smart cities

Article

Towards Municipal Data Utilities: Experiences Regarding the Development of a Municipal Data Utility for Intra- and Intermunicipal Actors within the German City of Mainz

Philipp Lämmel [1,*], Jonas Merbeth [2], Tim Cleffmann [2] and Lukas Koch [2]

1 Fraunhofer Institute for Open Communication Systems FOKUS, Kaiserin-Augusta-Allee 31, D-10589 Berlin, Germany
2 Daten-Kompetenzzentrum für Städte und Regionen DKSR GmbH, Winterdeldtstraße 21, D-10781 Berlin, Germany
* Correspondence: philipp.laemmel@fokus.fraunhofer.de

Abstract: This paper describes the requirements analysis phase towards the establishment and implementation of a municipal data utility (KDW = Kommunales Datenwerk, German) to facilitate data sharing between intra- and intermunicipal stakeholders. Against the backdrop of increasing digitisation and the growing importance of data-driven decision making in municipal governance, this paper aims to address the pressing need for efficient data management solutions within and across municipalities. Based on a structured self-developed methodology, the authors use a qualitative research approach: the paper examines the experiences and challenges encountered during the requirements phase, the design phase, and the development phase of the KDW. The findings indicate that the establishment of a robust KDW requires (1) extensive stakeholder engagement, (2) clear governance structures, and (3) a robust technical infrastructure. In addition, the study highlights the critical importance of establishing a sound legal framework that addresses data ownership, privacy, security and regulatory compliance. Addressing legal and regulatory barriers to data sharing is paramount to the successful implementation and operation of the KDW. The paper concludes by highlighting the potential benefits of KDWs and outlining future work. The overall methodology, approach, and outcome are validated within the city of Mainz, and the lessons learned are accommodated in the insights presented in the rest of the paper.

Keywords: smart city; data sharing; urban data platform; interoperability; security

check for updates

Citation: Lämmel, P.; Merbeth, J.; Cleffmann, T.; Koch, L. Towards Municipal Data Utilities: Experiences Regarding the Development of a Municipal Data Utility for Intra- and Intermunicipal Actors within the German City of Mainz. *Smart Cities* **2024**, *7*, 1289–1303. https://doi.org/10.3390/smartcities7030054

Academic Editor: Pierluigi Siano

Received: 22 April 2024
Revised: 7 May 2024
Accepted: 9 May 2024
Published: 28 May 2024

1. Introduction

A multitude of actors is responsible for the provision of services of general interest and welfare in a city/municipality. Despite local differences, the basic structure is often quite similar: the administration shares the provision of services of general interest with municipal enterprises. This leads to recurring challenges in different cities, which can be addressed with similar approaches thanks to comparable structures. This is particularly true for digitisation and data use. However, there is a lack of a market-ready solution that enables local actors to work together on a legally secure data basis to ensure the delivery of complex services and infrastructure within a municipality. From environmental agencies to social services, from transport to theatres, and from hospitals to energy and water, each actor provides 'its' services. However, complex challenges such as traffic changes, climate protection or pandemics require networked and jointly developed solutions. Evidence-based decisions in services of general interest can only be made if (partly) restricted data from a wide range of actors are accessible. Until now, however, these data have remained in the silos of individual organisational units and have not reached their full potential [1]. There are many reasons for this: lack of control, undefined legal responsibilities, data protection, including critical infrastructure data, and a lack of easily accessible tools for comprehensive data management. In addition, municipal data management solutions currently focus on the publication of specific data, such as open data or geospatial data. Current solutions do not dare to deal with (partially) restrictive data or high-value data. However, for efficient services of general interest, it is crucial for municipal companies to be able to share these data securely with each other. Until now, there has been no technical platform solution for this and, above all, no platform solution that is secure in terms of European and German municipal law. Thus, in the project »(Inter-) Kommunale Datenwerke: Konzeptionierung, Evaluierung und Umsetzung einer kommunalen Data Sharing Plattform« [2], we propose a municipal data utility (KDW = Kommunales Datenwerk, German) for the city of Mainz (Germany) that is designed as a virtual space for the exchange of restrictive data between intra- and intermunicipal actors. The KDW standardises data exchange for municipal actors from a legal perspective and implements an appropriate role and rights model. An additional focus of the envisioned open source-based solution is on user friendliness, thus increasing the acceptance of the proposed KDW. The main features of the proposed KDW are as follows:

- Build the KDW based on a legally secure foundation with legally secure contracts, which are developed in steps during the project.
- Secure data access through user rights model according to the requirements of municipal stakeholders.
- Secure inter- and intramunicipal data exchange and implementation of applications.

In their paper, He et al. examine the role and potential for improvement of legal governance in the construction of smart cities. Their paper describes the digital economy index of 31 provinces in China from 2014 to 2020 and analyses the function of legal governance in the development of local smart cities [3]. He et al. identify the data security issue, data alienation issue, public data opening, and sharing issue as the most critical challenges in the context of smart city development. Elvas et al. propose a novel approach to health data access in a hospital environment hereby using homomorphic encryption to ensure privacy and secure data sharing. Their proposed Information Sharing Infrastructure (ISI) framework enforces GDPR adoption and integrates artificial intelligence (AI) capabilities for data analysis as well as data-sharing agreements [4]. Šestak and Copot outline key challenges to data sharing in agri-food chains that hinder the establishment of an Agricultural Data Ecosystem (ADE). The challenges identified include stakeholder mistrust, inadequate data access and use policies, and unclear data ownership agreements, and interoperability [5]. Finally, Šestak and Copot propose a set of design principles for designing and implementing an agricultural data space that aims to mitigate the negative impact of the abovementioned challenges [5]. He et al. analyse the characteristics of the multi-combination of power grid data across services, and introduce a blockchain-based system in their paper. The overall architecture of the blockchain system is presented, implemented using the Hyperledger open source framework and evaluated [6].

