Downloaded from orbit.dtu.dk on: Apr 27, 2024

DTU Library

A Data Sharing Guideline for Buildings and HVAC Systems

Energy in Buildings and Communities Technology Collaboration Programme

White, Stephen M.; Marszal-Pomianowska, Anna Joanna; Jin, Guang Yu ; Madsen, Henrik; Candanedo, José; Harmelink, Mirjam ; Stensson, Sofia ; Gori, Virginia

Publication date:
2023

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
White, S. M., Marszal-Pomianowska, A. J., Jin, G. Y., Madsen, H., Candanedo, J., Harmelink, M., Stensson, S., & Gori, V. (2023). *A Data Sharing Guideline for Buildings and HVAC Systems: Energy in Buildings and Communities Technology Collaboration Programme*. International Energy Agency.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

ANNEX 81

EBC
Energy in Building and Communities Programme

International Energy Agency

A Data Sharing Guideline for Buildings and HVAC Systems

Energy in Buildings and Communities Technology Collaboration Programme

March 2023

ANNEX 81

EBC
Energy in Building and Communities Programme

International Energy Agency

A Data Sharing Guideline for Buildings and HVAC Systems

Energy in Buildings and Communities Technology Collaboration Programme

March 2023

Authors

Stephen White, CSIRO, Australia (stephen.d.white@csiro.au)
Anna Marszal-Pomianowska, Aalborg University, Denmark
Guang Yu Jin, Building Construction Authority, Singapore
Henrik Madsen, Danish Technical University, Denmark
José Candanedo, University of Sherbrooke, Canada
Mirjam Harmelink, Technical University Delft, Netherlands
Sofia Stensson, RISE, Sweden
Virginia Gori, University College London, UK

© Copyright CSIRO 2023

All property rights, including copyright, are vested in CSIRO, Operating Agent for EBC Annex 81, on behalf of the Contracting Parties of the International Energy Agency (IEA) Implementing Agreement for a Programme of Research and Development on Energy in Buildings and Communities (EBC). In particular, no part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of CSIRO.

Published by CSIRO, PO Box 330, Newcastle, NSW 2300, Australia

Disclaimer Notice: This publication has been compiled with reasonable skill and care. However, neither CSIRO, nor the Contracting Parties of the International Energy Agency's Implementing Agreement for a Programme of Research and Development on Energy in Buildings and Communities, nor their agents, make any representation as to the adequacy or accuracy of the information contained herein, or as to its suitability for any particular application, and accept no responsibility or liability arising out of the use of this publication. The information contained herein does not supersede the requirements given in any national codes, regulations or standards, and should not be regarded as a substitute for the need to obtain specific professional advice for any particular application. EBC is a Technology Collaboration Programme (TCP) of the IEA. Views, findings and publications of the EBC TCP do not necessarily represent the views or policies of the IEA Secretariat or of all its individual member countries.

ISBN (13-digit)

Participating countries in the EBC TCP: Australia, Austria, Belgium, Brazil, Canada, P.R. China, Czech Republic, Denmark, Finland, France, Germany, Ireland, Italy, Japan, Republic of Korea, the Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, Turkey, United Kingdom and the United States of America.

Additional copies of this report may be obtained from: EBC Executive Committee Support Services Unit (ESSU), C/o AECOM Ltd, The Colmore Building, Colmore Circus Queensway, Birmingham B4 6AT, United Kingdom
www.iea-ebc.org
essu@iea-ebc.org

Preface

The International Energy Agency

The International Energy Agency (IEA) was established in 1974 within the framework of the Organisation for Economic Co-operation and Development (OECD) to implement an international energy programme. A basic aim of the IEA is to foster international co-operation among the 30 IEA participating countries and to increase energy security through energy research, development and demonstration in the fields of technologies for energy efficiency and renewable energy sources.

The IEA Energy in Buildings and Communities Programme

The IEA co-ordinates international energy research and development (R&D) activities through a comprehensive portfolio of Technology Collaboration Programmes (TCPs). The mission of the IEA Energy in Buildings and Communities (IEA EBC) TCP is to support the acceleration of the transformation of the built environment towards more energy efficient and sustainable buildings and communities, by the development and dissemination of knowledge, technologies and processes and other solutions through international collaborative research and open innovation. (Until 2013, the IEA EBC Programme was known as the IEA Energy Conservation in Buildings and Community Systems Programme, ECBCS.)

The high priority research themes in the EBC Strategic Plan 2019-2024 are based on research drivers, national programmes within the EBC participating countries, the Future Buildings Forum (FBF) Think Tank Workshop held in Singapore in October 2017 and a Strategy Planning Workshop held at the EBC Executive Committee Meeting in November 2017. The research themes represent a collective input of the Executive Committee members and Operating Agents to exploit technological and other opportunities to save energy in the buildings sector, and to remove technical obstacles to market penetration of new energy technologies, systems and processes. Future EBC collaborative research and innovation work should have its focus on these themes.

At the Strategy Planning Workshop in 2017, some 40 research themes were developed. From those 40 themes, 10 themes of special high priority have been extracted, taking into consideration a score that was given to each theme at the workshop. The 10 high priority themes can be separated in two types namely 'Objectives' and 'Means'. These two groups are distinguished for a better understanding of the different themes.

Objectives - The strategic objectives of the EBC TCP are as follows:

- reinforcing the technical and economic basis for refurbishment of existing buildings, including financing, engagement of stakeholders and promotion of co-benefits;
- improvement of planning, construction and management processes to reduce the performance gap between design stage assessments and real-world operation;
- the creation of 'low tech', robust and affordable technologies;
- the further development of energy efficient cooling in hot and humid, or dry climates, avoiding mechanical cooling if possible;
- the creation of holistic solution sets for district level systems taking into account energy grids, overall performance, business models, engagement of stakeholders, and transport energy system implications.

Means - The strategic objectives of the EBC TCP will be achieved by the means listed below:

- the creation of tools for supporting design and construction through to operations and maintenance, including building energy standards and life cycle analysis (LCA);
- benefitting from 'living labs' to provide experience of and overcome barriers to adoption of energy efficiency measures;
- improving smart control of building services technical installations, including occupant and operator interfaces;
- addressing data issues in buildings, including non-intrusive and secure data collection;
- the development of building information modelling (BIM) as a game changer, from design and construction through to operations and maintenance.

The themes in both groups can be the subject for new Annexes, but what distinguishes them is that the 'objectives' themes are final goals or solutions (or part of) for an energy efficient built environment, while the 'means' themes are instruments or enablers to reach such a goal. These themes are explained in more detail in the EBC Strategic Plan 2019-2024.

The Executive Committee

Overall control of the IEA EBC Programme is maintained by an Executive Committee, which not only monitors existing projects, but also identifies new strategic areas in which collaborative efforts may be beneficial. As the Programme is based on a contract with the IEA, the projects are legally established as Annexes to the IEA EBC Implementing Agreement. At the present time, the following projects

have been initiated by the IEA EBC Executive Committee, with completed projects identified by (*) and joint projects with the IEA Solar Heating and Cooling Technology Collaboration Programme by (☼):

Annex 1: Load Energy Determination of Buildings (*)
Annex 2: Ekistics and Advanced Community Energy Systems (*)
Annex 3: Energy Conservation in Residential Buildings (*)
Annex 4: Glasgow Commercial Building Monitoring (*)
Annex 5: Air Infiltration and Ventilation Centre
Annex 6: Energy Systems and Design of Communities (*)
Annex 7: Local Government Energy Planning (*)
Annex 8: Inhabitants Behaviour with Regard to Ventilation (*)
Annex 9: Minimum Ventilation Rates (*)
Annex 10: Building HVAC System Simulation (*)
Annex 11: Energy Auditing (*)
Annex 12: Windows and Fenestration (*)
Annex 13: Energy Management in Hospitals (*)
Annex 14: Condensation and Energy (*)
Annex 15: Energy Efficiency in Schools (*)
Annex 16: BEMS 1- User Interfaces and System Integration (*)
Annex 17: BEMS 2- Evaluation and Emulation Techniques (*)
Annex 18: Demand Controlled Ventilation Systems (*)
Annex 19: Low Slope Roof Systems (*)
Annex 20: Air Flow Patterns within Buildings (*)
Annex 21: Thermal Modelling (*)
Annex 22: Energy Efficient Communities (*)
Annex 23: Multi Zone Air Flow Modelling (COMIS) (*)
Annex 24: Heat, Air and Moisture Transfer in Envelopes (*)
Annex 25: Real time HVAC Simulation (*)
Annex 26: Energy Efficient Ventilation of Large Enclosures (*)
Annex 27: Evaluation and Demonstration of Domestic Ventilation Systems (*)
Annex 28: Low Energy Cooling Systems (*)
Annex 29: ☼ Daylight in Buildings (*)
Annex 30: Bringing Simulation to Application (*)
Annex 31: Energy-Related Environmental Impact of Buildings (*)
Annex 32: Integral Building Envelope Performance Assessment (*)
Annex 33: Advanced Local Energy Planning (*)
Annex 34: Computer-Aided Evaluation of HVAC System Performance (*)
Annex 35: Design of Energy Efficient Hybrid Ventilation (HYBVENT) (*)
Annex 36: Retrofitting of Educational Buildings (*)
Annex 37: Low Exergy Systems for Heating and Cooling of Buildings (LowEx) (*)
Annex 38: ☼ Solar Sustainable Housing (*)
Annex 39: High Performance Insulation Systems (*)
Annex 40: Building Commissioning to Improve Energy Performance (*)
Annex 41: Whole Building Heat, Air and Moisture Response (MOIST-ENG) (*)
Annex 42: The Simulation of Building-Integrated Fuel Cell and Other Cogeneration Systems (FC+COGEN-SIM) (*)
Annex 43: ☼ Testing and Validation of Building Energy Simulation Tools (*)
Annex 44: Integrating Environmentally Responsive Elements in Buildings (*)
Annex 45: Energy Efficient Electric Lighting for Buildings (*)
Annex 46: Holistic Assessment Tool-kit on Energy Efficient Retrofit Measures for Government Buildings (EnERGo) (*)
Annex 47: Cost-Effective Commissioning for Existing and Low Energy Buildings (*)
Annex 48: Heat Pumping and Reversible Air Conditioning (*)
Annex 49: Low Exergy Systems for High Performance Buildings and Communities (*)
Annex 50: Prefabricated Systems for Low Energy Renovation of Residential Buildings (*)
Annex 51: Energy Efficient Communities (*)
Annex 52: ☼ Towards Net Zero Energy Solar Buildings (*)
Annex 53: Total Energy Use in Buildings: Analysis and Evaluation Methods (*)
Annex 54: Integration of Micro-Generation and Related Energy Technologies in Buildings (*)
Annex 55: Reliability of Energy Efficient Building Retrofitting - Probability Assessment of Performance and Cost (RAP-RETRO) (*)
Annex 56: Cost Effective Energy and CO2 Emissions Optimization in Building Renovation (*)
Annex 57: Evaluation of Embodied Energy and CO2 Equivalent Emissions for Building Construction (*)

Annex 58: Reliable Building Energy Performance Characterisation Based on Full Scale Dynamic Measurements (*)
Annex 59: High Temperature Cooling and Low Temperature Heating in Buildings (*)
Annex 60: New Generation Computational Tools for Building and Community Energy Systems (*)
Annex 61: Business and Technical Concepts for Deep Energy Retrofit of Public Buildings (*)
Annex 62: Ventilative Cooling (*)
Annex 63: Implementation of Energy Strategies in Communities (*)
Annex 64: LowEx Communities - Optimised Performance of Energy Supply Systems with Exergy Principles (*)
Annex 65: Long-Term Performance of Super-Insulating Materials in Building Components and Systems (*)
Annex 66: Definition and Simulation of Occupant Behavior in Buildings (*)
Annex 67: Energy Flexible Buildings (*)
Annex 68: Indoor Air Quality Design and Control in Low Energy Residential Buildings (*)
Annex 69: Strategy and Practice of Adaptive Thermal Comfort in Low Energy Buildings
Annex 70: Energy Epidemiology: Analysis of Real Building Energy Use at Scale
Annex 71: Building Energy Performance Assessment Based on In-situ Measurements
Annex 72: Assessing Life Cycle Related Environmental Impacts Caused by Buildings
Annex 73: Towards Net Zero Energy Resilient Public Communities
Annex 74: Competition and Living Lab Platform
Annex 75: Cost-effective Building Renovation at District Level Combining Energy Efficiency and Renewables
Annex 76: ☼ Deep Renovation of Historic Buildings Towards Lowest Possible Energy Demand and CO_2 Emissions
Annex 77: ☼ Integrated Solutions for Daylight and Electric Lighting
Annex 78: Supplementing Ventilation with Gas-phase Air Cleaning, Implementation and Energy Implications
Annex 79: Occupant-Centric Building Design and Operation
Annex 80: Resilient Cooling
Annex 81: Data-Driven Smart Buildings
Annex 82: Energy Flexible Buildings Towards Resilient Low Carbon Energy Systems
Annex 83: Positive Energy Districts
Annex 84: Demand Management of Buildings in Thermal Networks
Annex 85: Indirect Evaporative Cooling
Annex 86: Energy Efficient Indoor Air Quality Management in Residential Buildings

Working Group - Energy Efficiency in Educational Buildings (*)
Working Group - Indicators of Energy Efficiency in Cold Climate Buildings (*)
Working Group - Annex 36 Extension: The Energy Concept Adviser (*)
Working Group - HVAC Energy Calculation Methodologies for Non-residential Buildings (*)
Working Group - Cities and Communities
Working Group - Building Energy Codes

Summary

Digitalisation offers new opportunities for saving energy in buildings. Digitalisation fundamentally takes a data-driven approach to the management and control of energy consuming equipment in buildings. This data-driven approach includes steps of (i) data capture, (ii) data management (iii) data analysis and (iv) data-driven decision implementation.

Hence, the benefits of digitalisation are somewhat predicated on access to relevant data, in a way that is cost-effective, trustworthy, flexible, and consistent with obligations to manage privacy and commercial rights.

The Energy Efficiency Hub Digitalisation Working Group (2022) identifies access to data, interoperability, and privacy as three of the most important barriers to the uptake of digitalisation.

This guide aims to provide building owners and policy makers with concepts and language to unlock this access to data. Some of the questions it aims to address include:

- How do I know that a given data stream can be relied on?
- What information/attributes does a data stream need to have, to enable machines (and humans) to process the data efficiently?
- Who owns the data and what permissions are needed to use it?
- Who should/shouldn't be able to access relevant data?
- What protections need to be implemented to control/manage access to data (including management of possible privacy issues)?
- How can all this data be safely consolidated in one place and then distributed to service providers in a way that is scalable and avoids expensive bespoke solutions?
- What tools and technology services are available, and what commercial and institutional arrangements support effective implementation of data mangement services?

This guide aims to provide building owners and government policy makers with concepts and language relevant to these questions, with a focus on deploying data management services for saving energy in building mechanical-services applications. It is a high-level general-principles document. It is not intended to provide detailed engineering requirements for purchasing purposes, or legal advice on regulatory issues. The reader should seek their own independent advice.

The guide concurs on the fundamental importance of modern data management practices articulated in the FAIR data principles (Wilkinson, 2016). That is; data should be Findable, Accessible, Interoperable and Re-usable.

These principles highlight the need for interoperability, both at the device-level (through use of interoperable communication standards/protocols) and at the analytics-level (through use of semantic information standards/protocols). Rich meta-data, attached to each data stream, helps to provide explanatory context to better understand the identity and relevance of a given data stream. Use of meta-data schemas provides structure for storing data, streamlines extraction of data from storage for subsequent processing, and supports machines to automatically and logically 'reason with data'.

An appropriate licence should be obtained before using data. This can be complicated by the uncertainty associated with 'ownership' of data. While it is generally assumed that data belongs to the building owner, data has a practical tendency to find its way to service providers managing the data - and can therefore end up relatively inaccessible to the building owner.

While noting that much of the data used in non-residential building-services applications is unlikely to be highly sensitive, awareness and risk management strategies should be put in place for the appropriate treatment of

any personal data. The Australian Government's 'Best Practice Guide to Applying Data Sharing Principles' (Commonwealth of Australia, 2019a) describes five data sharing principles that could help to identify control strategies for satisfying obligations under relevant privacy legislation.

Data platforms (variously called IoT platforms, Energy Management Information Systems (EMIS), Distributed Energy Resource Management Systems (DERMS) - in different contexts/applications) provide cloud-based solutions for consolidating and processing relevant data. Users will typically need to comply with the data capture and data management standards employed by the platform. This may impact on the degree of interoperability that can be achieved.

Data platforms can empower building owners, by providing them with data sharing capability through role-based authentication and the ability for a designated administrator to grant/withdraw permissions to access specific data fields. This can help support enforcement of privacy restrictions and management of dataset ownership rights.

Unfortunately, data platforms are often relatively invisible infrastructure, to the purchaser of energy productivity services, because they are generally bundled up inside the provider's overall service. When multiple siloed services are deployed in a building, this may unknowingly lead to multiple data platforms being implemented in a building – an inefficient outcome resulting in higher costs for everyone.

An alternative approach is to separate (i) the procurement of data collection and data platform services from (ii) the procurement of energy productivity application services. This can reduce duplication of data management tasks, consolidate management of services, support data sovereignty and support competitive sourcing of third-party services.

There are various highly capable companies offering data platform services, typically using a Platform-as-a-Service (PaaS) business model. Noting the complexity of data governance, various government and co-operative models are also emerging to provide trustworthy data platform services. More broadly than just the buildings sector, the Open Data Institute has proposed the need for Data Institutions that steward data on behalf of others.

The guide concludes with a list of suggested topics and questions that prospective clients might consider as they begin a journey to source data platform services.

