

**A data ecosystem for data-driven thermal energy transition
Reflection on current practice and suggestions for re-design**

Diran, Devin; Hoppe, Thomas; Ubacht, Jolien; Slob, Adriaan; Blok, Kornelis

DOI

[10.3390/en13020444](https://doi.org/10.3390/en13020444)

Publication date

2020

Document Version

Final published version

Published in

Energies

Citation (APA)

Diran, D., Hoppe, T., Ubacht, J., Slob, A., & Blok, K. (2020). A data ecosystem for data-driven thermal energy transition: Reflection on current practice and suggestions for re-design. *Energies*, 13(2), Article 444. <https://doi.org/10.3390/en13020444>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Article

A Data Ecosystem for Data-Driven Thermal Energy Transition: Reflection on Current Practice and Suggestions for Re-Design

Devin Diran ¹, Thomas Hoppe ^{2,*} , Jolien Ubacht ² , Adriaan Slob ¹ and Kornelis Blok ²

¹ Department of Strategy & Policy, Netherlands Organisation for Applied Scientific Research TNO, Anna van Buerenplein 1, 2595 DA The Hague, The Netherlands; devin.diran@tno.nl (D.D.); adriaan.slob@tno.nl (A.S.)

² Faculty of Technology, Policy and Management (TPM), Delft University of Technology, Jaffalaan 5, 2628 BX Delft, The Netherlands; J.Ubacht@tudelft.nl (J.U.); K.Blok@tudelft.nl (K.B.)

* Correspondence: T.Hoppe@tudelft.nl; Tel.: +31-15-278-2783

Received: 28 November 2019; Accepted: 10 January 2020; Published: 16 January 2020



Abstract: The transition towards low-carbon thermal energy systems requires solid information provision to support both public and private decision-making, which is future proof and optimal in the context of the system dependencies. We adopt a data ecosystem approach to answer the following research question: How can a data ecosystem be analyzed and developed to enable the data-driven support of the local thermal energy transition, by capturing both social and technical aspects of the urban thermal energy system? A case study research design of the Netherlands, with an embedded case of the city of Utrecht therein, was used, including data collection involving 21 expert interviews representing a diversity of stakeholders, and qualitative data analysis using NVivo version 10. The data ecosystem includes the necessary elements, roles, and context for decision makers in a local heat transition and captures the social as well as technical aspects of an urban thermal energy system. Assessment of the data ecosystem pertaining to thermal heat transition in the city of Utrecht shows that it is still in its infancy phase, with challenges, barriers, and shortcomings in all its key elements. We present suggestions for the (re-)design of an inclusive and holistic data ecosystem that addresses the current shortcomings.

Keywords: data ecosystem; open data; sustainable energy; thermal energy system; energy transition

1. Introduction

The challenge posed by climate change to society can only be adequately addressed by drastically reducing human-induced greenhouse gas (GHG) emissions [1]. The Netherlands is on the verge of implementing its “Klimaatakkoord” (Climate Agreement). This governance framework aims to accelerate the energy transition by proposing and enabling the realization of measures to reduce GHG emission by 49% in 2030 relative to 1990 [2].

For households in the Netherlands, the generation of heat for a variety of purposes, such as space heating, domestic hot water, and cooking, contributes to over 82% to the final energy consumption and is largely dependent on natural gas [3–5]. From February 2018 onwards, the Dutch national government aims to end the use of natural gas in the urban environment. To establish this discontinuation and the required large-scale uptake of sustainable heating systems, local governments will have a pivotal role, together with the building owners and occupants [2].

In the planning and realization of sustainable urban heating systems to replace the natural gas grid, 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 considered decisive in

a transition that is intrusive and dependent on investments and behavioral change from a wide range of stakeholders. Because of the great share of energy consumption for thermal applications in urban neighborhoods, this paper focuses on the provision of sustainable heat to households at low cost and in a manner that is reliable and natural gas free. The increasing decentralization of the energy system due to a growing variety of sustainable sources (e.g., wind, solar, geothermal, or bio-gas) that needs to fit with the local technical and social conditions is thereby taken into account [6]. Here, it should be noted that the “technical context” of the present study pertains to the urban thermal supply chain, i.e., the natural gas supply chain (from resource suppliers and retailers, via network operators to end consumers) or alternatively a district heating supply chain (e.g., from residual heat supplier via the district heating network operator to the end consumer of heat).

To effectively fill in the novel responsibilities and tasks to kick-off the energy transition at the pace desired in the Climate Agreement, Dutch local governments and stakeholders are expected to innovate and become creative with the existing decision-making mechanisms [2,7]. This requires a sound knowledge base to support transparent, effective, and legitimate decision-making. However, current information provision is often inadequate, and many knowledge gaps are left unaddressed. A recent study presenting a data strategy for the information provision in the Dutch energy transition yielded a set of knowledge gaps. However, this study was predominantly focused on the supply side of energy [8]. The knowledge needs of energy system actors regarding the demand side, and the knowhow support to citizens, were not adequately addressed yet. Another study addressing information provision in the thermal energy transition revealed that it is predominantly the demand side of the energy system that entails most of the knowledge gaps [9]. This pertains to detailed, accurate, and topical information about citizens regarding their consumption behavior, attitude, willingness to participate, and preferences towards technologies and solutions [9–13]. This paper elaborates on the end-user side of the thermal energy system and how to effectively, reliably, and frequently capture and process data to develop appropriate strategies for citizen engagement in the sustainable heat transition.

This points to shortcomings in what is called a data ecosystem, a concept that refers to an overview of the relationships between data users, data providers, tools, and the data infrastructure. Such a data ecosystem can support decision-making processes [14] by facilitating access to sharing and using data [15], and by giving insight regarding the relative positions of the actors, such as data providers, users, sources, and resources in the ecosystem [16]. The data ecosystem approach matches the rise of big and open data to enhance society, affecting all aspects of human activity, by incentivizing citizen participation, transparency, economic growth, and innovation [17]. Technological advances in big and open data are driving the exponential increase in the volume of data. In line with the immense growth in the available data, Information and Communication Technology (ICT) infrastructures and technologies for data analytics have been introduced and developed at an unprecedented pace to exploit its potential. Data facilities and infrastructures are developed and maintained by many parties and utilized by many users [17,18]. They are part of a wider data ecosystem in which each instrument or tool can add value as part of the information puzzle. The data ecosystem of the thermal energy transition may lay the foundation for the provision of knowledge for decision-making among all stakeholders involved. To assess how a data ecosystem can support the thermal energy transition in the Netherlands we answered the following research questions in this paper: How can a data ecosystem be developed to enable the data-driven support of the local thermal energy transition, by capturing both social and technical aspects of the urban thermal energy system?

In Section 2, we provide a literature background on the concept of data ecosystems to develop a general data ecosystem framework. Next, in Section 3, we elaborate our research approach to apply the general data ecosystem towards a “data ecosystem 2.0” specifically for the Dutch heat transition. Subsequently, in Section 4, we present the results of the empirical analysis. These findings are used to evaluate the data ecosystem and to formulate additional elements, which are presented in Section 5, leading to the design of a data ecosystem 2.0 suited for the Dutch heat transition. Finally, Section 6 discusses the results and the paper ends with conclusions in Section 7.

2. Literature Research on the Data Ecosystem Approach

Based on a literature review, in this section, we introduce the data ecosystem approach along with the concepts of big and open data. From the data ecosystem literature, a general data ecosystem framework is derived as a first step towards the design of a data ecosystem framework for the Dutch thermal energy transition.

2.1. Data Ecosystem Conceptualization and Definitions

Following the immense growth in big and open data, the infrastructure and technologies to enable data publication and analytics have increased. Big data storage and processing facilities, open data portals and platforms, and data analytics tools and models have been introduced and developed at an unprecedented pace to exploit the potential behind big and open data. Zuiderwijk et al. [17] and Demchenko et al. [18] state that these facilities and infrastructure are utilized by many users, and are developed and maintained by a large numbers of parties. Hence, we view them as part of a broader data ecosystem in which each model or tool can add value as part of the information puzzle. In this research, data ecosystems play a pivotal role in understanding the involved stakeholders and their interaction in generating and utilizing open and big data for an inclusive heat transition. A literature review on data ecosystems by Oliveira et al. [19] highlighted the following benefits of the data ecosystem approach: (1) Improved political and social aspects, (2) improved economic aspects, (3) convenient data generation and utilization, (4) improved communication and interaction between actors, and (5) improved quality of data and services.

The term ecosystem is derived from its use in biology, where it represents a natural unit, functioning as a whole, consisting of plants, animals, and microorganisms together with the non-living physical resources in the environment [19]. For open data, the ecosystem metaphor is used to refer to an interdependent social system consisting of individuals, organizations, infrastructures, and resources that can be created in technology-enabled, information-intensive social systems [14,17,19]. In a study with a focus on the architectural elements of big data ecosystems, the ecosystem metaphor is used to state that the big data challenge is not only related to its core technological components, rather it is a complex whole of components to store, process, and visualize data, and deliver results [18]. This whole of the interrelated components is what is defined as a big data ecosystem, which deals with the evolving data, models, and supporting infrastructure over the data lifecycle.

The term data ecosystem is commonly used in studies; however, there is little consensus on its definition. A literature review [19] listed fifteen different definitions, stating that the first work on data ecosystems was published in 2011 by [20–22]. A commonly used definition for a data ecosystem is: “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 stake-holders” ([23], p. 3).

According to Pollock [22] a data ecosystem consists of data cycles in which intermediate data consumers, e.g., app developers and data wranglers, process the data and share the cleaned, integrated, and packaged data back into the ecosystem for use by the end consumers.

Alternatively, Ding et al. [20] view a data ecosystem from a linked data perspective and define it as, “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” ([20], p. 326). Despite the variety in definitions, common elements of these ecosystems are found, namely: (1) Stakeholders and roles, (2) relationships and interactions, and (3) (digital) resources.

According to Oliveira et al. [19], the theoretical basis for data ecosystems has been slow to develop, with no common agreement on how theories should be shaped. In the 29 articles on data ecosystems reviewed, a large variety of theories are applied to study and understand data ecosystems, namely: Socio-technical systems theory, resource dependency theory, value chain theory, and software ecosystems. Others relate data ecosystems to business ecosystems, Innovation ecosystems,

digital ecosystems, and open government ecosystems [17]. It can be concluded that in data ecosystems, elements are combined and linked using various theories to achieve novel and effective data applications. These elements are further elaborated in the following paragraph.

