

Review

Big Data Value Chain: Multiple Perspectives for the Built Environment

Gema Hernández-Moral ^{1,*}, Sofía Mulero-Palencia ¹, Víctor Iván Serna-González ¹, Carla Rodríguez-Alonso ¹, Roberto Sanz-Jimeno ¹, Vangelis Marinakis ², Nikos Dimitropoulos ², Zoi Mylona ³, Daniele Antonucci ⁴ and Haris Doukas ²

¹ CARTIF Technology Centre, Parque Tecnológico de Boecillo, Boecillo, 47151 Valladolid, Spain; sofmul@cartif.es (S.M.-P.); vicser@cartif.es (V.I.S.-G.); caraln@cartif.es (C.R.-A.); robsan@cartif.es (R.S.-J.)

² Decision Support Systems Laboratory, School of Electrical & Computer Engineering, National Technical University of Athens, 15780 Athens, Greece; vmarinakis@epu.ntua.gr (V.M.); ndimitropoulos@epu.ntua.gr (N.D.); h_doukas@epu.ntua.gr (H.D.)

³ HOLISTIC IKE, 15343 Athens, Greece; zmylona@holisticsa.gr

⁴ Institute for Renewable Energy, Eurac Research, 39100 Bozen/Bolzano, Italy; daniele.antonucci@eurac.edu

* Correspondence: gemher@cartif.es

Abstract: Current climate change threats and increasing CO₂ emissions, especially from the building stock, represent a context where action is required. It is necessary to provide efficient manners to manage energy demand in buildings and contribute to a decarbonised future. By combining new technologies, such as artificial intelligence, Internet of things, blockchain, and the exploitation of big data towards solving real life problems, the way could be paved towards smart and energy-aware buildings. In this context, the aim of this paper is to present a critical review and an in-detail definition of the big data value chain for the built environment in Europe, covering multiple needs and perspectives: “policy”, “technology” and “business”, in order to explore the main challenges and opportunities in this area.

Keywords: big data; artificial intelligence; machine learning; analytics; building stock



Citation: Hernández-Moral, G.; Mulero-Palencia, S.; Serna-González, V.I.; Rodríguez-Alonso, C.; Sanz-Jimeno, R.; Marinakis, V.; Dimitropoulos, N.; Mylona, Z.; Antonucci, D.; Doukas, H. Big Data Value Chain: Multiple Perspectives for the Built Environment. *Energies* **2021**, *14*, 4624. <https://doi.org/10.3390/en14154624>

Academic Editor:
Dimitrios Katsaprakakis

Received: 23 June 2021
Accepted: 26 July 2021
Published: 30 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Current climate change threats and increasing CO₂ emissions, in particular from the building stock, represent a context where it is necessary to act upon and provide efficient manners to manage energy consumption and generation in buildings and contribute to a decarbonised economy. Within buildings today, more and more data are being generated, taking into consideration the increasing adoption of information and communication technologies (ICT), such as artificial intelligence (AI), Internet of things (IoT), and distributed ledger technology (DLT)/blockchain [1].

By combining these new technologies and the exploitation of big data towards solving real life problems surrounding buildings, the way could be paved towards a decarbonised future [2]. To this end, it is necessary to understand what technologies and solutions already exist, as well as how existing methods, processes, data models, and platforms could be exploited in the building sector. The aforementioned series of steps and information are crucial to generate value and useful insights from building data. They constitute the big data value chain (BDVC), which is considered to play key role in the future energy economy, bringing opportunities of digitalisation and accelerating the energy transition via the enhancement of building sector energy performance. In terms of an increase in energy efficiency and building performance, as well as intelligent energy management of the built environment, big data can play an important role. As an example, although a discrepancy between predicted and actual performance is unavoidable, solutions are necessary to explain and narrow this gap to make more accurate forecasts and reduce energy demand as much as possible.

To this end, an in-detail definition of the BDVC for the built environment is required so as to understand, control, and eventually improve it. Big data in the built environment can cover different purposes such as monitoring and predicting energy performance, verification of energy improvement considering thermal comfort perspectives, building performance simulation, and policy and decision making.

Methodology

The aim of this paper is to define these multiple perspective approach of the built environment covering needs for “policy”, “technology”, and “business” fields and identifying opportunities towards the enhancement of the BDVC in Europe. These three fields aim to provide holistic view of the challenge by delving deeper into the high-level context (policies and trends), analysing existing technologies and data, and making a link towards real-life application and how the analytics could improve decision-making and business-as-usual practices. More in detail, the methodology can be described as follows:

- **Policy perspective:** the main objective followed was to present the main directives and initiatives in the field of energy, big data, and digitalisation that could play a role and have an impact on the BDVC in the European context.
- **Technology perspective:** this perspective focuses on frameworks, architectures, and datasets that are currently present and could support the BDVA. In particular, the focus has been placed on exploring the role of digital twins as a key element to help connect energy challenges among different scales (buildings to cities). To complement this vision, an in-depth but non-exhaustive analysis of building-related datasets and repositories is presented. Both EU and national-level repositories have been selected, and all of them are public and are related to the building stock, energy fields, or other contextual data (statistical, geographic, geometric, or meteorological data).
- **Business perspective:** given the current context and trends, as well as the technological possibilities, this perspective explores how this can be combined with powerful analytics to improve decision making. In this context, four main challenges are addressed, which correspond to the categories of pilots deployed in the MATRYCS H2020 project: (1) energy performance, (2) building and related infrastructure design and refurbishment, (3) policymaking, and (4) energy efficiency financing. In addition, cross-cutting support is explored through geo-clustering methods.

Thus, the paper is organised following this classification. In Section 2, a detailed review of the big data concept in the built environment, as well as directives and initiatives of the European Union (EU), is presented (policy perspective). In Section 3, an extensive review of existing frameworks, architectures, and existing datasets across EU and at the national level is provided (technology perspective). In Section 4, particular emphasis is placed on analytic services and business models for the built environment (business perspective). The contributions and conclusions of the current study are discussed in Section 5.

2. “Policy” Perspective in Europe: Big Data Concept in Buildings

A number of European Directives and initiatives are setting strict objectives on Member States in order to meet targets set for 2030. To this end, significant changes and investments are required such as an increase in the renewable energy sources (RES) capacity and buildings stock energy efficiency improvement [3].

2.1. European Directives and Regulations

In addition to the common framework established by the European Green Deal [4], of particular relevance are the Energy Performance in Buildings Directive (EU) 2018/844 and the Energy Efficiency Directive (EU) 2018/2002, both part of the ‘clean energy for all Europeans’ package. It is also worth highlighting the initiative of the Renovation Wave [5], which focuses on boosting building renovation for climate neutrality and recovery. Moreover, existing directives and regulations of the European Commission (EC) related to thematic areas, such as big data, AI, and IoT, are the “European Strategy for

Data” [6], “Proposal for a Regulation on European data governance” [7], “White Paper on Artificial Intelligence” [8], “Report on the safety and liability implications of Artificial Intelligence, the Internet of Things, and robotics” [9], “Ethics Guidelines for Trustworthy Artificial Intelligence” [10], “The Open Data Directive” [11], “General Data Protection Regulation” [12], and “Open-source software strategy 2020–2023” [13].

2.2. ECTP Vision for Buildings towards 2030

The European Construction Technology Platform (ECTP), as a leading organisation shaping the future of the built environment and construction sector in Europe, recently issued its Strategic Research and Innovation Agenda (SRIA) [14]. Long-term and intermediate goals (for 2050 and 2030) have been established. Climate changes and associated policies towards a CO₂ neutral society are among the main drivers for the transition in the building sector, which call for rethinking the design, maintenance, and management of buildings, as well as for accelerating the renovation of the housing stock and the integration of renewable energy sources.

2.3. BDVA Vision for Data towards 2030

The Big Data Value Association (BDVA) has been operative since 2014 and is now becoming DAIRO (standing for Data, AI, and Robotics), with the mission to develop the innovation ecosystem that will enable the data and AI-driven digital transformation in Europe [15]. In the BDVA position paper of November 2020 [16], they envisaged the main bottleneck for the exploitation of AI technologies, as a primary driver of the data economy, to be widespread, secure, and effective data sharing.

2.4. Common European Digital Platform and Collaborative Networks

A number of digital platforms and collaborative digital networks in Europe (Table 1) aim to gather information on good practices and achievements by each centre involved, as well as work in line with other European initiatives. Some have the ultimate goal of defining the future of Europe on the basis of data, as well as creating more transparency and visibility, while others focus on energy in buildings with the objective of gathering knowledge and fostering exchange, as well as taking up innovative and effective measures.

Table 1. EU digital platforms and collaborative networks.

Platform	Short Description	Key Stakeholders
BUILD UP [17]	The BUILD UP initiative supports EU Member States in implementing the Energy Performance of Buildings Directive (EPBD).	Professionals working in the building sector with an interest in energy efficiency
European Energy Efficiency Platform (E3P) [18]	The E3P among other tasks facilitates the practical implementation of the Energy Efficiency Directive at national, regional, and local levels, with data collection and analysis.	Energy efficiency experts in a wide range of thematic areas
Coalition for Energy Savings [19]	The Coalition for Energy Savings is a common advocacy platform to promote and mainstream energy efficiency at the European level, a centre of expertise on energy efficiency, and a forum to exchange intelligence on energy efficiency.	Businesses, professionals, local authorities, trade unions, consumers, and civil society organisations
Housing Evolutions Hub [20]	The Housing Evolutions Hub highlights the latest innovations in the field of social, public, affordable, and responsible housing.	Practitioners and policymakers in the area of housing/energy efficiency/sustainability of buildings and neighbourhoods
AI4EU [21]	AI4EU was established to build the first European Artificial Intelligence On-Demand Platform and Ecosystem.	Wide range of actors including scientists, entrepreneurs, SMEs, industries, funding organisations, and citizens

3. “Technology” Perspective: Digitalisation of the Built Environment

The digitalisation of the built environment is gaining ground and is exponentially growing today, even more so with the technological advancements and data exploitation capabilities that are becoming more and more available. On this basis, the generation of digital twins as digital data repositories to support specific processes is becoming more and more frequent. However, relevant challenges need yet to be overcome; building data are only partially available, rarely up to date, and almost never integrated into a single platform so that informed decisions can be made at the micro- and the macroscale. A wealth of opportunities exists through the combination of both static and dynamic data through the generation of analytics, leading to more accurate and cost-effective solutions, as well as the ease of analysis processes.