An important use-case, for the concepts discussed in this guide, is the establishment of an 'Energy Data Space', that provides an interoperability framework, governance requirements and software tooling for delivering demand flexibility as a distributed energy resource (as anticipated in the European 'Digitalisation of Energy Action Plan' (DoEAP)). This use case is a critical step toward enhancing reliability and affordability of future electricity grids. Attention to the principles in this guide, will help reduce industry fragmentation, provide building owners with equitable access to flexibility markets, and support increased competition.

Table of contents

Figures

Tables

Abbreviations

Abbreviations	Meaning
AI	Artificial Intelligence
AMQP	Advanced Message Queuing Protocol
API	Application Programming Interface
APP	Application
BIM	Building Information Modelling
BMS	Building Management System
CDR	Consumer Data Right
DER	Distributed Energy Resources
DERMS	Distributed Energy Resource Management System
DSO	Distribution Service Operator
EMIS	Energy Management Information System
EMS	Energy Management System
FAIR	Findable, Accessible, Interoperable and Reusable
GDPR	General Data Protection Regulation
GEB	Grid Interactive Efficient Building
HVAC	Heating Ventilating and Air Conditioning
ICT	Information and Communication Technology
IEA	International Energy Agency
IEQ	Indoor Environment Quality
IoT	Internet of Things
ISO	International Organization for Standardization
IT	Information Technology
LAN	Local Area Network
M&V	Measurement and Verification
MQTT	Message Queuing Telemetry Transport
OT	Operations Technology
PaaS	Platform as a Service
SFTP	Secure File Transfer Protocol
SSN	Semantic Sensor Network Ontology
TSO	Transmission Service Operator
VAV	Variable Air Volume

1. Introduction

The International Energy Agency (2017) explained the concept of Digitalisation as “the increasing interaction and convergence between the digital and physical worlds” where “the digital world has three fundamental elements:

- **Data**: digital information.
- **Analytics**: the use of data to produce useful information and insights.
- **Connectivity**: exchange of data between humans, devices and machines (including machine-to-machine), through digital communications networks.

The trend toward greater digitalisation is enabled by advances in all three of these areas; increasing volumes of data thanks to the declining costs of sensors and data storage, rapid progress in advanced analytics and computing capabilities, and greater connectivity with faster and cheaper data transmission”.

Utilising these elements, smart autonomous systems are able to ‘reason with data’ and implement optimal decisions (in real-time) to streamline business processes and to manage the operation of physical assets – ultimately leading to improved energy productivity.

The International Energy Agency (2017) found that digitalisation could cut energy use by about 10% by using real-time data to improve operational efficiency. They also found that ‘smart demand response’ could provide 185 gigawatts (GW) of system flexibility, roughly equivalent to the currently installed electricity supply capacity of Australia and Italy combined.

Delivering both energy efficiency and demand flexibility services, the US Department of Energy (2021) identifies the need for ‘Grid-Interactive Efficient Buildings’ (GEBs), which could harness the power of digitalization to save the US power system USD 100-200 billion over the next two decades and help reduce CO_2 emissions by 80 million tonnes per year.

1.1 Digitalisation and the Need for Data Sharing

AlphaBeta (2018) explains digitalisation as an automated process from data to decisions. This process includes steps of (i) data capture, (ii) data management (iii) data analysis and (iv) decision and action. This process, along with a sample of the relevant digital technologies involved in each of these steps, is illustrated in Figure 1.1.

One of the exciting opportunities arising from digitalisation, is the ability to apply artificial intelligence to improve the operation of buildings. With sufficient ‘intelligence’, the claim is that physical assets in a building can autonomously select an informed course of action for achieving higher-level objectives (eg optimization of energy, IEQ, occupant experience etc).

Fundamental to this ability to apply artificial intelligence, is the ability to source quality data as inputs to the relevant algorithms. Data can come from equipment and sensors in the building, and from a wide range of other external sources.

Without this data (situational awareness) it is impossible to make informed decisions. However, it is a significant technical IT task to access, consolidate and validate these diverse data sources. Furthermore, there is growing awareness of both the commercial value of data and the potential for adverse outcomes from misuse of data. This requires attention to the legal and governance arrangements associated with managing data.

Digitalisation can be restricted to on-site data processing (‘at the edge’), or data can be consolidated in the cloud. This guide generally assumes at least some use of the cloud, to (1) access key external data sources and (2) to enable information to be efficiently distributed to relevant people via remote PC and mobile devices.

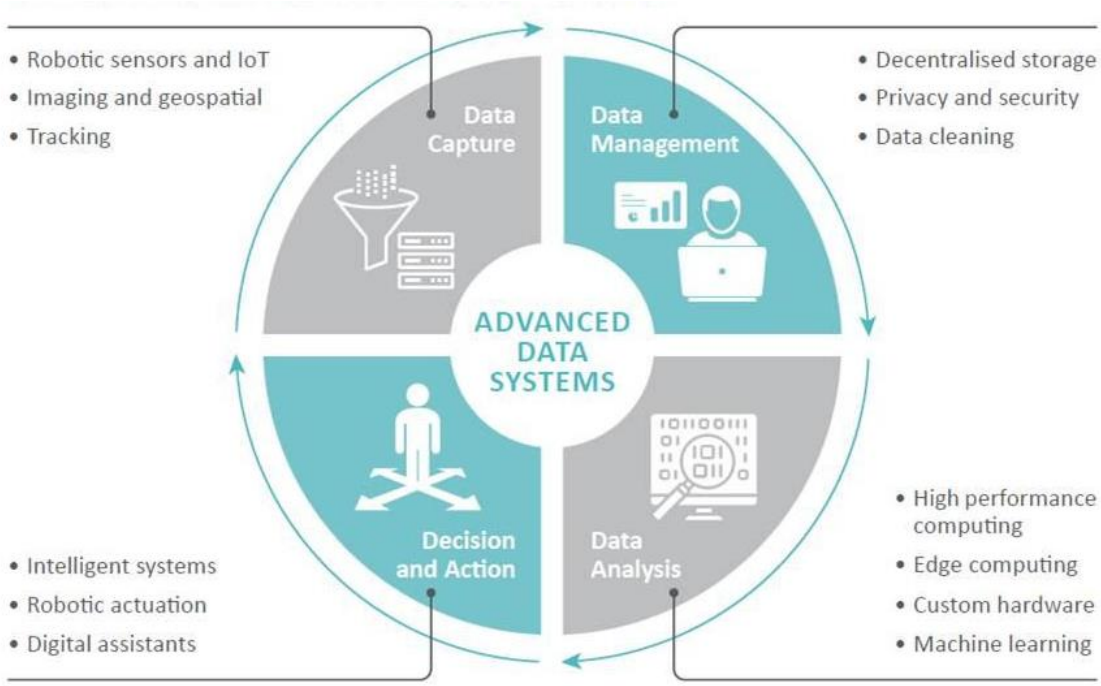

Figure 1.1: Data Innovation Relies on Specialised Systems for Data Capture, Management, Analysis and Action (Source: AlphaBeta, 2018)

1.1 The HVAC (Building Services) Data That Might be Shared

The focus of this guide is on digitalisation in non-residential buildings. Digitalisation of a building is assumed to result in a 'Smart Building'. The IEA Annex81 defines a 'Data-Driven Smart Building' as

> *'A building that uses digitalisation technologies to dynamically optimise its operation, where optimisation objectives typically relate to site energy use, IEQ, and occupant experience.*
>
> *Ideally, it is sufficiently connected and integrated with markets and processes, that it can adaptively respond to externalities and changing conditions (e.g. weather, electricity prices, energy supply constraints, equipment maintenance, etc). Ideally, it has sufficient memory of past events, and ability to anticipate future impacts, that it can select an informed course of action for achieving higher-level objectives – reminiscent of human intelligence.*
>
> *To achieve this vision, a Data-Driven Smart Building utilises live and historical data from relevant sensors, IoT equipment, mobile devices, and other sources to provide situational awareness for informed decision-making. Achieving the optimization objectives will often benefit from advanced supervisory-level automation, driven by computational analysis (eg Machine Learning, AI, etc) applied to available data.'*

Locatee and Memoori (2017) identify seven attributes of Smart Buildings (Figure 1.2). Each of these attributes provides the basis for a set of use-cases (applications), which can deliver tangible benefits in the form of (i) higher operational efficiency and resource utilisation, (ii) improved user experience and indoor environment for building occupants, (iii) connectivity between stakeholders and (iv) risk mitigation.

All of these applications will have different data needs, as inputs for automated decisions and actions. Of most interest to the International Energy Agency are the various energy productivity applications[1]. These applications are illustrated in Figure 1.3 (adapted from Kramer et al, 2020). They can reduce energy consumption in buildings by over 30% (Zhang et al, 2022).

[1] Energy productivity is a term used to include all the various forms of useful energy services – including energy efficiency, load shaping and flexible demand services

Figure 1.2: Seven Smart Buildings Attributes and Their Focus (Source: Locatee and Memoori, 2017)

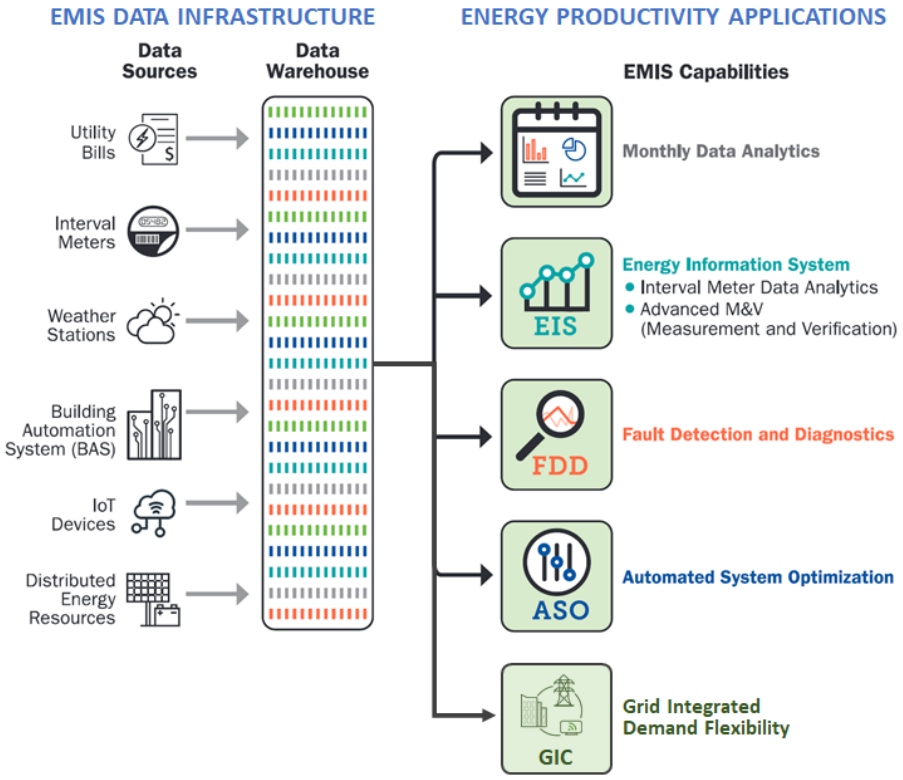

Figure 1.3: Typical energy productivity applications that can be hosted on a data platform (adapted from Kramer et al, 2020)

Some data sources (not exhaustive) that could be used as inputs for these applications are listed in Table 1.1. As indicated in the IEA Annex81 definition of a Data-Driven Smart Building, both current and historic data is required for each data source, in order to identify trends and deploy computational analytics.

Table 1.1: Data that is (or can be) used by energy analytics software services

Data Type	Data Points
Energy metering	• Site billing meter • Equipment submeters • Renewable energy source

Base building equipment	• HVAC equipment operational data from BMS
	  - Chillers
	  - Fans and pumps
	  - Cooling towers
	  - Valves and dampers
	• Lighting
	• Vertical transport

Indoor conditions	• Zone temperature and humidity • Supply and return air temperature • CO_2 level • VAV box/ diffusers
External conditions	• Outdoor temperature and humidity • Energy source price and carbon intensity
Occupant	• Occupant presence or count in rooms • Comfort settings and preferences
Distributed energy resources	• Battery charge • Electric vehicle charging • Standby gensets

1.2 This Guide

Quick inspection of these data sources and data requirements highlight a range of questions that need to be addressed before the benefits of digitalisation can be realised in buildings. Some of these questions include;

- How do I know that a given data stream can be relied on?
- What information/attributes does a data stream need to have, to enable machines (and humans) to process the data efficiently?
- Who owns the data and what permissions are needed to use it?
- Who should/shouldn't be able to access relevant data?
- What protections need to be implemented to control/manage access to data (including management of possible privacy issues)?
- How can all this data be safely consolidated in one place and then distributed to service providers in a way that is scalable and avoids expensive bespoke solutions?
- What tools and technology services are available, and what commercial and institutional arrangements support effective implementation of data mangement services?

This guide addresses these questions, with the aim of providing building owners and government policy makers with concepts and language relevant to data-sharing – in the context of building mechanical services. Ideally this guide will be sufficient to begin the process of choosing digitalisation tools and energy analytics solutions for the purpose of enhancing energy productivity in non-residential buildings.

The focus is on streamlining real-time data exchange between both humans and machines. It includes consideration of both technical standards for data management and the ethical and governance aspects of data sharing. It aims to provide more clarity on the business models and policy principles that can support improved scalability for the energy analytics and HVAC controls industry.

While of great importance, cyber security considerations are touched on but are generally outside the scope of this guide. For more detailed information on information security management systems (ISMS) and their requirements, the reader is referred to the ISO/IEC 27000 family of standards.

Similarly, the guide is not intended to be an engineering manual for IT/communications networking in buildings.

The guide is a high-level, general-principles document. It does not constitute legal advice about how an entity should comply with regulatory requirements relating to data sharing. The reader should seek their own independent legal advice where appropriate.

2. Data and Data Management Approaches

The information and communications technology (ICT) industry hosts and processes vast amounts of data, driving productivity in all manner of business activity. Technologies and open standards have been developed to streamline this work. The building mechanical services industry has been a relatively late adopter of these practices.

Some of the principles of modern data management practices are discussed in this chapter. A key aim is to embed the philosophy that underpins the FAIR data principles (Wilkinson, 2016), with specific emphasis on enhancing the ability of machines to automatically find and use data. The FAIR data principles (Findable, Accessible, Interoperable, Reusable) (Figure 2.1) have a goal that data can be discovered by potential future users and easily re-used either alone, or in combination with newly generated data.

Box 2 | The FAIR Guiding Principles

To be Findable:
F1. (meta)data are assigned a globally unique and persistent identifier
F2. data are described with rich metadata (defined by R1 below)
F3. metadata clearly and explicitly include the identifier of the data it describes
F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:
A1. (meta)data are retrievable by their identifier using a standardized communications protocol
A1.1 the protocol is open, free, and universally implementable
A1.2 the protocol allows for an authentication and authorization procedure, where necessary
A2. metadata are accessible, even when the data are no longer available

To be Interoperable:
I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
I2. (meta)data use vocabularies that follow FAIR principles
I3. (meta)data include qualified references to other (meta)data

To be Reusable:
R1. meta(data) are richly described with a plurality of accurate and relevant attributes
R1.1. (meta)data are released with a clear and accessible data usage license
R1.2. (meta)data are associated with detailed provenance
R1.3. (meta)data meet domain-relevant community standards

Figure 2.1: The FAIR guiding principles (Source: Wilkinson, 2016)

These principles underpin at least part of the pathway to a future industry, where deploying software becomes analogous to deploying 'Apps' on a mobile phone. Such a self-service, plug-and-play vision aims to reduce the cost and friction involved in implementing software services. It could also reduce barriers to entry for software developers, leading to more competition and more innovation.

2.1 Data Quality

It is apparent that data is at the heart of the digitalisation revolution. But the quality of any analysis or task is dependent on the quality of the data it uses.

At a basic level, data quality relates to trust that the respective sensors and devices are sending data, and that the data that is being sent correctly represents what it's purported to be. A wide range of issues can arise at this basic level. For example, sensors can go offline due to connectivity issues, sensors can fail leading to outlier readings, or sensors can just fail to update leading to a static signal. Procedures for data gap filling and data anomaly detection are required.

Higher level (more strategic) data quality issues relevant to the FAIR data principles include

- *Labelling and context*
    - *Data richness*: A simple unlabelled source of data is generally of limited value. Additional information (meta-data) on the source of the data, the physical meaning of the data, the units of measure, how the source of the data relates to other objects in its ecosystem, etc - all add context that can be used to infer causation of events and recommend remedial actions. By way of example; address-matching is often a means for linking records, so that analytics can be used to discover new correlations and drive administrative processes. In many digitalisation use-cases, time stamping is also required to ensure that diverse data sets can be correlated when examining time-bounded events.
    - *Ground truth*: Machine learning algorithms will often 'train' using 'ground-truth' data where the target event or condition is known to occur. After training, the algorithm is then able to detect the event/condition from other confounding factors. Memoori (2021) note that "*many AI models are trained through supervised-learning which requires data to be properly labelled and categorised*".

- *Structure and discoverability*: Memoori (2021) identifies that "*one of the key barriers to enabling widespread automation across the industry is that each building is unique*" and that "*developing sufficiently adaptable machine learning algorithms… is a huge challenge*". They also claim that "*storing, organising, structuring and labelling it in an appropriate fashion … can prove a challenge*". Metadata Schemas enable data to be structured and stored in a way that allows the logical relationships between data sources to be understood by machines. This can help to 'de-prototype' buildings from a software perspective. Databases can be queried, using an open database query language, to automatically discover and retrieve relevant data sources for subsequent processing by energy analytics software. The extent to which database structures can be aligned with industry-standard metadata schemas, will influence the ease with which third party software developers can access data. Furthermore, alignment and interoperability between different metadata schemas can support wider data discoverability across knowledge domains, opening up opportunities for innovative new smart building applications.