2.2. Data Ecosystem Characteristics and Elements

The review study on data ecosystems by Oliveira et al. [19] shows that the various views have one aspect in common: They identify these ecosystems as socio-technical systems with a large variety of technical and social elements involved and a high degree of interdependencies. Hence, the study and design of data ecosystems requires a thorough contextual understanding of human interactions, in relation to the technological, cultural, political, and economic context.

According to Van Schalkwyk et al. [15], the context of data ecosystems can be characterized in three categories. First, the regulatory context entails laws, policies, standards, and agreements, and guides the structure of the data ecosystem and the relation between the actors. Second, the institutional or environmental context is where the actors operate under certain values, rules, and norms. Third is the technological context, encompassing the ICT resources and operators but also other enabling technologies that contribute to connecting the elements.

Finally, Oliveira et al. [19] report on the organizational structure of data ecosystems, entailing the actors and their role in the data ecosystem, but also the interests and business models connecting these actors. The most common organizational structures found in the reviewed work are: (1) Keystone centric, where stakeholders are organized around the key actor responsible for directly or indirectly providing data; (2) intermediary based, where the data intermediaries are accountable for adding value to the data; (3) platform centric, where a platform shapes the organization; and (4) marketplace based, where the market provides the infrastructure, business models, rules, and services for the exchange of data between actors.

In addition to the characteristics discussed above, several studies propose data ecosystem elements or components, in order to interact and achieve the desired functionality. Table 1 presents key elements for both open data and big data ecosystems derived from the literature. First, Zuiderwijk et al. [17] propose a set of four core elements and three additional elements for data ecosystems with the focus on open government data.

Next, three categories of ecosystem elements are proposed by Mercado-Lara et al. [24] that particularly emphasize the stakeholders and their roles within the ecosystem.

Demchenko et al. [18] and Shin and Choi [25] studied the big data ecosystems, with some key differences. The latter take a broader approach in proposing the key elements, also considering the social aspects, in addition to the technological aspects presented by the former.

Finally, among these studies, none address both big data and open data coherently. The present study will take an integrated approach, addressing both big and open data. Given the characteristics and elements of data ecosystems, Davies [21] proposes a set of steps to facilitate data ecosystem development, namely: (1) To identify the entities that can take a role as being an essential data ecosystem component; (2) to comprehend the nature of the transactions and interactions that occur between those entities; (3) to identify what resources are needed by each entity to engage with each other in transactions and interactions; and (4) to observe the indicators used to measure the data ecosystem health and performance.

Although the body of literature on data ecosystems is growing, no application of data ecosystem research has thus far been conducted in the energy domain. This would be expected given the current trend of increasing digitization in the energy sector, with the introduction of smart metering and control technology to promote energy efficiency, demand side management, and to cope with the increasing diversity and share of (intermittent) renewable energy sources. Approaching the data-driven support of a thermal energy systems through a data ecosystem perspective fits the social complexity of the thermal energy transition and the need for a holistic approach considering the dynamic interaction between the different parts of the system.

Table 1. Overview of data ecosystem elements and components.

Ecosystem Type	Element or Component	Source
Open Data Ecosystems	(1) Releasing and publishing open data on the internet	[17]
	(2) Searching, finding, assessing, and viewing data and the associated licenses	
	(3) Cleansing, analyzing, enriching, combining, linking, and visualizing 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) Metadata to connect the elements	
Open Data Ecosystems	(1) Government policies and practices	[25]
	(2) Innovators as a combination of technology, business, and the government	
	(3) Users, civil society, and business	
Big Data Ecosystems	(1) Data models, structures and types	[18]
	(2) Big data management on, e.g., lifecycle, transformation and staging, and sourcing, curating, and archiving	
	(3) Big data analytics and tools, e.g., Big data applications, the target use, presentation, and visualization	
	(4) Big data infrastructure for storage, computation, exchange, and big data operational support	
	(5) Big data security, on in-rest and in-move data, and through trusted processing environments	
Big Data Ecosystems	(1) Infrastructure	[26]
	(2) Software and technology	
	(3) Service and applications	
	(4) Standards	
	(5) Users	
	(6) Social and cultural factors	
	(7) Government	
	(8) Industry	

2.3. The Data Ecosystem Framework: Actor Roles, Elements, and Context

Based on the literature reviewed on data ecosystems in the previous sections, we present a general framework, which will be used as the basis for developing the data ecosystem for the thermal energy transition in the Netherlands (see Figure 1). We start with the data stakeholder context in which the actors and their roles are presented. For data ecosystems, in general, three types of roles can be defined [26,27]:

- (1) The data producers and/or providers: The entities that produce and provide the data, whereby data can be acquired, stored, and processed over a myriad of technologies and formats.
- (2) The data users: The entities with the will and/or skills and technology for data analytics.
- (3) The data intermediaries: The entities that organize and facilitate the release and exchange of data and coordinate the participation of data users and providers, i.e., the organizational functions of the intermediary. Additionally, the intermediaries may have data-related functions,

e.g., data pre-processing and the development and provision of technology and facilities for data sharing and analytics.

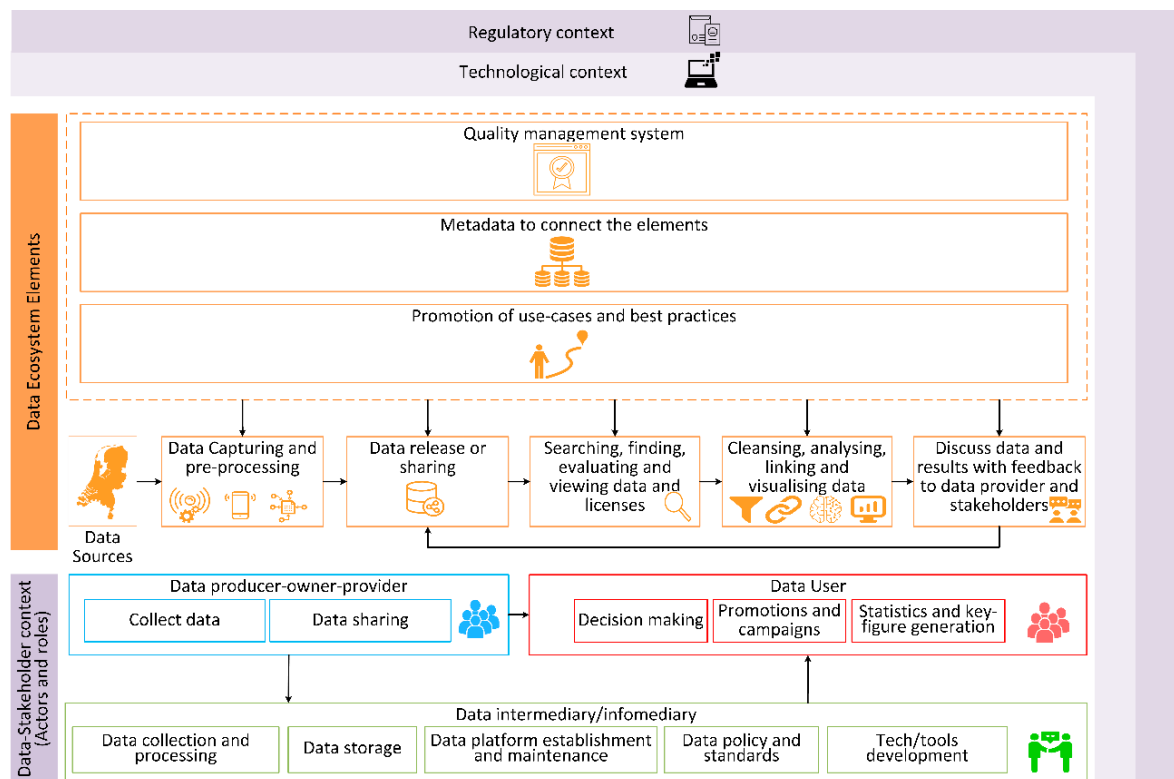


Figure 1. Framework for developing a data ecosystem.

In Figure 1, the lower part referenced the “data stakeholder context” depicts how these roles relate to each other and to the data ecosystem. For each role, which activities are encompassed is stated.

The mid-part of the framework, the “Data Ecosystem Elements”, represents the elements derived from the literature study. Elements from works that respectively focus on open government data and big data [17,18,24,25] have been combined to complement each other. The elements we included in the framework pertain to: (1) Data capturing and pre-processing to the desired data format; (2) data release or sharing; (3) searching, finding, viewing, and assessing data and data licenses; (4) cleansing, analyzing, linking, and visualizing data; (5) discussing data and results and providing feedback to data providers and stakeholders; (6) use-case promotion; (7) metadata to connect the elements; and (8) quality management. Elements 1 to 5 capture the data lifecycle management, which proposes steps on how to gather data and add value to the data as it passes through the different steps. On this subject, various studies present models, for instance:

- (1) Data acquisition; (2) data analysis; (3) data curation; (4) data storage; and (5) data usage [28].
- (1) Data collection and registration; (2) data filtering; (3) enrichment and classification; (4) data analytics, modelling, and prediction; (5) data delivery and visualization; and (6) a parallel step of (6a) data re-purposing; (6b) analytics re-factoring; and (6c) secondary processing [18].
- What these models commonly show is that the steps are organized in a linear way, whereby occasional feedback loops occur towards earlier steps. Elements 1 to 5 are a combination of the models proposed by [17,18,28].

In addition to the data stakeholder context, we include the regulatory and technological context, as proposed by Van Schalkwyk et al. [15]. The regulatory context refers to the institutional and environmental context in which the thermal energy transition takes place. The technological

context refers to the technological characteristics of the energy system with regards to ICT, thermal energy technologies, and other enabling technologies relevant for the heat transition and its increasing digitization.

This section has predominantly placed the focus on the first two elements of data ecosystems, namely (1) stakeholders and their roles and (2) interactions and relations. The third element, digital resources, will be elaborated in the following paragraphs.

