European Commission

DBL

DIGITAL BUILDING LOGBOOK

Technical guidelines for digital building logbooks

Guidelines to the Member States on setting up and operationalising digital building logbooks under a common EU framework.

10 November 2023 (Final Version)

DBL Study Team: Ecorys, TNO, Arcadis and Contecht

DG GROW

Glossary of terms and abbreviations

Term	Definition
AaaS	Authorisation as a Service
API	Application Programming Interface. Interfacing software that allows programmers to access existing software functionalities or data sources without needing to know their internal details (like data formats). Especially relevant are web-based API's (or Web API's) like RESTful API's over HTTP having no memory. See definition of REST below.
BIM	Building Information Model. Building information modelling is a process supported by various tools, technologies and contracts involving the generation and management of digital representations of physical characteristics of places.
BMP	Building Material Passport. A digital document listing all the materials that are included in a building to facilitate circularity decisions and reuse of materials.
CI/CD	Continuous Integration / Continuous Delivery.
CityGML	A standard to store and exchange 3D city models with semantics in the GIS domain. Source: Open Geospatial Consortium (OGC).
CKAN	Comprehensive Knowledge Archive Network, a free and open-source data management system
CQRS	Command and Query Responsibility Segregation design pattern.
DBL	Digital building logbook. This is "a common repository for all relevant building data. It facilitates transparency, trust, informed decision making and information sharing within the construction sector, among building owners and occupants, financial institutions and public authorities." A DBL can be seen as a simple "standalone" Digital Twin for buildings focussing on the data aspects.
DCAT	Data CATalog vocabulary recommendation by W3C. A catalogue is a collection of meta-data about data sets and data services.
DGA	European Data Governance Act. A key pillar of the European strategy for data, the DGA seeks to increase trust in data sharing, strengthen mechanisms to increase data availability and overcome technical obstacles to the reuse of data.
Digital Twin	Digital Twins are meant to bridge the gap between real-world physical objects and their virtual representations. They typically comprise data and software functionalities having a direct link to reality via sensors and sometimes even actuators. There are many kinds of digital twins from simple "stand-alone" via "predictive" to fully "autonomous" variants. For any kind of digital twin, the right data, including semantic alpha-numeric data, geometric data and linked documents, is essential.
DMP	Data Management Plan. A formal document that outlines how data are to be handled both during a project and after project completion.
DMS	Data Management System. A DMS is in the building sector also often referred to as a Common Data Environment (CDE). There is a large variety of DMSs. Some are document-based, and some are more semantic. Some treat data sets only as a whole, some can go into these data sets, etc.

Term	Definition
	A well-known open-source DMS is CKAN powering data hubs/portals worldwide, especially in the governmental sector for public data.
DPP	Digital Product Passport. A tool to enable sharing of key product-related information that is essential for products' sustainability and circularity. In a sense, a building is a complex product and a DBL or BMP could be considered a type of product passport.
EPBD	Energy Performance of Buildings Directive
EIF	European Interoperability Framework. The EIF gives specific guidance on how to set up interoperable digital public services. It provides 47 recommendations, which are organised around three main pillars (12 principles, four interoperability layers, and a conceptual model)
EPC	Energy Performance Certificate. They provide information on the energy efficiency of buildings and recommended improvements. They exist in all Member States but their implementation varies across countries. The EPBD revision aims to harmonise them.
GIS	Geographic Information System. It is a computer system for capturing, storing, checking, and displaying data related to positions on Earth's surface.
IaaS	Infrastructure as a Service
IAM	Identity and Access Management
IFC	Industry Foundation Classes (IFC) is a standard Data Model from the open BIM world by buldingSMART (bSI). It covered originally only buildings but the latest version (IFC4.3) also covers civil infrastructures (bridges, roads, rails). It has a semantic extension mechanism referred to as buildingSMART data Dictionary (bSDD). More info at: https://www.buildingsmart.org/.
HTTP	Hypertext Transfer Protocol: protocol for transmitting data such as HTML or JSON
JSON	JavaScript Object Notation: an open data interchange format. See: https://www.json.org/json-en.html
JSON-LD	A Linked Data variant of JavaScript Object Notation (JSON) that can be used as a JSON-based serialisation for RDF especially popular by web developers already familiar with JSON. JSON-LD can be regarded as JSON but introduces some extra key words to define linked data aspects like URI-based prefixes. See: https://json-ld.org/
Ontology	An ontology is a semantic data model in the W3C linked data/semantic web context. It defines the concepts, attributes, datatypes, relations and constraints/rules giving meaning to the actual (master/reference or project) data. Based on source: https://www.w3.org/standards/semanticweb/ontology.
OWL	Web Ontology Language – "The W3C Web Ontology Language (OWL) is a Semantic Web language designed to represent rich and complex knowledge about things, groups of things, and relations between things." Source: https://www.w3.org/OWL/.
PaaS	Platform as a Service
RDBMS	Relational Database Management System.

Term	Definition
RDF	Resource Description language (RDF) is the standard framework by W3C where all data (data sets and data models) is in the end represented by atomic 'subject-predicate-object' triples forming labelled graphs. The W3C recommendation can be found at: https://www.w3.org/TR/rdf11-concepts/.
RDFS	RDF Schema language (RDFS) is the simplest ontology language. It has standard vocabulary to defines concept (classes) and their links with RDF properties. OWL is a stronger language built on top of RDFS. https://www.w3.org/TR/rdf11-schema/.
REST	Representational State Transfer architectural style, which includes principles for developing APIs. See definition of API above. See: https://www.ibm.com/topics/rest-apis
RPO	Recovery Point Objective
RTP	Recovery Time Objective
SaaS	Software as a Service
SHACL	SHapes Constraint Language (SHACL) is a linked data language following the closed world assumption (all 'unknowns' become 'false'). Instead of a focus on logical derivation, it is more about useful data verification. https://www.w3.org/TR/shacl/.
SKOS	Simple Knowledge Organisation System (SKOS) is a vocabulary to define terms, multilingual labels and definitions. It can also define a 'weak' taxonomy in terms of higher and lower terms. It is a language to define thesauri. Technically it is itself define in OWL. SKOS is the linked data language to define dictionaries, providing the terms for ontologies. https://www.w3.org/TR/2009/REC-skos-reference-20090818/.
SML	Semantic Modelling and Linking standard, Part 1 (generic) and Part 2 (domain-specific) by CEN TC442 WG4 (support data dictionaries) ([11], [12]).
SPA	Single Page Application
SPARQL	SPARQL Protocol and RDF Query Language (SPARQL) is the linked data query language for RDF data. It can be compared to SQL for relational technology. The difference is that SPARQL is well-defined and on a higher semantic/logical level. https://www.w3.org/TR/sparql11-overview/.
SSL	Secure Sockets Layer.
SQL	Structured Query Language: programming language for relational databases.
Turtle	Turtle is a linked data format (also referred to as 'serialisation'). It is based on the RDF framework. Turtle introduces some modelling 'short-cuts' to efficiently model the triples to make it also human-friendly (next to machine -processable). The W3C recommendation can be found at: https://www.w3.org/TR/turtle/.
URI	Uniform Resource Identifier, which is a unique sequence of characters that identifies a logical or physical resource used by web technologies.
W3C	World Wide Web Consortium: the international community that develops open standards for the web.

Introduction

The *Technical guidelines for digital building logbooks* were developed on behalf of the European Commission's Directorate for the Internal Market, Industry, Entrepreneurship & SMEs (DG GROW) by a team of researchers from Ecorys, TNO, Arcadis and Contecht. The guidelines were developed over a time frame of 18 months in collaboration with stakeholders from the construction ecosystem.

The purpose of this handbook is to guide the Member States of the European Union (EU) in setting up and operationalising digital building logbooks (DBLs) under a common EU Framework. The ultimate goal is **to use this European-wide framework for exchanging built environment data seamlessly**. As such the document addresses primarily national, regional and local authorities. It can, however, also be of use to other organisations interested in developing DBLs or generally working with building data. The document provides step-by-step guidelines on implementing national-level DBLs as well as reusable materials and explanations to use the proposed DBL Framework.

This Framework consisting of a semantic data model, data management plan and data dictionary for European DBLs has been developed as part of this project. It suggests a DBL core model that can also be extended with additional elements. Moreover, complementing the technical aspects, you will find in this report guidance on the practical implementation of DBLs and socio-economic considerations.

As such, the handbook guides at three separate levels, namely 1) the **practical implementation steps**; 2) the **technical implementation**; and 3) the **economic implementation** of DBLs.

The work presented in this handbook and associated deliverables builds on a preceding study. This study provided among others a definition of DBLs (see Box 1) as well as various recommendations, two of which this study partially addresses. These were:

- to develop a standardised approach for data collection, data management and interoperability including its legal framework; and
- to develop guidelines for linking existing databases.

Box 1: Definition of a digital building logbook[1]

A digital building logbook is a common repository for all relevant building data. It facilitates transparency, trust, informed decision making and information sharing within the construction sector, among building owners and occupants, financial institutions and public authorities.

A digital building logbook is a dynamic tool that allows a variety of data, information and documents to be recorded, accessed, enriched and organised under specific categories. It represents a record of major events and changes over a building's lifecycle, such as change of ownership, tenure or use, maintenance, refurbishment and other interventions. As such, it can include administrative documents, plans, description of the land, the building and its surroundings, technical systems, traceability and characteristics of construction materials, performance data such as operational energy use, indoor environmental quality, smart building potential and lifecycle emissions, as well as links to building ratings and certificates. As a result, it also enables circularity in the built environment.

Some types of data stored in the logbook have a more static nature while others, such as data coming from smart meters and intelligent devices, are dynamic and need to be automatically and regularly updated. A digital building logbook is a safe instrument giving control to users of their data and the access of third parties, respecting the fundamental right to the protection of personal data. Data may be stored within the logbook and/or hosted in a different location to which the logbook acts as a gateway.

[1] Source: European Commission, Executive Agency for Small and Medium-sized Enterprises, Volt, J., Toth, Z., Glicker, J.et al., Definition of the digital building logbook – Report 1 of the study on the development of a European Union framework for buildings' digital logbook, Publications Office, 2020, https://data.europa.eu/doi/10.2826/480977.

The technical guidelines contain five chapters, several annexes and supporting documents. They are structured as follows:

1. **Why are digital building logbooks important**: An introduction to DBLs, their policy context, uses, benefits and relevance to policymakers and other actors.
2. **Implementation roadmap**: Step-by-step approach starting with developing use cases, identifying user groups and data needs, raising awareness and communicating benefits to stakeholders and implementing the DBL first at a small scale before extending.
3. **Technical implementation**: Guidance on implementing the DBL Framework including the proposed data architecture, technologies and the initial core DBL ontology and dictionary.
4. **Economic implementation**: Overview of the required costing and resourcing of setting up a national DBL platform focusing on skills, financial and technical resources.
5. **Legal aspects**: A review of risks to consider when implementing DBLs, such as legal, privacy and data risks including ideas on how to address them.

Next to these five chapters, the following annexes provide further information:

- **Annex I – Detailed information on the DBL Ontology**: It presents several diagrams that are fully generated from the lexical DBL Ontology in linked data form (RDFS)
- **Annex II – Overview of other supporting documents**: It presents other reports developed as part of this project that provide further context as well as the DBL Semantic Data Model.
- **Annex III – Data Management Plan template:** A simple template to follow when developing a data management plan.
- **Annex IV – Screened EU legislation and caselaw**: An overview of the documents screened for the review of legal aspects of DBLs.
- **Annex V – Relevant national datasets**: An overview of existing national datasets that can provide part of the input for national DBLs.

For further reading on the subject matter of building logbooks, we can also recommend the practical guidelines on building passports prepared by the United Nation's Global Alliance for Buildings and Construction.[2]

[2] The Building Passport: A tool for capturing and managing whole life data and information in construction and real estate, available here: https://globalabc.org/resources/publications/building-passport-tool-capturing-and-managing-whole-life-data-and.

1. Why are digital building logbooks important

Being a common repository for all building-related data, the potential for DBLs to become a unique tool that provides policymakers and private actors alike a **single point of entry to verified and trusted building data**. Having such access to verified and trusted building data will not only be a useful asset, but a must considering the increasing reporting requirements for buildings. A DBL supports policymaking and allows the unlocking of various use cases and business models.

In particular, the revision of the Energy Performance of Buildings Directive (EPBD), increases the reporting requirements for Member States and mandates the creation of national databases for Energy Performance Certificates (EPC). Similarly, the EU Taxonomy for sustainable activities has increased disclosure requirements on sustainable finance for financial institutions also related to investments in buildings and renovation. These policy developments further **stress the need for making building data more findable, accessible, interoperable and reusable** across the EU.

A central role of DBLs is provided in the European policy context with the EU supporting its development and implementation in Member States to act as **an integrator of built environment data required for policy goals** related to the green and digital transitions (see Box 2).

Box 2: EU Policy context for digital building logbooks

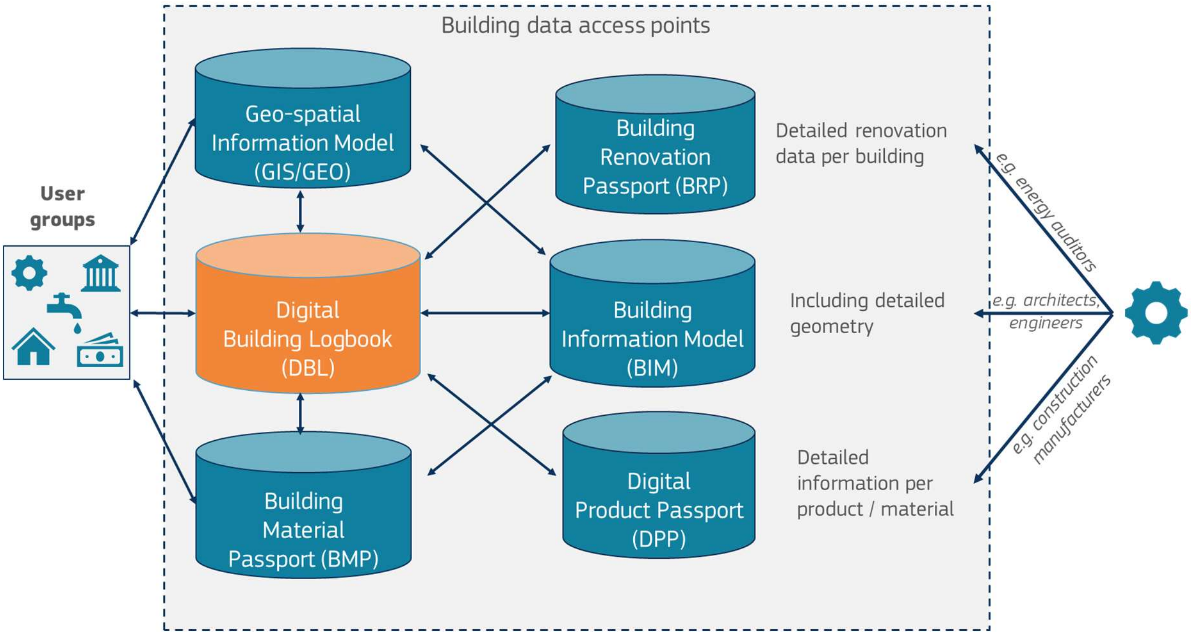
The Transition Pathway for Construction highlights that "Member States and regions are developing Digital Building Logbooks to track information on construction, renovations, material use and safety aspects". In particular, as part of the Green Deal, the 'Renovation Wave' communication outlines that the European Commission will introduce DBLs that will integrate all building-related data provided by the upcoming Building Renovation Passports, Smart Readiness Indicators, circularity indicators through Level(s) and EPCs to ensure compatibility and integration of data throughout the renovation journey. The aspect of circularity and the link with Leve(s) is also echoed by the 'Circular Economy Action Plan' which calls for developing DBLs to help promote circularity principles throughout the lifecycle of buildings in line with the 'Circular Economy Principles for Buildings Design'. Finally, the 'EU Climate Adaption Strategy' mentions DBLs as a tool to better integrate climate resilience considerations into the construction and renovation of buildings. For safety and quality of life, DBLs can support fire safety by providing an overview of electrical and gas installations and past safety inspections or regarding the prevalence of asbestos in buildings.

Regarding digitalisation and building data, DBLs are one of the tools enabling the digital transition of construction. They complement the digitalisation of building permits, the development of a construction data space, and the introduction of building information modelling (BIM) through public procurement. As such the Data Governance Act and Data Act are also key policies linked to DBL development as they aim to improve data sharing, availability and reuse and include measures for public authorities to access and use data as well as to develop interoperability standards. Similarly, for geospatial data, the INSPIRE Directive showcases the potential of creating European data infrastructures and therefore has been an inspiration for the DBL Framework. Other initiatives such as the revision of the Construction Product Regulation, the introduction of Digital Product Passports and the promotion of open standards also aim to further improve data availability and digitisation of information in the construction sector. In the medium to long term, these initiatives will therefore benefit DBLs by making more data available and accessible.

National **DBLs ultimately should act as the integrator bringing together all these separate initiatives and providing a one-stop-shop for building-related data**. This will not happen immediately but will be a step-by-step approach as many of these initiatives are being developed in parallel. However, In the long term DBLs can facilitate policymaking and contribute to research and monitoring by providing information to for example the EU Building Stock Observatory and the European Construction Sector Observatory.

A **Digital Building Logbook (DBL) should not replicate what is already there but connect and integrate existing data sources**, which are becoming increasingly available and accessible with improvements in digitalisation and standardisation. As such, it should be clearly distinguished from existing Building Information Models (BIM) and Geo-spatial Information Models (GIS/GEO), which often provide more detailed information about a building or its environment than a DBL would do. Similarly, parallel initiatives such as the Buildings Material Passport (BMP), Building Renovation Passport (BRP) and the Digital Product Passport (DPP)[3] with more detailed information on materials and the products used, can provide inputs to the DBL's global perspective. For this reason, the initiatives should co-exist, with a DBL linking to all others in some way. Figure 1 provides further context on these relationships.

Figure 1: Context of digital building logbooks



Note: The terms 'Building Passport' and 'Building Logbook' are often used interchangeably and refer to the same concept, however the logbook character of a DBL puts the time perspective, i.e. the ability to "log" changes into more focus. Please also note that this visualisation takes the perspective of a DBL and does not aim to be comprehensive as potentially many more interlinkages could be identified between each initiative. The linkages between DBL and BIM & GIS are further explained and detailed in report D2.8.

With these guidelines, the focus is not on the role of individual building logbooks, i.e. the use of your personal DBL to keep track of your house, but rather on the larger perspective of regional, national or EU-level DBL portals. There are also many benefits for individual owners or tenants in having easy access to their building data through a personal building logbook. For example, for individual reporting requirements or when selling a property. However, the **purpose of the proposed DBL framework is to connect these individual points** and provide a public good through easier access to trusted data on the wider building stock throughout its lifecycle.

The **long-term vision is a network of interoperable national DBL platforms that are connected at the EU level** through a European portal. The idea is to start small with basic building data and slowly add more datasets and functionalities to extend the potential use of these platforms. Public authorities will be able to develop their DBLs at their own pace and with their priorities in mind. Some might focus on energy performance and renovation while others focus on building permit

[3] The DPP aims to promote recyclability, circularity and compliance of products. DPPs need to be provided by product manufacturers for each of their products. Depending on how comprehensively the DBL is developed, it could contain (references to) many different DPPs for all the products and materials that are part of a building.

processes or the tax valuation of buildings. However, while doing so, this framework offers materials that Member State authorities can use to kickstart or improve their national DBL from a harmonised starting point, which could also facilitate interoperability between national DBLs in the future.

Ultimately, **DBLs are a tool that can have many uses by providing better access to better building data**. It is important that the Member States, supported by the EU, start the process of developing national DBLs as this will allow policymakers and private actors to unlock these use cases. In doing so, especially the urgency in meeting climate goals and renovation targets should be a key driver in the development of DBLs as managing this transition will be difficult without proper data.

1.1 Use cases for digital building logbooks

The reasons why users would use a DBL are numerous. Use cases relate to existing building processes as they showcase how these processes can be processed more efficiently and effectively with the help of DBLs. They might also relate to new services where a DBL could unlock new business models by combining information from different datasets and from data that was previously not (widely) available.

Often, it will be about extracting data from linked databases. Various use cases can be developed based on this data such as the analysis of data of a building or a group of buildings, retrieving initial data for the next step in the building's life cycle, and control of regulatory issues. Partly, the use of a DBL is government-driven and market-driven with actual use depending on data availability and data quality. Here, users, but also data owners, have a responsibility to get and keep this in order and to enrich the databases.

There are **five main user groups** or relevant stakeholders of DBLs:

- **Governmental agencies –** need a DBL to obtain insights for policymaking, issuing of licences and enforcement of regulation and disaster management. Sub-divided into national, regional and local levels of governance with different responsibilities but similar interests. National policymakers should be a driving force behind DBL implementation.

- **Construction sector –** need a DBL to obtain and report building-related information for the design and build. The sector encompasses varied sub-groups (e.g. architects, engineers, contractors, real estate, etc.) which one needs to consider for a full picture of the varying data needs and responsibilities.

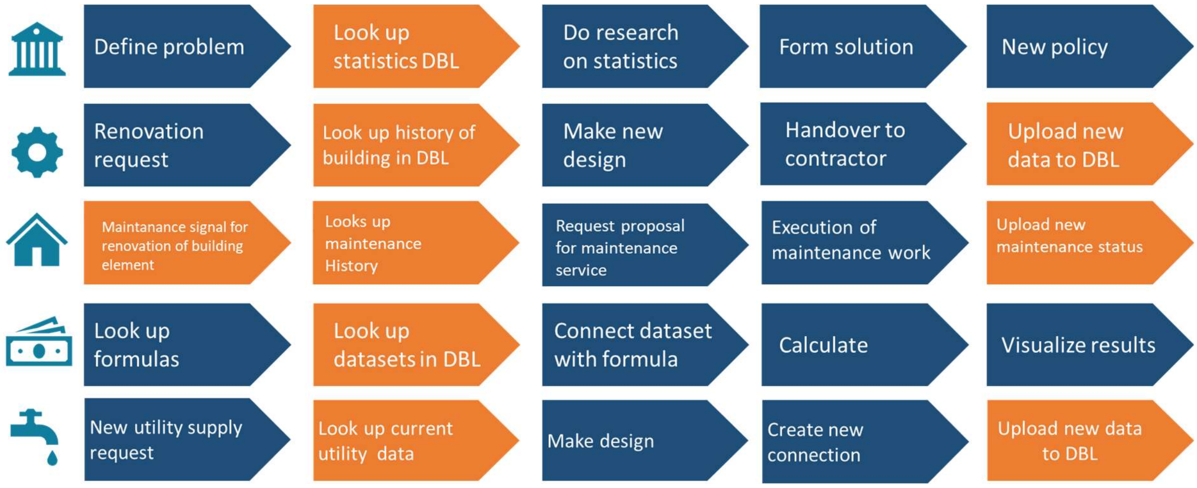
- **Building owners and users –** need a DBL as a basis for the collection of information about their buildings for exploitation and maintenance. The building owner often needs to be the linking pin to provide direct or indirect access to building information.

- **Financial institutions –** need a DBL to perform analysis based on formulas to gain insight into the assets market, and its developments and gain an understanding of building transactions. More transparency and quality of the data provided by the DBL benefits the financial sector

- **Utility companies –** need a DBL to obtain building information for the connection of utilities to the building analyses of its uses and provide information related to building use or performance.

The **benefits of DBLs differ by user group** as each of them has different uses for DBLs. Exploring all of these would be its own study, however, Figure 2 showcases one example for each of the groups and at which steps in their user journey they would be able to benefit from a DBL.

Figure 2: Exemplified user journeys of the five user groups (orange highlighted potential DBL use)



In general, developing user groups, use cases and user journeys around them is a helpful tool to understand who benefits and what the benefits are. Such an assessment can be useful in deciding what functionalities and building data to focus on and creating priorities in setting up a national DBL. You can learn more about how to develop a use case in Section 2.1.

1.2 European connectivity, Member State benefit

Ultimately, with this document, the aim is to guide Member States in setting up their national DBLs and in realising the benefits associated with improved building data while opening the opportunities for a future EU-wide DBL portal by ensuring interoperability across Member States. These benefits for Member States and stakeholders come in many facets, however, below you can find seven key reasons for implementing a national DBL following the European framework:

- **Interoperability**: A standardised implemented framework ensures consistency in data structure, formats, and terminology across different buildings and systems. It enables interoperability, making it easier to exchange data between various stakeholders, systems, and software applications. This promotes seamless collaboration, reduces information silos, and improves the efficiency of building management processes.
- **Keeping track of the state of the built environment**: A DBL provides a comprehensive standardised record of a building's design, construction, and operational history. It captures information about building construction, maintenance activities, renovations, and other relevant data. This data can be analysed throughout the whole of Europe and more specifically at the national level to identify inefficiencies, track performance trends, and make informed decisions to optimise energy consumption, reduce operational costs, and enhance overall building performance.
- **Regulatory Compliance**: The framework helps to ensure compliance with regulatory requirements related to building performance, energy efficiency, and sustainability. It provides a standardised method to collect, store, and report data that is necessary for demonstrating compliance with relevant regulations, certifications, and environmental targets. This streamlines the reporting process and facilitates auditing and verification procedures.
- **Transparency and Accountability**: A DBL framework promotes transparency by providing a repository on European and national levels of information about a building's performance and history. This transparency enhances accountability among building owners, facility

managers, and other stakeholders by enabling them to track and report on metrics related to energy consumption, emissions, and sustainability goals. It supports evidence-based decision-making and facilitates communication with regulators, tenants, and the public.

- **Lifecycle Management and Maintenance**: A logbook captures data throughout the entire lifecycle of a building, from design and construction to operation and maintenance. It helps stakeholders keep track of maintenance activities, equipment replacements, and warranties, which in turn supports proactive maintenance planning, reduces downtime, and extends the lifespan of building assets. The logbook also serves as a valuable resource for future renovations, retrofits, and decommissioning.
- **Data-Driven Insights and Innovation**: The two levels (European and national) of a DBL provide a rich source of data that can be analysed to gain insights into building performance, identify patterns, and inform decision-making. It supports data-driven innovation by serving as a basis for enabling the further development of advanced analytics, machine learning, and artificial intelligence applications for optimising energy consumption, predicting maintenance needs, and enhancing occupant comfort. The DBL can also facilitate research, benchmarking, and the sharing of best practices across the industry.
- **International Collaboration and Harmonisation**: The reference architecture and suggested implementation of the European DBL framework promotes collaboration and harmonisation of practices across Member States. It allows for the exchange of knowledge, experiences, and lessons learned, fostering a shared understanding of building performance and sustainability. This collaboration can lead to the development of common standards, guidelines, and benchmarks that drive continuous improvement and enable Europe to be at the forefront of sustainable building practices.

Overall, a European framework for DBLs is essential for promoting transparency, improving building performance, ensuring regulatory compliance, and driving innovation in the construction and facilities management sectors. It facilitates data-driven decision-making, supports sustainable practices, and contributes to Europe's broader goals of energy efficiency, carbon reduction, and sustainable development.

Beyond these seven reasons for implementing a national DBL one can further identify specific types of benefits related to data, economic and policy aspects (See Table 1).

Table 1: Overview of potential data, economic and policy benefits

Benefit	Description	
Data quality	A data governance structure and management plan ensure high-quality data in terms of its relevance, completeness, consistency, precision, simplicity, traceability, scalability and adaptability.	**Data aspects**
Data accessibility	Unlocking the use of existing data for various use cases.	
Data availability	Accelerating the digitisation of building data and making more data available and findable to be used.	

Benefit	Description	
Data verification and trust	With clear and harmonised definitions and checks of compliance with data norms, a DBL can become a single authoritative source of building data and reducing duplicates or errors.	
Level playing field	Improving the level playing field for building information and access to data for all market parties (transparency and accountability).	**Economic aspects**
Efficiency	Making existing services related to buildings and building data such as permits more efficient due to ease of access to information.	
Better services	Providing better services to buildings and building owners due to the better quality or availability of information (e.g. on past works).	
New services	New types of services and business models can develop around building data, due to the additional data about buildings being available.	
Regulatory compliance	Easing the implementation of existing rules and regulations as well as of new EU-level regulations such as the EPBD.	**Policy aspects**
Policy making	Improving future decision-making by policymakers through improved availability, access and quality of building data.	
Better analysis	Allowing for better analysis of risks (fire, gas, electricity, etc.), sustainability (circularity, energy performance, CO2, etc.), quality of life (living space, asbestos, etc.) and other aspects of buildings.	
Improved building performance	In the long-term, the above benefits can lead to an overall improvement in the building stock and the quality of life of citizens.	

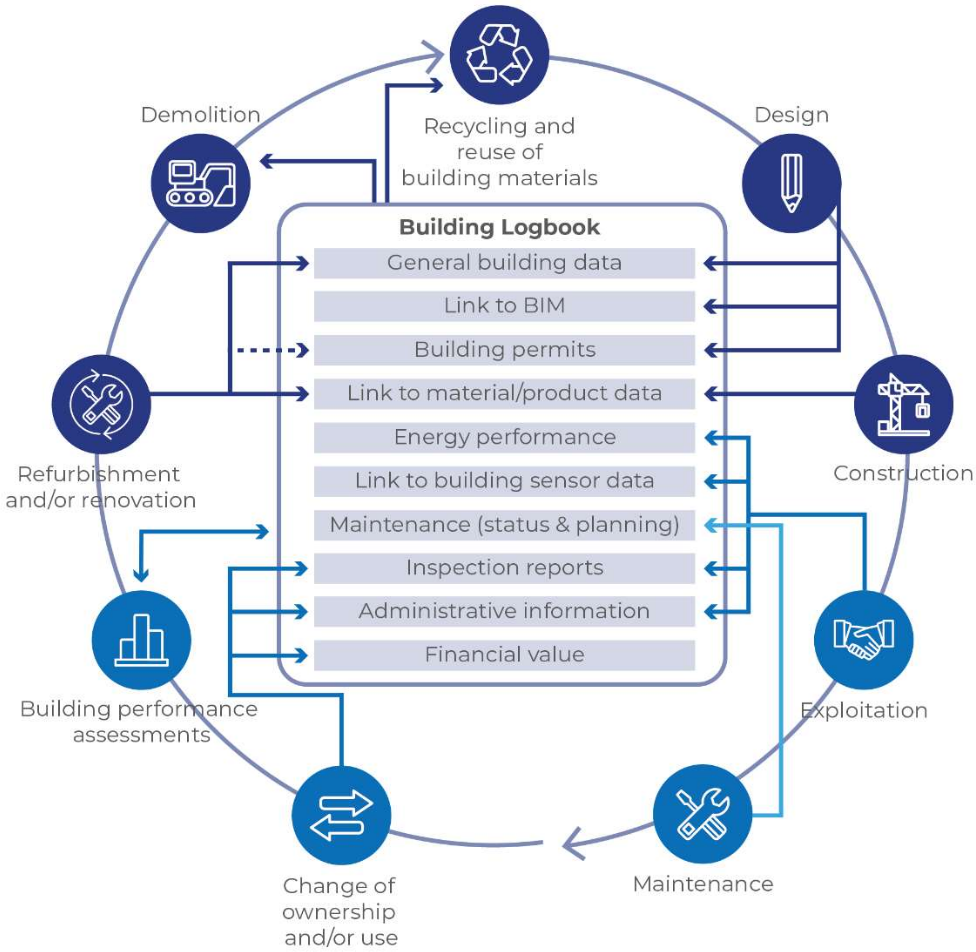
1.3 Digital building logbooks and building life cycle management

Currently, **the quality of available data in existing (governmental) databases varies and cannot always be trusted as a single source of truth**. Prominent stakeholders that use these databases are aware of their shortcomings and have ways to verify the data before usage. Data quality very much depends on which sources one looks at - most administrative data is of high quality - but when it comes to data related to performance, conditions and material use verification is usually needed.

In future building management, it would be preferable to have a single source of truth for building data. Every stakeholder can find their required data to analyse, renovate, connect, or maintain a building or a group of buildings. A DBL can act as such a single source of truth and in time, the **DBL should help to produce more trustworthy data as it will capture all (meta)data needed to increase reliability and to work on more standardising of data**. The DBL should enable insight into the level of data quality based on various metrics (see also Section 3.3.1). In addition, the development of use cases and business models as described in Chapter 2 will help clarify data requirements and therefore help raise data quality over time.

Regarding future building management, a building owner could benefit from the DBL through ease in regulatory compliance by for example accessing and sharing information on energy performance. Similarly, a building owner can more easily ask market parties for services such as financing, insurance or maintenance contracts based on the information provided directly or indirectly by the DBL. As such **a DBL can support building management throughout the life cycle of a building** from its inception following its design and construction, over its exploitation by recording changes in ownership, function and renovations, and finally to its end-of-life demolition and deconstruction (see Figure 3).

Figure 3: Overview of the building life cycle and use of a DBL



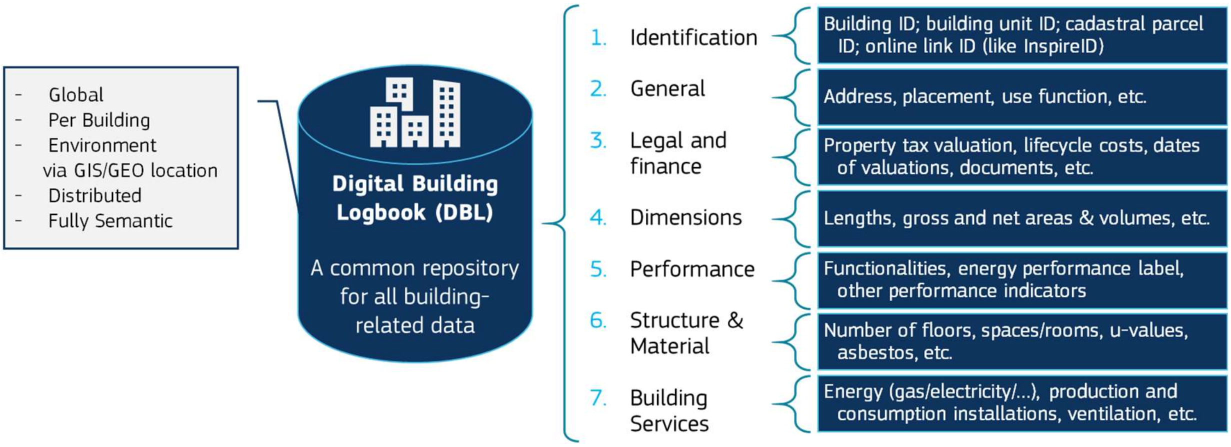
Throughout the life cycle phases, there are various moments in time when there is a dependence on DBL data and when various stakeholders create information that should be uploaded back into the DBL. A DBL should facilitate a smooth and easy entry and retrieval of data for all authorised parties.

1.4 The proposed DBL Framework

The European Framework for DBLs is designed to capture and manage comprehensive information about a building throughout its lifecycle. It distinguishes between the EU level with a European portal as a centralised point for access to the DBL and national portals for every Member State to implement their own, possibly enriched national logbook.

The ambition is to create a common repository for all building-related data as a single point of entry to verified and trusted building data. As such the DBL Framework provides a structure and core model for Member States to replicate. The main components of the DBL Framework are structured around seven aspects each containing various properties (see Figure 4). As data availability might vary across Member States, only the building identification aspect is required for the model to work, however, the more information can be added the more use cases can be unlocked through the DBL.

Figure 4: The main components of the DBL framework



Beyond its components, the DBL Framework also follows several key principles that should be taken into account when implementing a national DBL. Figure 5 on the next page summarises these seven principles. In the following chapters, you will find further details about the DBL Framework, its technical aspects as well ideas on how to implement it.

Figure 5: Principles of the European Framework for DBLs

Standardised Data Structure: The framework defines a standardised minimal mandatory data structure (see Annex I for the DBL Ontology) to ensure consistency and interoperability across different buildings and systems. This structure includes relevant information such as building specifications, construction details, energy performance, maintenance records, and any modifications or renovations. If needed, this Ontology could be enriched with the extra properties, important for the specific Member State.

Open Data Formats: The framework promotes the use of open data formats, such as JSON and JSON-LD (JSON for Linked Data), to facilitate data exchange and integration with other building management systems. This enables easy access and analysis of the logbook data by building owners, managers, and other stakeholders.

Logbook Security: Strong measures are in place to ensure the security of the logbook data. This includes user authentication and authorisation mechanisms, access control, and compliance with relevant data protection regulations within the EU.

Lifecycle Considerations: The logbook data structure in terms of ontology captures data throughout the entire lifecycle of the building, including design, construction, commissioning, operation, maintenance, and decommissioning. This comprehensive approach enables better understanding of the building's performance over time and facilitates effective decision-making for optimisation and sustainability.

Stakeholder Engagement: The framework involves collaboration with various stakeholders, including building owners, facility managers, architects, engineers, contractors, and regulatory bodies. The proposed solution provides the reference architecture, choice on tooling for the development standardisation and the suggestions on hardware and configuration management.

Accessibility and User-Friendly Interfaces: The digital logbook is designed to be accessible through user-friendly interfaces, such as web-based portal, making it easy for stakeholders to access, update, and analyse the logbook data. Visualisation tools and analytics capabilities enhance the usability and value of the logbook. The proposed tools are discussed further in this section.

Scalability and Future-Proofing: The framework is designed to accommodate future advancements in technology and evolving industry requirements. In long term it should be scalable to handle large amounts of data and flexible enough to incorporate new data types or standards as they emerge. That is why also the long-term vision discussed to show how more complex and scalable solution could look like.

2. Implementation roadmap

To ensure implementation and acceptance, it is important to not only focus immediately on the technical implementation but to also consider legal, political, and social aspects. To guide Member States and public authorities in implementing DBLs, this chapter will outline various steps to consider during the process. Please note, that each of these might need to be adapted to your unique situation. For example, Member States with federal systems or otherwise strong involvement of regional or local authorities in the management of their building stock will need to consider their involvement.