- *Provenance*: A data-source can be compromised in a range of ways. For example, technicians can inadvertently alter some hardware or software configuration, and fail to log changes. Secure digital identities may be appropriate for equipment assets to manage traceability and data provenance, helping to ensure that decisions are being made based on correctly identified and operational information. In relation to distributed energy resources (DER), the European Commission's Group of Experts on Energy Efficiency (2021) calls for the "*identity of equipment within buildings (as well as their associated rights) to be verified electronically in real time to enable decentralised units to dynamically respond to its environment and markets by sovereignly switching between modes of operation*".

- *Privacy*: Just because data is available does not mean that it can be ethically used. It is particularly important to note that relevant digitalisation use-cases may involve interaction with occupants and perhaps inadvertent collection of occupant personal data. Any restrictions on use of data should be clearly communicated and ideally controlled, to prevent misuse.

The European Commission's Group of Experts on Energy Efficiency (2021) identifies that "*privacy and cyber security considerations require management of data such that it is protected against unauthorised use (confidentiality), that it can be relied upon to be correct (integrity), and that it is accessible when required (availability)*"

While data-sources will rarely be perfect, there is value in ensuring that deviations from perfect are documented and communicated with the data, to avoid inappropriate use of the data.

O'Reilly (2020) surveyed 1,900 people working in the field of Artificial Intelligence, to get their perspectives on the data quality issues they face. A wide range of data quality issues were identified (Figure 2.2).

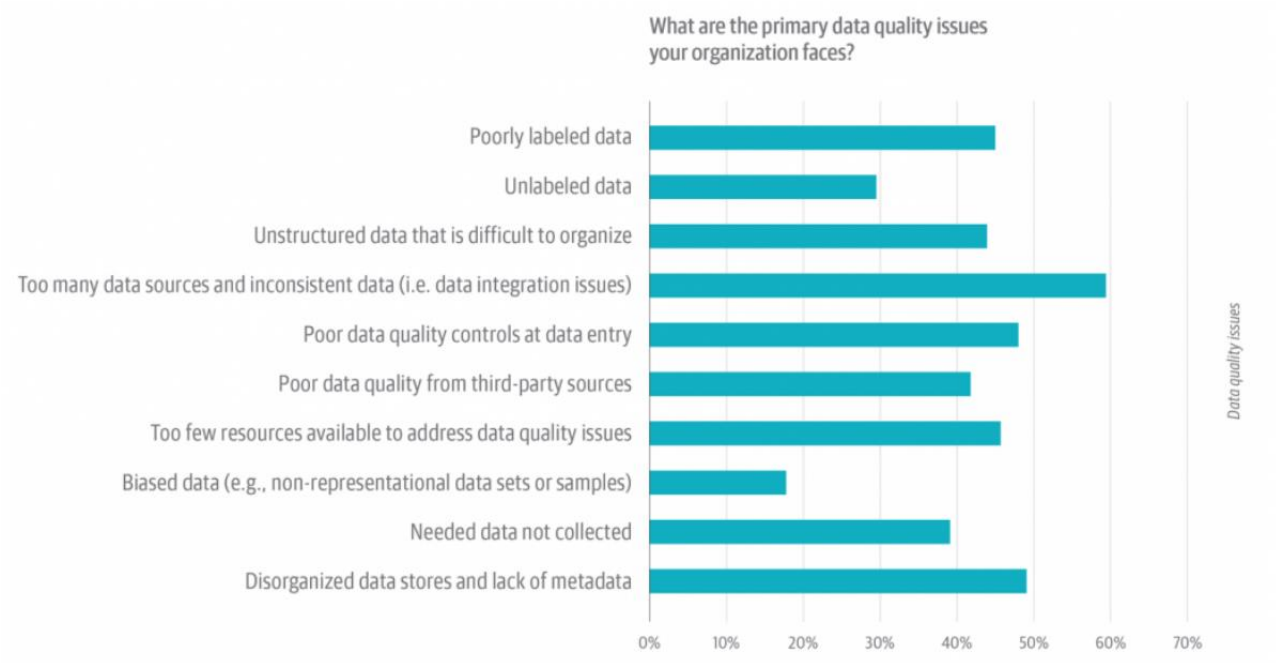

Figure 2.2: Primary data quality issues (in relation to AI) faced by respondents' organisations (Source: O'Reilly, 2020)

2.2 Data Standards

Virtually all studies on the barriers to digitalisation point out the significance of interoperability issues and the need for data standards to help overcome them. The International Energy Agency (2021) claim that "*a key issue for creating market value is interoperability, or the ability of devices to communicate with each other and work in an integrated system*". Valdez et al (2020) identify that "*The issue of interoperability is not only an issue of communication between different types of platforms, but also an issue of the interpretation of commands on a central platform which communicates between different types of devices*". The US Department of Energy (2021) identifies that "*the current lack of interoperability results in expensive integration efforts"* and that the desired *"seamless connectivity [*between devices*] is not yet widespread*". The European Commission's Group of Experts on Energy Efficiency (2021) calls for "*technically open and not proprietary interfaces to operational information systems*". Harbor Research (2020) state that "*the Intelligent Building Energy Management Systems (IBEMS) market landscape is fragmented, with many startups attempting to disrupt entrenched incumbents, whose systems are outdated, difficult to integrate, and do not incorporate emerging technologies*". They suggest that "*plug-and-play interoperability between devices and systems*" is required to enable software service providers to thrive.

Interoperability issues relate to both

1. *Device level (technical/syntactic) communications protocols*: Sensors, BMS controllers and other HVAC equipment communicate between themselves using communications protocols applied by the manufacturer of the respective hardware. Proprietary or closed protocols will generally make it difficult and expensive for third party service-providers to gain access to information from the devices and implement solutions. This can lead to vendor lock-in and high on-going service costs. BACnet was introduced as an open communications protocol to address this issue. BACnet is both an international (ISO) and ANSI standard. It is maintained by ASHRAE.

   BACnet divides the task of device interoperability into three distinct areas: Objects (information), Services (action requests), and Transport systems (internetworking, electronic messages). BACnet data is organized by "objects" and each object has "properties". This enables devices to be polled and respond, so that they can be discovered. BACnet "services" are formal requests that one BACnet device sends to another BACnet device to ask it to do something. This model of *objects* and

services is realized by encoding messages into a stream of numeric codes that represent the desired functions or services to be performed. The "language" of this encoding is common to all BACnet devices. BACnet devices exchange information and do things by sending and receiving electronic messages containing this coded language. BACnet provides flexibility by allowing multiple types of transport systems to be used to convey these coded messages between devices.

It should be noted that implementation of BACnet is not always uniform, and interoperability issues can still exist in some cases.

2. *Analytics level (informational/semantic) data modelling protocols*: While BACnet and other communications protocols have some ability to represent meta data from specific telemetry and automated devices, this will not typically be sufficient to describe all the relevant features of a building and how they interrelate. Analytics software (eg fault detection and diagnosis, model predictive control) will typically draw on additional data/metadata relating to building properties and installed equipment systems. Contextualising all the data and describing the structure of a building is the role of the metadata schema. As discussed in Section 2.1, metadata schemas also provide a key tool for supporting storage and exchange of the data required for data-driven analytics.

   Details of a number of metadata schemas, relevant to the operation of building mechanical services, are provided in complementary IEA Annex81 work (Fierro and Pauwels, 2022). This work provides guidance on the overall structure, features and trade-offs behind different metadata schemas, along with context on how metadata schemas are applied in practice. Similar to the discussion on communications protocols above, the focus is on "open" (rather than proprietary) metadata schemas - ie schemas that are permissively licensed, open-source or otherwise widely available.

Harbor Research (2020) point out that "*buildings have trouble adopting a standard data labelling or naming convention*" and "*while BACnet adoption is increasing, further protocol consolidation and data labelling standardisation needs to occur*". Hardin et al (2015) provides more extensive discussion of interoperability.

2.3 Data Consolidation and Data Sharing

Data can be exchanged locally, between devices on-premises, using various LAN technology options. Modern BMS products provide good interfaces for engineering staff to interact with the BMS for maintaining building HVAC systems. Details of this operations technology (OT) data sharing is outside the scope of this guide.

Importantly, data management functionality can also operate as an integrated IT solution, drawing in data feeds from a wide variety of sources including (i) metering infrastructure, (ii) building management system data, (iii) distributed sensor networks, (iv) cloud databases (eg weather data) and (v) from external markets (eg electricity price data).

Unfortunately, existing buildings with legacy controls hardware may not have the IT infrastructure and connectivity for the more sophisticated IT based approach. Otte et al (2021) citing Zimmerman (2021) notes that "*most commercial buildings were built before 2000 … so these buildings contain a mix of new and obsolete systems, while new systems are added in a piecemeal fashion*". Harbor Research (2020) also identify that "*most buildings are not currently equipped for grid interactivity*".

Legacy technology issues are compounded by industry structures which tend to result in separate vendors maintaining separate databases, leading to data siloes. Memoori (2021) identifies that "*building systems data is notoriously siloed*" and note that "*gaining access to [diverse data] … can prove a challenge*". Furthermore, building owners generally don't collect data unless there is an immediate and obvious use-case. Retrofitting new data collection (eg submetering) to service new analytics services may be more costly and difficult.

The challenge then is to bring diverse data sources together in a 'single pane of glass', so that trends can be compared, and analytics applied, to yield new insights. This is a challenge of data-consolidation or data-federation

The focus of this guide is on data consolidation and management in the cloud. While cloud computing brings with it potential for cyber security risks, it provides software developers with access to a greater diversity of data sources. It also provides non-specialist staff, in the building owner's organization, with streamlined access to data over the enterprise IT network, and data can be shared with external contractors and suppliers (including by communication to mobile devices such as smartphones and tablets). This capability enables innovative 21st century business models and logistics solutions that would not otherwise be possible.

2.3.1 Data Platforms

Data sharing between field devices and the cloud (end to end operation of cyber-physical systems) is achieved using an "IoT platform" or "data platform", a middleware platform hosted in the cloud. Typical objects and functionalities of an data platform architecture are illustrated in Figure 2.3. The data platform is (for digitalisation) analogous, at least partly, to what a computer operating-system is for a personal computer; guiding computational workflows and exchanging data to/from storage.

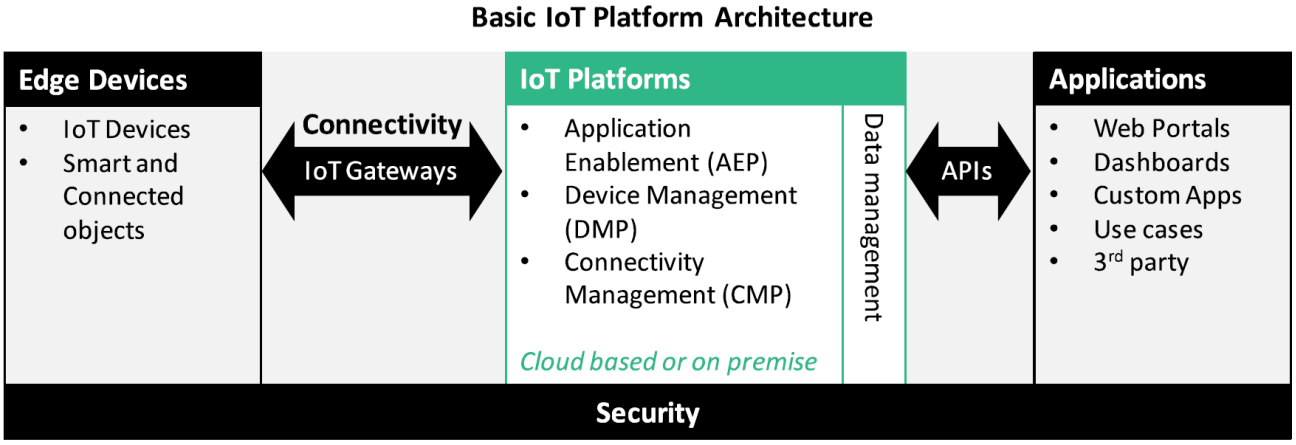

Figure 2.3: Basic IoT/Data Platform Architecture

Importantly, the data platform provides a means for collecting data once, only, and then reusing the data multiple times by sharing the data, via APIs, with an ecosystem of other applications. This is much lower cost than having every application needing to establish connectivity with various on-premises edge devices. Unfortunately, many platforms are restricted to a single vendor's services, limiting the number and diversity of applications to which the data can be shared.

Various alternative industry-specific names can be given to a data platform. For energy-productivity applications in the property industry, the data platform could be called an Energy Management Information System (EMIS). Alternatively, when an energy-utility attempts to manage loads from multiple buildings (and other grid resources), it does so through what is often called a Distributed Energy Resource Management System (DERMS).

There can be a symbiotic relationship between data standards and data platforms. Data platforms will generally use a specific data standard(s). Consequently, selection of the data platform may inadvertently lock in data standard(s) that subsequent suppliers must then comply with (particularly if there are limited interoperability drivers between standards). Ultimately an ecosystem of service providers will gravitate toward the data standards most widely adopted by the relevant data platform providers. This is a good outcome for industry if the selected data standards are non-proprietary.

2.3.2 Data Platform Functionality for Enabling Applications and Business Models

A diverse mix of smart building software applications (eg those suggested in Figure 1.2) could be hosted on a data platform. This highlights potential for the functionality of a data platform to be significantly greater than simply gathering and storing engineering measurements from field sensors and devices.

For example, Ofgem (2019) considered the software functionalities required to deliver '*a twenty first century electricity system that is more decentralised, more flexible, more responsive to changing demand and more accommodating of variable renewable generation*'. Relevant functionalities they identified for a data platform are illustrated in Figure 2.4.

Coordination	Coordinating platform tasks Facilitating data flows Harmonisation of standards and principles Alignment with external platforms and markets Conflict avoidance
Flexibility Procurement	Attracting flexibility providers and purchasers to the market Communicating requirements and availability Matching providers and purchasers
Dispatch and Control	Sending signals to dispatch assets Notification of asset dispatch Verification of asset dispatch
Platform Transaction Settlement	Verification of service against transaction Settlement of transactions
Platforms Market Services	Credit checking Asset pre-qualification
Analytics and Feedback	Network analytics, response times etc. Counterparty scoring and review Identification of market faults

Figure 2.4: Six tasks envisaged of Flexibility Platforms, for providing grid services (Source: Ofgem, 2019)

Not all of the functionalities in Figure 2.4 will be required from a typical smart buildings data platform (and associated application specific software). However, the Figure gives an useful flavour of what's potentially involved in integrated information management, hosted on a data platform, for the purpose of implementing a value-adding business service.

Importantly, many of the functionalities will exchange information, using relevant template data-fields, as a means of supporting standardised administrative processes and for enforcing governance policies. For example, input fields that require a practitioner's accreditation/identity number could be used to unlock access to certain restricted fields in a database, and so on.

Role-based authentication and the ability for a designated administrator to grant/withdraw permissions to access specific data fields is an important characteristic, required to satisfy commercial imperatives and to enforce privacy restrictions. This can also be used to enforce agreed dataset ownership rights.

Other generic functions could include data entry 'wizards' that help guide users toward correct use of software applications, including use of relevant anomaly detection algorithms and suggestions for default values in data fields. Search functionality (supported by relevant meta-data schemas) can support data discovery on the platform.

2.4 Data Sharing for Grid Integrated Control of Buildings: A Case Study

Control of building HVAC equipment can provide valuable flexible demand resources to the electricity grid. This is one of the energy productivity use-cases illustrated in Figure 1.3. Various data platform functionalities, that can help facilitate trading of flexible demand resources, are further illustrated in Figure 2.4. Flexibility will be a crucial resource for ensuring the reliability and affordability of future electricity grids that contain high levels of variable renewable energy generation.

Digitalisation and associated exchange of data is inherently required to enable this integration, through (amongst other things) (i) the communication of rolling forecasts of price signals, electrical infrastructure constraints, flexible demand asset availability (ii) real-time automated dispatch processes for activation of flexible demand capacity and (iii) measurement, verification and settlement of delivered flexible demand resources.

In this way, digitalisation plays a fundamental coordinating role for large numbers of distributed energy resources to participate in, and to support, the electricity grid. Importantly, the procedures, conventions and standards for data/information exchange (enshrined in platform software) become a key part of how the rules of market participation are implemented.

However, individual building owners (who could provide valuable flexible demand resources) are in no place to set these data exchange requirements or to build bespoke platforms for managing the necessary real-time data exchange. Indeed, it would be vastly inefficient for each flexible demand buyer and seller to set up their own digital platform and to insist on using their own proprietary standards and protocols.

Consequently, there is a critical need for common data platform software tooling, built using open interoperable standards. Such a common approach ensures

- Equitable access is provided to all building owners to enable them to provide their flexible demand resources into the market, and so that building owners are not locked out of the market by proprietary data standards and software
- There is a competitive market for energy services that values demand-side flexibility
- The market for demand side flexibility is coordinated rather than fragmented, and duplication is minimised.

This need was identified by the European Union leading to the 'EU Action Plan on the Digitalisation of the Energy Sector' (2022). The Action Plan supports the creation of a common European "Energy Data Space", that fosters the adoption of an interoperability framework, and addresses the governance requirements of the data space. The Action Plan is aligned with the planned Implementing Act for data interoperability requirements and procedures stated in article 23 and 24 of the Electricity Directive (EU) 2019/944.

In this plan, 'Data Spaces' is "*an umbrella term corresponding to any ecosystem of data models, datasets, ontologies, data sharing contracts, and specialized management services (i.e., as often provided by data centers, stores, repositories, individually or within "data lakes"), together with soft competencies around it (i.e., governance, social interactions, business processes)*' (Scerri et. al., 2022).