2.4. Digital Resources in Data Ecosystems: Big and Open Data

Data is increasingly gaining value as an economic asset. In contrast, the use of data for public decision-making in general is still underdeveloped. In public policy contexts, data analytics can offer new insights into behaviors and patterns, with less dependence on cumbersome surveying methods to retrieve data. The technological innovation achieved in the last decades, in particular the ICT domain, such as the use of social media, smart phones, and Internet of Things (IoT), is driving the exponential increase in the volume of data, hence big data. Moreover, open data is considered a major enabler of public service innovation [29].

Big data, and its role in both public and private decision-making, has attracted significant research in the past decade [30,31]. The literature is diverse regarding the definition of big data; a comprehensive definition, which also emphasizes the increase in scale and scope of data, addresses big data as “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” ([32], p. 349). This working definition will be applied in this paper due to its comprehensiveness and specificity relative to the more common definition of big data in industry, where it is defined by means of five characteristics, i.e., volume, value, variety, velocity, and veracity [33]. 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 utilized in the context of smart and micro grid research to enable renewable energy generation and demand side management, asset management, and collaboration [34].

The Open Knowledge Foundation defines 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” ([35], p. 6). For data to be classified as open data, the data should comply with the ten criteria established in 2010 by the Sunlight Foundation, namely: The data should be complete, primary, timely, accessible, machine processable, non-discriminatory, non-proprietary, permanent, license-free, and preferably free of charge [16]. The open access to data sets, and advanced analytical methods and tools have opened up endless possibilities to generate new knowledge [29].

2.5. Challenges and Barriers for Big and Open Data

Reis et al. [36] provide an overview of recurring challenges and barriers in the application of open data and open data portals worldwide, based on a systemic literature review. The following is an overview of these challenges, complemented with challenges from other work: 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 metadata; poor data lifecycle management; poor data traceability; high-paced technological development [37]; inconsistent data [38]; lack of data standards [39–41]; and communication of results [42]. Next to the numerous challenges and barriers for big data and open data stands the potential to support decision-making in the thermal energy transition. In the next paragraph, we briefly address the application of big and open data in the energy domain.

2.6. A Review of Big and Open Data in the Energy Domain

Big and open data provide ample potential for improved planning, implementation, and operation of energy infrastructure as presented in [38,39,43]. The following paragraphs will address the current

state of data utilization in climate and energy policy, with the focus on the demand side, where measures for energy efficiency and sustainable heating are necessary, but where high-quality and detailed data is commonly lacking [8,9].

For evidence-based policy-making, the data-driven approach utilizes data on energy, the environment, and the economy to establish a knowledge base and subsequently recognize problems, set priorities, establish and implement policy, and finally monitor and assess the efficacy of policy and the need for an adaptive response [44–46]. Big and open data can be sourced from both the energy infrastructure for generation [43] and distribution [47], as from the energy use, driven by end-user behavior and building characteristics [29,30]. In the heat transition, climate policy has a significant role to enable and guide the transition. Although the relevance and necessity for big and 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 [32,48].

3. Research Design and Methodology

3.1. Research Design

We used a case study research design, in particular an embedded single-case study design [49]. It encompasses the Netherlands as the main case, the stage where the thermal energy transition has to be realized on a national level, and the municipality of Utrecht as an embedded case. The case study is suitable to enable the collection of qualitative data from which rich and detailed information can be derived on the local challenges regarding the thermal energy transition, its information provision, and the associated data ecosystem, via the stakeholder interactions and complex social phenomena [50]. The case study data were collected with the goal of representing the real-life context of the thermal energy transition with regards to the actors involved; the current and future decision-making processes; the knowledge gaps based on the technological, social, and economic context; and the data needs and availability. As a framework to gather the case data, the data ecosystem framework presented in Section 2.3 was utilized. On its turn, the empirical data yielded two results pertaining to the data ecosystem framework: (1) The extent to which the framework elements can also be found in the case and how these elements perform in the facilitation of the information provision in the thermal energy transition, and (2) additions to the framework in terms of new elements, which can be identified in the case, and elements that are necessary in the case but not yet in the framework.

3.2. Case Description: The City of Utrecht

In order to include as much as possible of the local heating system context and to grasp the actual challenges regarding the thermal energy transition and the associated knowledge gaps and barriers that should be addressed by the data ecosystem framework, the focus of the embedded case study was the city of Utrecht. Utrecht is located in the Dutch province of Utrecht, with a surface area of 99 km². In 2018, Utrecht counted 348,000 inhabitants, divided over 178,000 households, in 151,000 dwellings [51]. With the ambition to become climate neutral by 2030, Utrecht developed its policy to make drastic changes to its energy landscape, including the heat supply, in the coming years [52,53]. The building stock in Utrecht is dominated by apartment buildings and row houses. When compared to other cities in the Netherlands, Utrecht has an overrepresentation of relatively old dwellings, of which many originate from before 1945 (i.e., pre-war dwellings) [54].

In 2017, the residential dwellings in Utrecht consumed a total of 6.39 PJ of energy, of which around 80% was used as heat for space heating, cooking, and warm tap water. The remaining energy use is in the form of 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% [55]. Over 58% of the dwellings has energy label C or better, on a range from G to A, with A having a higher energy sustainability performance compared to B [55]. In the Netherlands, after Rotterdam, Utrecht is the

leader in terms of the installed district heating capacity. In 2015, 52,800 dwellings were connected to the district heating network. This system is operated by energy provider Eneco, and is expected to increase to 58,000 connections in 2020 [4].

3.3. Data Collection: Expert Interviews

To gather the data for the analysis of Utrecht, both open and semi-structured interviews were used over two rounds. In the first round of problem orientation, we conducted open interviews to gather empirical information on the perception of the problem, possible solutions, and the stakeholders involved. In the second round of problem analysis and data ecosystem design, we conducted semi-structured interviews. The set of stakeholders derived from the first round of interviews (see Table A1 in the Appendix A) were used to form the pool of interviewees in the second round (see Table A2 of the Appendix A). In total, 21 expert interviews were conducted.

3.4. Data Treatment and Analysis

For the data treatment and analysis, we used the qualitative data analysis software NVivo version 10 [56]. Data treatment consisted of two coding iterations where the simultaneous coding or co-occurrence coding method was applied, initially on a predefined code set derived from the literature, and second, on a code set enriched with the codes derived from the empirical data. Each iteration was followed by a filtering round to remove the rarely used codes and by merging codes that showed significant similarity or correlation. Data analysis consisted primarily of cluster analysis, operationalized through a coding matrix, for the definition of themes and interrelations between themes [57,58].

In addition to the definition of themes, the cluster analysis aided the classification of the data, by grouping data in classes related to various characteristics of the respondents [58]. This revealed respondent characteristics vis-à-vis the data, e.g., which challenges or knowledge gaps were commonly mentioned by a specific category of respondents. This led to a better understanding of the complexity and variation in the perceptions and attitudes of various stakeholders when studying the socio-technical system behind the thermal energy transition. In turn, this knowledge was utilized to specify the adjustment or addition of elements to the data ecosystem and shape the roles and interactions in the data stakeholder context for the specific case of the thermal energy transition.

4. The Current State of Stakeholders and the Knowledge Provision in the Dutch Heat Transition

4.1. The Socio-Technical System

The first round of interviews describes the socio-technical system of the Dutch heat transition as extensive, where stakeholders can be divided into seven categories, i.e., citizens, government, government authorities, market (construction, technical installation, energy utility and service, etc.), intelligence (research and advisory), real-estate (developers, intermediaries, and housing corporations), and others (from network operators to financial institutions and local citizen initiatives). This extensive field of stakeholders evolves around an equally extensive field of technologies at different stages of maturity, sustainability, and affordability, with different characteristics on the energy source and temperature, and which can be deployed as individual or collective systems. The social domain meets the technological domain to make decisions, invest, and adopt alternatives for natural gas, and is challenged by the technological uncertainty and the allocation of costs among public and private parties. This decision-making strives for the optimization of the thermal energy transition, whereby CO₂ emissions are minimized, at minimal societal costs, while the reliability and fairness of the thermal energy system needs to be guaranteed. Additionally, the decision-making of the various stakeholders needs to be aligned in order to have the benefits of scale and time with regard to changes in the infrastructure, e.g., joint replacement of the sewage system, old natural gas grid and a new district heating network, and the dwellings, e.g., collective insulation and installation of heat pumps.

In the stakeholder domain, a notable finding is that many stakeholders are in the process of comprehending the challenge of disconnecting from natural gas in the built environment and searching for their role to realize this. Given that these stakeholders are not yet aware of their role, it is subsequently unclear what resources they have available for the thermal energy transition while their attitudes towards the transition are widely varying and unstable. It is commonly mentioned by the respondents that the government is responsible for establishing the facilitating conditions; this will provide the stakeholders clarity on their role and reduce uncertainty in the decision-making.

From a data ecosystem perspective, some clear roles can be distinguished, e.g., Stedin as a distribution system operator (DSO) takes the role of a raw data supplier, because their legal status forbids the addition of value to data. The municipality with the thermal energy transition vision, and utility companies are taking the role of data users in supporting the planning and development activities in the thermal energy transition. Cadastre, responsible for the real estate register, and CBS (Statistics Netherlands) are positioning themselves as data intermediaries, whereby the role entails, on the one hand, organizational aspects pertaining to the establishment and maintenance of databases and the facilitation of the stakeholders involved in that process. On the other hand, they are involved in data aspects where they add value to data, by means of (pre-)processing, and releasing data for further analysis and utilization by users. The role of “standards organizations” is important in the data ecosystem to ensure data quality and interoperability; however, such an authority currently only exists for government geo-data.

4.2. The Knowledge Gaps and Data Needs

Stakeholders often address challenges in co-occurrence with knowledge gaps. In other words, many of the challenges in the thermal energy transition are a consequence of lacking or sub-optimal knowledge, entailing inadequate information provision to the decision-makers in the thermal energy transition. The identified knowledge gaps can be summarized in five main themes: (1) The energy system and environment, (2) dwellings and end-users, (3) market and economic aspects, (4) the decision-making process, and (5) the data ecosystem. Of these categories, the majority of knowledge gaps focus on: (1) The energy system and environment, and (2) dwellings and end-users. Within these two categories, the following knowledge needs are derived:

- (1) The detailed characteristics of dwellings that impact the costs of retrofit and the applicability of thermal installation upgrades;
- (2) Perceptions and attitudes of citizens and building owners influencing their willingness to act;
- (3) The preferences of stakeholders for (a) natural gas alternatives and (b) their role; and
- (4) Investment and planning cycles of actors to align these efforts in order to gain momentum and to reduce the societal costs of the transition.

