# AI and Data Democratisation for Intelligent Energy Management

**Vangelis Marinakis \*** , **Themistoklis Koutsellis** , **Alexandros Nikas** and **Haris Doukas**

School of Electrical & Computer Engineering, National Technical University of Athens, 15780 Zografou (Athens), Greece; tkoutsellis@epu.ntua.gr (T.K.); anikas@epu.ntua.gr (A.N.); h_doukas@epu.ntua.gr (H.D.)
**\*** Correspondence: vmarinakis@epu.ntua.gr; Tel.: +30-210-772-3612

**Abstract:** Despite the large number of technology-intensive organisations, their corporate know-how and underlying workforce skill are not mature enough for a successful rollout of Artificial Intelligence (AI) services in the near-term. However, things have started to change, owing to the increased adoption of data democratisation processes, and the capability offered by emerging technologies for data sharing while respecting privacy, protection, and security, as well as appropriate learning-based modelling capabilities for non-expert end-users. This is particularly evident in the energy sector. In this context, the aim of this paper is to analyse AI and data democratisation, in order to explore the strengths and challenges in terms of data access problems and data sharing, algorithmic bias, AI transparency, privacy and other regulatory constraints for AI-based decisions, as well as novel applications in different domains, giving particular emphasis on the energy sector. A data democratisation framework for intelligent energy management is presented. In doing so, it highlights the need for the democratisation of data and analytics in the energy sector, toward making data available for the right people at the right time, allowing them to make the right decisions, and eventually facilitating the adoption of decentralised, decarbonised, and democratised energy business models.

**Keywords:** artificial intelligence; data democratisation; energy data spaces; interoperability; data sharing; energy management; decarbonisation; decision support

## 1. Introduction

Currently, there is increasing debate over the importance of data democracy [1,2]; press pieces, academic articles, and debates over mainstream media expand or redefine the term and continuously reshape its meaning and framework. The growth in data volumes (i.e., big data [3]), the expansion in data types/formats, and the plethora of data analysis and process methodologies (e.g., data mining [4,5] and machine learning [6–8] procedures), combined with the constantly increasing momentum of Artificial Intelligence (AI) [9,10], indicate that the 21st century marks the era of data science [11–14]: everybody imprints data somewhere (on the Web, Internet of Things (IoT), etc.) almost about everything.

AI-based analytics can identify hidden, potentially useful information patterns within large datasets that can be transformed into actionable outcomes and knowledge to support improved decision making [15]. However, in most cases, few possess these data.

This is particularly evident in the energy sector. The energy system is going through an unprecedented transformation that can be summarised in two main factors: supply side changes mainly orienting on the installation of renewable energy sources [16] and demand-side solutions orbiting around behavioural change and further highlighting a new, more active role for consumers [17]. Such a transformation is significantly affecting the business structure and determining the creation of completely new concepts of utility. In such a context, one of the main factors is that we are straying away from the concept of delivering a commodity and instead entering a new economy of services.

### 1.1. Barriers to the Diffusion of AI and Data Democratisation in Energy Sector

Energy utilities tend to believe that AI will have a big impact on their business: AI spreading is perceived to eventually allow them to align their business processes to novel optimal energy management operations [18], within a context of rising share of deployed renewable energy [19]. However, they are far from integrating AI in key business processes [20]. This is mostly due to energy utilities—despite having high technical skill in power systems—very often lacking on their workforce Information and Communication Technologies (ICT) background to fully exploit the AI potential in their business, and hence progress toward integrating AI applications or services within their processes [21].

In addition to poor-ICT-skilled stakeholders, a high number of leading-edge Small and Mid-size Enterprises (SMEs) are emerging to provide a vast number of smart energy management services and technologies to energy stakeholders. However, the fuzzy and often unclear regulation on access to data has been instead hindering the potential of sustainable business cases and return on investment for SMEs.

This includes the General Data Protection Regulation (GDPR), which is in force in the European Union (EU) and the European Economic Area (EEA). The GDPR is liable for keeping energy consumer data, as well as fragmented regulation on data access and cross-country harmonised or standardised requirements for services provisioning to energy stakeholders. The GDPR also addresses the transfer of personal data outside the EU and EEA areas, while one of its main aims is to simplify the regulatory environment for international business by applying a unified regulatory framework within the EEA. Despite the plethora of such regulations worldwide, there are steps forward to align different legal frameworks among countries in the context of data sharing. It is worth mentioning that the GDPR became a role model for many national laws outside the EEA (e.g., in Chile, Japan, Brazil, South Korea, Argentina and Kenya). Moreover, the California Consumer Privacy Act (CCPA) and the Personal Information Protection and Electronic Documents Act (PIPEDA) of Canada also share many features with the GDPR.

Additionally, the presence of consolidated functional/organisational silos, combined with the lack of semantic and business interoperability across data stream providers, has too been hindering the exploitation of the high potential of AI-based analytics in the energy sector. While most value chains cross the boundaries of individual sector, existing big data and IoT applications in Business to Business (B2B) energy settings tend to be "siloed": they do not provide easy ways for sharing, reusing, and repurposing data assets, cognitive models, and AI services across domains/stakeholders. It becomes evident that combined data from different domains (energy, climate, economics, and financing; user behaviour and societal trends; regulations; and many more) can open a wide spectrum of unprecedented AI energy analytics services [22], if effectively tackled with appropriate technological solutions.

### 1.2. Overarching Objective of This Research

The aim of this paper is to analyse AI and data democratisation, in order to explore the strengths and challenges in terms of data access problems and data sharing, algorithmic bias, AI transparency, privacy, and other regulatory constraints for AI decisions, as well as to review novel applications in different domains, giving particular emphasis on the energy sector. In this context, a data democratisation framework for intelligent energy management is presented, integrating seven components for: (1) data discovery and interoperability, (2) data quality compliance, (3) data privacy and sharing, (4) AI-based library, (5) analytics service builder and visualisation, (6) AI energy analytics services, and (7) security. In doing so, it highlights the need for the democratisation of data and analytics in the energy sector, toward making data available for the right people at the right time, allowing them to make the right decision, and eventually facilitating the adoption of decentralised, decarbonised, and democratised energy business models.

The rest of the paper is organised as follows. In Section 2, the progress toward democratisation of data in the era of data science is analysed. Section 3 presents the

growing importance of AI and data democratisation through different applications for insights-driven decisions. A data democratisation framework is presented in Section 4 that enables AI-based cross-sector analytic tools for intelligent energy management, based on seamless data information and knowledge exchange under respective sovereignty and regulatory principles. Finally, the last section summarises the key issues that have arisen in this paper.