The Smart Energy Operating System (Smart Energy OS) is a Data Space designed around a hierarchy of data handling with associated technical and governance rules that aim to ensure coherence across all relevant aggregation levels (wholesale energy markets, transmission networks, distribution networks, and end-use consumers) (Figure 2.5), with a focus on providing solutions that include multi-objective criteria like energy efficiency and flexibility. The Smart Energy OS includes a hierarchy of controllers that are able to utilise the flexibility at the building, district and community levels to solve grid related issues for both the TSO and the DSO.

The Smart Energy OS is analogous to an operating-system for a personal computer; guiding fundamental workflows, as well as information, between various hierarchies of the energy system. It utilises 'Minimum Interoperability Mechanisms' (MIMs) as a common and interoperable set of information requirements that ensure that necessary data is accessible between stakeholders.

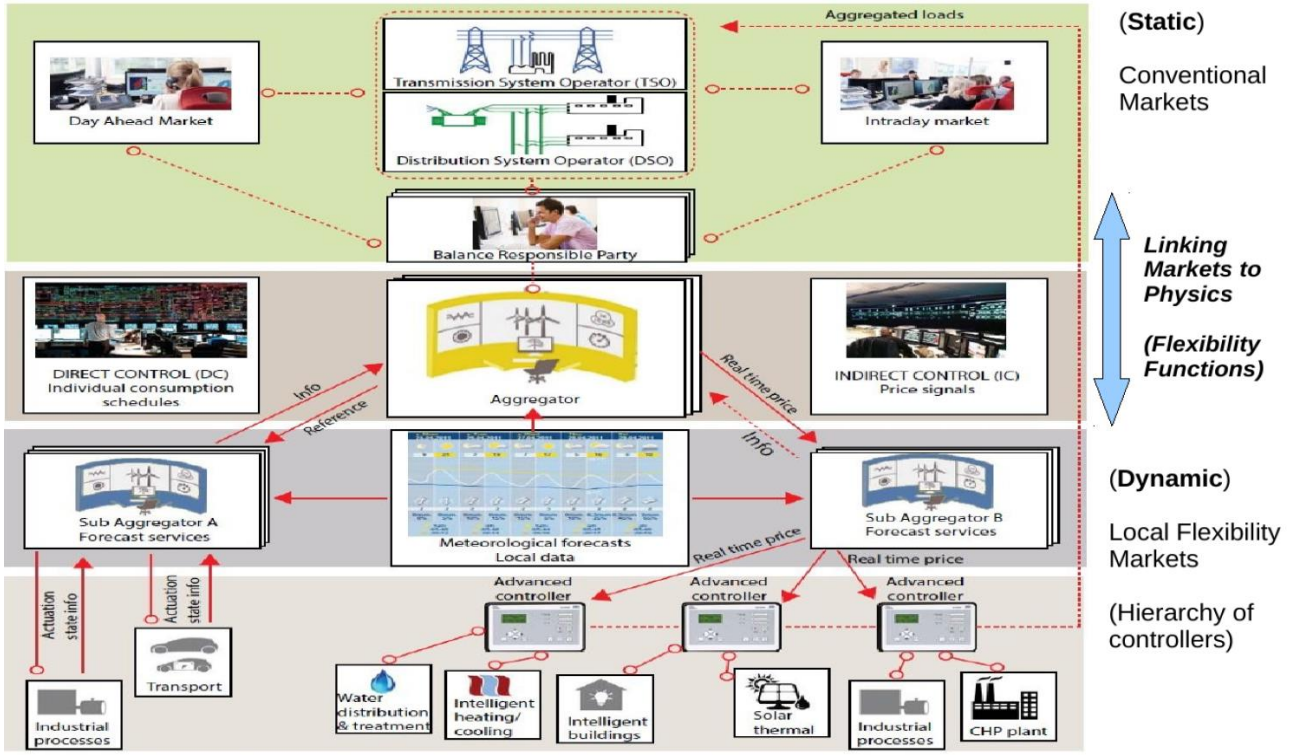

Figure 2.5: The Smart Energy OS for digitalization of energy systems using the Flexibility Functions as a fundamental MIMs for linking markets to the physics

For example, the Smart Energy OS contains a framework of spatial and temporal hierarchies for ensuring that forecasts of customer demand, and available flexibility, are coherent across relevant aggregation levels (from a home energy management system (HEMS), to the transformer/substation to the national grid etc). This is achieved using the so-called 'Flexibility Function' (Junker et al., 2020).

The Flexibility Function is one of the fundamental MIMs within the Smart Energy OS setup. It provides a condensed data exchange framework for linking market signals with the physical ability of buildings to provide flexibility. The signal is typically a dynamic price signal (Corradi et al., 2013), which is formed by solving market and grid related optimization and control problems (see Madsen et al. (2015) for details). The forecast demand in response to a step-change in price is provided by the flexibility function (Figure 2.6). The flexibility function can be used to calculate the so-called Flexibility Index for a building with given HVAC control strategies in a digital environment (see Junker et al., 2018).

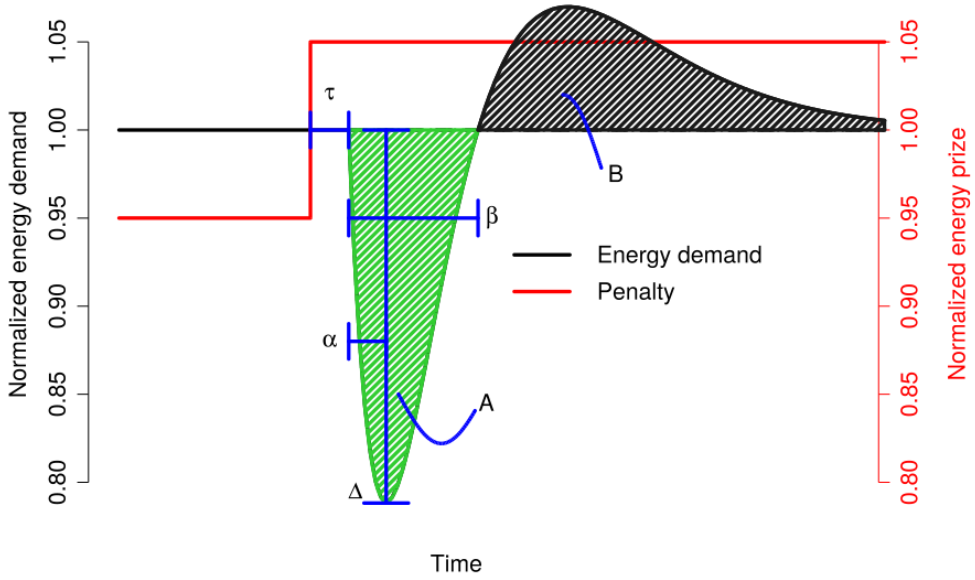

Figure 2.6: The Flexibility Function describes the relation between price (penalty) and demand. In this figure the Flexibility Function is depicted as a step-response function. The characteristics of the flexibility can be specified by using the values indicated indicated in the graph

Importantly, the Smart Energy OS can utilise data-driven approaches for near real-time forecasting and control, and real-time assimilation of sensor/metering data into digital-twin models (Thilker et al., 2021). This ensures that the models have a potential for self/auto-calibration as well as built-in real-time adaptation to changes of the systems.

Beyond the technical data exchange requirements for operating the electricity system, the Smart Energy OS also manages data processing requirements, including cyber security and privacy. In this way, it provides trusted data-sharing to empower end-users to focus on energy efficiency and flexibility - without being subject to disproportionate technical requirements (e.g. software), administrative requirements, procedures, data sharing and charges.

The Smart Energy OS is designed to keep privacy-related information at the edge, through the formulation of the Flexibility Function MIM which contains the needed information for the grid operators to be able to control the grids as well as for the aggregators and BRPs to provide a balance between demand and supply at a minimum cost.

For a more technical description of the Smart Energy OS setup for data-driven smart buildings we refer to the IEA Annex 81 state-of-the-art report on data-driven smart buildings (Candanedo et al., 2023).

The role of Center Denmark as a not-for-profit organisation and smart energy hub, providing institutional arrangements consistent with trusted data sharing, is further discussed in Section 4.3.

3. Constraints on Data Sharing

Many studies have investigated barriers that constrain the adoption of digital technologies in the non-residential buildings sector. The Australian 'RACE for 2030' Cooperative Research Centre scoping study on digitalisation (Trianni et al, 2022), conducted a thorough review of the literature to identify relevant barriers. They categorised them under the general headings of technological barriers, economic barriers, social barriers and regulatory barriers. The identified barriers are illustrated in Figure 3.1.

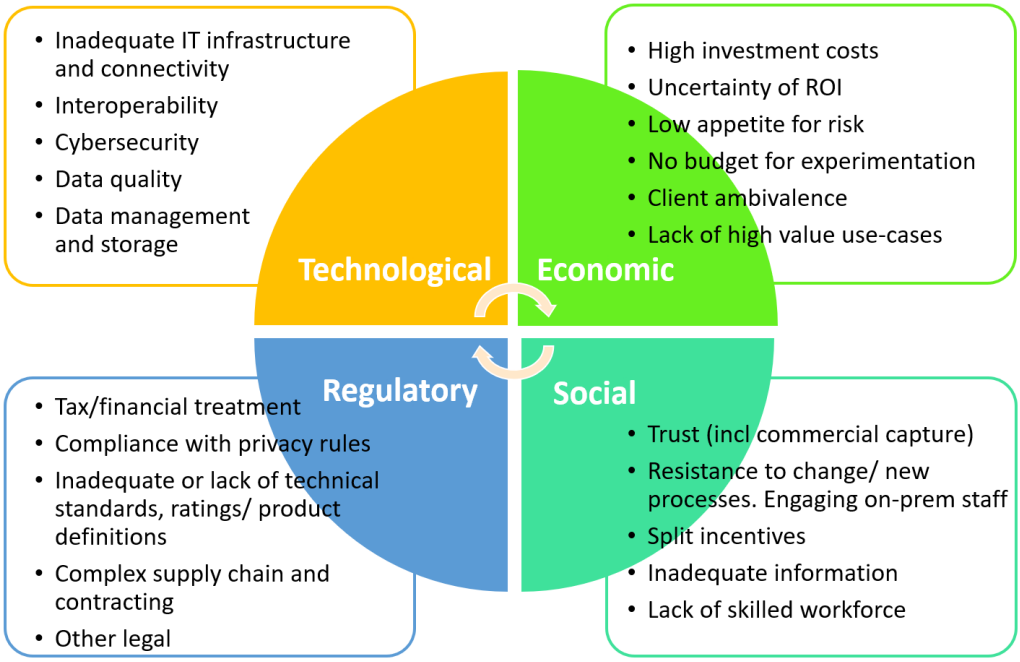

Figure 3.1: Identified barriers to the adoption of Digitalisation (Source: Trianni et al, 2022)

Interestingly, the identified technological barriers relate almost exclusively to data sharing. Some of these barriers have been discussed in previous Sections. Further discussion on data sharing barriers is provided in this Section, along with a framework for addressing these barriers.

3.1 Data Ownership

Care should be taken to ensure that permission has been granted prior to using data. Generally this is obtained through a license from the data owner. However the concept of the data owner is not always obvious. It is common for people and businesses to refer to data as if it is something that can be owned. For example, individuals and businesses very commonly refer to 'my data' or 'our data'. However, in general, there are no property rights in individual points of data. Therefore, data cannot be owned.

This is sensible considering the ubiquitous unintentional generation and transmission of data linked to, for example, the use of a smart phone. Extending this example, it's difficult to imagine allocating data rights between the smart phone user who initiated the creation of data, the smart phone manufacturer or App developer that provided the ability to create the data or the service provider that used the data to provide the desired service, and so on.

However, deliberate collection and curation of data as a dataset can be viewed as a creative work that can be subsequently analysed and exploited. Hence, the maker of such datasets can typically obtain protection under copyright law. Identifying the maker can be complex. The maker will probably be the entity who made the commercial decision to collect the data and made the commercial investment in carrying out the collection and curation of the datasets.

In the case of non-residential buildings, there are a number of actors that may have some claim to rights over datasets and/or need beneficial access to data at various points in the lifecycle of a building. These actors may include

- The Building Owner owns the assets that generate data (the Building Management System (BMS), Internet of Things (IoT) sensors and other enterprise sources), and they need access to these data sets for their business operations (for overseeing building operations and for completing various reporting requirements). However, the building owner may not directly own data sets created by third party service providers unless it is explicitly included in the service contract. When contracting, the building owner must also consider future scenarios where they will need to share data with other third party service providers.
- Software Providers use data to create meaningful insights. These insights may be in the form of a derived output data stream that gives instructions (manual or automated) on how to manage energy consuming assets. Given the software provider made the commercial investment to develop the algorithms, they may have rights in relation to the insights and data outputs produced by the software.
- BMS Contractors and Facilities Managers install requisite automation and communications hardware to establish the ability to collect data from buildings. They subsequently use that data to maintain relevant building assets. As part of installing the data collection capability, they may apply their own data management structures and host relevant databases.
- IoT Sensor Providers may directly ingest data from their sensor hardware to their own proprietary cloud platform, where the data is structured and stored as a dataset. This data will often need to be combined with other data sources to extract value for stakeholders.
- Building Occupants will generate data through their mobile devices and apps aimed at providing them with a superior user experience from the building. This data may include personal data, with accompanying ownership and utilization constraints.

Given that the building owner ultimately pays for all the services, and needs the ability to competitively source providers at regular intervals, it is generally assumed that data should 'belong' to the building owner. However, data has a practical tendency to find its way to the service provider and be inaccessible to the building owner. Indeed, the uplift of data to access-controlled external IT systems is often used as part of a service provider's business model for ensuring that the building owner retains their services.

There is a movement to try and unlock data through 'Consumer Data Right' (CDR) protections and initiatives. For example, the Green Button Alliance aims to provide electricity customers with easy access to their energy usage data in a consumer-friendly and computer-shareable format. CDR protections give consumers more control over their data, enabling them to access and share their own data with accredited third parties, and to thereby access better deals on products and services. It enables more choice of providers, simpler setup of transactions, more competition and diversity of product offerings

3.2 Privacy and Other Commercial Sensitivities

Most energy productivity applications will use base building data sources that relate to the operation of the whole building as it services the aggregate (rather than individual) needs of occupants. Hence, this data is typically not considered to be personal data.

However, some applications could potentially utilise occupant data which is personal. For example, occupant movement data can be used as an input to drive allocation of energy consuming HVAC services (eg switching off HVAC to unoccupied rooms, or allocating staff to office spaces in ways that minimize overall building energy consumption, etc). Occupants in the building may also wish to interact with building services using Apps on their mobile phone (eg to improve thermal comfort conditions, make meeting room bookings, or access other resources etc), which has the potential to lead to the collection of personal data.

While much can be done without using personal data, each data-driven application needs to be considered carefully for its potential to encroach on privacy issues.

The General Data Protection Regulation (GDPR) regulates data protection and privacy in Europe. It has also been adopted, in full or part, in various other jurisdictions.

The GDPR (Principles of the GDPR | European Commission) defines personal data as any information that relates to an identified or identifiable living individual. Different pieces of information, which collected together can lead to the identification of a particular person, also constitute personal data. Examples of personal data could include (i) a name and surname; (ii) a home address; (iii) an email address; (iv) movement records (for example the location data function on a mobile phone); (v) a cookie ID; (vi) access logs in a building and (vii) personal preference data (such as that collected via social media).

Sensitive data is a subcategory of personal data which is subject to specific, more stringent processing conditions.

The GDPR protects personal data regardless of the technology used for processing that data – it's technology neutral and applies to both automated and manual processing.

The GDPR protects personal data by placing requirements on companies and organisations which 'process' personal data as part of their activities. 'Processing' covers a wide range of operations that could be performed on personal data – including the collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction of personal data.

The type and amount of personal data a company/organisation may process depends on the reason for processing it and the intended use.

Data relating to companies or other legal entities is not personal data, and so is not covered by the GDPR. However, information in relation to one-person companies may constitute personal data where it allows the identification of a natural person. The rules also apply to all personal data relating to natural persons in the course of a professional activity, such as the employees of a company/organisation, business email addresses like 'forename.surname@company.eu' or employees business telephone numbers.

By definition, meta-data schemas seek to link data. For example, Bhattacharya et al. (2015) note that a meta-data schema could be used to combine relationships (location-person, e.g 'occupantOf(room)'), with an identifiable device (gadget-person, e.g. 'macAddrOfPhone(user)') and location ('gadget-location, e.g. 'computerIn(room)') to essentially target specific groups or individuals and track them. This could allow for a broad query such as 'show me all cctv footage of the room that contained the phone of person x yesterday'. Care should be taken to minimize the chance of mundane, and otherwise beneficial applications, yielding data that is unexpectedly identifiable.

While less subject to regulation, commercial sensitivities may also require restrictions on data sharing. For example, details about a building (that might be gleaned from semantic models), could plausibly expose sensitive business intelligence (such as equipment degradation, or building future expansion plans) to third parties, or pose a security risk.

3.3 Data Processing and Data Sharing Principles

While noting that much of the data used in building-services applications, in non-residential buildings, is unlikely to be highly sensitive, awareness and risk management strategies should still be put in place for the appropriate treatment of any personal data.

There are six legal bases for processing personal data that are recognised by the GDPR (Article 6(1))

1. The data subject has given consent
2. Performance of a contract, to which the data subject is party

3. Compliance with a legal obligation
4. Protecting the 'vital interests' of the data subject
5. Public interest or acting under official public authority
6. 'Legitimate interests'

While 'performance of a contract' would be the default basis for processing data, individual occupants of a building are unlikely to be a party to the various contracts between building owners and service providers. Consequently, consent will be required from relevant occupants in the building, to use their personal data. Consent should be a voluntary 'opt-in' choice, based on sufficient information and adequate understanding of the data processing activity.