4.3. Databases and Data Platforms for the Heat Transition

An explorative inventory of databases relevant for the thermal energy transition resulted in a total of 24 databases, which can be classified as: Supply side data, demand side data, building stock data, and energy statistics data (see Figure A1 for a schematic overview). Most of the data openly available is on the supply side and the infrastructure for distribution and storage. However, on the demand side, including dwelling characteristics, end-user characteristics, and behavior, little data is available and released as open data. Data with great potential, but not yet captured adequately, relate to (1) citizen preferences and attitudes towards the alternatives for natural gas and retrofit measures, and (2) the dwelling’s structural state and the retrofit and thermal installation measures implemented.

Besides the databases, we identified and assessed data platforms or portals targeting energy. This yielded nine platforms with varying degrees of functionality and data feeds (see Tables A3 and A4 in the Appendix A for a description of these platforms). When placing these platforms and databases in the context of a data ecosystem, it can be concluded that they represent a very rich ecosystem. However,

little is known about the links and interactions between the platforms and databases. This means that several platforms have redundant functionality, whereas other platforms complement each other. Consequently, this poor overview of ecosystem opportunities hampers the linked utilization of the platforms. Data-driven strategies have the potential to address knowledge needs, but significant barriers are reported by the stakeholders in the development and execution of effective data-driven strategies. They pertain to: (1) Restricting (privacy) legislation; (2) data ecosystem barriers, e.g., difficult and cumbersome data search and acquisition, poor data quality, and detail level; (3) barriers found among stakeholders, e.g., poor willingness to share data; and (4) high perceived cost. In the next section, these barriers are further elaborated.

4.4. Barriers to Data-Driven Approaches in the Thermal Energy Transition

There is a large amount of data currently available or that could become available in the foreseeable future through, among others, the data platforms [9]. However, at the present stage of data and platform applications for the thermal energy transition, reoccurring challenges and barriers are encountered. The interviews with stakeholders in the Dutch heat transition revealed many barriers encountered:

- (1) Privacy legislation, namely the “Algemene Verordening Gegevensbeheer” or 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 reoccurring barrier stated by the interviewees (11 out of the 18 interviewees). Privacy legislation impacts data release, linking, and analytics. The application of big and open linked data (BOLD), by linking data to derive richer information, is perceived as illegal and thus limits these activities (Business Developer, Eneco, Rotterdam, the Netherlands, 2018).
- (2) Difficult and lengthy processes to find and access the necessary data via the appropriate platforms or portals is the second most mentioned barrier. A lot of time and expertise is required to gain access to the data and to utilize the data effectively. Among other aspects, this is caused by the dispersed distribution of data and inconvenient interfaces. This process is experienced as inconvenient, devious, and complex by several interviewees. HoogravenDuurzaam, a sustainable neighborhood citizen initiative, questions whether the added value of insights from the data and platforms outweigh the hassle (Chairman, HoogravenDuurzaam, 2018). In addition, the data may be subject to restrictions and costs, which is also experienced as a barrier (Strategic Adviser, Kadaster, 2018).
- (3) The immature state of organizations with regards to data-driven strategies. Organizations, such as the municipality of Utrecht, Heijmans, Rosmalen, the Netherlands, (a construction company), and the citizen initiatives are struggling to comprehend the technological complexity, and lack the facilitating data ecosystem to support them in developing and executing data-driven strategies. In the commercial sector, business firms lack the skills and a clear business case for the added value of utilizing and sharing data; this is a barrier towards the participation of the commercial sector in the data ecosystem (Project Director, Heijmans, Rosmalen, the Netherlands, 2019).
- (4) Poor data quality and incomplete data, e.g., missing entries in the BAG (abbreviation for basic registration of addresses and buildings; translation by the authors), and poor database compatibility when combining data (Business Developer, Eneco, Rotterdam, the Netherlands, 2018).
- (5) Lack of data at the ultra-local level. Most open data that is available on platforms and portals is on the regional and municipality level; there is little outreach to the lower levels of detail.
- (6) Distrust among citizens leads to hesitation and a low willingness to share data, for instance, citizens opt to turn off the smart meter or do not register their photovoltaic solar panels. This is enforced by the perception that the purpose of the data is not always clear. If citizens are not convinced of the purpose the willingness to share data is low (Business Developer, Stedin, Rotterdam, the Netherlands, 2018; Statistical Officer, CBS, 2018).
- (7) There is a mismatch between the prevailing legal framework and the design and implementation of novel policy instruments. On the one hand, policy instruments impose the data and services that

- specific actors need to deliver, but on the other hand the restrictions posed by legislation, such as privacy legislation, limit their ability to comply with these requirements. For example, due to their legal position, DSOs are not allowed to add value to the data they own, e.g., visualizing the data is not allowed. Hence, they are limited to releasing raw open data (Business Developer, Stedin, 2018).
- (8) For organizations looking to organize a data ecosystem and improve the release of data, a barrier mentioned is the lack of cooperation by data owners, such as the energy providers (Statistics Officer, CBS, 2018).
 - (9) Finally, perceived high costs necessary for the training and acquisition of skilled labor and the ICT infrastructure and software are commonly mentioned as a barrier withholding the large-scale roll-out of data-driven strategies.

5. A Data Ecosystem to Drive the Dutch Thermal Energy Transition

In order to tackle the barriers and challenges, we focus on the following three building blocks: (1) The information needs in the thermal energy transition; (2) the existing and potential databases and platforms with the encountered challenges and barriers; and (3) potential technologies to improve the data capturing, exchange, and utilization. We applied the empirical data into refinement and additions to the generic data ecosystem framework derived from the literature in Section 2. After refinement with the interview and case study data, the improved data ecosystem, entitled “Data Ecosystem 2.0”, was assessed during an expert validation session, and discussed with a leading researcher in the field of open data and data ecosystems at Delft University of Technology and with experts at TNO (Netherlands Organization for Applied Scientific Research). Data Ecosystem 2.0 is presented in Figure 2. In the following sections, we first address the data ecosystem elements in Section 5.1, the data stakeholder context in Section 5.2, the regulatory context in Section 5.3, and the technical context in Section 5.4.

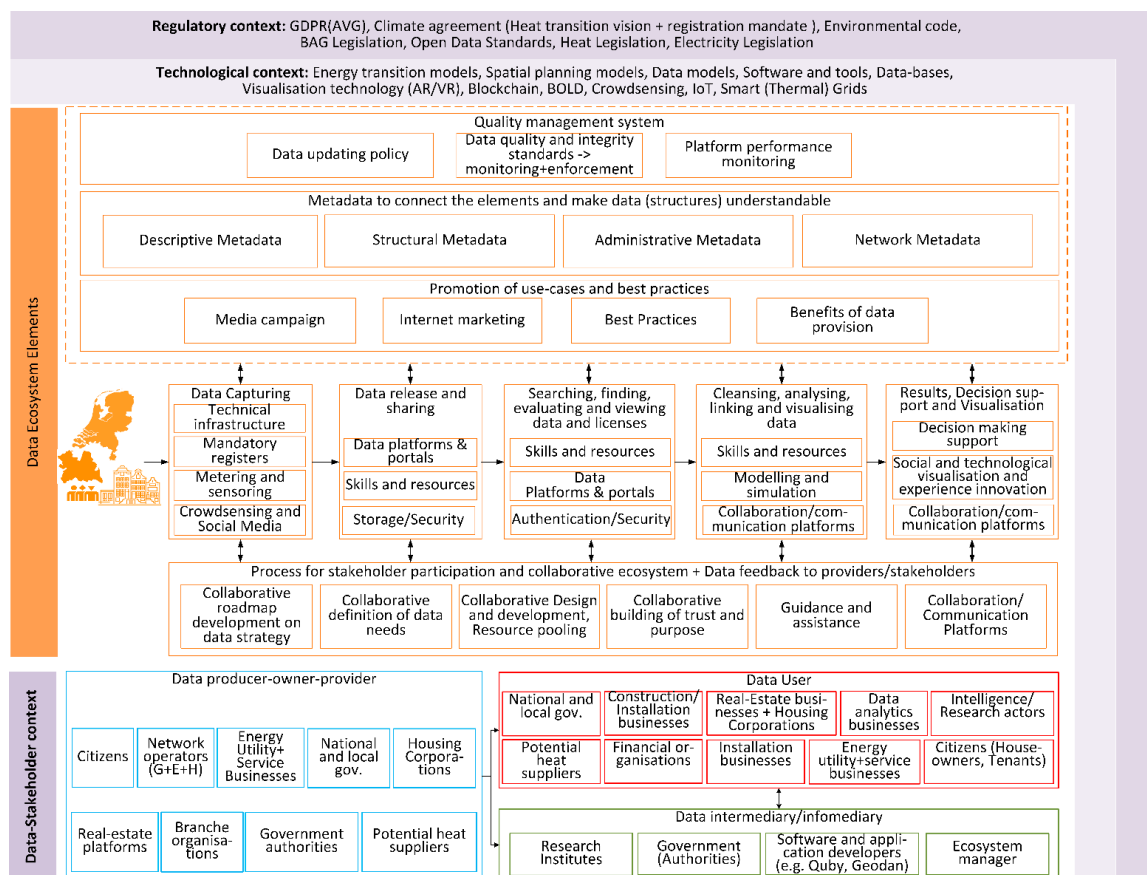


Figure 2. The data ecosystem 2.0 design to support the heat transition in the Netherlands.

5.1. The New or Adjusted Elements in the Proposed Data Ecosystem 2.0

In the following paragraphs, we present the main challenges that were mentioned pertaining to the data ecosystem elements of the quality management system, the metadata, and the promotion of use cases and best practices. Furthermore, in this layer of the framework, we add the new aspect of a process for stakeholder participation and collaboration.