In contrast to the previously mentioned papers, the contributions of this paper are the development of a holistic methodology for the creation of Municipal Data Utilities to facilitate data sharing between intra- and intermunicipal stakeholders. The overall methodology, approach, and outcome have been validated within the city of Mainz, hereby addressing the needs of cities in Germany. Germany is a federation and comprises sixteen constituent states; each has its own constitution, and is largely autonomous in regard to its internal organisation. Thus, a legally secure foundation with legally secure contracts is of paramount importance in the context of secure data sharing in Germany. Given the many uncertainties related to data sharing [7], such a legally secure framework is not yet implemented in Germany and symbolises a new way to developing data-sharing solutions. Initiatives relevant to the development of such a framework will be considered during the development of our framework hereby ensuring longevity and its relevancy. The aim of this paper is to contribute our methodology to the body of knowledge and to describe the status of the project implementation after the requirements analysis phase and to present the first lessons learned: Section 2 briefly describes the concept of an open urban platform and its relevance for the development of the envisaged KDW. The main characteristics of such a platform are discussed. Section 3 outlines the methodology used to develop the KDW. It describes the overall process and the steps involved in the design and implementation of the data utility. Section 4 then presents a specific case study that demonstrates the application of the proposed methodology. It will show how the municipal data utility is currently being developed in the city of Mainz in Germany, covering various aspects such as requirements analysis, legal framework, and system design. Section 6 will summarise the main findings and conclusions of the research paper. Areas for further research will also be discussed.

2. An Open Urban Platform—The Foundation of Our Municipal Data Utility

According to Hess and Koch, urban data platforms enable the realisation of various use cases in fields of action of urban and regional development (e.g., mobility or climate protection) by making data available in a city or region [8]. The term "urban data platform" is used in many different and inconsistent ways. In practice, urban data platforms can have many different implementations that differ in terms of user groups, data types and use cases. In general, however, urban data platforms collect, store and process data or information generated in a municipality and make this information available to the urban community or selected user groups [9]. In the context of this paper, an urban data platform is a set of software components that aim to aggregate data from different sources in an urban context into a logically centralised data platform. Hess and Koch state that an urban data platform typically enables the following use cases [8]:

- **The data platform as a data catalogue:** The data platform allows data providers to share their data in a way that users of the data can access it. For example, city departments can store their data on the platform so that the city community can access it.
- **The data platform as a data presentation interface:** The data platform provides access to target group-specific data. For example, interested citizens can compile and use visual representations of the data for their specific needs, with accompanying text if necessary.
- **The data platform as an enabler of further services:** Providers of software solutions and services can access the data on the data platform to enhance their services. This can be used to optimise processes within the administration or to enable applications for the urban community.

As mentioned by Brutti et al, in order to be able to fully exploit the Smart City vision potentials, solutions should be based on two fundamental concepts: open data and interoperability [10]. Thus, to bring to market an open urban platform that achieves these required characteristics and to provide an open source tool for cities and municipalities, DKSR has developed an open source version, namely *DKSR Open UrbanPulse (OUP)* [11]. It is an open real-time urban platform for cross-domain data integration and matching. The OUP is a real-time sensor data platform that follows the vision of Open Urban Platforms as expressed by the *European Innovation Partnership Smart Cities and Communities EIP* [12] and defined in *DIN SPEC 91357* [13], hereby also focusing on technical aspects such as interoperability and security by design [14]. OUP is based on open standards which can be seen as an additional crucial enabler for interoperability. However, the rigorous use of open standards is still a challenge in the context of open urban platforms [15] and thus, OUP is providing an essential tool for solving this challenge. OUP is aimed at cities, municipalities, utilities and businesses that want to use new and existing urban data sources to create innovative value-added services. To this end, OUP enables both small and large cities and companies to efficiently integrate new and existing data sources, process and analyse data in near real time, and finally share the data with various stakeholders. In addition to these capabilities, OUP also allows existing static data to be combined with near real-time data from sensors. Unlike many other platforms, OUP is cloud agnostic and can run in different clouds, in container environments such as *Kubernetes*, or in the local data centre.

Figure 1 shows the structure of the OUP Core and its related modules. The OUP Core consists of the following main modules:

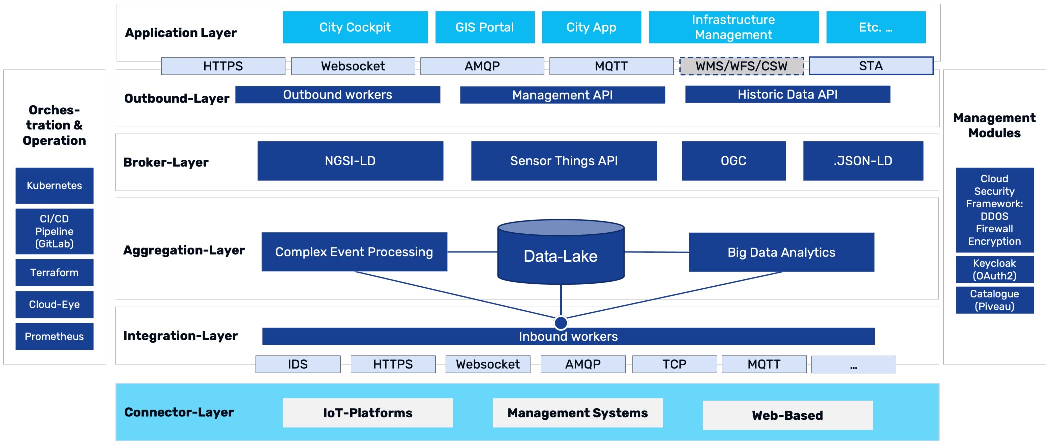

Figure 1. Software Architecture (high-level) of DKSR OUP Core Platform. Own representation. © DKSR GmbH.

UrbanPulse has a scalable micro service architecture for data processing and analysis, with a dedicated connector framework for the easy integration of sensors and other urban management systems. Ultimately, OUP provides fully integrated access to urban sensor data from different urban domains based on intelligent services. In this way, OUP combines a city's multiple data sources to better visualise, understand, and respond to all of the city's needs. Municipalities, businesses and individuals can use the information collected, generated and delivered by OUP to optimise their individual decisions and improve their digital services and processes. The core services are written in Java and can run on different [virtual] infrastructures. The platform consists of different types of modules that provide specific services to the other modules of the platform. To provide a service, the modules can use existing Software-as-a-Service (SaaS) technologies (e.g., databases, storage, and connectivity). Due to its modular architecture, the platform supports a scale-up strategy by increasing the virtual resources of a node (CPU, memory, etc.) as well as a scale-out strategy, where individual modules can be deployed multiple times on different nodes. The main functional requirements are a holistic view of data from different urban domains (e.g., mobility, energy, environment and administration) in order to identify optimisation opportunities.