3.1. Digital Building Twin

The innovative services, applications, and technologies surrounding the digital building twin are analysed in the paragraphs below (i.e., existing data models, BIM and digital infrastructures, end-to-end process digitalisation, and how other scales can be tackled).

3.1.1. Data Models

Existing data models, relevant in the context of the building sector, are FIWARE Smart Data Models [22], the Industry Foundation Classes (IFC) [23], the international standard City Geography Markup Language (CityGML) [24], the INSPIRE Directive (2007/2/EC) [25], the Smart Applications REference (SAREF) ontology [26], and BRICK ontology-based metadata schema [27]. These data models have been harmonised in such a way as to enable data portability for different applications in different domains, when observed from different perspectives: as individual buildings, as a group of buildings (districts), as a group of districts (cities), etc.

3.1.2. BIM and Digital Infrastructures

Building information modelling (BIM) can be defined as a digital toolset used to digitally map buildings or infrastructure facilities for various purposes: visualisation, scheduling, communication, and collaboration between stakeholders through a building/facilities life cycle [28]. BIM allows various users to extract and analyse the data to make decisions and improve the process of delivering the building [29]. However, implementing BIM within the company requires a considerable build-up expertise, especially appropriate employee training and substantial ICT upgrading. This can be especially challenging for small and medium enterprises (SMEs), creating the need for extra investment into new technologies and training which may be difficult and not always possible. According to Poljansek [30] in order to successfully implement BIM in the construction sector, interoperability must be ensured; in this way, various stakeholders can share the information and cooperate the planning. Building owners and investors can slowly adopt the BIM technology until they manage to understand all the benefits of using BIM.

3.1.3. IoT and Integration with BIM-Enabled Platform

IoT devices make use of several protocols to connect to the internet and send data to the data repository (usually, data repositories are proprietary software from companies). Examples of these protocols are LoraWan, Sig-Fox, MQTT, NB-IoT, etc. What they have in common is the light use of bandwidth; usually, devices transfer little information, such as temperature, in a not very high frequency.

These data are very useful in the development of a digital twin of a building (district or city when available) because it is possible to know the behaviour of the building by means of the study of data evolution in time. However, there is a quite important lack of information regarding the location of the IoT devices. Usually, this information is not used by the IoT companies/platforms, and it is not incorporated into the datasets available to

download. Commonly, only information about the name (that is selected by the particular user as desired) is available in the dataset, and it is usually not representative.

To make a proper match among devices, their measurements, and the digital twin, it is necessary to know where these devices are installed and what they are measuring. In the particular case of a building, while considering the standard IFC4 and the definition of spatial structures and the definition of devices included, it is needed to match devices to spaces.

3.1.4. City Digital Twin

Digital twin technology has also been applied at city and regional scales. It is important to highlight that, at these scales, the level of detail expected is not the same that at the building scale; it is unlikely that all the needed information will be available for a perfect BIM model for all the buildings in a city or region. Using a digital twin as a city model is an exceedingly difficult concept. In product and plant design, digital twins have been used for decades, to create, test, and build everything first in a virtual environment. The use of digital twins in a city concept is a larger vision compared to industrial design or a building digital twin. This can help with simulations and analysis of present and planned city environments, maintenance and administration systems, emergency planning and management, and the construction production chain. Additionally, the impact of the COVID-19 pandemic is forcing cities to accelerate their digital transformation.

City digital twins are based fundamentally on topographic and geometric models of the city infrastructure. CityGML has been used as a data model for the representation of urban objects in 3D. It defines classes and relationships for the most relevant topographic objects in city models with respect to properties: geometric, topologic, and semantic data. An example of this is Helsinki's 3D data model [31] (offered also as Virtual Helsinki [32]), a digital city twin that in part was created using CityGML 2.0 (especially for the Kalasatama digital twin [33]).

3.2. Building-Related Datasets and Repositories

Available open datasets in Europe were identified and classified on the basis of the level that the reported information refers to: public EU repositories (Table 2) and public national repositories (Table 3). The identified datasets are not only datasets containing building information, but also datasets related to the energy performance of the buildings, districts, and regions. The majority of the datasets that report information at an EU level also contain information about each individual Member State.

Table 2. Public EU repositories.

Repository	Description
EU Building Stock Observatory [34]	It monitors the energy performance of buildings across Europe. Among the features under supervision, the energy efficiency levels in buildings (EU countries /EU as a whole), certification schemes, financing aspects, and levels of energy poverty throughout the EU can be mentioned.
EU Energy Poverty Observatory [35]	It is an initiative by the EC to help Member States combat energy poverty. The approach is to use a set of indicators that individually capture a slightly different aspect of the phenomenon.
EUROSTAT [36]	It is responsible for publishing high-quality statistics and indicators at the European level that allow comparisons between countries and regions.
Statistical Review of World Energy [37]	This report analyses data on world energy markets from the prior year. The review has provided timely, comprehensive, and objective data to the energy community since 1952.
TABULA EPISCOPE [38]	A concerted set of energy performance indicators is given. It is focused on residential building typologies and contains data about buildings' energy needs, costs, demand, emissions, etc. as a function of climate zone, construction year classes, and buildings' characteristics.
ENTRANZE [39]	It aims to support policymaking process, by providing the required data, analysis, and guidelines to achieve a fast and strong penetration of nZEB and RES-H/C within the existing national building stock.

Table 2. Cont.

Repository	Description
ODYSSEE—Full database [40]	It contains energy and macroeconomic data at an economy-wide level and environmental indicators at an economy-wide and sectoral level (industry, transport, residential, services, and agriculture) over 2000–2018.
ODYSSEE—Key indicator tool [41]	It contains saving rates and consumption data at an economy wide and sectoral level and offers the results in a geolocated form.
ODYSSEE—Decomposition tool [42]	This online-web tool decomposes the energy use into various explanatory effects.
ODYSSEE—Market diffusion tool [43]	This tool reports indicators reflecting the market diffusion of various energy efficient technologies.
ODYSSEE—Comparison tool [44]	This tool enables the comparison of two countries in terms of their energy efficiency performance at an economy-wide and sectoral scale.
ODYSSEE—Energy saving tool [45]	This tool displays the trends and targets for the primary and final energy consumption, as well as the energy savings at a national level.
ODYSSEE—EU energy efficiency scoreboard [46]	This tool scores EU countries on (a) the energy efficiency level, (b) the energy efficiency progress, (c) the energy efficiency policies, and a combination of all these criteria.
MURE database [47]	It provides information on energy efficiency policies and measures that have been carried out in the Member States of the EU (as well as Norway, Switzerland, and Serbia).
EnergyPlus Weather Data [48]	EnergyPlus is an open-source whole-building energy-modelling engine. Weather data for more than 2100 locations are available in its weather format, which are arranged by World Meteorological Organisation region and country.
Climatic Research Unit (CRU) [49]	The objective of the CRU is to improve the scientific understanding of the climate system and its interactions with society. It contains weather data, monthly and annually.
European Environment Agency (EEA) [50]	EEA focuses on providing data related to environmental policies and other topics related to the environment, taking advantage of its extensive network.
Our World in Data [51]	It contains information about macroeconomic and energy-related variables.
European Open Data Portal [52]	It represents the access point to data institutions, agencies, and other bodies of the EU.
European Data Portal [53]	It collects the metadata of the public sector information available on the public data portals of EU countries.
World Bank [54]	It contains information about the majority of macroeconomic and energy-related variables.
Organisation for Economic Co-operation and Development (OECD) [55]	Its website contains information about the majority of macroeconomic and energy-related variables.
Covenant of Mayors [56]	Its website contains information about the climate mitigation measures and targets set per municipality, as well as the estimated impacts in terms of estimated greenhouse gas emissions reduction per sector.
The shift data portal [57]	Its website contains information about national macroeconomic and energy statistics.
United Nations Statistics Division website (UNSD) [58]	Its website contains information about national macroeconomic and energy statistics.
Ember [59]	Its website offers interactive tools that report statistics about energy systems.
Climate Fund Inventory Database [60]	It supports recipient countries, least developed ones in particular, by providing consolidated information on the number and types of climate funds that are available.
The Carbon Centre [61]	It aims to support cities, towns, and regions tackling climate change (CDP and ICLEI are partnering to present one unified process for subnational climate action reporting). This site contains information about cities' climate mitigation measures, climate targets, and performance in terms of carbon emissions
Open Street Map (OSM) [62]	Maps are created using geographic information captured with mobile GPS devices, orthophotos, and other free sources.
Copernicus data [63]	It offers information services based on Earth observation and “in situ” data covering six thematic areas: atmosphere monitoring, marine environment monitoring, land monitoring, climate change, emergency management, and security.
HotMaps [64]	Values related to final energy consumption and useful energy demand for space heating, space cooling and domestic hot water, construction materials and methodologies, technologies used, and building stock data/information can be found both for the residential and the non-residential sectors per building types and construction vintages.
ZEBRA [65]	It contains information related to energy performance certificates, materials employed for the buildings, energy performance, and final energy consumption, among others.