Even with a legal basis for processing personal data, a company/organisation must respect key rules, including:

- personal data must be processed in a lawful and transparent manner, ensuring fairness towards the individuals whose personal data is being processed ('lawfulness, fairness and transparency');
- there must be specific purposes for processing the data and the company/organisation must indicate those purposes to individuals when collecting their personal data. A company/organisation can't simply collect personal data for undefined purposes ('purpose limitation');
- the company/organisation must collect and process only the personal data that is necessary to fulfil that purpose ('data minimisation');
- the company/organisation must ensure the personal data is accurate and up-to-date, having regard to the purposes for which it is processed, and correct it if not ('accuracy');
- the company /organisation can't further use the personal data for other purposes that aren't compatible with the original purpose;
- the company/organisation must ensure that personal data is stored for no longer than necessary for the purposes for which it was collected ('storage limitation');
- The company/organisation must install appropriate technical and organisational safeguards that ensure the security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technology ('integrity and confidentiality').

The GDPR takes a risk-based approach, where companies/organisations processing personal data are encouraged to implement protective measures corresponding to the level of risk of their data processing activities.

One form of processing, that a company/organization may wish to do, is to share data. The Australian Government's 'Best Practice Guide to Applying Data Sharing Principles' (Commonwealth of Australia, 2019a), is one source that discusses how personal data can be shared while respecting privacy rights. It has a particular focus on sharing potentially sensitive data collected by Government. The Guide notes "*the growing imperative for public sector data to be used more effectively to improve government service delivery*" in ways "*that can't be addressed when data remains in siloes*". The Guide aims to "*support agencies 'responsibility to share' by providing an approach for effectively managing the risks associated with data sharing*".

Importantly, it tries to avoid a binary approach of data being either 'closed' (unable to be shared) or 'open' (unrestricted public access). Instead, it aims for controlled access to the right people, for the right reasons, with safeguards in place. To do this, it applies a set of Data Sharing Principles based on the Five Safes Framework.

The Principles (Figure 3.2) enable a privacy-by-design approach to data sharing by balancing the benefits of using data with a range of risk-management controls and treatments. By focusing on controls and benefits, instead of merely reducing the level of detail in the data to be shared, the Principles can assist with maximising the usefulness of the data.

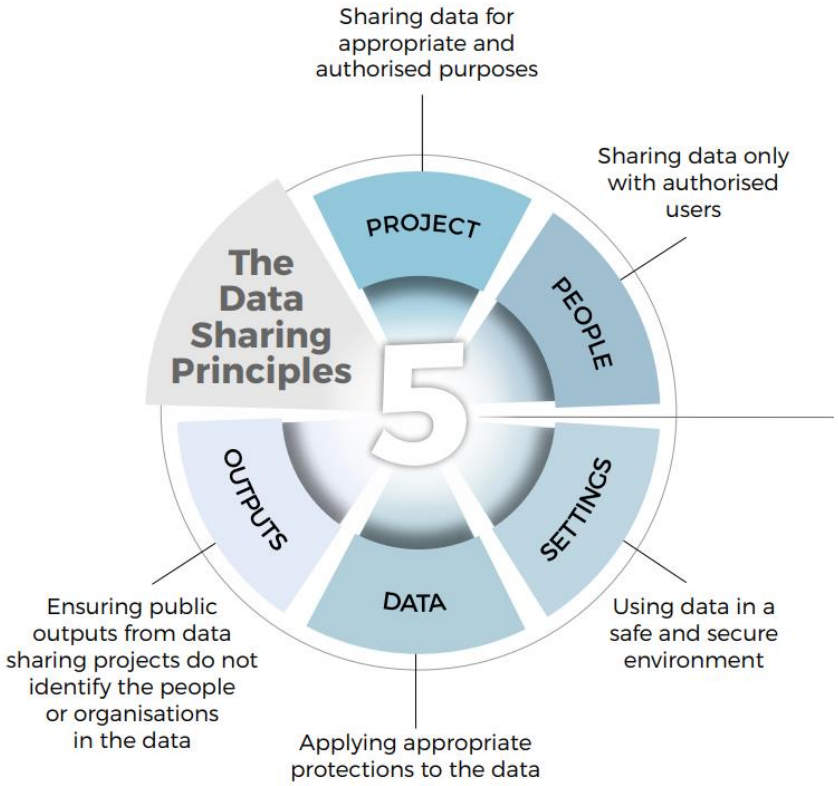

Figure 3.2: Principles for Sharing Data Safely (Source: Commonwealth of Australia, 2019b)

Each of the Principles can be considered as a focus area, where the stringency of the required control mechanisms (risk management choices) can be adjusted to achieve an appropriate balance between openness and the level of sensitivity of the data being shared.

While each Principle can be considered independently, all five Principles should be considered jointly to evaluate whether a particular instance of personal data sharing is a safe arrangement. Controls should be based on a realistic assessment of the likelihood and consequence of a risk occurring and be made in the context of organisational risk tolerance, rather than based on hypothetical worst-case scenarios.

A range of possible control options, under each Principle, are identified in the Australian Government's 'Best Practice Guide to Applying Data Sharing Principles' (Commonwealth of Australia, 2019a).

For example, under the 'Data Principle', only a limited set of certain non-sensitive personal data might be selected for release to certain types of organisations. And under the 'People Principle' a more fulsome dataset might be released only to a research group that has gone through an accreditation process to prove that they have the capability and IT infrastructure necessary to manage the data appropriately. In another example, under the Outputs Principle, publication of aggregated data or deidentified data may be sufficient to manage risks (noting that for data to be truly deidentified, the deidentification step must be irreversible).

Deidentification is a common control. Privacy algorithms, such as k-anonymity (Sweeney, 2002), l-diversity (Machanavajjhala et al, 2006), or differential privacy (Dwork, 2006) can be used to protect data agents record-linkage, attribute linkage and probabilistic attacks, respectively. Typically, the more manipulation done to protect the data (eg deidentification, aggregation), the less fidelity and useful the data becomes. Kjærgaard et al. (2020), identify that there are several frameworks for protecting time-series data, and suggest that the 'PAD framework' can protect building-related time series data with the privacy model of k-anonymity. The unique property of the framework is that the data publisher can specify how the data is to be used. These specifications are then considered as part of the anonymization phase of the data

3.4 Risk Mitigation Examples for Some Relevant Data Streams

Based on the discussion above, this Section provides some examples where data-risk has been assessed and some level of controls have been put in place. These examples start with consideration of a particular data stream, that is typical of those encountered in relevant building services applications. These examples are intended as a means for illustrating practical examples of data risk management in building management applications. However, they may not be appropriate in individual circumstances and should not be relied on as a deemed to satisfy solution for addressing privacy requirements.

It is also noted that treating individual data sources in isolation may not provide a complete picture for managing privacy, as privacy protection controls can potentially be compromised in a range of ways. For example, consideration should be given to the potential for multiple data sources to be combined in ways that create personal data and/or increase the sensitivity of the data. Physically entering and observing building features, or activity in a building, could create sufficient knowledge to reidentify data. Similarly, when dealing with time-series data relating to occupants in a building, there can often be repeating patterns (e.g. an occupant will likely get into the office at a fairly similar time each day) that could be used to infer meaning.

Example 3.1: Residential Heat Meters	
Description	Remotely readable heat meters are used extensively by district heating networks for measuring energy consumption for billing purposes. Measurements are typically taken hourly or even sub-hourly. Heat meter data is generally held by the district heating utility company. Records include residential address. This can be correlated with publicly available data on building type, construction year and, where available, the energy label. Important information can be derived from heat meter data-sets including (i) correlation of consumption patterns with occupant densities and building characteristics (eg U-value of the building envelope), and (ii) the internal temperature in winter or the building's dependency on solar gains. Analysis of this data can be used to help manage supply of energy to the district heating system, particularly to balance demand with supply from intermittent renewable energy sources. Heat meter data was provided by a district heating utility to a university for analysis. Processed data was anonymised and published.
Data Sensitivities (Personal/ Sensitive/ Commercial)	The data is considered personal because it relates to the energy consumption behaviour of identifiable living individuals. However, the data is not considered sensitive because the temporal resolution is low (hourly) and heat demand does not correlate with sensitive lifestyle activities that may occur in a home.
Basis for Processing (where relevant)[+]	The basis for processing is acting in the **Public Interest** and **Legitimate Interest** (Article 6(1)(e) and (f)) provided that the basic processing principles in Article 5 of the General Data Protection Regulation are observed (Danish Energy Agency, 2018)
Data Sharing Controls[++]	Address linked data was provided to the university as an authorised user, trained in managing data privacy. Address data was used solely for obtaining relevant meta-data on the building physical properties, before being discarded. The unique meter and customer identifier were anonymised by replacing them with random integers to prevent particular customers from being identified. Furthermore, the files were renamed with random numbers to remove postal code information.

[+] "Basis for Processing" under the GDPR
1. The data subject has given consent
2. Performance of a contract, to which the data subject is party
3. Compliance with a legal obligation
4. Protecting the 'vital interests' of the data subject
5. Public interest or acting under official public authority
6. 'Legitimate interests'

[++] Data Sharing Controls
1. Project: Sharing data for appropriate and authorised purposes
2. People: Sharing data only with authorised users
3. Settings: Using data in a safe and secure environment
4. Data: Applying appropriate protections to the data
5. Ensure public outputs from data sharing projects do not identify the people or organisations in the data

Example 3.2: Data from User-Generated Requests for Building Services	
Description	Occupants in a non-residential building will have need, from time to time, to log a call with facilities management. This could be to improve comfort conditions (eg change the thermostat set point), book a meeting room in the building or ask for help to fix a maintenance issue. These requests are progressively being moved to job-logging software and to mobile device 'Apps'. The software collects data relating to the location of the issue, date and time of request and will typically collect the name and contact details of the requestor. The data is used by facilities management to schedule staff to attend to a matter, or possibly to provide an automated solution (eg socially driven HVAC Control (General Services Administration, 2015)) A range of new building occupant tools are providing innovative new ways of customising user experience in buildings.
Data Sensitivities (Personal/ Sensitive/ Commercial)	The collected data is personal data as it contains information that relates to an identified or identifiable living individual. Even if the name of the individual is not collected, the location and time of request could potentially be used to identify a particular person. The data is not deemed sensitive as it relates to operation of the building.
Basis for Processing (where relevant)[+]	The primary legal grounds for processing of personal data, is for the purposes of providing services under a contract. The job logging software platform, or mobile App service is provided to the building owner, under contract, and then to the tenant. The tenant may then offer the service to their staff, the privacy of which would be dealt with under staff employment contracts (similar to other internal software-enabled business services). Depending on the HVAC related service provided, the service may be optional opt-in only, achieved by downloading the App. In this case a direct contractual relationship with the occupant could be established with the software service provider (as part of the App download process). In some circumstances, the grounds for processing of personal data may be legitimate interest, as a provider of the Services (for example, to protect the security and integrity of systems).
Data Sharing Controls[++]	A publicly available privacy policy gives details of what data is collected and how the data will be used in order to provide the desired services. In this way the Data Project Control is to only share data for agreed/authorised purposes, appropriate to the service. Data is retained for 5 years Disclaimer is made that the service is not for children under the age of 16. As a 'People Control' the privacy policy requests children under 16 not to use the service.

Example 3.3: Wifi Data for Building Occupancy	
Description	The TU Delft gathers WiFi data, which are available for research purposes. The data is used, for example, to detect building occupancy and relate these with data from the building management systems and energy meter data to reduce energy wastage in utility buildings. The following topics are available within the WiFi dataset: • WiFiclientCounts • WiFiclientsessions_Month • WiFiclientsessions_Vendor • WiFiclientstats_Day, WiFiclientstats_Month and WiFiclientstats_Vendor are created.
Data Sensitivities (Personal/Sensitive/ Commercial)	The collected WiFI data (Mac address of mobile devices) are personal data as it contains information that relates to an identified or identifiable living individual. But it is not considered privacy sensitive.

Basis for Processing (where relevant)[+]	Storing the data in the building management system is on a 'legitimate interest' basis because (i) it is being used as part of the normal expected operation of the building, (ii) there is minimal privacy impact, and (iii) extra steps have been taken to protect people's rights and interests through pseudonymisation of the data, prior to use in research.
Data Sharing Controls[++]	Within the TU Delft a Data Protection Impact Assessment (DPIA) was executed to ensure that all potential impacts on the use of WiFi data are properly assessed. This plan needs to be officially approved by the management. As a 'Data' control, only specific WiFi data is extracted from the Cisco management system and anonymised/pseudonymised before storing them at the TU Delft data platform. After that, they become available for research purposes. There are three API calls whose data is pseudonymised in different ways. • For the topics *_Day, the MAC addresses and IP addresses are written away with a hash function that changes per day. • For *_Month topics, the MAC addresses and IP addresses are written using a hash function that changes monthly. • For the *_Vendor topics, the last half of the MAC addresses are overwritten with zeros, and the IP addresses are written away with a daily changing hash function. As a 'Project' control, data is only available for appropriate and authorised purposes. This is achieved by requiring that all researchers at TU Delft create a Data Management Plan (DMP) before being given access to the data. In the DMP researcher describe how data will be collected, managed, stored, and made available during the study, and how it will be shared after the study is completed. In this plan researchers indicate which WiFi items they want to use.

Example 3.4: Residential Smart Thermostats Datasets	
Description	This dataset comprises historic data from over 100,000 homes in North America. The dataset includes high-resolution setpoint and indoor air temperatures, outdoor temperature, HVAC equipment daily usage profile, and relative humidity. The metadata includes approximate location (city, state/province, country), floor area, age of the house, number of floors, number of occupants, type of heating system and the style of dwelling.
Data Sensitivities (Personal/Sensitive/ Commercial)	The data is personal since it is correlated with the occupants' airconditioner and heating usage behaviour.
Basis for Processing (where relevant)[+]	The basis for processing is that consent has been obtained. The program is volunteer based (i.e., the homeowners sign a voluntary agreement to make their datasets available to researchers/control engineers). The owners agree that this information could be used for research purposes, without enabling their identification.
Data Sharing Controls[++]	Applying Project Controls**,** the nature of the shared data is restricted to that which is more useful for extracting general/aggregate trends and conclusions, rather than for gleaning information about individual users. Applying Data Controls, each house is anonymized, and the exact location of the dwelling is not available. Applying People Controls, users of the datasets (universities and research centres) sign a non-disclosure agreement, whose provisions included using only the data for scientific research purposes, not sharing the data with unauthorized users, nor providing details about the structure of the dataset storage and making sure no results would contain any reference that enables the identification of the occupants.

4.Institutional Arrangements for Data Sharing

The GDPR identifies a 'Data Controller' as an entity that determines the purposes for which, and the means by which, personal data is processed. Hence, the Data Controller is the company/organisation that decides 'why' and 'how' personal data should be processed.

It also identifies 'Data Processors' who process personal data on behalf of the Data Controller. The Data Processor is often a third party external to the Data Controller. A typical activity of a Data Processor is to offer IT solutions, including cloud storage. In the context of cloud-based applications, a Data Processor can also be known as a 'data custodian'.

Data Controllers must actively demonstrate full compliance with all data protection principles and are responsible for the GDPR compliance of any Data Processors that they might use to process the data. Data Processors process personal data according to the data controller's instructions, but become responsible when they operate outside the Data Controller's instructions.

4.1 Platform Intermediaries

Data platforms are discussed in Section 2.3. Significant investment and expertise is required to develop and implement data platform software. Consequently, specialist companies have developed and commercialised data platform services to avoid the need for building owners to develop their own bespoke solutions.

These data platforms provide a common software infrastructure for hosting data and enabling the processing of data by value adding analytics services, as illustrated in Figure 4.1.

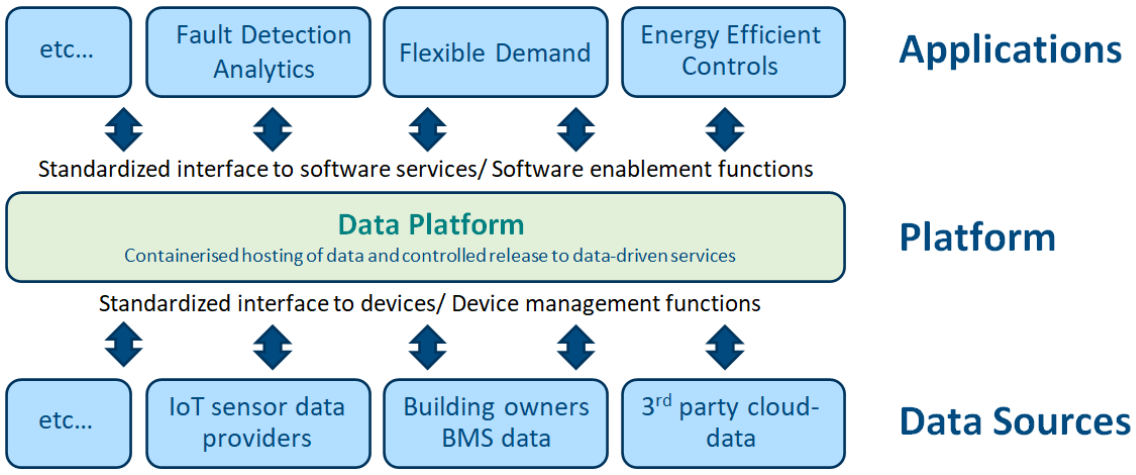

Figure 4.1: The role of a data platform to connect and enable software analytics services

The assembly of consistent yet individually curated datasets, on the platform, enables third-party applications software to be cost effectively developed and deployed. Without this enabling aggregation, individual software application developers would need to engage with each individual data owner, potentially having to develop tailored applications based on the structure and semantics of each source of data. In this way, the platform can play a key role in brokering the relationship between data providers and software application developers who are able to use the data to deliver value back to the providers. This brokerage entails:

- Setting consistent standards for data uploaded or deposited to the platform
- Qualifying/validating/authorising the applications that interact with the data on the platform

Where personal data is involved, it is clear that a data platform is inherently tasked with processing data. In this way the owner/operator of the data platform will be a Data Processor, if not a Data Controller (depending on circumstances). Therefore, in addition to enabling technical aspects of data-driven smart building services,

a fit-for-purpose data platform would be able to help implement controls relating to the five data sharing principles. This may include

- Cyber security to protect stored data from malicious attack
- Authentication of users/role-based access to data
- Tracking consent, contractual and/or other conditions that may be attached to data
- Tracking data usage and provenance (who, where, how often)
- Aggregation and/or deidentification of data, and/or other forms of metadata suppression
- Encrypted data matching processes and secure environments/location for data processing

Noting the potential complexity of these issues and the risks involved, the emergence of specialist data platform providers is helpful - as a means of consolidating industry practice around common standards and professional expertise.