## 2. Data Democracy in the Era of Data Science

Thousands of terabytes per day [23], containing insightful information and knowledge, may be lost, or stored without being processed properly. Moreover, few individuals or institutions in the world possess critical amounts of data [24,25], thereby having advantage over others; to them, data democracy is far from a 'lucrative' concept. However, there exist more barriers than the latter's reluctance to share their data. It seems that the concept of data open access follows the footprints of a past, yet insistent, controversy: open-source versus commercial software applications and operational systems. The difficulties and corresponding lessons learned from their coexistence could pave the way for establishing data democracy.

Open-source software inspirators sought to break the barriers established by commercial software monopolies and to make software accessible to all. However, people still use commercial software on a large scale, and IoT is not yet framed with open-source products. Similarly, smart devices, like smartphones and tablets, are exclusively using operating systems of oligopoly companies, which use but do not give the collected data back to society. Moreover, it is controversial to see many open-source software products being utilised by tech and consultancy giants to eliminate their product cost and increase their profits, without them giving at least credits to the developers of these products, or the corresponding communities.

These limitations hindering the expansion of open-source software resemble the barriers that data democratisation is facing: complexity, limited education or training from non-experts, lack of innovation and product evolution, and limited product support post-installation. In other words, the voluntary work of open-source communities is not enough in complex societies, with limited spare time, and large global division of labour. In essence, progress toward data democracy dictates that the public and private sectors alike be convinced for its importance and that stakeholders (including civil society and institutions) be transformed, rather than build up antagonistic institutions/communities to break into the existing regime.

### 2.1. Conceptualising Data Democratisation

Considering the above, people tend to define data democracy as something more than data open access [1]. Interestingly, data democracy is all about data open access. Extended definitions only seek to further elaborate its concept trying to establish it against real-world barriers. In brief, data democratisation is the process of bringing digital information to the average, non-expert end user. By doing so, the latter would be able to gather, process, and analyse data to reach critical conclusions (i.e., make decisions) without requiring outside help. The end goal is to transform people or employees into well-educated 'data science citizens' (or 'lay data scientists').

Education is not the only prerequisite, however. Data must be easily accessed from safe, secure, organised repositories at any time by anyone—with respect to legal confidentiality and privacy issues. Hence, the quality of the provided data is of paramount importance. It cultivates the necessary trust among the end users, who want to use unquestionably (the content of) data platforms implementing the data democracy concepts and principles. In that context, data democracy requires extra personnel (engineers, developers, business analysts, legal staff, etc.) to guarantee all the above. Even if the motivation lies in bypassing current data gatekeepers (IT personnel), data democracy may indirectly

empower other sectors (data science, data mining, data analytics, cloud computing, legal advisors, and so on).

Therefore, data democratisation should not be viewed as a stand-alone procedure that simply reduces labour cost and mitigates the barriers of an old-fashioned division of labour. It is rather an entirely new concept, which restructures the economy (i.e., production, logistics, management, etc.) and the social structure in terms of efficiency. This signifies a novel point of view regarding the relationship and power balance among citizens, institutions, governments and policymakers, employers, technocrats, and many more.

Data democracy seems to bring to life an economic structure that encompass the benefits, and transcends the boundaries, of both 'free market' and 'monopolistic/oligopolistic' economic models, in terms of access to data. The former resembles a decentralised system. The latter a hierarchical, strictly vertically organised system. A decentralised system is closer to the widespread globalisation economic process. However, the decentralised system lacks a fundamental assumption of the 'free-market' model: the absolute knowledge of all stakeholders over what is going on in the entire market (i.e., nobody has benefit over the others). This assumption though is only (theoretically) realistic in the 'monopolistic/oligopolistic' system. All production and related information/data are concentrated to the few vertically structured enterprises. The flexibility of the decentralised global economy could be combined with fuller/holistic access to data and knowledge, which is yet only conceivable in the 'monopolistic/oligopolistic' systems. In that context, stakeholders can have a holistic, fully informed point of view over different economic sectors.

### 2.2. Data Democracy in the Public Sector

Yet, the progress toward democratisation of data is slow, with data distribution remaining non-homogeneous along the global public sphere. The private sector is, broadly speaking, more reluctant to carry out radical reforms in this direction. On the other hand, Open Government Data (OGD) is gradually gaining acceptance [26–28]; see for examples the United States [29], India [30], and the United Kingdom [31]. OGD is data produced by governments or government-controlled entities and can be freely used and reused or redistributed by anyone. OGD promote transparency as well as accountability of governments (and their affiliated entities) to citizens, and can support scientists/researchers, politicians, and a diversity of other stakeholders to elaborate evidence-based policies and encourage citizens to take better-informed personal decisions.

The importance of the latter is ever evident during the COVID-19 virus pandemic, and it will be in its aftermath. Intuitively, it is considered that OGD will result in economic growth: companies are expected to make use of OGD, thereby promoting business creation and innovative, citizen-centric services. For example, the European Commission (EC) in 2011 expected that its Open Data Strategy for Europe would have improved the European economy by offering an annual 40-billion-euro boost [32]. However, this implied that companies would invest in the utilisation of this strategy (and open data in general) and reform their structural organisation accordingly [33]. The motivation, however, behind such an adaptation would be increased efficiency and profits.

### 2.3. Data Democracy in the Private Sector

Companies should be convinced that data democracy is worth the trouble. Despite the barriers, though, some companies have shown considerable interest in open data and some novel business models have emerged [34,35]. When it comes to data democracy in the private sector, the concept is significantly different than in the public sector: while OGD is freely available to all members of society (including the private sector), this is not the case in the private sector, where data democracy is perceived as the act of opening organisational data to as many employees as possible [36]. This limitation (i.e., this intra-organisational definition of open data) is restrictive for companies and citizens/consumers alike. In a decentralised way of production, where too many different companies are

involved in the production of final products and services, there must be inter organisational implementations (i.e., cloud platforms), which share open data. This would benefit the enterprises belonging to the same sector and/or chain of production and would likely provide consumers with the capacity to provide feedback about the end products.