In short, for proper implementation, the following steps are suggested:

1. Defining the role and purpose of the DBL;
2. Determining the data needs of the DBL;
3. Quantifying benefits;
4. Identifying stakeholders, raising awareness and acceptance;
5. Analysing barriers to implementation;
6. Developing a working DBL prototype;
7. Preparing the legislative framework and an enforcement strategy;
8. Starting to implement and realising benefits and feedback loops.

Before implementing these eight steps, as step zero, it is recommended to **assess how pre-existing work within and outside of your Member State can support your development**. Existing components, tools and data can greatly facilitate the work of implementing a DBL. Some work might have already been implemented at the EU level, internationally, by market parties or in another Member State. Such existing tools can be used as building blocks and can be adapted to your purpose, thereby facilitating any or all of the eight steps.

2.1 Defining the role and purpose of the DBL

As a first step, one needs to define the role and purpose of the foreseen DBL. A national or EU legal requirement to collect or publish building data could provide a good starting point for the development of a national DBL. In most EU countries the national Cadastre collects data to meet certain legal requirements. Making these publicly accessible could be a good way to kickstart the development of a national DBL. However, a so-called use case methodology will still help to define the overarching goals (i.e. why we collect and publish the data) and define the pertinent data to be collected and published.

While DBLs could potentially cover all kinds of use cases throughout the lifecycle of buildings, it is recommended to **start small with one specific use case in mind** (e.g. supporting the process of building permits, energy performance, renovation roadmaps, building control, utility, sales, financial assets, circular economy, etc.). Defining such a use case allows the development of a value proposition and facilitates highlighting benefits for different stakeholders or so-called user groups. It will also help articulate the specific purpose and scope of a DBL and define the intended goals, objectives and functionalities, which in turn provide clarity to stakeholders and guide the development and implementation process.

It is also important to **identify and understand the needs of stakeholders** involved in the building management process, which means that governmental agencies, the various actors within the construction sector, building owners and users, financial institutions and utility companies are all represented.[4] Understanding their collective needs and requirements ensures that the logbook is designed to address all their specific needs and delivers value to all stakeholders.

[4] It should be noted that only a subset of these stakeholders is involved in most DBL initiatives and that which stakeholders to involve depends on the purpose and use case of the foreseen DBL.

2.1.1 Use case approach

A use case approach is used in system analysis to identify, clarify, and organise system requirements. The use case is made up of a set of possible sequences of interactions between systems, users and data in a particular environment and related to a particular goal. The method creates a document that describes all the steps taken by a user to complete an activity.

Depending on how in-depth and complex those interactions are, use cases describe a combination of the following elements:

- **User** – anyone or anything that performs a behaviour (who is using the system).
- **Stakeholder** – someone or something with vested interests in the behaviour of the system under discussion.
- **The goal** – the final successful outcome that completes the process.
- **The process** – the process and steps taken to reach the end goal, including data requirements and data sources.
- **Preconditions** – what must be true or happen before, during and after the use case runs.
- **Triggers** – these are the event that causes the use case to be initiated.

The general method to construct a use case typically involves the following steps:

Figure 6: Steps for implementing a use case analysis

1. Identify the user groups	2. Define the scope	3. Identify the goals	4. Customer journey mapping	5. Analyse the use case	6. Review and revise
Identify the different user groups involved in the system, including any users, systems, or external entities that interact with the system.	Define the boundaries of the system and determine what functions and features are included in the use case.	Identify the goals or objectives that each type of user is trying to achieve by interacting with the DBL.	Define the specific steps that each type of user (persona's) must take to achieve the goal, including any inputs, outputs, painpoints and/or interactions with the system.	Analyse the use case to identify any potential problems or issues and determine how they can be addressed.	Review the use case with stakeholders and revise as necessary to ensure that it accurately reflects the requirements of the DBL.

Overall, the purpose of a use case is to provide a clear, concise description of how each type of user interacts with the DBL to achieve a specific goal or objective. It helps to ensure that the system meets the needs of the user and that all functional requirements are properly defined and documented.

2.1.2 Customer Journey Mapping

Customer journey mapping is a visual representation of the entire customer experience with a system. It aims to capture the customer's interactions, emotions and touchpoints at various stages of their total journey. To find out where the DBL can add value for the user groups, their current experiences and processes need to be understood. Mainly the 'moments of misery' need to be indicated to identify use cases for the DBL that can lead to an improved user experience. A tool that is used to visualise actions, thoughts, and emotions during a process is called a customer journey map[5]. Moreover, it is also a tool to develop personas (descriptions of user groups and their needs).

User groups are a way of categorising the different types of users who will interact with the DBL or service. This helps to better understand their needs and behaviours, and ultimately design a product that meets those needs. To communicate what type of user groups are used to base the use cases, personas are used. A persona represents the goals and personal characteristics of the user group of interest. It contains information that is needed to understand the needs and desires of the users[6].

The general method of a customer journey involves the following steps:

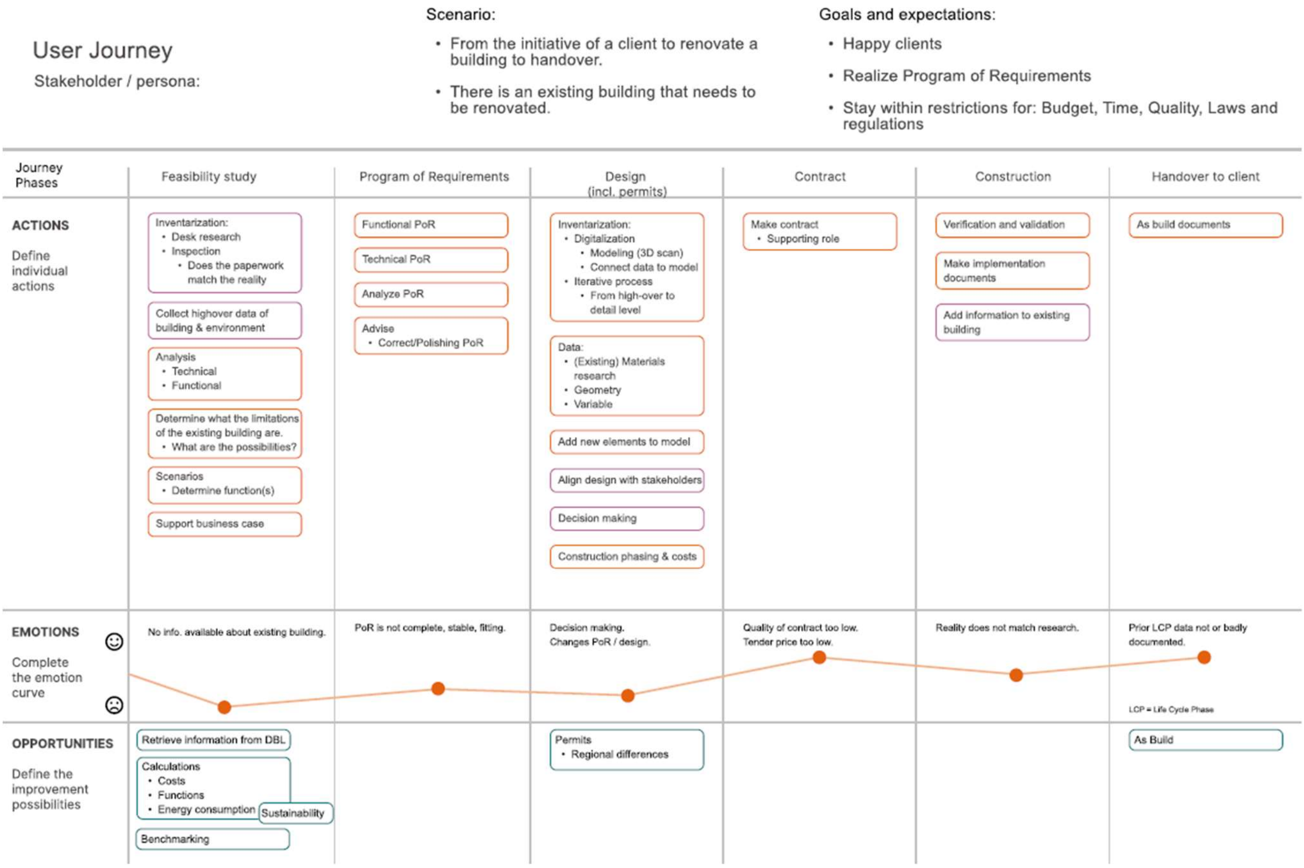
1. **Define the user persona:** Identify the different types of users and create a persona for each one to better understand their needs and behaviour.
2. **Map out the touchpoints**: Map out the different touchpoints or interactions that the customer has with the DBL.
3. **Identify the pain points:** Identify the pain points or areas where the customer may experience frustration or dissatisfaction.
4. **Define the goals**: Define the goals or objectives that the persona is trying to achieve through their interaction with the DBL.
5. **Analyse the customer/user journey:** Analyse the customer journey to identify any potential problems or issues and determine how they can be addressed.
6. **Review and revise:** Review the customer journey with stakeholders and revise as necessary to ensure that it accurately reflects the customer's needs and expectations.

Overall, the purpose of a customer journey is to provide a clear, concise visualisation of the users' experiences and identify areas where improvements can be made to better meet their needs and expectations.

An example of a customer journey map for the design manager for a renovation project can be found below to illustrate what makes up a customer journey map.

[5] Lewrick, M., Link, P., & Leifer, L. (2020). The design thinking toolbox: A guide to mastering the most popular and valuable innovation methods. John Wiley & Sons.
[6] Gudjonsdottir, R., & Lindquist, S. (2008). Personas and scenarios: design tool or a communication device? COOP, 165–176. https://dl.eusset.eu/bitstream/20.500.12015/2795/1/00474.pdf.

Figure 7 - Customer Journey Map of Design Manager in the Renovation process



2.1.3 Improving overall user experience

User groups, customer journeys, and use cases are all related to the process of designing and developing the DBL service that meets the needs of the user or customer. Customer journey mapping provides a holistic view of the customer experience, highlighting all the touchpoints and emotions along the way, while use cases focus on specific interactions between a user and the system to achieve particular tasks or goals.

Customer journey mapping and use cases are related in the sense that they both aim to improve the overall customer experience, but they do so from different angles:

- **User-centred design**: customer journey mapping provides a broader context for understanding the customer's overall experience, while use cases provide detailed insights into specific user-system interactions.
- **Improving user experience**: customer journey mapping is focused on identifying pain points and opportunities for improvement at a macro level, while use cases offer more detailed information to refine the functionality and user interface design for the specific interactions between them.
- **Identifying touchpoints**: customer journey mapping can help inform the creation of use cases by highlighting the various touchpoints where customers interact with the product or system. Use cases can then be developed to ensure that each touchpoint is effectively addressed and optimised.

All three of these concepts are related in that they help to ensure that the DBL meets the needs of the user or customer. User groups and customer journeys help to identify those needs, while use cases help to define the functional requirements of the system to meet those needs. By considering all three of these concepts together, designers and developers can create a product or service that is user-centred and meets the needs of the customer.

2.2 Determining the data needs of the DBL

In order to determine the data needs of a DBL use case one first needs to gain a comprehensive understanding of the use case. The previous paragraph described how to gain such understanding by following steps using the use case description and customer journey mapping.

The following steps outline a structured approach to identifying and defining the data requirements and needs of the DBL to realise a specific use case:

1. Understanding the DBL Use Case
2. Identify the data requirements based on the primary process of the use case
3. Identify aspects
4. Define properties
5. Specify data requirements
6. Prioritise data requirements
7. Document the aspect model
8. Identify responsibilities and data sources

Step 1. Begin by gaining an **understanding of the use case**. As discussed above, this includes defining the user persona, mapping out the touchpoints, identifying the pain points in the primary process of the user, defining the goals of the user, analysing the user journey, and reviewing and revising with the stakeholders. Doing so will give an insight into the main opportunities to alleviate the pain points in the primary process and define a use case that provides value to the user.

Step 2. The next step is to **identify the current data needs and requirements** of the primary process of the user to successfully deliver on the primary process. As an example, we can take the Technical Due Diligence process that is frequently used in high-value asset transactions. This process requires a lot of (technical) asset information to gain insight into the condition of the building and how it has been managed throughout the life cycle of the building. A pain point for the user is in this example the difficulty to get certain types of building-related information. The use case should in that case focus on solving this issue for specifically these types of building-related information. This and other data needs should be organised in a template that is called an aspect model (see the example below, and Section 3.3 on semantic operability for more detail on a proposed classification of the aspects).

Step 3. To **identify the aspects (data dimensions)** that are relevant to the use case, one should start by dissecting the primary process of the user and plotting it in an aspect model (see further below for an example). Based on the various process steps of the user, the data requirements throughout the process and the identified pain points, one can identify aspects that represent different dimensions or categories of information.

Step 4. For each identified aspect, **break it down into further properties**. Properties provide more granular detail within each aspect. For instance, within the energy consumption aspect, properties might include electricity usage, heating and cooling consumption, or renewable energy sources. Similarly, within the maintenance records aspect, properties could be equipment maintenance, repair history, or service contracts.

Step 5. For each sub-aspect, **determine the specific data elements** or attributes that are required to capture the necessary information. This involves identifying the data types, formats, and any additional metadata or contextual information needed. For example, within the electricity usage sub-aspect, data requirements may include total consumption, date and time stamps, and the cost of electricity that is provided.

Step 6. Evaluate and **prioritise the identified data requirements** based on their importance and relevance to the DBL use case together with stakeholders. Consider the impact on decision-making, compliance requirements, operational efficiency, or stakeholder needs. This prioritisation helps ensure that critical data needs are addressed first.

Step 7. Create and **document a comprehensive aspect model** that captures all the identified aspects, properties, and associated data requirements. This model serves as a reference document for the DBL development and provides a clear structure for data collection, storage, and retrieval. The data model should outline the relationships between aspects, properties, and any dependencies between data elements.

Step 8. The final step is to **identify and document the stakeholders** responsible for providing and documenting the data, the source from which the data stems and the stakeholders responsible for verifying the data. These responsibilities can be defined on the level of the aspect, or for each of the properties depending on specific aspects and properties. The verification process will use the data requirements for verification. Some of the data may already be available at the national level. Annex V provides an overview of relevant national datasets that have been identified in this study.

Example of an aspect model

In Figure 8 below you find an example of an aspect model for the construction sector that is focused on the primary process of building renovation process. This is an example to showcase the function of the aspect model in relation to the defined persona and customer journey.

In the customer journey of the design process for a building renovation, the various process steps have been defined: proposal stage, feasibility study, program definition, design permit, contract, construction, handover with client and exploitation. The various aspects under consideration were: building identification, general, legal & finance, dimensions, performance, structure & material and building services. For each of the aspects, various properties have been defined.

The process steps each have different data needs, which can be populated in the overview below, where a green cell identifies a data need and a blue cell identifies a possibility to share the information and data that is produced during the steps of the user's primary process.

For each of the data needs (in green) various data are required to complete the step of the primary process. This can then be defined in additional rows below the aspect row.

Figure 8: Example of an aspect model for the building renovation use case

Construction - Renovation	Role: Design Manager							
Retrieve data from DBL Provide data to datasets _ DBL	Proposal phase	Feasibility study	Program definition	Design	Permit	Contract	Constru c-tion	Handover
Aspect: Building identification								
Building ID; building unit ID; cadastral parcel ID; online link ID like InspireID								
Aspect: General								
properties: Relationships between key DBL entities								
Data Requirement: Building, BuildingUnit, CadastralParcel, etc.								
properties: Indication of types or subtypes Data Requirement: (via a simple attributes not full taxonomy)								
properties: Address Data Requirement: placement indicators (like geo coordinates)								
categories) ***Data Requirement***: at building level and building unit level; user profile like students, seniors, asylum seekers, etc.								
properties: Dates of construction, permits, renovation etc.								
properties: Documents Data Requirement: Urban licenses, renovation proposals, renovation passports etc.								
properties: Documents Data Requirement: BIM model, technical drawings; etc.								
Aspect: Legal & Finance								
properties: tax valuation; lifecycle cost; annual maintenance cost; rental value and maximum rental value; sales transaction value; etc.								
Dates of valuations								
Documents: sales deed, tenancy agreement, insurance policy, clean soil statement, rule violations								
Aspect: Dimensions								
Data Requirement: Lengths, gross and net areas & volumes								
linked geometric representations (0D, 1D, 2D, 3D)								
Aspect: Performance								
properties: Functionality offered, connection to utility services, indoor health & comfort levels								
Data Requirement: Energy performance label, circularity label, energy and water use label, CO2 and N2 emissions label, smart readiness indicator etc.								
Actual energy and water consumption and production levels								
Document: Energy Performance Certificate (EPC)								
Aspect: Structure & Material								
properties: Number of / breakdown in zones, floors, spaces/rooms, elements, components, products, materials								
Data requirement: U-values for various element types								
Year of latest materials inspection, asbestos check etc.								
Certain materials as asbestos for authorised DBL users only								
Aspect: Building Services								
properties: Energy (gas/electricity/solar/thermal/city heating, …), production and consumption installations								
Ventilation system								
Water and sewerage installations								
Building automation parameters								
Security for authorised DBL users only								
Telecommunications connectivity								
Data Requirement: Year of latest installation or repair								
Year of latest fire security check, elevator and balcony safety check, swimming pool salmonella check etc.								
Documents: elevator and fire safety inspection certificates, maintenance contract, utilities contracts								

2.3 Quantifying benefits

For a decision to develop a DBL, it can help sometimes **to demonstrate that the benefits are larger than the costs**. The financial costs of developing a DBL are discussed in Section 4.2 further below. If it is accepted that the qualitative benefits discussed in Section 1.3 justify the costs of certain millions of euros to invest in the development of a DBL, there is no need to quantify the benefits.

However, if stakeholders doubt that the benefits are larger than the costs or demand financial compensation for sharing data a so-called **Standard Cost Model** helps to quantify benefits. A Standard Cost Model is essentially a method to define units and unit costs. Multiplying the unit costs with the number of units then gives the total costs. The same approach can be applied to benefits.

The following steps can be used to quantify benefits:

1. Identify the relevant units
2. Estimate unit benefits (or avoided costs)
3. Estimate the number of units (typically: per year)
4. Multiply the number of units by the unit benefit
5. Add future annual benefits while applying a discount factor

Step 1. **The unit is the entity or the event that generates a benefit**. For example, an entity could be a municipality that uses the DBL to develop a renovation strategy. An event could be an application for a building permit, people or a company moving to another building, a safety inspection, etc. Benefits are often costs that are avoided, in the context of a DBL by accessing data from the DBL. The use case analysis discussed above will help to identify the relevant units that generate benefits (or avoid costs).

Step 2. **To estimate unit benefits**, it helps to identify the activities and out-of-pocket costs that are avoided by accessing data from the DBL. Again, the use case analysis discussed above helps identify the relevant activities. For example, in a building permit procedure, the main activities include the submission of data by the application, the check of those data and discussions about missing or incorrect data. By estimating the duration of the avoided activity in hours, the number of people involved in that activity and multiplying these with an hourly pay rate, one can calculate the benefits of one entity or event. It generally suffices to collect data for this estimate for a "typical" user. For example, an average-sized municipality, an average-sized construction company, etc.

The duration of the activity and any out-of-pocket costs that can be avoided by using a DBL, for example, a fee to obtain design documents from the architect or construction to get information that a DBL aims to include, will vary per event. For instance, the duration of a building permit procedure may depend on the type of building permit procedure and the number of buildings involved. A practical solution is to use historical averages, under the assumption that future costs per application will follow past trends. Also, it may suffice to collect data on historical costs for a dominant type of building permit procedure. If the avoided costs for the dominant building permit procedure alone already exceed the costs of developing and maintaining the DBL, then the benefits of avoiding costs across all types of building permit procedures will exceed the costs even more.

Step 3. Some events take place many times in one year, such as an application for a building permit, people moving to another building and safety inspections. Other events take place only once every few years, such as a major renovation. For events that take place many times in one year, one can use historical numbers, assuming that future numbers will be similar.

To **estimate the frequency of events** that take place once every few years, one may use plans, expert opinions or even historical data. However, an event that will not take place within the time horizon of the DBL (see further step 5 below), should not be taken into account. If an event occurs randomly every few years, such as an earthquake or a flood, that necessitates a repair for which building data need to be looked up, one can estimate an annual probability. Using historical data or design data is therefore advised – in case buildings or dykes are designed to withstand an earthquake or a flood with a magnitude that occurs every x-number of years.

Step 4. For each type of entity and event, the **multiplication of the (annual) number of units and the unit benefits** or unit costs avoided gives the (annual) total benefit per type of entity or event. Building data may reduce costs for different types of entities and events. For example, data on asbestos in a building is useful every time people consider moving to that building, to reduce the number of asbestos inspections and in the end for the removal of asbestos.

Step 5. Annually occurring benefits can be added up, to estimate the "present value" of all future benefits. However, it is **customary practice to limit the number of years taken into account** (the so-called time horizon) and to give smaller weights to benefits further in the future (called discounting). In a DBL context, the time horizon could be 10 or 20 years but should not be longer. The reason is that technical developments in the digital area render current technologies obsolete in the intermediate future. For example, at the start of this millennium mobile phone apps did not exist, Internet Explorer was the dominant Internet search engine, and while programming languages have a longer lifetime, even they do not last forever. This also means that the benefits of an event that is certain not to take place in the next 10 or 20 years, should not be included.

The weighting of a benefit 'B[t]' that occurs 't'-years in the future, is done by multiplying it with a so-called discount factor which depends on an appropriate discount rate. The discount rate reflects a preference for current benefits over future benefits because future benefits are more uncertain. For example, fewer people may use the DBL than expected, there may be delays in the development of the DBL, priorities may shift from building data to health data, and so on. For companies, the future is often more uncertain than for governments due to competition, for instance. For companies, an appropriate discount rate may vary between 6 and 12% per year depending on the market environment. For governments, a discount rate of 2 or 4% is typically used. In the formula, the discount factor for a benefit that is '[t]' years in the future, is '1/(1+r)^t'.

If the future annual benefits are assumed not to change over time, the following table shows the value with which to multiply the annual benefits to obtain the total present value of benefits. Although a discount rate of 0% should not be used, we present the results to show how the discount rate relates to simply adding up future annual values.

Table 2: Discount rates for different time horizons

Discount rate \ Time horizon	10 years	15 years	20 years
0% per year	10	15	20
2% per year	9.2	13.1	16.7
4% per year	8.4	11.6	14.1

Hypothetical example for quantifying benefits

As an entirely hypothetical illustration of the type of assumptions needed in practice, we consider demolition permits. In France, 17,000 demolition permits are issued per year.[7] Suppose that the use case has identified the following pain points without a DBL:

- Obtaining copies of the building drawings (construction company)
- Looking up paper files and data in paper files (authority)
- Discussion about missing or incorrect data (both construction company and authority)

For obtaining copies of the building drawings an estimate is needed on how often these are lost. Let us assume this is the case for one in ten demolition permit procedures. Let us further assume that architects charge EUR 1,000 for providing a copy of their drawings. Let us assume that a DBL initiates

[7] See https://www.statistiques.developpement-durable.gouv.fr/liste-des-permis-de-construire-et-autres-autorisations-durbanisme.

a process of digitising building data, and without the DBL, paper files would need to be looked up. A further assumption could be that that takes 30 minutes per application. Lastly, suppose that discussions about missing or incorrect take 1 hour per application. Additional assumptions are needed about hourly pay rates (including or excluding employer social security contributions) – let us say EUR 30 and EUR 50 per hour for authorities and construction companies, respectively.

Then, the unit costs for authorities that a DBL could save is EUR 30 times 1.5 hours = EUR 45. For construction companies, it is out-of-pocket payments of EUR 1,000 in 10% of the applications (on average EUR 100 per application), plus 1 hour of discussions at EUR 50 per hour, or in total an expected EUR 150 per application. The total unit benefits (of authorities and companies) are then EUR 45 + 150 = 195 per application. For 17,000 demolition procedures per year, this implies an annual benefit of more than EUR 3 million. Discounted over 10 years at 4% per year, this number can be multiplied by 8.4 to obtain an estimated present value of close to EUR 28 million (see Table 3).

Of course, a use case analysis may reveal additional activities that may be avoided with a DBL, and the values used in this hypothetical example will be different for each country – including for France where only the historical number of demolition procedures is a fact that was looked up. Nevertheless, it serves as an illustration of how to quantify benefits and how large benefits can be.

Table 3: Hypothetical example of benefits of a DBL for more efficient demolition procedures, France.

Benefit factor	Total	Authorities	Companies
Need to obtain drawings			10%
Fees charged for drawings			1,000
Hourly pay rate		30	50
Hours to look up paper files		0.5	
Hours to discuss the application		1	1
Unit benefits	195	45	150
The annual number of demolition procedures	17,000		
Annual benefits	3,315,000	765,00	2,550,000
Present value factor *	8.4		
Present value	27,846,000	6,426,000	21,420,000

* Time horizon = 10 years, discount rate = 4% per year

2.4 Identifying stakeholders, raising awareness and acceptance

The functional use of the DBL depends on a combination of people, processes, and data. It is therefore important to not lose sight of the human component when realising a shared database. Moreover, accountability and transparency are key for the success of projects like a national DBL. Therefore, one should involve all stakeholders in the development of a DBL to get the buy-in from the industry and stakeholders involved in all the different life cycle phases of buildings. These actors must see the benefit of the DBL and there needs to be a common understanding on aggregating built environment data and its utilisation for the DBL to succeed.

We already explained the importance of developing use cases for, among others, the purpose of identifying users. During this process, involving these users will already be key to properly understanding their needs and expectations towards a DBL and raising awareness and acceptance about the initiative.

Based on the use case analysis, one should have a general understanding of which stakeholders to involve, however, one should also consider indirectly impacted stakeholders. These might be stakeholders to which the chosen use case does not directly apply, but that nonetheless have a stake in the construction industry and thereby may be affected indirectly by the changes.

In general, one can differentiate between stakeholders that:

- Are **required** as they will need to contribute data to the DBL;
- Are **involved** as they might use the DBL;
- Are **indirectly affected** by any developments in the built environment as a result of other actors' activity.

The first two groups should be involved in activities deciding on the exact scope, requirements and functionalities of the DBL, while the third needs to be informed.

Taking this knowledge, as an implementing authority the first step of this exercise should be to draw up a stakeholder overview and map relevant stakeholders. For this purpose, a simple table in an Excel file can be created that captures all relevant information. Such a mapping is a continuous exercise as during implementation additional stakeholders, not previously thought of, might surface.

Table 4 below provides an example of a hypothetical case for a DBL focusing on an energy performance use case. In Germany This can be extended by additional information, e.g. contact details and be used throughout the project.

Table 4: Hypothetical example of a stakeholder mapping for a use case on energy performance in Germany

User group	Sub-groups	Stakeholder, example for Germany	Role
Governmental agencies	EU level (for alignment with EU legislation)	European Commission, DG ENER (Primary), DG GROW and DG ENV (Secondary)	Involved
	National level	Federal Ministry for Housing, Urban Development and Building, Federal Ministry for Digital and Transport, Federal Ministry for the Environment, Nature Conservation, Nuclear Safety and Consumer Protection Institute for Federal Real Estate, German Energy Agency, Federal Office for Information Security (BSI)	Required

User group	Sub-groups	Stakeholder, example for Germany	Role
	Regional level, if applicable	Corresponding responsible ministries and agencies at the state level, e.g. the Bavarian State Ministry for Housing, Construction and Transport	Involved
	Local level	Representative organisations of local governments, e.g. the German Association of Towns and Municipalities and the Association of German Cities	Required
Construction sector	Architectural and engineering services	Architects' Delegation Germany (includes Bundesarchitektenkammer, Bund Deutscher Architekten, Bund Deutscher Baumeister Architekten und Ingenieure, Vereinigung freischaffender Architekten) German Association of Consulting Engineers	Required
	Construction products	German Building Materials Association	Indirectly affected
	Construction on-site (incl. renovation)	Central Association of the German Building Trade (ZDB), Main Association of the German Construction Industry (HDB), German Confederation of Skilled Crafts (ZDH)	Required
	Building surveyors / evaluators	Federal Association of German Building Surveyors (BBauSV) and Association of Building Surveyors in Germany	Required
	Real estate and Facility managers	Federal Working Group Real Estate Industry Germany (combines various associations such as Bundesverband Freier Immobilien- und Wohnungsunternehmen, Bundesfachverband der Immobilienverwalter, Immobilienverband Deutschland, etc.), Bundesverband Deutscher Wohnungs- und Immobilienunternehmen, Deutscher Verband für Facility Management	Involved
	Standardisation organisation	European Level: CEN/TC 442 National level: DIN Standards Committee Building and Civil Engineering (DIN NABau) Industry: buildingSMART Germany	Involved
Building owners and users	Owners	Central Association of German House, Apartment and Property Owners (Haus & Grund Deutschland)	Required
	Tenants	German Tenant Organisation (Deutscher Mieterbund)	Indirectly affected

User group	Sub-groups	Stakeholder, example for Germany	Role
Financial and legal institutions	N/A	The Association of German Banks (BdB), the German association of savings banks and money transfers (DSGV), KfW, Gesamtverband der Deutschen Versicherungswirtschaft, Bundesnotarkammer	Involved
Utility companies	N/A	German Association of Local Public Utilities of municipally determined infrastructure undertakings and economic enterprises	Indirectly affected

In general, the stakeholders who will be the potential contributors and users of the DBL should already be consulted during the use case analysis. However, once the use case(s) have been analysed and the initial stakeholder mapping is conducted, the DBL initiative should be announced to the wider group. This includes the required and involved stakeholders who should be informed on the scope, purpose, timeline, and possibilities to get involved. An official communication directly targeting the identified stakeholders through their representatives is expected but the information could also be announced publicly on a website so others can learn about the initiative. Preferably, this announcement is combined with a first information session to provide more context and to provide a forum for questions.

Following the announcement, a good way to ensure proper involvement of stakeholders during the development and implementation of the DBL is to set up an advisory or working group which can be regularly consulted. This can be one larger group or separate groups focusing on specific aspects (e.g. data responsibilities, DBL functionalities, etc.). This will allow the development team to regularly present progress and establish feedback loops that can be used for an iterative development process. It also gives stakeholders the ability to voice concerns and make suggestions for improvements.

Box 3: Examples of stakeholder involvement in similar projects

For the development of the Kadaster Knowledge Graph, the team of the Dutch Cadastre involved various stakeholders working in several stages. In the initial stages, collaboration took place with stakeholders from different sectors such as the utility sector, security of the built environment, and spatial planning. Workshops were conducted to understand their specific use cases and questions, which helped shape the knowledge graph. In subsequent stages, it became essential to involve stakeholders in the standardisation process to formalise and enhance the model with proper semantics. This iterative approach ensured that stakeholder input was incorporated throughout the development and led to the evolution and refinement of the Knowledge Graph.

For Estonia’s e-construction platform, industry associations played a crucial role. By collaborating with architects, engineers, general contractors, maintenance professionals, utility owners, and planners, different perspectives and expertise were brought together. Estonia successfully initiated a joint roadmap called the "Long Term View on Construction 2035", which set out a path to increase productivity and improve the principles for developing and improving the built environment. This collaborative document has gained recognition and support from various stakeholders, becoming a significant milestone in the industry. The involvement of diverse stakeholders and the endorsement of the document by the Estonian Parliament have contributed to the success of their approach.

Source: DBL Workshop of 27 June 2023 on “Technical guidelines for Member States”

2.5 Analysing barriers to implementation

When starting to implement a DBL, you should **identify the potential barriers** and consider how to overcome these barriers to achieve a successful implementation.

A starting point can be existing research such as this study (see Chapter 5) or the preceding study[8] which identified the following types of barriers: costs, static nature of DBLs and need for manual updates, privacy, data management, access to information, administrative burden, and a fragmented regional approach.

Contrasting the barriers, the study also lists success factors which include clarity both on the legal framework and the scope for the DBL (see also the use case analysis in Section 2.1). It also calls for alignment with other initiatives and standards (see Chapter 3 and D2.1), proper testing and inclusion of stakeholder experiences (see Sections 2.4 and 2.6), ease of use, regular updates, properly explained process for data validation and clarity on data input responsibilities (see Sections 2.2 and 3.3).

As outlined in the short paragraph above, there are a multitude of potential barriers and success factors to implementing DBLs. It is important to be aware of these and **analyse them in your national or regional context** (in federally organised countries) as there may be differences in the legal framework or other aspects such as additional layers of governance.[9] Therefore, it is recommended to first assess the barriers, success factors and potential solutions internally and then to discuss them in a structured manner with stakeholders (see Section 2.4).

Generally, we can differentiate between the following types of barriers:

- Political barriers;
- Legal barriers;
- Technical barriers; and
- Economic barriers.

Political barriers

While apart from the questions regarding data ownership and privacy, the topic of DBLs and building data is in itself not a political topic that might divide political parties, there are nonetheless political barriers. The first set of barriers refers to barriers related to a lack of political will to implement such an initiative. Large-scale and long-term projects such as a national DBL **require initiative among public authorities**. While the status quo may work adequately, it is not necessarily most efficient and there may be untapped potential for proper building management and policymaking. Such political initiative often requires one institution to take charge to create a long-term vision for future data management in the built environment.

Similarly, the **political capacity might also be dispersed**. Different regional authorities or government agencies, or even different departments within the same agency might work on separate initiatives or have separate and overlapping competencies and responsibilities that all relate to data in the built environment. Therefore, DBL implementation requires a dynamic and forward-thinking public authority that is also characterised by a shared vision to initiate change and combine efforts.

[8] European Commission, Executive Agency for Small and Medium-sized Enterprises, Dourlens-Quaranta, S., Carbonar, G., De Groote, M., et al., Study on the development of a European Union framework for digital building logbooks : final report, Publications Office, 2021, https://data.europa.eu/doi/10.2826/659006.

[9] Barriers can vary greatly between Member States in both the type and their intensity. For example, a country such as Germany will face the complexity of having to deal with different responsibilities across federal layers which complicate standardisation. In addition, in Germany there are widespread concerns regarding privacy and data protection, making Germans particularly reluctant to provide additional data to their governments. In contrast, the high costs combined with long term benefits is a more transversal barrier. Though even here, a more fiscally restrained country (by choice or involuntary) will be more hesitant and favour more short term focused policies.

A starting point could be a roadmap for the built environment that sets some overarching goals and a timeline. This vision should be a common vision that is agreed upon with stakeholders and therefore requires effective communication with the affected parties and ensuring their buy-in.

Finally, the **sheer size and complexity of a DBL project** might discourage policymakers from initiating such a project. To ensure successful implementation it is therefore recommended to start small with basic functionalities that can be extended over time. Starting with all sorts of complex solutions might overwhelm users and could be seen as an administrative burden by stakeholders. It therefore could also lead to implementation barriers down the road. Moreover, government accountability simply requires tangible results in a not-too-distant future. Therefore, it is again advised to start with small goals, to implement a DBL step-by-step and only expand after a prototype DBL works well.

Legal barriers

There is a link between legal and political barriers due to the need for **a legal foundation for a DBL** to allow for data to be collected, processed and published in a DBL.[10] Irrespective of the varying degrees of data security and their advancement, all countries face the same **privacy safety and data security** concerns in relation to DBL and will need to address these in consultations with their respective stakeholders.

Legal barriers also entail the **issue of responsibilities**, for example for the quality and timeliness of data. Finally, also legislation or policy initiatives regarding **open standards** should be considered as DBLs can also support the move towards more open standards in the built environment.

We discuss these legal risks in more detail in the risk inventory in Chapter 5. However, in general, Member States should carefully map existing building codes and legislation related to the built environment, construction, databases and privacy. Based on that mapping a legal assessment needs to be made if any legal changes or clarifications are required (see also Section 2.7).

Technical barriers

Technical barriers relate to challenges related to the technologies involved with implementing a DBL. Chapter 3 on the technical implementation provides guidelines for this purpose. In general, an **agile approach** with several iterations is recommended to deal with unforeseen problems. However, apart from these, there might be technical barriers unique to the situation specific to a Member State.

These could relate to the **availability of data** and the simple issue of most information being on physical carriers that first need to be digitised. It could also relate to the use of **different data formats** and incompatible legacy systems or legacy data that require translation or conversion to be useful, or. There are various examples of rich data sets in Member States, which used different data formats or were collected for different purposes and therefore could not be easily linked with a new platform such as a DBL. **Legacy issues** may also relate to past changes in building classifications[11], energy performance classifications etc., which require additional work to understand and use.

Technical barriers also relate to the **skills and capacities** needed regarding data scientists and similar experts. The level of digitalisation varies greatly across EU authorities and access to experienced software developers can be a bottleneck for many countries. This requires a review of

[10] For example, in Germany public sector owners of buildings have certain legal requirements, as do private owners when it comes to operational liability of buildings, which could serve as a starting point in this case for a legal foundation.

[11] Generally it is recommended that the building type is separated from the building identifier, so that changes to the type of building do not affect the identifier.

the types of skills needed and identifying experts if not already available within the team (see also Section 4.1).

Finally, **data storage** can also be a barrier as it needs not only a technical solution but also political agreement. It is preferred that data stays at the owner and is shared through a common data environment, however, this requires agreements on data formats, authentication systems and trust. To overcome technical barriers, sound data management principles (see Section 3.5) are key and it is recommended to start with a data management plan once data needs are identified to better understand the current and required future situation.

Economic barriers

Finally, economic barriers relate on the one hand to the simple **need to invest resources** into this endeavour. This, therefore, requires the political will to allocate these resources to public budgeting. On the other hand, they relate to the **different economic interests** of actors within the built environment as well as data owners and users.