Connector Layer

The connector layer is the data ingestion layer for the OUP Platform. It enables the consumption of data from the various IoT platforms, management systems, and web-based services. It communicates the data to the OUP core using HTTPs.

Integration Layer

This layer is the interface to the connector layer, which uses HTTPs connections to allow the OUP Core to consume the data and events coming from the connector layer.

Aggregation Layer

This layer allows the data consumed to be stored in a database built into the platform. It allows the stored data to be queried, processed, and exposed. It also allows further analysis, manipulation, aggregation, and refinement of the stored data using complex event-processing mechanisms.

Broker Layer

This layer is used to structure distributed software systems with decoupled, remote components. It thereby sits in between the aggregation layer and the outbound layer.

Outbound Layer

Once the data are stored, the outbound layer allows the data to be accessed using the REST and WebSocket interfaces. The OUP Management API provides access to the various modules and entities exposed by the OUP, and the Historical Data API provides access to the historical data stored. Real-time data stored on the platform are accessible via WSS and HTTPs interfaces, while historical data are available via HTTP interfaces.

Application Layer

The application layer enables specific digital applications, such as a CityApp or an existing GIS portal, to display the collected and processed data. It helps to connect to other systems. Throughout the system, the OUP is orchestrated by a functioning CI/CD pipeline and several software tools. This allows the system to run and operate smoothly and efficiently. Additional management modules help to organise the system and provide secure authentication.

3. Proposed Methodology for the Development of a Municipal Data Utility within the City of Mainz

The following paragraphs explain the integrated methodology (depicted in Figure 2) for the development of the municipal data utility (KDW). As stated above, the overall objective of the project is to implement the technical and legal design of an inter- and intramunicipal data utility based on a concrete use case within the city of Mainz. For this purpose, a municipal data utility based on open-source technology will be developed and tested from a technical perspective as well as being evaluated from both a technical and legal perspective. The practical use will provide further requirements. The resulting software will provide the technical basis for municipal data utilities and thus facilitate the use of data for new services of general interest. To this end, the project will pursue the following milestones:

1. The functional and organisational data governance for a municipal data-sharing platform is identified, understood, and implemented according to the needs of the practice partners. Included is an identity and access management system that is appropriate to the local government structure and allows for fine-grained management of data access.
2. Access to different data can be granularly controlled by the data owner.
3. A legally stable framework for a municipal data system has been developed.
4. The live operation of the municipal data system in the city of Mainz will be implemented using the example of a mobility-related application, including the connection of at least four data sources (at least one partially restrictive or restrictive data set).
5. The use case PoC (public transport + parking) examines the added value at the interface between private and public transport, which is created by combining the data and, if necessary, implemented outside the project via an existing MaaS application.
6. A concept for a data system and a business model for the municipal sharing platform (software) is developed.
7. The (partial) results of the project will be published as open documentation/open source code.

The methodology is designed to be as close as possible to the needs and requirements of real-world deployments, as the project stakeholders are the utilities and municipalities that are the direct customers of the municipal data utility. The first step is to create and align the overarching shared vision of all relevant stakeholders. This process also includes identifying, understanding, and implementing the functional and organisational data governance for the KDW in line with the needs of the relevant stakeholders. The identified governance structure is well documented and available to all relevant stakeholders. This first, initial step also aligns with research [3,5].

Secondly, the requirements analysis starts in parallel with the definition of the legal framework. The requirements analysis is based on the results coming from the initial step and includes typical activities such as desk research on data security and privacy, potential data sets, definition of user stories and use cases, and several other topics. Based on the results of the desk research, workshops are to be conducted to establish a common understanding of the legal framework. In addition, functional and non-functional requirements are derived and well documented in a requirements catalogue. In line with customer expectations, these requirements will be prioritised according to the Kano model of customer satisfaction [16]. In parallel, a legally stable framework for a municipal data system is being developed and documented, which will further support the acceptance of the planned KDW. As mentioned in Section 1, the definition of a legally secure foundation with secure contracts is of paramount importance in the context of secure data sharing in Germany. Given the fact that Germany is a federation, a legally secure foundation with legally secure contracts is important to establish the KDW.

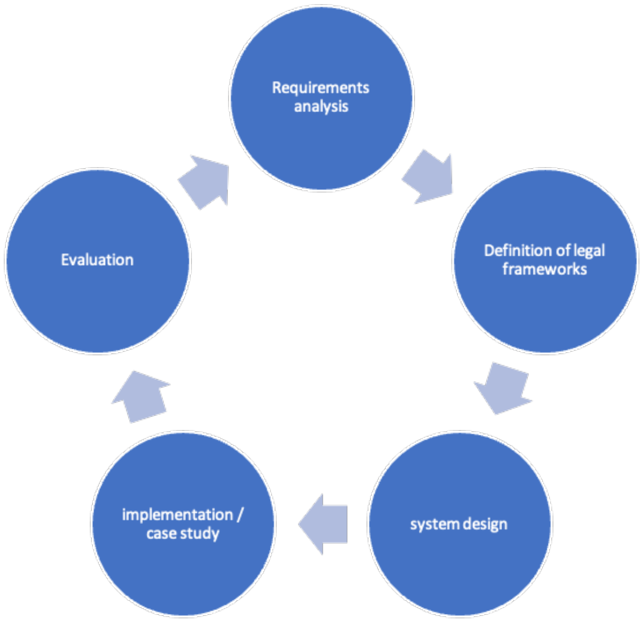

Figure 2. Methodology for implementing the municipal data utility.

Thirdly, based on the identified requirements and the legal framework, the development team designs the overall system, including the definition of the system architecture, the identification of the required software tools and development processes. An Identity and Access Management System (IDMS) based on Keycloak [17] will be integrated for securing data access and for supporting user federation, strong authentication, user management as well as role-based access control. The backend structure of the software will also be discussed and first drafts/mock-ups and wireframes for the frontend are about to be created in an iterative way. These mock-ups will be discussed with relevant stakeholders to ensure user acceptance.

Once the design phase is completed and the expected frontend is aligned with users' expectations, in the fourth and fifth steps, the implementation of the above system design is carried out and evaluated in a real production environment. The lessons learned from real-world deployment and operations are then fed back into the requirements analysis, enabling continuous improvement of the KDW.