### 2.4. Data and AI Democracy and the EU's Strategy to Overcome Related Barriers

The process of transforming non-experts into 'lay data scientists' must be discussed in parallel with the democratisation of data [37–39] and AI/machine learning [40–44]. Open data can be used to feed predictive models allowing non-experts to reach crucial data-driven decisions. Understanding the algorithms behind AI models nowadays, however, requires a basic grasp of computer science and a solid understanding of mathematics or statistics [45].

The lack of appropriate tools for capturing the real-time dynamics, the scarcity of and competition for AI experts, the need for knowledge transfer to new contexts and accordingly for training new AI for each different context, have started being addressed by the EC, through:

- The recent initiative to boost AI [46] to support everyday life, while creating a fertile ground for EU entrepreneurs and SMEs to create innovative AI-based services.
- The AI4EU platform [47], which has been funded by the European Union (EU) with a view to set up and manage a one-stop-shop environment where know-how resources, algorithms, solutions, and services related to AI are channelled.

However, the success of the above EU-level AI strategy strongly depends on the successful implementation of the European Strategy for Data [48], since AI solutions/services could be successfully adopted to the extent of facilitated and increased EC-level data sharing, while fully respecting privacy and other regulatory constraints.

Another critical barrier to be solved for AI to be expanded, which follows the selection of an AI algorithm for a given application, lies in the requirement of advanced programming skills for its utilisation in most commercial AI frameworks. Therefore, democratising AI requires that the corresponding frameworks, and especially their Graphic User Interfaces (GUI), be upgraded. In that sense, AI platforms that require minimum coding, if any, are also a much-needed step toward data democracy. Besides, it was only after mitigating the complexity of interaction between end users and computing machines that computing science (and its products and services) became popular, proliferated, and widely used by non-experts.

## 3. The Growing Importance of AI and Data Democratisation

AI and data democratisation is gaining ground and acceptance, yet the pace has hitherto been slow. This underexplored area has significant potential in both theory and practice. In that context, and to our knowledge, there are only a few press pieces and academic articles in this field of study.

### 3.1. A Brief Literature Overview

In the housing market, data democracy provides both sellers and consumers with insightful data and the corresponding indices to make vital decisions over future alternatives [49–51]. In the domain of healthcare, data democracy means more than decision support [52,53]: shared, big data repositories are perceived to strengthen synergies among research teams, laboratories, and the broader science community, facilitating cross-validation of scientific conclusions. In this case, however, data democracy should encompass responsible, legitimate, and secure sharing of personal, sensitive information of patients [54], as well as tracing medical conclusions to both data and expert opinion, to avoid rendering exaggeration or irresponsible advice authoritative (e.g., during the COVID-19 pandemic [55]).

Several applications of data democratisation can also be found in the field of energy resource, including data democratisation in offshore drilling [56], oil and gas operations [57,58],

as well as earth [59] and climate [60] sciences. These research activities are orbiting around efforts to gain useful insights into—and improve—resource and sectoral efficiency (e.g., in digital agriculture [61]).

More specifically, in climate change science and policy, there are increasing calls for 'unboxing' the 'black box' of the complex and diverse [62] quantitative systems modelling activities [63], which are predominantly used to inform decisions [64], and for opening them to non-experts [65]. This entails making science comprehensible and digestible [66], which goes beyond developing open-source tools [67] and instead aims to promote science co-production [68], inclusivity [69], and pluralism [70].

### 3.2. Data Democracy in the Energy Sector

Despite the energy sector being generally considered as rather conservative investment-wise, the advent of AI is boosted by the rising availability of much larger datasets being generated at an unprecedented rate at the edge of the grid. Grid-owned assets, such as Intelligent Electronic Devices (IEDs), transformers, feeders, Phasor Measurement Units (PMUs), as well as non-grid-owned assets, such as decentralised generation and smart buildings, flexible industrial and residential building-level or community-level aggregated loads, and model-based (simulated) data, are opening up to more distributed architecture, which incorporates more local-level data distributed integration, control, and applications, thus creating new innovation challenges.

Moreover, off-domain data, such as Energy Performance Certificates, building stock auditing, Energy Performance Contracts, weather conditions, climate data, geographical imagery, multimedia unstructured data sources, financial data on energy efficiency invest-ments, socioeconomic data, social media, energy end-user's characteristics and comfort levels, etc., may allow novel energy analytics to provide energy stakeholders with more robust actionable insights, if suitably integrated [22,71]. This creates the need for disruptive, cutting-edge approaches towards a trustful and legally binding data sharing culture, where data value is shared and exchanged amongst the members of the ecosystem, following a fair distribution of resources, hence facilitating a seamless integrated energy value chain and more efficient business processes [72].

Dealing with cross-domain data requires a precise understanding of their nature and radically new digital technologies, such as AI, IoT, as well as cloud and big data technologies, to efficiently process large quantities of data within tolerable elapsed times, with safety and security [73]. The lack of energy-tailored data ownership and sovereignty management policies has prevented in many cases from making available and sharing data while respecting privacy, protection and security, and appropriate learning-based modelling capabilities.

Distributed Ledgers and Blockchain technologies are ideal candidate technologies to implement decentralised privacy- and regulation-aware data governance and sovereignty, due to their intrinsic support for data immutability, transparency, accountability, and trace-ability, enabling full data provenance tracking (e.g., [74,75]). However, such technologies do not currently offer the necessary scalability to address the short latency requirements typical of the smart energy grids environments [76,77].

Finally, much like energy, climate, and biophysical limits, big data and data democracy have also been studied and/or promoted as concepts in the context of transitioning from a linear to a circular economic model [78–80], where user-friendly data can provide insights for improving decision-making in terms of efficient resource use. Democratised AI tools can help predict trends of future supply and demand, weather fluctuations, demographics, biophysical limitations, as well as rates of reducing waste/externalities, reusing materials, and recycling products.

It is evident that extensive innovative approaches are needed to assess the knowl-edge and tools developed over the years and extend them to a holistic, AI-powered, human-centric framework, aimed to optimise business processes and ensure the long-

term competitiveness across the energy sector [81] and intertwined domains (including environment, society, and economy) [82].

## 4. A Data Democratisation Framework for Intelligent Energy Management

In this section, we propose a framework that enables AI-based analytic tools for intelligent energy management, based on seamless data/information and knowledge exchange, under respective sovereignty and regulatory principles. It aims at evolving and scaling up innovative AI energy analytics services, which significantly contribute to achieve techno-economically optimal management of the energy value chain, especially for SMEs and non-tech end-users.