Regarding the **need for funding**, it is important to highlight the overall benefits of DBLs as explained in Section 1.3 and where needed to secure the required funding. For **realistic budgeting**, it is recommended to start with small goals and work towards a more basic version of the required DBL that already showcases some of its benefits to users, stakeholders and lawmakers (see Section 2.6). It is not advised to underestimate costs to increase the political will and instead, when necessary, quantify the benefits as discussed in Section 2.3. Chapter 4 gives an overview of the types of costs and the rough order of magnitude, although the costs are dependent on the ambition level. Additional **EU resources** can also be identified, for example, the European Commission's Technical Support Instrument does not require co-financing and provides Member States with technical expertise to design and implement reforms[12].

While it is envisaged that public authorities bear the cost of developing the DBL, **private-sector data providers may also incur costs**. For example, engineering companies increasingly develop their own building information models and may need to convert data into the format required by the national DBL (if this requirement is imposed on them by regulation, etc.). Some DBL initiatives in the past have failed to materialise due to **administrative burdens** incurred or perceived by market players.[13] Legal requirements, government support for private investments and minimising the cost for market players may all help to overcome economic barriers for market players who need to contribute to the national DBL. Costs could be minimised for example by linking the requirement to update data to existing processes (e.g. building valuation) or by accepting that the public authority converts formats.

Nevertheless, even if market players bear no administrative burden, **private actors may feel no incentive to share data** from their own databases. This is a main barrier in many DBL developments and it is therefore recommended as much as possible to involve stakeholders early on (see Section 2.4) to address these concerns and propose solutions. For example, private stakeholders may not be willing to share all detailed information in their possession but may be willing to share certain key data to keep the building data up-to-date, to reduce inspection costs, or other benefits that the use case analysis reveals. Additionally, a DBL should facilitate existing business or regulatory processes for companies and allow private actors to develop private use cases and business models around it thereby giving them a clear incentive to support the development.

[12] For more information, see: https://commission.europa.eu/funding-tenders/find-funding/eu-funding-programmes/technical-support-instrument_en.
[13] See e.g., R2M et al., Study on the Development of a European Union Framework for Digital Building Logbooks, December 2020, https://op.europa.eu/en/publication-detail/-/publication/cacf9ee6-06ba-11eb-a511-01aa75ed71a1/language-en

2.6 Developing a working DBL prototype

From the experience of DBLs that have been implemented and general software development, it is usually best to start small and use an iterative approach that allows for regular feedback[14]. It is therefore **recommended to first start with a proof of concept and then develop a working prototype** as this supports the understanding of the use and purpose among stakeholders and colleagues indirectly impacted (e.g. lawyers that might need to assess any legal implications). Therefore, once one has determined the role and purpose of the DBL (Section 2.1) and also comes to a first understanding of the requirements and needs regarding data and stakeholders (Sections 2.2 and 2.4), one can assign a development team to start the development of the DBL.[15]

Prototyping is a common phase in software and platform development, and examples have been shared on the internet and in literature.[16] In developing the prototype it is suggested to use an **agile software development approach**. Agile software development requires less time, allows for adapting to changing requirements, as well as more efficient collaboration with stakeholders and identifying issues at early stages.[17]

In general, we can distinguish three phases, however not all of these might be required for your use case if knowledge can be drawn from existing projects[18]:

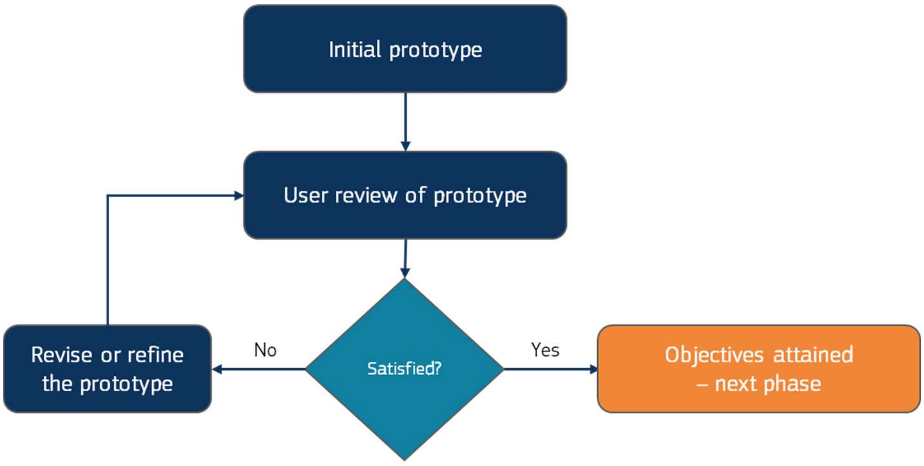
1. Proof of concept to validate the technical capabilities for the foreseen DBL project;
2. Prototype to showcase your concept;
3. Minimum viable product to test the reception by users and stakeholders (see Section 2.8).

The **proof of concept** does not necessarily require any development but should flow out of the exercises done as part of 2.1 and 2.2 to prepare a short text document combined with mock-ups describing and visualising how the DBL would work. A proof of concept is often a small and internal project, but its outcomes can also be a helpful tool to engage with stakeholders about the ideas and foreseen functionalities behind the DBL. It allows us to analyse the possibility of the DBL. This provides information on the feasibility of the idea and helps guide the discussions with stakeholders (Section 2.4). For a proof of concept, one can check initiatives in other countries, keeping in mind that most countries start with making Cadastre data public and then add data to build up a proper DBL, although France started a DBL-like initiative to document energy performance-related data.[19]

Based on initial feedback from stakeholders, colleagues and developers on the proof of concept, one can move to internally developing a first software **prototype**. This first version of a DBL demonstrates or simulates how the final product would look and function. The prototype needs to contain data for only a few buildings. The prototype also demonstrates what data are collected, which helps to concretely weigh the legitimate interests of various stakeholders whose data will be collected, or on whom data will be collected, and thus to justify the data collection, or revise the data to be collected. The prototype DBL should have input and output modules, but does not need to operationalise the online input from various stakeholders, or data quality check automatisms. After internally prototyping the DBL and confirming that the data to be collected can be justified in a legal framework (see also Section 2.7), prototyping may involve the collection of feedback from selected stakeholders on a confidential basis (see Figure 9).

[14] This feedback process could be aligned with a working group of stakeholders set up as part of Section 2.4.
[15] For more information on the required people and skills, see Section 4.1.
[16] See for example: https://github.com/LandRegistry/find-property-information-prototype or https://www.mdpi.com/2078-2489/10/10/310.
[17] For a description of what agile software development entails, see: https://www.agilealliance.org/agile101/.
[18] A proof of concept can help validate the technical capabilities required, which might not be needed if these have been validated through other projects.
[19] Estonia (https://xgis.maaamet.ee/xgis2/page/app/adsavalik), France (https://data.ademe.fr/datasets/dpe-tertiaire/full), the Netherlands (https://vastgoedloep.nl/), Slovenia (https://ipi.eprostor.gov.si/jv/), Spain (https://www1.sedecatastro.gob.es/Cartografia/mapa.aspx?buscar=S)

Figure 9: The prototyping process



Source: Carr, M. (1998). Prototyping and Software Development Approaches.

Throughout the development, rather than using custom solutions, the use of **off-the-shelf solutions** in line with **European standards** and especially **open standards** is recommended.[20] In addition, one should where feasible develop the DBL using **open-source** components. This allows other users to extend the national DBL with their use cases and allows Member States to use and replicate each other's work.[21]

2.7 Preparing the legislative framework and an enforcement strategy

Before further developing a DBL, a legal assessment is needed. As advised in Chapter 5, a DBL should be established in national law. The General Data Protection Regulation (GDPR) in that case requires a legal impact assessment. In this impact assessment, the legality of collecting personal data should be assessed by weighting various legitimate interests, for example, building owners versus people considering buying or renting a building (or building unit such as an apartment or a shop in a shopping mall). The risk of using personal data profiling needs to be assessed, and for which data it is necessary to restrict processing and access to avoid profiling. Likewise, it needs to be assessed what data are protected in the meaning of Article 3(1) of the EU Data Governance Act (also further discussed in Chapter 5).

The preparation of the legislative framework and an enforcement strategy consists of several steps:

1. Assessing the legality of each data item.
2. Assessing compatibility of data sharing with national law.
3. Regulating who needs to provide data while minimising the administrative burden.
4. Regulating the responsibility for the quality and timeliness of data.
5. The legal analysis, concerning the amendment of existing legal provisions and the formulation of new legal provisions.

Step 1. To prepare the legislative framework, it is necessary to **know exactly which data will be collected and published**. For example, to assess compliance with the GDPR it is not sufficient

[20] This includes standards developed by CEN TC/442, buildingSMART, ISO/TC 205, ETSI (e.g. SAREF) and other organisations. For a discussion on relevant standards see the separate deliverable 'D2.1 Semantic Data Model'.
[21] For example, Estonia's e-construction platform has apart from two components relied fully on open-source components, which allows other Member States to reuse their work. The components that are not using open source relate to BIM-based components, where a proprietary solution was necessary.

that “contact data” will be collected because email addresses and mobile phone numbers carry a greater risk of profiling than fixed-line phone numbers and street address data. Likewise, it may not be sufficient to grant the police or universities access to energy use data (if they are to be included in the DBL), but only to authorised persons within these institutions for the specific purpose of an investigation or research, and universities perhaps only in anonymised form.

To give another example, to assess compliance with the EU Data Governance Act, it is not sufficient to that “data on balconies” will be collected, because the number of balconies is likely not commercially confidential, but the design, structure and materials used for the balcony might be (and certainly are if they are patented).

In addition, if documents are to be included in the DBL, these documents need to be screened for personal data and otherwise protected data, to justify whether these documents can be included and to remove or blacken data that should not be public.

Step 2. The **legality of data sharing needs to be checked with national law**, which is not discussed in Chapter 5. These include for example regulations on professional confidentiality and incompatibility of professional activities. For example, notaries may be prohibited from sharing information except for specific purposes, and the DBL purpose may need to be added to the national legislation on notaries. Or architects may not engage in real estate activities and thus the collection of building data from different market players in different phases of the building life cycle may need to be specified.

Step 3. Once it is determined what data is not personal data or otherwise protected it is possible to **regulate who needs to provide these data unconditionally**. For personal data and data that is protected for other reasons (e.g. commercial interests), conditions for sharing data need to be considered, as discussed in Chapter 5. To facilitate the data collection, it is recommended to assess who needs to provide which data to whom according to existing law.[22] If other pieces of legislation already require that certain data must be provided, this reduces the administrative burden of providing data for the DBL.

To reduce the administrative burden, it is furthermore advised to limit any requirement to provide data as much as possible to so-called “moments of change”. Examples of such moments include the sale of a building, a movement to a different building, a renovation, an inspection, etc.

One needs to make a principled choice between making the owner of a building or building professionals responsible for providing data. If one starts from private building logbooks, it makes sense to make the owner responsible for sending updates to the national DBL for example once per year. For example, building professionals would in this case update the private building logbook at “moments of change”. In practice, most national DBL-like initiatives start by making cadastral building data public, and then it may make more sense to require building professionals to directly provide data to the DBL.

Step 4. Concerning the **quality and timeliness of data**, it makes sense to set norms for these and to make the data provider responsible for ensuring that the data provision meets these norms. Of course, if the building owner is made responsible for updating the DBL and building professionals have entered incorrect data in the private building logbook, the building owner may hold in turn the building professional responsible. The responsibility to provide correct data does not absolve the DBL authority from checking the quality of data if only to enforce the data quality norms.

Step 5. Following the previous steps, one should conduct a **regulatory study** of the proposed DBL and the existing laws to assess which specific regulations require updating. For example, for the Estonian e-construction platform, a legal analysis was conducted which aimed to identify legal barriers and analysed limitations regarding personal information and whether one can distinguish

[22] Examples of relevant pieces of legislation could include the law on national cadastres, the national building code, energy efficiency in building regulations, public procurement rules for construction works, building permit regulations, and relevant safety regulations for example concerning earthquakes, noise pollution norms etc., property law and laws regarding land use, territorial planning regulations, real estate (taxation) law and professional regulations such as on notaries.

technical from personal information.[23] In the Estonian case, the study found that none of the data that was proposed to be included in the DBL was privacy-sensitive information.

Beyond this, the Estonian legal analysis covered:

- A review of the need for changes in the construction legislation in the context of the proposed platform;
- An assessment of the administrative and judicial practice of the current Building Code and the regulations issued under it; and
- A proposal of amendments, with wording and justification, to the Building Code, construction works and the obligation to submit a building permit or a building permit.

Based on such an assessment, one can prepare the national legislation to align with the foreseen requirements of the DBL and adjust legislation to avoid any legal barriers.

Finally, to facilitate implementation it is recommended:

1. to have a prototype ready so it is not just theoretical and easier to justify legislation; and
2. to combine where possible the implementation with another regulatory action related to the built environment and the use case (e.g. the digitalisation of the building permit process or new reporting requirements for EPCs).

Box 4: Using existing building processes and regulatory actions to ease DBL implementation

As outlined earlier, to reduce the regulatory burden and facilitate DBL implementation one should consider how the provision of data for the DBL can be integrated into existing building processes and upcoming regulatory actions. For example, in 2023 the new Environment and Planning Act is coming into force in the Netherlands. This Act will combine several existing laws on spatial planning and will digitalise processes such as building permits. It will also dictate which building data needs to be made available upon completion of a building providing clear guidelines and rules to construction companies, owners and the government. Similarly, an Italian draft law decree requires owners who obtain their building permit to provide at completion a DBL. For updating legacy data on existing buildings, a solution is often to make the receipt of subsidies (e.g. for insulation work) conditional on the provision of certain information about these renovations. In general, our feedback survey showed, that moments of change for buildings can be a great tool to ensure building data is up to date. These can include a change in ownership, major renovations, installation of new energy or heating systems, inspections, technical assessments and energy performance evaluations.

2.8 Starting to implement and realising benefits and feedback loops

Once the preceding steps have been completed and in particular a prototype has been developed (Section 2.6), the legislative framework prepared (at least the first four steps, as discussed in Section 2.7), one can move to the actual DBL implementation. The implementation includes the following steps:

1. Turning your prototype into a minimum viable product;
2. Development and implementation of a communication plan;
3. Agreement on internal roles and responsibilities;

[23] More information regarding the study is available here: https://eehitus.ee/timeline-post/ehitise-elukaare-oigusruumi-digitaliseerimiseks-kohandamine-uuring/ (in Estonian).

4. Limited rollout with selected municipalities or regions;
5. Provision of training and providing support to users;
6. Wider rollout and communication of the development roadmap.

Step 1. The prototype should be further developed into what is called a **minimum viable product**. This is a DBL version with the minimum features necessary to solve the identified problems during the use case analysis. The minimum viable product should facilitate the data provision for all types of stakeholders that need to provide data, and for all types of end-users to access the data to which they are authorised. It also should include automatic checks on missing data and checks against data norms. It is advised to use off-the-shelf tools (see Section 4.3 for more detail) to ensure technical interoperability from the start.

Step 2. In preparation for the rollout, a **communication plan and communication materials** should be developed to inform stakeholders about the rollout, data norms and the responsibilities of different actors such as authorities and building owners. For example, a launch webinar could be scheduled with the identified and targeted stakeholders (under 2.3). This event could be opened to other interested stakeholders as well as recorded and more widely communicated.

Step 3. In parallel, internally clear **roles and responsibilities** need to be assigned. For example, a lead data manager needs to be assigned to monitor compliance of building data with data norms through automatic and manual checks and communicate about data norms and errors in data. Their responsibility will also be to ensure overall compliance with the DBL framework presented in Chapter 3 to ensure future EU-wide interoperability. Similarly, software developers are needed to further develop the DBL and ensure its maintenance. For more information on the team, see Section 4.1.

Step 4. Following the first three steps, one can start with a **limited rollout with a few selected municipalities or regions**. A minimum viable product version of the DBL can already be used in practice and thereby tested by early adopters. Such frontrunners can help not only to test the DBL, but also showcase its benefits and thereby help with the future uptake. A limited rollout also facilitates especially in larger Member States the monitoring and follow-up with any issues encountered by users or authorities. From interviews, it has become clear that comments from stakeholders in the initial months of the roll-out can be overwhelming. To increase the acceptance of the DBL, comments should be addressed timely and adequate. One should not hesitate to pause the rollout if critical issues need to be resolved. And an adequately large development team should be in place.[24] The wider rollout can be announced at a later point in time or based on voluntary participation and grow naturally.

Depending on the requirements towards data providers, i.e. if new data standards are to be applied, **training, or at least training material** should be provided in parallel to the rollout. This helps to inform on how to properly use the DBL and provide the data in the right format. Implementing bodies at the local or regional level should be targeted through training sessions as they can then further communicate any relevant knowledge to local stakeholders.

Step 5. Finally, a **development roadmap** should be published and communicated. This plan outlines how the DBL will be extended in different phases and what other features can be expected in future versions. It can also indicate the timing of the wider rollout. This plan should provide a step-by-step approach to the development team on what aspects to work on and splits the larger ambition of a DBL covering all possible aspects of building management into manageable parts. It also helps to align with simultaneous national or EU-level projects such as EPC labels, circularity and smart readiness indicators.

For example, in Estonia, they started with a building registry, which has been slowly extended to their e-construction platform with additional public and private use cases being developed and aligned

[24] A development team peaking at 40 persons has been reported in one interview (see also Section 4.1).

with it[25]. A similar approach is followed by the Cadastre Knowledge Graph of the Dutch Kadaster, which integrates cadastral data with other data to provide new market solutions such as a privately owned real estate platform.

This development plan should also include regular internal and external review points to allow for feedback loops and identify ways to further improve the DBL.

[25] The Estonian e-construction platform aims to enable a lossless exchange of standardised and trustworthy data throughout the building lifecycle. It encompasses various services, including the development of utility network databases, implementation of BIM-based building permits, and the creation of a 3D digital twin for visual representation of built environment data. The platform seeks to connect both public and private sector services and data, fostering collaboration and efficiency.

3. Technical implementation

For the technical implementation, these guidelines and the DBL Semantic Data Model, provide a framework that one should follow when implementing a digital building logbook. The framework provides a core model, which should be replicated by Member States as only an EU-wide framework can support interoperability and a common European data space for construction. However, this framework is also flexible and can be extended with additional elements depending on national or regional requirements.

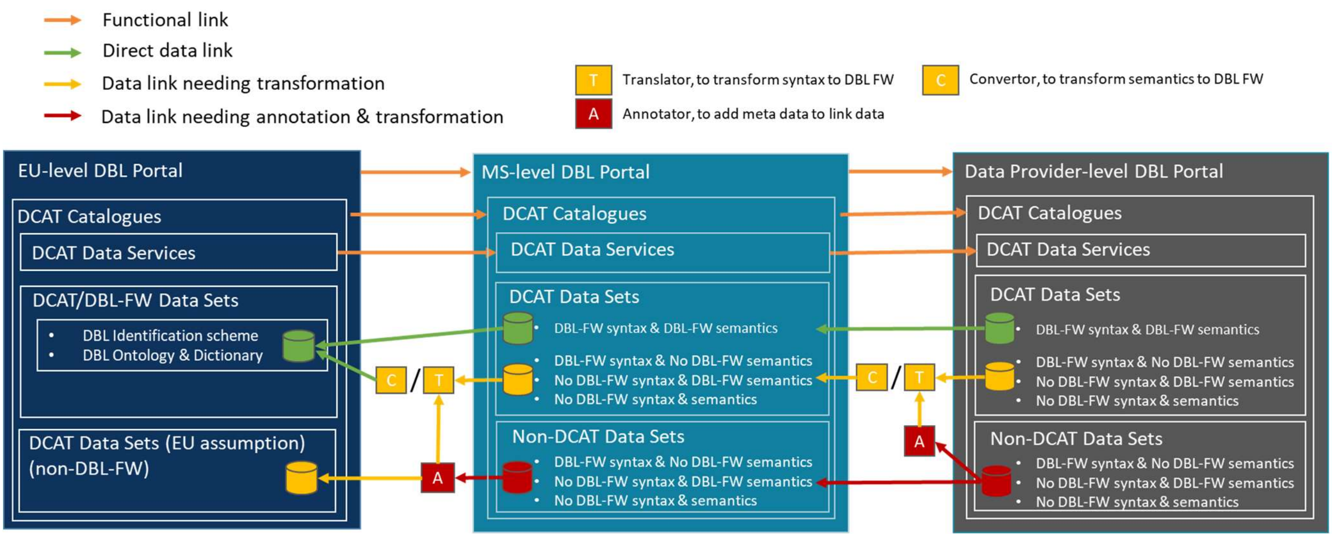
The following sections outline the main concepts of the framework, how to apply it to your context and what its limitations are. More information in graphical form is provided in Annex I. In addition, a more detailed description of the DBL Semantic Data Model is provided in the separate deliverable D2.1 and the linked data implementation in D2.2 (see Annex II for more details).

3.1 Introducing the EU Digital Building Logbook framework

The proposed DBL framework distinguishes between two levels: a European portal as a centralised point for access to the DBL, and national portals for every Member State to implement its own, possibly enriched national DBL. As a third level, there are also the actual data providers, which could be the owners or someone providing the data on their behalf and as such constitute the DBL owner level. Therefore, three levels can be distinguished:

- **DBL EU Level:** general access point for the EU building stock with the three branches of the cadastral parcel, building and building unit.
- **DBL Member State Level:** Using the EU DBL framework to ensure interoperability with the EU level, but with a possibility to add more branches so that it can serve legal or technical due diligence requirements with national legislation.
- **DBL Provider Level:** The level of the data provider, possibly supported by a personal DBL, which can set, for example, maintenance plans and provide information about one's building.

Figure 10: The levels of interaction between the EU DBL portal, Member State portals and data providers



These guidelines represent the reference framework and suggest the instructions to follow in the implementation of a national DBL for the Member State level.

3.1.1 Key principles for implementing a DBL

When implementing a national DBL, five key principles should be taken into consideration. These are simplicity, FAIRness, data quality, Levels of Information Need, and keeping data at its source (see Figure 11).

Figure 11: The five principles for DBLs

Simplicity means to start simple before adding complexity. There are two trade-offs to consider:

- **Usability vs. comprehensiveness:** A simple tool is easier to use and more accessible and will therefore be taken up more quickly. This however means initially leaving out details and foregoing a richer tool with more functionalities and possibilities.
- **Flexibility vs. harmonisation:** Similarly, the more building aspects a DBL covers and the more comprehensive it is, the more it can harmonise data in the built environment. However, this also reduces the flexibility of market actors and requires them to adapt, which again highlights the need to use a step-by-step approach to not force too much adaption at once.

Fairness refers to the **FAIR principles** which aim to optimise the use and reuse of data by making data concise and concrete. This is achieved with data that is: 1) **findable** via a data identification scheme and metadata; 2) **accessible** through user identification, authentication and authorisation; 3) **interoperable** by using agreed data formats or data access mechanisms; and 4) **reusable** via data specifications defining the intended meaning of data. Interoperability and reusability are provided by our proposed DBL framework (see Sections 3.2 and 3.3), while findability and accessibility depend on where and how the data is stored (e.g. in the cloud).

On top of requiring concise and concrete data (implied by the data formats) and clear (implied by reusability), one relies also on data being correct, complete and timely. So beyond defining the data well, **data quality** is highly important. It is therefore important to ensure high-quality data in terms of its relevance (usefulness and timeliness), completeness, consistency, precision, simplicity, traceability, scalability and adaptability. To ensure this, a data management plan and proper data governance agreements are needed (see Section 3.5).

Another important guiding principle is the EN 17412 **Level of Information Need (LOIN)**, which refers to the three levels of data:

1. Alphanumerical data sets, also known as semantic data;
2. Geometric data sets, describing explicit location/shape representations;
3. Documents, with unstructured information content.

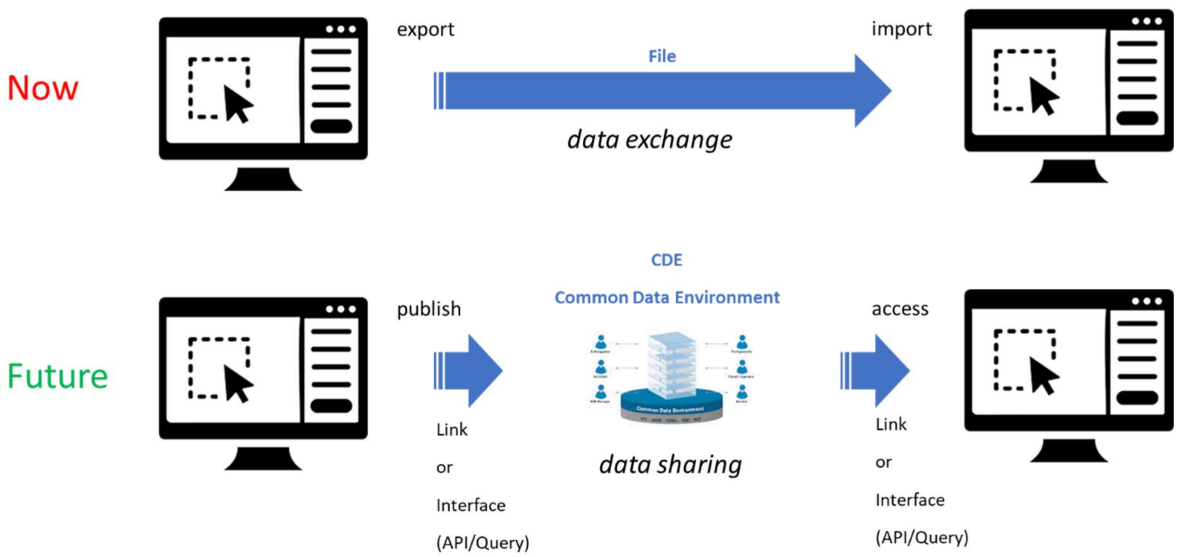
All three levels should be addressed by a DBL and can be treated at the 'meta-level' described by meta-data (data about data sets) or one can go into their content assuming some predefined data structure. In general, the trend is that more and more documents and geometries are replaced with machine-processible and interpretable data.

Finally, the fifth guiding principle is to **keep data at its source** to avoid multiple copies. This requires a distributed solution where data is not copied but made available and it relates to the difference between data sharing and data exchange, which is explained further in the next section.

3.1.2 The modes of interaction of DBLs

The generic use case refers to the fact that our DBL agreements are used for any data exchange and/or data sharing situation between a data provider/owner and data users (Figure 12). In general, data sharing is preferable unless there is a change of ownership of the data. In all other situations, data sharing helps to keep a single source of truth where data is consistently governed, updated and extended. Clarity about the source in addition creates clarity about the responsibility for the quality of the source data.

Figure 12: Forms of data communication



Depending on the more specific use case, a national DBL does not just link to source data, but also validates it (e.g. with a quality flag, and a message about data issues to the source), and combines or integrates data from different sources all linked together and where possible enriched via data derivation and aggregation. The DBL is also a portal that where necessary manages authorisation procedures to access some data and facilitates payment for some data.

3.2 Types of interoperability

The digital transformation of public administrations is one of the top priorities of the EU and the Member States. Likewise, for economic sectors such as construction and the built environment digitalisation and access to data are becoming increasingly important. This is also reflected in the development of common European data spaces that follow design principles that include a common technical infrastructure, building blocks as well as interconnections and interoperability.[26]

The key to such a successful future-proof digitalisation on a national, European (or even beyond on a worldwide) scale is "Interoperability". The proposed **Interoperable Europe Act**[27] highlights this intending to strengthen cross-border interoperability and cooperation between public authorities.

[26] Commission Staff Working Document on Common European Data Spaces, see also: https://joinup.ec.europa.eu/collection/semic-support-centre/data-spaces.
[27] The Interoperable Europe Act proposes an EU cooperation framework on interoperability for national administrations on data exchanges and IT solutions, see: Interoperable Europe Act Proposal.

Furthermore, Article 12(i) of the **Data Governance Act** (DGA)[28] requires measures to ensure interoperability for "data intermediary services". In a true "logbook" setting wherein building owners upload updates on their building, notaries document transactions, or in general collect part of their data from third parties, DBLs should be considered to be a data intermediary service. Therefore, EU law to ensure interoperability applies. The DGA does not define interoperability, but amongst others, it requires the use of open standards in the sector in which the data intermediary service operates.

When ensuring interoperability, it is recommended to take a broad view. A DBL should not only be compatible with widely used open standards but also with legal requirements and the policies of data providers. Therefore, we use the **definition of interoperability** as the ability of systems to exchange and make use of information. The purpose of interoperability is further defined in the 'New **European Interoperability Framework** (EIF)'[29] as:

> *"The ability of organisations to interact towards mutually beneficial goals, involving the sharing of information and knowledge between these organisations, through the business processes they support, by means of the exchange of data between their ICT systems."*

The EIF defines four layers that build upon each other (see Figure 13). These four layers help us understand what interoperability is further as it highlights that it not only relates to cooperation and agreements around technical issues but also organisational, legal and semantic aspects that need to be considered.

Figure 13: EIF forms of interoperability

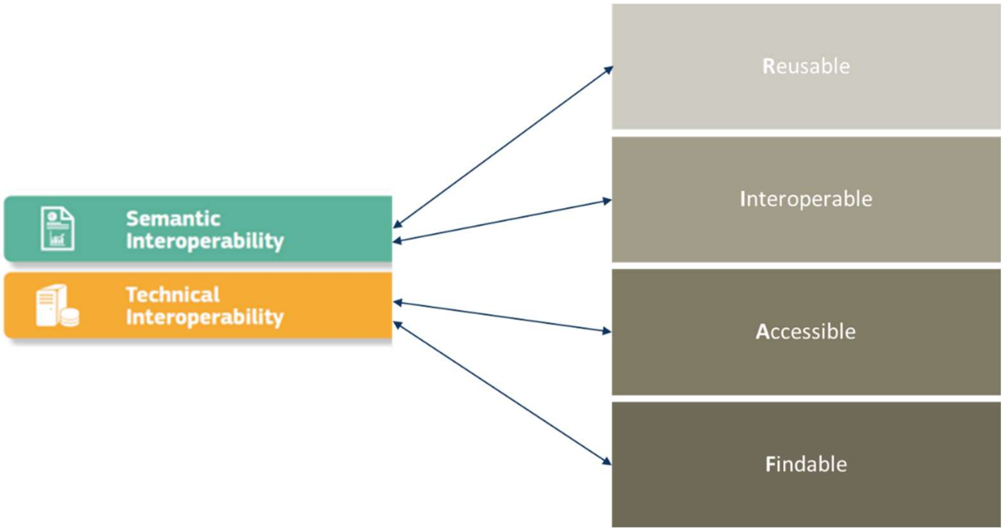
The EIF has furthermore a strong connection with the **FAIR principle**, which is also recommended as one of five guiding principles in this handbook (see Section 3.1). Specifically, the technical interoperability layer of the EIF refers to the two bottom layers of FAIR, i.e. 'Findability' and

[28] See https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52020PC0767.
[29] *"The EIF gives guidance, through a set of recommendations, to public administrations on how to improve governance of their interoperability activities, establish cross-organisational relationships, streamline processes supporting end-to-end digital services, and ensure that existing and new legislation do not compromise interoperability efforts."*, New European Interoperability Framework, Promoting seamless services and data flows for European public administrations, European Union, 2017. https://op.europa.eu/en/publication-detail/-/publication/bca40dde-deee-11e7-9749-01aa75ed71a1/language-en.

'Accessibility'. In turn, the EIF semantic interoperability refers to the two top layers of FAIR, namely the 'Interoperability' and the 'Reusability' (Figure 14).

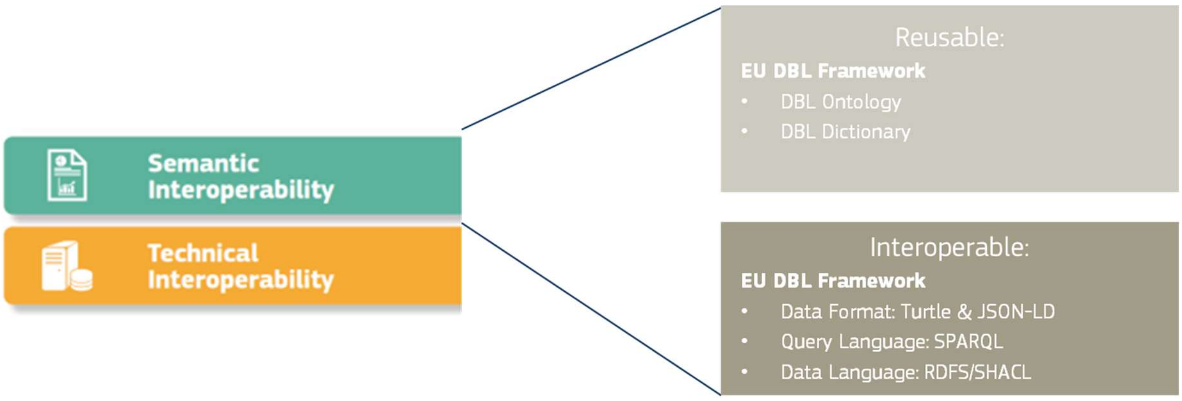
Beyond the scope of these guidelines, legal interoperability is about ensuring that organisations operating under different legal frameworks, policies and strategies can work together and organisational interoperability means documenting and integrating or aligning business processes and relevant information exchanged.

Figure 14: EIF and its general linkages with the FAIR principles



Following this logic, the understanding of FAIR on interoperability is a bit narrower than the EIF concept of interaction between organisations, with FAIR Interoperability reflecting only syntactic issues while the FAIR principle of reusability focuses on the idea of agreeing on common semantics in the form of dictionaries and ontologies. The **DBL Framework provides this semantic interoperability**. Using the DBL Ontology and the DBL Dictionary as well as the agreed data formats (Turtle, JSON-LD), query language (SPARQL) and data language (RDFS/SHACL) ensures both reusability and interoperability and thereby semantic interoperability (Figure 15).

Figure 15: EIF semantic interoperability versus FAIR - more detailed

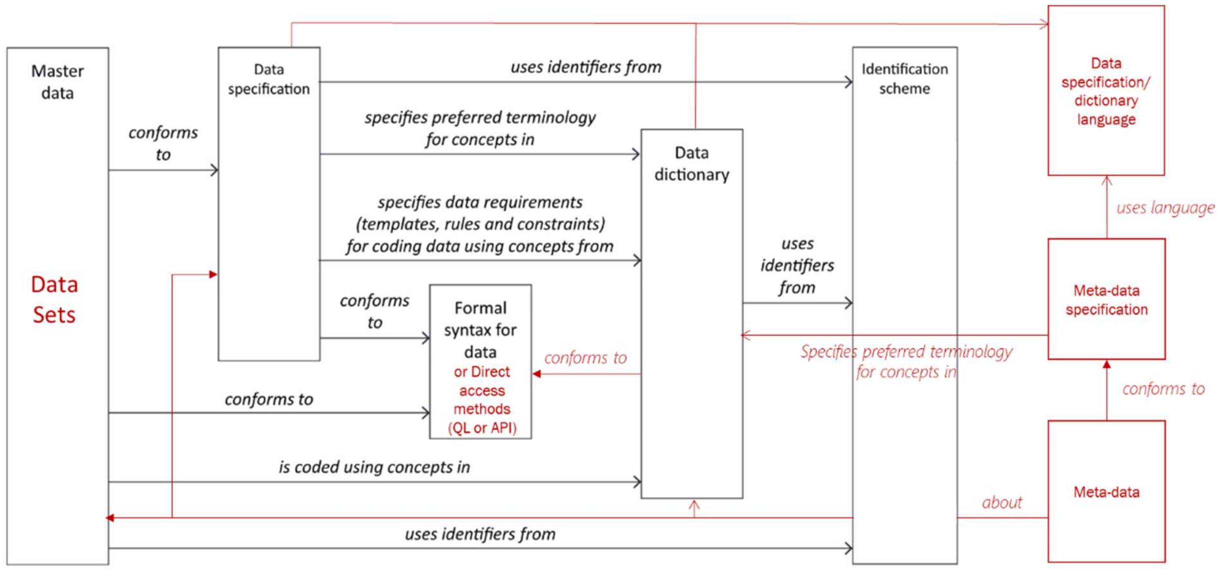


The guidelines in this handbook focus on the **Semantic Interoperability layer** (Section 3.3), covering both purely semantic aspects related to the meaning of data but also syntactic aspects of how the data (and its related semantics) is represented. In simple terms: it addresses the data format and the data structure. For the **Technical Interoperability layer** dealing with underlying data storage, data transport and data access aspects, some of the options to make a national DBL work and interoperable with other databases as well as an EU-level platform are discussed in Section 3.4.

3.3 Semantic interoperability

The data architecture for European and national DBLs embodies a structured approach to capturing, organising, storing, exchanging and sharing data related to building information. Our proposed data architecture is based on (an extension of) the standard ISO 8000 Part 110 Data Architecture (see Figure 16).

Figure 16: Data Architecture (in red are extensions on the original one)



3.3.1 Key Elements

Without going into too many details[30], the key elements of semantic interoperability in the FAIR principles are presented below:

Data specification: The data architecture defines a (possibly standardised) data specification or data model that represents the different aspects of building information. This model is based on aspects defined in existing data sources from different Member States and includes properties such as buildings, building units, parcels, systems, energy consumption, maintenance records, and other relevant attributes and relations.

- In the DBL Framework, this is the DBL Ontology (the term used for a specification in case linked data/semantic web technologies are used).

Data dictionary: where the data specification defines the semantic 'things', the dictionary defines the human-interpretable 'strings' used to name and denote different objects such as a building and its components.

Metadata[31] plays a crucial role in describing and organising the logbook data. It includes information such as the data source, timestamps/periods of validity, data quality, units of measurement, and other contextual details. Effective metadata management ensures data integrity, supports data

[30] The full explanation can be found in Deliverable D2.1 The Semantic Data Model.