Table 2. Cont.

Repository	Description
CommONEnergy [66]	It includes building sector data and final energy demand data for non-residential buildings, especially focusing on the trade sector.
Integrated Database of the European Energy System (JRC IDEES) 2015 [67]	JRC IDEES offers a set of disaggregated energy–environment–economy data, compliant with the EUROSTAT energy balances, as well as widely acknowledged data on existing technologies. It also contains a plausible decomposition of final energy consumption.
ExcEED [68]	A European database for measured and qualitative data on beyond the state-of-the-art buildings and districts.
iNSPiRe [69]	Building stock analysis and data gathering exercise focusing its attention on published literature and other sources, aiming to extrapolate information about the current residential and office building stock.
ZENSUS 2011 [70]	This dataset contains disaggregated data concerning a building stock analysis for Germany, information about the occupancy of the buildings, and socioeconomic-related data. Information concerning the type of heating systems used is also reported.
Towards a sustainable Northern European housing stock—Sustainable Urban Areas 22 [71]	It contains complete data for a building stock analysis with data varying from state to state between 2000 and 2006. Data concerning material used and heating, ventilation, and cooling systems installed are also reported. Construction/demolition rates (1980–2004) have been added to the report.
DEEP [72]	DEEP is an open-source database for energy efficiency investment performance monitoring and benchmarking. It provides an exhaustive analysis of the performance of energy efficiency investments in order to support the assessment of the related benefits and financial risks.
D’Agostino et al. [73]	It provides an overview on the results of the data collected by the Green Building Programme (GBP) and its main results from the launch in 2006 up to its completion in 2014. It focuses on building characteristics, energy performance, efficiency measures, and energy savings.
National Housing Census: European statistical System [74]	This dataset contains a variety of data collected in relation to the national census performed in 2011 by EU27 + UK Member States. It is possible to find data concerning households, such as the number of components of single households at a granularity until NUTS3 level.
EDGAR [75]	Carbon dioxide (CO ₂) emissions by country and sector (buildings, transport, other industrial combustion, power industry, and other sectors) have been collected for the years between 1970 and 2018 and are reported expressed in MtCO ₂ /year.
CORDEX [76]	Climatic data for Europe expressed as daily, monthly, and seasonal mean values, as well as at 3 or 6 h resolution. Data for air temperature at 2 m, wind speed, atmospheric pressure, and humidity can be found.
PVGIS—Photovoltaic Geographical Information System	This GIS dataset contains data related to the solar radiation. It considers both day- and night-time periods, expressing the solar radiation raster map in W/m ² .

Table 3. Public EU national repositories.

Country	Repository	Description
Italy	Open Data Hub Italia [77]	It provides the most complete catalogue of Italian open data.
	EPC cadastre of Lombardia Region [78]	This database provides information related to energy verification, primary energy demand, transmittance (u-value) of façade elements, thermal production systems and emission systems, and photovoltaic and solar panels.
	GreenDataset [79]	It includes detailed power usage information, obtained through a measurement campaign in households in Austria and Italy
Slovenia	Portal energetika [80]	National portal where data on energy efficiency, RES production, energy certificates of buildings, energy management, etc. are collected.
	OPSI [81]	It is a single national website for the publication of open data for the entire public sector.
Poland	Geoportal [82]	Data (cloud point) from Airborne Laser Scanning for Poland (ALS), land development, land developments plans, and cadastral data.
	EPC register [83]	This database covers only public office buildings.
Spain	Spanish Cadastre data [84]	It makes the cadastral data of the territory under its jurisdiction available to citizens (almost the entire national territory). Information about properties’ cadastral information is organised by municipality, and it is INSPIRE-complaint.
	AEMET OpenData [85]	AEMET OpenData is a system for the dissemination and reuse of AEMET information. The State Meteorological Agency of Spain is a state agency whose objective is the provision of meteorological services, which are the responsibility of the State.

Table 3. Cont.

Country	Repository	Description
	INE Open data [86]	The National Statistics Institute has created the Open data space in order to include the public information resources generated in it.
	CNIG Download Centre [87]	This website provides digital geographic information produced by the National Centre of Geographic Information.
	BIM Document Library [88]	This site contains a dashboard with different reference documents published by the main actors, as guides, manuals, standards, reports, etc. in different countries.
Portugal	Open data Portal [89]	Dados.gov is the Portuguese Public Administration's open data portal. Its function is to aggregate, reference, and store open data from different public administration bodies and sectors, thus creating the central catalogue of open data in Portugal.
Germany	Bauwerksdatenbank [90]	Physical constitution of the built environment—Database of Buildings and Infrastructure.
	Greece National Data Portal [91]	Datasets for central government, local authorities, and public bodies.
Greece	Open Data Greece [92]	It provides open geospatial data and services for Greece, serving as a national open data catalogue, an INSPIRE-conformant spatial data infrastructure, and a powerful foundation for enabling value added services from open data.
Latvia	Latvia's Open Data portal [93]	The aim of this data portal is to gather and circulate government institution and organisation collected data in one place for public use.
Belgium	Belgium Government Open data [94]	This website is related to the Belgian Open Data Initiative.
	ENEX [95]	Czech national database of energy audits and EP certificates.
Czech Republic	Energo 2015 [96]	The Czech Statistical Office provides information on energy consumption of households, available for the year 2015.
United Kingdom	Live tables on EPC [97]	Data from the Energy Performance of Buildings Registers since 2008 (nondomestic and domestic properties), including average energy efficiency ratings and energy use.

Other relevant public repositories can be found and exploited at the regional level, such as the Energy DataHub [98] and EPC Register [99] from the region of Castilla y León. Existing private data sources that could provide useful information are “Enerdata” related to national energy statistics (e.g., energy demand, CO₂ emissions) and the International Energy Agency (IEA, Paris, France), which contains information about most energy-related variables (e.g., energy consumption, energy intensity, energy prices) at a national and EU level.

4. “Business” Perspective: Data Analytics for the Built Environment

Data analytics in the activities and businesses related to the building sector and its life-cycle can identify useful information patterns within large datasets that can be transformed into actionable outcomes/knowledge to support improved decision making [100,101]. Data analytics applications can support energy management and optimisation, building and related infrastructure refurbishment and design, policymaking, investment de-risking, and a big data vision through geo-clustering.

4.1. Data Analytics for “Energy Performance”

A number of analytic services have been developed, focused on the operational stage of the buildings (Table 4). These services are aimed at monitoring, analysing, and improving the energy performance of buildings, as well as incorporating predictive capabilities related to comfort evaluation, energy demand, consumption, generation, and other building capabilities. The scope is to improve building energy performance and infrastructure energy performance, as well as optimise building energy efficiency considering thermal comfort perspectives, among others.

Table 4. Overview of energy performance and energy management services.

Reference	Service	Features	Algorithms
Xuemei et al. [102]	HVAC system operation improvement	Date; daily average/lowest/highest temperature	SVM (RBF) PCA-SVM KPCA-SVM
Xuemei L. et al. [103]	HVAC system optimisation	Dry-bulb temperature; relative humidity; solar radiation	LS-SVM (RBF) ANN (BPNN)
Solomon et al. [104]	HVAC system efficiency improvement	Temperature; wind speed; humidity; pressure; dew point temperature; wind direction; precipitation	SVM (RBF)
Dagnely et al. [105]	Green electricity production management	Occupancy; recency; temperature; irradiance; time	OLS SVM (RBF)
Massana et al. [106]	Daily Power system operation and control	Temperature; relative humidity; solar radiation; indoor temperature; indoor light level; occupancy; date	MLR ANN (MLP) SVM (PUK)
Zhao et al. [107]	Energy conservation	Holiday day; weather; zone mean air temperature; infiltration volume; heat gain through window/lights; zone internal total heat gain	SVM (RBF)
Liu et al. [108]	Abnormal energy usage identification	Occupancy; solar radiation	SVM (RBF)
Mena et al. [109]	Energy demand management	Date; outdoor temperature/humidity; solar radiation; outdoor wind speed/wind direction; state of pumps/boilers/absorption machine/cooling tower	ANN (NAR)
Yang et al. [110]	Building daily operation	On/off status of compressors; temperature of water entering ice tank/evaporator; outdoor relative humidity/temperature; chilled water; date; electric current in chiller; percentage of chilled water	Building daily operation
Lam et al. [111]	Heating load management	Outside temperature, solar radiation, workday, occupancy profiles, operational power level characteristics, transitional characteristics	ANN (MLP)
Farzana et al. [112]	Energy supply side management	Locale; population; people per household; electrification rate; type of devices/lighting bulbs; lighting energy fuel; fuel type; space heating and cooling	ANN (BPNN)
Jovanović et al. [113]	Above normal energy consumption detection	Heating consumption of previous day; mean daily outside temperature; date	ANN (FFNN) ANN (RBFN) ANN (ANFIS)
Kwok et al. [114]	Energy auditing	Relative humidity; outdoor temperature; bright sunshine duration; solar radiation; occupancy area; rainfall wind speed; occupancy rate	ANN (MLP)
Nunzio et al. [115]	Chiller soft detection	Outdoor air temperature; supply/return chilled water temperature; supply condenser water temperature; cooling tower fan VFD signal; supply cooling tower water temperature	PCA
Marinakis et al. [116]	Photovoltaic (PV) production	Temperature; humidity; pressure; wind direction degrees; solar radiation; dew point; wind speed; date	MLR
Marinakis et al. [114]	Energy consumption	Outdoor temperature; indoor temperature; degree days (heating or cooling); humidity; pressure; wind direction degrees; solar radiation; dew point; wind speed; date; working/nonworking day; envelope characteristics; occupancy profile	MLR
Singh and Yassine [117]	Behavioural analytics and energy consumption forecasting	Smart meter datasets, weather data, appliances	SVM (MLP)
Marinakis et al. [114]	Indoor air temperature	Outdoor temperature; indoor temperature; month; day; hour; working/nonworking day; on/off of the heating system scheduling	MLR
Marinakis et al. [118]	Thermal comfort validator	Temperature; air velocity lighting; clothing; activity	Decision Support System

4.2. Data Analytics for “Building and Related Infrastructure Design and Refurbishment”

Data analysis allows the design, refurbishment and development of buildings and their related infrastructure, phenomena that are traditionally considered subjective, complex, or unpredictable. Thus, standardised catalogues and Energy Conservation Measure (ECM)-based scenarios evaluation have been used, offering appropriate retrofitting solutions to the specific typology of analysed buildings and related infrastructure according to their current state (Table 5). Evaluating the correct application of an ECM in the energy efficiency of buildings is an increasingly popular topic in construction.