The data democratisation process for intelligent energy management can take several forms. More specifically, the framework consists of seven main components: (1) data discovery and interoperability, (2) data quality compliance, (3) data privacy and sharing, (4) an AI-based library, (5) analytics service builder and visualisation, (6) AI energy analytics services, and (7) security (Figure 1).
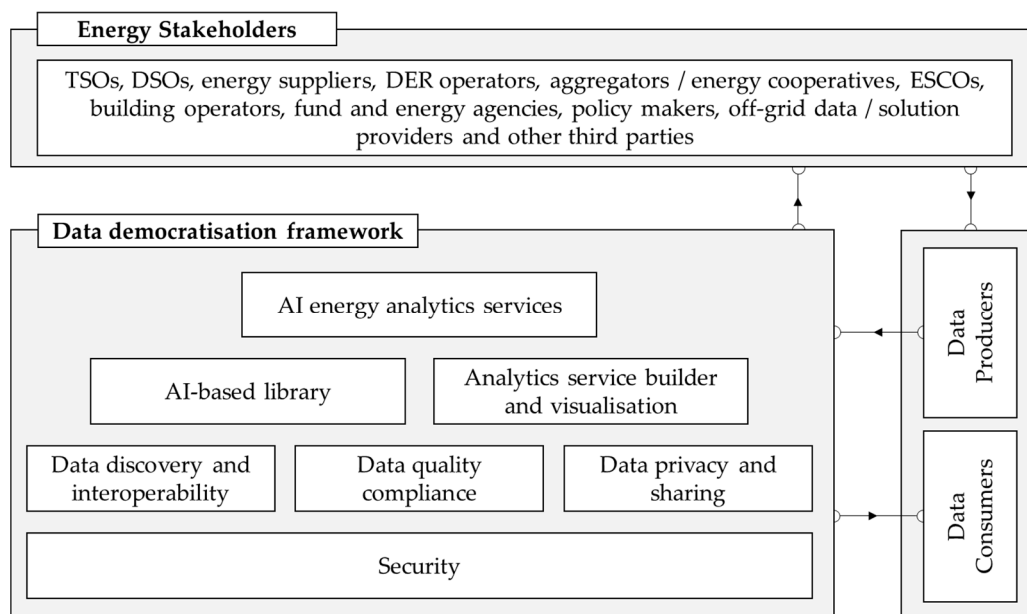


**Figure 1.** Data democratisation framework for intelligent energy management.

AI energy analytics services address the needs of a large number and diversity of energy stakeholders, including Transmission System Operators (TSOs), Distribution Service Operator (DSOs), energy suppliers, Distributed Energy Resource (DER) operators, aggregators and/or energy cooperatives, Energy Service Companies (ESCOs), building operators, funding and energy agencies, policymakers, off-grid data/solution providers, and other third parties. These AI-based services are grouped into the following categories:

- Increasing the efficiency and reliability of the electricity network.
- Optimising the management of DER assets connected to the grid.
- De-risking investments in energy efficiency and increasing the efficiency and comfort of buildings.

### 4.1. Data Discovery and Interoperability

This component deals with the discovery and interoperability of the various data from different sources and/or platforms. It provides the necessary functionalities to support end users (including data producers and/or data consumers, as well as non-expert end-users) to join the platform and search for and discover services matching their goals and business needs. A field for orchestrating semantic interoperability (e.g., based on the International

Data Spaces [83] for data sharing) describes the data exchange interfaces for read/write access and import/export of available data and output analytics.

Moreover, this component provides standardised data gates relying on existing or promising standardised data models. More specifically, it is based on open standards, open Application Programming Interfaces (APIs) like NGSI-LD CIM APIs, and open data models, such as FIWARE [84], BIM [85], SAREF/SAREF4Building [86], BRICK [87], and HAYSTACK [88]. This component includes connectors (plugins and IT interfaces) for already International Electrotechnical Commission (IEC) Common Information Model (CIM) standardised data and/or data platforms, as well as connectors to non-electrical/non-grid assets (gas, water, etc.), non-energy external data sources (sensing systems, weather data, etc.), and market data. Connection to the specific data hubs/systems is also driven.

### 4.2. Data Quality Compliance

The data quality compliance component is responsible for converting primary data collected from different sources into a format that renders them available for specific uses and suitable for generating new information and supporting decision-making. It allows integration, pre-processing, and querying of heterogeneous data. The ingestion of the datasets is followed by a series of data pre-processing actions. Data cleansing and curation focuses on the research, design, and development of the optimal data pre-processing pipelines.

### 4.3. Data Privacy and Sharing

Data sharing is an important step for democratising data. In this respect, this component deals with the integration of blockchain, enabling a secure, trusted, and flexible data/model sharing for B2B cross-stakeholders. This component handles smart energy cross-domain heterogeneous data and metadata after a process of harmonisation and integration. The data are maintained off-chain, ensuring privacy/GDPR compliance, and data sovereignty. Its core functionalities offer trusted mechanisms of data sharing and re-using.

A full governance cycle is implemented in order to support the creation, specialisation, update, storage, and maintenance of the Data Access Policies allowing also to dynamically derive access control policy rules based on contextual information (e.g., volume or frequency of data access, location of request, etc.). The resulting connector enables and facilitates fully interoperable, cross-stakeholder data access and sharing, while allowing data owners to maintain full control and sovereignty of their data, and fulfilling regulatory, privacy, business, and security constraints.

### 4.4. AI-Based Library

A library of reusable AI-based machine learning models is made available with a view to promote quick adaptation and re-use of machine learning models along different contexts. This component elaborates Automated Machine Learning (AutoML) methods and processes that render ML available to non-experts, as well as provide documentation and tutorials for each algorithm. In that context, various comprehensible metrics indicating, for example, the degree of accuracy, bias, and/or fit of a predictive model to training data, as well as robustly measuring the sensitivity of an AI model to these, can be important steps toward AI democratisation, and vital for the end-user to understand which algorithm best fits their needs.

This component also leverages on privacy-preserving federated learning, where data models are shared at the level of a subset of training models, hence yielding significant reduction of communication costs, and fully addressing data privacy concerns. Moreover, when partial data sharing is allowed, or data providers are willing to share their own data, the component combines unsupervised deep learning/reinforcement learning models for cross-context transfer learning (e.g., cross-context artificial intelligence by design) with inline adaptive self-learning to improve the resulting ML models.