5.1.1. Quality Management System

Poor data quality and incomplete data were explicitly mentioned as being problematic by several data-users whom we interviewed, when utilizing open data. Subsequently, the quality of insights gained from the data is poor. The existence of a quality management system is therefore considered 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, whereas other data types lack such standards. Also relevant for the thermal energy transition are energy potential data, data on technology performance and costs, data on the dwellings, and data on the consumer's demand and behavior. A fair share of this data is not geo-data, leaving a significant share of current data without quality standards. For a solid data ecosystem reaching further than geo-data, this quality system requires expansion to other data types.

Moreover, a future data management system would benefit from not being limited to the development of standards, but also to communicate these standards towards the data ecosystem participants, and to monitor and enforce the compliance to these standards.

For the provision of data for the energy transition, the following elements must be included in a quality management system: Reliability, completeness, topicality, continuity, independence, veracity (source integrity and quality), and interoperable data formats [8,9]. By ensuring these quality factors for the data ecosystem 2.0, the barriers and challenges encountered in the current ecosystem, as presented in Section 4.4, can be resolved.

5.1.2. Metadata

Metadata, or data on data, is important for users to gain a thorough understanding of the data and to assess whether the data fits their needs. Moreover, metadata can benefit the data ecosystem in improved storing, preservation, accessibility, visualization, and interoperability of open data [17]. The three types of required metadata are [59]: (1) Descriptive metadata on the characteristics of the data, which enables convenient data identification and discovery; (2) structural metadata on the composition of the database, enabling transparency and architecture improvement; and (3) administrative metadata on the intellectual property of data and data archiving, enabling clarity in the conditions for use.

In addition to the above-mentioned metadata, we propose to establish metadata standards on the landscape of databases and platforms, and the associated providers and users. The term proposed is network metadata, and this network metadata should provide clarity to providers, users, and intermediaries on the links between databases and platforms. Network metadata may contribute to the interoperability between databases 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 effective open data systems, having sufficient users and data providers is crucial. The expert interviews revealed that the utilization of platforms and availability of portals is still low. In order to improve the use of the data and get the most out of its potential, use-case promotion informs and inspires users about data applications. Therefore, we propose to promote best practices and innovative use-cases in which the added value and purpose of data utilization becomes apparent to (potential) data sources, such as the citizens; data intermediaries, such as parties, developing software and tools;

and data users. Being aware of the added value of the data yields benefits, which are twofold: On the one hand, it motivates decision-makers to utilize data-driven tools and methods to support decision making, and on the other hand, it incentivizes data providers to share data.

5.1.4. Result Communication and Visualization

According to Reinhart and Davila [42], communicating immense amounts of data to stakeholders as comprehensive and actionable information is challenging. This aspect was first embedded in the element of cleaning, linking, analyzing, and visualizing data in the generic data ecosystem framework. However, the results presentation is proposed as a separate element in the data ecosystem 2.0, because it is derived to be of a different nature and depends on different theories and processes 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 visualize insights from data to citizens, but is less relevant vis-à-vis data cleaning, linking, and analysis (Product Owner, Geodan, 2018). Nevertheless, these two elements are strongly intertwined, and need to align activities.

5.1.5. Data Ecosystem 2.0 Addition: Process for Stakeholder Participation and Data Ecosystem Collaboration

A new part of the data ecosystem 2.0 is the process for stakeholder participation and collaboration, due to the strong need for process support to decision-makers, such as the municipalities, housing corporations, and citizens (Business Developer, Stedin, 2018; Consultant, Overmorgen, 2018; Chairman, HoogravenDuurzaam, 2018; Strategic Adviser, Municipality of Utrecht, 2018). This process support ranges from technical knowledge to financial aspects, and the identification and utilization of data-driven decision support methods and tools. This pertains to the added element of stakeholder participation and collaboration, which is merged with the element on the discussion of data and provision of feedback to the data providers [17].

This element should include data providers and users early in the data ecosystem and over the complete data lifecycle, in order to improve utilization. On the data demand side, the aim is to create: (1) Platform designs with functionalities in which the needs of the users are met according to a demand-driven approach; and (2) increased familiarity of the platform from the early start.

On the data supply side (as presented in Section 4.4), there is a lack of trust by citizens in authorities and businesses that aim to capture, acquire, and utilize citizen data. This distrust is caused by privacy concerns, and an unawareness of the purpose of data utilization. Without knowing what the data will be used for, and what data is specifically collected, 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 new element for stakeholder participation and collaboration aims to build trust and awareness on the supply side by engaging (potential) data suppliers in the identification and definition of the added value of data utilization and the functionality of platforms and portals. Ultimately, the enhanced trust can contribute to an increased willingness of citizens and other data holders to share data.

Finally, the current Dutch data ecosystem is lacking data assessment and feedback to improve the data quality. The two-way feedback between users and providers over the data lifecycle, as included in this element, stimulates the continuous improvement of data and platform quality.

5.2. Data Stakeholder Context

In Figure 3, the current data stakeholder context is presented in detail. The data ecosystem 2.0 stresses the importance to initiate the data ecosystem manager, who can take up the responsibility for the data quality system. This role is currently missing in the data ecosystem of the thermal energy transition. Several organizations pick up tasks on the organization and management of the ecosystem, but there is no coherence in the data ecosystem management. This incoherence leads to duplicate functionalities of data platforms, and poor interoperability. Data ecosystem-wide metadata

management and coordination of the stakeholder participation process are tasks that can basically be performed by the data ecosystem manager.

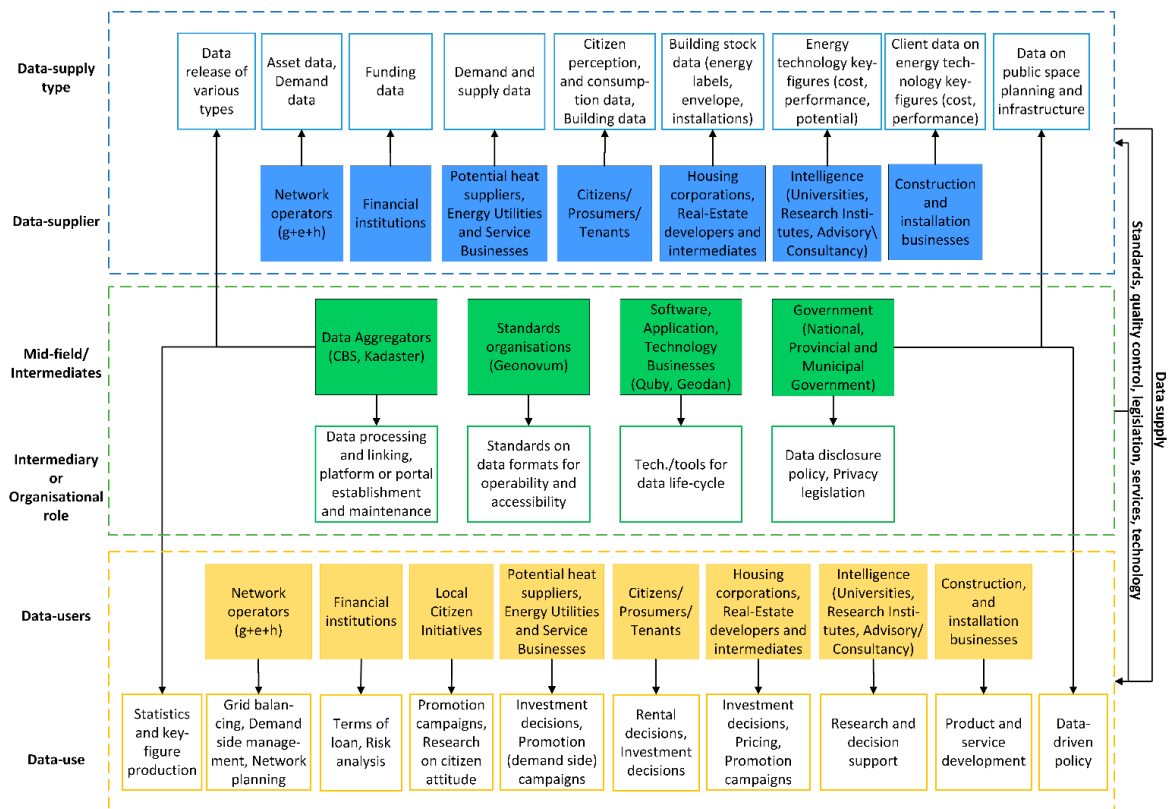


Figure 3. The data stakeholder context for the thermal energy transition in the Netherlands.

In the Dutch 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, such as ministries and government authorities [8]. Here, the sole participants in the data commission are public parties and data intermediaries according to the role definition in this study. However, from the data ecosystem approach, it is recommended to expand the composition of this data ecosystem manager with representatives from: (1) The data suppliers (e.g., the grid operators); (2) the data users (e.g., the energy utility and service businesses); and (3) the norms and standards organizations (e.g., Geonovum). With this composition, the relevant interest is represented in the data ecosystem manager, benefiting from comprehensive data ecosystem organization and management.

5.3. Regulatory Context

In the Dutch regulatory context, the Climate Agreement has an important role. It can provide the legal basis to impose mandates on households, e.g., to register the installed PV panels and thermal insulation measures in dwellings. In the data ecosystem, this can stimulate an increase in data availability by addressing the currently missing data. Moreover, the Environmental Code or “Omgevingswet”, a new scheme on environment and spatial regulations, will be relevant for the thermal energy transition data ecosystem when enacted by 2021. By means of the “Construction archive” (or “Bouwdossier” in Dutch; translation by the authors), encompassing a digital archive on each dwelling, containing relevant building data, under the Digital Scheme of the Environmental Code (or “Digitaal stelsel Omgevingswet” in Dutch; translation by the authors), novel instruments are provided to enrich the data availability on dwellings.