Table 5. Analysis of ECM tools.

CERPLAN	EnergOpt	Winwatt	EnconCalc 3.0	ECOCITIES	PRoCasaClima
Define: heating demand; primary energy demand; CO ₂ emissions (e.g., from energy pass, performance cert., etc.); base shapes; usage types; walls, windows (layers, u-value or other)	Define: usage type; wall and window area; wall mounting (layer by layer from a library); windows by u-value, dimensions, etc.; ventilation and cooling strategy and efficiency	Define: usage type; wall and window area; wall mounting layer by layer; windows by u-value, dimensions, etc.; ventilation and cooling preferences; heating distribution system; enter positive energy impacts (PV production)	Define: enter positive energy impacts (e.g., PV production)	Define: building type (by usage); priorities of the refurbishment; wall and window area; wall mounting (layer by layer from a library); windows by u-value; heating distribution system; enter positive energy impacts (e.g., PV production)	Define: walls, windows (layers, u-value); usage types
Enter: max. investment costs obtained offers	Enter: type of heating distribution; heating source; mechanical ventilation performance percentage of renewables	Enter: heating distribution heating source ventilation performance share of renewables obtained offers	Enter: energy demand (e.g., from energy performance certificate); heating demand; cooling demand; ventilation demand; heating source; total area of walls; primary energy factors CO ₂ factors	Enter: energy demand (e.g., EPC); energy demand for ventilation and cooling; enter heating source	Enter: ventilation preferences for mechanical ventilation systems; cooling preferences for the cooling system (if existing)
Calculate: cooling demand; refurbishment cost; generate most cost-effective alternatives	Calculate: heating demand CO ₂ emissions; share of renewables	Calculate: heating, cooling and ventilation demand	Calculate: most cost-effective measures (financially)	Calculate: alternatives for refurbishment; shows costs and future savings for each refurbishment measures	Calculate: primary energy demand; CO ₂ emissions; heating demand; cooling demand

Some other tools are available at the national level, such as URSA from Slovenia, EnCert-HR from Croatia, and KI Expert Plus from Croatia. All of them require building information regarding building refurbishment, refurbishment costs, and usability. On the international level, the National Renewable Energy Laboratory (NREL) of the US Department of Energy developed a tool to prioritise energy efficiency investments. The tool uses established methodologies to evaluate the energy savings and cost of those savings of these investment opportunities [119].

4.3. “Policymaking” Data Analytics

4.3.1. EPC Harmonisation and Analytics

Although the Energy Performance Certificates (EPCs) for the Member States come from the implementation of the same European directive (EPBD), EPCs datasets are very heterogeneous for different countries and regions of the EU, and their comparison is very challenging. The availability of harmonised EPC datasets across Europe would be beneficial firstly at the EU level, allowing the comparison of EPCs datasets from different regions and countries, and secondly for regional energy agencies, as a valuable support to energy efficiency policies. As expected, there is limited material referring to applications that use ML and DL algorithms for EPCs (Table 6).

Table 6. Overview of EPC AI-relevant studies.

Reference	Service	Features	Learning Algorithm
Khayatian et al. [120]	Energy performance	Degree days; net floor area; year of construction; thermal conductivity; average floor height; opaque surface area; dispersant surface; opaque-to-glazed ratio; glazed surface area; construction period; nonlinear features	ANN
Hardy et al. [121]	Error Analysis of EPCs	Energy efficiency rating; inspection date; lodgement date; property type; built form; floor description; wall description; roof description; total floor area	Random forest, linear Regression
Garcia-Nieto et al. [122]	Thermal power efficiency	Useful surface; thermal power; CO ₂ emissions; primary energy consumption; opaque enclosures; holes and skylights	Gaussian process regression
Cozza et al. [123]	Energy consumption prediction analysis from EPC	Building type; construction year; ERA; envelope factor, energy label; mechanical ventilation; heating system construction year; u-value ground; u-value roof/ceiling; u-value external walls; u-value windows; construction type	Lasso regression

4.3.2. Supporting Policymaking Impact Assessment and SECAPs Implementation

State-of-the-art methods and applications related to Sustainable Energy and Climate Action Plans (SECAP) decision making and policy evaluation were gathered for the purposes of this section. The findings are separated in two tables; the first one includes studies that analysed mainly textual information (Table 7), and the second one includes studies that took a more vertical approach in climate change relating issues that affects policymakers (Table 8).

Table 7. Overview of policymaking AI-relevant studies.

Reference	Purpose of the Study/Expected Results	Method/Technique	Learning Algorithm
Lesnikowski et al. [124]	Identifying adaption policies regarding climate issues (water treatment, sustainable growth, etc.).	Topic modelling	Latent Dirichlet analysis
Lesnikowski et al. [124]	Identifying environmental issues in Canada's local governments.	Topic modelling	Robust latent Dirichlet analysis
Rana and Miller [125]	Proving that machine learning approaches can help understanding natural resource policy and predicting socio-economic effects of them.	Socioeconomic Systems and econometrics methods	Causal tree (CT) and causal forest (CF) decision-tree algorithms
Biesbroek et al. [126]	Mapping actions regarding climate change adaptation by from policy texts and identifying high confidence blocks of adaptation.	Sorting and topic modelling	ANN
Debnath et al. [127]	Deep-narrative analysis in energy politics.	Topic modelling, grounded theory	Latent Dirichlet analysis
Hanchen et al. [128]	Identifying patterns and trends regarding hydro energy research and contributing towards strategy planning for hydro production growth.	Topic modelling	Latent Dirichlet analysis
Tavana et al. [129]	Identifying key issues in energy sector and the techniques that policymakers use for risk assessment.	Text clustering, topic modelling	K-means clustering
Boussanis and Coan [130]	Introducing methodology to identify and record key issues regarding policy issues and climate change topics.	Text analysis	Latent Dirichlet analysis
Kreif and Ordaz [131]	To provide an overview and an illustration of machine learning methods for causal inference, with a view to answer typical causal questions in policy evaluation	Overview	Several ML methods

Table 8. Overview of ML/DL applications in policymaking.

Reference	Purpose of the Study/Expected Results	Features	Learning Algorithm
Magazzino et al. [132]	Examining CO ₂ emissions, renewable energies, coal consumption, and economic growth relation.	Solar energy generation, wind energy generation, coal consumption, economic growth (GDP), and environmental pollution	D2C causality model
Mardani et al. [133]	Developing an efficient multistage methodology to predict carbon dioxide emissions.	Energy consumption, GDP, CO ₂ emissions	Hybrid of ML techniques (SOM, SVD, ANFIS, ANN)
Mason et al. [134]	To predict future energy demand, wind generation, and carbon dioxide emissions in one country (Ireland).	Historical time series of energy demand, wind power generations, CO ₂ intensity levels	CMA—ES, PSO, DE, BP, MA, RWF, LR
Nam et al. [135]	Forecasting electricity demand and renewable energy generation and renewable energy scenario suggestions to guide energy policy	Electricity demand, electricity supply, wind power generation, photovoltaic power generation	GRU, LSTM, DNN, SARIMA, MLR
Acheampong and Boateng [136]	Predicting carbon emissions for five countries and identifying significant contributory variables	Carbon emissions intensity, energy consumption, financial development index, foreign direct investment, economic growth, industrialisation, R&D (total trademark applications), population, trade openness, urbanisation	ANN
Abrell et al. [137]	To analyse climate policy effects by providing an ex post evaluation of a real-world policy experiment of carbon pricing: the UK carbon tax, also known as the Carbon Price Support.	Fossil-fuel power plant output, fuel prices, carbon prices, emissions factors and plant-specific heat efficiencies, plant capacity, demand, temperature	Causal inference

4.4. Data Analytics for “Energy Efficiency Financing”

When planning investments in the field of increasing the energy efficiency of a building, it is extremely important to obtain substantive support [138]. In the context of Triple-A [139], the Deep platform’s data are analysed for constructing some predefined performance classes about energy efficiency projects. Accordingly, in the context of SPEEDIER [140], the patterns and trends of the building’ data (e.g., user’s behaviour, energy consuming equipment on site, building fabric characteristics, etc.) are analysed to find potential areas for implementing energy saving retrofits.

In the context of CityInvest [141], the data of the energy efficiency projects that have been implemented through the project are analysed, in conjunction with the data of the local authorities in pilot regions, to recognise the key success factors of energy efficiency retrofits. This feedback is then integrated in future cases. EEnvest [142] and Industrial Energy Accelerator (IEA) [143] analyse the data of the inspected energy efficiency projects (e.g., experience and capability of the technical staff, model used for baseline estimation) to quantify their risk of failing to achieve their predicted performance. Launch [144] followed the same approach, while also analysing the macroeconomic data of the country in which the investment takes place, such as the energy prices trajectory, for calculating the risk of an energy efficiency project.