*4.5. Analytics Service Builder and Visualisation*

The aim of this component is to develop and validate descriptive, predictive, and prescriptive/anticipatory data analytics along a variety of significant applications, by making use of the AI-based library. It incorporates the analytic service builder, in order to enable reuse of business logic capabilities on the selected datasets. It offers an environment to create customised services, based on a set of available analytical features, implementing different statistical and analytics algorithms, along with the functionality to perform custom queries on the available datasets.

Moreover, the visualisation engine, where the end user may have the opportunity to select one of more visual presentation modes, is enriched by innovative multi-dimensional visual analytics. These visualisations vary from simple statistical charts to map views and more, aiming to deliver useful understanding and insights coming from the data to the end users.

*4.6. AI Energy Analytics Services*

The AI energy analytics services component spans from analytics-based applications, tailoring power network improved decision-making to improve its reliability, to applications tailored to commercial aggregators/suppliers aimed to optimise management of flexibility assets, to applications addressing building-scale, energy-efficient comfort management, and dynamic renewable investment risk assessment.

4.6.1. Increasing the Efficiency and Reliability of the Electricity Network

The ability to analyse and manage vast volumes of data is increasingly important to improving operational efficiency and reliability of the electricity network. Proactive cross-functional, cross-domain analytics services for integrated grid-owned asset and grid operation management, tailored to power network operators, namely TSOs and DSOs, include:

1.  Predictive and prescriptive TSO/DSO grid-owned asset maintenance to facilitate and support grid optimal operation and/or planning, by trading off maintenance cost against accelerated asset ageing due to network overloads, and by combining and integrating grid endogenous and exogenous context-based information, such as Light Detection and Ranging (LIDAR), weather, and geographic.
2.  Edge-level network load and renewable energy generation prediction, by leveraging and integrating heterogeneous data from network assets, consumers and DERs smart meters, weather forecasting, geographical information, with a view to providing actionable insight to enable optimised grid operation and planning.

4.6.2. Optimising the Management of DER Assets Connected to the Grid

In the last two decades, the energy sector has undergone a significant transition. Government incentives and advances in technology have resulted in a proliferation of distributed renewable energy resources, particularly solar and wind generation. In addition, plug-in electric vehicles are now penetrating the market. As a result, the power load has an entirely different character and, in the resulting distributed grid, the energy and information flows are bidirectional. Therefore, it is critical to leverage available data to better understand energy demand, and align this to energy generation and distribution, to maximise operational efficiency. In particular, DER-level predictive analytics services include predictive analytics for DER/prosumer flexibility potential forecasting for: (i) optimal DER coordination and operation, (ii) optimal cross-energy (including electric vehicles) or cross-infrastructure (water network) management, (iii) voltage peak-to-peak (VPP) portfolio optimised management, and (iv) optimised power market operation.

### 4.6.3. De-Risking Investments in Energy Efficiency and Increasing the Efficiency and Comfort of Buildings

The built environment is responsible for 40% of final energy consumption [89]; therefore, it is important to increase the energy efficiency and comfort of buildings. Furthermore, there is a need to increase investments in energy efficiency measures, which can be achieved by de-risking investments, in particular by reliably predicting and monitoring energy savings. Predictive analytics for enhanced reliability and reduced risks of energy efficiency investments, tailored to ESCOs and financing institutions, are aimed at performing finer-grained prediction for building comforts. This is done by integrating a variety of historical data on energy efficiency investments, with near-real-time metered energy consumption, thereby contributing to better defining energy performance contract conditions.

### 4.7. Security

This last component is responsible for researching, maintaining, and reinforcing the legal and security policies and mechanisms employed (authentication, authorisation, auditing, policy-based management, and data encryption) including but limited to proactive vulnerabilities/risk assessment, secure session management, cross-site scripting protection, command injection flaws, buffer overflows protection, input sanitisation, and protection against logging breaches. For every new functionality of the platform, a validation process against existing data protection requirements takes place to ensure compliance with EU and national regulations and data protection acts. This component also ensures that the data transferred to the cloud for processing meet the security constraints—i.e., they are encrypted and anonymised where necessary.

## 5. Discussion

Despite the large number of technology-intensive organisations, their corporate know-how and underlying workforce skill are not mature enough for a successful rollout of AI services in the near future. However, things have begun changing, owing to the increased adoption of data democratisation processes, and the capability offered by emerging technologies for data sharing, while respecting privacy, protection, and security, as well as appropriate learning-based modelling capabilities for non-expert end-users.

In the energy sector, the constantly increasing momentum of data technologies and the growing trend toward a data economy, which leverages on exploiting the untapped economic value of legacy data assets, represents an unprecedented opportunity for energy stakeholders. These trends, combined with AI, constitute a catalyst toward conceptualising and generating innovative energy management services and applications.

They can gain significant improvements in the reliability and efficiency of electricity grid operation and planning; enable better and more efficient management of the grid-owned, behind-the-meter assets connected to the grid; while at the same time enabling a more reliable assessment of energy efficiency and renewable energy financial investments, taking into due consideration operational performance monitoring and comfort of citizens.

At the same time, such a trend paves the way for Energy Data Spaces, as the energy service data-driven collaborative ecosystem of the future.

However, energy stakeholders (including related SMEs and non-tech industries) are facing challenges in adopting AI, which may reduce the speed with which AI is adopted and thus limit its economic potential, with major challenges being represented by access to AI enablers, access to human capital, and AI-skilled talented workforce.

These challenges create the need for the democratisation of data and analytics in the energy sector, to make data available for the right people at the right time, enabling them to make the right decisions, thereby facilitating a seamless integrated energy value chain and more efficient business processes. Moreover, it is necessary to provide the appropriate framework toward a trustful and legally binding data sharing culture in the energy sector, where data value is shared and exchanged among the members of the ecosystem participants, following fair distribution of resources.

## 6. Conclusions

The data democratisation framework presented enables cross-sector analytic tools for intelligent energy management, based on seamless data/information and knowledge exchange under respective sovereignty and regulatory principles. These capabilities render AI as a new factor of production that can drive growth in at least three directions, those of energy commodities networks, DERs (renewable energy generation, buildings, districts, and communities), as well as synergies with and implications on other energy and non-energy domains (e.g., heat, transport, water, personal safety, and finance).

AI-based applications range from optimal risk assessment for energy efficiency investments planning, to optimised management of grid- and non-grid-owned assets, improved efficiency and reliability of electricity networks operation, as well as new socially and environmentally sustainable business models. These applications include a large number of energy stakeholders, including TSOs, DSOs, DERs operator, aggregators, ESCOs, building operators, and off-grid data/solution providers.

