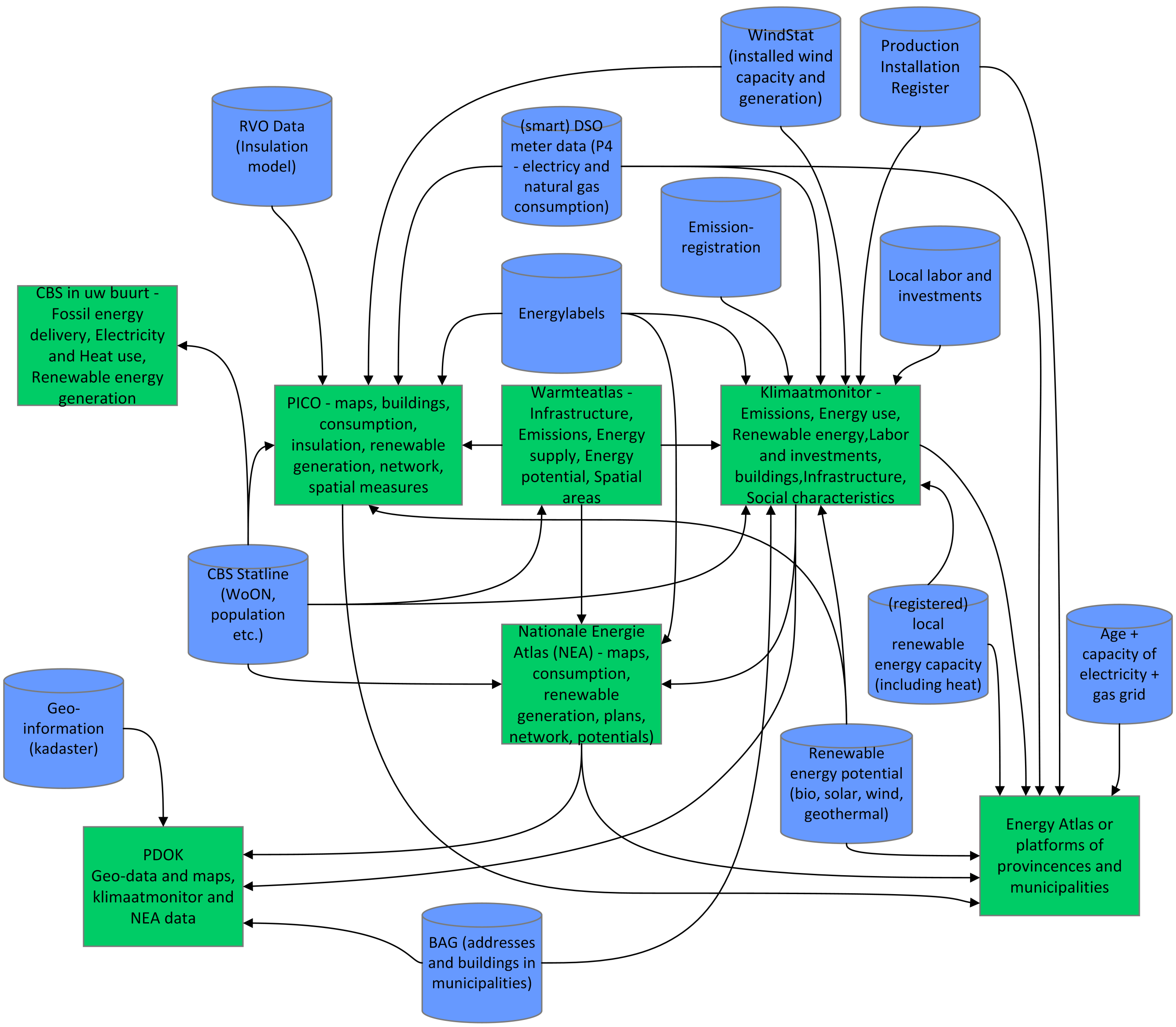

Figure A2. Visualised how the data platforms utilise the data from figure A1, own figure