Fowlie [145] conducted an experimental evaluation of the Weatherisation Assistance Program using a sample of approximately 30,000 households in Michigan. R2A [146] cre-

ated a database that contains big data, both real records and modelled data, about buildings household's energy efficiency measures and their performance across eight countries. In the context of the Energy and Environmental Policy Analysis (EEPA) [147], a technology-based methodology for controlling the energy efficiency project's performance was implemented. In the same context, Fan and Fu [148] performed a case study on building energy efficiency retrofits, mining operational data of the building.

Two data sources were used by Garay-Martinez et al. [149] for monitoring the building's energy consumption and comfort conditions. One was the radiator energy use and the other was a weather station on the roof of the building, for capturing the climate conditions. Heo and Zavala [150] presented a Gaussian modelling framework for measurement and verification (M&V) practices. Gallagher et al. [151] employed powerful ML regression algorithms to maximise the effectiveness of available data.

4.5. Geo-Clustering Service as Support to the BD Vision

Clustering technologies take advantage of different data sources and process them in order to group the different elements in sets with similarities, which can provide very valuable information. Geo-clustering applications use also the geographical information available in order to include the geographical component in the algorithms, providing results that consider the location of the elements (Table 9).

Table 9. Overview of ML/DL applications in geo-clustering.

Reference	Service	Features	Learning Algorithm
Kuster et al. [152]	Geo-mapping methodology	Definition of 116 clusters using 16 parameters on building domain.	Not defined: use Matlab + Excel and "some criteria to select cluster" not specified
Sesana et al. [153]	Geo-cluster mapping tool	Geo-cluster concept is based on the possibility to locate similarities across enlarged EU by correlating single or multiple parameters and indicators organised in homogeneous layers and sublayers.	Correlation and cross-correlation between variables.
Exceed Project. [154]	Geo-cluster tool	Geo-clustering of building performances according to both energy and comfort aspects through the identification of specific KPIs. Classification of buildings by filtering them with building metadata. Benchmarking of building using specific indicators.	K-means algorithm
Fatiguso et al. [155]	Building geo-cluster	Collection of geographic and climatic data, simulation of solar radiation and wind exposure, mapping of typologies, materials, construction techniques, and historic architectural values of all the buildings.	ArcGIS mapping cluster toolset (not specified what)
Gangoellés et al. [156]	Building geo-cluster	Novel approach for identifying and defining a set of reference buildings by applying the <i>k</i> -means clustering method to energy performance certificate database.	K-means algorithm

5. Conclusions

Buildings are producing increasing data on energy production and consumption from various sources (e.g., smart meters, building management systems). Collecting, processing, analysing, and provisioning of reliable building data are key challenges for the built environment. There are currently a number of barriers which are actually hampering the exploitation of its full potential in improving building energy performance and management.

- The **digitalisation of the built environment** is a critical issue in the architecture, engineering, and construction (AEC) industry, which is rapidly increasing. However,

there is still a great amount of effort to be exerted. In particular, in the existing building stock, finding digital data to characterise buildings, their materials, or their energy consumption has been an unsuccessful endeavour.

- The **lack of quality and accuracy of data** is also a challenge. Robust data quality strategies and methodologies are needed for data imputation, covering data uncertainty, data quality, reliability and data consistency, and data cleansing. For an improvement or increase in data stock, not only monitoring and digitalisation are essential. In some cases, there is no possibility of integrating hardware for digitalisation. In these cases, deep learning strategies for extracting data with more granularity are necessary and enable digital strategy.
- At the same time, a **standardised data-driven architecture** for buildings is missing. In this respect, sector-wide asset schemas should be defined, stemming from extended domain-specific ontologies to allow the curation, normalisation, and homogenisation of diverse content types and artefacts, such as FIWARE, SAREF (including extensions of SAREF4Buildings), BRICK, and IFC (BIM).
- A systematic approach to **organising and managing data** is largely missing, taking into consideration that most information is not available in one place. The lack of interoperability across repositories leads to additional costs.
- One of the main challenges concerns **trusted mechanisms of data sharing and re-using**, in order to maximise the value of AI-based analytics.

A solution to pave the way for an improved BDVC can be found in the deployment of digital twins. However, the notion of compiling multidimensional digital models that can support the entire lifecycle of the asset and that such models can be mined for data and used for ML/DL requires a leap of imagination. New initiatives are necessary to address the four main pillars that will unlock the deployment of digital building twins: (1) modelling and integration of information; (2) data enrichment; (3) assuring their interoperability with different data hubs; and (4) linking them with real business cases.

The data-driven architecture should lay the foundation for facilitating seamless B2B multilateral interoperability and supporting cross-stakeholders and cross-domain-based near-real-time edge-based analytics services tailored to the stakeholders of the building value chain. In this way, new data-driven paradigms can be developed, to analyse and manage vast volumes of heterogeneous data, whose management becomes essential with a view to address increasingly complex energy systems. Operational efficiency and the reliability of buildings as an active node of bigger ecosystems (i.e., districts, power systems) can be employed to inform real-time decisions, provide early warning for abnormal conditions, and avoid potential failures. Moreover, available data can be leveraged to better understand energy demand and align this to energy generation and distribution to maximise operational efficiency in buildings or groups of buildings, thus optimising the management of assets and grids connected to the building.

Author Contributions: Conceptualisation, G.H.-M. and V.M.; methodology, G.H.-M., S.M.-P., V.I.S.-G. and C.R.-A.; formal analysis, S.M.-P., V.I.S.-G., C.R.-A., R.S.-J., N.D., Z.M. and D.A.; investigation, G.H.-M., S.M.-P., V.I.S.-G., C.R.-A., R.S.-J., N.D., Z.M. and D.A.; writing—original draft preparation, V.M.; writing—review and editing, G.H.-M. and S.M.-P.; supervision, H.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the H2020 European Commission project MATRYCS (Grant Agreement No. 1010000158).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the references of sources analysed are placed in the document already.

Acknowledgments: The work presented is based on research conducted within the framework of the H2020 European Commission project MATRYCS under contract No. 101000158. All information related to the MATRYCS project is available on the website <https://matrycs.eu/> (accessed on 28 July 2021). The authors wish to thank all the consortium partners, especially Marco Pau (RWTH), Pasquale Andriani (ENG), César Valmaseda (CARTIF), Julien Dijol, and Alice Pittini (Housing Europe), whose contribution, helpful remarks, and fruitful observations were invaluable for the development of this work. The content of the paper is the sole responsibility of the authors and does not necessarily reflect the views of the EC.

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

Abbreviations

AEC	Architecture, engineering, and construction
AEMET	National Meteorology Agency (Spain)
ALS	Airborne laser scanning
AI	Artificial intelligence
ANFIS	Adaptive neuro-fuzzy inference system
ANN	Artificial neural networks
BD	Big data
BDVA	Bid data value association
BDVC	Big data value chain
BEM	Building energy model
BIM	Building information modelling
BP	Backpropagation
BPNN	Back propagation neural network
B2B	Business to business
CF	Causal forest
CityGML	City Geography Markup Language
CMA-ES	Covariance Matrix Adaptation Evolution Strategy
CNIG	National Centre for Geographic Information (Spain)
CO ₂	Carbon dioxide
CRU	Climatic Research Unit
CT	Causal tree
DAIRO	Data, AI, and robotics
DE	Differential evolution
DEEP	De-risking energy efficiency platform
DL	Deep learning
DLT	Distributed ledger technology
DBT	Digital building twins
DNN	Deep neural network
D2C	Dependency to causality
EC	European Commission
ECM	Energy Conservation Measure
ECTP	European Construction Technology Platform
EEA	European Environment Agency
EEPA	Energy and Environmental Policy Analysis
EPBD	Energy Performance of Buildings Directive
EPC	Energy Performance Certificates
EU	European Union
E3P	European Energy Efficiency Platform
FFNN	Feed forward neural network
GBP	Green Building Programme
GDP	Gross domestic product
GIS	Geographical information system
GPS	Global positioning system
GRU	Gated recurrent unit
HVAC	Heating, ventilation, and air conditioning
ICT	Information and communication technologies

IDEES	Integrated Database of the European Energy System
IEA	International Energy Agency/Industrial Energy Accelerator
IFC	Industry Foundation Classes
INE	National Institute of Statistics (Spain)
IoT	Internet of things
JRC	Joint Research Centre
KPCA	Kernel principal component analysis
LR	Left-to-right, rightmost derivation in reverse
LSTM	Long short-term memory
LS-SVM	Least squares support vector machines
MA	Memetic algorithm
ML	Machine learning
MLP	Multilayer perceptron
MLR	Multiple linear regression
MQTT	Message Queuing Telemetry Transport
M&V	Measurement and verification
NAR	Nonlinear autorregressive
NB-IoT	Narrowband IoT
NREL	National Renewable Energy Laboratory
nZEB	Near-zero energy building
OECD	Organisation for Economic Cooperation and Development
OLS	Ordinary least squares
OSM	Open Street Maps
PCA	Principal component analysis
PSO	Particle swarm optimisation
PUK	Pearson VII universal kernel
PV	Photovoltaics
RBF	Radial basis function
RBFN	Radial basis function network
RES	Renewable energy sources
RWF	Reshaped Wirtinger flow
R&D	Research and development
SAREF	Smart Applications Reference
SARIMA	Seasonal autoregressive integrated moving average
SECAP	Sustainable Energy and Climate Action Plan
SME	Small and medium-sized enterprise
SOM	Self-organising map
SRIA	Strategic Research and Innovation Agenda
SVD	Singular value decomposition
SVM	Support vector machine
UNSD	United Nations Statistics Division