4. Case Study

The following section presents the application of the proposed methodology to the city of Mainz and some lessons learned from the requirements analysis phase.

4.1. Requirements Analysis

As part of the project, interviews were conducted with experts and with potential users and stakeholders who are expected to benefit from the results of the project. The interviews were held digitally via Microsoft Teams and lasted between 30 and 45 min. For the evaluation of the results, the obtained answers were clustered into groups. The first step was to identify the potential users of the KDW. They belong to the following interest groups:

- Administrative and municipal authorities from Mainz:
  - Statistics and survey office;
  - Department of digitalisation;
  - Energy management of building industry;
  - IT-related departments and Geoinformation system.
- Municipal utilities and related companies of Mainz:
  - Publicly owned undertaking for economics;
  - Process and informational management of networks.

After identifying the groups of potential stakeholders, six qualitative interviews were conducted. Therefore, the so-called Mayring method [18,19] was used to analyse the answers. A pool of potential questions about their expectations and willingness but also their concerns about data sharing formed the basis of the survey. The questionnaire was supplemented with questions about existing software solutions within the municipality, data security requirements, and existing data formats. The interviews were conducted digitally, and each interview lasted between 45 and 60 min. All interviewees agreed to be contacted about the outcome of the project and possible further evaluation (qualitative or quantitative). All interviewee statements were co-authored and clustered into four different categories of requirements: functional, technical, organisational and legal. These clusters or codes were formed within the process and were used to subdivide and analyse the text. According to Mayring, the basic idea is that one "wants to get as close as possible to the object of research in order to derive new, differentiated questions" [18,19]. Therefore, the explorative approach is especially well suited for topics that have solely been rudimentarily researched [18,19].

Within the functional cluster, interviewees called for a central data-sharing portal where different data sets from different municipal stakeholders are collected. They asked for a standardised system of roles and permissions that would allow easy and efficient data sharing between different departments within the municipality.Advantages are seen in an increase in efficiency of the municipality, a better and easier fulfilment of administrative tasks and an easier coordination of responsibilities within projects. Another mentioned advantage was its expected output of evidence-based decision making on a quantitative (instead of qualitative) basis. Disadvantages were seen in the potential misuse of data for 'wrong' purposes and a potential misinterpretation of data, which can potentially lead to 'wrong' assumptions. The technical and managerial clusters revealed not only a need for a common data-sharing solution within the municipality but also a need to standardise data formats and reduce inefficiencies in data sharing. Geospatial data were considered particularly important. This and other data formats should be linked to administrative data from other levels, such as federal, state and regional data sets. The possibility of sharing data with other cities was seen as an important asset. A technical link to existing software solutions was also requested by some respondents. There was also a desire to link data sets from other departments using a so-called harvester, which would reduce manual effort when responsibilities or staff in a particular department change.

Finally, the legal cluster revealed expectations and requirements. Respondents pointed out that some of the data sets used (depending on the specific case) must be anonymised and pseudonymised. Legal frameworks such as the EU General Data Protection Regulation (GDPR) [20], European Data Governance Act [21] as well as Data Act [22] (enforcable sind January 2024 [23]) are to be considered European or national frameworks. At the state level, the so-called *Landestransparenzgesetz RLP* (State Transparency Law) must be taken into account [24]. By summarising the most important statements and implementing them in the created municipal data utility, users can easily understand the required data security of their data. In this context, one participant considered whether it would be appropriate to appoint a data protection officer within the municipality.

In addition, interviews were conducted with experts and initiatives working in the same field. Initiatives with comparable or similar projects and objectives, such as the municipality of Mönchengladbach/MobiData Baden-Württemberg and the city of Munich, helped to shape the project's objectives, conceptual, technical and legal requirements and to synchronise its steps with similar projects.

An interim report presented the current status of the project and the results of the qualitative expert interviews conducted. In addition, about 40 participants, mainly municipal actors, had the opportunity to give their opinion on the results and ideas from the expert interviews. Their votes and answers were collected using the Slido software [25]. The number of participants varied depending on the question. The answers were then sorted in descending order of frequency of mention.

When asked about the preferred legal framework for such a municipal data work, 41% (of 17 participants) answered that a nationwide GmbH (i.e., limited liability company) would be the best option for their municipality; 35% favoured the creation of classical so-called AGBs (terms of use); and only 24% voted for the development of municipal GmbHs (i.e., municipal limited liability companies).

In the next question, the participants had to choose three features that were preferred for the development of a municipal data works. The 25 participants had several options for their answer, but the preferred features were as follows:

1. The KDW is directly an open data portal (48%).
2. The KDW has an integrated component for analysis of data (48%).
3. The KDW allows map-based visualisations (40%).
4. The KDW has, at its disposal, data-sharing agreements (40%).
5. The KDW has integrated legal checklists (36%).

The open question on use cases beyond the sharing of so-called KRITIS (critical infrastructure data) data across organisational boundaries was answered differently. While some see use cases such as circular economy or neighbourhood development as important, others see use cases in the areas of disaster response, crisis management and extreme weather events. Other use cases mentioned were energy efficient neighbourhoods, developing climate models for a city, digital twins, optimising mobility and transport planning, scenario techniques for urban planning, and climate actions such as environmental sensing or climate change mitigation. When asked whether it should be possible to restrict the free publication of data including metadata, a majority of 62% said that it should be possible. This clear result indicates that a municipal data process should include the possibility to withhold information about some metadata.

In summary, the interviews helped the project team to understand and accommodate the different requirements of their potential customers. Through the use of this method, the expert interviews and the open workshop, the functional, technical, managerial and legal requirements were understood and categorised. The results formed the basis for the subsequent technical workshops.

Technical Workshops

A technical workshop aimed to collect the needed requirements and prioritise them according to the Kano model [16]. As the municipal data utility is a research project, the requirements were not fully clear right from the start. Therefore, requirements from similar projects, the provided answers of the conducted interviews (see above), and the expectations within the scope of the German mFund mobility smart city program [26] were brought together. In the next step, these high-level requirements were categorised in four groups and subdivided into smaller categories. This was performed by using the basic clustering technique of typical software engineering or requirement engineering processes:

1. **General** (with subgroups of general functionalities, expected results for concepts and guidelines and overall functional requirements);
2. **Technical** (with subgroups of infrastructure, usability, data mining and structure of metadata, roles and rights concept, data upload, data download, data handling and pipelines, APIs);
3. **Organisational** (with subgroups of management-structure and technical structure);

4. **Juridical** (with subgroups of general information and guidelines).