[31] It should be noted that 'metadata' is often a confusing concept. Some definitions also see a data model itself as 'structural metadata'. In this study we use it in a narrower sense as being data about data.

governance, and enhances data discovery and understanding. It should be noted that, like normal data, metadata also requires its own (meta)data specification.

- Metadata describes both data sets as a whole and data on elements in those:
- For the first category, it is proposed to use the existing standard W3C DCAT metadata ontology also implemented by existing general data management systems such as CKAN.
- For the second category, an initial set of meta-attributes is proposed in our DBL Ontology/Dictionary (such as valid times, registration times, life-cycle phase/status etc.).

Finally, the data architecture of Figure 16 shows various **data technology elements** such as identification schemes, data formats/APIs/Query Languages and data languages to be used to define and access the data, metadata and their specifications and dictionaries.

- For these data technology elements, W3C Linked Data / Semantic Web technology is proposed. The key reasons include:
- Fully web-based
- Efficient at linking various data sources
- Well-defined formats and languages that are computer-processible/interpretable
- CEN TC442 WG4's EN-17632 Semantic Modelling and Linking (SML) standard is utilised for the standard application of these linked data technologies.

It is noted that there are alternative technological/modelling routes like bSI and CEN TC442 WG4 Data Templates (involving ISO 23386/23387 or the currently being developed EN ISO 16757) that need further integration development internally at CEN TC442 and externally via alignment between bSI and CEN TC442. The data templates are especially in scope for Digital Product Passport (DPPs).

These technologies are then used to define the semantic data model that can be extended by Member State authorities, building owners and construction professionals to control their actual data content (Table 5).

Table 5: Technological recommendations

Geographic level (Stakeholder)	Formats & Direct access methods	Identification scheme	Language	DBL Semantic Data Model: ontology & multi-lingual data dictionary (incl. DCAT 'meta-data' specifications)	DBLs (content incl. IDs and meta-data)
EU level (European Commission)	Like JSON-LD & SPARQL	Like Unique Object Identifiers (UOIs)	Like RDF, SKOS, RDFS, OWL, SHACL	Key result of this DBL project	Asserted/inferred by the European Commission
Member States level (National authorities / agencies)				Reused by agencies	Asserted/inferred by agencies
Data Provider level (building owners, construction professionals)				Reused by data providers	Asserted/inferred by data providers

3.3.2 The DBL Identification scheme

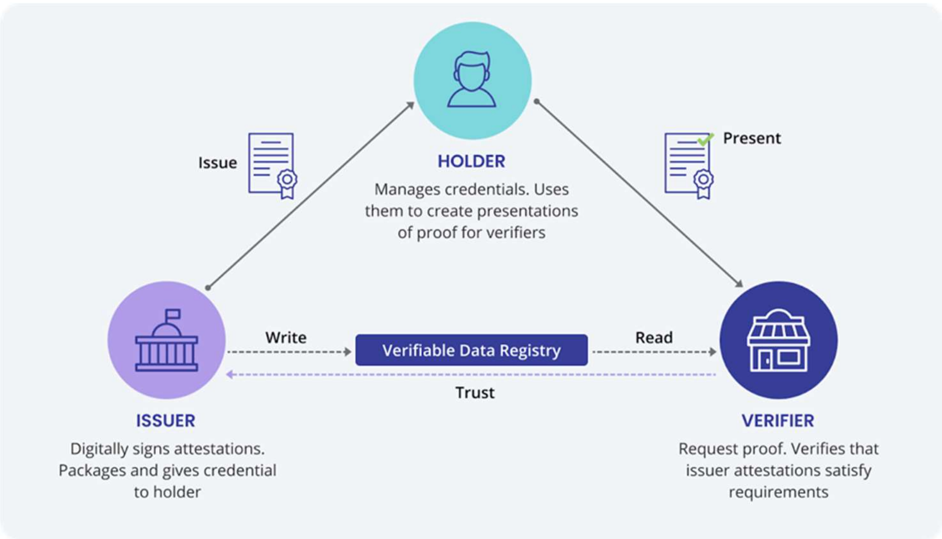
The purpose of a DBL is to make building data available. To that end, the data should be findable: the "F" from the FAIR principles. To find data, identifiers are needed at several levels. Firstly, users of a DBL should be enabled to identify individual buildings. Secondly, if authorisation is needed to access data, the DBL machine needs to be able to identify whether the user is an authorised user. And thirdly, the DBL machine must be able to extract the relevant building data. To the latter end, identifiers are needed for the DBL machine to find the right attributes and relations[32]. At the same time, data modellers need to be able to maintain the DBL ontology (definitions and structure of data). Therefore, user-friendly names for IDs are proposed such as "Building" and "energyPerformance", also following the Semantic Modelling and Linking (SML) guidelines for names.

Person identification schemes

For authorised access, a third mode of identification is relevant, namely the identification and authentication of individuals. On both demand (EU) and supply (standardisation bodies) sides, there is an enormous interest in electronic IDs (eIDs) as a key ingredient for better cross-border communication. At the European level, for example, the eIDAS Regulation and the concept of Self-Souverain Identities (SSIs) facilitate the authentication of individuals. However, this is not further discussed as the identification of individuals is not specific to DBLs.

At W3C technical solutions are offered for distributed eIDs like WebIDs and Distributed IDs (DIDs). The latter is now a W3C Recommendation and is also considered at eIDAS. Both approaches are web-based in the sense that they use URIs patterns. However, DIDs have the benefit of not needing a central registry and they provide verifiers for both issuers and holders of credentials (see Figure 17).

Figure 17: Key participants for DIDs



Currently, the CEN SML URI pattern/strategy and INSPIRE-compliant identifiers will be used for the DBL. This allows for an easy adaptation to any stable alternative approach emerging in the future.

[32] This follows the European data standard developed in WG4 ("Support Data Dictionaries")/TG3 of CEN TC442: EN 17632-1:2022 - Semantic Modelling and Linking (SML).

Building identification scheme – INSPIRE principles

Building identification – a definition

Identification of buildings is critical for a DBL because if one building is confused with another, having the wrong data may create more problems than having no data. Building identification simply means that the DBL machines correctly identify certain data as those belonging to the building that the user is looking for. This means that buildings should be uniquely identifiable. This may sound simple, but a national identification system will fail if different municipalities allocate the same number to different buildings. Building identifiers also should not change over time, because if a building has a new ID and the building data is still stored under the old ID, the building data may no longer be findable. This means for example that a building ID should remain the same even if the building is repurposed. This section first discusses how building identification is done in most countries and what are the issues. The section then discusses INSPIRE principles to identify building data and concludes with the final guideline.

Following INSPIRE[33] principles ensures that buildings are uniquely identified. Essentially, the INSPIRE principle is that local units should be free to assign numbers to buildings in their area and that more aggregate administrative units (for example the country) should include codes for less aggregate administrative units (for example a municipality).

Common practices across countries

In most countries, municipalities assign consecutive numbers to buildings independently of each other. Thus, two municipalities may assign the same number to two different buildings. To ensure that building identifiers are unique at the country level, national cadastres include a code for the municipality in the building identifier. If multiple districts within a city assign numbers, the city should include district identifiers in the municipal building identifier system. For this reason, national cadastres should allow a large number of digits for municipal building identifiers. Most countries reserve 12 or 16 digits for the local ID.

In federally organised countries, regions ensure that building identifiers are unique within their region, a national building identifier should include a code for the region in which the building is located.

Most countries have one national building identifier system. However, in some countries such as Spain, two national systems of building identifiers co-exist.[34] In that case, an overarching national building identifier system should include an identifier of the building identifier system, or the different authority managing their own national building identifiers.

Some countries such as the Netherlands include an object identifier in the building identifier. In the Netherlands, positions 5-6 of the building identifier for the code 10 “Buildings”, to differentiate the building identifiers from other identifiers such as for the public spaces (street, square, etc.), other accommodation objects, houseboat berths, caravan pitches, phone booths etc. Stakeholder opinions are divided about whether to include an object code in the identifier of the building (and other object codes in the identifiers of other types of objects).

While the inclusion of object codes in a building identifier is generally not problematic, a building identifier should <u>not</u> include the type of building. The reason is that a building may be repurposed, after which:

- either the identifier is revised to reflect the new type of building in which case two codes (the new and the old one) circulate for the same building,
- or the old identifier is not changed and no longer accurately indicates the type of building.

[33] See: Annex H in https://inspire.ec.europa.eu/documents/Data_Specifications/D2.5_v3.4rc3.pdf
[34] See: https://www.rodenasabogados.com/cru-idufir/#:~:text=El%20CRU%20es%20el%20C
Código,la%20propiedad%20en%20toda%20España.

A better alternative to indicate the type of building is to include the type of building in the building data.

At the EU level, building identifiers should include a country code. Hence, at the EU level, the identifier of a building in Spain could look like[35]:

es_catastro_navarra/pamplona-ayuntamiento/123456789

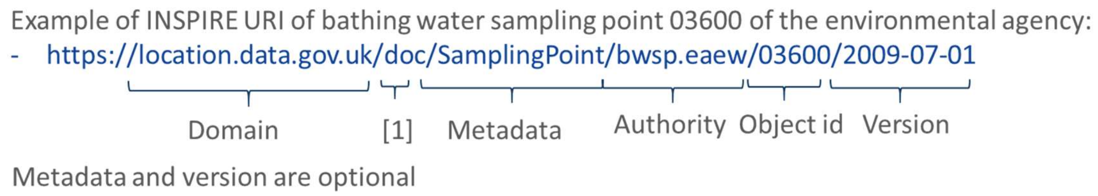
Where “es” refers to the country, catastro_navarra refers to the cadastral building identification scheme of the Navarra region and Pamplona-ayuntamiento to the municipal authority. The code 123456789 is the so-called local ID.

The INSPIRE principle and its application in building identification

Following INSPIRE[36] principles ensures that buildings are uniquely identified. Essentially, the INSPIRE principle is that local units should be free to assign numbers to buildings in their area and that more aggregate administrative units (for example the country) should include codes for less aggregate administrative units (for example a municipality). National cadastres already apply this principle, but as the case of Spain illustrates, buildings can still have different identifiers if more than one national institute has its own building registration system with its own building identifiers. The INSPIRE principle would then just be to add an identifier of the institute or the registration system to the building identifier.

Another aspect of the INSPIRE principle covers how users can find information on a building (or any other information) by browsing through menus. This is not the way buildings are identified in current DBL-like applications, where users can zoom in on a map to select a building, can type in an address or a cadastral number. If a DBL contains limited data on each building, the data could be shown if a mouse hovers over the building, or by right-clicking to open a menu with subfolders for different types of information on the building.

Building data can be presented in many ways, e.g. with a drop-down list, through menus, through tree-search etc. However, **if a DBL has a lot of data on a building** that can only be fully shown in a document with many pages, it might make sense to generate a separate document for that building. In that case, it becomes relevant how the DBL machine or the user can identify that document. The INSPIRE identification principle to identify a document is to build up a URI link to that document in components in the order of decreasing significance from left to right, for example (based on the Inspire Guideline cited above):



In the above example, the “doc” classification in the URL indicates that documents can be found under that component. All metadata that describes sampling point data can be found under the component “SamplingPoint”, and actual sampling point data can be found under the component “bwsp.eaew”.

The real-world identifier should be in a list that can be found on the web as in the following example.

[35] The written-out names serve illustrative purposes and could be replaced by shorter codes.

[36] See Annex H in https://inspire.ec.europa.eu/documents/Data_Specifications/D2.5_v3.4rc3.pdf

Here, "id" indicates that identifiers are listed under this component, "bathing-water" indicates a list of bathing water identifiers, and "ukc2102-03600" is one of the identifiers listed under "bathing-water".

As long as documents have a clear structure, documents can be machine processed. A practical example of a portal where building information is provided in the form of documents is the Slovak cadastre portal (see also Annex V).[37]

Other ways of uniquely identifying buildings

The most convenient way for users to search for a building is by zooming in on a map or by typing an address. It is recommended to allow both ways to search for a building, as most current DBL-like portals do.

If a DBL portal features a map to select buildings, buildings should be identified by **geographical coordinates**. In Europe various coordinate systems have been developed, however, the worldwide GPS coordinate system is most widely used in applications which makes it more future-proof and its wide applicability is thus recommended.[38]

For **addresses**, INSPIRE guidelines provide a system to avoid issues such as mixing up house numbers and floor numbers which is discussed further below. Addresses can also uniquely identify buildings. The INSPIRE Guidelines provide the following address components[39]:

1. Address identifier general = Address identifier composed of numbers and/or characters
2. Address number = Address identifier composed only of numbers
3. Address number extension = Extension to the address number
4. Address number 2nd extension = Second extension to the address number
5. Building identifier = Building identifier composed of numbers and/or characters.
6. Building identifier prefix = Prefix to the building identifier
7. Corner address 1st identifier = Address identifier related to the primary thoroughfare name in a corner address
8. Corner address 2nd identifier = Address identifier related to the secondary thoroughfare name in a corner address
9. Entrance door identifier = Identifier for an entrance door, entrance gate, or covered entranceway
10. Floor identifier = Identifier of a floor or level inside a building.
11. Kilometre point = A mark on a road whose number identifies the existing distance between the origin point of the road and that mark, measured along the road
12. Postal delivery identifier = Identifier of a postal delivery point
13. Staircase identifier = Identifier for a staircase, normally inside a building
14. Unit identifier = Identifier of a door, dwelling, suite or room inside a building

---

[37] See https://zbgis.skgeodesy.sk/mkzbgis/sk/kataster/detail/kataster/parcela-c/852228/1097_6?bm=zbgis&pos=48.713652,20.130300,20

[38] See: https://www.crs-geo.eu/, or https://www.techtarget.com/whatis/definition/GPS-coordinates#

[39] See: https://inspire.ec.europa.eu/codelist/LocatorDesignatorTypeValue or https://inspire.ec.europa.eu/id/document/tg/ad

Two towns in the same country may share the same name, such as two municipalities called "Bergen" in the Netherlands, and they may happen to use the same name for one of their streets. In that case, the address should pop up twice, with the name of the respective region in which the municipalities are located (Limburg and North-Holland in the case of Bergen), and the user should select the appropriate address.

The DBL proposal

To summarise the various proposals the following guidelines are proposed for identification:

1. Eliminate the risk of duplicate building identifiers by identifying all authorities that could allocate the same number to different buildings and include the code for those authorities in the building identifier.
2. Follow a uniform structure for building identifiers and also for addresses, to avoid mixing up for example codes for regional and local authorities in identifiers, or house numbers and floor numbers in addresses.
3. Allow for user-friendly building searches based on typing in an address or zooming in on a map to select a building. In case of multiple hits, show all hits to allow the user to select the one they are looking for.
4. If all data on one building takes multiple pages to present, generate a document.

If a document is generated, it is important to fill the document with only the information that the user is authorised to access. This limitation is not specific to showing data in one document: the same care needs to be taken if the data is shown in for example a screen pop-up. Hence, if a DBL contains information that is only available to authorised users, it is advisable to generate the document after each user request, depending on his degree of authorisation.

Generating a document has the additional advantage that it can easily be downloaded. However, this comes with the risk that an authorised user shares the downloaded document with an unauthorised user (willingly or unwittingly). To mitigate this risk, the Data Governance Act provides that a public authority (such as the DBL manager) needs to require authorised users to take appropriate measures to avoid leakage of authorised data, as further discussed in Section 5.3.

Lastly, it is noted that the most demanded feature by professional users is the selection of all buildings that meet certain criteria. For example, a query feature to give all buildings in Brussels a floor space of at least 5,000 m^2 and built after 1990. However, the selection of buildings and the identification of buildings are two separate issues. The selection of all buildings meeting certain criteria requires a fast search engine (Section 3.4 and 4.3) and a semantic data model to extract information from building records, where the extraction should also include the building identifier so that later additional information for the selected buildings can be extracted. Generating documents with all requested information per building would be a third advantage in managing the information exchange. It should also be noted that the selection of all buildings that meet certain criteria could lead to the risk of profiling, an aspect that is discussed in Section 5.2.

3.3.3 The DBL Semantic Data Model (dictionary and ontology)

A key component of the proposed DBL approach is a **Semantic Data Model** for all relevant DBL data sources. This data model consists of two parts:

- A **dictionary**, defining the building-related terms used, and
- An **ontology** that uses these terms as names for concepts, attributes, relations and all kinds of constraints that should hold for cardinalities and actual values.

Using the dictionary, **the ontology controls the DBL data sets, their possibilities and impossibilities**. Standards used by the ontology such as SML (CEN TC442) are introduced in 'D2.1 The DBL Semantic Data Model'.

Below, an example is provided to illustrate how this would work in practice for the concept referred to by the term 'Building'. The example first showcases how the ontology expresses the term 'Building' in linked data language and provides its definition as well as the definition of the term 'gross volume'. It then shows which terms are linked to both concepts (e.g. quantity and units for the concept 'gross volume').

In the example below, the ontology describes the dbl-terms "building" and "gross volume", which can be used to inform the user of the definitions of those terms. The ontology defines a building as a class (a set of things), and gross volume as a property of a building that should have a numerical value in cubic metres. Each building should have an identifier (_1 in the example), and a field and format for the gross volume, where the field needs to be filled with a number in the specified format (for example, 5000 or 5,000).

Box 5: Ontology example of the term 'Building'

The resulting ontology in linked data language (in the Turtle format) becomes:

```
dbl-term:Building
  a skos:Concept ; -- note: 'a' is a shortcut for rdf:type
  skos:prefLabel "Buiding"@en ;
  skos:definition "Enclosed constructions above and/or underground which are intended or used for the shelter of humans, animals, things or the production of economic goods and that refer to any structure permanently constructed or erected on its site."@en ;
.
dbl-term:grossVolume
  a skos:Concept ;
  skos:prefLabel "grossVolume"@en ;
  skos:definition "The total volume of all interior spaces in a building or building unit over the gross floor area. This total volume is enclosed by the outer boundary surfaces of the foundation, the exterior walls and the roof (including the dormers and skylights) (DIN 277-1 2005)."@en ;
.
```

These terms are used by the concept definition:

```
dbl:Building
  a rdfs:Class ;
  rdfs:seeAlso dbl-term:Building .

dbl:grossVolume
  a rdf:Property ;
  rdfs:seeAlso dbl-term:Building ;
  rdfs:domain dbl:Building ;
  rdfs:range sml:QuantityValue ;
  qudt:unit unit:M3 .
```

This can be used to model a specific instance in a data set:

```
ex:Building_1
  a dbl:Building ;
  dbl:grossVolume_1 .

ex:QuantityValue_1
  a sml:QuantityValue ;
  qudt:numericvalue “364”^^xsd:float .
```

CEN TC442 SML introduces several variants of property modelling. As the example shows, constructs from the European CEN SML standard[40], such as ‘sml:QuantityValue’ and ‘unit:M3’, are directly reused. The elements ‘rdfs:Class’ and ‘rdf:Property’ are language constructs from, respectively, the RDFS and the RDF languages.

Please note that these definitions of a building are combinations of ‘triples’ separated by semi-colons. A ‘triple’ in RDF languages codifies a statement about semantic data. Doing so, follows the “subject – predicate – object” logic. For example, two triples with the subject ‘Building_1’ are presented below. In the triplet of the first example, a specific building identified as “Building_1” is the subject. The term “is” or “belongs to the class” is the predicate and Building is the class of all buildings. In the second triplet, the building is again the subject, the predicate is “has” and the object is the gross volume.

Triple 1:	ex:Building_1 a dbl:Building .	The object ‘Building_1’ is a building as defined in the dictionary.
Triple 2:	ex:Building_1 dbl:grossVolume ex:QuantityValue_1 .	The object ‘Building_1’ has a certain gross value.

“Triples” are combined by introducing a semicolon in Turtle:

```
ex:Building_1
  a dbl:Building ;
  dbl:heightAboveGround ex:QuantityValue_1 .
```

This combination indicates that the ‘grossVolume_1’ key-value pair also belongs to ‘Building_1’.

A core common ontology was built, which is described in the next paragraphs. More information can be found in Annex I.

[40] The DBL Ontology follows all syntactic, semantic and pragmatic recommendations by the European data standard developed in by CEN TC442 WG4 (“Support Data Dictionaries”)/TG3: EN 17632-1:2022 -Semantic Modelling and Linking (SML) in the built environment. This means that it selects its data languages, data serialisations, direct access methods and modelling patterns for example for IDs from this standard. *See also: Semantic Modelling and Linking standard, Part 1 (generic) and Part 2 (domain specific) by CEN TC442 WG4 (support data dictionaries).*

Key concepts of the DBL Framework

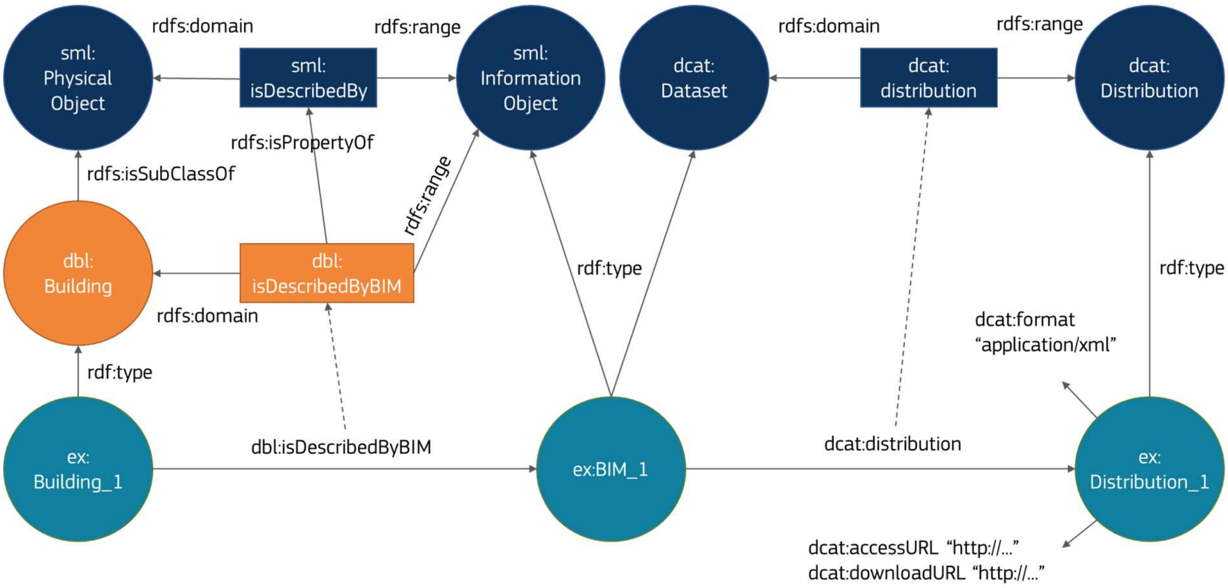
A building logbook includes instance data classified according to the following three key concepts heavily based on the European INSPIRE specification[41]:

1. Building;
2. BuildingUnit, being a part of a building related to one facility/activity complex (INSPIRE-term) or functional entity (SML-term);
3. CadastralParcel.

In terms of the top-level ontology of SML, all three concepts are seen as subclasses of both 'sml:SpatialRegion' and 'sml:TechnicalEntity'. In addition, various generic "helper concepts" exist such as Address, Agent (Person or Organisation), Location, Geometry etc.

The three concepts have sets of (pre)defined attributes and relations where relations can also target a whole data set (instances of 'sml:InformationObject' and/or 'dcat:Dataset' to be precise). Because of the use of Linked Data, such a link can be any local or remote data resource depending on its URI-based ID. This is shown in Figure 18.

Figure 18: Physical and information object



Bi-temporal logic

For all three key entities (building, building unit and cadastral parcel) a 'bi-temporal logic' is applied. This logic grasps their various lifespan states in time but also provides relevant assertion times on the meta-level.

Many current approaches apply these ideas on the object/instance level enabling traceability and reproducibility and avoiding data loss. For flexibility, the bi-temporal logic was applied on the (lower) property level since properties can come from many data sources all having their independent life span states and assertion times from different data sources and parties. This way, a more continuous situation for the object is obtained which in this case must be derived. In other words, with this approach the object state becomes emergent. It should be noted that properties cover both attributes and relations. Therefore, a change in building owner is also considered a state change for a building.

[41] For more information on INSPIRE, see: https://inspire.ec.europa.eu/data-specifications/2892.

This approach matches well with the decision to model the properties in the SML complex variant way allowing for not only units but also other kinds of meta-data like these various timing aspects.

As a result, a kind of "property-states" is introduced. In practice, when properties are combined at a certain point in time, the logbook can show an on-the-fly result in an overall object state (sml:State in SML) typically having "beginning-" and "end-point" relations from some event (sml:Event).

NOTE 1: SML states are only dealing with object states, not property states. SML states (and events connecting them) can still be used where:

- *Object states are not asserted but derived when needed (from the various more detailed/atomic property value states)*
- *Object states can differ in multiple property values*

NOTE 2: Other metadata properties than actual value or unit result in parallel property-states: a new property state is only changed based on its value change; not its life cycle 'status' indication or 'reference' indication.

The seven aspects of the DBL Framework

All properties (attributes or relations) included in the DBL Framework are grouped according to a set of seven aspects. These aspects have been chosen based on a review of existing data sets as well as based on stakeholder feedback. For flexibility, all properties except the identifiers, are optional. The seven aspects are presented in Table 6 below.

Table 6: Overview of the seven DBL aspects and their properties

Aspects	Properties
1. Identification	a. Building ID; building unit ID; cadastral parcel ID; online link ID like InspireID
2. General	a. Relationships between key DBL entities (Building, BuildingUnit, CadastralParcel) b. Indication of types or subtypes (via simple attributes, not full taxonomy) c. Address, placement indicators like geo coordinates d. Use function (residential, non-residential and sub-categories) at the building level and building unit level; user profiles like students, seniors, asylum seekers etc. e. Dates of construction, permits, renovation etc. f. Documents: Urban licenses, renovation proposals, renovation passports etc. g. Documents: BIM model, technical drawings etc.

Aspects	Properties
3. Legal and Finance	a. Property tax valuation; lifecycle cost; annual maintenance cost; rental value and maximum rental value; sales transaction value etc. b. Dates of valuations c. Documents: sales deed, tenancy agreement, insurance policy, clean soil statement, rule violations
4. Dimensions	a. Lengths, gross and net areas & volumes b. Linked geometric representations (0D, 1D, 2D, 3D)
5. Performance	a. Functionality offered incl. connection to utility services, indoor health & comfort levels b. Energy performance label, circularity label, energy and water use label, CO2 and N2 emissions label, smart readiness indicator etc. c. Actual energy and water consumption and production levels d. Document: Energy Performance Certificate (EPC)
6. Structure & Material	a. Number of / breakdown in zones, floors, spaces/rooms, elements, components, products, materials b. U-values for various element types c. Year of latest materials inspection, asbestos check etc. d. Certain materials as asbestos for authorised DBL users only
7. Building Services	a. Energy (gas/electricity/solar/thermal/city heating, ...), production and consumption installations b. Ventilation system c. Water and sewerage installations d. Number of elevators, balconies, swimming pools etc. e. Building automation parameters f. Security for authorised DBL users only g. Telecommunications connectivity h. Electrical (&gas) safety inspection report i. Year of latest safety inspections j. Year of latest installation or repair k. Year of latest fire security check, elevator and balcony safety check, swimming pool salmonella check etc. l. Documents: elevator and fire safety inspection certificates, maintenance contract, utilities contracts

These aspects are all coded as standard SML groups represented by ‘rdfs:Container’ having the mentioned properties as their members (via ‘rdfs:member’).

NOTE 3: Links to unstructured documents or non-LD structured data sets are modelled via sml:InformationObject and/or dcat:Dataset as a whole, linking from there to the actual files via rdfs:seeAlso respectively dcat:distribution/(dcat:accessURL and/or dcat:downloadURL).

Life cycle phase indication

Information on buildings, building units and parcels is typically associated with, or even created in a specific life cycle phase. These life cycle phases can be summarised as:

- Programming (As-required);
- Design (As-designed);
- Build (As-built); and
- Use (As-used).

Considering that buildings experience changes during their lifetime due to maintenance, renovation, repurposing and demolishment, one should note that:

- **Maintenance** is regarded as a new as-built & new as-used;
- **Renovation** is regarded as new as-designed & new as-built & new as-used;
- **Repurposing** is regarded as a new as-required & new as-designed & new as-built & new as-used; and
- **Demolishment/Recycling** is regarded as an ultimate form of repurposing ("no more purpose").

As with the unit information, this status is optionally attached to the property value level via a dbl:status property that points to an enumeration type with the following allowed values:

- dbl:As-required
- dbl:As-designed
- dbl:As-built
- dbl:As-used

Please note that simply changing the status of a property and not the property value itself is not regarded as a new property state since a required value stays relevant in parallel if an as-designed or as-built value is added. Formulated the other way round: there can be multiple property states with the same status (but in different time periods) reflecting maintenance, renovation or repurposing events.

Extending the core DBL Ontology/Dictionary

Depending on specific data needs anybody can define their dictionary/ontology by importing the current common core DBL Framework and extending them by following 10 steps. These 10 steps are illustrated in Figure 20 and thereafter explained with an example.

Since all ontologies and dictionaries in linked data are "textual", one can make such changes in principle in any text editor. However, a semantic development tool is typically used, such as the commercial TopBraid Composer (TC) by TopQuadrant or the open-source Protégé (see Figure 199).However, a semantic development tool is typically used, such as the commercial TopBraid Composer (TC) by TopQuadrant or the open-source Protégé (see Figure 19).

Figure 19: Example of semantic tooling

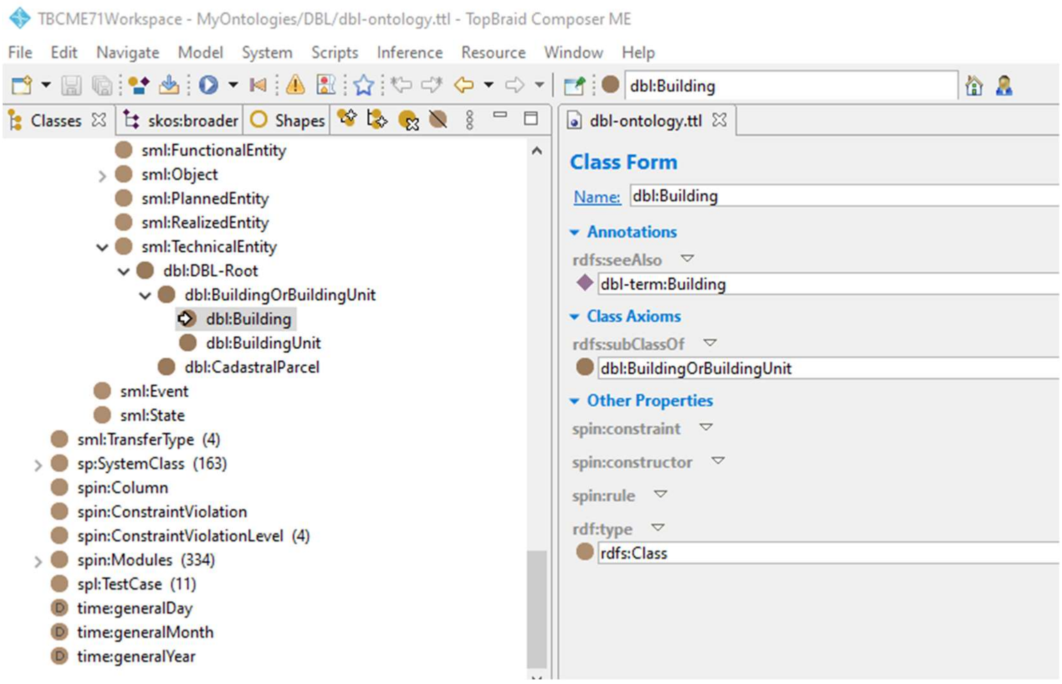
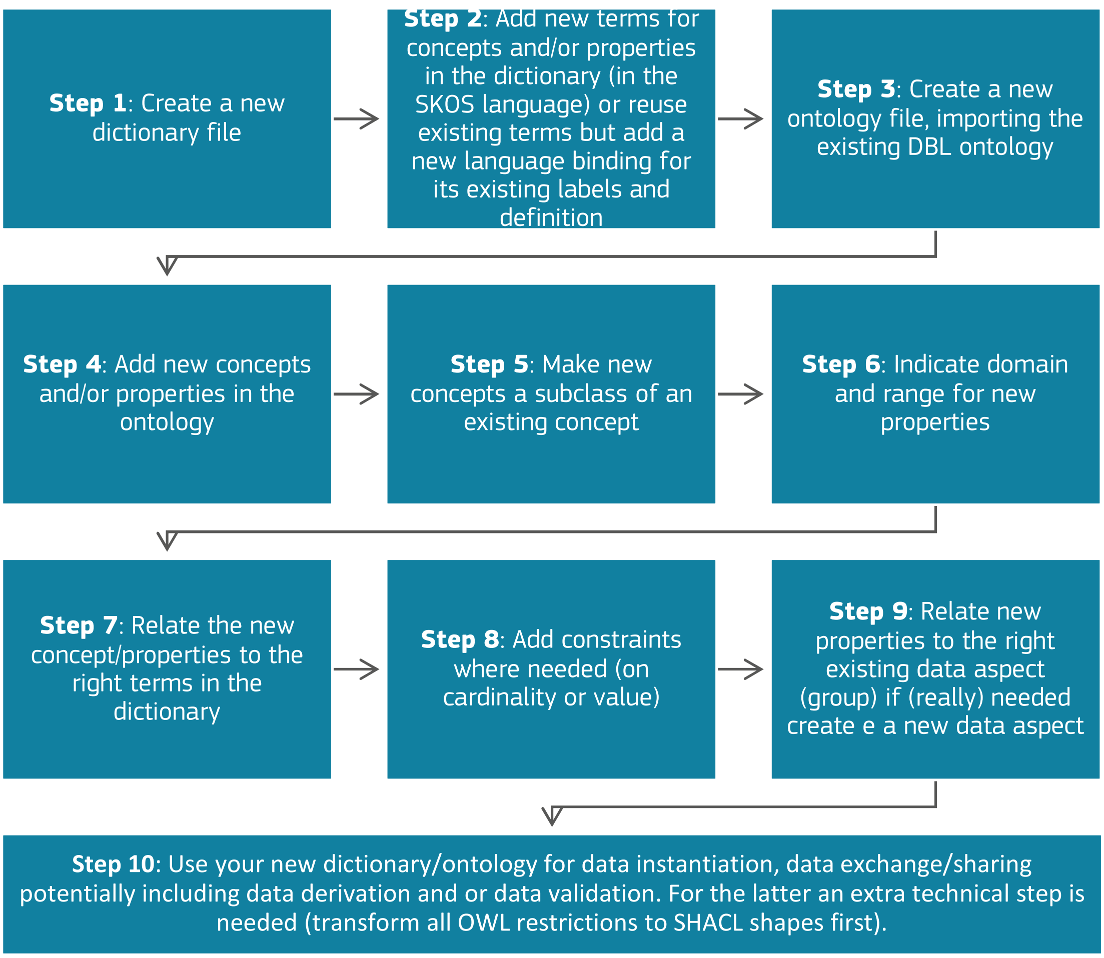


Figure 20: 10 steps for extending the core DBL Ontology/Dictionary



This 10-step approach is illustrated below with a simple example. Suppose a Member State like The Netherlands has a lot of special (technical) buildings like “Greenhouses” with the obligatory (non-optional) attribute “horticultureType” that can be “FloriCulture” or “FoodHorticulture”. Let’s follow the **10 steps** for this extension of the core model.

Step 1 – Define a file NL[42]-dbl-dictionary.ttl with the following Turtle header content.

```
# baseURI: https://data.overheid.nl/dbl/skos/term
@prefix nl-dbl-term: <https://data.overheid.nl/dbl/term#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .

<https://data.overheid.nl/dbl/skos/term>
  a skos:ConceptScheme .
```

Note 1: Because of modelling no weak taxonomy on the dictionary level and add (not change) existing terms there is no need to import the existing core DBL dictionary.