However, unfortunately, the distinct value of the data platform role has not yet been widely recognized in the industry. As a result, the data platform role is typically bundled with specific application services as a single market offering. This business model continues the industry's tendency to form siloes and to enshrine interoperability (data access) barriers.

These factors indicate that data platforms can (at best) facilitate desirable access to data but (at worst) can frustrate access to data and create a position of market power for platform owners. Separating out the data platform role from individual application services would be a useful step toward enabling FAIR data management principles, supporting innovation, and driving value in the market.

4.2 Governance Structures

Governance of data resources relates to the framework of rules, relationships, systems and processes by which data is managed and decisions are made in relation to data access.

Some of the relevant industry stakeholders are detailed in Section 3.1, where they are discussed in relation to data ownership (and confusion around data ownership). Some of the rules and approaches for managing privacy are also discussed in Section 3. These are all important considerations in the management of data governance.

However, the focus of this Section is on the commercial and institutional roles of different stakeholders in the ecosystem, and how that might be governed to maximise benefit. The aim is to find a balance between solutions that (i) best support FAIR data principles while managing privacy and (ii) maximize data value while minimizing risks.

Some approaches and motivations, for accessing and managing data from non-residential buildings, are described in Table 4.1. Table 4.1 identifies some considerations relating to the relationships and resulting power dynamics that can potentially arise.

There is no unique best approach amongst those listed in Table 4.1. Traditionally, the relevant services are provided as commercial services, and there are various highly capable companies offering these services. Companies typically offer data platforms using a Platform-as-a-Service (PaaS) business model.

Table 4.1: Approaches for sharing data for enabling data-driven energy services

Initiating Stakeholder /Purpose	Technology Provider /Action	Outcome
Building owner seeks energy savings from their BMS contractor.	BMS contractor implements advanced control functions on-premises through the BMS.	Low capex solution but with minimal improvement to broader building data collection and management functions. Gives limited visibility of impact to the building owner (except through monthly energy bills). BMS contractor is in a position of power in relation to data access/visibility.
Building owner brings in a specialist energy analytics company to identify savings opportunities.	Energy analytics company installs their hardware gateways to access data from the BMS and EMS, and applies software analytics tools to identify faults and energy saving opportunities	The energy analytics company is potentially perceived as having a policing function over the incumbent BMS contractor. This may create an unhelpful relationship between the two. Data on the operation of the building ends up sitting on the servers of the energy analytics company, with potentially limited access for the building owner and other third-party providers. The energy analytics company may be in a position of power relating to data access/visibility. Data collection and use is dealt with separately, for each use-case in the building, resulting in siloes of data.
Building owner establishes their own 'data-lake' for capturing, storing and managing data on the operation of their buildings. Data is stored on the building owners servers and separately distributed to service providers by the building owner.	Building owner pays software developers to build their bespoke data-lake and APIs to distribute data to where it needs to go (presumably as a white labeled version of an existing data-platform).	The building owner maintains sovereignty over their data and full access to the data. However, the bespoke solution may be expensive and inhouse skills may be required to maintain the system and enable data sharing with third parties. The marketplace for services (that could use the data) may be opaque and any non-standard features would increase costs of interacting with third party service providers.
Building owner utilizes a 3rd party data platform service to collect and manage their data, as a separate service to any analytics providers that the building owner may subsequently use.	The platform provides the technology for streamlining data exchange with service providers nominated by the building owner. A marketplace/ ecosystem of service providers has standard products that have been designed to access the data platform	Data platform costs can be shared across many users of the platform, resulting in lower costs compared with a bespoke data-lake. To the extent that the data platform is independent from analytics services, and provides the building owner with tools for accessing and self-managing their data, the data platform could be a trusted marketplace for transacting services. Such platform-based business models are common in the IT industry. While creating great potential for innovation and competition, data inherently resides on the data platform owner's infrastructure with potential for data harvesting.

However, there is growing realization that data sharing services have unique challenges that may require consideration of alternative business models. Specifically in relation to data platforms (the fourth option in Table 4.1), these challenges include

- Data Sovereignty: By collecting data and storing data on behalf of building owners, a data platform could potentially gain significant visibility of confidential or sensitive data, and/or have capability to use the data to provide services to others. Indeed, the business model of some IT platforms, in other domains, is to provide services for free on the condition that the collected data can be used to generate separate revenue streams for the platform owner.
- Vendor lock in: Once data is stored on an external provider's data platform, it may be difficult to recover that data if there is reason to change provider. This potential to lose historical data creates a disincentive to change provider, stifling competition and innovation. Similarly so-called 'network effects' have the potential to create natural monopolies for platform providers. The more a platform is adopted by users the more useful it becomes and the harder it is for others to compete.

Addressing these challenges, the Open Data Institute has proposed the need for Data Institutions. They define Data Institutions (https://theodi.org/article/what-are-data-institutions-and-why-are-they-important/) as organisations that steward data on behalf of others, often towards public, educational or charitable aims. They identify two opposing forces that tend to restrict data sharing and decrease value from data, being (i) the tendency of data owners to hoard data for their private benefit and (ii) the tendency of data owners to fear data sharing in case of unintended negative consequence.

The role, therefore, of a data institution is to steward data in a way that minimises these opposing forces and unlocks value from data. The Open Data Institute identifies six roles involved in stewarding data (Figure 4.2).

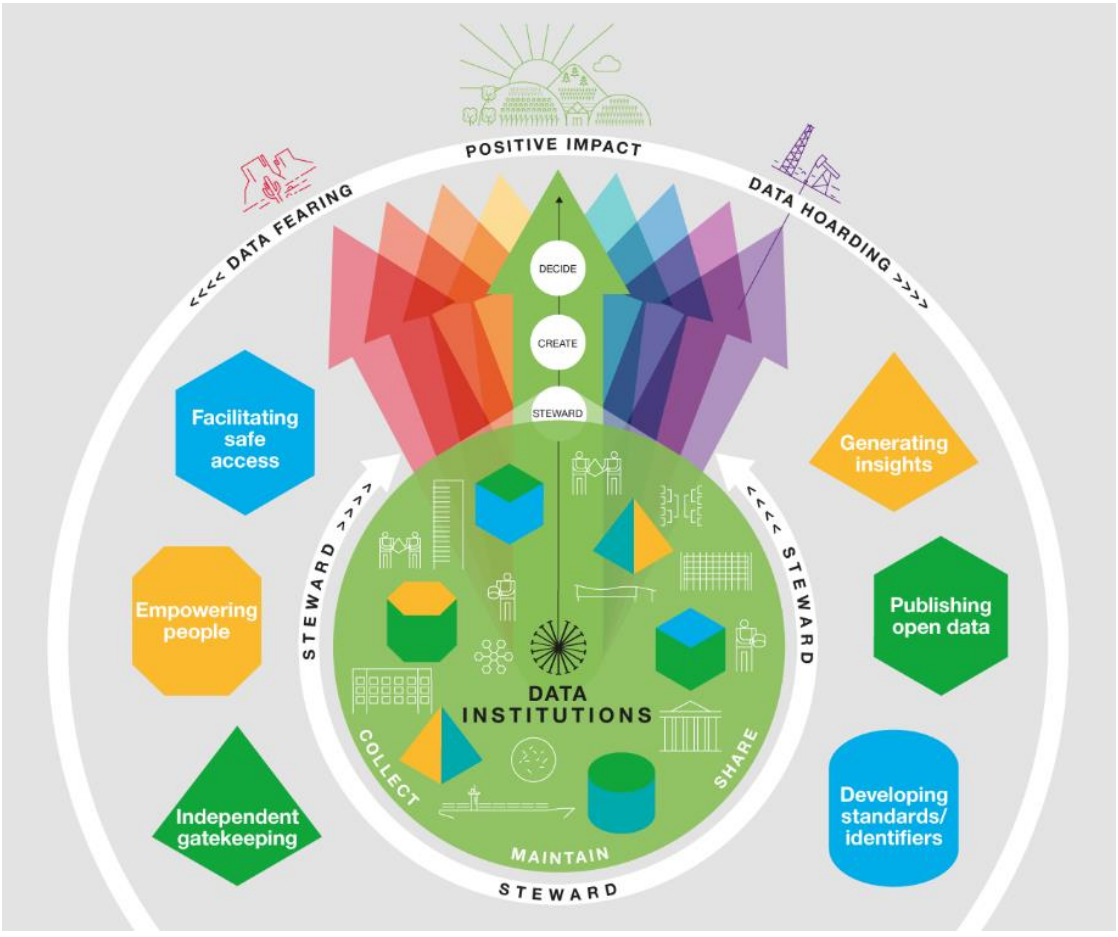

Figure 4.2: How data institutions unlock value from data (Source: The Open Data Institute)

Stewardship addresses data hoarding by adding new value through

1. *Generating insights*: Combining or linking data from multiple sources, and providing insights and other services back to those that have contributed data.
2. *Publishing open data*: This could include creating 'sandpits' of deidentified data to enable development of new algorithms and value adding services, that can further a particular mission or cause.
3. *Developing and maintaining data standards and identifiers*, and other infrastructure for a sector or field, to support discoverability and reusability of data.

Stewardship addresses data fearing by implementing controls through

4. *Facilitating safe access*: Protecting sensitive data and granting access under restricted conditions.
5. *Empowering people*: Enabling people to take a more active role in stewarding data about themselves and their communities.
6. *Independent gatekeeping*: Providing a gatekeeper service for data held by other organisations.

The Open Data Institute notes that various national infrastructure agencies (eg statistics agencies) have played the role of a data institution, on behalf of the public, for hundreds of years. However, the digital age requires discussion on how far these concepts can or should be extended.

Whether a data platform is provided by a commercial entity, a not-for-profit entity or a government agency, the simple act of providing the data platform technology requires the platform provider to exercise many of these data stewardship functions. There is clearly potential synergy between the functionality provided by the data platform technology provider and the roles required of a data institution.

Lawrence and Oh (2021) further identify 'Data Trusts' as a type of data institution that allows individuals or groups to pool resources, tasking an independent 'trustee' to manage those resources for the benefit of the trust's members, where data trusts are characterised by their focus on:

- enabling data-driven innovation for social and economic benefit, by creating a trustworthy environment for data sharing.
- re-balancing power asymmetries in data exchanges, by encouraging and empowering the originators of the data to play an active role in setting the terms of data use – and the distribution of the value that creates – and providing a platform for collective negotiation; and
- anticipating, preventing, and managing the vulnerabilities associated with data use, through professional data stewardship.

Data trusts offer a vehicle for individuals or groups to choose how they want data about them to be used. The trustee acts for their benefit, and provides a mechanism for bottom-up engagement that empowers individuals and communities in decisions about data use.

Lawrence and Oh (2021) provide various case studies of data trusts, including data trusts operating on a commercial basis on behalf of its members.

In the context of non-residential buildings, building owners could, for example, establish a cooperative to share the cost of developing data platform services in a way that ensures best practice stewardship while giving building owners control over their data.

Similarly, many government policy initiatives for the non-residential buildings sector (eg ratings and certification agencies, subsidy programs) require common data sharing infrastructure that could plausibly be used as a public service for the sector.

4.3 Case Studies

A range of data sharing services exist in the sector. Some examples of these are described below (with a focus particulary on entities that support not for profit and/or pubic-good outcomes). The authors of this report are not endorsing any of these services, but provide them as a means of illustrating possible solutions to some of the challenges of data sharing. The not-for-profit case studies and their respective features of interest are

1. Green Button Alliance: is an example of a not for profit organisation providing a set of open data sharing standards, and certification of commercial data platforms, with the aim of putting control in the hands of consumers (enabling them to authorise 3rd party access to their data).
2. Centre Denmark: is an example of a not-for-profit collaboration between universities, TSOs, DSOs, public organizations and private companies providing trusted data sharing and streaming services.
3. Super Low Energy Building Smart Hub: is an example of an online government data sharing resource, hosting national building energy data sets and AI-assisted energy analysis tools. It offers a variety of use cases for industry including green building certification and green financing.
4. Data Clearing House: is an example of an independent, government-developed IoT platform providing a hosting service for building owners to control 3rd party access to their data and for analytics companies to deploy software applications. It has been developed, at least partly, to support independent measurement and verification of demand flexibility, and it utilises the open standard Brick ontology.

Green Button Alliance, USA	
Context	Access to energy use data is required to enable consumers to take action to reduce their overall energy use and save money. For example, access to energy use data enables consumers to identify energy-saving home improvements (e.g., insulation, high performance glazing, new appliances, solar PV etc). It could also help evaluate opportunities to save money where time-of-use (TOU) pricing for electricity creates incentives for off-peak shifting of energy consumption. Often consumers will need help from third party companies and Apps, to interpret their data into actionable advice. Unfortunately, energy usage data has typically been inaccessible to consumers, and difficult to share with third parties - despite the data arguably being the consumers private information.
Vision of the Entity	The Green Button initiative is an industry-led effort to respond to a White House call-to-action to provide electricity customers with easy access to their energy usage data in a consumer-friendly and computer-friendly format. The Green Button Alliance is the single, definitive go-to-place for all things related to the Green Button initiative — from certification of implementations to marketing and education.
Stakeholders: their problems, their role and expected interaction with the Entity (including who are the data generators and data consumers?)	Green Button Connect My Data (CMD) is an open-data standard designed to unlock access to utility interval usage and billing data—providing easy, seamless access for software applications. Green Button CMD enables consumers to authorize third-party solutions to obtain interval meter data quickly and securely. This enables accurate and detailed analysis of usage data, while ensuring customer data is protected and privacy is maintained. The Green Button standard is implemented by utilities, whose platform/capability is then certified by the Green Button Alliance **Energy consumers** can use a certified "Green Button" platform to download or connect to their utility-usage data in order to gain better insight of waste and inefficiencies; adjust consumption patterns, streamline ESG reporting, support competitive negotiation of tariffs and save money. **Electricity Utilities** can use the Green Button standard to streamline processing of requests for data, support administration of energy efficiency programs, and to provide avenues for greater customer engagement. **Control vendors, technology developers and system integrators** can access data to develop and install products and systems that provide value to customers.

	They can also participate in a growing ecosystem of users, thus expanding the market for their products and applications.
Legal form of the Entity	The Green Button Alliance ("The GBA") was established in February 2015 as a 501(c)(3) non-profit organization, to foster the development, compliance, and adoption of the Green Button standard. Previously Green Button had been run by government.
Decision making positions and processes	Membership is composed of leading utilities, governmental departments and agencies, solution providers, and affiliate organizations that collaborate to advance the Green Button initiative. The GBA is governed by a board consisting of a representative from each 'sponsor member' and some additional elected members from 'participating members' (participating members include core collectors/distributors of information using the standard)
Technology used to underpin data exchange	Green Button is based on the Energy Services Provider Interface (ESPI) data standard released by the North American Energy Standards Board (NAESB) in 2011. The data standards development process was facilitated by the Smart Grid Interoperability Panel, a public private partnership that is facilitated by the National Institute of Standards and Technology (NIST). Green Button 'Download My Data' enables utility customers to download their own energy consumption data directly to their own computer, and if they so choose, upload their own data to a third-party application. Green Button 'Connect My Data' allows utility customers to automate the secure transfer their own energy usage data to authorized third parties, based on affirmative (opt-in) customer consent and control. The Green Button standard utilizes an Extensible Markup Language (XML) including Atom Syndication Format, which enables it to support complex data structures that allow highly detailed data
Management of data sensitivities	Green Button is consistent with current privacy and security practices. Customers have to first authenticate themselves on a utility portal with a login and password before they see and download their own information. If they want, customers can share their own data that they have downloaded, by independent choice and action, with those they trust.
Commentary (eg (i) why the structure was chosen (ii) how the structure unlocks value, and (iii) financial self-sustainability)	The Green Button initiative was officially launched in January 2012. Over 50 utilities and electricity suppliers have signed on to the initiative. In total, these commitments ensure that over 60 million homes and businesses will be able to securely access their own energy information in a standard format. This number will continue to grow as more utilities voluntarily make energy data more available to their customers in this common, machine-readable format. Adoption was boosted when the US Federal government adopted the standard for its own portfolio of buildings in 2013. Green button has been mandated in Ontario, Nova Scotia and New Hampshire.