References

1. Marinakis, V. Big Data for Energy Management and Energy-Efficient Buildings. *Energies* **2020**, *13*, 1555. [CrossRef]
2. Marinakis, V.; Doukas, H.; Tsapelas, J.; Mouzakitis, S.; Sicilia, Á.; Madrazo, L.; Sgouridis, S. From big data to smart energy services: An application for intelligent energy management. *Future Gener. Comput. Syst.* **2020**, *110*, 572–586. [CrossRef]
3. Koutsandreas, D.; Spiliotis, E.; Doukas, H.; Psarras, J. What is the macroeconomic impact of higher decarbonization speeds? The case of Greece. *Energies* **2021**, *14*, 2235. [CrossRef]
4. European Commission. A European Green Deal. Available online: https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal_en (accessed on 31 May 2021).
5. European Commission. *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: A Renovation Wave for Europe—Greening Our Buildings, Creating Jobs, Improving Lives*; COM(2020) 662 Final; European Commission: Brussels, Belgium, 2020.
6. European Commission. *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: A European Strategy for Data*; COM(2020) 66 Final; European Commission: Brussels, Belgium, 2020.
7. European Commission. *Proposal for a Regulation of the European Parliament and of the Council on European Data Governance (Data Governance Act)*; COM(2020) 767 Final; European Commission: Brussels, Belgium, 2020.

8. European Commission. *White Paper on Artificial Intelligence—A European Approach to Excellence and Trust*; COM(2020) 65 Final; European Commission: Brussels, Belgium, 2020.
9. European Commission. *Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee: Report on the Safety and Liability Implications of Artificial Intelligence, the Internet of Things and Robotics*; COM(2020) 64 Final; European Commission: Brussels, Belgium, 2020.
10. European Commission. *Ethics Guidelines for Trustworthy AI*. Available online: <https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1> (accessed on 31 May 2021).
11. European Commission. *Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on Open Data and the Re-Use of Public Sector Information (Recast)*; European Commission: Brussels, Belgium, 2019.
12. European Commission. *Regulation (EU) 2016/679 of the European Parliament and of the Council on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation)*; European Commission: Brussels, Belgium, 2016.
13. European Commission. *Communication to the Commission Open Source Software Strategy 2020–2023, Think Open*; C(2020) 7149 Final; European Commission: Brussels, Belgium, 2020.
14. ECTP. *ECTP Strategic Research and Innovation Agenda 2021–2027*. 2019. Available online: http://www.ectp.org/fileadmin/user_upload/documents/ECTP/Miscellaneous_doc/ECTP_SRIA_FINAL_20-11-2019.pdf (accessed on 31 May 2021).
15. BDVA. *BDVA Strategic Research and Innovation Agenda*. Available online: https://www.bdva.eu/sites/default/files/BDVA_SRIA_v4_Ed1.1.pdf (accessed on 31 May 2021).
16. BDVA. *Towards a European-Governed Data Sharing Space. BDVA Position Paper*. Available online: https://www.bdva.eu/sites/default/files/BDVA%20DataSharingSpaces%20PositionPaper%20V2_2020_Final.pdf (accessed on 31 May 2021).
17. BUILD UP. *The European Portal for Energy Efficiency in Buildings*. Available online: <https://www.buildup.eu/> (accessed on 31 May 2021).
18. European Energy Efficiency Platform (E3P). Available online: <https://e3p.jrc.ec.europa.eu/> (accessed on 31 May 2021).
19. Coalition for Energy Savings. Available online: <http://energycoalition.eu> (accessed on 31 May 2021).
20. Housing Europe. *Housing Evolutions Hub*. Available online: <https://www.housingevolutions.eu/> (accessed on 31 May 2021).
21. *A European AI on Demand Platform and Ecosystem (AI4EU)*. Available online: <https://www.ai4eu.eu/> (accessed on 31 May 2021).
22. *FIWARE Smart Data Models*. Available online: <https://smartdatamodels.org/> (accessed on 31 May 2021).
23. *Industry Foundation Classes (IFC)*. Available online: <https://www.iso.org/standard/70303.html#:~:text=The%20Industry%20Foundation%20Classes%2C%20IFC,or%20facility%20management%20industry%20sector> (accessed on 31 May 2021).
24. *City Geography Markup Language (CityGML)*. Available online: <http://www.opengis.net/spec/citygml/2.0> (accessed on 31 May 2021).
25. *INSPIRE Directive (2007/2/EC)*. Available online: <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32007L0002/> (accessed on 31 May 2021).
26. *Smart Applications REFERENCE (SAREF)*. Available online: [https://saref.etsi.org/#:~:text=The%20Smart%20Applications%20REFERENCE%20\(SAREF,ontology%20depending%20on%20specific%20needs](https://saref.etsi.org/#:~:text=The%20Smart%20Applications%20REFERENCE%20(SAREF,ontology%20depending%20on%20specific%20needs) (accessed on 31 May 2021).
27. *BRICK*. Available online: <https://brickschema.org/> (accessed on 31 May 2021).
28. Sacks, R.; Eastman, C.; Lee, G.; Teicholz, P. *BIM Handbook: A Guide to Building Information Modeling for Owners, Designers, Engineers, Contractors, and Facility Managers*, 3rd ed.; Wiley: Hoboken, NJ, USA, 2018; p. 688.
29. Succar, B. *Building information modelling framework: A research and delivery foundation for industry stakeholders*. *Autom. Constr.* **2009**, *18*, 357–375. [[CrossRef](#)]
30. Poljansek, M. *Building Information Modelling (BIM) standardization*; EUR 28977 EN; Publications Office of the European Union: Luxembourg, 2017.
31. *Helsinki's 3D City Models*. Available online: <https://www.hel.fi/helsinki/en/administration/information/general/3d/3d/> (accessed on 31 May 2021).
32. *Virtual Helsinki*. Available online: <https://www.virtualhelsinki.fi/> (accessed on 31 May 2021).
33. *The Kalasatama Digital Twins Project*. Available online: https://www.hel.fi/static/liitteet-2019/Kaupunginkanslia/Helsinki3D_Kalasatama_Digital_Twins.pdf (accessed on 31 May 2021).
34. *EU Building Stock Observatory*. Available online: https://ec.europa.eu/energy/eu-buildings-database_en (accessed on 31 May 2021).
35. *EU Energy Poverty Observatory*. Available online: <https://www.energypoverty.eu/indicators-data> (accessed on 31 May 2021).
36. *EUROSTAT*. Available online: <https://ec.europa.eu/eurostat/web/energy/data> (accessed on 31 May 2021).
37. *Statistical Review of World Energy*. Available online: <https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html> (accessed on 31 May 2021).
38. *TABULA EPISCOPE*. Available online: <http://webtool.building-typology.eu/> (accessed on 31 May 2021).
39. Available online: <https://entranze.enerdata.net/> (accessed on 31 May 2021).
40. *ODYSSEE—Full Database*. Available online: <https://www.odyssee-mure.eu/> (accessed on 31 May 2021).
41. *ODYSSEE—Key Indicator Tool*. Available online: <https://www.indicators.odyssee-mure.eu/online-indicators.html> (accessed on 31 May 2021).