Step 2 – Add the new terms to the content.

```
nl-dbl-term:Greenhouse
  a skos:Concept ;
  skos:prefLabel "Greenhouse"@en ;
  skos:prefLabel "Kas"@nl ;
  skos:definition " A greenhouse is a building with glass walls and a glass roof used to grow plants, such as tomatoes.@en ;
  skos:definition "Een kas is een gebouw met glazen wanden en een glazen dak waarin planten worden gekweekt, zoals tomaten.@nl .

nl-dbl-term:horticultureType
  a skos:Concept ;
  skos:preflabel “horticultureType"@en ;
  skos:preflabel “tuinbouwType"@nl ;
  skos:definition "The possession of the kind of product that is grown."@en ;
  skos:definition "Het hebben van het soort product dat geteeld wordt.”@nl .

nl-dbl-term:HorticultureType
  a skos:Concept ;
  skos:preflabel “HorticultureType"@en ;
  skos:preflabel “TuinbouwType"@nl ;
```

[42] Netherlands/Dutch

```
  skos:definition "The kind of product that is grown. The choice of allowed values is FloriCulture or FoodHorticulture"@en ;
  skos:definition "Het soort product dat geteeld wordt. De keuze in mogelijke waarden is Sierteelt of VoedingsTuinbouw"@nl .

nl-dbl-term:FloriCulture
  a skos:Concept ;
  skos:preflabel “FloriCulture"@en ;
  skos:preflabel “Sierteelt"@nl ;
  skos:definition "Growing flowers, plants, trees, sod, etc."@en ;
  skos:definition "Het telen van bloemen, planten, bomen, graszoden etc.”@nl .

nl-dbl-term:FoodHorticulture
  a skos:Concept ;
  skos:preflabel “FoodHorticulture"@en ;
  skos:preflabel “Voedingstuinbouw"@nl ;
  skos:definition "Growing food like fruit and vegetables.”@en ;
  skos:definition "Het telen van voeding als groente en fruit.”@nl .
```

Step 3 – Define a file nl-dbl-ontology.ttl with the following Turtle header content.

```
# baseURI: https://data.overheid.nl/dbl/rdfs/def
# imports: https://data.europa.eu/dbl/rdfs/def
# imports: https://data.overheid.nl/dbl/skos/term

@prefix nl-dbl: <https://data.overheid.nl/dbl/def#> .
@prefix dbl: <https://data.europa.eu/dbl/def#> .
@prefix nl-dbl-term: <https://data.overheid.nl/dbl/term#> .
@prefix dbl-term: <https://data.europa.eu/dbl/term#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix quantitykind: <http://qudt.org/vocab/quantitykind/> .
@prefix qudt: <http://qudt.org/schema/qudt/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix sh: <http://www.w3.org/ns/shacl#> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix sml: <https://w3id.org/sml/def#> .
@prefix unit: <http://qudt.org/vocab/unit/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix dcat: <http://www.w3.org/ns/dcat#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix wgs84_pos: <http://www.w3.org/2003/01/geo/wgs84_pos#> .
@prefix geo: <http://www.opengis.net/ont/geosparql#> .
```

```
@prefix locn: <http://data.europa.eu/m8g/> .

<https://data.overheid.nl/dbl/skos/term>
  a owl:Ontology ;

  owl:imports <https://data.europa.eu/dbl/rdfs/def> ;
  owl:imports <https://data.overheid.nl/dbl/skos/term> ;
```

Step 4 – Add to the content the new concept and (complex) property.

```
nl-dbl:Greenhouse
  a rdfs:Class .

nl-dbl:horticultureType
  a rdf:Property .

nl-dbl:HorticultureType
  a sml:EnumerationType .

nl-dbl:FoodHorticulture
  a nl-dbl:HorticultureType .

nl-dbl:FloriCulture
  a nl-dbl:FloriCultureType .
```

Step 5 – Add the statement on the supertype:

```
nl-dbl:Greenhouse
  a rdfs:Class ;
  rdfs:subClassOf dbl:Building ;
  rdfs:seeAlso dbl-term:Greenhouse .
```

Step 6 – Add the statements on domain and range:

```
nl-dbl:horticultureType
  a rdf:Property ;
  rdfs:domain nl-dbl:Greenhouse ;
  rdfs:range dbl:RelationReference ;
  rdfs:seeAlso nl-dbl-term:horticultureType .
```

Step 7 – Add the links to the corresponding terms:

```
nl-dbl:Greenhouse
  a rdfs:Class ;
  rdfs:subClassOf dbl:Building ;
```

```
  rdfs:seeAlso dbl-term:Greenhouse .

nl-dbl:horticultureType
  a rdf:Property ;
  rdfs:domain nl-dbl:Greenhouse ;
  rdfs:range dbl:RelationReference ;
  rdfs:seeAlso nl-dbl-term:horticultureType .

nl-dbl:HorticultureType
  a sml:EnumerationType ;
  rdfs:seeAlso nl-dbl-term:HorticultureType .

nl-dbl:FoodHorticulture
  a nl-dbl:HorticultureType ;
  rdfs:seeAlso nl-dbl-term:FoodHorticulture .

nl-dbl:FloriCulture
  a nl-dbl:FloriCultureType ;
  rdfs:seeAlso nl-dbl-term:FloriCulture .
```

Step 8 – Add an open world constraint in the form of an OWL restriction (minimum cardinality being 1 i.s.o. default 0).

Note 2: This is done via a superclass. Instead, the class for which the restriction holds should be a subclass of the class of all things having that restriction.

```
nl-dbl:Greenhouse
  a rdfs:Class ;
  rdfs:subClassOf dbl:Building ;
  rdfs:subClassOf [
    a owl:Restriction ;
    owl:onProperty nl-dbl:horticultureType ;
    owl:minCardinality "1"^^xsd:nonNegativeInteger ;
  ] ;
  rdfs:seeAlso dbl-term:Greenhouse .
```

Step 9 – Add the grouping information for the property added.

```
dbl:General
  a rdfs:Container ;
  rdfs:seeAlso dbl-term:General ;
```

```
            rdfs:member nl-dbl:horticultureType .
```

Step 10 – Start instantiating the greenhouse specification for a specific greenhouse (a greenhouse you can or could point at in reality). Including inherited property defined for relevant superclasses.

```
          ex:MyGreenhouse
            dbl:inspireID NL-BAG-1ghgh$$hhhhh999 ;
            dbl:dateOfConstruction “2012-03-24” ;
            nl-dbl:horticultureType nl-dbl:FoodHorticulture .
```

3.4 Technical interoperability

3.4.1 Introduction

This section discusses several technical aspects of managing a DBL portal. A DBL portal needs to perform many functions. It needs to collect data, validate data, manage access to data (in particular authorisation of users and payment for certain data, if applicable), make data accessible, and explain data (metadata). The nine key features of a DBL that ensure these basic functionalities are discussed below in this introduction.

A DBL portal has three layers: a front-end layer managing the data sharing with the end users, a back-end layer managing the data collection from data providers, and a data layer for the validation, processing and storage of the actual building data. These three layers are also discussed shortly below in this introduction.

A DBL will necessarily encompass a lot of technology. Without going into detail, the first subsection after this introduction briefly defines eight data technology concepts that portal developers should provide clear information on to their clients, in the context of these guidelines the public authority responsible for the DBL.

The remainder of this section discusses the frontend setup, the backend setup and maintenance aspects of the data layer.

Key features and components that a national gateway might include:

1. **Data Integration and Aggregation:** The gateway acts as a central integration point for collecting logbook data from various sources, including individual buildings, facility management systems, energy monitoring systems, and maintenance management systems. It aggregates the data to create a unified view of the building portfolio at a national level.
2. **Data Validation and Quality Control:** The gateway performs data validation and quality control checks to ensure the accuracy, completeness, and consistency of the logbook data. It verifies data integrity, performs data cleansing, and applies data quality rules to identify and correct any discrepancies or anomalies.
3. **Data Standardisation and Harmonisation:** The gateway enforces data standardisation and harmonisation across different logbook sources. It aligns diverse data formats, units of measurement, and terminology to establish a common representation and ensure interoperability between systems. This standardisation facilitates data exchange and enables meaningful analysis.

4. **Data Security and Access Control:** The gateway incorporates robust security measures to protect logbook data. It implements access controls, encryption, and user authentication mechanisms to ensure that only authorised individuals or systems can access the data. Compliance with data protection regulations, such as the General Data Protection Regulation (GDPR), is a crucial aspect of the gateway's design.
5. **API and Data Exchange:** The gateway provides APIs and data exchange mechanisms to facilitate seamless integration and data sharing between the DBL and other systems. It enables a secure, standardised data transfer using protocols such as REST. The gateway supports both data ingestion into the logbook and data retrieval for external systems.
6. **Metadata Management:** The gateway manages metadata associated with logbook data. It stores information such as data sources, timestamps, data quality metrics, and other contextual details. This metadata enables better understanding, searchability, and traceability of logbook data across different stakeholders and systems.
7. **Data Analytics and Reporting:** The gateway may incorporate data analytics and reporting capabilities to generate insights and reports from logbook data. It can provide visualisations, key performance indicators (KPIs), and customisable dashboards to support decision-making, performance monitoring, and compliance reporting at European and national levels.
8. **Interoperability with National Systems:** For a national logbook gateway, interoperability with existing national systems becomes crucial. The gateway should be designed to integrate with national databases, reporting platforms, and regulatory systems to enable efficient data exchange and compliance monitoring at a national level.
9. **User Interface and Stakeholder Engagement:** The gateway may include a user interface that allows stakeholders, such as building owners, facility managers, regulatory bodies, and auditors, to access and interact with logbook data. The user interface provides intuitive navigation, search functionalities, and collaboration features to engage stakeholders and promote adoption.

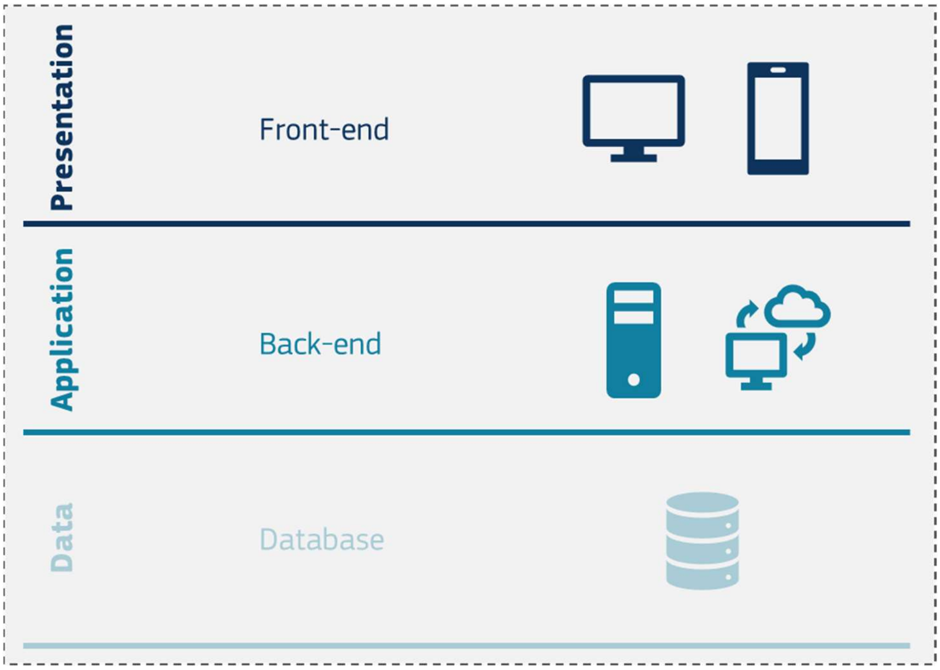
It is important to note that the exact architecture and functionality of a national gateway for a DBL may vary based on specific regional requirements, standards, and existing infrastructure. The gateway's design should prioritize scalability, security, interoperability, and ease of use to ensure effective management and exchange of logbook data across Europe or a specific country.

The DBL application architecture should be a traditional three-tier architecture that consists of:

1. Presentation (front-end) layer.
2. Application (back-end) layer.
3. Data layer.

The presentation layer communicates with the application layer over the HTTPS protocol. For that purpose, the logbook back-end will comprise several microservices that provide REST APIs and serverless handlers to provide data stored in relational and graph databases for the front end.

Figure 21: The three-tier architecture of Logbook portal



Separation of the concerns is important in terms of the development and maintenance of the application, allowing it to debug, improve and extend functionalities more easily.

3.4.2 The data technology

A DBL will encompass a lot of data technology, which would cost a lot of time to develop from scratch. Therefore the general recommendation is to use broadly adopted off-the-shelf data technology. Data technology means the tools, technologies, and platforms to manage, search and browse the logbook data. Because the DBL is not a fixed concept and should evolve, an overview of types of tools that could help develop a DBL in the near or further future is presented:

- **Database Management Systems (DBMS):** DBMS is the foundation of data storage and management. Relational databases, such as MySQL, PostgreSQL, or Microsoft SQL Server, are commonly used for structured logbook data. NoSQL databases like graph databases, MongoDB or Cassandra may be employed for handling unstructured or semi-structured data.
- **Data Governance and Metadata Management:** Data governance tools and metadata management systems help enforce data quality standards, ensure compliance, and maintain metadata associated with logbook data. These tools can support data lineage, data cataloguing, and data stewardship processes.
- **Cloud Computing:** Cloud platforms offer scalable and cost-effective infrastructure for hosting logbook data and related applications. Cloud services provide flexibility, high availability, and data redundancy.
- **Data Integration and ETL (Extract, Transform, Load):** Integration tools and Extract, Transform, Load (ETL) processes enable the extraction of data from various sources, transformation to a common format, and loading into the logbook database. Tools like Apache Kafka, Apache Nifi, or Sparql may be used for data integration.
- **APIs and Web Services:** Application Programming Interfaces (APIs) and web services allow for data exchange and integration between the digital logbook and external systems,

such as building management systems, energy monitoring systems, or maintenance management systems. REST, a standard protocol, is commonly used for API implementation.

- **Data Visualisation and Reporting:** Data visualisation tools and reporting platforms enable the creation of interactive dashboards, charts, and reports to present logbook data in a meaningful way. Tools like Tableau, Power BI, or Grafana can be employed to visualise key performance indicators, energy consumption trends, and maintenance metrics.
- **Machine Learning and Artificial Intelligence:** Advanced data technologies like machine learning and artificial intelligence can be utilised to analyse logbook data, identify patterns, and derive insights. Techniques like predictive maintenance, anomaly detection, or energy optimisation can be applied to optimise building performance. Libraries and frameworks like TensorFlow, PyTorch, or scikit-learn are commonly used in machine learning applications.
- **Data Security and Privacy Measures:** Robust data security measures, including encryption, access controls, and user authentication mechanisms, are crucial for protecting logbook data. Compliance with data protection regulations, such as GDPR in the European Union, is essential to ensure the privacy of personal information.

3.4.3 Front-end setup

Where the back-end setup manages the interaction between the DBL platform and the data providers, the front-end setup manages the interaction between the DBL platform and the users. It should manage aspects such as site browsing, searching a building, accessing building information, and where applicable authorisation if certain data is only accessible for authorised users, and payment if the user needs to pay for certain data.

In this section we discuss the current standard for building websites (Single Page Applications – SPA), how users can search for buildings, and what current services can be used to manage authorisation and payment. Specifically for an EU DBL portal, we discuss a short-term vision and a long-term vision. Finally, we present a mock-up to show the possible look and feel of such an EU portal.

Single-page applications

Over the last decade, the popularity of single-page applications has been increasing. Nowadays, several mature and powerful JavaScript libraries and frameworks for SPA are available and they allow the building of powerful, fast and functional websites displaying dynamic data. To easily fulfil the standards for modern, responsive websites, the logbook front-end should be built upon one of the SPA libraries, such as Angular, React or Vue.

The main difference between SPA and traditional ways of rendering websites is the way the browser retrieves data from the server. In a traditional approach, the entire website content is downloaded to render a page. To render another page, the entire content must be downloaded again. In SPA, the browser first downloads the entire application containing HTML, CSS and JavaScript that can render any page within the application. Data from the server required to render proper content is retrieved using asynchronous requests when needed. It makes the navigation faster – similar to a native application.

Authorisation and payment management

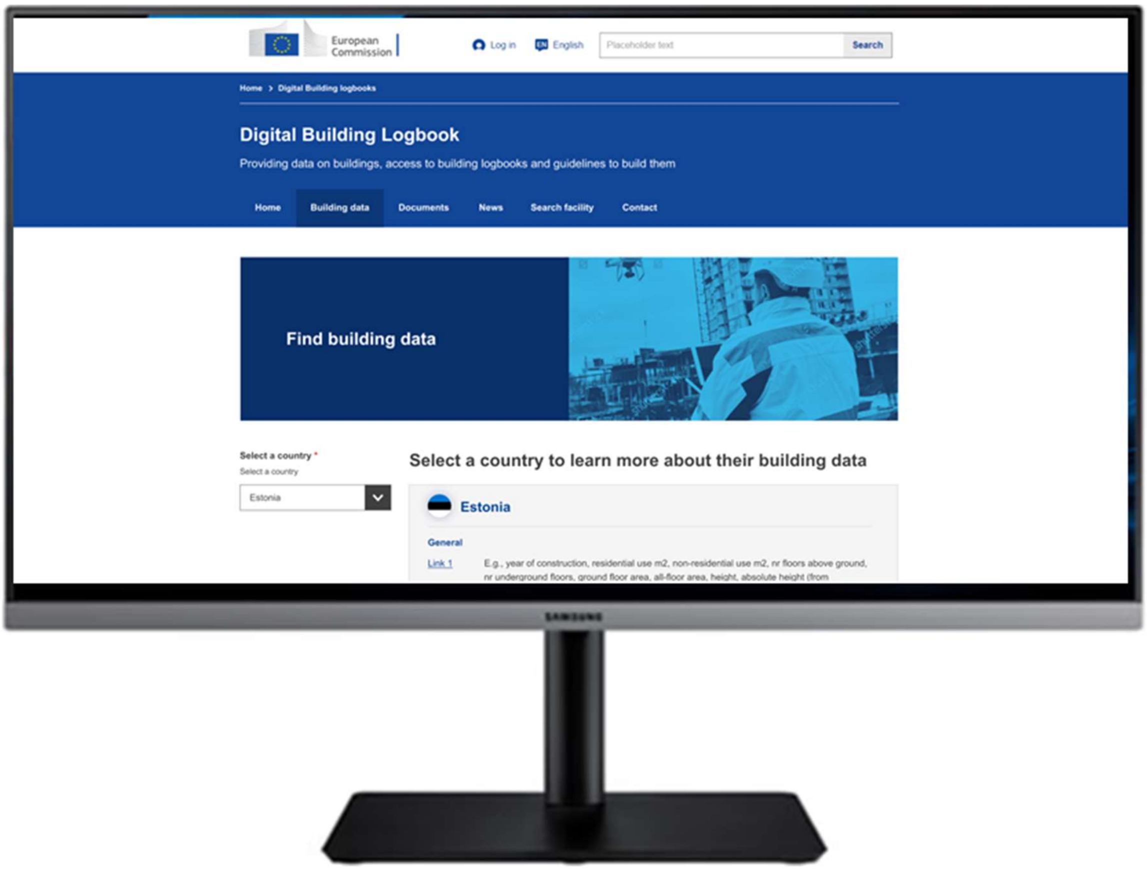
Authorisation and payment management is only necessary if certain data are accessible to authorised users only, or if users need to pay for certain data. Perhaps this is self-evident, but it is advised to use off-the-shelf solutions to ensure interoperability and to ensure that these critical features work flawlessly from the start.

Authorisation procedures, called IAM (Identity and Access Management), are integrated with cloud services that follow the SaaS (Software as a Service) model such as Datopian and LacesHub. The SaaS model, Datopian and LacesHub are further briefly discussed in Section 4.3. If the DBL platform does not use a cloud service in which IAM is integrated, dedicated IAM software can be used.[43] For online payments, online payment systems software exists.[44] Of course, if in the end data providers need to be paid in turn, an accounting system is needed to keep account of the payments received and sent.

EU-portal, short-term vision

In the short run, the EU portal is conceptualised as a repository of hyperlinks to national DBLs combined with a depository of guidelines and metadata. The current EU portal is thus a start page facilitating the search for national DBL portals. In this short run, the metadata need only describe the national datasets or DBL portals, such as the purpose and the type of data that is available. However, in the short run, the EU portal should not include definitions of data items because they are different across the national datasets and DBL portals.[45]

Figure 22: Example of how such EU-portal linking to national DBLs could look like in the short-term



[43] See e.g. https://www.g2.com/categories/identity-and-access-management-iam
[44] See e.g. https://www.g2.com/categories/payment-processing and many other online payment system software comparison sites.
[45] For a mock-up, see https://xd.adobe.com/view/09af13dc-c80c-482c-a4f5-1377fc65821b-5cfc/?fullscreen

EU-portal, long-term vision discussion

In the long run, the envisaged EU portal would allow users to search for buildings directly. This would mean that national portals are bypassed if a user accesses the EU portal. Many variants to implement this concept are conceivable in the long run.

For example, the EU portal could be limited to certain core data that are publicly accessible, such as perhaps the year of construction, various floor areas energy performance labels and other data that are proposed in the EU core ontology (see Linked Data Implementation in D2.2). For building data that Member States add to the EU core ontology, especially data that require authorisation or payment, Member State portals would still need to be visited in this concept.

In this concept, Member States need to make core data available to the European Commission and use a certain EU format, also proposed in the EU core ontology of Deliverable D2.2. Such data could be made available periodically through a download facility, in which case the EU downloads the data, does some minor translation or conversion of data, and stores the data centrally to make them accessible to users. An advantage of the central copy is that certain data are still accessible through the EU portal even if the national portal is down (and of course, data are still accessible through the national portal if the EU portal is down). A disadvantage is that data on the EU portal may be older than the source data.

Alternatively, the EU portal could send a query to the national portal after a user request, the national portal sends queries where necessary to national data sources and passes the result back to the EU portal. An advantage of decentralised data extraction is that always the most recent data from all the sources are presented to the user. A disadvantage is that national portals should adhere strictly to the EU format for the EU core data for this approach to work. Also, if a national portal is down, it would take some time before data are available again, depending on the backend disaster management discussed further below.

A possible solution to implement the long-term approach easily is if national portals generate documents for each building that contain only open and free data – no authorisation or payment necessary. The national portal could make the document available through a hyperlink, for example:

https://es/doc/catastro_navarra/pamplona-ayuntamiento/123456789.doc

Then, after a user of the EU portal has selected a building (by typing an address or zooming in on a map) that the EU portal identifies as a building with a certain identifier, the EU portal can just link the user to the relevant hyperlink from the national portal. Of course, more advanced solutions are possible. An advantage of the above approach is that national and EU portals could be developed independently of each other. A disadvantage is that this approach does not automatically ensure that data are harmonised and machine-processible if different Member States prepare such documents in slightly different formats. This is not a problem as long as one searches data for only one building but could be a problem if data on many buildings are searched, for example, all buildings with sufficient floor space.

Our final proposal is, therefore, to start developing an EU portal based on the short-term vision, for Member States to “copy” useful features of other DBL portals to improve their national portal, and in the meantime, the European Commission and Member States discuss a vision for the longer term.

EU-portal mock-up

The mock-up[46] is not a prototype EU portal but presents the look and feel of a possible EU portal. It is recommended to use the European Component Library (ECL) to design the landing page of the EU

[46] See https://xd.adobe.com/view/09af13dc-c80c-482c-a4f5-1377fc65821b-5cfc/?fullscreen

portal. The advantage is that the landing page can then be easily integrated into an existing EU website. The mock-up also adheres to website accessibility guidelines.[47]

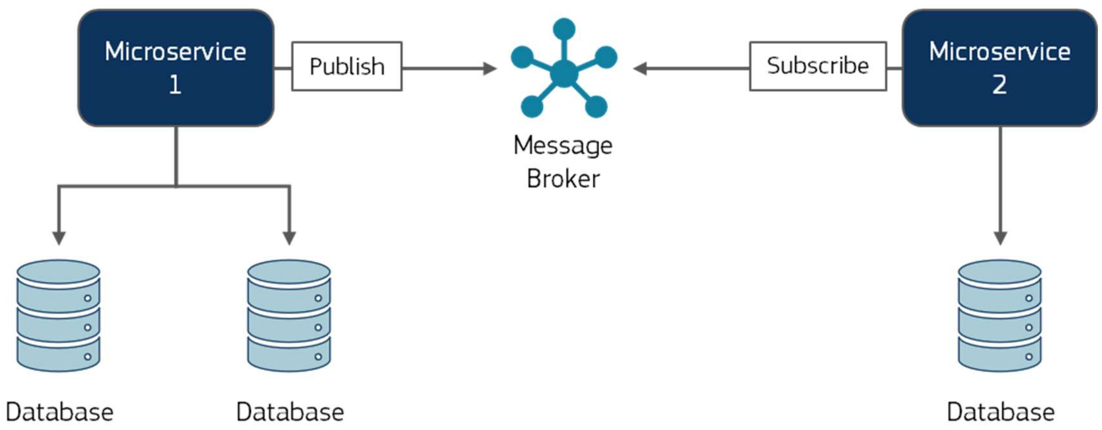
3.4.4 Back-end setup

A DBL is a complex system, and several principles should be followed to facilitate efficient development. The principle of microservices is to develop different parts independently of each other, with clearly defined inputs and outputs. The reason is that one does not need to know the whole system to improve one element, which makes the development and maintenance of the DBL much more manageable. A second principle is to use well-known programming languages. Experimenting with new languages is nice for small projects, but not for a DBL. A third principle is to develop, maintain and run different pieces of the DBL in separate hardware environments so that a crash in one environment only affects part of the DBL rather than the whole DBL. This principle is called containerisation and has the additional advantage of improving the speed of operations. The fourth principle is to develop a backup system according to the so-called SOLID principles. Lastly, it is important to manage the authorisation of the DBL developers, so that unauthorised users cannot accidentally (or malignantly) corrupt the DBL. All these principles are discussed in more detail in this subsection.

Microservices

Microservices are a widely used architectural pattern in the IT field. They organise the application into many independent, loosely coupled services that can be owned and managed by different teams. They allow to build and frequently extend large-scale applications by delivering small deployable pieces of the software. Using microservices architecture in the logbook gives an advantage, especially in terms of a long-term vision of the platform and potential improvements and new features. The alternative as a monolith application, where Logbook is one single deployable component, could result in a less scalable and bigger application whose codebase is more difficult to maintain.

Figure 23: Loose coupling of services in a microservice architecture



All the services should be loosely coupled, which means that they must communicate over a reliable message broker using a PubSub pattern (publish-subscribe). They should not access the same database, following a “database per service” pattern which means that one database can communicate with only one microservice.

[47] https://digital-strategy.ec.europa.eu/en/policies/web-accessibility/.

The main microservice that provides **REST API endpoints** should be the backbone of the logbook backend that communicates with the main databases. A reliable, fast and easy-to-use relational database, e.g., open-source PostgreSQL or proprietary Microsoft SQL Server which can serve as a primary database that stores information about all the basic entities of the logbook platform. Relational databases give an advantage of data integrity and support for complex queries in comparison with popular NoSQL databases like MongoDB. For knowledge graphs, RDF triple stores GraphDB or Neo4j are recommended.

In addition to the traditional server applications, downstream microservices can be built using serverless technologies to reduce costs to a minimum. Serverless technologies from popular cloud providers like Azure, Amazon or Google allow to building of fully functional REST APIs and support many different handlers like HTTP, different types of events, streams etc. A serverless functions service provides flexible pricing where the costs are incurred based on the resources' usage, without any necessity of maintaining virtual machines.

Programming languages

The main microservice and any required downstream services should be built upon a relatively popular language that offers a mature framework. The purpose of using a programming language that is broadly adopted in the industry is its potential for future work, community support and ease of hiring new developers. A framework providing all the necessary tools and components to build REST APIs, and popular design patterns, is necessary to reduce the time and costs of software development.

The programming language choice should depend on the architectural and design patterns that will be implemented in each microservice. It is possible to use a different language or framework for each service. In general, the programming language should be well established and relatively popular which ensures maturity of available tools and libraries as well as community support. Moreover, it is easier to hire developers to maintain and improve the platform as more popular languages have usually more active developers on the market. It is recommended to pick specific frameworks that provide ready-to-use components, and multiple tools that are implemented efficiently and safely. Building such components from scratch may lead to the risk of low performance and/or insecurity.

The following are examples of recommended technologies that could be used to create any of the logbook microservices:

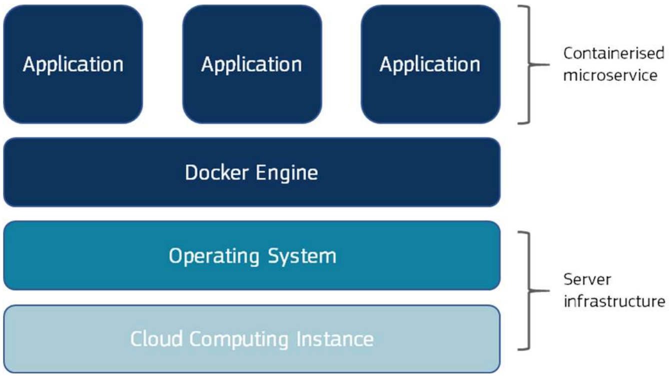
- **NestJS framework with TypeScript programming language.** NestJS is a well-established framework for building scalable server-side applications. It supports different programming paradigms like Object Oriented Programming or Functional Programming. The framework requires Node.JS version 16 as a runtime environment. It supports microservice architecture natively (meaning in a fully integrated way) and a vast array of transport layers like HTTP, TCP, MQTT and others.
- **Spring Boot framework with Java programming language.** Spring Boot is a well-established framework for building scalable server-side applications. Java language supports Object Oriented Programming of which the principles help in the development process. Spring Boot applications run on Java Virtual Machine.

Picking less popular languages or frameworks may result in unnecessary issues, particularly with regard to integrating the most recent technologies that have no proper support and libraries written for that particular language.

Containerisation

To improve the development and deployment processes, it is recommended to use containerisation to allow the building and running of software with all required dependencies in an isolated hardware environment. Docker Engine is a mature solution widely supported on every type of system such as Gnu/Linux, Windows, UNIX, and MacOS, and it is an industry-standard for containerisation nowadays.

Figure 24: Containerisation using Docker Engine



There are several advanced tools available for container orchestration and management – one of them, and currently the most popular is the open-source Kubernetes. It allows scaling containerised applications to handle huge workloads. Even if such tools are not used in the early stage of the project, providing containerisation for every microservice in the DBL will help with scaling out in the future and automating the deployment process.

Backup principles

To provide high availability and secure data, the database in the main geographical region should have a 'read replica' in a different region. A read replica is an exact copy of the database into which the entire data from the main database is continuously replicated. It can be used to improve the performance of heavy workloads and disaster recovery. Read replicas can be promoted to a main database to failover in the case of an incident.

To separate concerns and provide more performant applications and more efficient scaling, the main service can implement the CQRS (Command and Query Responsibility Segregation) design pattern. Since it is expected to have more frequent reads than writes, CQRS allows to easily optimise database reads as queries can be separated from commands.

REST endpoints should accept and respond to JSON format. The specification of the available endpoints, with required parameters, response structure, error codes and messages should be described using a standard specification, such as OpenAPI 3.0, that provides full documentation for developers and users.

SOLID principles are considered a good practice in software development, that allows ensuring an understandable and maintainable code. SOLID stands for:

- **S**ingle-responsibility Principle
- **O**pen-closed Principle
- **L**iskov Substitution Principle
- **I**nterface Segregation Principle
- **D**ependency Inversion Principle

The microservices in the logbook platform should follow the SOLID principles, especially in terms of long-term maintenance and improving functionality. The SOLID principles were introduced in 2000 by Robert Martin in his paper "Design Principles and Design Patterns"[48] and they have since been widely adopted in the IT industry.

Authorisation of platform developers

In terms of security, it is strongly recommended to use Authorisation as a Service from an external authentication provider. It provides a mature, safe and relatively easy-to-implement solution that works with client applications as well as with the back-end and can protect endpoints using a JWT token approach. Building authorisation from scratch requires extensive knowledge and is time-consuming, and it also leaves room for vulnerabilities. Examples widely used include AaaS: Microsoft Active Directory B2C, AWS Cognito, Okta, and Auth0.

Not all the endpoints in the logbook back-end require authentication. Most of them will be public like endpoints to retrieve data for the client application. Endpoints that require authorisation are mostly POST, PUT, PATCH and DELETE methods which trigger database writes (commands - using CQRS notions).

A roles and permissions system for the application should be implemented to introduce different levels of access for different roles in the organisation. For authorised users, access to non-public endpoints can be limited depending on the assigned role. Three basic roles are defined:

- **User:** who has basic permissions to interact with the platform resources, and can create, delete and update certain entities in the database.
- **Admin:** has all the user's privileges and additional unlimited access to every function in the admin panel.
- **Guest:** can interact with public endpoints only.

Accessing protected endpoints without a valid token will result in an authorization error 401 or 403 forbidden error in case the token is valid but has no permissions.

To increase overall infrastructure security, a proper networking structure should be configured with separate subnets that have access to the public internet or some without access that can communicate only over internal network addresses.

3.4.5 Maintenance of the platform and cloud infrastructure

Software and infrastructure maintenance requirements depend on various factors such as technology stack, level of complexity or cloud service provider. Several important requirements are defined to build, ship and run software that is robust, highly available and has no significant downtimes.

A DBL will always require maintenance and further development, for example, because of security updates, end-users use new software or data are provided in a new format. Development, testing and actual operation of the DBL should be done in different environments to avoid negative effects of testing in particular on the developers and end-users (for example, data getting lost). In case data gets lost, a plan should be in place to recover the data. Further consideration should be given to the choice between self-hosting and cloud services. Using a certain cloud service has the advantage that the provider takes care of the technical interoperability of that service but may come with a certain degree of vendor lock-in. Lastly, the use of virtual networks with their own access protocols is

[48] Robert C. Martin, 2000, Design Principles andDesign Patterns, available at: http://staff.cs.utu.fi/~jounsmed/doos_06/material/DesignPrinciplesAndPatterns.pdf

recommended to improve the security of the DBL system. All these maintenance aspects are discussed in the remainder of this subsection.

CI/CD principles

To improve building, testing and deploying code during the development and maintenance phases, automation tools should be used to introduce CI/CD which is considered a best practice in software engineering. Continuous Integration allows for the automation of software building and testing using specific cloud tools. Every change to the software is checked and validated by CI pipelines. Continuous Delivery works with the CI results being responsible for deploying the code to specific environments.

The logbook platform should have a minimum of three separate environments where infrastructure resources are isolated from each other:

- **Development:** for software developers who implement new features.
- **Test:** for software testers who check the features implemented in the development environment and validate whether they are ready to be deployed to production.
- **Production:** main environment available to end users. It has only stable code deployed which was tested and accepted in a test environment.

Deployment to test and production environments should be scheduled and restricted to only certain features that have been confirmed as operational and stable.

Many platforms support CI/CD, and some of them can be self-hosted. The most popular solutions on the market are Microsoft AzureDevOps, Gitlab, and AWS CodePipeline.

Disaster management

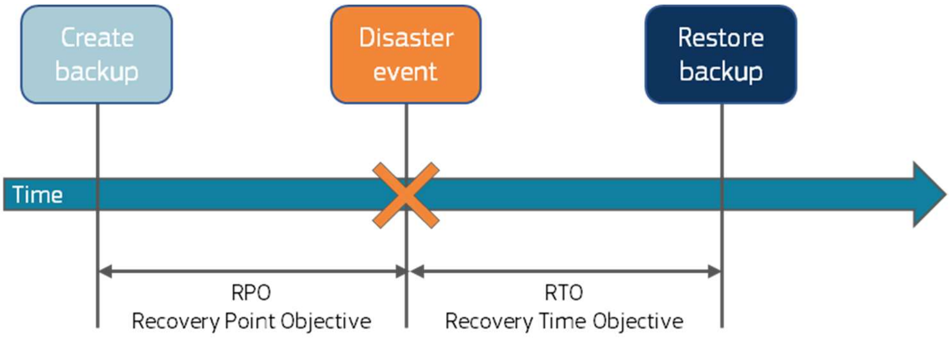
Every working piece of software tends to fail due to different reasons. A complete solution should have a plan to minimise business downtimes and render the service operational again, for instance, in the case of a disaster (i.e., natural disasters, such as floods that break the data centre where data is stored, and more common events of failure like software problems). Every component of the system should be included in this plan, particularly persistent storage like databases or file storage services.

A disaster recovery strategy is determined by two important notions:

- **RPO (Recovery Point Objective):** essentially the maximum time between backups – data before the RPO is accepted to be lost).
- **RTO (Recovery Time Objective):** defines how much time is needed to return the service to up and running.

Both values are given in the amount of time.

Figure 25: RPO and RTO timeline



The figure above illustrates a timeline. At a certain point in time, a backup of the persistent storage is created. In case of a disaster, all the data that has been created since that backup is lost, therefore, a smaller amount of data is lost with more frequent backups. The system is down from the disaster event up to the restore backup point and it is not operational. After restoration, data can be read and written again.

RPO and RTO are usually defined by stakeholders and strictly depend on business requirements. While RPO is defined as how frequently a backup of persistent storage is created. For example, creating a backup once a week could result in a 7-day data loss.

RTO defines how many resources are maintained to bring back the functionality within a given time range. There are many strategies to restore a working application, ranging from long downtimes when the team manually deploys the whole infrastructure again, to no downtimes when a fully developed emergency and failover infrastructure is maintained in case of a disaster event. There are also some hybrid solutions whereby a minimum emergency infrastructure is maintained which can serve at least during the most critical operations.

All the approaches generate different costs, i.e., smaller RTO and RPO result in higher costs. In the case of the logbook app there are a few remarks that may help to define both variables:

- **RPO**: Write operations on the Graph Database are not so frequent because unauthorized users are not allowed to overwrite data. Only the administration team will have the necessary permissions to extend graphs, therefore, the RPO does not need to be smaller than 48 hours.
- **RTO**: The logbook is most of the time not a critical service for stakeholders and users. Potential downtimes, even when they last over 24 hours, should not significantly affect any business or operation.

To ensure that backups are safe and available, it is recommended to store them in a minimum of two geographical regions, in case of a natural disaster in one of them. A multi-region approach is not required for web servers, but it can greatly improve the availability of the service. Geographical regions can be selected for every infrastructure component purchased from the most popular cloud vendors, e.g., Microsoft Azure Cloud offers multiple European regions in countries like Germany, Belgium, France, Denmark etc.

Cloud services versus self-hosting

Maintenance of cloud infrastructure usually differs for each component and can change depending on the service type. The main difference for logbook platform components is using self-hosted vs. fully managed solutions. In general, cloud providers have a shared responsibility model for the services they offer, and they are responsible for maintaining the hardware, networking and software required to run the services. They are usually not responsible for customer data, access management, operating systems, network and firewall configurations and security. Those responsibilities lie on the customer side, with some exceptions like certain fully managed services.

Self-hosted solutions give certain flexibility and in most cases are cheaper than fully managed ones. Customers can purchase cloud computing instances and deploy their own services on the operating system level. It is recommended that in the DBL, components like CKAN and GraphDB will be self-hosted, which means that maintainers of the platform will be responsible for installing and patching software, security and data management.

Fully managed solutions are offered for certain services only. Database services are a good example, e.g., Microsoft Azure SQL Database or Amazon AWS RDS provide a service where provisioning, updating, configuration and backups can be fully automated and carried out by the provider. Scaling of the database is easier and automated as well. thus, less overhead in this case means higher costs.