In an additional technical workshop, all collected requirements were discussed. The outcome was a prioritisation of them according to the Kano model [16]. This model helped to prioritise the clustered requirements into five categories:

1. **Must-be quality:** Requirements that are expected and taken for granted.
2. **One-dimensional quality:** Requirements that result in satisfaction when fulfilled and dissatisfaction otherwise.
3. **Attractive quality:** Requirements that provide satisfaction when achieved but do not otherwise cause dissatisfaction.
4. **Indifferent quality:** Requirements that are neither good nor bad and do not result in customers' satisfaction or dissatisfaction.
5. **Reverse quality:** Requirements that refer to a high degree of achievement resulting in dissatisfaction and to the fact that not all customers are alike.

In a final documentation step, the requirements are divided into functional and non-functional requirements and prioritised according to RFC 2119 [27]. For the English terms (in brackets), the German translations MUSS, DARF NICHT, SOLLTE (SHOULD), SOLLTE NICHT (SHOULD NOT) and KANN (MAY) are used. The term MUST therefore defines requirements that must be met. MUST NOT defines properties or events that must not occur. Requirements covered by the terms SHOULD and SHOULD NOT correspond to concrete recommendations. The term MAY, on the other hand, expresses optional "nice-to-have" requirements, the implementation of which is optional.

4.2. Legal Framework

The success of an intermunicipal data framework depends on ensuring a sufficient level of legal certainty. This is crucial to incentivise municipal partners to share data (on the donor side) and to motivate third parties to query data from the framework (on the recipient side). The "legally validated" label is essential for the data framework. The validation process needs to be designed "from the bottom up", meaning that the desired data access and use define the legal requirements for data collection and input. This is particularly relevant for aspects such as data selection (sources, granularity, format, and personal reference), data categorisation, the pseudonymisation or anonymisation of data, and the preferred justification norm (Art. 6 GDPR [20]). It should be noted that data do not necessarily lose value when personal references are removed; on the contrary, the statistical interpretability of anonymised data can often add value. It is essential to develop a clear understanding of the intended uses of the data framework. The evolving data concept needs to be validated, taking into account various legal areas, including data protection law (GDPR and possibly region-specific regulations such as the so-called 'Landestransparenzgesetz RLP' and more), IT security, tax law (considering data as an asset and the value of databases), intellectual property and commercial rights (software, database rights, trade secrets, and copyrights), as well as the newly enforceable Data Act [22,23]. Municipal and antitrust law may also be relevant, particularly in relation to competition between municipal enterprises and private companies. Given the dynamic regulatory environment at the EU level, it is imperative to keep abreast of the ever-changing legal frameworks in the data economy. Processes must then be defined to adequately reflect the legal requirements and ensure compliance by all stakeholders in the data space. The development and validation of a robust data policy depends on close collaboration between project leaders, IT specialists and legal experts. All three parties must share a common understanding of the capabilities of the data framework. On this basis, technical, procedural, and contractual requirements can be finalised and implemented. The final step is to draft the relevant terms and conditions, policies, and contracts.

As introduced in Section 1, He et al. examine the importance of legal governance in the construction of smart cities. Their paper focuses on China; the authors identify several challenges, such as data security issue, data alienation issue, public data opening, or sharing issue. This aligns well with the objectives of our proposed legal framework [3].

However, the 31 provinces examined in the paper do have a different legal background than cities in Germany, hereby highlighting the importance of a locally adapted framework. Another interesting discussion is provided in [7]: the authors discuss the need to reform existing legal frameworks and provide policy options. These options relate—among others—to the Data Act [22], the Digital Markets Act [28] and Section 19a GWB (German Act against Restraints of Competition) [29] as well as the legal framework for data intermediaries [7]. The provided options will be rigorously considered in the development of our proposed framework.

4.3. System Design

Figure 3 presents a high-level architecture of the envisioned KDW instance. It consists of three major layers, namely the inbound layer, data management layer, and the outbound layer. The layers are described below.

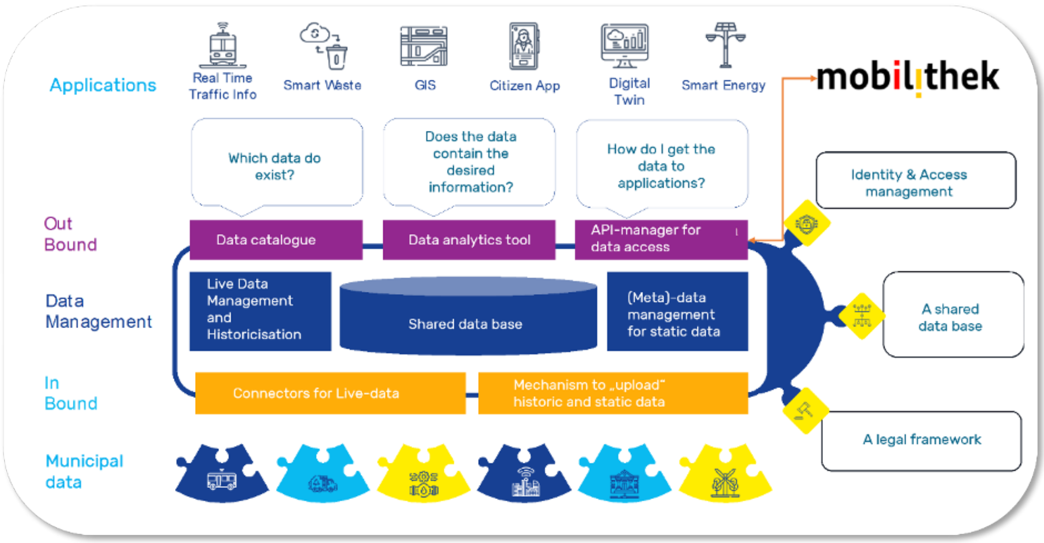

Figure 3. Software Architecture (high level) of KDW Instance. Own representation. © DKSR GmbH.