5.4. Technological Context

The technological context of the current data ecosystem can be described as sparsely populated and in its infancy. Currently, energy transition models, such as Vesta MAIS, developed and used for spatial energy planning by the Netherlands Environmental Assessment Agency (PBL) [60], make use of available databases. Moreover, the available platforms are casually used by stakeholders to visualize and comprehend the challenges in the thermal energy transition. However, to this point, actual decision-making has not been adequately supported in a data-driven way. Interviewees argued that governance processes have not yet embraced data-driven support. In addition, they mention the lack of critical data; this is data on the demand side of energy markets. Although the technology is steadily improving, technologies in the field of data analytics and visualization remain underutilized and even irrelevant in the absence of adequate data. Hence, the experts interviewed suggest the technological context to be expanded with innovations like 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 visualization. We therefore refined the framework element of the technological context with a list of innovative technologies that can be explored to enhance the user friendliness of data analyses.

6. Discussion

The present study takes a novel approach and addresses the thermal energy transition and its associated knowledge needs and energy data flows from both the supply and demand side. In general, academic research on these themes in the Netherlands is lacking, or mainly focusing on the infrastructural and supply side of energy (i.e., [8]). On the supply side, potential data sources are found regarding the generation potential and cost factors of sustainable heat sources, such as geothermal and solar energy, and distribution infrastructure, such as heat networks. However, results from this study indicate that although data on the supply side of energy markets is necessary, data availability on the demand side is more underdeveloped. This is a problem because the bottom-up and highly diverse nature of the thermal energy transition requires ample knowledge on the behavior, preferences, and attitude of end users. This holds in particular for energy transitions where not every end user is equally motivated or aware of its needs and the urgency of the transition.

A fair share of the common challenges and barriers in the literature on big and open data in energy systems are also encountered in the current data ecosystem of the Dutch thermal energy system. For instance, recurring barriers, such as inconvenient data access, availability and retrievability, inadequate or poor data collections, and lacking data standards, as presented in leading works for open and big data in energy systems [29,37,40–42], are confirmed by the interviewees in the present study. However, we also uncovered new barriers, such as distrust among citizens leading to poor willingness to share data, contradicting policy requirements and legislation, and the dispersed character of data over various platforms and owners.

This study contributes to the data ecosystem literature in a methodological way as case studies are rare in this field. Moreover, this study highlights the application of data ecosystems, inspired by [17–19,24–26], to the new domain of thermal energy systems. We show that data ecosystem elements (as presented in the data ecosystem literature) can indeed be considered relevant when applied to the case of the thermal energy transition in the Netherlands. In addition, the study revealed that although the importance of these elements is acknowledged, real life data ecosystems are lacking the organizational completeness. The case study showed that the critical quality management system element is absent, resulting in recurring challenges with respect to data quality and integrity.

Among studies proposing elements that contribute to effective working data ecosystems [17,18,24,25], we observe that none address both big and open data integrated in data ecosystems. This study used an integrated approach, looking into both big data and open data due to the nature and scale of data in Dutch urban thermal energy systems. As a result, the data ecosystem framework that was conceived

forms a more comprehensive view, in which both big and open data are considered, and where the scope reaches further than technology, also paying attention to the data stakeholder context, and the technical as well as the regulatory context.

Finally, we see an analogy to the concept of entropy from physics. When applied to a data ecosystem, we indicate that under the present data ecosystem on heat, in the Netherlands (access to), data is currently in a state as if it were an uneven distribution of quantities in an isolated system that tends to flatten inequalities (i.e., suggesting increasing entropy). In short, applied to a data ecosystem for a local heating system, this could mean that with increasing entropy, all data sources within a data ecosystem could ultimately be unlocked and actively play a role within that system, leading to a situation in which people have equal access to the relevant information.

7. Conclusions

7.1. Main Conclusions

7.1.1. Answering the Main Research Question

For the climate goals to be realized, the energy landscape in the built environment will have to change drastically. In this process, challenged by technological novelty, high costs, and social complexity, local governments along with citizens and stakeholders need to jointly work towards prioritizing, aligning, and executing decisions to shape the thermal energy transition from the household level up to the municipal level. However, the present study found that municipalities, citizens, and stakeholders are lacking the specific, accurate, and objective knowledge and capacities to support and engage in complex decision-making and investments to foster the thermal energy transition. This present study particularly elaborated insights on the end-user side of the thermal energy system and how to effectively, reliably, and frequently capture and process data to support decision-making towards a supported and optimal heat transition, according to the following research question: How can a data ecosystem be developed to enable the data-driven support of the local heat transition, by capturing both social and technical aspects of the urban thermal energy system?

7.1.2. The Socio-Technical System

The socio-technical system of the heat transition is very extensive, with significant interdependencies, and evolves around an equally extensive and dynamic field of technologies. These fields converge to make decisions, invest, and adopt alternatives for natural gas, whereby CO₂ emissions are minimized, at minimal societal costs, while reliability and fairness of the thermal energy system needs to be guaranteed. Additionally, the decision-making of the various stakeholders needs alignment to benefit from the scale and time with regard to changes in infrastructure, e.g., joint replacement of the old natural gas grid and a new district heating network, and the dwellings, e.g., collective thermal insulation. This is met with challenges on the uncertainty pertaining to both technological and policy measures.

When addressing perceptions held by stakeholders, many are still in the process of comprehending the challenge of disconnecting from natural gas in the built environment and searching for their role in the transition. Subsequently, it is also unclear what resources they have available for the thermal energy transition while their attitude towards the transition is widely varying and unstable. Many stakeholders expect the government to establish the facilitating conditions and financing measures; this should provide clarity on their role, after which they can proceed to decision-making in the thermal energy transition.

7.1.3. Knowledge Gaps among Stakeholders

The present study showed that technological and social knowledge gaps among the decision-makers are perceived as the most pressing challenges in the thermal energy transition.

The lacking or sub-optimal knowledge entails inadequate information provision to the decision-makers responsible and accountable for decision-making in the transition. The majority of the identified knowledge gaps refer to the themes of the energy system and environment, and the dwellings and end users. Within these categories, the most addressed knowledge needs are: (1) The detailed characteristics of dwellings, which impact the costs of retrofit and the applicability of thermal installation upgrades; (2) perceptions and attitudes of citizens and building owners influencing their willingness to take action; (3) the preferences of stakeholders for (a) natural gas alternatives, and (b) their role in the thermal energy transition; and (4) investment and planning cycles of actors to align these efforts in order gain momentum and reduce the societal costs of the transition.

7.1.4. Towards the Design of a Data Ecosystem 2.0 for the Thermal Energy Transition

Next to analyzing the current data ecosystem of Dutch heat transition, the present study sought to develop a data ecosystem to enable the data-driven support of the local heat transition, by capturing both social and technical aspects of the urban thermal energy system. The current data ecosystem for the thermal energy transition was found to be in its infancy phase at this point, with challenges, barriers, and shortcomings in all elements as proposed by the data ecosystem framework derived from the literature and presented in Section 2.3. Of the eight elements, the current data ecosystem is particularly lacking in: (1) The data discussion and feedback; (2) metadata; (3) use-case promotion; and (4) quality management system. In addition, the data feed is sub-optimal, whereby little to no data is included for the demand side on the citizens.

Currently, the energy transition models used, such as Vesta MAIS, utilize available databases, and the available platforms are casually used by stakeholders to visualize and comprehend challenges in the heat transition. So far, actual decision-making has not been supported adequately in a data-driven way. This can be assigned to relevant data technology being in its infancy, governance and decision-making processes not fully embracing data-driven support, and a lack of critical data, mainly on the demand side of thermal energy systems.

To conceive a data ecosystem that effectively supports the thermal energy transition decision-making, it is important to solve problems related to the lack of data on heat demand, dwelling characteristics, and citizen attitude and perceptions. This requires a stable and adequate data infrastructure and the use of technologies, such as blockchain-based technologies, crowd sensing, and big and open linked data.

Moreover, an adequate quality management system, which includes new roles, such as a data ecosystem manager, and a process for stakeholder participation and data ecosystem collaboration is essential. The data ecosystem manager is not only responsible for data standards and quality assurance but also needs to coordinate the process of stakeholder participation and collaboration. This new process element aims to contribute to an increased involvement of the stakeholders, leading to a better familiarity with the ecosystem, and that will also foster trust among data suppliers for an improved willingness to share data.

Finally, the data ecosystem 2.0 includes data discussion and feedback. By doing this along the data ecosystem, first, the quality of the data and infrastructure is continuously assessed and improved, and second, continuous interaction enables a data ecosystem that is aware of the specific data needs, and targets those data needs effectively and efficiently.

7.2. Limitations

The data ecosystem framework presented in this paper is a novel approach that takes a holistic approach on objects, actors, and interactions for data-driven strategies. A strength of this approach enables a comprehensive understanding. However, a limitation is that this approach stays on the meta-level, and further research is necessary to operationalize findings from this approach on the data infrastructure and activity level.

The data ecosystem framework derived from the literature and the application of this framework did not undergo the desired thorough validation. The validation was twofold. First, both the data ecosystem framework and designs for a future data ecosystem 2.0 were made subject to expert validation, where the completeness and topicality were assessed. Second, the data ecosystem framework underwent empirical validation by applying it to the Utrecht case and by assessing if its elements were 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 assess the representativeness of the case study results.

Finally, the study targeted the data-driven support in the thermal energy transition and aimed to provide a picture that was as complete as possible on relevant data and technical aspects of urban thermal energy systems. However, due to the meta-level of the analysis, the total of data types included in the study is non-exhaustive and is subject to fast development.

7.3. Recommendations for Future Research

Given the main findings of this research and its limitations, we present the following recommendations towards future research: The present study proposed particular technologies, e.g., blockchain-based technologies, crowd sensing, and big and open linked data, that have the potential to close the gaps between the current and desired data ecosystem, e.g., missing data, trust among data providers including citizens, data security and safety, and equal access to all. However, these technologies 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 thermal energy transition, regarding both the technological and operational aspects, as well as the societal embeddedness of the technologies.

Businesses in the energy sector encounter significant challenges to quantify the value of releasing their data, and subsequently struggle to develop profitable business propositions to release relevant data as a commercial asset. Consequently, it is recommended to study potential strategies and business cases for commercial parties in the energy sector on how data can become a business asset, in order to incentivize the commercial parties to release their data, enrich the data ecosystem, and improve the information provision.

We also suggest further elaboration of the data ecosystem by researching more case studies and eventually modelling the data system framework. It would basically require a different type of study of modelling a data ecosystem and trying to optimize it. We suggest to also address uncertainty when modelling the data ecosystem.