Containerisation described in the back-end section helps significantly with deployment and the maintenance process as it can be used to run both the logbook microservices and other parts of the infrastructure that will be self-hosted. Containerisation also makes the platform future-proof as it allows it to scale using container orchestration tools such as Kubernetes. Management of the Kubernetes cluster and configuration is usually a relatively difficult task that requires specific knowledge thus, overhead is expected to increase.

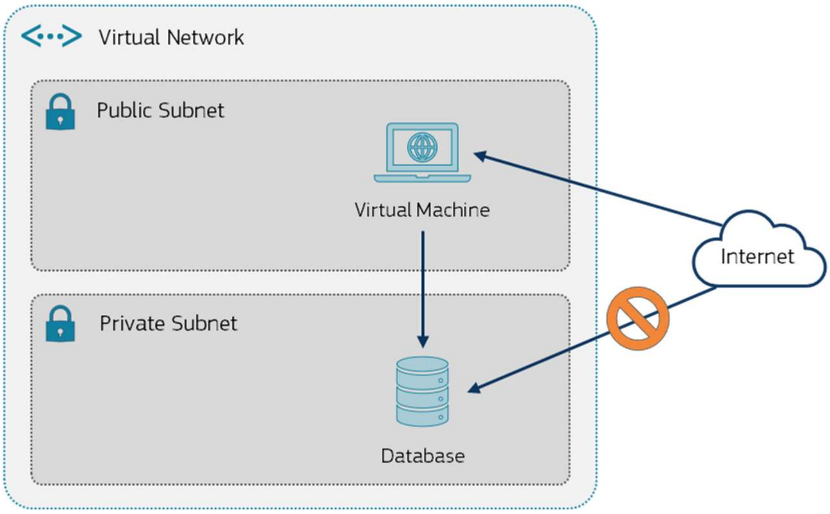
Virtual networks

To ensure proper isolation of the resources and improve the security of the introduced three-tier architecture, virtual networks in the cloud with all the features should be established and configured. Any of the main cloud service providers offer virtual network features that resemble a traditional network structure where the resources live (Azure Virtual Network, Amazon Virtual Private Cloud).

A virtual network can have strictly defined access rules, internal address ranges, access to the public internet, and communication with other virtual networks. It is important for cloud computing instances to place them in proper subnets within the network and define access rules for specific protocols like SSH, HTTP or any other that is required by the software deployed on it. It is an important aspect of security to properly configure firewalls and limit access to the data layer. Communication between the back-end application and the database can use public internet but this solution has a serious impact on database security as it is publicly accessible.

Within virtual networks, every resource belongs to a specific subnet. Every subnet can be exposed to public internet or it can be completely disconnected. If the subnet has no access to the internet, it can still communicate with other resources using internal addresses within the virtual network. Application tier resources and Data tier resources will be created in separate subnets, where the former is connected to the public internet and the latter is not. That allows for access to the database only from the applications running within the first subnet.

Figure 26: Example of public and private subnets in virtual networks



Depending on the service type, instances from the data layer can use a NAT gateway if needed to establish an outgoing connection.

Passwords and credentials for every application, service and architecture component should be stored in a dedicated service for secure credentials management. It can provide necessary non-public credentials to any service. Those include database passwords, connection strings, tokens, IDs and any configuration parameters that should not be stored as plain text in the virtual machine or control version. Example services are Azure Key Vault or Amazon Secrets Manager.

3.5 Data Management Plan

Besides all the data modelling aspects addressed in the previous sections, it is important to deal with the management of all data. A DBL is a common repository of relevant building data. It describes the building from various data sources provided by various data source owners, having specific characteristics per data type. Considering the multi-stakeholder involvement, **a DBL needs an associated Data Management Plan (DMP) that describes how DBL data over its life cycle is collected, stored and provided**. This document discusses elements of national DBLs, which are also relevant for an EU portal providing access to national DBL data. Considering a DBL is not a BIM or BMP, but stores higher level data, this DMP focuses on data relevant for example building permit procedures, assessing renovation subsidies, assessing the taxable value of the building, etc.

A DMP should cover all three stages in the data-sharing process:

1. **Collection of data**: Need to define minimum requirements per data type to be admitted to the (local) building logbook, to be able to fulfil the functionalities required (stage-gate). It also requires addressing how the data availability and quality are maintained (i.e. via validation and updating), and who is responsible for the data demand and supply side.
2. **Storage of data**: How and where the data is stored (accessibility, security, size, format, etc.).
3. **Provision of data**: Information on the data source for users, to always interpret and re-use the data in the proper way (through the use of metadata, i.e. data about data).

Across the data life-cycle, relevant data aspects include information on: ownership, context, purpose, scope, size, quality, structure, version, availability, rights/licensing, accessibility, confidentiality, privacy, security, intellectual property, time, financial aspects, etc. in the appropriate granularity.

A DMP should clarify in detail how the DBL data is managed both tactically (which standards and systems to be used) and operationally (given those standards and systems, how to communicate etc.). For the DBL data, a **data management template** needs to be defined. Such a template shows how to deal with several key aspects of data management, namely:

- Purpose of data;
- Availability of data;
- How to process and use data;
- Storage of data;
- Quality of data;
- The organisation of data;
- (Cyber)security & privacy of personal data.

For these aspects, a template (see Annex III) is provided as a guideline on how to consider each DBL data aspect, including a checklist that summarises ten main questions to consider (see Table 7 below). It does not give a full plan in the form of 'solutions' but it is merely a checklist of potential data issues to be dealt with.

Table 7: Checklist for a DBL Data Management Plan

Question	Checklist
1. Is the data findable and available?	Yes / No
2. Is the data open or restricted?	Open / Restricted
3. Is the data free or paid?	Free / Paid
4. Is restricted and/or paid data securely accessible (downloadable or direct access) via some agreed trust/pay framework?	Yes / No
5. Is the data in the correct format, or is translation needed?	Correct format / Requires translator
6. Is the data well-defined by available and sufficiently clear structural metadata?	Yes / No
7. Is the data in the right (common, agreed) semantics, or is conversion needed?	Correct semantics / Requires convertor
8. Is the data published under the correct license for reuse?	Yes / No
9. Is the data complete, sufficiently precise, accurate and up to date?	Yes / No
10. Is the data relevant to the intended process?	Yes / No

The following sub-sections discuss how the DBL Framework deals with the (1) collection of data, (2) storage of data, and (3) provision of data.

3.5.1 Collection of data

Cloud-based data

The **DBL Framework assumes a cloud-based (Internet/www-based) data infrastructure**. So, it is typically assumed that all data relevant for a DBL is findable (the F in the FAIR principle) in the cloud and can be communicated (exchanged or shared) via this cloud as the 'transport mechanism' when needed.

Still, there might be reasons not to organise data collection via the cloud. For one, the volume of data sources might be too large to be exchanged in the cloud. For example, high-frequency/resolution of time/space measurements can involve terabytes of raw data that need even physical hard disks to be transported because the exchange of terabytes would take too much time. In any case, a DBL would not contain but only link to such data sets or aggregate them first before inclusion. Another reason for exchanging data offline and not cloud-based is sometimes optimising security. Sometimes having no online access at all to the data is the best option for security. For DBLs, this could be the case for defence-related buildings or buildings that may contain safety risks (e.g., nuclear power plants).

To fulfil the required functionalities, the collection of data for admission into the DBL must meet certain minimum requirements per data type. This includes specifying the necessary data fields, formats, and standards for each type of data. Additionally, it is essential to address how data availability and quality will be maintained. This involves ensuring that the data is up-to-date, accurate, and reliable, and establishing procedures for data validation and verification.

The responsibility for data demand and supply should also be defined. This includes identifying the stakeholders involved in data management, such as data providers, data users, and data custodians. The data demand side refers to the users or stakeholders who require access to the data stored in the DBL. The data supply side refers to the individuals or organizations responsible for providing and maintaining the data within the DBL.

By addressing these aspects, the DBL can ensure that the collected data meets the minimum requirements, maintains data availability and quality, and establishes clear responsibilities for data demand and supply. This will enable the DBL to effectively fulfil the required functionalities and provide valuable data for various purposes.

Open versus restricted and free versus paid

A **DBL can include both open and restricted data**. In contrast to restricted data, public data does not require a trust framework involving identification (and 'registration') nor authentication and authorisation (IAA). Independently of this, both open and restricted data can be free (of charge) or paid.

Table 8: Hypothetical example of data types for combined categories

	Free	Paid
Open	Year of construction	Topographic data
Restricted	Owner info	Taxation values

To decide whether data can be opened, an assessment must be made as to whether reuse poses risks to fundamental values and privacy. Typically, open data is protected from the rights of third parties by being released with a Creative Commons (CC) - 0 statement[49]. If it is considered important that users acknowledge the source, open data should be released with a CC-BY statement[50]. In most cases, 'open data' is explicitly associated with the 'FAIR' principle for data: easy to find, accessible, interoperable, using open standards and reusable by specifying agreed (common, federated) metadata. In the case of DBL, **the FAIR principle is already adhered to via the proposed DBL framework**. On the other hand, restricted data (that are privacy or commercially sensitive) should be identified (like ownership of a building hence the building data on it) and be made available via authorisation procedures and with Creative Commons licenses that protect commercial rights, as discussed in chapter 6 of the technical guidelines.

It is recommended to provide meta-data for any data aspect/property in the DBL whether the data is:

- open or restricted; and restricted, how it can be securely accessed; and
- free or paid; and if paid, how it can be securely paid for.

Volume, variety and velocity

For any data source, it is important to know its **volume, variety, and velocity** (time window of relevance). In the case of DBL, volume, variety and velocity refer to:

Volume

For DBL, the volume of a data source is especially important when semantically defined data is exchanged (instead of shared). However, generally, such large sets of data (like building sensor measurements) and building geometry data (like point cloud representations resulting from 3D scanning) should not be included in the DBL directly (only linked by it).

Variety

The variety for the semantic/geometric data is controlled by the proposed DBL ontology and dictionary. However, linked documents can be of any format or any semantic structure and hence can be varied.

Velocity

Velocity is an important issue for DBL since there needs to be a match between requirements and provisions. This velocity determines the required update strategy. Legal data requires a high velocity to avoid fraud, as do measurements for real-time control of the building. On the other extreme, static data like the construction year of a building is not likely to change at all. However, most of the aspects/properties are somewhere in between having their typical update schedule or triggering events. For each data aspect/property, an appropriate update frequency should be determined.

Push versus Pull

The initiative for the exchange or sharing of data can either be a push (i.e. information is "pushed" from the data source to the other system, i.e. when the data source is generated or updated it is also pushed to the DBL) or a pull (i.e. information is "pulled" by the system from the data source, i.e. the DBL retrieves the data from the source). For a DBL, stakeholder feedback showed a **preference for the pull approach** and data being extracted from the source whenever the end user requests them.

In either case, it is the **responsibility of the DBL user to check for updates** and the data provider should make sure the latest updates are available at the publication point. **Metadata should**

[49] For more information, see: https://creativecommons.org/publicdomain/zero/1.0/deed.
[50] For more information, see: https://creativecommons.org/licenses/by/2.0/.

indicate in all cases the update frequency, which is either periodical or dependent on 'moments of change'. Appropriate 'moments of change' could be:

- The initial set-up of a DBL, so it can be used;
- Change in the building ownership;
- At the 'reception' of the building after construction or major renovation;
- When changes are made to the building energy systems (by the owner);
- When there are changes in building regulations (by the authorities);
- After inspections;
- Validity of technical assessment reports expires;
- After an energy performance evaluation;
- Other events such as a fire, earthquake, etc.

Such moments of change are the typical time that a DBL user would check for updates (when data is exchanged) or the other way round, a data provider would publish an update of the data (when data sharing).

3.5.2 Storing of data (when exchanged by source)

The **storing of data is more of an internal issue that is already 'internally' resolved** as long as the DBL framework agreements (concerning syntax and semantics) are respected when collecting and publishing data. To provide an example, when a relational Database Management System (DBMS) is used for data storage (e.g., open-source SQLite or PostgreSQL) then **a data provider needs to make sure that the data is translated/converted correctly** without data loss. Furthermore, the availability, security and quality of data should be maintained.

In any case, a DBL manager, the person with overall responsibility for the setup and maintenance of a national or EU-level DBL, should implement the DBL in semantic/relational/etc. servers including the needed data translators and converters.

Minimum data requirements

The current proposed DBL framework is flexible in the sense that it does not provide many constraints regarding the data: **only the IDs of buildings, building units and parcels are obligatory minimum data requirements**. All other properties (attributes and relations) are optional and based on data availability. The DBL can be further constrained by Member States and/or data providers, owners, etc. by making further properties required. Making properties obligatory typically requires laws and regulations. As a consequence, data verification of DBL data against the DBL ontology is currently limited.

In case this becomes relevant, a so-called 'Closed World Assumption (CWA)' would be helpful as defined in SHACL or related to a buildingSMART Information Delivery Specification (IDS).

A CWA means that you assume that all data is available. If data is not there ('unknown') it is assumed to be false. The latter information can be used to derive or check more data. [51] CWA is the opposite of the Open World Assumption (OWA) where unknown data stays unknown (there might be someone out there saying something about it). To provide an example: If an obligatory property value is missing in OWA there is no verification error. In CWA you would get an error since that property value is not part of the 'closed world' that you know of.

[51] For more information, see: https://www.w3.org/wiki/ClosedWorldAssumption.

Data quality

Regarding the quality of data, the FAIRness of DBL data can be formally checked:

- The DBL can check the 'well-formedness' (whether the data conforms to the required exchange format) of the data (is it in the right format); alternatively, whether the shared data can be accessed via the required access method;
- The DBL can check the semantics of the data (does the data conform to the ontology, given also the human semantics provided in the dictionary). This includes checking for obligatory elements such as attributes, relations, parts, etc.

As outlined above, to facilitate this check, a closed-world assumption (CWA) could help since otherwise missing data can never be checked.

Beyond FAIRness and the validation of data against the DBL ontology/dictionary, data should satisfy further quality aspects.

One key aspect here is **accuracy** and whether the data describes the real-world situation accurately. For national registries, existing processes aim to ensure this accuracy of data, but for other sources like local BIM models, this is much more difficult. The reason is that it is difficult to know whether the building was constructed according to its design as reflected in its BIM model. Similarly, questions about whether one can trust an assigned energy or circularity label value can arise. Besides accuracy, there is also the aspect of **precision** which indicates something about the level of detail that is addressed. For example, when a building height is indicated as 12 meters, is this the precise height or has the value been rounded up or down? Finally, a third aspect is **completeness**. As for measurements, data might simply not be available for certain years or aspects (caused by many reasons). What to decide i.e. if the embedded building units for some floors in a building are set as 'unknown'? A DMP could describe that feedback should be given to the data provider in some form.

3.5.3 Provision of data (publishing)

Similar to the collection of data, **DBL data can be published in multiple ways**. For example, via download of data sets or via providing 'RESTful HTTP(S)-based APIs' embedding queries in some kind of query language. This could be SPARQL as already promoted by the European Commission[52]. This is not only the case for the actual data sets but also for the semantic data models such as the DBL Ontology or DBL Dictionary (in general these are all 'data sets').

Especially for published data, as further discussed in Chapter 5, one has to consider:

- **The context and purpose of the data**: What may the data be used for? For example, the taxation value of a building could be used for taxes, to determine maximum rent prices, or by private parties to negotiate a sales price. The data provider may want to put restrictions on the purpose for which data can be used, by limiting authorisation to certain users. Also, a system needs to be implemented for paid data. Both authorisation and pay procedures should take the information about which data aspects are restricted and/or paid from the semantic model. Our recommendation is to include for every data aspect a field that indicates whether only authorised users may view the data and whether users need to pay
- **The privacy of personal data**: Is it necessary to anonymise the data first before publishing? It is recommended that Member States determine what type of data is privacy sensitive. The General Data Protection Regulation (GDPR) requires the data subject (whom the data is about) to be informed about the storage of data and its justification.

4. Economic implementation

In this chapter, we discuss what is needed for the implementation of a national digital building logbook in terms of people, funds and technical requirements. The information from this chapter is based mostly on interviews and the feedback survey.

Regarding the people needed, some needs are implicit in the other chapters. Obviously, a legal expert is needed for a legal impact assessment (Section 2.7 and in more detail in Chapter 5). Data modellers are needed for the data modelling discussed in Section 3.3. Platform engineers are needed to develop a platform as discussed in Section 3.4. A knowledge manager is needed for the development and implementation of a data management plan (Section 3.5). Section 4.1 presents an overview of these and other people needs.

Regarding financial costs, Sections 2.6 through 2.8 discussed the more costly steps conceptually. Section 4.2 provides a breakdown and estimates of the range of financial costs involved. Regarding technical requirements, approaches and solutions were discussed in Section 3.4. Sections 4.3 and 4.4 discuss practical aspects of hardware needs, software choices and associated costs in more detail.

4.1 People

While most of the team should be software developers, other skills are needed as well:

- A **project manager**, ideally with a background in the construction sector and with experience in software development in both the public sector and the private sector, to negotiate the project needs and manage the team, planning, budgets and deliverables
- A **software architecture specialist**, to communicate about DBL features, design solutions and make decisions about the best solution
- A **knowledge manager**, to take stock of which information the team needs and ensures they can access that information
- A **legal expert** in privacy law, to act as a data protection officer in the meaning of the GDPR
- A **platform engineer**, for basic platform configuration and platform availability, data backup and incident management
- An **IT management and support team**, to ensure that the hardware and operation system work, to manage software security updates and to provide general IT support
- A **team of data modellers**, who know how to develop an ontology that is behind the data in the DBL, for example, building on the core ontology discussed earlier in Section 3.2
- A **team of software engineers**, to develop and maintain database software, DBL features, interfaces to communicate with other applications (APIs) and the DBL portal
- A **team of data managers**, who ensure compliance of building data with data norms through automatic and manual checks, communicate about data norms and errors in data, and reformat legacy data

No architects or engineers have been included in the people needs. In the opinion of the Consortium, they are not needed for the day-to-day development of a DBL, as opposed to the development of a BIM. The reason is precisely that DBL content should be accessible and understandable for laymen. However, it would be good to include architects and engineers in a steering committee. And the team will need to co-operate with architects and engineers because information from BIM is useful input for the DBL, and way to share data between BIM and DBL will need to be discussed.

For the economic aspects of implementing a DBL, it is useful to distinguish between **a development phase** and an **operational phase**. In reality, these phases are not entirely separate, especially because an incremental development of a national DBL is recommended.

In the development phase, a larger team will be needed than in the operational phase. In the development phase, a team of 40 people (including external developers) may be needed, as has been the case in Estonia. The core team to keep the DBL operational and do some occasional development may consist of up to 20 people, with a smaller permanent core team and the others hired as needed for ongoing development of the DBL, as is the case in Estonia.

Most of the above skills are needed early in the development phase already. For example, the legal expert can help decide which data are privacy sensitive, and what is the appropriate procedure to collect and process such data if at all, with the appropriate justification.

4.2 Financial

The cost of developing a DBL depends on the ambition level. It also depends on the degree to which off-the-shelf solutions are used. These impact both the monetary cost and the duration of the (initial) development. As discussed in Chapter 2 on the implementation roadmap, it is a good practice to start small to achieve the first tangible results in a reasonable time.

According to interviews with building logbook developers from Bulgaria and Estonia, the cost of developing a DBL was invariably answered to be a few million euros or even a more specific amount like two million euros.[53] However, it should be recognised that a DBL requires continued maintenance and small-scale development. Maintenance is not only needed to update building information but also to keep up with security updates, performance optimisation, dealing with error reports etc.

In a feedback survey, stakeholders from across the EU ranked different types of costs from high to low, provided cost estimates in terms of duration from start to finish, man-days of work and out-of-pocket expenses, and reported which costs would also have been incurred without a DBL initiative. The result are in line with the expert views of the Consortium. The answers to the first question indicate that collecting (legacy) data on existing buildings is the most costly, and licensing costs and the development of a dictionary are the least costly.

Table 9: Final ranking of types of costs based on the total score (N=20 respondents from BE x5, CZ x 2, DE x 2, ES x 2, NL x 2, PT x 2, AT, IT, LU, RO, SE)

Score	Type of cost
82%	Collecting data on existing buildings
69%	Verification of data
67%	Development of an online platform
65%	Developing a semantic model
59%	Agreement about data norms with stakeholders
42%	Costs for licensed data
41%	Developing a dictionary

[53] Two million euros is also the cost of each of three Horizon projects to develop prototype digital building logbooks.

In terms of **lead time**, which is the duration between the initiation and the completion of the initial development, up to 6 or 7 years should be expected. The most time-intensive activities are the collection and verification of data, which both take up to 5 years but can take place in parallel. The development of an online platform is reported to take one to two years. The other activities are reported to take 6 to 12 months each. Purchasing licensed data need not take more than two days.

In terms of **man-days**, the sum of minimum reported values is 350, and the sum of maximum reported values is 8,000. The average reported cost in man-days is about 3,000. However, in reality, there are some trade-offs. For example, it is not necessary to invest in agreements about data norms with stakeholders. However, this can save time in developing a dictionary. Developing a semantic model can save time in verifying data. Thus, 1,800 man-days is a more accurate estimate of developing a fully functional national DBL. At labour costs of EUR 500 per man-day, 1,800 man-days correspond to a **labour cost equivalent to EUR 9 million**. It should be noted, that in reality the costs of collecting and verifying data depends on what data is already available, for how many buildings, whether the data is digitalised, and how many buildings share the same design.

For **out-of-pocket costs**, the sum of minimum reported values is EUR 52,000. However, especially the reported cost of collecting data on existing buildings varies a lot. If the ambition level of the DBL is limited to making digital data from the national cadastre public, the cadastre might only charge a few thousand euros to prepare the dataset. If one outsources the manual processing of legacy files that still need to be digitalised, this could cost EUR 2 million per 100,000 unique buildings – assuming that a legacy file can be processed in 15 minutes. Of course, some savings are possible for terraced houses that are almost identical and it is certainly advisable to identify terraced houses between which only a few characteristics differ. The total costs of actually collecting legacy data depend very much on the number of buildings, the number of sources from which various building data need to be collected and the percentage of digital files. It also depends on whether technical drawings are already in the municipal archives or need to be purchased from the architect. The necessity for this depends on the ambition level of the national DBL. Thus, many factors determine the cost of collecting legacy data. For a country with millions of buildings, the costs could thus run into hundreds of millions of euros. It is very much for this reason that we would advise starting the DBL with buildings for which digital data are already available, which tend to be newer buildings. Leaving out the cost of collecting legacy data on existing buildings, the sum of out-of-pocket costs varies between 50,000 and EUR 150,000 with an **average of EUR 100,000**.

In addition, a **legal impact assessment** (see Section 2.7 and Section 5.2) could cost **EUR 80,000**. It was the cost of a legal impact assessment of the Estonian construction e-portal.[54]

However, it should be kept in mind that in multiple interviews it was also reported that the development of a first prototype DBL, with data on one or two buildings, need not take more than 2 years and need not cost more than EUR 2 million.

Table 10: Can you give an estimation in terms of lead time in months (the time between the initiation and completion of initial development), man-days, or outsourced costs (expert estimate based on responses)?

	Lead time (months)	Man-days	Cost in EUR
Collecting data on existing buildings	up to 60 (N=4)	500 – 1,200 (N=3)	Up to tens or hundreds of millions – see text (N=3)
Verification of data	up to 60 (N=4)	800 – 2,000 (N=4)	20,000 – 50,000 (N=2)

[54] See: https://www.mercell.com/da-dk/udbud/112620050/hanked-ehitise-elukaare-oigusruumi-digitaliseerimiseks-kohandamine-80-000-eur---muudatus-udbud.aspx.

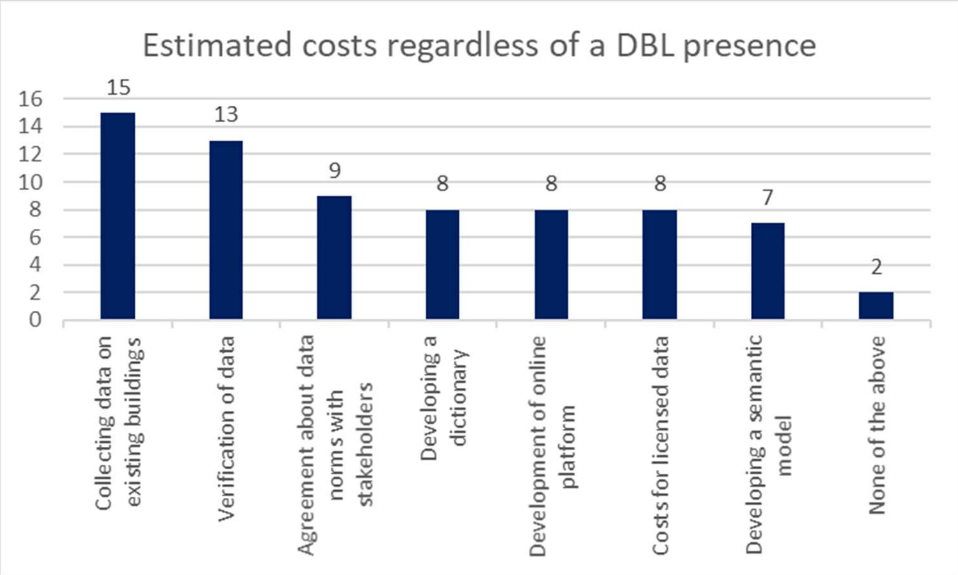
	Lead time (months)	Man-days	Cost in EUR
Development of an online platform	12 –24 (N=4)	800 – 2,000 (N=3)	Up to 50,000 (N=1)
Developing a semantic model	6 –12 (N=4)	400 – 1,000 (N=3)	5,000 – 10,000 (N=2)
Agreement about data norms with stakeholders	12 (N=3)	400 - 800 (N=4)	25,000 – 50,000 (N=2)
Developing a dictionary	6 –12 (N=3)	400 - 800 (N=2)	0 (N=1)
Costs for licensed data	0.5 (N=1)	2 (N=1)	Up to a few thousand (N=1)
Total	72 (N=1)	1,800 (N=1)	100,000 excl. for collecting building data (N=2)

Although the development of a national DBL is costly, it should be noted that 40% or more of the respondents consider **most costs as unavoidable**. Without a national DBL, engineering companies would still develop their own DBLs to improve the design and construction processes and would submit data for new buildings in digitised form. A national DBL would harmonise industry trends.

Digitisation of building data is a key trend in the construction sector but is also necessary to improve government services. For example, with digitised building data, the quality of the building stock can be better assessed based on the year of construction, the date of the latest major renovation, a logbook of inspection approvals and/or other data in the DBL. Also, digitised data make building permit procedures (including for renovations and demolitions) more efficient. Digitised building data on energy performance, solar panels and other installations help monitor progress towards climate goals. Indeed, respondents of the feedback survey reported the collection and verification of building data most often to be unavoidable.

To ensure the efficient processing of data updates, data norms need to be agreed upon with stakeholders. When digitising building data, data from multiple buildings can efficiently be extracted. Then coherence of definitions and thus a dictionary (see also Section 3.2) becomes important to compare data between buildings. An online platform improves the accessibility of data (Section 3.3) and a semantic model improves the interoperability of shared data (also Section 3.2).

Figure 27: Which costs would have been incurred or do you expect to incur even in the absence of developing a DBL as a platform for collecting and sharing data? (N=20)



The table below summarises the potential financial costs and lead time.

Table 11: Summary of financial costs and planning

Activity\year	1	2	3	4	5	6	7
	€ 2 mln.		Wide range, average € 7 mln.				
Data norm discussions							
Prototype DBL							
DBL legislation							
Development of an online platform							
Collecting and verifying data							

4.3 Technical - hardware

A third implementation aspect that also has economic implications concerns technical aspects. In general, we would advise using off-the-shelf solutions, preferably open-source, to address the technical interoperability requirements of a DBL portal as discussed in Section 3.4. Hosting a DBL requires software to manage data requests, but also hardware is needed, whether the DBL is self-hosted on one's server or a cloud service is used.

We first discuss the amount of data storage needed, and then the types of cloud storage services that are offered: always with infrastructure but depending on the choices also with other services. Lastly, we discuss technical choices for sharing and maintaining a semantic core model to which users can add their own components.

In terms of **hardware**, a server or a cloud computing instance is needed. A server or cloud computing instance processes data updates and handles queries on building data. Cloud infrastructure is in addition helpful to backup data.

Ultimately, if a DBL would merely link to source data then hardware would be needed mostly at the source data and hardly for the DBL. An estimate for the necessary hardware is more meaningful for the variant where a central copy of the data is made. If data are harmonised, this reduces **data storage needs**. The reason is that then much of the building data can be reduced to simple combinations of numbers, letters and/or dates. If the data is in addition stored in CSV format and compressed (or actually, in a CKAN database as discussed later), the data storage could be reduced to one or two kilobytes per building. **Per million office and residential buildings, only 2 gigabytes would then be needed**. For buildings where data are difficult to standardise such as stadiums and hospitals, more storage space per building is necessary, but there are much fewer of them. Of course, much more data storage space is needed if some fields include long descriptions of each building, but for a first working prototype, this is not recommended. Also, much more data storage space is needed for (signed) documents which can take 100 or 200 kilobytes of data and cannot be compressed much further. Thus per million buildings one type of document could add up to 200 gigabytes.

Another feature that could take up tens of gigabytes is maps. Such maps are useful to view, zoom in and then select a building. Some portals[55] also offer the possibility to download map segments, which could be useful for users of a DBL portal. Such maps are nowadays often created with third-party tools and loaded to the cloud, in which case no storage space is needed on the own server. To use third-party maps as a data source, one needs to work with map protocols such as WMS, WMTS, WFS etc.

It is also worth noting that the proposed DBL framework internalises historical values, by just adding any change to the building to existing data instead of overwriting them. Thus, there is no need to store monthly backups over a long period, although it is still advisable to store one additional backup in case the core data gets corrupted (due to a hack, software failure, hardware crash, or human error).

Multiple **cloud storage alternatives** exist, which fall in the categories of self-hosted solutions and fully managed services. Such types of service models are known as IaaS, SaaS and PaaS (Infrastructure as a Service, Software as a Service, and Platform as a Service).[56] For example, Microsoft Azure Cloud offers all three types of service models.[57] With IaaS services, the DBL manager hires one or more so-called cloud computing instances on which software can be installed, whether it is software built on its own or software licensed by a third party. With **IaaS** services, one can for example install off-the-shelf software like CKAN and assume own responsibility to manage resources such as RAM, CPU, SSD and the operating system. The IaaS option is discussed further below under Cloud storage option 1 and further below on semantic model hosting tools. With **SaaS**, the customer is fully responsible for the software configurations only. Examples of SaaS are "Active Directory" in Azure, Datopian (discussed under Cloud storage option 2) GitHub and LacesHub are also considered as examples of SaaS as discussed further below on semantic model hosting tools. With **PaaS**, the vendor shares responsibility for Identity and Access Management (IAM) and storage. Examples of PaaS are "Azure Functions" and "Event Hubs" in Azure. Because PaaS is essentially SaaS without IAM, this option is not separately discussed.

Cloud storage option 1: IaaS

The Infrastructure as a Service (IaaS) option is illustrated with Azure Cloud[58], but other providers such as Amazon offer similar solutions. In IaaS, providers offer cloud computing instances, which are called

[55] For example, https://mapy.geoportal.gov.pl/imap/Imgp_2.html.
[56] For example https://www.nigelfrank.com/insights/iaas-vs-saas-vs-paas-a-guide-to-azure-cloud-service-types.
[57] See: https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-saas/.
[58] See https://azure.microsoft.com/en-us/pricing/vm-selector/

Virtual Machines (VMs) in Azure and for example Elastic Compute Cloud (EC2) in Amazon. It is recommended to run DBL components on different VMs (following the Azure example). For example, a CKAN database has six components. With CKAN, it is recommended to run the PostgresDB component as the main database that stores data on a separate VM. Another VM could run a search engine like Solr. Solr takes every single entry or update in Postgres and creates an indexed entry to speed up queries. Optionally one can use Azure PostgreSQL and a VM dedicated to that, which costs extra (not in the table below) but saves time configuring the cloud computing instance. To further increase speed when installing CKAN in the IaaS option, it is recommended to work with six VMs, one for each component of CKAN:

1. Nginx
2. Main CKAN app
3. datapusher tool
4. PostgresDB
5. Redis cache
6. Solr search engine

In the table below, it should be noted that 730 hours corresponds to one month of hours. It is also noted that a price reduction is possible if one combines the cloud solution with own SQL servers with an Azure licence.

Table 12: Azure Cloud cost estimate

<table>
<tr><th>Type</th><th>SKU</th><th>N</th><th>Pay as you go per month</th><th>One-year contract per month</th><th>Three-year contract per month</th></tr>
<tr><td>Virtual machine</td><td>A2 v2 (2 Cores, 4 GB RAM, 730 hours, Linux</td><td>6</td><td>$381.06</td><td>$284.92</td><td>$204.06</td></tr>
<tr><td>Virtual machine</td><td>A2 v2 (2 Cores, 4 GB RAM, 730 hours, Linux; 1 managed disk – E6</td><td>1</td><td>$63.51</td><td>$47.49</td><td>$34.01</td></tr>
<tr><td>Storage</td><td>Block blob storage General Purpose V2, Hierarchical Namespace, GZRS Redundancy, Hot Access Tier, 500 GB Capacity</td><td>1</td><td colspan="3">$48.25</td></tr>
<tr><td colspan="3">Total</td><td>$497.62</td><td>$385.45</td><td>$291.12</td></tr>
</table>

Cloud storage option 2: Datopian[59]

Datopian specialises in cloud-based hosting for CKAN and can be considered as a SaaS (software as a service). A basic setup costs $399 per month. However, the basic setup does not have solutions to ensure the interoperability of reading data from multiple sources (in different formats). To include this, a standard Datopian setup will be needed which costs $899 per month. The standard Datopian setup is comparable with the above Azure Cloud setup in terms of functionality, but Datopian offers more personal assistance (at a price).

4.4 Technical - software

For sharing building data and the semantic model describing the structure and definitions of the building data, supporting software is needed.

[59] See https://www.datopian.com/pricing

In terms of **software**, different types of software are needed:

1. A tool to build a semantic model will save a lot of time
2. Software to host the semantic model on the internet
3. Software to store and search building data
4. API integration software to collect and/or share data

In all cases, we can provide examples of several software tools, without suggesting which company or trademark should be preferred over others.

For sharing building data, there is a large choice between website and app development tools. Front-end aspects are discussed in Section 3.4. Below, we suggest some tools for semantic modelling and for hosting semantic models.

Semantic modelling tools

To develop a semantic model, we suggest the following tools:

a. Protégé ontology developer, which is a free and open-source developer[60]

b. CKAN, which is a free and open-source data management system[61]

Semantic model hosting tools

a. (Microsoft) Azure blob storage,[62]which costs about 1.3 to 2.5 cents per gigabyte per month

b. GitHub[63], which is open source and freely downloadable and installable

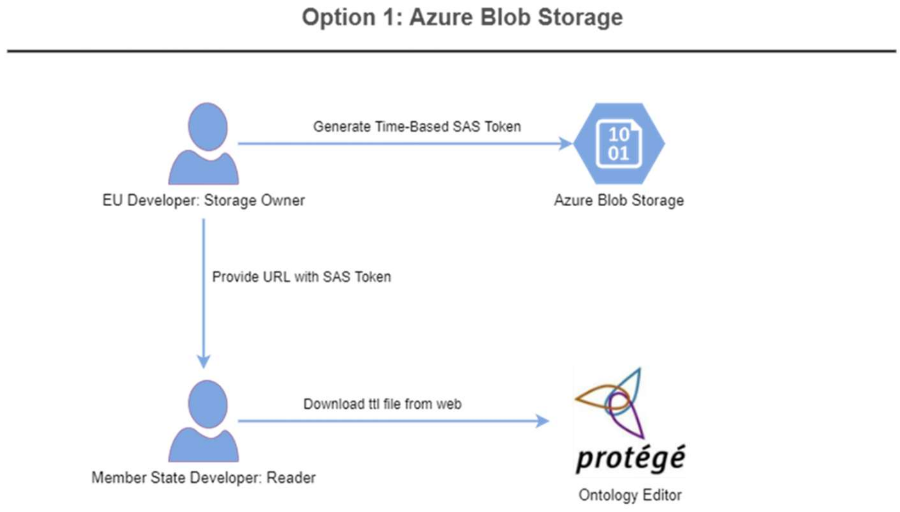
c. LacesHub[64], with a cost of around EUR 100 per user per month[65]

Azure blob storage can be considered Infrastructure as a Service (IaaS). GitHub and LacesHub are generally considered to be Software as a Service (Saas). The concepts of IaaS and SaaS were discussed in the previous section.

With **Azure blob storage**, the developer of the ontology provides a URL with a secure access token from which the user of the ontology can download it. Any communication about changes or extensions of the ontology would take place out-of-system. The figure below illustrates this option for a situation where Member States use the EU ontology.