4.3.1. General Considerations

The project was set up by using a development and staging environment for the instances. Firstly, the data catalogue with the embedded filtering options was developed, according to the standardisation of DCAT-ap.de [30]. In the next step, the Identity and Access Management System (IAM-System) based on Keycloak [17] was developed by using the previously defined roles and rights, in line with the requirements of the city of Mainz. In the next step, a Frontend, developed with Figma [31], was created and was from then on used as the basis for future development. From this project phase on, the basic frame of the system was set, and additional tiles for different functionalities were added, such as a tile for the visualisation of dashboards for different KPIs and urban domains, and a tile for digital twins with precise requirements that come from research projects, such as Twinby in the German state of Bavaria [32], or the INTERREG-project Twins4Resilience [33]. Lastly, these tiles were embedded within the described frame of the original KDW and will give cities the possibility to integrate multiple and combining cases, domains, and projects within one software.

4.3.2. The In-Bound

Firstly, historic and static data from municipal stakeholders from different departments of Mainz (such as environment, traffic, economics, social, urban planning, building, etc.) are used as the basis of further analysis. These data sets will be brought together in the KDW and can be provided directly by users and data providers within the user-friendly frontend. Additional stakeholders outside the municipality can provide additional data sets after they are considered trustworthy (according to existing data security and privacy standards) and relevant data providers. Near real-time data and live data from devices within the built environment can be added to the existing sources. The thus needed connectors to merge the collected data into the KDW are provided by DKSR. The already developed 170 connectors will help to merge these data rapidly and efficient within the new data space.

4.3.3. The Data Management

A user-centric frontend and a structured backend with a connecting middleware allow easy data management. Live data and historic/static data will be combined within a shared database. These data sets can be researched quickly and easily with metadata filters according to DCAT-AP.de [30]. An Identity Management System (IDMS) allows only designated users to up- and download restricted data sets. Sets of open data will be provided for all users of the portal. A common database will store all data sets. A legally secure framework will enable security from a legal perspective.

4.3.4. The Out-Bound

The outbound within the frontend is a catalogue with metadata about the contained data sets and the underlying data sets themselves. An integrated data analytics tool helps to analyse the input data. On the basis of historic, static, and real-time data, synthetic data based on algorithms can be generated. An API-Manager helps to manage the data sets and connectors. The imported, combined, and analysed data can be used for several use cases, such as real-time traffic information, smart waste management, urban planning with geographic information system, the development of a citizen app or a digital twin, or the analysis of smart energy districts.

The municipal data utility connects interests of several stakeholders in one digital space. The user-friendly frontend with its integrated IAM-system, its data catalogue and the implemented connectors of DKSR-OUP will be the central data-sharing space for stakeholders in Mainz and surroundings. The publication of its code as Open Source, and some data as Open Data will allow to scale up the results and experiences of this project for other cities and regions. Thereby, the expected main users will be the city municipality of Mainz with its various departments and resorts, the public services of Mainz, such as Mainzer Stadtwerke (Mainz utility companies in Mainz, Germany)) or Mainzer Verkehrsgesellschaft (mobility/transportation provider in Mainz, Germany), and other local and regional actors, such as PMG Parken in Mainz GmbH (local company managing municipal parking lots). An exchange of data with other portals, such as Mobilithek for traffic-related data (nationwide) [34] or Kommunale Datenzentrale Mainz (KDZ) (citywide) [35], is within the aim of this project.

4.3.5. Scaling

The KDW could be very effective in real-world scenarios, as it is combining different streams of information (such as e-Governance, digitization for local authorities, digital twins, digital and 'smart' solutions for various urban domains but mainly mobility and environment) within one system. Thereby, and in accordance of each respective cities' challenges, the KDW can provide a digital tool for the following:

1. Combine existing static data sets from municipalities with newly created dynamic data sets from IoT devices, sensors, cameras, and LoRaWAN networks.
2. Create a portal for the easy access and sharing of restrictive, semi-restrictive and open data sets.
3. Combine these data sets into dashboards or embed them into digital twins to find possible solutions for urban and regional planning departments on the basis of quantitative data.
4. Increase municipal efficiency and hasten planning processes due to the easy-to-access and easy-to-provide data sets.

Within the real world, the KDW can be used for multiple cases for all domains that shape urbanity, such as environment, mobility, economy, and sociological parameters and many more. A short exemplary case study on how the KDW can be used for urban planning (or more precisely traffic planning) will be explained in the following.

5. The Proof of Concept

While the overarching aim of the project strives for conceptualizing, designing, developing, and testing intermunicipal data works, a proof of concept aims in showcasing its usability and benefits. The proof of concept is about mobility stations or mobility hubs. Therefore, the aim is to identify and analyse potentially suitable locations within the city of Mainz for the construction of mobility stations or hubs. These stations combine different modes of transportation, such as shared mobility, public transport, bicycle and scooter rentals, and offers for car sharing in a single location. This allows users to quickly and efficiently change from one mode of transport to another. As the city of Mainz strives for constructing several of them within the urban realm, the project provides the possibility of using urban analytics and existing static, dynamic, and synthetic data to identify potentially suitable locations within urban areas by intersecting different data sets. The data sets used thusly can be seen in Table 1. It can be seen that in the context of this PoC, the combination of (especially closed) data sets provides a better holistic view and enables municipal stakeholders to better plan their endeavours, to better serve their citizens while at the same time being on the legally secure side regarding data sharing. A full, quantitative evaluation of the proof of concept will be conducted in the future.

Table 1. Identified data sets for the PoC.

Urban Indicator	Data Set	Organisation	Available
Geography	Base map for standard geodata	OpenStreetMap (OSM)	Open in OSM
Traffic	Public transport network of lines and passenger numbers	Rhein-Main-Verkehrsverbund (RMV)	Capacity utilization figures for passenger counting points/tickets sold on request, routes for public buses and streetcars
Traffic	Locations of sharing offers for bicycles, scooters, scooters, General Bikeshare Feed Specification (GBFS)	Private sharing providers such as VOI, LIME, BOLT, TIER, SÜWAG2GO	Closed and only on request (private providers), Open Data (MainRad stations)
Traffic	Car sharing offers, parking space register and locations/utilization of parking garages	Private car sharing providers such as Share Now, Flinkster, Book-n-Drive, MainzRIDER	Closed and only on request
Traffic	Locations of already built mobility stations and hubs, e-charging stations	Mobility stations and hubs (City of Mainz, Mainz Mobility), charging stations (NOW GmbH/BMDV)	Open in Mobilithek and Geodata Office Mainz
Traffic	Parking spaces for various means of transportation (bicycle, motorcycle, car)	City of Mainz	Open in Geodata Office Mainz