42. ODYSSEE—Decom-Position Tool. Available online: <https://www.indicators.odyssee-mure.eu/decomposition.html> (accessed on 31 May 2021).
43. ODYSSEE—Market Diffusion Tool. Available online: <https://www.indicators.odyssee-mure.eu/market-diffusion.html> (accessed on 31 May 2021).
44. ODYSSEE—Comparison Tool. Available online: <https://www.indicators.odyssee-mure.eu/benchmarking.html> (accessed on 31 May 2021).
45. ODYSSEE—Energy Saving Tool. Available online: <https://www.indicators.odyssee-mure.eu/energy-saving.html> (accessed on 31 May 2021).
46. ODYSSEE—EU Energy Efficiency Scoreboard. Available online: <https://www.odyssee-mure.eu/data-tools/scoring-efficiency-countries.html> (accessed on 31 May 2021).
47. MURE Database. Available online: <https://www.measures.odyssee-mure.eu/energy-efficiency-policies-database.html#/> (accessed on 31 May 2021).
48. EnergyPlus Weather Data. Available online: <https://www.energyplus.net/weather> (accessed on 31 May 2021).
49. Climatic Research Unit (CRU). Available online: <https://sites.uea.ac.uk/cru/data> (accessed on 31 May 2021).
50. Available online: <https://www.eea.europa.eu/data-and-maps> (accessed on 31 May 2021).
51. Our World in Data. Available online: <https://ourworldindata.org/> (accessed on 31 May 2021).
52. Available online: <https://data.europa.eu/euodp/en/home> (accessed on 31 May 2021).
53. European Data Portal. Available online: <https://www.europeandataportal.eu/en> (accessed on 31 May 2021).
54. World Bank. Available online: <https://data.worldbank.org/> (accessed on 31 May 2021).
55. Organization for Economic Co-Operation and Development (OECD). Available online: <https://data.oecd.org/> (accessed on 31 May 2021).
56. Covenant of Mayors. Available online: <https://www.covenantofmayors.eu/plans-and-actions/action-plans.html> (accessed on 31 May 2021).
57. The Shift Data Portal. Available online: <https://www.theshiftdataportal.org/energy> (accessed on 31 May 2021).
58. United Nations Statistics Division Website (UNSD). Available online: <https://unstats.un.org/unsd/energystats/> (accessed on 31 May 2021).
59. Ember. Available online: <https://ember-climate.org/data/> (accessed on 31 May 2021).
60. Climate Fund Inventory Database. Available online: <https://qdd.oecd.org/subject.aspx?subject=climatefundinventory> (accessed on 31 May 2021).
61. The Carbon Center. Available online: <https://carbonn.org/> (accessed on 31 May 2021).
62. Open Street Map (OSM). Available online: www.openstreetmap.org (accessed on 31 May 2021).
63. Copernicus Data. Available online: <https://www.copernicus.eu> (accessed on 31 May 2021).
64. HotMaps. Available online: <https://gitlab.com/hotmaps/building-stock/-/tree/master/data> (accessed on 31 May 2021).
65. ZEBRA. Available online: <https://zebra-monitoring.enerdata.net/> (accessed on 31 May 2021).
66. CommONEnergy. Available online: <https://eeg.tuwien.ac.at/commonenergy/> (accessed on 31 May 2021).
67. Mantzos, L.; Matei, N.A.; Mulholland, E.; Rózsai, M.; Tamba, M.; Wiesenthal, T. *Integrated Database of the European Energy System*; European Commission, Joint Research Centre (JRC): Luxembourg, 2018.
68. ExCEED. Available online: <http://www.exceedproject.eu/> (accessed on 31 May 2021).
69. iNSPiRe. Available online: https://zenodo.org/record/3256270#.X_wayHqSnIU (accessed on 31 May 2021).
70. ZENSUS 2011. Available online: <https://ergebnisse.zensus2011.de/#> (accessed on 31 May 2021).
71. Itard, L.; Meijer, F. *Towards a Sustainable Northern European Housing Stock: Figures, Facts and Future*; Delft Centre for Sustainable Urban Areas: Delft, The Netherlands, 2008.
72. De-risking Energy Efficiency Platform (DEEP). Available online: <https://deep.eefig.eu/overview> (accessed on 31 May 2021).
73. D'Agostino, D.; Cuniberti, B.; Bertoldi, P. Energy consumption and efficiency technology measures in European non-residential buildings. *Energy Build.* **2017**, *153*, 72–86. [[CrossRef](#)]
74. National Housing Census: European Statistical System. Available online: <https://ec.europa.eu/CensusHub2/query.do?step=selectHyperCube&qhc=false> (accessed on 31 May 2021).
75. EDGAR (Emissions Database for Global Atmospheric Research) CO2 Emissions. Available online: <https://github.com/openclimatedata/edgar-co2-emissions> (accessed on 31 May 2021).
76. CORDEX—Regional Climate Model Data on Single Levels for Europe. Available online: <https://cds.climate.copernicus.eu/cdsapp#!/dataset/projections-cordex-single-levels?tab=overview> (accessed on 31 May 2021).
77. Open Data Hub Italia. Available online: <https://sciamlab.com/opendatahub/dataset> (accessed on 31 May 2021).
78. EPC Cadastre of Lombardia Region. Available online: http://www.cened.it/statistiche_cened (accessed on 31 May 2021).
79. GreenDataset. Available online: <https://sourceforge.net/projects/greend/> (accessed on 31 May 2021).
80. Portal Energetica. Available online: <https://www.energetika-portal.si/> (accessed on 31 May 2021).
81. OPSI. Available online: <https://podatki.gov.si/> (accessed on 31 May 2021).
82. Geoport. Available online: <https://www.geoport.gov.pl/o-geoportalu/aktualnosci> (accessed on 31 May 2021).
83. EPC Register. Available online: <https://rejestrcheb.mrpit.gov.pl/rejestr-budynkow> (accessed on 31 May 2021).

84. Spanish Cadastre Data. Available online: <https://www.sedecatastro.gob.es/Accesos/SECACCDescargaDatos.aspx> (accessed on 31 May 2021).
85. AEMET OpenData. Available online: <https://opendata.aemet.es/> (accessed on 31 May 2021).
86. INE Open Data. Available online: <https://www.ine.es/uc/qLXuzTM8> (accessed on 31 May 2021).
87. CNIG Download Center. Available online: <http://centrodedescargas.cnig.es/CentroDescargas/index.jsp> (accessed on 31 May 2021).
88. BIM Document Library. Available online: <https://www.buildingsmart.es/observatorio-bim/biblioteca-bim/> (accessed on 31 May 2021).
89. Open Data Portal. Available online: [Dados.gov.pt](https://dados.gov.pt) (accessed on 31 May 2021).
90. Database of Buildings and Infrastructure. Available online: <http://ioer-bdat.de/en/> (accessed on 31 May 2021).
91. Greece National Data Portal. Available online: <https://data.gov.gr/> (accessed on 31 May 2021).
92. Open Data Greece. Available online: <http://geodata.gov.gr/> (accessed on 31 May 2021).
93. Latvia's Open Data portal. Available online: <https://data.gov.lv/eng> (accessed on 31 May 2021).
94. Belgium Government Open Data. Available online: <https://data.gov.be/en> (accessed on 31 May 2021).
95. ENEX. Available online: <https://www.mpo-enex.cz> (accessed on 31 May 2021).
96. Energo. 2015. Available online: <https://www.czso.cz/csu/czso/energo-2015> (accessed on 31 May 2021).
97. Live Tables on Energy Performance of Buildings Certificates. Available online: <https://www.gov.uk/government/statistical-data-sets/live-tables-on-energy-performance-of-buildings-certificates> (accessed on 31 May 2021).
98. Energy DataHub. Available online: <https:// analisis.datosabiertos.jcyl.es/pages/eren/?flg=es> (accessed on 31 May 2021).
99. EPC Register. Available online: <https://datosabiertos.jcyl.es/web/jcyl/set/es/energia/certificados-eficiencia/1284543386412> (accessed on 31 May 2021).
100. Wang, Y.; Chen, Q.; Kang, C.; Xia, Q. Clustering of electricity consumption behavior dynamics toward big data applications. *IEEE Trans. Smart Grid* **2016**, *7*, 2437–2447. [[CrossRef](#)]
101. Sheng, G.; Hou, H.; Jiang, X.; Chen, Y. A novel association rule mining method of big data for power transformers state parameters based on probabilistic graph model. *IEEE Trans. Smart Grid* **2018**, *9*, 695–702. [[CrossRef](#)]
102. Xuemei, L.; Lixing, D.; Jinhu, L.; Gang, X.; Jibin, L. A Novel Hybrid Approach of KPCA and SVM for Building Cooling Load Prediction. In Proceedings of the 2010 International Conference on Knowledge Discovery and Data Mining, Phuket, Thailand, 9–10 January 2010; p. 1.
103. Li, X.; Lu, J.-H.; Ding, L.; Xu, G.; Li, J. Building cooling load forecasting model based on LS-SVM. In Proceedings of the 2009 Asia-Pacific Conference on Information Processing, Shenzhen, China, 18–19 July 2009; Volume 1, pp. 55–58.
104. Solomon, D.M.; Winter, R.L.; Boulanger, A.G.; Anderson, R.N.; Wu, L.L. *Forecasting Energy Demand in Large Commercial Buildings Using Support Vector Machine Regression*; Tech. Rep. CUCS-040-11; Department of Computer Science, Columbia University: New York, NY, USA, September 2011.
105. Dagnely, P.; Ruetter, T.; Tourwé, T.; Tsiorkova, E.; Verhelst, C. Predicting hourly energy consumption. Can you beat an autoregressive model. In Proceedings of the 24th Annual International Conference on Machine Learning, Delft, The Netherlands, 9 June 2015; Volume 19.
106. Massana, J.; Pous, C.; Burgas, L.; Melendez, J.; Colomer, J. Short-term load forecasting in a non-residential building contrasting models and attributes. *Energy Build.* **2015**, *92*, 322–330. [[CrossRef](#)]
107. Zhao, H.; Magoulès, F. Parallel support vector machines applied to the prediction of multiple buildings energy consumption. *J. Algorithms Comput. Technol.* **2010**, *4*, 231–249. [[CrossRef](#)]
108. Liu, D.; Chen, Q. Prediction of building lighting energy consumption based on support vector regression. In Proceedings of the 2013 9th Asian Control Conference (ASCC), Istanbul, Turkey, 23–26 June 2013; pp. 1–5.
109. Mena, R.; Rodríguez, F.; Castilla, M.; Arahal, M.R. A prediction model based on neural networks for the energy consumption of a bioclimatic building. *Energy Build.* **2014**, *82*, 142–145. [[CrossRef](#)]
110. Yang, J.; Rivard, H.; Zmeureanu, R. On-line building energy prediction using adaptive artificial neural networks. *Energy Build.* **2005**, *37*, 1250–1259. [[CrossRef](#)]
111. Lam, J.C.; Wan, K.K.W.; Wong, S.L.; Lam, T.N.T. Principal component analysis and long-term building energy simulation correlation. *Energy Convers. Manag.* **2010**, *51*, 135–139. [[CrossRef](#)]
112. Farzana, S.; Liu, M.; Baldwin, A.; Hossain, M.U. Multi-model prediction and simulation of residential building energy in urban areas of Chongqing, South West China. *Energy Build.* **2014**, *81*, 161–169. [[CrossRef](#)]
113. Jovanović, R.Ž.; Sretenović, A.A.; Živković, B.D. Ensemble of various neural networks for prediction of heating energy consumption. *Energy Build.* **2015**, *94*, 189–199. [[CrossRef](#)]
114. Kwok, S.S.K.; Yuen, R.K.K.; Lee, E.W.M. An intelligent approach to assessing the effect of building occupancy on building cooling load prediction. *Build. Environ.* **2011**, *46*, 1681–1690. [[CrossRef](#)]
115. Cotrufo, N.; Zmeureanu, R. PCA-based method of soft fault detection and identification for the ongoing commissioning of chillers. *Energy Build.* **2016**, *130*, 443–452. [[CrossRef](#)]
116. Marinakis, V.; Doukas, H. An advanced IoT-based system for intelligent energy management in buildings. *Sensors* **2018**, *18*, 610. [[CrossRef](#)]