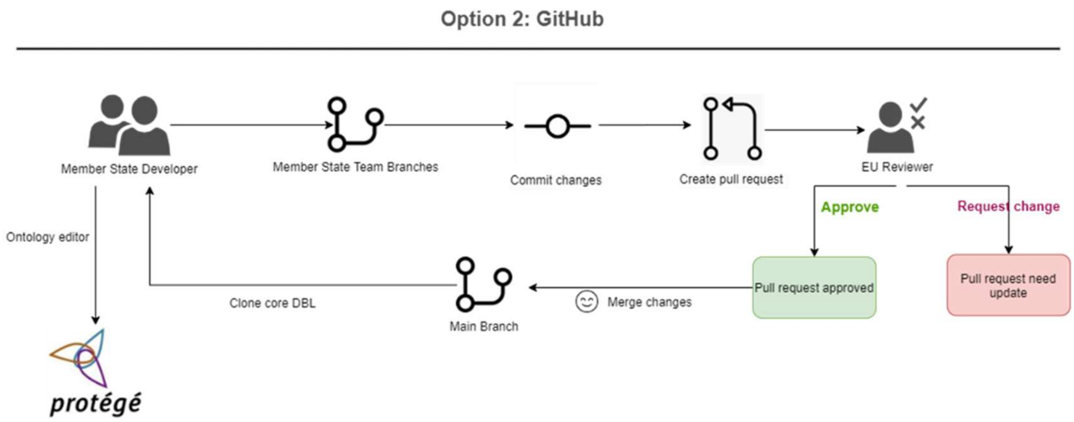
[60] See https://protege.stanford.edu/.
[61] See https://ckan.org/.
[62] See https://azure.microsoft.com/en-us/pricing/details/storage/blobs/.
[63] See https://github.com/meta-toolkit/meta-toolkit.org/blob/master/setup-guide.md.
[64] See https://laceshub.com/laces-suite/.
[65] See https://laceshub.com/pricing/.

Figure 28: Example hosting option of Azure Blob Storage



With **GitHub**, users can propose changes that the developer can approve or where the developer can request the user to revise or undo the change.

Figure 29: Example hosting option of GitHub



A service such as Laces Hub makes the ontology (or any other data) freely available to other users, from many operating systems, at a price of around EUR 100 per month.

Software to store and search data

For storing and searching data, graph databases are best suited. All graph databases come with very flexible query languages, allowing you to create complex queries and execute them fast.[66] Wikipedia lists many alternatives. We suggest one open-source alternative and one graph database that seems commonly used in the construction sector

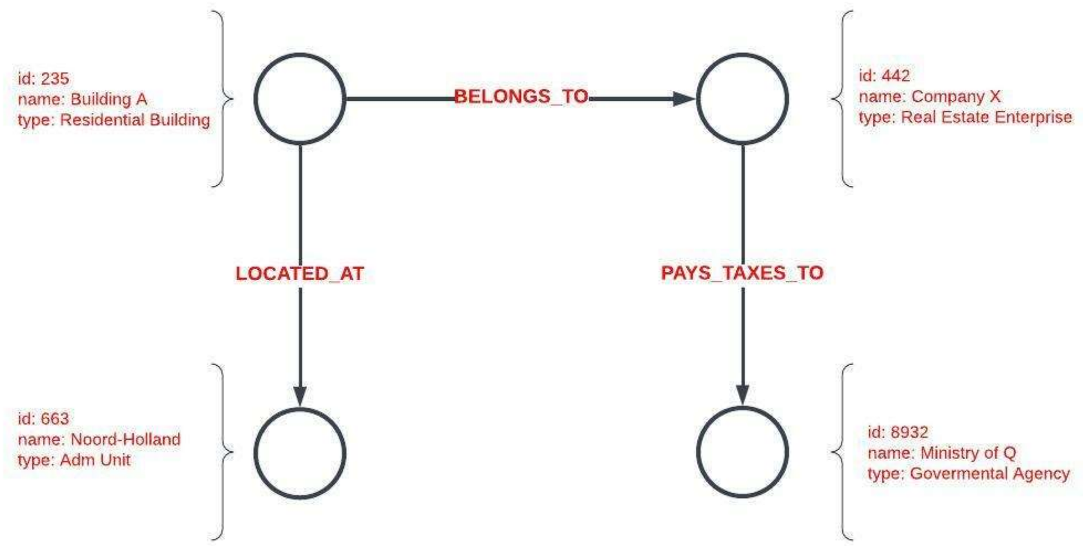
a. Neo4j, which is open-source[67] where a free version has limited capacity

b. Ontotext GraphDB, which is proprietary with a free and paid version[68]

[66] For a video explanation, see https://neo4j.com/blog/ontologies-in-neo4j-semantics-and-knowledge-graphs/.
[67] See https://neo4j.com/product.
[68] See https://www.ontotext.com/products/graphdb/?ref=menu.

The following figure presents an example of a graph structure that can be created with such a graph database, where all property keys and values and names of relationships can be customised.

Figure 30: Example of a graph structure created with a graph database



An example query for the above model can look in pseudo-code like:

RETURN ALL Residential Building WHICH BELONGS_TO Company X AND LOCATED_AT Noord-Holland

API software to collect and/or share data

If one chooses the cloud solutions discussed above in the hardware section, API solutions are integrated (Azure cloud) or a version with API solutions is recommended (Datopian). If one does not use a cloud solution to store and manage data, Capterra provides an overview of dedicated API integration software.[69] Each software has a different pricing plan, but one should expect a cost of upwards of $200 per month.

[69] See https://www.capterra.com/sem-compare/integration-software/.

5. Legal implementation

5.1 Introduction

Legal risks are some of the key challenges in developing a DBL. Other risks have been discussed in other chapters, notably lack of stakeholder support (Chapter 2), lack of interoperability (Chapter 3), technical challenges (also Chapter 3) and financial challenges (Chapter 4). Section 2.5 also provided an overview of the general barriers to implementing a DBL.

This chapter focuses on legal risks for DBL managing authorities specifically linked with DBL and does not address other potential legal risks, for example, risks linked with so-called "smart" buildings, i.e. hi-tech buildings automatically and continuously collecting data of building occupants, with sensors and other technologies. As many such data are likely to be considered "personal", there are potential risks to the privacy of the occupants' personal data rights.

Identification of the legal risks linked with DBL

To identify the legal risks, a scoping survey was held among a diverse group of stakeholders of which 114 started the survey and 36 completed it. Feedback on initial results was collected in a second survey which 80 stakeholders started and 28 completed. In addition, three interviews were held with people who were involved in the development of a national or regional DBL or a building logbook-like initiative were interviewed (Estonia, Basque country, Bulgaria).

Based on the results, various critical legal risks were identified. A number of legal documents were screened on relevance for these legal risks (see Annex IV), to select the main pieces of legislation for further analysis. The most relevant EU law (amongst others the GDPR (General Data Protection Regulation)[70] and the 1996 Directive on legal protection of databases)[71] and national law of some countries (Estonia, Germany, and the Netherlands) were then checked for potential solutions to address these risks. This document aims to provide both guidance concerning the identified legal risks surrounding DBLs and suggestions for:

- organising the processing of personal data so that it complies with EU data privacy law
- classifying degrees of privacy sensitivity
- types of licences to deal with commercial rights

Some of these topics overlap with other domains. For example, authorisation to access data and ensuring data security also touches on data management and technical solutions that are discussed in other chapters of the technical guidelines. However, this chapter focuses on challenges from a legal point of view.

In a scoping survey, stakeholders identified three main legal risks:

- The rights of individuals to access, rectify and transfer personal data
- Requirements that documents must contain privacy-sensitive data
- Uncertainty about who is responsible for data security

The first two questions relate to the processing of **personal data** and will be discussed first.

[70] 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, https://eur-lex.europa.eu/eli/reg/2016/679/oj

[71] Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A31996L0009

While the short answer to the third question is that the holder of data is responsible for data security, further legal provisions regarding **data security** will also be discussed.

A legal risk that was less often mentioned in the survey but was mentioned in several interviews are rights of private owners to commercially exploit the building data of their buildings, including the refusal to share data to be included in a DBL or to demand payment for sharing data. Such rights of **commercial** nature are discussed in the last section of this chapter.

Other legal barriers such as specific regulations on professional confidentiality and incompatibility of professional activities depend on national legislation and are not discussed in these guidelines.

It should be noted that the legal risk assessment focuses on risks for the implementation of a DBL as a whole and not on technical legal requirements such as those following from INSPIRE or LEVEL(s), and whether CPR requirements should be met with in a DBL or only in DPPs and building materials passports (see chapter 1).

In this chapter we cover personal data, commercially sensitive data, and public data. How personal data need to be processed, is the same for data that are owned by the private sector actors and the public sector. For private-sector owned data, we discuss which rights owners of commercially sensitive data have. For publicly owned data, any copyright of the private data owners (whose data are in the DBL, such as possibly architect drawings) still needs to be respected. In addition, we advise that public-sector owned data are always made available under a license that states that users may only further share the DBL for free (and with a proper reference).

5.2 Personal data rights

a) GDPR and personal data in DBLs

The General Data Protection Regulation (GDPR)[72] Article 4(1) defines personal data as "*any information relating to an identified or identifiable natural person ('data subject')* […]". In addition, **Article 9 prohibits the processing of certain "*special categories of data*"** such as ethnic origin, political opinions and genetic or health data unless for specific purposes or if the data subject has given explicit consent or has manifestly made those data public. Apart from this, the GDPR does not define explicitly what data are personal, although Recital (18) provides home addresses as an example.

It is, therefore, important to differentiate between those building data referring to natural persons, potentially protected by GDPR, and those referring to legal persons, not covered by GDPR.

In the next step, it is important to check which personal data referring to natural persons allows the **identification** of individual persons. So long, as such identification is not possible, neither directly nor indirectly by linking several different data, personal data rights are respected. In cases in which individual persons are, directly or indirectly, identifiable, the use or otherwise of such data depends on the consent (defined in Art 4 (11) GDPR) given by that person, the "data subject".

The question of who owns a building, whether a natural person or a legal person, is not relevant in this context, as only the qualification as personal data is decisive. If a building with dozens of tenants is owned by a company, then that does not change the character of tenant-related data as "personal".

Based on interviews, desk research[73] and workshop feedback, the general opinion seems that **data purely relating to a building do not "relate" to a person.** Consequently, data such as year of construction, total floor area, number of balconies or architectural plans and drawings would not be considered personal data, as long as they do not include information that identifies the owner or allows doing so. However, no law or court decision was identified that settles this question beyond doubt.

[72] See https://eur-lex.europa.eu/eli/reg/2016/679/oj.
[73] See e.g., https://www.owenboswarva.com/blog/post-hou5.htm.

Also, in the final event, representatives of some countries responded to a question about personal data that they did not consider for example fire alarms as personal data, while the representative of another country considered for example the type of isolation material personal data. What data Member States will classify as personal data will depend on how much the government wants to protect the interests of building owners and other parties such as possibly prospective buyers and tenants of the building. Therefore, the classification of personal data will differ between Member States, and hence the public availability of certain data.

It is proposed to classify each data item in terms of privacy sensitivity before deciding to include it in a DBL or under which conditions (for example authorisation requirements) to include it in the DBL. The following example illustrates what the authors believe to be a borderline case, although we realize that others may have a distinct opinion the one way or the other.

b) Target type of tenants

Under a target type of tenants we understand the target group of tenants or apartment owners for whom the building is constructed, as for example in the case student homes, senior citizens' homes or nursing homes. Responses on a proposal to include the target type of tenants in the DBL revealed a **border case** of what is considered personal data. On the one hand, this could be argued not to be privacy sensitive because it may also be open to certain people not meeting the criteria, such as partners or carers. On the other hand, even aggregate information on actual tenants could be privacy-sensitive. For example, the percentage of tenants that are registered at a university reveals data on persons if the percentage is close to 0% or 100%. Whether the publication of the target type of tenants is problematic from a personal data perspective, may thus vary between Member States. If national law prohibits letting apartments to persons not in the target type, the target type of tenants would reveal personal data. If national law does not prohibit letting or selling building units to people who do not belong to the target group, the target type of tenants could be included in the DBL. However, data on the actual type of tenants, even at an aggregate level, still runs the risk of revealing personal data.

c) Reasons to legislate DBL processing of personal data

Article 6 of the GDPR provides that personal data may be lawfully processed if one more or conditions apply. For example, Article 6(1)(a) states that personal data may be processed for one or more specific purposes if the data subject has given consent. This condition comes with the risk that some people may not consent, in which case a national DBL would not cover all buildings. An alternative condition that is more useful for a DBL is Article 6(1)(c), namely that "***processing is necessary for compliance with a legal obligation to which the controller is subject***". This condition suffices to lawfully process personal data, as well as another condition in Article 6(1)(e): "*processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority invested in the controller*". However, the condition of Article 6(1)(e) comes with a right of data subjects to object to the processing of data according to Article 21(1) discussed later. Thus, enshrining the processing of personal data in law is more effective in ensuring coverage of all buildings in a national DBL.

Articles 12-14 of the GDPR also provide that generally, the data subject needs to be informed about which data are collected and why. However, Article 14(5)(c) provides that if data is not obtained from the data subject itself, the data subject does not need to be informed if "*obtaining or disclosure is expressly laid down by Union or Member State law to which the controller is subject and which provides* ***appropriate measures to protect the data subject's legitimate interests***". Hence, if for example, a law states that the Cadastre or Civil Registries must collect certain personal data, then there is no need to obtain consent from the data subject to include the personal data in the DBL.

d) Legitimate interests of data subjects

Although mentioned in several articles, the GDPR does not explicitly define the legitimate interests of data subjects.[74] In practice, a careful balance needs to be made between various legitimate interests of the building owner and other people. In the case of building owners as data subjects, one might for example define legitimate interests as the right to own and use a building (legal interests) and/or as beneficial interests such as living or working in the building, a share of rental income, and a share in the proceeds of a sale if the building is sold.[75] While not informing prospective buyers or tenants of deficiencies is generally unlawful, it may be in the interest of the building owner that information on deficiencies is not publicly accessible. For example, public information about the presence of **asbestos** in a house may negatively affect the sales value and thus the data subject. However, for potential buyers this information would be useful. In Germany, the Federal Supreme Court has decided that the seller needs to inform prospective buyers of asbestos.[76] However, for example in Belgium, sellers do not yet have to inform buyers about asbestos, although from 2032 each house in Flanders needs to have an asbestos certificate.[77] Hence, the implication of legitimate interests of data subjects for the inclusion of data in a national DBL may differ between countries.

e) The right to be forgotten

Besides the refusal of consent, the right to be forgotten is another potential risk that a national DBL does not cover all buildings. Again, enshrining the processing of personal data in law reduces this risk. The GDPR provides in Article 17(3) **the conditions under which data subjects have no right to be forgotten**, including if "*data processing is necessary to comply with Union or Member State law or is carried out in the public interest or in the exercise of official authority vested in the controller*".

In addition, Article 17(1)(a) provides that the data subject has a right to be forgotten if "*the personal data are no longer necessary in relation to the purposes for which they were collected or otherwise processed*." However, if personal data is necessary, a right to be forgotten may still follow from Article 17(1)(b)-(f). Hence providing the necessity to process certain personal data in law is a surer way to ensure that the national DBL covers all buildings.

Thus, from a GDPR perspective, after a data protection impact assessment and confirming which (personal) data may be lawfully included in a DBL, **explicitly listing all (personal) data** that a DBL must contain in a national law, or in a decree that is mentioned by a law, suffices for the lawful processing and even disclosure of (personal) data – as long as the data controller and processor take appropriate measures to protect the data subject's rights and legitimate interests. It is advisable to regulate the specific processing operation of the DBL in the law that establishes the DBL. If a data protection impact assessment is done in the context of the adoption of that law, Article 35(10) of the GDPR provides that no further impact assessments before individual processing activities are needed (unless the Member State deems it necessary). All data that the DBL shall collect, can then be listed in a decree or statute mentioned by that law.

The approach outlined above could be feasible for some Member States. It is possible to list all personal data in one law or statute and update this list when new personal data are collected (for example, in Estonia), or to mention any new addition in a different piece of legislation as the need arises (for example, in the Netherlands). In other countries, additional potential restrictions to process personal data may need to be considered (for example, Germany).

---

[74] Not to be confused with the legitimate interests of the data controller or a third party mentioned in GDPR.
75 See https://www.rocketlawyer.com/gb/en/family-and-personal/manage-personal-property/legal-guide/legal-interest-and-beneficial-interest-in-property
[76] See http://juris.bundesgerichtshof.de/cgi-bin/rechtsprechung/document.py?Gericht=bgh&Art=pm&pm_nummer=0066/09
[77] See https://aspadvies.be/blog/asbest-melden-bij-verkoop-woning-verplicht.

f) Application to some Member States

Which data are considered personal data and which personal data can be published, varies between Member States. In the final event, representatives from Czechia, Lithuania and the Netherlands were asked which data they considered to be personal data. Where representatives from Czechia and Lithuania did not consider for example fire alarms in a building to be personal data, the representative from the Netherlands considered the use of isolation materials in a building to be personal data. Also, whether (personal) data are published, depends on the weighting of legitimate interests of the building owner as data subject and the legitimate interests of other people, as noted above in the case of asbestors in Belgium and Germany. Below, we provide examples of legal considerations in three Member States. These examples may not be transferable to other Member States, but provide insight in the type of considerations for the processing of building data.

Estonia. For example, in Estonia the Art. 58(2) of the Construction Code provides that a building register shall be established. Article 9 of the Estonian statute of the building register provides that it shall include the first and last name, the personal identification number or the date of birth of the owner of the building, and Article 39(1) that "***the data are public, except for cases provided for in legislation***".[78] Although the building register is a regular cadastre, it contains more data than cadastres in most other countries including for example energy performance labels (Article 12 of the statute) and can thus be seen as a cadastre that is being developed into a national DBL.

The Netherlands. In the Netherlands, Article 3a of the Cadastre Act[79] provides that the Cadastre shall collect personal data for the four purposes determined in Article 2a: legal certainty, effective geo-information structure, public tasks, and support of economic activities, except for supporting a charitable or commercial relationship. Article 8 provides that the real estate register (and other registers) shall be public, and Article 48 provides that the key register shall include personal data of the real estate owner, and amongst others the mortgage value or the maximum mortgage value that is allowed. However, the Dutch Cadastre also collects additional building data as required in other laws of which Article 41 of the Cadastre Regulation provides an overview.[80] A central place providing an overview of personal data collected in the DBL is recommended for the sake of transparency.

Germany. In other Member States, **national law or jurisprudence** may further restrict the collection of personal data. For example, in Germany, the Federal Constitutional Court decided in 1983 that citizens have a right to self-determination concerning their personal data.[81] Of course, there have been several EU instruments since 1983 that govern personal data, but this decision may still present an additional legal risk to include personal data, that to our knowledge has not been resolved in legislation or jurisprudence. Because this self-determination on the provision of personal data severely limits a nationwide collection of personal data, a first step regarding personal data for the establishment of a national DBL should be a formal decision that data on buildings, or at least certain data on buildings, are not personal data. Personal data in the DBL would then remain incomplete where people do not assent to provide these data, but at least a DBL could be populated with data such as year of construction, total floor area, the number of balconies, etc.

g) Data subjects' right to object

The GDPR provides for certain rights of data subjects to object to the processing or specific uses of personal data. Where a data subject has lawfully objected to data processing, this means that his personal data may not be included in a DBL, rendering a DBL incomplete depending on what data the Member State considers to be personal data. Objections to specific uses of personal data have implications for how personal data can be published to ensure that the personal data cannot be used for the purposes that the data subject objects to.

---

[78] See https://www.riigiteataja.ee/akt/126062015013.
[79] See https://wetten.overheid.nl/BWBR0004541/2023-05-01.
[80] See https://wetten.overheid.nl/BWBR0027695/2021-02-24.
[81] See www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/DE/1983/12/rs19831215_1bvr020983.html.

Article 21(1) of the GDPR grants data subjects the right to **object to data processing** that is "*necessary for the performance of a task carried out in the public interest or in the exercise of official authority invested in the data controller*" (Article 6(1)(e)), or data processing "*necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject* […]" (Article 6(1)(f)). In both cases the controller shall no longer process the personal data "*unless the controller demonstrates compelling grounds for the processing which override the interests, rights or freedoms of the data subject, or for the establishment, exercise or defence of legal claims*". It should be noted that Article 21 does not provide a right to object to data processing for the specific purpose of a national DBL if the personal data to be processed for that purpose are enshrined in law, the condition of Article 6(1)(c).

In addition, data subjects may object to data processing for **direct marketing purposes** according to Article 21(2) GDPR and may object to (automatic) **profiling** in the context of the use of information society services in general according to Article 21(5) GDPR, regardless of the conditions of Article 6 for lawful processing of personal data.

h) Profiling

As noted above, Article 21(5) provides that data subjects may object to automatic profiling in the context of the use of information society services. Article 4(4) GDPR defines profiling as "*any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements*". For example, if an automatic process groups people by income this already profiling. This would actually affect building owners if for example only those in certain income classes are targeted for renovation proposals. To avoid the risk of profiling, measures need to be taken to prevent automatic processing of data that could be used for profiling. To address the issue of profiling, the following two steps are proposed:

- Firstly, define categories of privacy sensitivity[82]:
- Not privacy sensitive (e.g., building data)
- Somewhat privacy sensitive (e.g., name of the owner, annual energy consumption)
- With a risk of automatic profiling (e.g., email address, income class, real-time energy consumption)
- Secondly, limit "bulk access" to somewhat privacy-sensitive data (cat. ii) and limit any access to selected authorities for very privacy-sensitive data (cat. iii).

In the feedback survey, the **most demanded data extraction feature was bulk extraction** of data on all buildings meeting certain criteria, for example all buildings with a certain minimum total floor area, or all buildings constructed after a certain year.

Although indicative examples are given above of what data might be considered somewhat or very privacy sensitive (cat. b and c), there is no definition of what type of personal data allows automatic profiling, and the Member States or even the data controller should assess which personal data constitute a risk of automatic profiling.

[82] The special categories of personal data defined in Article 9(1) of the GDPR are not considered relevant for a DBL – if they are nevertheless considered for inclusion in the DBL they would form a fourth category to which the other provisions of Article 9 apply,

In principle, it makes no difference whether privacy-sensitive data can be directly extracted from the DBL or indirectly from documents obtained from the DBL.

The principle behind limiting "**bulk access**" for somewhat privacy-sensitive data (cat. b) is to limit profiling by users of the DBL. However, if the data extraction includes the name of the owner, this could be combined with data from other sources for direct marketing purposes. Thus, bulk extraction of building data could be limited to data that are not privacy-sensitive and to documents that contain no privacy-sensitive data. In the specific case when somewhat privacy-sensitive data (cat. b) can be numerically aggregated, they could be presented as averages at the street level or city district level.

In the feedback survey, a majority of 26 out of 28 agreed fully (17) or partially (9) that a DBL should include privacy sensitive data if appropriate authorisation procedures are in place. The one respondent who answered not to know the answer, clarified that online platforms require users to accept their rules if they use the platform – which seems to be a different issue (such as tracking data) from the publication of building data. The one respondent who did not agree mentioned that information that affects national security should not be in the DBL at all.

If there is a legal requirement that certain documents such as perhaps a sales deed or a registration of ownership must contain very privacy-sensitive data such as an email address (cat. c), there is a risk that inclusion of documents in the DBL reveals such highly sensitive personal data. Therefore such documents should be made accessible with only severe restrictions, for example **access limitation to selected authorities on a case by case basis**, for example in the context of law enforcement on a case by case basis, or in the context of academic research.

5.3 Data protection and security

Data protection is implicitly defined in Article 25(2) GDPR as "*appropriate technical and organisational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed*". It means for example that people working at the DBL managing authority who verify or classify building data do not have access to contact data of the owner of the building.

According to a definition used by IBM, "*data security is the practice of protecting digital information from unauthorized access, corruption or theft throughout its entire lifecycle. It's a concept that encompasses every aspect of information security from the physical security of hardware and storage devices to administrative and access controls, as well as the logical security of software applications*"[83].

The GDPR implicitly defines data security as *"appropriate technical and organisational measures" (Article 32(2) GDPR) to address "the risks that are presented by processing, in particular from accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to personal data transmitted, stored or otherwise processed"* (Article 32(1) GDPR).

This subsection discusses the requirements for data protection and data security.

a) Data protection requirements

The specific risk that stakeholders saw with regard to data protection is uncertainty about who is responsible for this. Legislation would be the most obvious solution to address these risks. For personal data, the responsibility in the GDPR is attributed to the controller by Art. 24(1,2) GDPR and

[83] See: https://www.ibm.com/topics/data-security.

to the processor by Art. 28(1) GDPR. For personal data, the GDPR defines different responsibilities for a **data controller** and a **data processor**. The data controller is the entity that determines how and why personal data are processed. The data processor is the entity that processes the data on behalf of the data controller. If a public sector body (such as a national DBL managing authority) holds other data that are protected on certain grounds, Chapter 1 of the Data Governance Act[84] applies and will be discussed later in this section.

Article 25 of the GDPR makes the data controller responsible for the implementation of the processing of personal data, amongst others in particular the following aspects of data protection:

- Ensuring data protection principles such as pseudonymisation and data minimisation
- Ensuring that personal data are not made accessible without the individual's intervention to an indefinite number of natural persons

For these responsibilities, the data controller shall take into account "*the state of the art, the cost of implementation and the nature, scope, context and purposes of the processing as well as the risks of varying likelihood and severity for rights and freedoms for natural persons by the processing*".

Article 25 also provides that a certification mechanism may be used to demonstrate compliance, and Article 42(3) confirms explicitly that certification shall be voluntary. This means that compliance of a DBL with GDPR may be proven by a certification from an accredited certification scheme. The GDPR does not require certification, but nothing in the GDPR prevents Member States from requiring such certification.

b) Data security definition and risks

With regard to personal data, the GDPR deals with data security. The European Data Governance Act[85] discussed further below applies to data security of commercially confidential data and is discussed further below.

Risks linked with "personal data breach", defined in Art. 4(12) GDPR, concerns any structured set of data, personal or otherwise, "*which are accessible according to specific criteria, whether centralised, decentralised or dispersed on a functional or geographical basis*" (Art. 4 (6) GDPR for personal data), including DBL, but they are not specific for DBL.

Article 32 GDPR states that data controllers and data processors are responsible for data security.

To address the data security aspect of personal data processing, it suffices to include in the law that establishes the national DBL an **article that defines the data controller and data processor**. For example, a data controller could be a Ministry or public authority, and the data processor could be an employee of that Ministry or the National Cadastre. Their responsibilities then follow from the GDPR and any additional national regulations (and must be complied with).

For public authorities or bodies processing personal data in the context of a DBL, they must in their role of data controller and/or data processor designate an expert in data protection law as a **data protection officer**, who needs to be involved in all issues relating to the protection of personal data (Article 37 and 38).

The GDPR does not provide explicit rules by which "appropriate technical and organisational measures" the data controller and data processor shall ensure the data security, except in general terms, to take account of the **state of the art and the cost of implementation** (Article 32(1)), and non-exhaustively:

- The pseudonymisation and encryption of personal data

[84] EU Regulation 2022/868, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32022R0868.

[85] Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32022R0868

- Secure processing systems and services
- Timely restoration of incidentally lost personal data
- Regular evaluations of the security of the processing
- Steps to ensure that only the designated processor can process the personal data

c) Transfer to third countries

Also relevant is that Article 46(1) makes the data controller or data processor responsible for **not transferring personal data to third countries that the Commission has not approved** unless they provide appropriate safeguards and data subject rights are enforceable. Measures against direct access to personal data from such third countries include geo-blocking, measures against VPN and regular monitoring of internet traffic to the DBL website or app. In addition, such transfers may happen involuntarily if servers in such third countries are used either directly or indirectly through the use of certain IT software or platforms, and authorities from those countries require the transfer of personal data. Chapter V (Article 44-50) of the GDPR only regulates intentional transfers to third countries. However, Article 32(2) of the GDPR requires in general an "*appropriate level of security*" for amongst others "*unauthorised […] access to personal data transmitted, stored or otherwise processed*".

d) The European Data Governance Act

As noted earlier, Chapter 1 of the Data Governance Act (Regulation 2022/868, DGA)[86] applies to public authorities holding data that are protected on certain grounds. These grounds are according to Article 3(1) of the DGA:

- Commercial confidentiality
- Statistical confidentiality
- Intellectual property rights of third parties
- Protection of personal data that fall outside the scope of EU Directive 2019/1024 on public sector information

Statistical offices, municipal registrations, patent offices and tax authorities are typical public bodies that hold protected building data which might be of interest for inclusion in a DBL. In such cases, the DGA might apply to this re-use of certain such building-related data, if their data subject or rights holder considers them commercially confidential, such as the rental price or sales price of an office building, or the design of an architect.

The conditions for re-use of data that relate to data security are provided in Article 5(3) of the DGA. The public sector body holding the data is responsible for preserving the protection of data following Union and national law, by one or more of the following means:

- Ensuring anonymisation in the case of personal data
- Modifying, aggregating or treating by another means of disclosure control, in the case of commercially confidential information, including trade secrets or content protected by intellectual property rights
- Providing remote access in a secure processing environment that is controlled by the public sector body [that holds the data, red.]
- If remote access would jeopardise the rights and interests of third parties, access and re-use of data within the physical premises in which the secure processing environment is located in accordance with high-security standards

[86] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R0868.

This means what is most appropriate, could be different for each protected data item. Rental prices per square metre could be averaged per city district, and documents or architectural designs could be made available to authorised users of selected authorities through secure remote access.

In addition, Article 5(4) provides that the public body holding the protected data shall impose conditions “that preserve the integrity of the functioning of the technical systems of the secure processing environment used” by the person re-using the data (the re-user for short). The public body shall also reserve the right to verify the processing of data by the re-user and reserve the right to prohibit the use of data if that jeopardises the rights and interests of third parties. For example, the public body checks every result before releasing them out of the secure (remote or within-premise) environment.

Article 5(5) provides the additional safeguard that the public body holding the protected data shall impose a confidentiality obligation on the re-user, in particular the re-identification of any data subject (natural person) to whom the data relates.

If the above means to protect data cannot be allowed, Article 5(6) says that the public sector shall provide assistance to contact the data subject or rights holder.

Article 5(8) provides that data that is considered confidential according to EU or national law, may only be re-used with consent of the data subject or rights holder.

Article 5(9) requires that anyone intending to transfer protected data to third countries must notify the public body holding these data when he requests the re-use of data, and conditions and procedures for permission for a transfer of re-used data are described in Article 5(9)-5(14).

Article 5(7) provides that the re-use of data shall only be allowed in compliance with intellectual property rights, discussed below under the header “Commercial rights”.

5.4 Copyright and *sui generis* rights

a) Introduction

In the context of a national DBL, the use of data from databases in which others have substantially invested may be most relevant. However, some other aspects of commercial rights are covered as well, which may be relevant if the aim is that the DBL includes copy-righted content such as perhaps architectural designs or device-generated data such as perhaps smart meter data.

In this section, the following aspects are discussed:

- Copyright and sui generis rights
- a potential general provision allowing public sector buyers to require an exchange of building data inspired by the proposed Data Act and the existing right to require this in case of public procurement
- a proposal that data owners and providers should have free access to their own data
- Licenses to set conditions on sharing DBL data

Commercial rights of data owners are protected in various pieces of EU legislation, depending on how the data are created. The most relevant is Directive 96/9/EC[87] on the legal protection of

[87] Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, see https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX:31996L0009. The partial amendments by the copyright Directive 2019/790/EU with regard to text mining, educational purposes and cultural heritage do not affect the matter discussed in these guidelines.

databases. It distinguishes between **copyright**-protected databases and databases protected by '***sui generis***' rights. For *sui generis* rights the only criterion is a substantial investment in either the obtaining, verification or presentation of the contents. Examples of databases with *sui generis* rights could be a database with data compiled from other sources, an online version of offline data, or a database with data extracted from documents such as sales deeds, building ownership registrations etc.

An architect's folder of technical drawings could be **copyright-protected**. If someone has made substantial investments in a platform to make architectural designs available under the conditions of an architect, or someone invests substantially in collecting sales prices from many real estate managers, these could be examples of databases that are protected by ***sui generis*** rights. If a national DBL manager wishes to publish data taken from a database of other online platforms, then such a database is most likely protected by *sui generis* rights rather than intellectual property rights (although the data may still be protected by intellectual property rights, if the data include for example architectural design data).

b) Copyright and *sui generis* rights protection

Copyright-protected databases enjoy the highest degree of commercial protection. Article 5 of Directive 96/9/EC provides (here in slightly summarised form), that no part of the database may be reproduced, translated, adapted, arranged, or otherwise changed, distributed, communicated, displayed or performed (these are all called "restricted acts") without authorization from the author (creator) of the database. However, as an exception to these restricted acts, Article 6 (2)(c) of this Directive provides that Member States can limit these rights for an administrative procedure.

Databases protected by ***sui generis*** rights enjoy a slightly lower degree of commercial protection, namely that without permission of the maker of the database, the database may not be extracted, re-utilised or evaluated in whole or substantial part, and that insubstantial parts may not be repeatedly and systematically extracted or re-utilised if that conflicts with normal exploitation or otherwise unreasonably prejudices the legitimate interests of the maker of the database (Article 7 of Directive 96/9/EC). A DBL that **systematically** facilitates end users to access data that is obtained from a database protected by sui generis rights, must therefore seek permission from the maker of that database. Again, Member States may restrict the rights of the maker of the database for an administrative procedure according to Article 9(c) of Directive 96/9/EC. In addition, Article 5(7) of the Data Governance Act (DGA) provides that a public sector body holding data from a database protected by sui generis rights, may not use this as an argument to prevent the re-use or to restrict the re-use beyond the limits set in the DGA, which have been discussed under the previous header "Data security".

The Member State's right of **limiting author and maker rights** means for example that the Member States can provide that certain data must be provided in building permit procedures. The degree to which author and maker rights may be limited in Article 6(2)(c) and Article 9(c) respectively is not further specified. Thus, if national law provides that the data obtained in the administrative procedure shall be publicly available, this does not directly contravene Directive 96/9/EC. However, it remains important to check commercial data protection rules at the national level and ensure that the DBL also respects national rules.

In addition, it should be noted that the rights of database owners and the rights of Member States to limit the rights of database owners still do not prejudice the **rights of the data owners**, who own (part of) the content of the database. The copyright directive 2001/29/EC[88] applies to creative

[88] Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32001L0029

works only and should not apply to building data. However, certain building data may be protected by patents. For example, in 2020 about 1,600 BIM-related patents were applied for.[89]

It should also be noted that Article 13 of Directive 96/9/EC provides that this Directive does not prejudice **copyright provisions in other legislation**, including for example confidentiality, access to public documents and the law of contract. Thus, even if EU law allows Member States to limit the rights of database owners, database owners may still be granted additional rights by national law, which is not further discussed in these guidelines.

c) Declaration of pre-existing rights

For **copyright-protected data** such as likely the design of non-public buildings, and other data protected in the sense of Article 3 of the Data Governance Act (discussed above under the header "Data protection"), the safest approach is that the DBL manager requires a **declaration on pre-existing rights** or their absence from any data provider. If pre-existing rights apply, the declaration should provide under which conditions the privately owned data may be published. It is difficult to define in general when data are commercially confidential or copyright protectable. For example, for the design (structure, building materials) of for example balconies the authorship of an architect or engineer is evident and such a design document is copyright protectable. But for example for the number of balconies there is probably no clear authorship and this datapoint is probably not copyright-protectable. To clarify pre-existing rights and conditions under which data may be shared, a standard form could be developed with the following formulation:

> "*I [name person], representing [name organisation]*
>
> *[warrant that the data listed below are free of rights or claims from creators or from any third parties for publication through the Digital Building Logbook /*
>
> *clarify that the data are owned by [...] and may only be published under the condition(s) [that only ... have access] / [the data owner is paid [...] per view/download] / [of license [...]] / [other conditions].*"

d) Proposed Regulation on device-generated data

The proposed Data Act[90], on which the Council[91] and the European Parliament[92] presented their positions and reached a provisional agreement on 27 June 2023,[93] will legislate how data generated by the use of products or related services (**device-generated data** for short, such as smart meter data) may be made available. Data that is automatically generated by, for example, smart meters or solar panel monitors would be covered by the proposed Data Act.

The proposed Data Act provides that so-called data holders (defined in Art. 2(6)) who have access to data generated by products or related services and who are not micro or small enterprises (Art. 7(1)) are required to make those **data available to users** of the product (Article 4). In addition, on request of the user or a party acting on his behalf, the data holder shall share the data with a third party (Article 5(1)). Which parties can act on behalf of a user is not clearly defined, and thus it is not certain whether housing associations or homeowner associations could be regarded as such. This third party could be a DBL platform.

Although the proposed Data Act only applies to data generated by products like smart meters and is thus less relevant for typical prototype DBLs, Member States could also use it as a **template** for national legislation on other protected building-related data. For example, it could perhaps be used

[89] See https://www.researchgate.net/figure/BIM-related-patents-from-2011-to-2020_fig1_361008249
[90] Proposed Data Act: see https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2022%3A68%3AFIN
[91] See https://data.consilium.europa.eu/doc/document/ST-7413-2023-INIT/en/pdf.
[92] See https://www.europarl.europa.eu/doceo/document/TA-9-2023-0069_EN.html.
[93] See https://www.eu-data-act.com/.

as a template to give building owners the right to view technical drawings even if they are copyright-protected.

e) Sharing device-generated data

The Proposed Data Act, in addition, prohibits third parties, according to Art. 6 (2):

- to impair the decision autonomy of users
- to use the data for profiling purposes
- to share the data with other third parties and in particular with gatekeepers
- to develop a product that competes with the product that generates the data it received
- to prevent the user from making data available to other parties

Article 8 of the proposed Data Act provides that if the data holder is obliged to make device-generated data available, he shall agree to **fair, reasonable and non-discriminatory** terms with the data receiver. According to Article 9(1), this includes reasonable compensation, although Article 9(3) provides that Union law or national implementing law may exclude or reduce compensation for making data available.[94]

In the specific case of **outsourcing** the operation of a national DBL platform to a private company with a dominant position in multiple countries[95], Article 5(2) of the proposed Data Act provides that the gatekeeper may not solicit or commercially incentivise users of the device to provide device-generated data. Consequently, as long as national DBL platforms are run by public authorities or by private companies that run a DBL platform in only one country, the data holder (for example the DBL managing authority) is not prohibited from sharing data with the DBL platform upon request from the user of the device. If the private company runs a DBL plantform in multiple countries, the data holder would need investigate if the private company has a dominant position and has not solicited or incentivised users.

f) Public procurement provision to require building data exchange

For **public buildings** such as schools, hospitals, town halls, electricity plants etc., the construction is typically publicly procured. The 2014 “classical sectors” Public Procurement Directive[96] provides general rules on communication in Article 22. In addition, for public works and design contests, this Directive provides in Article 22(4) that Member States may require that bidders to public procurement use BIM (Building Information Modelling), as long as (until the BIM tool becomes commonly available) the contracting authority is required to provide unrestricted and full direct access to such a tool, to provide free provisional tokens for access, or to support an alternative channel for electronic submission of tenders.