Finally, research can be conducted on how to embed the data ecosystem findings in policy making, in other words, research on policy measures that enable the realization of a data ecosystem in the operational spheres of the Dutch thermal energy transition. Relevant aspects here are models for policy instruments like subsidies or alternative incentives, organizational models for the data ecosystem manager, or governance models for shared data infrastructure.

Author Contributions: Conceptualization, D.D.; methodology, D.D., T.H., J.U.; analysis, D.D., data collection, D.D.; writing—original draft preparation, D.D., T.H., J.U.; writing—review and editing, D.D., T.H.; visualization, D.D.; supervision, T.H., J.U., A.S., K.B.; funding acquisition, A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors like to thank 21 interviewees and five independent reviewers for their valuable contribution to the present paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

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

#	Role and Subject	Organization	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-Organization 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

Table A2. Cont.

#	Role and Subject	Organisation	Stakeholder Type	Interview Type
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 advice 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

Table A3. Overview of data platforms or portals that can be utilized 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, Labor 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 neighborhood level.
PDOK Platform and Viewer-Kadaster	Energy use, Sustainable generation, Energy potential, Spatial area mapping, Spatial planning, (Subsurface) infrastructure, Hydrological system	PDOK, or Public-service on the map, is a national geographical data portal or platform which combines, releases and visualizes 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”

Table A3. Cont.

Data-Platform/Portal	Theme	Description
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 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 visualization and analysis of that data. The data utilized for the Energy platforms and the visualization 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-Cadaster	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. Overview of the data platform or portal functionality.

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

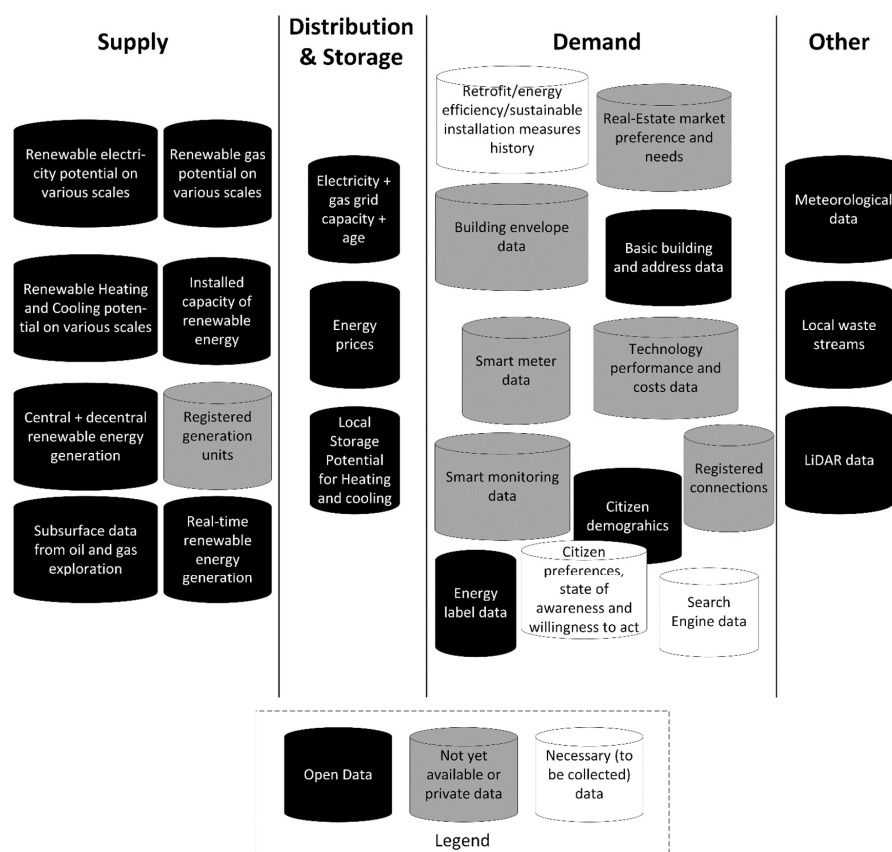


Figure A1. Overview of data sources (current and potential) with relevance for the heat transition.

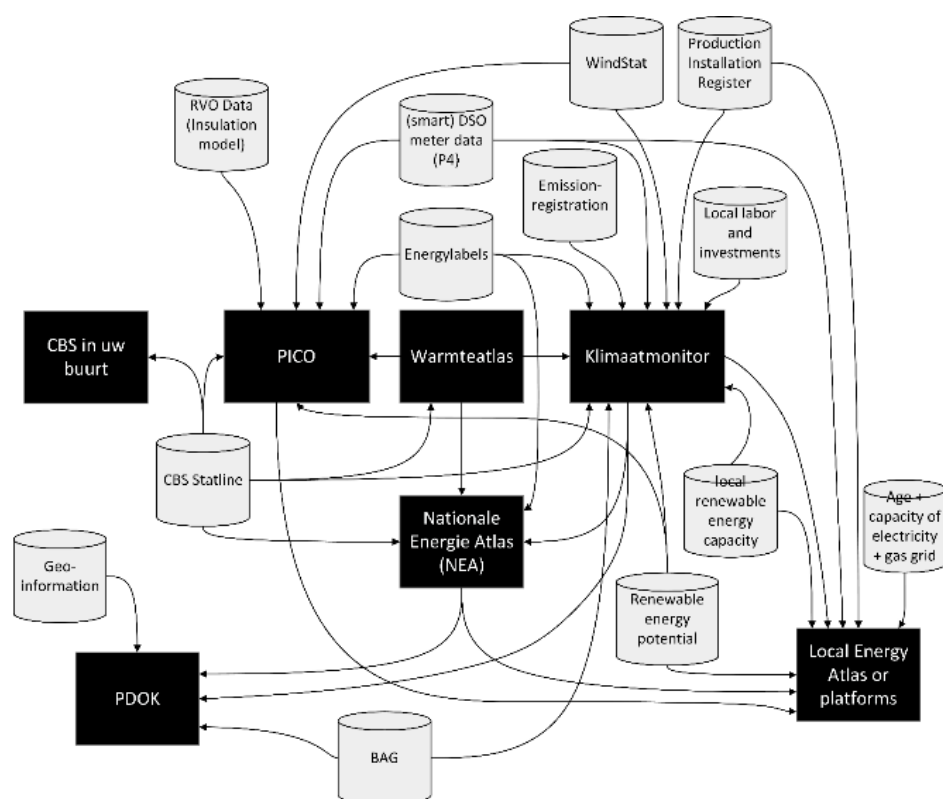


Figure A2. Visualization of how the platform uses the data from Figure A1.

References

1. United Nations. *Paris Agreement* [Internet]; United Nations: New York, NY, USA, 2015; Available online: https://unfccc.int/sites/default/files/english_paris_agreement.pdf (accessed on 14 January 2020).
2. Klimaatberaad. *Voorstel Voor Hoofdpijnen van Het Klimaatakkoord*; Sociaal-Economische Raad: The Hague, The Netherlands, 2018.
3. Energie Beheer Nederland. *Energie in Nederland* [Internet]. 2016, p. 2030. Available online: <https://www.energieinnederland.nl/2017/energieverbruik/gebouwde-omgeving> (accessed on 14 January 2020).
4. Menkveld, M.; Matton, R.; Segers, R.; Vroom, J.; Kremer, A.M. *Monitoring Warmte 2015*; ECN: Amsterdam, The Netherlands, 2017; p. 66.
5. Niessink, R.; Rösler, H. *Developments of Heat Distribution Networks in the Netherlands*; ECN: Petten, The Netherlands, 2015.
6. Young, J.; Brans, M. Analysis of factors affecting a shift in a local energy system towards 100% renewable energy community. *J. Clean Prod.* **2017**, *169*, 117–124. [CrossRef]
7. Broto, V.C. Urban governance and the politics of climate change. *World Dev.* **2017**, *93*, 1–15. [CrossRef]
8. Dekker, G.; Keller, K.; Swertz, O.; Vroom, J.; Mink, M.; van den Hoek, A.; Noordegraaf, L.; Hoogervorst, N.; Matthijssen, J.; Baltussen, J.; et al. VIVET: Voorstellen om de Informatievoorziening Energietransitie te Verbeteren. Available online: <https://www.cbs.nl/nl-nl/achtergrond/2019/14/vivet-betere-informatievoorziening-energietransitie> (accessed on 16 January 2020).
9. Diran, D. Data-Driven Decision Making towards Inclusive and Sustainable Urban Heating in the Netherlands: The Data Ecosystem Approach. Master's Thesis, Delft University of Technology, Faculty of Technology, Policy and Management, Delft, The Netherlands, 2019.
10. Grafakos, S.; Flamos, A.; Enseñado, E. Preferences matter: A constructive approach to incorporating local stakeholders' preferences in the sustainability evaluation of energy technologies. *Sustainability* **2015**, *7*, 10922–10960. [CrossRef]
11. Koirala, B.P.; Araghi, Y.; Kroesen, M.; Ghorbani, A.; Hakvoort, R.A.; Herder, P.M. Trust, awareness, and independence: Insights from a socio-psychological factor analysis of citizen knowledge and participation in community energy systems. *Energy Res. Soc. Sci.* **2018**, *38*, 33–40. [CrossRef]
12. Kalkbrenner, B.J.; Roosen, J. Citizens' willingness to participate in local renewable energy projects: The role of community and trust in Germany. *Energy Res. Soc. Sci.* **2016**, *13*, 60–70. [CrossRef]
13. Hatzl, S.; Brudermann, T.; Reinsberger, K.; Posch, A. Do public programs in 'energy regions' affect citizen attitudes and behavior? *Energy Policy* **2014**, *69*, 425–429. [CrossRef]
14. Harrison, T.M.; Pardo, T.A.; Cook, M. Creating open government ecosystems: A research and development agenda. *Future Internet* **2012**, *4*, 900–928. [CrossRef]
15. Van Schalkwyk, F.; Willmers, M.; McNaughton, M. Viscous open data: The roles of intermediaries in an open data ecosystem. *Inf. Technol. Dev.* **2016**, *22*, 68–83. [CrossRef]
16. Welle Donker, F.; van Loenen, B. How to assess the success of the open data ecosystem? *Int. J. Digit. Earth* **2017**, *10*, 284–306. [CrossRef]
17. Zuiderwijk, A.; Janssen, M.; Davis, C. Innovation with open data: Essential elements of open data ecosystems. *Inf. Polity* **2014**, *19*, 17–33. [CrossRef]
18. Demchenko, Y.; De Laat, C.; Membrey, P. Defining architecture components of the Big Data Ecosystem. In Proceedings of the 2014 International Conference on Collaboration Technologies and Systems (CTS), Minneapolis, MN, USA, 19–23 May 2014; pp. 104–112.
19. Oliveira, M.I.S.; Lima, F.B.; Lóscio, B.F. Investigations into Data Ecosystems: A systematic mapping study. *Knowl. Inf. Syst.* **2019**, 1–42. [CrossRef]
20. Ding, L.; Lebo, T.; Erickson, J.S.; DiFranzo, D.; Williams, G.T.; Li, X.; Michaelis, J.; Graves, A.; Zheng, J.G.; Shangguan, Z.; et al. TWC LOGD: A portal for linked open government data ecosystems. *Web Semant. Sci. Serv. Agents World Wide Web* **2011**, *9*, 325–333. [CrossRef]
21. Davies, T. Open Data: Infrastructure and ecosystems. *Open Data Res.* **2011**, *1*, 1–6.
22. Pollock, R. Building the (open) data ecosystem. *Open Knowl Found Blog*, 31 March 2011.
23. Zuiderwijk, A.; Janssen, M.; van de Kaa, G.; Poulis, K. The wicked problem of commercial value creation in open data ecosystems: Policy guidelines for governments. *Inf. Polity* **2016**, *21*, 223–236. [CrossRef]