117. Singh, S.; Yassine, A. big data mining of energy time series for behavioral analytics and energy consumption forecasting. *Energies* **2018**, *11*, 452. [CrossRef]
118. Marinakis, V.; Doukas, H.; Spiliotis, E.; Papastamatiou, I. Decision support for intelligent energy management in buildings using the thermal comfort model. *Int. J. Comput. Intell. Syst.* **2017**, *10*, 882–893. [CrossRef]
119. Farese, P.; Gelman, R.; Hendron, R. *A Tool to Prioritize Energy Efficiency Investments*; National Renewable Energy Laboratory (NREL): Golden, CO, USA, 2015.
120. Khayatian, F.; Sarto, L.; Dall’O’, G. Application of neural networks for evaluating energy performance certificates of residential buildings. *Energy Build.* **2016**, *125*, 45–54. [CrossRef]
121. Hardy, A.; Glew, D. An analysis of errors in the energy performance certificate database. *Energy Policy* **2019**, *129*, 1168–1178. [CrossRef]
122. García-Nieto, P.J.; García-Gonzalo, E.; Paredes-Sánchez, J.P.; Sánchez, A.B. A new hybrid model to foretell thermal power efficiency from energy performance certificates at residential dwellings applying a Gaussian process regression. *Neural Comput. Appl.* **2021**, *33*, 6627–6640. [CrossRef]
123. Cozza, S.; Chambers, J.; Deb, C.; Scartezzini, J.-L.; Schlüter, A.; Patel, M.K. Do energy performance certificates allow reliable predictions of actual energy consumption and savings? Learning from the Swiss national database. *Energy Build.* **2020**, *224*, 110235. [CrossRef]
124. Lesnikowski, A.; Belfer, E.; Rodman, E.; Smith, J.; Biesbroek, R.; Wilkerson, J.D.; Ford, J.D.; Berrang-Ford, L. Frontiers in data analytics for adaptation research: Topic modeling. *WIREs Clim Chang.* **2019**. [CrossRef]
125. Rana, P.; Miller, D.C. Machine learning to analyze the social-ecological impacts of natural resource policy: Insights from community forest management in the Indian Himalaya. *Environ. Res. Lett.* **2019**, *14*, 024008. [CrossRef]
126. Biesbroek, R.; Badloe, S.; Athanasiadis, I.N. Machine learning for research on climate change adaptation policy integration: An exploratory UK case study. *Reg. Environ. Chang.* **2020**, *20*, 85. [CrossRef]
127. Debnath, R.; Darby, S.; Bardhan, R.; Mohaddes, K.; Sunikka-Blank, M. Grounded reality meets machine learning: A deep-narrative analysis framework for energy policy research. *Energy Res. Soc. Sci.* **2020**, *69*, 101704. [CrossRef] [PubMed]
128. Jiang, H.; Qiang, M.; Lin, P. A topic modeling based bibliometric exploration of hydropower research. *Renew. Sustain. Energy Rev.* **2016**, *57*, 226–237. [CrossRef]
129. Tavana, M.; Shaabani, A.; Santos-Arteaga, F.J.; Vanani, I.R. A review of uncertain decision-making methods in energy management using text mining and data analytics. *Energies* **2020**, *13*, 3947. [CrossRef]
130. Boussalis, C.; Coan, T.G. Text-mining the signals of climate change doubt. *Glob. Environ. Chang.* **2016**, *36*, 89–100. [CrossRef]
131. Kreif, N.; DiazOrdaz, K. Machine Learning in Policy Evaluation: New Tools for Causal Inference. Available online: <https://arxiv.org/abs/1903.00402> (accessed on 31 May 2021).
132. Magazzino, C.; Mele, M.; Schneider, N. A machine learning approach on the relationship among solar and wind energy production, coal consumption, GDP, and CO₂ emissions. *Renew. Energy* **2021**, *167*, 99–115. [CrossRef]
133. Mardani, A.; Liao, H.; Nilashi, M.; Alrasheedi, M.; Cavallaro, F. A multi-stage method to predict carbon dioxide emissions using dimensionality reduction, clustering, and machine learning techniques. *J. Clean. Prod.* **2020**, *275*, 122942. [CrossRef]
134. Mason, K.; Duggan, J.; Howley, E. Forecasting energy demand, wind generation and carbon dioxide emissions in Ireland using evolutionary neural networks. *Energy* **2018**, *155*, 705–720. [CrossRef]
135. Nam, K.; Hwangbo, S.; Yoo, C. A deep learning-based forecasting model for renewable energy scenarios to guide sustainable energy policy: A case study of Korea. *Renew. Sustain. Energy Rev.* **2020**, *122*, 109725. [CrossRef]
136. Acheampong, A.O.; Boateng, E.B. Modelling carbon emission intensity: Application of artificial neural network. *J. Clean. Prod.* **2019**, *225*, 833–856. [CrossRef]
137. Abrell, J.; Kosch, M.; Rausch, S. *How Effective Was the UK Carbon Tax?—A Machine Learning Approach to Policy Evaluation*; CER-ETH—Center of Economic Research at ETH Zurich Working Paper 19/317; CER-ETH: Zürich, Switzerland, 2019.
138. Doukas, H.; Xidonas, P.; Mastromichalakis, N. How Successful are Energy Efficiency Investments? A Comparative Analysis for Classification & Performance Prediction. *Comput. Econ.* **2021**. [CrossRef]
139. Triple-A. Available online: <https://aaa-h2020.eu/> (accessed on 31 May 2021).
140. SPEEDIER. Available online: <https://speedierproject.eu/> (accessed on 31 May 2021).
141. CITYnvest. Available online: <http://citynvest.eu/> (accessed on 31 May 2021).
142. EEnvest. Available online: <http://www.eenest.eu/> (accessed on 31 May 2021).
143. Industrial Energy Accelerator. Available online: <https://www.industrialenergyaccelerator.org/> (accessed on 31 May 2021).
144. LAUNCH. Available online: <https://www.launch2020.eu/> (accessed on 31 May 2021).
145. Fowlie, M. Do energy efficiency investments deliver? Evidence from the weatherization assistance program. *The Quarterly Journal of Economics. Q. J. Econ.* **2018**, *133*, 1597–1644. [CrossRef]
146. Costanzo, E.; Weatherall, D.; Androutsopoulos, A.; Gomez Onate, V. Can big data drive the market for residential energy efficiency? In *ECEEE 2017 Summer Study Proceedings, Panel: 8. Monitoring and Evaluation: Building Confidence and Enhancing Practices*; ECEEE: Stockholm, Sweden, 2017.
147. Byrne, J.; Taminau, J.; Carretero, D.; Shin, S.; Xu, J. *Risk Mitigation in Energy Efficiency Retrofit Projects Using Automated Monitoring and Verification Techniques*; Technical report prepared for the Delaware General Assembly; CEEP, University of Delaware: Newark, DE, USA, 2019.

148. Fan, C.; Fu, X. Mining big building operational data for improving building energy efficiency: A case study. *Build. Serv. Eng. Res. Technol.* **2017**, *117*–128. [[CrossRef](#)]
149. Garay-Martinez, R.; Arregi, B.; Lumbreras, M.; Zurro, B.; Gonzalez, J.M.; Hernandez, J.L. Data driven process for the energy assessment of building envelope retrofits. *E3S Web Conf.* **2020**, *172*, 25001. [[CrossRef](#)]
150. Heo, Y.; Zavala, V.M. Gaussian process modeling for measurement and verification of building energy savings. *Energy Build.* **2012**, *53*, 7–18. [[CrossRef](#)]
151. Gallagher, C.V.; Leahy, K.; O'Donovan, P.; Bruton, K.; O'Sullivan, D.T.J. Development and application of a machine learning supported methodology for measurement and verification (M&V) 2.0. *Energy Build.* **2018**, *167*, 8–22.
152. Kuster, C.; Hippolyte, J.L.; Rezqui, Y.; Mourshed, M. A simplified geo-cluster definition for energy system planning in the Europe. *Energy Procedia* **2019**, *158*, 3222–3227. [[CrossRef](#)]
153. Sesana, M.M.; Cuca, B.; Iannaccone, G.; Brumana, R.; Caccavelli, D.; Gay, C. Geomapping methodology 191 for the GeoCluster Mapping Tool to assess deployment potential of technologies for energy efficiency in 192 buildings. *Sustain. Cities Soc.* **2013**, *17*, 22–34. [[CrossRef](#)]
154. Geo-Clustering to Deploy the Potential of Energy Efficient Buildings across EU. Available online: <https://cordis.europa.eu/project/id/285501/reporting> (accessed on 31 May 2021).
155. Fatiguso, F.; De Fino, M.; Cantatore, E.; Sciotti, A.; De Tommasi, G. Energy models towards retrofitting of the historic built heritage. *WIT Trans. Built Environ.* **2015**, *153*, 159–170.
156. Gangolells, M.; Casals, M.; Ferré-Bigorra, J.; Forcada, N.; Macarulla, M.; Gaspar, K.; Tejedor, B. Office representatives for cost-optimal energy retrofitting analysis: A novel approach using cluster analysis of energy performance certificate databases. *Energy Build.* **2020**, *206*, 109557. [[CrossRef](#)]