New, energy-based social service might enable pairing of public electric vehicle charging stations with adjacent renewable energy generation assets, and providing incentives for drivers to charge their vehicles in time intervals with surplus renewable energy capacity. It may also provide incentives to alleviate energy poverty at the local level, while offsetting for example energy surplus credits from free delivery to poor people from the energy bills in a peer-to-peer fashion, or targeting optimal trade-off among personal comfort and optimal energy management, or otherwise deliver personal safety/security and ambient assisted living services.

**Author Contributions:** Conceptualization, V.M. and H.D.; methodology, V.M. and A.N.; formal analysis, V.M., A.N. and T.K.; investigation, V.M., A.N. and T.K.; resources, V.M.; writing—original draft preparation, V.M.; writing—review and editing, V.M., A.N. and T.K.; supervision, H.D. All authors have read and agreed to the published version of the manuscript.

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

## References

1. Batarseh, F.A.; Yang, R. *Data Democracy: At the Nexus of Artificial Intelligence, Software Development, and Knowledge Engineering*, 1st ed.; Academic Press: Cambridge, MA, USA, 2020; p. 266.
2. Awasthi, P.; George, J.J. A case for data democratization. In Proceedings of the Americas Conference on Information Systems (AMCIS) 2020 Proceedings, Salt Lake City, UT, USA, 10–14 August 2020; p. 23. Available online: https://aisel.aisnet.org/amcis2020/data_science_analytics_for_decision_support/data_science_analytics_for_decision_support/23 (accessed on 14 June 2021).
3. Helbing, D.; Frey, B.S.; Gigerenzer, G.; Hafen, E.; Hagner, M.; Hofstetter, Y.; Van Den Hoven, J.; Zicari, R.V.; Zwitter, A. Will democracy survive big data and artificial intelligence? In *Towards Digital Enlightenment*; Helbing, D., Ed.; Springer: Cham, Switzerland, 2017.
4. Pujari, A.K. *Data Mining Techniques*, 1st ed.; Universities Press: Hyderabad, India, 2001; p. 288.
5. Olson, D.L.; Delen, D. *Advanced Data Mining Techniques*, 1st ed.; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2008; p. 180.
6. Kotsiantis, S.B.; Zaharakis, I.D.; Pintelas, P.E. Machine learning: A review of classification and combining techniques. *Artif. Intell. Rev.* **2006**, *26*, 159–190. [CrossRef]
7. Dey, A. Machine learning algorithms: A review. *Int. J. Comput. Sci. Inf. Technol.* **2016**, *7*, 1174–1179.
8. Dutton, D.M.; Conroy, G.V. A review of machine learning. *Knowl. Eng. Rev.* **1997**, *12*, 341–367. [CrossRef]
9. Brunette, E.S.; Flemmer, R.C.; Flemmer, C.L. A review of artificial intelligence. In Proceedings of the 2009 4th International Conference on Autonomous Robots and Agents, Wellington, New Zealand, 10–12 February 2009; pp. 385–392.

10.  Oke, S.A. A literature review on artificial intelligence. *Int. J. Inf. Manag. Sci.* **2008**, *19*, 535–570.
11.  Dhar, V. Data science and prediction. *Commun. ACM* **2013**, *56*, 64–73. [CrossRef]
12.  Provost, F.; Fawcett, T. Data science and its relationship to big data and data-driven decision making. *Big Data* **2013**, *1*, 51–59. [CrossRef]
13.  Waller, M.A.; Fawcett, S.E. Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *J. Bus. Logist.* **2013**, *34*, 77–84. [CrossRef]
14.  Van Der Aalst, W. Data science in action. In *Process Mining*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 3–23.
15.  Bhattarai, B.P.; Paudyal, S.; Luo, Y.; Mohanpurkar, M.; Cheung, K.; Tonkoski, R.; Hovsapian, R.; Myers, K.S.; Zhang, R.; Zhao, P.; et al. Big data analytics in smart grids: State-of-the-art, challenges, opportunities, and future directions. *IET Smart Grid* **2019**, *2*, 141–154. [CrossRef]
16.  Siozios, K.; Anagnostos, D.; Soudris, D.; Kosmatopoulos, E. *IoT for Smart Grids: Design Challenges and Paradigms*; Springer: Cham, Switzerland, 2019; p. 282.
17.  Caramizaru, A.; Uihlein, A. Energy communities: An overview of energy and social innovation. In *EUR 30083 EN*; Publications Office of the European Union: Luxembourg, 2020.
18.  Xu, Y.; Ahokangas, P.; Louis, J.N.; Pongrácz, E. Electricity market empowered by artificial intelligence: A platform approach. *Energies* **2019**, *12*, 4128. [CrossRef]
19.  Elavarasan, R.M.; Afridhis, S.; Vijayaraghavan, R.R.; Subramaniam, U.; Nurunnabi, M. SWOT analysis: A framework for comprehensive evaluation of drivers and barriers for renewable energy development in significant countries. *Energy Rep.* **2020**, *6*, 1838–1864.
20.  Henzelmann, T.; Hammermeister, F.; Wurm, B.; Nonnenmacher, L.; Preiss, S.; Schroer, K. *Artificial Intelligence: A Smart Move for Utilities*; Roland Berger: Munich, Germany, 2018.
21.  Davenport, T.; Loucks, J.; Schatsky, D. Bullish on the business value of cognitive. Leaders in cognitive and AI weigh in on what's working and what's next. In *The 2017 Deloitte State of Cognitive Survey*; Deloitte Development: New York, NY, USA, 2017.
22.  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]
23.  Desjardins, J. How Much Data Is Generated Each Day? World Economic Forum. 2019. Available online: https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f/ (accessed on 14 June 2021).
24.  Stucke, M.E. Should We Be Concerned about Data-Opolies? 2 Georgetown Law Technology Review 275 (2018), University of Tennessee Legal Studies Research Paper No. 349. Available online: https://ssrn.com/abstract=3144045 (accessed on 14 June 2021).
25.  McIntosh, D. We need to talk about data: How digital monopolies arise and why they have power and influence. *J. Technol. Law Policy* **2018**, *23*, 185.
26.  Magalhaes, G.; Roseira, C.; Manley, L. Business models for open government data. In Proceedings of the International Conference on Theory and Practice of Electronic Governance, Guimarães, Portugal, 27–30 October 2014.
27.  Martin, C. Barriers to the open government data agenda: Taking a multi-level perspective. *Policy Internet* **2014**, *6*, 217–240. [CrossRef]
28.  Janssen, K. The influence of the PSI Directive on open government data: An overview of recent developments. *Gov. Inf. Q.* **2011**, *28*, 446–456. [CrossRef]
29.  Data.gov. Available online: https://www.data.gov/ (accessed on 14 June 2021).
30.  Data.gov.in. Available online: https://data.gov.in/ (accessed on 14 June 2021).
31.  Data.gov.uk. Available online: https://data.gov.uk/ (accessed on 14 June 2021).
32.  European Commission. *Digital Agenda: Turning Government Data into Gold*; European Commission: Brussels, Belgium, 2011.
33.  Kaasenbrood, M.; Zuiderwijk, A.; Janssen, M.; de Jong, M.; Bharosa, N. Exploring the factors influencing the adoption of open government data by private organisations. *Int. J. Public Adm. Digit. Age* **2015**, *2*, 75–92.
34.  Ferro, E.; Osella, M. Eight business model archetypes for PSI re-use. In *Open Data on the Web Workshop*; Google Campus: London, UK, 2013.
35.  Janssen, M.; Zuiderwijk, A. Infomediary business models for connecting open data providers and users. *Soc. Sci. Comput. Rev.* **2014**, *32*, 694–711. [CrossRef]
36.  Zuiderwijk, A.; Janssen, M.; Poulis, K.; van de Kaa, G. Open data for competitive advantage: Insights from open data use by companies. In Proceedings of the 16th Annual International Conference on Digital Government Research, Phoenix, AZ, USA, 27–30 May 2015; pp. 79–88.
37.  Streeter, L.A.; Kraut, R.E.; Lucas, H.C.; Caby, L. How open data networks influence business performance and market structure. *Commun. ACM* **1996**, *39*, 62–73. [CrossRef]
38.  Pournaras, E.; Nikolic, J.; Omerzel, A.; Helbing, D. Engineering democratization in internet of things data analytics. In Proceedings of the 2017 IEEE 31st International Conference on Advanced Information Networking and Applications (AINA), Taipei, Taiwan, 27–29 March 2017; pp. 994–1003.
39.  Pournaras, E.; Gaere, E.; Kunz, R.; Ghulam, A.N. Democratizing data analytics: Crowd-sourcing decentralized collective measurements. In Proceedings of the 2019 IEEE 4th International Workshops on Foundations and Applications of Self* Systems (FAS*W), Umea, Sweden, 16–20 June 2019; pp. 265–266.