Center Denmark (Hub for Smart Energy Systems), Denmark	
Context	Center Denmark was set up as a neutral platform where universities, TSOs, DSOs, public organizations and private companies can work together to accelerate the green transition through intelligent data solutions. Center Denmark provides several services for data exchange for both national (Danish) and European projects, services and initiatives. Center Denmark provides data exchange services across the entire energy sector. Data can, for instance, show consumer consumption data for water, heat and electricity. Center Denmark offers data collection, sharing and data-driven services for supporting the green transition. Center Denmark provides data in near real-time and offers two-way streaming data exchange services, e.g., for controlling heat pumps in residential buildings. Center Denmark also offers solutions related to grid services; examples being voltage control and dynamic transformator rating. Using these services network operators can use the energy grids more efficiently. Center Denmark is becoming an European TEF (Test and Experimental Facility) for smart energy systems.
Vision of the Entity	Center Denmark is a non-profit organisation, and the vision is to support an acceleration of the green transition by providing smart and data-driven solutions for buildings and the industry. The solutions are built using open source software.
Stakeholders: their problems, their role and expected interaction with the Entity (including who are the data generators and data consumers?)	**Energy Utilities** can use Center Denmark for data collection and for assistance in providing digitalized operations of energy grids and systems. **Universities and Research Organisations** can get access to a broad spectrum of data related to energy and water. Examples of such data are energy meter data, meteorological data, grid and market related data. **End-users (both industry and residential house owners)** can use Center Denmark for optimizing the energy and emission efficiency. As an example a smart control of heat pumps for indoor pools in summerhouses can save approximately 30 pct. **Technology developers and systems integrators** can get access to data and tools for a development of new solutions which can be tested and validated.
Legal form of the Entity	Center Denmark was established in 2019, and the strategy and vision is controlled by the Center Denmark Fund.
Decision making positions and processes	Center Denmark Fund is an independent and non-profit organisation led by all relevant Danish Universities and Utility Operators for Electricity, Heat and Water.
Technology used to underpin data exchange	Center Denmark is using the Data Lake principles for storing the data. Data exchange is supported using a wide range of technologies such as RESTful API (or WEB API) and Apache Kafka.
Management of data sensitivities	Center Denmark has a Trusted Data Sharing system where the owners of the data are able to control who can access their data.
Commentary (eg (i) why the structure was chosen (ii) how the structure unlocks value, and (iii) financial self-sustainability)	Center Denmark started as a national (Danish) initiative, but now Center Denmark provides data sharing and streaming services for more than 10 countries (mostly in Europe). The Danish Energy Agency will, starting from January 2023, use the Center Denmark for Energy and Emission Accounting.

Super Low Energy Building (SLEB) Smart Hub, Singapore	
Context	In support of realising the targets of the Singapore Green Building Master Plan and Singapore Green Mark certification scheme (Singapore green building rating system), SLEB Smart Hub (www.sleb.sg) is set up under Green Buildings Innovation Cluster (GBIC) as an online resource centre that provides a variety of digital services to stakeholders in the green building value chain. The key features of SLEB Smart Hub include: • Data repository and business intelligence dashboards to store and visualise the national building energy data, green building data, green building-related research projects data, green technology data, and company data; • AI-assisted energy analysis tools – help the green building project team to do compliance check for meeting Green Mark certification requirements; • Energy efficiency self-assessment tools - help borrowers (building owners, business owners, and homeowners) to do self-assessment for their green building plan and submit the assessment reports to banks for green loan applications; • Energy-efficiency advisor – a recommendation system that recommends suitable green solutions based on the building profile; • Common energy dashboard – helps building owners and managers monitor, verify, and report their buildings' operational data and operational carbon emissions; • Data dictionary – standardises the data terminology and format to address the data interoperability challenges. With its abundant datasets and comprehensive data analytics tools, SLEB Smart Hub enables a variety of use cases for green building certification and green financing. It also helps overcome the information asymmetry among green building project stakeholders and thus facilitates more effective collaborations based on trusted data and tools.
Vision of the Entity	SLEB Smart Hub is envisioned as the leading digital platform for green buildings in the tropical region. It enables deep energy and carbon emissions savings in the built environment in the topics.
Stakeholders: their problems, their role and expected interaction with the Entity (including who are the data generators and data consumers?)	The stakeholders include building owners and developers, consultants and contractors, solution suppliers, research institutions, and financial institutions. **Building owners and developers** can refer to their buildings' energy benchmark reports auto-generated based on their buildings' data to gain insights of their buildings' energy efficiency performances and thus identify the potential areas for improvement. And if they want to source solutions and find companies to improve their buildings' energy performance, they can access the neutral database of green solutions and energy service companies. **Consultants and contractors** can learn green ideas from featured green building projects and published green building knowledge. And they can use energy efficiency analysis tools and benchmark data to streamline preparing the reports for green building certification submissions or green loan applications for their clients. **Solution suppliers** can use the listing service to create awareness of their emerging technologies and solutions. In the meanwhile, they can access to green building database to identify business opportunities. **Research institutions** can use the data to facilitate their research activities. And they can publish their research outcomes on the SLEB Smart Hub website which triggers collaboration with industries to help them translate their technology from laboratories to a commercially viable product. **Financial Institutions** can include self-assessment tools in their green loan framework. The tools are neutral and easy to use and help them save time and costs in evaluating the green loan application while preventing greenwashing. In addition, financial institutions can access green building data to facilitate their green loan and green bond issuance.

Legal form of the Entity	The SLEB Smart Hub has been run by BCA International Pte Ltd since 1 Jan 2022, a company fully owned by Singapore Building and Construction Authority (BCA). Previously SLEB Smart Hub had been run by BCA directly.
Decision making positions and processes	The SLEB Smart Hub is funded by Singapore National Research Foundation (NRF) under the BCA GBIC programme. GBIC programme office is in charge of developing and operating it.
Technology used to underpin data exchange	The data exchange standard is based on the Common Energy Dashboard (CED) data requirement released as part of the BCA Green Mark 2021 Standards in the year 2021. The data standards development process was facilitated by a few of the leading building management system (BMS) companies and utility companies. Application programming interfaces (APIs) and SFTP services are adopted to facilitate the data exchange between SLEB Smart Hub and 3rd partis systems and platforms. SLEB Smart Hub also allows users to report the data by uploading Excel or CSV files or by filling out web forms.
Management of data sensitivities	SLEB Smart Hub is consistent with government privacy and security practices as well as data protection and governance policies. A data classification mechanism is adopted. Users can access the “Open” data without logging in. Users have to log in to their authenticated user account to update and view their own data, which is classified as “Sensitive” data, by agreeing to the Terms of Use of the website and providing consent to sharing data.
Commentary (eg (i) why the structure was chosen (ii) how the structure unlocks value, and (iii) financial self-sustainability)	The SLEB Smart Hub was officially launched by then-Singapore National Development minister Mr. Lawrence Wong in September 2019. Over 400 organisations have signed on to be members. The take-up rate was boosted when Singapore banks adopted the SLEB Smart Hub self-assessment tools for their green building loan evaluation. SLEB Smart Hub enables more than 5 GWh energy savings and mobilises more than $ 1.5 billion (SGD) of green loans per annum. And it provides free data access to Singapore green building projects data which enables various use cases and services. SLEB Smart Hub adopted a multi-tenancy architecture, if 3rd party require customised data services or developing new Apps on top of SLEB Smart Hub platform, 3rd party needs to pay development fee and maintenance fee for such efforts.

Data Clearing House, Australia	
Context	Smart buildings utilise some combination of data, connectivity and AI to dynamically optimise energy use, IEQ, and occupant experience.Unfortunately, the property sector is a laggard in the adoption of such digitalisation technology. There is considerable uncertainty in the industry over cyber security and privacy. Access to data is hampered by interoperability issues and commercial lock-in to proprietary systems. The cost of retrofitting digital infrastructure in existing buildings is considered too high. The Data Clearing House platform (DCH) is an IoT platform. It uses open-standard, interoperable methods for data ingestion, and for data and metadata management. It addresses key barriers by providing a trustworthy, secure IT platform as a service (PaaS) that can be used by an ecosystem of innovators in the data analytics and building automation industry.
Vision of the Entity	The Data Clearing House aims to support growth of the data analytics and building automation industry, by reducing the risk and effort of integrating and maintaining disparate building data sources.
Stakeholders: their problems, their role and expected interaction with the Entity (including who are the data generators and data consumers?)	**Building Owners** have difficulty consolidating data into the one location and utilising it for value-adding services. They worry about the ongoing cost and risks associated with lock-in to the products of a single commercial vendor and the potential for data leakage and data loss. Building owners will host data from HVAC systems and other sensors and devices in the DCH data platform. This will give them streamlined access to data-driven services (energy, maintenance, occupant experience etc). **Data-Driven Software Service Providers** have a high cost associated with recruiting buildings to their services and with maintaining relevant digital infrastructure. Service providers will access buildings that have already been digitally enabled and onboarded onto the DCH platform. This will reduce their cost to provide energy saving analytics and electricity demand flexibility services.
Legal form of the Entity	The DCH Platform is owned and operated by CSIRO, Australia's National Research Agency. Day-to-day operation of the platform will be outsourced, and the plan is for management of the platform to be governed by industry users, as a data cooperative.
Decision making positions and processes	The governing board of the Data Clearing House will consist of elected users of the DCH platform, with representatives of building owners, software service providers and installers (system integrators). The Data Clearing House will have a sub-committee to prioritise ongoing improvements to the DCH data platform.
Technology used to underpin data exchange	The DCH is a real-time cloud-based data platform. The platform has drivers for ingesting data from diverse sources including MQTT, AMQP, RESTful APIs, FTP, BACnet, Modbus. Each user's data is stored on the platform in isolation from other users, with role-based authorisation functionality so that explicit permission can be provided for data sharing. The Brick Schema is used to model the relationships between data points and support automated querying of the database. Tools are provided for constructing the building data model when the building is onboarded onto the DCH platform. The DCH allows arbitrary data analysis workflows to be developed, containerised, and deployed on dynamically allocated cloud computing resources as software 'applications'. The application's execution environment is isolated from core platform services and other applications. The platform hosts a self-service measurement and verification application (M&V App), that can be used to support independent financial settlement of dispatchable flexible loads.
Management of data sensitivities	Template data sharing agreements are available. Users own their data, and control their data on the platform by explicitly providing or withdrawing permission for third parties to access their data. The Data Clearing House software platform is merely providing the functionality to process the data.

Commentary (eg (i) why the structure was chosen (ii) how the structure unlocks value, and (iii) financial self-sustainability)	The origins of the platform (from a public sector Research agency), and the anticipated data-cooperative governance arrangements, are designed to overcome key trust related barriers that have hampered development of the industry. Aspirationally, the structure supports delivery of possible government policy solutions including (but not limited to) (i) providing R&D infrastructure for digital innovation, (ii) providing independent measurement and verification services, and (iii) providing digital connectivity layer for an independent low voltage DER market operator.

Two further case studies, below, are commercial services that highlight certain open source technology characteristics. There are many other commercial platforms available, that could equally have been included. So inclusion here should not be taken as endorsement of these commercial products.

5. Idun Real Estate Solutions AB: is an example of a commercial data platform established to instantiate the open source RealEstateCore ontology as a standardized way of communicating real estate information. It enshrines customer data sovereignty principles in agreements.
6. Sensative AB: is an example of a commercial IoT platform provider that has built it's platform utilising the FIWARE framework of open-source software platform components. The Yggio IoT platform aims to unify different technologies and services, to address interoperability barriers.

Idun Real Estate Solutions AB	
Context	Idun's beginnings can be traced back to an idea sparked between a data scientist, an engineer, and a real estate company with a shared passion for making buildings smart. The real estate company was tired of compromising their smart building vision due to reliance on external providers. An alternative didn't exist, so they decided to join forces and create it themselves. The key to leveraging the full potential of smart buildings—and making smart cities happen faster—they realized, is giving property owners control over their own data. The result of this conversation was RealEstateCore, an open source semantic language developed in 2017 and sponsored by a consortium between Vasakronan AB, Akademiska Hus AB, Klipsk AB, Jönköping University, Rise AB and Willhem AB. Serving as a facilitator, RealEstateCore prepares buildings to interact with one another in the smart cities of the future. A set of integrated applications, powered by RealEstateCore, help facilitate smart buildings that contribute to greater sustainability, well-being, productivity, and better business. Idun's ProptechOS platform manages data and applications for your property portfolio . It is the easiest way to get started using the RealEstateCore semantic language and to transform your property portfolio into an application platform ready for optimization, analysis and new services.
Vision of the Entity	We make buildings good inhabitants of the smart city
Stakeholders: their problems, their role & expected interaction with the Entity	When property owners have access to this shared language, they are able to connect their buildings with new services and possibilities on a large scale—without having to worry about building or technology-specific implementation details and formats.
Legal form of the Entity	Idun Real Estate Solutions AB is a private company. Idun's ProptechOS platform uses RealEstateCore, an open source ontology.
Decision making positions and processes	RealEstateCore is published under MIT open-source licence, that means that RealEstateCore is free to use for anyone. As a member you get the opportunity to be part of a technical committee to and work actively within your specific domain of the development of RealEstateCore. As a member you support the development of RealEstateCore and get the right to use the RealEstateCore membership logo in your public communication.

Technology used to underpin data exchange	ProptechOS is an IoT Operating System designed for and by real estate owners that is making the world's buildings smarter faster. Data control is the key to leveraging the full potential of smart buildings. ProptechOS is built on the RealEstateCore ontology. RealEstateCore is a standardized way of naming and categorizing real estate data, making it possible to use the information of different building systems with each other. It also enables standardized communication from different technical real estate and external IT systems. This creates opportunities for advanced data analysis, intelligent control, and the monitoring of buildings, as well as visualization of property data in e.g. 3D models. RealEstateCore is open source and free to use without costs, limitations, or license requirements. For instance, all relevant stakeholders, such as architects, property owners, property managers, system suppliers, and construction contractors, can use the RealEstateCore-standard to similarly describe the interaction, data reading, and central control of several different properties. What is RealEstateCore ontology, and why is it important? Property owners can use RealEstateCore to describe data of interaction inside the properties they manage, along with storage, management, and the sharing of this data. Modular ontologies, such as REC, are a collection of data schemas. These data schemas describe different concepts and relations, referring to data generated in spatial models and/or technical building systems. It can also refer to data from external sources such as weather services or energy grid demand. Having a shared language allows property owners to connect their buildings with new services on a large scale. They will not have to worry about the details or formats of technology- or building-specific implementation. RealEstateCore aims to bridge existing industry standards The content of RealEstateCore is not new but based in part on existing standards applied with a pragmatic approach to finding the least common denominator. In this way, the gap between different existing industry standards is bridged. RealEstateCore focuses on binding together and bridging four different domains for standards: • Digital representation of the building's construction elements (e.g., BIM/IFC) • Control and operation of the building (e.g., Brick Schema, Project Haystack) • IoT-technology (e.g., SSN, WoT) • Business data for processes and agreements (e.g., CDM/IBPDI)
Management of data sensitivities	The Customer Owns all data. Customer shall solely own all data and other information, including but not limited to building data, which the Customer can collect through the use of the ProptechOS. Customer's right to hold replicate of data. The Customer has the right to hold and maintain a replica of all data collected through the use of ProptechOS in a system directly owned and controlled by the Customer. The Company has the obligation to make all data available to the Customer for such purposes. The Company's right to use the data. The Company shall keep the data in strict confidence. The Company has no right to use the data for any other reasons than what is necessary for the proper fulfilment of its obligations under this agreement. Upon expiry or termination of this agreement, the Company agrees to return to the Customer all data and any documentation and other material produced by the Company linked to the ProptechOS services pertaining to the Customer, and destroy all data and material from its ProptechOS services and archives with the exemption of what is required to maintain subject to law. The Customer has full access and availability to extract a copy of all data from ProptechOS at any time without any extra cost from the Company. We will maintain certain data that you transmit to ProptechOS for the purpose of managing the performance of ProptechOS, as well as data relating to your use of ProptechOS. Although we perform regular routine backups of data, you are solely responsible for all data that you transmit or that relates to any activity you have undertaken using ProptechOS. You agree that we shall have no liability to you for any loss or corruption of any such data, and you hereby waive any right of action against us arising from any such loss or corruption of such data.