Table 1. *Cont.*

Urban Indicator	Data Set	Organisation	Available
Traffic	Registered vehicles by fuel type per stat. District	Statistical Office Mainz	Closed and only on request
Traffic	Floating Car Data (FCD)	INRIX	Closed and only on request
Economy and social affairs	Public WLAN hotspots, restaurants, supermarkets and leisure facilities	Mainz Geodata Office, OpenStreetMap	Locations of POIs are open
Economy and social affairs	Educational institutions (universities, schools, educational establishments)	Mainz geodata office	Locations of POIs are open
Economy and social affairs	Residents per stat. District	Mainz statistics office	Closed and only on request
Environment	Parks and green spaces	City of Mainz	POI locations are open

6. Conclusions and Future Work

The paper showcases the work conducted in our research project *"(Inter-) Kommunale Datenwerke: Konzeptionierung, Evaluierung und Umsetzung einer kommunalen Data Sharing Plattform"*, which aims to design and develop a municipal data utility (KDW) for the city of Mainz. The primary objective of this paper was to showcase the results of the requirements phase of the project and highlight the proposed methodology, which we believe to be applicable and extendable to various cities. After motivating the need for the proposed KDW, the paper briefly introduces the concept of an Open Urban Platform, which acts as the foundation of the KDW. We then introduce the methodology that is used during the project implementation, and provide a case study in which we applied the methodology to a real-world setting. We successfully applied the methodology to the current implementation of the KDW, which has allowed us to gain valuable insights.

Although we have already made significant progress in implementing the KDW, there are several areas that will be further explored: (1) Completing the legal framework: As we move towards the wider implementation and use of the KDW, it is critical to establish a robust legal framework that addresses issues related to privacy, security and ethical considerations. As mentioned above, this step will ensure compliance with relevant regulations and standards and build trust among stakeholders. Other, very relevant initiatives such as the provided policy options by Schweitzer et al. will also be considered during the finalisation of the legal framework [7]. (2) Continue implementation: While we have made significant progress in implementing the KDW, there are still additional features and functionalities that need to be developed. These include improving the data collection process, refining the knowledge discovery algorithms and improving the user interface to enhance usability and accessibility. (3) Evaluation of the KDW: It is essential to carry out a comprehensive, quantitative evaluation of the KDW in real-world scenarios. This evaluation will test the performance, scalability and reliability of the system. In addition, user feedback and satisfaction will be collected via User Acceptance Testing (UAT) to assess the system's usability and effectiveness in supporting decision making. It is envisioned to conduct UAT with actual users of the KDW, stakeholders involved in the development of the product, and business analysts, as well as end-user specialists. (4) Evaluate/adapt the methodology: Based on the findings, the proposed methodology will be further evaluated and adapted to address any limitations or challenges encountered during the implementation of the KDW. This evaluation process will include refining the methodology based on feedback from users and experts to ensure its effectiveness and adaptability in different cities. Besides the mentioned use case of mobility, other case studies for urban domains, such as environment, resilience, economy, or social aspects, are also thinkable. Besides using other urban indicators or parameters as the basis for testing the applicability of the developed software, wider geographical scopes are also conceivable for the respective object of investigation. Further interest was thereby already shown by one national-wide ministry, one German county, and various other municipalities. This could be seen as an indicator for the widely spread potential applicability of the developed software for various domains and geographical scales.

Author Contributions: P.L. was responsible for establishing the methodology and authoring the initial drafts for Sections 1, 2 and 4, ensuring a robust analytical framework. J.M. provided the conceptual foundation and expanded the narrative through his drafting of Sections 3–5. T.C.'s role was related to validation and review as well as editing. L.K. tasked with gathering data and resources, also contributed to the refinement of the manuscript through careful review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the German Federal Ministry for Digital and Transport under the funding number 19F2223. The authors are responsible for the content of this publication.

Data Availability Statement: Due to business reasons, restrictions apply to the availability of these data. Data were obtained from involved third parties and may be available on request from the third party.

Conflicts of Interest: Authors Jonas Merbeth, Tim Cleffmann, Lukas Koch are employed by the Data Competence Centre for Cities and Regions DKSR GmbH. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Raghavan, S.; Simon, B.Y.L.; Lee, Y.L.; Tan, W.L.; Kee, K.K. Data integration for smart cities: Opportunities and challenges. In *Computational Science and Technology, Proceedings of the 6th ICCST 2019, Kota Kinabalu, Malaysia, 29–30 August 2019*; Springer: Singapore, 2020; pp. 393–403.
2. DKSR. (Inter-)Kommunale Datenwerke–KDW. 2023. Available online: https://bmdv.bund.de/SharedDocs/DE/Artikel/DG/mfund-projekte/kwd.html (accessed on 2 February 2024).
3. He, W.; Li, W.; Deng, P. Legal Governance in the Smart Cities of China: Functions, Problems, and Solutions. *Sustainability* **2022**, *14*, 9738. [CrossRef]
4. Elvas, L.B.; Ferreira, J.C.; Dias, M.S.; Rosário, L.B. Health Data Sharing towards Knowledge Creation. *Systems* **2023**, *11*, 435. [CrossRef]
5. Šestak, M.; Copot, D. Towards Trusted Data Sharing and Exchange in Agro-Food Supply Chains: Design Principles for Agricultural Data Spaces. *Sustainability* **2023**, *15*, 13746. [CrossRef]
6. He, Q.; Liu, Y.; Jiang, L.; Zhang, Z.; Wu, M.; Zhao, M. Data Sharing Mechanism and Strategy for Multi-Service Integration for Smart Grid. *Energies* **2023**, *16*, 5294. [CrossRef]
7. Schweitzer, H.; Metzger, A.; Blind, K.; Richter, H.; Niebel, C.; Gutmann, F. Data Access and Sharing in Germany and in the EU: Towards a Coherent Legal Framework for the Emerging Data Economy. Report Humoldt-Universität zu Berlin. 2022. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4270272 (accessed on 15 April 2024). [CrossRef]
8. Hess, S.; Koch, M. *Urbane Datenplattformen. Von der Idee bis zur Umsetzung: Entscheidungshilfen für Kommunen*; Bundesinstitut für Bau-, Stadt-und Raumforschung (BBSR) im Bundesamt für Bauwesen und Raumordnung (BBR): Berlin, Germany, 2023.
9. Cuno, S.; Bruns, L.; Tcholtchev, N.; Lämmel, P.; Schieferdecker, I. Data governance and sovereignty in urban data spaces based on standardized ICT reference architectures. *Data* **2019**, *4*, 16. [CrossRef]
10. Brutti, A.; De Sabbata, P.; Frascella, A.; Gessa, N.; Ianniello, R.; Novelli, C.; Pizzuti, S.; Ponti, G. Smart city platform specification: A modular approach to achieve interoperability in smart cities. In *The Internet of Things for Smart Urban Ecosystems*; Springer: Cham, Switzerland, 2019; pp. 25–50.
11. DKSR. DKSR Open-UrbanPulse. 2023. Available online: https://github.com/DKSR-Data-Competence-for-Cities-Regions/DKSR-Open-UrbanPulse (accessed on 28 April 2024).
12. Commission, E. EIP SCC Market Place. 2023. Available online: https://smart-cities-marketplace.ec.europa.eu (accessed on 15 March 2024).