40. Allen, B.; Agarwal, S.; Kalpathy-Cramer, J.; Dreyer, K. Democratizing AI. *J. Am. Coll. Radiol.* **2019**, *16*, 961–963. [CrossRef]
41. Bagrow, J.P. Democratizing AI: Non-expert design of prediction tasks. *PeerJ Comput. Sci.* **2020**, *6*, e296. [CrossRef]
42. Banifatemi, A.; Miailhe, N.; Çetin, R.B.; Cadain, A.; Lannquist, Y.; Hodes, C. Democratizing AI for humanity: A common goal. In *Reflections on Artificial Intelligence for Humanity*; Springer: Cham, Switzerland, 2021; pp. 228–236.
43. Montes, G.A.; Goertzel, B. Distributed, decentralized, and democratized artificial intelligence. *Technol. Forecast. Soc. Chang.* **2019**, *141*, 354–358. [CrossRef]
44. Moreau, E.; Vogel, C.; Barry, M. A paradigm for democratizing artificial intelligence research. In *Innovations in Big Data Mining and Embedded Knowledge*; Springer: Cham, Switzerland, 2019; pp. 137–166.
45. Mixson, E. Make Data Accessible to Everyone with Data Democratization. Available online: https://www.aidataanalytics.network/data-democratization/articles/making-data-accessible-to-everyone-with-data-democratization (accessed on 14 June 2021).
46. European Commission. *White Paper on Artificial Intelligence—A European Approach to Excellence and Trust*; COM(2020) 65 Final; European Commission: Brussels, Belgium, 2020.
47. AI4EU—A European AI on Demand Platform and Ecosystem a European. Available online: https://www.ai4eu.eu/ (accessed on 14 June 2021).
48. 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.
49. McLaughlin, R.; Young, C. *Data democratization and spatial heterogeneity in the housing market. A Shared Future: Fostering Communities of Inclusion in an Era of Inequality*; Harvard Joint Center for Housing Studies: Cambridge, MA, USA, 2018; pp. 126–139.
50. Grey, J. The Democratization of Data. Housing Wire. Available online: https://www.housingwire.com/articles/40946-the-democratization-of-data/ (accessed on 14 June 2021).
51. Williams, D. How Big Data Will Impact Real Estate Buying, Selling and Developing. Mansion Global. Available online: https://www.mansionglobal.com/articles/how-big-data-will-impact-real-estate-buying-selling-and-developing-210771 (accessed on 14 June 2021).
52. Lewis, K.; Pham, C.; Batarseh, F.A. Data openness and democratization in healthcare: An evaluation of hospital ranking methods. In *Data Democracy*; Academic Press: Cambridge, MA, USA, 2020; pp. 109–126.
53. Kuiler, E.W.; McNeely, C.L. Knowledge formulation in the health domain: A semiotics-powered approach to data analytics and democratization. In *Data Democracy*; Academic Press: Cambridge, MA, USA, 2020; pp. 127–146.
54. Minielly, N.; Hrincu, V.; Illes, J. Privacy challenges to the democratization of brain data. *iScience* **2020**, *23*, 101134. [CrossRef]
55. Koch, T. Welcome to the revolution: COVID-19 and the democratization of spatial-temporal data. *Patterns* **2021**, *2*, 100272. [CrossRef]
56. Yoder, R.T. *Digitalization and Data Democratization in Offshore Drilling*; Offshore Technology Conference (OTC): Houston, TX, USA, 2019.
57. DiChristopher, T. Oil Firms Are Swimming in Data They Don't Use. CNBC. Available online: https://www.cnbc.com/2015/03/05/us-energy-industry-collects-a-lot-of-operational-data-but-doesnt-use-it.html (accessed on 14 June 2021).
58. Husseini, T. Big Data in Oil and Gas Operations and Other Tech Advancements: Seven Expert Opinions. Offshore Technology. Available online: https://www.offshore-technology.com/features/big-data-in-oil-and-gas-tech/ (accessed on 14 June 2021).
59. Yuan, K.; O'Neil, P.; Torrejon, D. Landsat's past paves the way for data democratization in earth science. In *Data Democracy*; Academic Press: Cambridge, MA, USA, 2020; pp. 147–161.
60. Faghmous, J.H.; Kumar, V. A big data guide to understanding climate change: The case for theory-guided data science. *Big Data* **2014**, *2*, 155–163. [CrossRef] [PubMed]
61. Chen, Y.; Li, Y.; Li, C. Electronic agriculture, blockchain and digital agricultural democratization: Origin, theory and application. *J. Clean. Prod.* **2020**, *268*, 122071. [CrossRef]
62. Nikas, A.; Doukas, H.; Papandreou, A. A detailed overview and consistent classification of climate-economy models. In *Understanding Risks and Uncertainties in Energy and Climate Policy*; Doukas, H., Flamos, A., Lieu, J., Eds.; Springer: Cham, Switzerland, 2019; pp. 1–54.
63. Pfenninger, S.; Hirth, L.; Schlecht, I.; Schmid, E.; Wiese, F.; Brown, T.; Wingenbach, C. Opening the black box of energy modelling: Strategies and lessons learned. *Energy Strategy Rev.* **2018**, *19*, 63–71. [CrossRef]
64. Doukas, H.; Nikas, A. Decision support models in climate policy. *Eur. J. Oper. Res.* **2020**, *280*, 1–24. [CrossRef]
65. Doukas, H.; Nikas, A. Involve citizens in climate-policy modelling. *Nature* **2021**, *590*, 389. [CrossRef]
66. Nikas, A.; Gambhir, A.; Trutnevyte, E.; Koasidis, K.; Lund, H.; Thellufsen, J.Z.; Doukas, H. Perspective of comprehensive and comprehensible multi-model energy and climate science in Europe. *Energy* **2021**, *215*, 119153. [CrossRef]
67. Huppmann, D. Open science has to go beyond open source. In *Nexus—The Research Blog If IIASA*; Intentional Institute for Applied Systems Analysis: Laxenburg, Austria, 2020.
68. Galende-Sánchez, E.; Sorman, A.H. From consultation toward co-production in science and policy: A critical systematic review of participatory climate and energy initiatives. *Energy Res. Soc. Sci.* **2021**, *73*, 101907. [CrossRef]
69. Nikas, A.; Elia, A.; Boitier, B.; Koasidis, K.; Doukas, H.; Casetti, G.; Chiodi, A. Where is the EU headed given its current climate policy? A stakeholder-driven model inter-comparison. *Sci. Total. Environ.* **2021**, in press. [CrossRef]