24. Mercado-Lara, E.; Gil-Garcia, J.R. Open government and data intermediaries: The case of AidData. In Proceedings of the 15th Annual International Conference on Digital Government Research, Aguascalientes, Mexico, 18–22 June 2014; pp. 335–336.
25. Shin, D.-H.; Choi, M.J. Ecological views of big data: Perspectives and issues. *Telemat. Inform.* **2015**, *32*, 311–320. [\[CrossRef\]](#)
26. van den Homberg, M.; Susha, I. Characterizing Data Ecosystems to Support Official Statistics with Open Mapping Data for Reporting on Sustainable Development Goals. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 456. [\[CrossRef\]](#)
27. Susha, I.; Janssen, M.; Verhulst, S. Data collaboratives as a new frontier of cross-sector partnerships in the age of open data: Taxonomy development. In Proceedings of the 50th Hawaii International Conference on System Sciences, Hilton Waikoloa Village-Waikoloa, HI, USA, 4–7 January 2017.
28. Curry, E. The Big Data Value Chain: Definitions, Concepts, and Theoretical Approaches. In *New Horizons for a Data-Driven Economy*; Springer: Cham, Switzerland, 2016; pp. 29–37.
29. Toots, M.; McBride, K.; Kalvet, T.; Krimmer, R. Open data as enabler of public service co-creation: Exploring the drivers and barriers. In Proceedings of the E-Democracy and Open Government (CeDEM), Krems, Austria, 17–19 May 2017; pp. 102–112.
30. Androutsopoulou, A.; Charalabidis, Y. 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, Galway, Ireland, 4–6 April 2018; pp. 575–583.
31. Clarke, A.; Margetts, H. Governments and citizens getting to know each other? Open, closed, and big data in public management reform. *Policy Internet* **2014**, *6*, 393–417. [\[CrossRef\]](#)
32. Poel, M.; Meyer, E.T.; Schroeder, R. Big data for policymaking: Great expectations, but with limited progress? *Policy Internet* **2018**, *10*, 347–367. [\[CrossRef\]](#)
33. Poel, M.; Schroeder, R.; Treperman, J.; Rubinstein, M.; Meyer, E.; Mahieu, B.; Svetachova, M. *Data for Policy: A Study of Big Data and Other Innovative Data-Driven Approaches for Evidence-Informed Policymaking. Rep about State-of-the-Art Amsterdam Technopolis, Oxford Internet Institute; Center for European Policy Studies: Brussels, Belgium, 2015.*
34. Zhou, K.; Fu, C.; Yang, S. Big data driven smart energy management: From big data to big insights. *Renew Sustain Energy Rev.* **2016**, *56*, 215–225. [\[CrossRef\]](#)
35. Dietrich, D.; Gray, J.; McNamara, T.; Poikola, A.; Pollock, R.; Tait, J.; Zijlstra, T. *Open Data Handbook Documentation*; Open Knowledge Foundation: Cambridge, UK, 2012.
36. Reis, J.R.; Viterbo, J.; Bernardini, F. 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, Delft, The Netherlands, 30 May–1 June 2018; p. 73.
37. Linder, L.; Vionnet, D.; Bacher, J.-P.; Hennebert, J. Big Building Data—a Big Data Platform for Smart Buildings. *Energy Procedia* **2017**, *122*, 589–594. [\[CrossRef\]](#)
38. Di Corso, E.; Cerquitelli, T.; Piscitelli, M.S.; Capozzoli, A. Exploring energy certificates of buildings through unsupervised data mining techniques. In Proceedings of the 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), Exeter, UK, 21–23 June 2017; pp. 991–998.
39. Mathew, P.A.; Dunn, L.N.; Sohn, M.D.; Mercado, A.; Custudio, C.; Walter, T. Big-data for building energy performance: Lessons from assembling a very large national database of building energy use. *Appl. Energy* **2015**, *140*, 85–93. [\[CrossRef\]](#)
40. Göçer, Ö.; Hua, Y.; Göçer, K. A BIM-GIS integrated pre-retrofit model for building data mapping. In *Building Simulation*; Tsinghua University Press: Beijing, China, 2016; pp. 513–527.
41. Chen, Y.; Hong, T.; Piette, M.A. Automatic generation and simulation of urban building energy models based on city datasets for city-scale building retrofit analysis. *Appl. Energy* **2017**, *205*, 323–335. [\[CrossRef\]](#)
42. Reinhart, C.F.; Davila, C.C. Urban building energy modeling—A review of a nascent field. *Build Environ.* **2016**, *97*, 196–202. [\[CrossRef\]](#)
43. Noussan, M.; Jarre, M.; Poggio, A. Real operation data analysis on district heating load patterns. *Energy* **2017**, *129*, 70–78. [\[CrossRef\]](#)
44. van Veenstra, A.F.; Kotterink, B. Data-driven policy making: The policy lab approach. In *International Conference on Electronic Participation*; Springer: Cham, Switzerland, 2017; pp. 100–111.

45. European Commission. *Quality of Public Administration: A Toolbox for Practitioners*; Publications Office of the European Union: Brussels, Belgium, 2017.
46. Wang, D.Y.C.; Trappey, A.J.C.; Trappey, C.V.; Li, S.J. 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*; National Chiao Tung University Institutional Repository: Hsinchu, Taiwan, 2014; pp. 781–789.
47. Visser, J. *Barriers in Open Energy Data: An Exploratory Study into Open Energy Data Barriers and Their Mitigation Strategies*; Delft University of Technology: Delft, The Netherlands, 2015.
48. Pfenninger, S.; DeCarolis, J.; Hirth, L.; Quoilin, S.; Staffell, I. The importance of open data and software: Is energy research lagging behind? *Energy Policy* **2017**, *101*, 211–215. [CrossRef]
49. Yin, R.K. *Case Study Research and Applications: Design and Methods*; Sage publications: Southend Oaks, CA, USA, 2017.
50. Miles, M.B.; Huberman, A.M. *Qualitative Data Analysis: An Expanded Sourcebook*; Sage: Southend Oaks, CA, USA, 1994.
51. Waarstaatjegemeente.nl. Bouwen en Wonen: Utrecht [Internet]. 2019. Available online: <https://www.waarstaatjegemeente.nl/dashboard/Bouwen-en-wonen> (accessed on 14 January 2020).
52. Gemeente Utrecht. Visie op de Warmtevoorziening in Utrecht: Naar een Klimaatneutrale Stad [Internet]. 2017. Available online: <https://omgevingsvisie.utrecht.nl/thematisch-beleid/energie> (accessed on 14 January 2020).
53. Utrecht, G. Utrecht Monitor: Energie [Internet]. 2018. Available online: <https://www.utrecht-monitor.nl/fysieke-leefomgeving-groen/milieu-duurzaamheid/energie> (accessed on 14 January 2020).
54. Van den Wijngaart, R.; van Polen, S.; van Bommel, B.; Harmelink, M. *Potentieel en Kosten Klimaatneutrale Gebouwde Omgeving in de Gemeente Utrecht*; PBL: Den Haag, The Netherlands, 2018.
55. Klimaatmonitor. Totaal Energiegebruik Woningen [Internet]. Available online: <https://klimaatmonitor.databank.nl/dashboard/Dashboard/Woningen/Totaal-energiegebruik-woningen--781/> (accessed on 14 January 2020).
56. QSR International. NVivo [Internet]. 2019. Available online: <https://www.qsrinternational.com/nvivo/what-is-nvivo> (accessed on 14 January 2020).
57. Henry, D.; Dymnicki, A.B.; Mohatt, N.; Allen, J.; Kelly, J.G. Clustering methods with qualitative data: A mixed-methods approach for prevention research with small samples. *Prev. Sci.* **2015**, *16*, 1007–1016. [CrossRef]
58. Macia, L. Using clustering as a tool: Mixed methods in qualitative data analysis. *Qual. Rep.* **2015**, *20*, 1083–1094.
59. Riley, J. *Understanding Metadata: What Is Metadata, and What Is It For*; NISO Press: Baltimore, MD, USA, 2017.
60. Van den Wijngaart, R.; van Polen, S.; van Bommel, B.; Harmelink, M. Het Vesta MAIS Ruimtelijk Energiemodel voor de Gebouwde Omgeving [Internet]. 2017. Available online: <https://www.pbl.nl/publicaties/het-vesta-mais-ruimtelijk-energiemodel-voor-de-gebouwde-omgeving-algemene-beschrijving> (accessed on 14 January 2020).