Sensative AB	
Context	The lack of standardization in IoT is a problem plaguing several industries. With an enormous number of connected devices implemented everywhere, the need for data quality control and seamless collaboration between those devices has never been greater.
Vision of the Entity	Creating order out of chaos: turning IoT confusion into clarity
Stakeholders: their problems, their role and expected interaction with the Entity (including who are the data generators and data consumers?)	Today, most IoT systems are complete end-to-end solutions, providing everything from devices to the cloud and the application that creates a strong vendor and data lock-in effect. This contrasts with the digitalization vision, where access to actionable real-time data from operations is critical. The IT architect needs IoT capability to fit existing and future IT architecture and manage any protocol or technology. Sensative's game-changing Digitizing infrastructure Management System (DiMS), Yggio, acts as an open horizontal IoT platform for any Smart domain, like Smart Cities and Buildings. It provides integration of any sensor, IoT, or IT technology or supplier, contextualized and normalized data, brokering, unified IoT device and infrastructure management, scalability for massive IoT, and a standardized API for service developers, making IoT services vendor and technology-neutral. The Yggio platform gives the organization the control and ownership of the data they need for their digitalization journey, making data accessible and actionable.
Legal form of the Entity	Sensative is a private company founded in 2013. It was included in Sweden's 33 hottest tech companies in 2017, and one of 50 semi-finalists of the PropTech Startup and Scale-up Europe Awards 2021. Sensative was also recognized with Frost & Sullivan's 2021 Europe Technology Innovation Leadership Award for the IoT sensors market. Sensative's Yggio IoT platform has been officially approved as a **Powered by FIWARE-platform** The FIWARE Foundation is a non-profit association with the mission: "to build an open, sustainable ecosystem around public, royalty-free and implementation-driven software platform standards that eases the development of new Smart Applications in multiple sectors."
Technology used to underpin data exchange	**A common Digitalization infrastructure Management System bringing** • A single API to all smart things, legacy, and data • Supporting API & Data models • Complete control and data ownership • Manage any things and data and share access • Automation through rules, scenarios, Machine Learning and AI • Built for control, security, scalability, and robustness **FIWARE API and data model** FIWARE provides a curated framework of open source software platform components which can be assembled together and with other third-party components to build platforms that support the development of Smart Solutions faster, easier and cheaper. The main and only mandatory component of any "Powered by FIWARE" platform or solution is a FIWARE Context Broker Generic Enabler, supplying a cornerstone function required in any smart solution: the need to manage context information, enabling to perform updates and bring access to context. Yggio implements the FIWARE API and other standards and open source libraries (gray boxes) that minimizes lock-in effects. **Security:** The security layer handles identity and access management that protects the core services as well as the devices. **Modular:** Yggio consists of a set of micro services that all communicate over an AMQP standard bus. **Protocol connectors:** Yggio has a number of protocols implemented today, and new protocols are added continuously

Management of data sensitivities	**Personal Data Processing Agreement** In a relationship between Sensative and our customers, we sign a so-called Personal Data Processing Agreement. The agreements do not mean that our customer can transfer responsibility to Sensative, but that our customer delegates parts of managing their responsibility to Sensative. Sensative's responsibility is to follow the customer's guidelines and instructions as specified in the agreement and associated appendices. Sensative operates with a clear understanding that the data is our customer's property, with all that entails. Our IoT solution, Yggio, is developed according to "privacy by design," and our processes are well documented. The relationship with our customers and partners is always based on mutual respect and trust.
Commentary (eg (i) why the structure was chosen (ii) how the structure unlocks value, and (iii) financial self-sustainability)	The FIWARE open source components and open data models are building blocks to build Smart Solutions from, but it's not an IoT platform in itself. If a platform, like Sensative's Yggio, is based upon FIWARE, it has a high level of interchangeability, minimizing vendor lock-in. It makes FIWARE-based solutions a strong contender to the large technology vendors' IoT solutions.

5. Data Platform Design Recommendations Summary

Based on the discussion in the previous sections, the IEA 'Data-Driven Smart Buildings' Annex81 envisages an ideal world where operational data from non-residential buildings is (i) managed according to the FAIR data principles, (ii) respects ownerhip and privacy principles if/where personal data may exist, (iii) utilizes fit-for-purpose data platform technology to streamline utilization of data, and (iv) is supplied by an independent data institution to ensure non-competing access to an ecosystem of software analytics solution providers.

Noting these aspirations, this Section aims to provide a preliminary list of topics and questions that prospective clients could use to begin discussions with candidate data platform providers. It aims to help the client to identify a suitable data platform for their needs. The topics/questions are clustered into 11 thematic areas:

1. Governance
2. Data access controls and security
3. Data upload/building onboarding
4. Data capture
5. Data storage
6. Data exchange and on-platform programming
7. Data and application code recovery
8. Output signals and control
9. Applications marketplace
10. Screens and visualization
11. Platform development

Questions are both of a quantitative and qualitative type.

1. Governance

This theme seeks to understand the extent to which selecting the platform locks the client into a proprietary ecosystem that restricts future commercial options. Questions include

- Who owns the various layers of data platform IP? How is the IP protected?
- Does the platform use an external Infrastructure-as-a-Service platform such as Azure or AWS? What is the underlying storage (eg. AWS Timescale, AWS S3(or azure equivalents)? What is the storage costs (eg. $/GB for read and write)?
- Is any of the platform IP open source? If the platform is open source, please provide the link(s) to the repository/repositories.
- If the platform is not open source, is there any arrangement where the software can be unlocked sufficiently for a third party to manage platform operations and make improvements on behalf of the client?
- Are there any other ways that the platform can provide assurance that ongoing value for money can be provided, and that the user would not become commercially hostage?

2. Data Access Controls and Security

This theme seeks to understand the extent to which the platform protects the clients data from leaking outside of the intended recipients. Questions include

- Does the platform enable the client to have complete discretion/control over who has access to their data?
- Does this include the ability for the client to make parts of the data (within a building) available to a given third party (as opposed to all or nothing)?
- Does the platform enable the data client to make parts of the data (within a building) available in an anonymised manner on a data commons? If so, please provide additional comment to who or what is being anonymised, and how it is done.
- Does the platform include provisions for different tiers of access based on role, internally within the client organization and externally within authorised third parties?
- Can data security levels be set from building management level? Please describe the levels
- Is the platform hosted in the cloud?
- Where is data held? Could separate instances be hosted in other countries?
- Who is responsible for the security of the platform?
- What measures are in place to ensure that the risk of malicious attack is minimized? Does the platform include intrusion detection?

3. Data Upload/Building Onboarding

This theme seeks to understand how a client would get started bringing their buildings into the data platform. Questions include

- Does the platform have a process for registration and connection of the equipment, sensors and devices that exist in a building, to the platform (building onboarding) in order to establish real-time data acquisition?
- What connectors and device communication protocols does the platform support?
- Is there a standard hardware/software solution available to facilitate easy registration of assets and configuration of data upload at a site level? Please describe.
- Who can connect a building to the platform? Is the process of connecting the platform to a building well documented and standardised or is it undocumented or bespoke? Please describe and provide relevant links.

4. Data Capture

This theme seeks to understand the forms of data that can be captured and relevant data capture features and processes. Questions include

- What is the maximum sampling frequency for logging data to the platform? How soon after data collection can the data be reviewed by the user?
- Can the platform log digital data on change of state with an exact timestamp?
- Are there any limitations (other than cost) on the data capture approach?
- What data cleaning functions and data anomaly detection functions can be applied to data streams?
- Can the platform hold static/semi-static data such as asset databases, productivity information, documents, and drawings?

5. Data Storage

This theme seeks to understand how data is stored and managed and how that supports the FAIR (Findable, Accessible, Interoperable, Reusable) data principles. Questions include

- Can historical time series data sets, from diverse data sources, be compared at common time intervals?
- Is the data sorted and held in a structured schema (such as Brick) that enables data to be linked to physical spaces and systems within the building? What schema preferences do you promote?
- Are tools available to assist in sorting unstructured or poorly structured data into this schema?
- What programming language or tools are used to query the database?
- Are there any limitations (other than cost) on the amount of data stored?

6. Data Exchange and On-Platform Programming

This theme seeks to understand how third parties (particularly independent software analytics providers) can utilize the platform to deliver their innovative services. The theme identifies two pathways for analytics services to access data; (i) utilizing the platform simply as a source of data but processing that data on separate servers/IT infrastructure and (ii) direct hosting of software applications on the platform itself (similar to Apps on a mobile phone). Questions include

- Does the platform provide APIs to assist in data export to external platforms?
- Does the platform host third party data (eg Bureau of Meteorology) that can be accessed on platform?
- Does the platform support on-platform programming in one or more common open programming languages and software development kit (SDK)?
- Does the platform provide secure programming space? How does the platform protect app developer IP from other platform users?
- Does the platform provide for a programming-commons (microservices) so that common routines and programs can be shared by parties wishing to do so?

7. Data and Application Code Recovery

This theme seeks to understand the freedom that the user has to shift between platforms without loss of data or IP, and whether there are any implications associated with changing supplier. Questions include

- Can a client remove their data from the platform in a structured format so that they can take it somewhere else without losing information?
- Can an application developer with programming on the platform copy down compiled and uncompiled programming without losing information so that they can take it somewhere else?

8. Output Signals and Control

This theme seeks to understand whether the data platform supports two way communication (read/write), and whether that can be used to provide some level of remote supervisory automation of devices and/or push notifications. Questions include

- Can the platform send alerts and pop up notifications to assigned users?
- Can the platform provide simple alerts and flag level outputs to external software and devices? How many levels of priority of flag are available?
- Can the platform provide high-level-interface level outputs to control systems to enable more complex cloud control?

9. Applications Marketplace

This theme seeks to understand the extent to which the platform supports an ecosystem of software providers and software applications. Questions include

- Are third party software providers using the platform to provide data-driven software services? If so how is that governed?
- Does the platform provide a means for app developers to list/advertise their programs to clients?
- Does the platform provide a payment gateway for the charging of app fees?
- Does the platform have processes for vetting applications to ensure they meet minimum standards of functionality and security before they are accepted?

10. Screens and Visualization

This theme seeks to understand the user interface and what/how information and data is presented for users. It seeks information on what basic GUI screens are provided. This information could be obtained through a demonstration.

11. Platform Development

This theme seeks to understand the future journey and ongoing fit of the platform with future client needs. Questions include

- What organizations or partners are leading the development of the platform?
- What is the vision for the role of the platform in the industry, and how will that develop over time?
- Do the platform management arrangements support data stewardship roles consistent with that of a data institution?
- What capability developments are planned for the platform over the next three years, and how is that expected to improve the performance and value of the platform?

6. References

Bhattacharya, A., Ploennigs, J., & Culler, D. (2015). “Short paper: Analyzing metadata schemas for buildings: The good, the bad, and the ugly”. In Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments (pp. 33-34).

Candanedo, J., et al. (2023), “State-of-the-Art Report on Data-Driven Smart Buildings”, IEA Annex 81, Report (work in progress).

Commonwealth of Australia (2019a), “Best Practice Guide to Applying Data Sharing Principles”, https://www.pmc.gov.au/sites/default/files/publications/data-sharing-principles-best-practice-guide-15-mar-2019.pdf

Commonwealth of Australia (2019b), “Sharing Data Safely”, https://www.pmc.gov.au/sites/default/files/publications/sharing-data-safety-brochure-march-2019.pdf

Corradi, O., Ochsenfeld, H. P., Madsen, H., & Pinson, P. (2013). “Controlling Electricity Consumption by Forecasting its Response to Varying Prices”. IEEE Transactions on Power Systems, 28(1), 421–430. https://doi.org/10.1109/TPWRS.2012.2197027

Danish Energy Agency (2018), “Danish clarification on remote reading in relation to the data protection regulation”, https://ens.dk/sites/ens.dk/files/Forsyning/fjernaflaesning.pdf

Dwork C. (2006), “Differential privacy”, Proceedings of the 33rd International Conference on Automata, Languages and Programming - Volume Part II, ICALP’06, 2006, pp. 1–12.

Energy Efficiency Hub Digitalisation Working Group, (2022), “Roadmap on Digitalisation for Energy Efficiency in Buildings”, https://energyefficiencyhub.org/wp-content/uploads/2022/11/DWGRoadmap.pdf

European Commission, “Digitalising the energy sector – EU action plan”, https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13141-Digitalising-the-energy-sector-EU-action-plan_en, accessed January 2023

Fierro G. and Pauwels P., (2022), “Survey of metadata schemas for data-driven smart buildings”, report for IEA Annex81 ‘Data-Driven Smart Buildings’, https://annex81.iea-ebc.org/Data/publications/IEA%20Annex%2081%20Survey%20of%20Metadata%20Schemas.pdf

General Services Administration (GSA), 2015, “GPG-025, Socially Driven HVAC for Personal Control”, https://www.gsa.gov/cdnstatic/GPG_Findings_025-Socially-driven_HVAC_1535393905.pdf

Group of Experts on Energy Efficiency (2021) “Improving Efficiency of Buildings through Digitalisation – Policy Recommendations from the Task Force on Digitalisation in Energy”, Economic Commission for Europe, Eighth session, Geneva, 20-21 September 2021, https://unece.org/sites/default/files/2021-06/ECE_ENERGY_GE.6_2021_5_Policy%20recommendations_final.pdf

Harbor Research (2020), “Intelligent Building Energy Management Systems”, report for the Continental Automated Buildings Association, https://www.caba.org/intelligent-building-energy-management/

Hardin D.B., Corbin C.D., Stephan E.G., Widergren S.E., and Wang W., (2015), “Buildings Interoperability Landscape”, Pacific Northwest National Laboratory, Report for the U.S. Department of Energy under Contract DE-AC05-76RL01830, https://www.energy.gov/sites/default/files/2016/01/f28/BuildingLandscapeReport.pdf

International Energy Agency (2017), “Digitalisation & Energy”, https://iea.blob.core.windows.net/assets/b1e6600c-4e40-4d9c-809d-1d1724c763d5/DigitalisationandEnergy3.pdf

International Energy Agency (2021), “Energy Efficiency 2021”, https://www.iea.org/reports/energy-efficiency-2021

Junker, R. G., Azar, A. G., Lopes, R. A., Lindberg, K. B., Reynders, G., Relan, R., & Madsen, H. (2018). Characterizing the energy flexibility of buildings and districts. Applied Energy, 225, 175–182. https://doi.org/10.1016/j.apenergy.2018.05.037

Junker, R. G., Kallesøe, C. S., Real, J. P., Howard, B., Lopes, R. A., & Madsen, H. (2020). “Stochastic nonlinear modelling and application of price-based energy flexibility”. Applied Energy, 275(1), 115096. https://doi.org/10.1016/j.apenergy.2020.115096

Kjærgaard M.B., Ardakanian O., Carlucci S., Dong B., Firth S.K, Gao N., Huebner G.M., Mahdavi A., Rahaman M.S., Salim F.D., Sangogboye F.C., Schwee J.H., Wolosiuk D., and Zhu Y. (2020), “Current practices and infrastructure for open data based research on occupant-centric design and operation of buildings”, Building and Environment 177, 106848

Kramer H., L.G., Curtin C., Crowe E, and Granderson J. (2020), “Proving the Business Case for Building Analytics”. Lawrence Berkeley National Laboratory, https://eta-publications.lbl.gov/sites/default/files/kramer_provingbuildinganalytics_october2020.pdf

Lawrence N., and Oh S., (2021), “Enabling data sharing for social benefit through data trusts - An Interim Report for the 2021 GPAI Paris Summit”, produced by the AAPTI Institute and the Open Data Institute for the Global Partnership on Artificial Intelligence, https://gpai.ai/projects/data-governance/data-trusts/enabling-data-sharing-for-social-benefit-data-trusts-interim-report.pdf

Locatee and Memoori (2017), “Navigating the Complex Smart Building Landscape: A Comprehensive Use Case Guide for Corporate Real Estate Professionals”, https://locatee.com/en/blog-post/navigating-the-complex-smart-building-landscape/

Machanavajjhala A., Gehrke J., Kifer D., Venkitasubramaniam M. (2006), “L-diversity: privacy beyond k-anonymity”, in: ICDE’06, 2006, 24–24.

Madsen, H., Parvizi, J., Halvgaard, R. F., Sokoler, L. E., Jørgensen, J. B., Hansen, L. H., & Hilger, K. B. (2015). “Control of Electricity Loads in Future Electric Energy Systems”. Handbook of Clean Energy Systems.

Memoori (2021), “AI & Machine Learning in Smart Commercial Buildings”, https://memoori.com/portfolio/ai-machine-learning-in-smart-commercial-buildings/

O’Reilly (2020), “The state of data quality in 2020”, https://www.oreilly.com/radar/the-state-of-data-quality-in-2020/

Otte K., Stelmach T., Chandan V., Evans M., Delgado A. (2022) “Digitalisation in the Buildings Sector: A Literature Review”, PNNL, Prepared for the U.S. Department of Energy under Contract DE-AC05-76RL01830

Scerri S., Tuikka T., Lopez de Vallejo I., and Curry E. “Common European Data Spaces: Challenges and Opportunities”, Data Spaces, pp337-357, https://doi.org/10.1007/978-3-030-98636-0_16

Sweeney L., (2002) “k-anonymity: a model for protecting privacy”, International Journal of Uncertainty, Fuzziness Knowledge-Based Systems. 10, 2002, 557–570, 05.

The Open Data Institute, “Data Institutions”, https://www.theodi.org/service/consultancy/data-institutions/

Thilker, C. A., Madsen, H., & Jørgensen, J. B. (2021). “Advanced forecasting and disturbance modelling for model predictive control of smart energy systems”, Applied Energy, 292, 116889. https://doi.org/10.1016/j.apenergy.2021.116889

Trianni A., Bennett N., Hasan A.S.M., Katic M., Lindsay D., Cantley-Smith R., Wheatland F.T., White S., Dunstall S., Leak J., Pears A., Cheng C.-T. And Zeichner F., (2022), “RACE for Business - Opportunity Assessment: Industry 4.0 for Energy Productivity”, https://www.racefor2030.com.au/opportunity-assessment-reports/

US Department of Energy (2021), “A National Roadmap for Grid-Interactive Efficient Buildings”, https://www.energy.gov/eere/buildings/grid-interactive-efficient-buildings

Valdez A., Nubbe V., Thakkar M., Reich J., Goetzler W. (2020), “Policy Guidance for smart energy saving consumer devices”, The Electronic Devices & Networks Annex of the IEA 4E Technology Collaboration Programme https://www.iea-4e.org/wp-content/uploads/publications/2020/12/Policy_Guidance_for_Smart__Energy-Saving_Consumer_Devices_May_2020.pdf

Wilkinson M. D. et al. (2016), “The FAIR Guiding Principles for Scientific Data Management and Stewardship”, Scientific Data, 3:160018, https://www.nature.com/articles/sdata201618

Zhang K., Blum D., Cheng H., Paliaga G., Wetter M. and Granderson J., 2022, “Estimating ASHRAE Guideline 36 energy savings for multi-zone variable air volume systems using Spawn of EnergyPlus”, Journal of Building Performance Simulation, 15(2), 215-236, https://www.tandfonline.com/doi/full/10.1080/19401493.2021.2021286

Zimmerman G. (2021), “How to Make Smart Building Data Smarter”, https://www.facilitiesnet.com/buildingautomation/article/How-to-Make-Smart-Building-Data-Smarter--19396

ANNEX 81

EBC

Energy in Building and Communities Programme

www.iea-ebc.org