[94] The Proposed Data Act also provides that data holders may be required to share data with public bodies on the grounds of exceptional needs (Art. 14 ff). However, because the purpose of DBL is to make building data systematically available, this ground for data sharing is not relevant for developing a DBL.

[95] This is the definition of a “gatekeeper” defined in Article 3 of the Digital Markets Act, to which Article 5(2) of the proposed Data Act refers. For the Digital Markets Act, see https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32022R1925.

[96] Directive 2014/24/EU of the European Parliament and of the Council of 26 February 2014 on public procurement and repealing Directive 2004/18/EC, Art. 22(4). The Utilities Directive 2014/25/EU has a similar provision in Art. 40(4).

g) Access of data owners to their own data in DBL

Although it is perhaps more an aspect of data management, it is a good principle that **data owners and providers** always have free access to their own data in the DBL that they own or have provided, even if the same data is not freely accessible to other users. For example, a building owner, but also the construction company, real estate manager etc. providing data should have free access to the energy performance label, soil conditions report, sales deed etc. The reason is that the data provider should be able to verify the data in the DBL and propose changes.

> In the feedback survey, 26 out of 28 respondents agreed fully (21) or partly (5) that data owners and providers should have free access to the data they own or provide (though not necessarily to other DBL data). The one respondent who did not agree, argued that maintenance of data is costly and everyone should pay, but this is clearly the view of a minority.

h) Licenses to set conditions on sharing DBL data

Regarding the use of **publicly owned** DBL data, it is recommendable that all service providers using these data must accept a licence which provides that they:

- may only share the DBL data for free; and
- shall be required to provide any newly generated data that used DBL data as input under an "Attribution-NonCommercial-ShareAlike" Creative Commons license, or under a license that grants more free access.

> This type of **creative commons license** (https://creativecommons.org/licenses/by-nc-sa/4.0/) means that appropriate credits to the source must be given and adaptions must be indicated without suggesting endorsement from the licensor (Attribution), that the material may not be used for commercial purposes (Noncommercial), and that any adaption of the material must be distributed under the same license as the original (Sharealike).

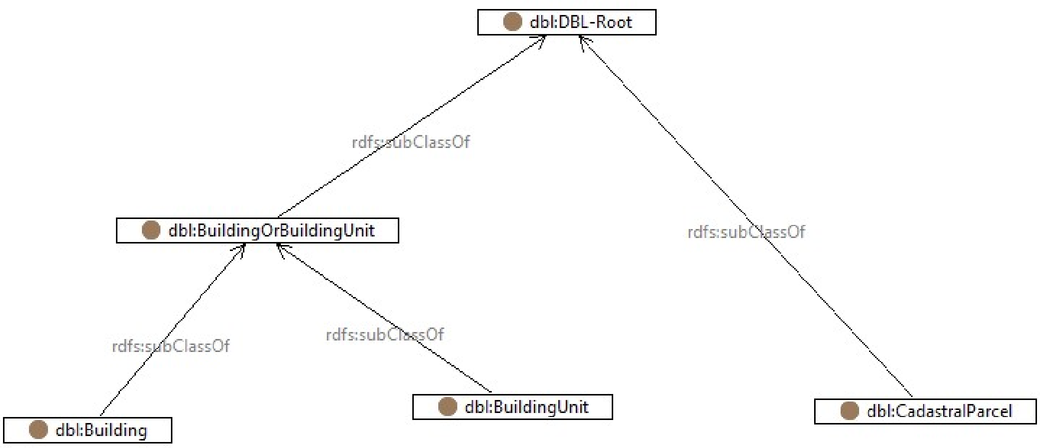
For example, if publicly owned DBL data include that the energy performance label of a particular building is "C", an engineering company creating a renovation cost calculator must provide the energy label for free, but may charge a fee for the calculation of renovation costs using the energy label. The developer of the renovation cost calculator must allow free public non-commercial use of the renovation cost after being paid EUR 10 per building, for example, to grant renovation subsidies. However, a renovation service provider seeking to approach potential customers, or the developer of a renovation price comparison tool would need to ask for a different license which the developer of the renovation cost calculator may grant or not.

> A majority of 24 out of 28 respondents in a feedback survey agreed fully (15) or partially (9) with the above principles. Partial responses underlined that right of access is not necessarily right of free access, and that authorised persons should be well defined.

Annex I: Detailed information on the DBL Ontology

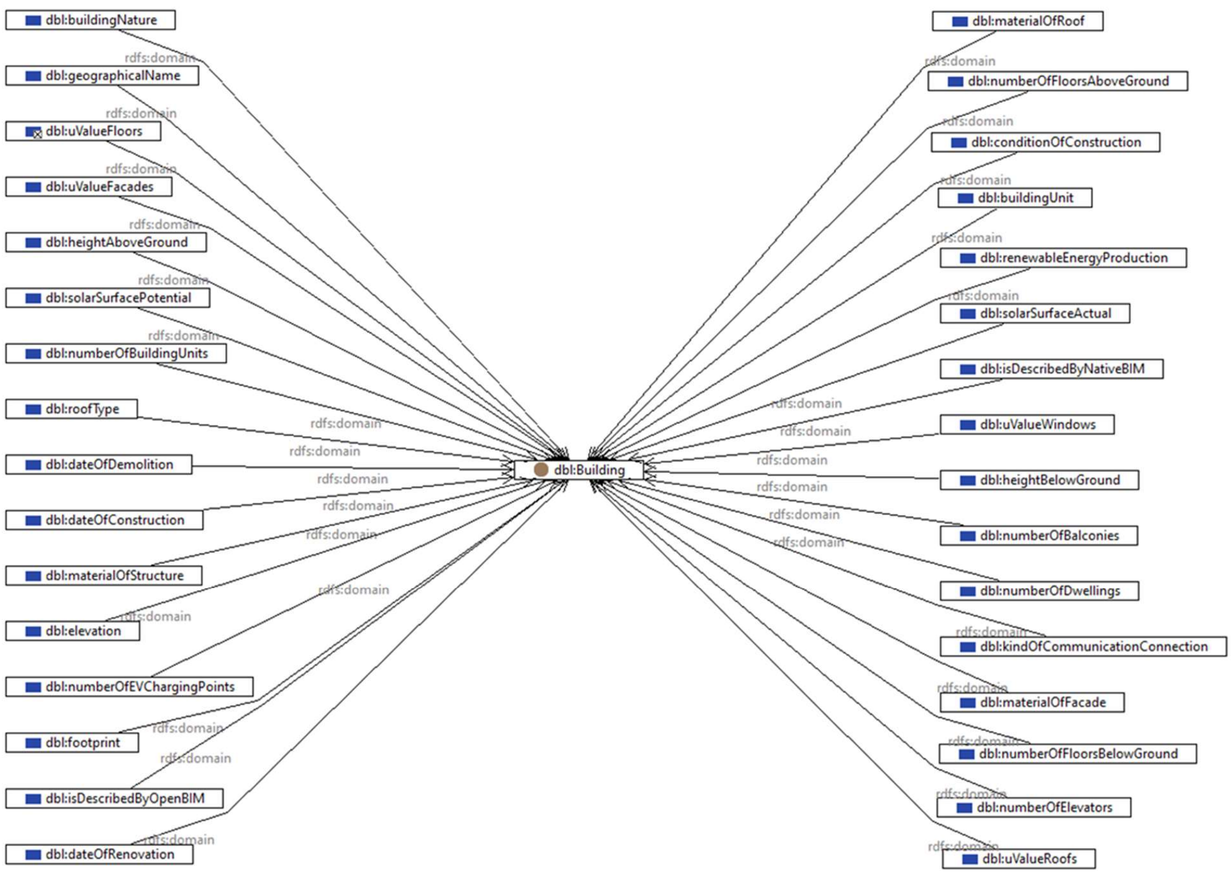
Below several diagrams are presented that are fully generated from the lexical DBL Ontology in linked data form (RDFS). The first diagram starts with the idea that there is a top-level ("root"), most generic, concept/class "DBL-Root". This root is subclassed into the three main DBL concepts: "Building", "BuildingUnit" and "CadastralParcel". For modelling efficiency, an intermediate concept "BuildingOrBuildingUnit" can collect properties relevant to both buildings and building units.

Figure 31: The top level of the DBL Ontology



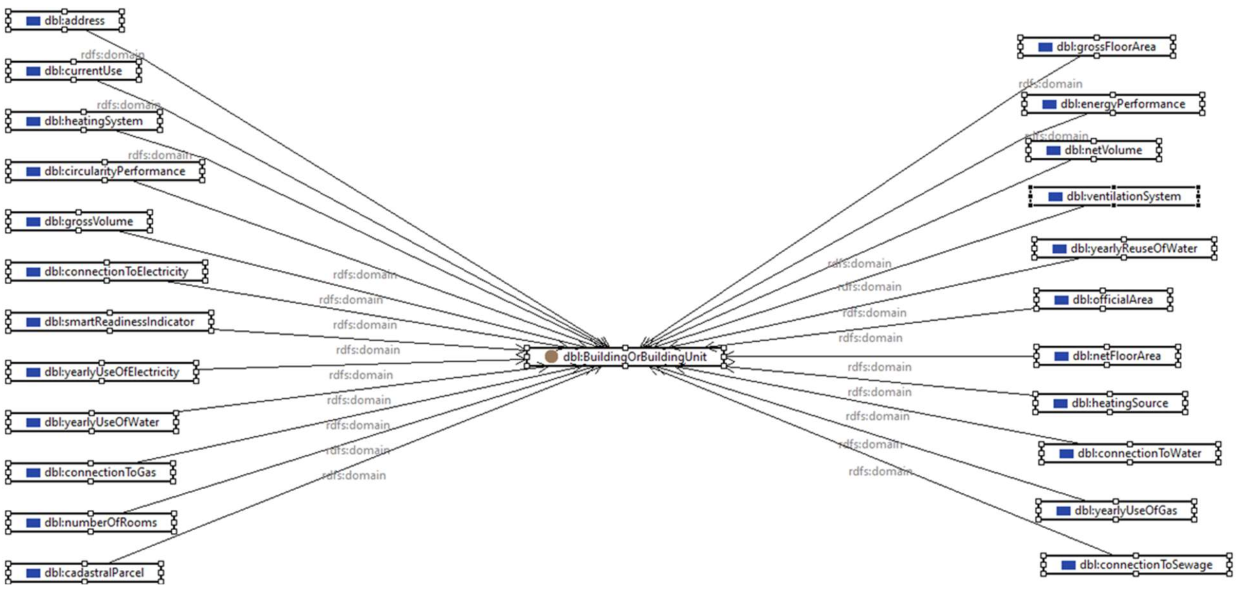
Next, a set of common properties (attributes and relations) is defined as having these concepts as a "domain". In the following figure, all (initial) common properties relevant to a building are depicted.

Figure 32: Properties for buildings



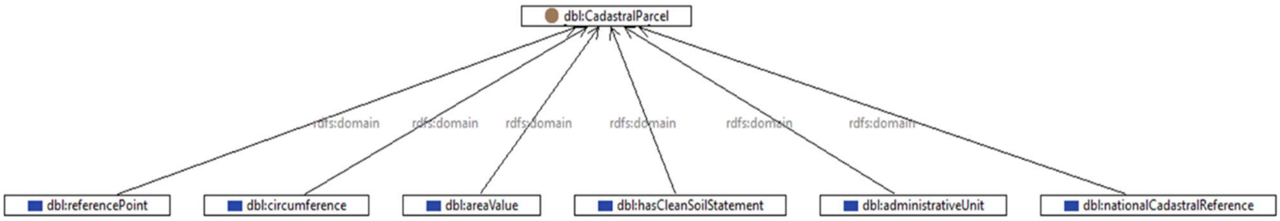
Note that all properties relevant for building units are also relevant for buildings. This means that we do not have a separate diagram for building units, but only for the higher level "BuildingOrBuildingUnit" concept.

Figure 33: Properties relevant for buildings and building units



Finally, we have a set of common properties for the concept of “CadastralParcel”.

Figure 34: Properties relevant for cadastral parcels

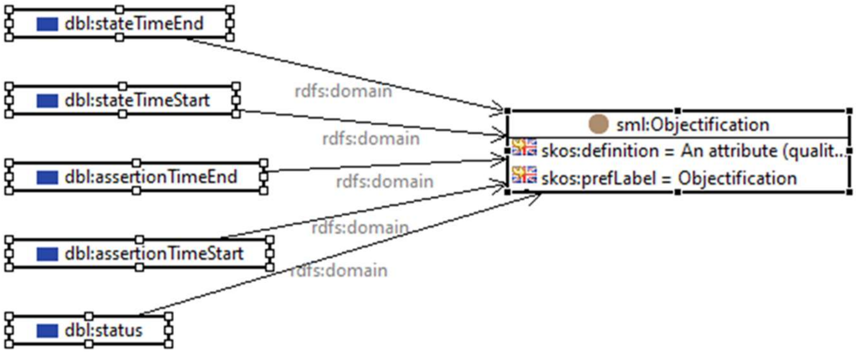


Most of these properties are attributes which are modelled as complex properties in the actual linked data code. While some of them are modelling interrelationships, such as:

- buildingUnit: from Building to BuildingUnit (many to many)
- cadastralParcel: from BuildingOrBuildingUnit to CadastralParcel (many to many)

Besides these ‘normal’ properties the following meta-properties are defined for the reused sml:Objectification (which is generally used to model complex properties):

Figure 35: Meta-properties (attributes) for objectifications used for complex property values



Annex II: Overview of other supporting documents

In this annex, you can find a short overview of other deliverables developed as part of this project and how they can guide you further in developing and implementing a DBL.

Table 13: List of relevant project deliverables

Report	Description
The DBL Semantic Data Model - Providing standard form and meaning to digital building logbook data.	This report defines the DBL Semantic Data Model and its underlying principles and technologies. This semantic data model is proposed as a European 'core model' for EU Member States and related data providers. It provides agreed common meaning and terminology for the various relevant logbook data sets.
DBL Semantic Data Model, Linked Data implementation – In RDFS and SKOS code.	This report describes the implementation (i.e. the "coding") of the core DBL ontology and dictionary in W3C Linked Data technology (formats and languages). It builds upon the conceptual description of the DBL Semantic Data Model and includes the core DBL ontology.
Key Data Sets and Functionalities – Relevant for the DBL Semantic Data Model.	This report confronts the DBL 'core model' with some key available (public) building data sets and their related software functionalities. The focus is on the most basic data that is in principle relevant for all EU Member States and that is the bare minimum a DBL should have content for.
DBL Use Cases – Showcasing how DBLS would work in practice.	This report provides a more in-depth description of how to implement a use case for DBLs than Section 2.1. It also presents more detailed findings on the use case analysis done for this project across the five main user groups.
DBL Data Management Plan – For the collection, storing and provision of DBL-relevant data.	A short report on the use of a data management plan to describe how DBL data over its life cycle is collected, stored and provided. It has mostly been integrated into Section 3.3 of these guidelines.
Development of databases – The expected potential for DBL in the near future.	A short report on digitalisation in the built environment and the increased availability of building-related data with a particular focus on BIM and GIS data.

Annex III: Data Management Plan Template

Below you can find a simple template to follow when developing a data management plan. This template builds on the descriptions in section 3.5. The template is still currently being developed.

Data Management aspect	Guiding questions	Further inputs
Purpose of data	• What is the purpose of the data collection/generation and its relation to the objectives of the DBL? • What types and formats of data will the DBL generate/collect? • To whom might it be useful ('data utility')? • Will you re-use any existing data and how?	Use case and customer journey analysis (2.1)
Collection of data		
Availability of data	• What is the origin of the data? • Is the data findable and available? • Is the data open or restricted? • Is the data free or paid? • Is the data relevant to the intended process?	Property set for data needs (2.2); Findability in the cloud
How to process and use data	• Is restricted and/or paid data securely accessible (downloadable or direct access) via some agreed trust/pay framework? • Is the data well-defined by available and sufficiently clear structural metadata? • Is the data published under the correct license for reuse? • Is the data complete, sufficiently precise, accurate and up to date?	
Storing of data		
Storage of data	• Is the data safely stored in certified repositories for long-term preservation and curation?	Cloud storage and DBL Framework alignment
Quality of data	• Is the data well-defined by available and sufficiently clear structural metadata? • Is the data in the correct format, or is translation needed? • Is the data in the right (common, agreed) semantics, or is conversion needed?	Alignment with DBL Framework: data formats (Turtle, JSON-LD), query language (SPARQL), data language (RDFS/SHACL)

	• Are data quality assurance processes described?	
The organisation of data		
Provision of data (publishing)		
(Cyber)security & privacy of personal data	• What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?	Technical interoperability (3.4) and Legal implementation (Chp. 5)

Annex IV: Screened EU legislation and caselaw

For the legal risk screening, the following documents have been considered, to ultimately select what the authors considered the main pieces of legislation for further analysis.

Nr	Access to data and use of data
1.1	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**), OJ L 119, 4.5.2016
1.2	Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the **legal protection of databases,** OJ L 77, 27.3.1996
1.3	Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (**Directive on privacy and electronic communications**) OJ L 201, 31.7.2002
1.4	Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the **free flow of non-personal data** in the European Union, PE/53/2018/REV/1
1.5	Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on **open data** and the re-use of public sector information, PE/28/2019/REV/1
1.6	Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (**Data Governance Act**), OJ L 152, 3.6.2022
1.7	**Proposal** for a Regulation of the European Parliament and of the Council on harmonised rules on fair access to and use of data (**Data Act**) of 23 February 2022 COM(2022)68, Factsheet of 23 February 2022 (19/07/2023, approval in committee of the text agreed at 1st reading interinstitutional negotiations. Awaiting Parliament's position in 1st reading)
1.8	Proposal for a Regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (**Regulation on Privacy and Electronic Communications**) COM/2017/010 final - 2017/03 (COD)
1.9	Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions “Digitising European Industry, Reaping the full benefits of a **Digital Single Market**” COM/2016/0180 final
1.10	Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions "**Building a European Data Economy**" COM/2017/09 final
1.11	Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions a **European strategy for data** COM/2020/66 final, Factsheet of February 2020

Nr	Access to data and use of data
1.12	**Data governance** and data policies at the European Commission, July 2020 Brochure
1.13	Communication Shaping Europe's **Digital Future**, Factsheet of 19 February 2020
1.14	Declaration 26 January 2022 on European **Digital Rights and Principles**
1.15	Commission Staff Working Document **Common European Data Spaces**, 23 February 2022, SWD(2022) 45 final
	Case law
1.16	Judgment of the Court (Grand Chamber) of 6 October 2015 in Case C-362/14: **Maximillian Schrems v Data Protection Commissioner** ECLI:EU:C:2015:650
1.17	Judgment of the Court (Second Chamber) of 19 October 2016 in Case C-582/14: **Patrick Breyer v Bundesrepublik Deutschland** ECLI:EU:C:2016:779
1.18	Judgment of the Court (Fifth Chamber) of 3 June 2021 in Case C-762/19: **SIA "CV-Online Latvia" v SIA "Melons"** ECLI:EU:C:2021:434
1.19	Judgment of the Court (Grand Chamber) of 16 July 2020 in Case C-311/18: **Data Protection Commissioner v Facebook Ireland Limited and Maximillian Schrems** ECLI:EU:C:2020:559
1.20	Judgment of the Court (Grand Chamber) of 2 March 2021 in Case C-746/18: **Criminal proceedings against H. K.** ECLI:EU:C:2021:152
1.21	Judgment of the Court (Grand Chamber) of 15 June 2021 in Case C-645/19: **Facebook Ireland Ltd and Others v Gegevensbeschermingsautoriteit** ECLI:EU:C:2021:483
1.22	Judgment of the Court (Fifth Chamber) of 17 June 2021 in Case C-597/19: **Mircom International Content Management & Consulting (M.I.C.M.) Limited v Telenet BVBA** ECLI:EU:C:2021:492
1.23	Judgment of the Court (Grand Chamber) of 22 June 2021 in Case C-439/19: **Proceedings brought by B** ECLI:EU:C:2021:504
1.24	Judgment of the Court (Third Chamber) of 25 November 2021 in Case C-102/20: **StWL Städtische Werke Lauf a.d. Pegnitz GmbH v eprimo GmbH** ECLI:EU:C:2021:954
1.25	Judgment of the Court (First Chamber) of 24 March 2022 in Case C-245/20: **X and Z v Autoriteit Persoonsgegevens** ECLI:EU:C:2022:216
1.26	Request for a preliminary ruling of 14 January 2022 in Case C-33/22: **Datenschutzbehörde (Austria)** Document 62022CN0033 (questions referred) Update: opinion Adv.Gen.of 11 May 2023, ECLI:EU:C:2023:397

Source: Ulrich Paetzold, EU-Consulting (internal document)

Annex V: Relevant national datasets

Existing national datasets can provide part of the input for national DBLs. This annex provides an overview of relevant national datasets. The following five types of datasets are considered:

- General building information
- Building monuments
- Building installations
- Energy efficiency labels
- Property valuation

Relevant datasets could not be identified for all types of datasets in all countries. The table below lists the relevant datasets that were identified per Member State.

MS	Type	Link and description
AT	General	https://www.data.gv.at/suche/ ; various building specific data in different files and formats (e.g. type Gebäude in the search filter)
AT	General	For example, https://www.data.gv.at/katalog/dataset/gebaeudeinformation-wien, City of Vienna, various public buildings: select downloads and then e.g. WFS GetFeature (CSV); data on address, construction year, architect, intended use, type of building, floors [roof floor, cellar, basement], and whether parcel structure, facade, core, windows and roof are as originally built
AT	Monuments	https://de.wikipedia.org/wiki/Portal:Österreichische_Denkmallisten ; Portal to municipal lists of building monuments per region
AT	Installations	https://www.statistik.at/atlas/?mapid=them_energie_klimafonds ; Number of solar panel installations per 1000 residents in six categories, per municipality
AT	Valuations	https://www.statistik.at/statistiken/volkswirtschaft-und-oeffentliche-finanzen/preise-und-preisindizes/immobilien-durchschnittspreise ; Average building plot price per m2 in five value classes, per municipality
BE	General	https://woningpas.vlaanderen.be/ ; Flanders, only accessible to owners of the building and persons to whom they share access. Includes amongst others information on year of construction, area and volume, energy performance certificate, various installations, information on soil pollution attest, information on asbestos attest, building quality conformity attest, information on completed renovations entered by the building owner, heritage value of the building, tool to assess solar panel potential, energy performance calculation tool, building quality assessment tool, checktool for soil pollution
BE	General	https://geoportail.wallonie.be/walonmap ; Wallonia: altitude and elevation. After checking "Catalogue du Géoportal", additional information on purpose of the building, nature environment, applicable regulations and risks (floods, nitrate pollution, erosion by water), renovation plans, availability for rent

MS	Type	Link and description
BE	General	https://www.totem-building.be/ ; Only accessible to authorised professional users (click on Login). Tool to assess the impact of building materials used in a building on the environment
BE	Monuments	https://www.onroerenderfgoed.be/ ; Click on GeoPortal or on Inventory. Heritage data on archeological, architectural and environmental monuments
BE	Energy efficiency	https://authenticatie.vlaanderen.be/stb/html/ssologin ; Flanders, only accessible to owners of the building. Includes energy performance data
BE	Energy efficiency	https://energie.wallonie.be/fr/acces-a-la-base-de-donnees-peb.html?IDC=9593 ; Wallonia, only accessible to owners of the building or after request. The webpage http://peb.energie.wallonie.be/bddpeb may take long to load
BE	Valuations	See BE General
BG	General	https://portal.registryagency.bg/pr/Services/252 ; Open to anyone but need to apply and access is registered. People can request acts and deeds on buildings and building certificates.
BG	Monuments	http://registersofia.bg/index.php ; select Декоративен елемент = Decorative elements ; Sofia city: list of buildings with decorative elements
CY	---	No relevant datasets in the five categories were identified
CZ	General	https://www.cuzk.cz/Katastr-nemovitosti/Poskytovani-udaju-z-KN/Dalkovy-pristup/Vystupy-z-KN-poskytovane-prostrednictvim-DP.aspx ; Cadastral extract, with mainly m2 per owner, names of building owners & information about apartments and non-residential premises (units): nr of units) (50 CZK per page = approximately EUR
CZ	Monuments	https://pamatkovykatalog.cz/uskp ; Central list of cultural monuments ; Architectual monuments
CZ	Installations	https://fdrive.cz/mapa-nabijecich-stanic ; Map of vehicle charging stations
DE	General	http://www.grundbuch-portal.de/stufe1-by.htm ; Only available to users who can prove secure data management (typical participants are notaries, banks and authorities). National portal to regional land registers
DE	Monuments	http://www.denkmalliste.org/denkmallisten.html ; Portal to regional lists of building monuments
DK	General	https://bbr.dk/se-bbr-oplysninger ; Year of construction, last renovation or addition year, outer wall materials, roofing materials, the building's areas and floors, heat and energy installations, use limitations, number and condition of toilets, bathrooms, kitchen
DK	Monuments	https://www.kulturarv.dk/fbb/index.htm ; FBB contains information about the country's approx. 7100 listed buildings and about the approx. 370,000 buildings whose conservation value has been assessed. In addition, FBB contains basic information about the country's more than 4

MS	Type	Link and description
		million buildings. This information originates from the Building and Housing Register (BBR) , from which it is updated automatically.
DK	Installations	See DK General
EE	General	https://livekluster.ehr.ee/ui/ehr/v1/infoportal/buildingdata ; E.g., year of construction, residential use m2, non-residential use m2, nr floors above ground, nr underground floors, ground floor area, all-floor area, height, absolute height (from Amsterdam zero), length, width, depth (below ground floor, incl. depth from a well
EE	Monuments	https://register.muinas.ee/public.php?menuID=monument ; Search tool for many types of cultural heritage buildings
EE	Installations	For example: https://livekluster.ehr.ee/ui/ehr/v1/building/120704140 ; Technical indicators of individual buildings, such as type of ventilation or energy source (under 'Ehitise tehnilised näitajad')
EE	Energy efficiency	For example: https://livekluster.ehr.ee/ui/ehr/v1/building/120704140 ; Energy label information per building (under 'Ehitise energiamärgised')
EE	Valuations	https://geoportaal.maaamet.ee/est/Ruumiandmed/Kinnisvara-tehingute-andmebaas-p81.html ; Accessible only to selected authoritised parties (land appraisers, statistical office, public research institutes, credit institutes). It has data on transaction values of land sales (not of building sales)
EE	Valuations	https://geoportaal.maaamet.ee/eng/Spatial-Data/Cadastral-Data-p310.html ; Taxable price of the land parcel calculated by the Land Registry in € (not of buildings). Download and extract an SHP file of county cadastral units, open spreadsheet software like Excel and open the *.dbf file from there
ES	General	https://www1.sedecatastro.gob.es/Cartografia/mapa.aspx?buscar=S ; Data on building surface area, cadastral parcel surface area, primary use, year of construction
ES	Monuments	https://www.culturaydeporte.gob.es/bienes/cargarFiltroBienesInmuebles.do?layout=bienesInmuebles&cache=init&language=es ; a list of monuments with registration codes and date of declaration
ES	Energy efficiency	https://edificioseficientes.gob.es ; Energy efficiency label, building surface area, date of expiration of label, links to renovation potential and potential financial support
FI	General	https://www.maanmittauslaitos.fi/rakennustietojen-kyselypalvelu ; Building Information Query Service (WFS), including info on floor area, volume, graduation date, apartments, apartment area , management basis (e.g. owner manages the apartment), use status of apartment (e.g. permanent residence or vacant), number of rooms, type of kitchen, equipment such as sauna, balcony etc.
FI	Monuments	https://kartta.paikkatietoikkuna.fi/?lang=en# ; Under Map Layers, click on Protected Sites - Protected Buildings. Only provides building ID, name and geocoordinates of the building.

MS	Type	Link and description
FI	Installations	https://dvv.fi/kiinteisto-rakennus-ja-paikkatiedot ; Population information system, one can find e.g. the heating method and fuel used of buildings and apartments.
FI	Energy efficiency	https://www.energiatodistusrekisteri.fi/ethaku ; Contains data on energy efficiency label and renewable energy production
FR	General	https://www.data.gouv.fr/fr/datasets/base-de-donnees-nationale-des-batiments/ ; Primary use, number of floors, number of apartments, residential surface, address data, year of construction, electricity and gas components, heating/cooling network components (this for less than 7,000 buildings), energy efficiency label prediction
FR	Monuments	https://www.data.gouv.fr/fr/datasets/immeubles-proteges-au-titre-des-monuments-historiques-2/ ; Data on address, building code, history of building, construction period, ownership status, date of protection, common name of the building, various administrative data
FR	Energy efficiency	https://data.ademe.fr/datasets/dpe-tertiaire ; Only offices/commercial buildings and with some missing data: energy efficiency inspection date; annual energy consumption ; energy consumption class ; estimated annual greenhouse gas emissions ; greenhouse gas emissions class ; type building ; sector of building user ; year of construction ; area for residential use ; thermal area of parcel ; souterrain Y/N ; number floors ; nr vertical circulations ; nr shops ; glass roof (Y / N) ; type of glass roof ; nr entrances with airlock ; nr entrances without airlock ; north-facing bay area Y/N ; west-facing bay area Y/N ; south-facing bay area Y/N ; surface of high floors for heat loss ; surface of low floors for heat loss ; surface of opaque vertical walls ; certification body ; date of DPE = energy performance diagnostic ; geo coordinates ; DPE model ; residential building Y/N ; type of building [shops, restaurants, hotels, public offices, other offices, utilities, housing, care homes, mixed use]; climate zone data)
GR	---	No relevant datasets in the five categories were identified
HR	General	https://dgu.gov.hr/pristup-informacijama/eu-fondovi/uspostava-e-registra-zgrada/5076 ; This website posts a notice that a building register is being developed
HR	Monuments	https://registar.kulturnadobra.hr/#/ ; The Register of Cultural Property of the Republic of Croatia website shows results for immovable and intangible cultural property.
HU	General	https://www.foldhivatal.hu/content/view/72/118/ ; Only owners of the land can apply for the data. Documents that they can view include the property sheet I (area of plot and legal nature, e.g.: monument) ; register of documents (floor plan)
HU	Monuments	See HU General
IE	Monuments	https://data.gov.ie/dataset/national-monuments-service-monuments-to-visit?package_type=dataset ; The 'National Monuments Service - Monuments

MS	Type	Link and description
		to Visit' pinpoints national monuments around the country which are accessible to the public.
IE	Energy efficiency	https://data.gov.ie/dataset/eba02-domestic-building-energy-rating?package_type=dataset ; Datasets including Domestic Building Energy Ratings.
IE	Valuations	https://opendata.housing.gov.ie/dataset/price-of-new-property-by-area-by-year ; Annual New Property prices by city by year
IT	General	https://www.catastosemplice.it/visure-catasto/visura-catastale-online/ ; The Urban Building Cadastre is the place where the data relating to all public or private real estate (civil, industrial and commercial buildings) present in a territory are archived. It includes info such as the exact geographical location; the size of the property; the intended use; land tax value.
IT	Monuments	https://www.monumentinazionali.it/inventario/inventario.htm ; General inventory of monuments present, subdivided by region.
IT	Valuations	https://www.agenziaentrate.gov.it/portale/web/guest/area-riservata ; A Spid , National Service Card , Electronic Identity Card and Revenue Agency credentials are needed to access building tax data on a building one owns
LT	General	https://www.registrucentras.lt/ntr/ ; Public data include only an address. Residents can also access data on tax valuation, energy efficiency certification, and mortgage data
LT	Monuments	https://www.geoportal.lt/geoportal/en/web/en/open-data-lt#savedSearchId={AF329EFA-A99D-40F3-B7DA-D98FBC7E73AF}&collapsed=true ; Datasets including those for museums, religious heritage and 'Data of the Tourist map' which includes UNESCO sites, museums, galleries, monuments, palaces, hillforts, observation decks, etc.)
LT	Energy efficiency	See LT General
LT	Valuations	See LT General
LU	Monuments	https://inpa.public.lu/fr/restauration/patrimoine_protege/liste-actualisee.html ; Address and date of declaration
LV	General	https://www.kadastrs.lv/ ; Select "portālā". Cadastral data available with a fee or free for owners with an account). It includes access to building floor plans and room group plans
LV	Monuments	https://mantojums.lv/cultural-objects ; List of cultural monuments
LV	Monuments	https://geolatvija.lv/geo/p/613 ; Register of cultural monuments (map search)
LV	Valuations	See LV General
MT	---	No relevant datasets in the five categories were identified

MS	Type	Link and description
NL	General	https://vastgoedloep.nl/ ; Contains data on primary purpose, surface area, year of construction, protected sight status, district living quality
NL	Monuments	https://www.monumenten.nl/monumenten?size=n_12_n ; Central government monuments register (yes/no indicator and short description of the monumental feature)
NL	Energy efficiency	https://www.energielabel.nl/woningen/zoek-je-energielabel/ ; Inventory to find energy label by zip code
NL	Valuations	www.woz-waardeloket.nl ; Real estate value of a building as assessed by the municipality
NL	Valuations	https://www.kadasterdata.nl/zakelijk ; Info available for purchase prices, property value etc. Need to pay for data
PL	General	https://mapy.geoportal.gov.pl/imap/Imgp_2.html ; It includes potentially interesting map layers such as real estate price (of properties) and existence of asbestos. Real estate prices are available for a limited number of counties: https://www.geoportal.gov.pl/o-geoportalu/aktualnosci/-/asset_publisher/HCHq0YGNRszn/content/29-10-2019-nowa-usluga-gugik-dotyczaca-cen-nieruchomosci for real estate. Asbestos data are published in the form of asbestos clearing activities at the level of municipalities: https://bazaazbestowa.gov.pl/pl/
PL	Monuments	https://nimoz.pl/baza-wiedzy/bezpieczenstwo-zbiorow-i-ludzi/ochrona-przeciwpozarowa/wykaz-muzeow-i-zabytkow-budowlanych-w-ktorych-wymagane-jest-stosowanie-systemu-sygnalizacji-pozarowej ; A list of names of buildings and address data of building monuments that require a fire alarm system
PT	Monuments	https://www.patrimoniocultural.gov.pt/en/patrimonio/patrimonio-imovel/pesquisa-do-patrimonio/classificado-ou-em-vias-de-classificacao/geral/result/?name=&situation=321895&catprot=&invtema=&type=&concelho=&records=100 ; List of monuments, sites etc.
RO	Monuments	https://patrimoniu.ro/ro//articles/lista-monumentelor-istorice ; List of historical monuments by county (basic information available)
SE	General	https://www.lantmateriet.se/en/real-property/my-property-/the-service-my-property/#anchor-0 ; With Swedish bank ID data, people can access the following standard data: who the legal owner of the property is, mortgages and tax information as well as information about community facilities and communities that are connected to the property. Owners of a building can view additional data such as property map (if there is one scanned), area, boundaries, and tax assessment value.
SE	Monuments	Search function for national monuments
SE	Valuations	See SE General
SI	General	https://ipi.eprostor.gov.si/jv/ ; Select the desired information under Sloji = Layers. After selecting an address the following pop up: gross floor space, nr

MS	Type	Link and description
		floors, nr apartments, nr business premises, building type (e.g. detached building), lowest and highest elevation angle, building height, year of construction, net floor area, living space, area of open terrace/balcony/loggia, floor plan, …
SI	Monuments	http://giskd2s.situla.org/rkd/Zacetek.asp ; The Register of cultural heritage consists of: register of immovable cultural heritage, register of intangible cultural heritage and register of movable cultural heritage.
SI	Monuments	https://gisportal.gov.si/portal/apps/webappviewer/index.html?id=df5b0c8a300145fda417eda6b0c2b52b ; Map search of building monuments
SI	Valuations	https://eprostor.gov.si/EV_EMV/emv/EMV.html ; Insight into the record of valuation models (select a region, and for that region download a pdf file with a table of tax value by floor area, plot area and year of construction)
SI	Valuations	https://eprostor.gov.si/ETN-JV/ ; Purchase and sales transactions (need to select an area, type of building and price value, and then relevant addresses pop up)
SK	General	https://data.gov.sk/sk/dataset/register-adries-register-budov ; Data on whether the building contains flats (Y/N) and the primary purpose of the building (often missing)
SK	General	https://zbgis.skgeodesy.sk/mkzbgis/sk/kataster/detail/kataster/parcela-c/852228/1097_6?bm=zbgis&pos=48.713652,20.130300,20 ; After searching for a building click on "Copy from the cadastral map". A pdf of a document appears which includes information on the parcel: register number, area in m2, type of property, method of land use, type of protected real estate, shared property status, land location, type of legal relationship. It includes the following information on the building: register number, type of building, description of the building (e.g., family house), type of protected real estate, location of the building, identification data of: the owners, the co-owned shares and the reference number of the ownership document; and identification data on the administrator, the tenant if not the owner, and other other authorised persons (whenever relevant)
SK	Monuments	https://www.culture.gov.sk/wp-content/uploads/2021/04/Zoznam_NKP_s_prioritou_ochrany_31.12.2020.pdf ; A list of municipalities, names of buildings and building IDs, by type of monument
SK	Energy efficiency	https://data.gov.sk/sk/dataset/centralny-register-energetickych-certifikatov-budov ; Central register of energy certificates

DBL

DIGITAL BUILDING LOGBOOK

LinkedIn | @EU_Growth

European Commission