13. Heuser, L.; Lacroix, G.; Müller, C.; den Hamer, P.; Schouten, S.; Welmers, M.; Fastenrath, U.; Kraus, T.; Knaup, V.; Schonowski, J.; et al. DIN SPEC 91357, Reference Architecture Model Open Urban Platform (OUP). 2017. Available online: https://www.din.de/en/wdc-beuth:din21:281077528?sourceLanguage&destinationLanguage (accessed on 15 March 2024).
14. Schieferdecker, I.; Tcholtchev, N.; Lämmel, P.; Scholz, R.; Lapi, E. Towards an open data based ICT reference architecture for smart cities. In Proceedings of the 2017 Conference for E-Democracy and Open Government (CeDEM), Krems, Austria, 17–19 May 2017; IEEE: Piscateville, NJ, USA, 2017; pp. 184–193.
15. Hernández, J.L.; García, R.; Fischer, M.; Schonowski, J.; Atlan, D.; Ruohomäki, T. An interoperable open specifications framework for smart city urban platforms. In Proceedings of the 2019 Global IoT Summit (GIoTS), Aarhus, Denmark, 17–21 June 2019; IEEE: Piscateville, NJ, USA, 2019; pp. 1–7.
16. Kano, N. Attractive quality and must-be quality. *J. Jpn. Soc. Qual. Control* **1984**, *31*, 147–156.
17. Keycloak. 2024. Available online: https://www.keycloak.org (accessed on 17 March 2024).
18. Mayring, P.; Fenzl, T. Qualitative Inhaltsanalyse. In *Handbuch Methoden der Empirischen Sozialforschung*; Baur, N., Blasius, J., Eds.; Springer Fachmedien Wiesbaden: Wiesbaden, Germany, 2019; pp. 633–648. [CrossRef]
19. Mayring, P. *Handbuch Qualitative Forschung in der Psychologie: Band 2: Designs und Verfahren-Qualitative Forschungsdesigns*; VS Verlag für Sozialwissenschaften: Wiesbaden, Germany, 2020. [CrossRef]
20. The European Parliament; The Council of the European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation). 2023. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016R0679 (accessed on 20 March 2024).
21. The European Commission. European Data Governance Act. 2023. Available online: https://digital-strategy.ec.europa.eu/en/policies/data-governance-act (accessed on 23 March 2024).
22. The European Commission. Data Act: Commission Proposes Measures for a Fair and Innovative Data Economy. 2022. Available online: https://ec.europa.eu/commission/presscorner/detail/en/ip_22_1113 (accessed on 19 April 2024).
23. The European Commission. European Data Act Enters into Force, Putting in Place New Rules for a Fair and Innovative Data Economy. 2024. Available online: https://digital-strategy.ec.europa.eu/en/news/european-data-act-enters-force-putting-place-new-rules-fair-and-innovative-data-economy (accessed on 20 April 2024).
24. Rhineland-Palatinate, L. Landestransparenzgesetz (LTranspG) Vom 27. November 2015. Available online: https://landesrecht.rlp.de/bsrp/document/jlr-TranspGRPrahmen/part/R (accessed on 8 May 2024).
25. Slido. Audience Interaction Made Easy. 2024. Available online: https://www.slido.com/?experience_id=240323-z (accessed on 10 March 2024).
26. BMDV. mFUND–Unsere Förderung für die Mobilität der Zukunft. 2023. Available online: https://bmdv.bund.de/DE/Themen/Digitales/mFund/Ueberblick/ueberblick.html (accessed on 20 February 2024).
27. Bradner, S. RFC2119: Key words for use in RFCs to Indicate Requirement Levels, 1997. RFC Editor. Available online: https://www.rfc-editor.org/info/rfc2119 (accessed on 20 March 2024).
28. European Commission. Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on Contestable and Fair Markets in the Digital Sector and Amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act). 2022. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32022R1925 (accessed on 10 April 2024).
29. der Justiz, B. Gesetz Gegen Wettbewerbsbeschränkungen (GWB). 1958. Available online: https://www.gesetze-im-internet.de/gwb/ (accessed on 11 April 2024).
30. Zedlitz, J.; Priefer, E.; Hauptvogel, J.; Panic, D.; Volkening, N.; Kochmann, P. DCAT-AP.de Spezifikation 2.0. 2022. Available online: https://www.dcat-ap.de/def/dcatde/2.0/spec/ (accessed on 10 April 2024).
31. Figma. Figma: The Collaborative Interface Design Tool. 2024. Available online: https://www.figma.com (accessed on 19 April 2024).
32. Bayerische Staatsministerium für Digitales. Digitale Zwillinge für Bayern. 2024. Available online: https://twinby.bayern/de/startseite (accessed on 15 February 2024).
33. City of Utrecht; T4R-Transforming Territorial Planning with Local Digital Twins. 2024. Available online: https://t4r.nweurope.eu/ (accessed on 10 April 2024).
34. BMDV. Mobilithek.info-Mobilitätsdaten Deutschland. 2024. Available online: https://mobilithek.info (accessed on 1 May 2024).
35. Mainz. Kommunale Datenzentrale Mainz. 2024. Available online: https://kdz.mainz.de (accessed on 11 November 2023).