70. Nikas, A.; Skalidakis, S.; Sorman, A.H.; Galende-Sanchez, S.; Koasidis, K.; Serepas, F.; Doukas, H. Integrating integrated assessment modelling in support of the Paris Agreement: The I2AM PARIS platform. In Proceedings of the Twelfth International Conference on Information, Intelligence, Systems and Applications (IISA 2021), Chania, Crete, 12–14 July 2021.

71. Marinakis, V.; Doukas, H. An advanced IoT-based system for intelligent energy management in buildings. *Sensors* **2018**, *18*, 610. [CrossRef] [PubMed]

72. Marinakis, V. Big data for energy management and energy-efficient buildings. *Energies* **2020**, *13*, 1555. [CrossRef]

73. Cichocki, A. Era of big data processing: A new approach via tensor networks and tensor decompositions. *arXiv* **2014**, arXiv:1403.2048.

74. Kontzinos, C.; Kontoulis, M.; Kapsalis, P.; Markaki, O.; Mouzakitis, S.; Manta, R.; Thireos, E. Methodology for secure storage and information exchange of medical data based on blockchain. *Arch. Hell. Med.* **2020**, *37*, 542–554.

75. Kontzinos, C.; Markaki, O.; Kokkinakos, P.; Karakolis, V.; Skalidakis, S.; Psarras, J. University process optimisation through smart curriculum design and blockchain-based student accreditation. In Proceedings of the 18th International Conference on WWW/Internet, Cagliari, Italy, 7–9 November 2019; pp. 93–100.

76. Pop, C.; Antal, M.; Cioara, T.; Anghel, I.; Sera, D.; Salomie, I.; Raveduto, G.; Ziu, D.; Croce, V.; Bertoncini, M. Blockchain-based scalable and tamper-evident solution for registering energy data. *Sensors* **2019**, *19*, 3033. [CrossRef] [PubMed]

77. Pop, C.; Cioara, T.; Antal, M.; Anghel, I.; Salomie, I.; Bertoncini, M. Blockchain based decentralized management of demand response programs in smart energy grids. *Sensors* **2018**, *18*, 162. [CrossRef]

78. Jato-Espino, D.; Ruiz-Puente, C. Fostering circular economy through the analysis of existing open access industrial symbiosis databases. *Sustainability* **2020**, *12*, 952. [CrossRef]

79. Gupta, S.; Chen, H.; Hazen, B.T.; Kaur, S.; Santibañez Gonzalez, E.D.R. Circular economy and big data analytics: A stakeholder perspective. *Technol. Forecast. Soc. Chang.* **2019**, *144*, 466–474. [CrossRef]

80. Perella, M. Big Data and Circular Economy—The Revolution Will be Circular. Reuters Events. 2016. Available online: https://www.reutersevents.com/sustainability/big-data-and-circular-economy-revolution-will-be-circular (accessed on 14 June 2021).

81. Marinakis, V.; Doukas, H.; Psarras, J. Energy Management 4.0. In *Handbook of Research on Artificial Intelligence, Innovation and Entrepreneurship*; Grigoroudis, E., Elias, G.C., Eds.; Edward Elgar Publishing: Cheltenham, UK, 2021; in press.

82. Nikas, A.; Lieu, J.; Sorman, A.; Gambhir, A.; Turhan, E.; Baptista, B.V.; Doukas, H. The desirability of transitions in demand: Incorporating behavioural and societal transformations into energy modelling. *Energy Res. Soc. Sci.* **2020**, *70*, 101780. [CrossRef]

83. International Data Spaces Association. Available online: https://internationaldataspaces.org/ (accessed on 14 June 2021).

84. FIWARE. Available online: https://www.fiware.org/ (accessed on 14 June 2021).

85. Lee, S.-K.; Kim, K.-R.; Yu, J.-H. BIM and ontology-based approach for building cost estimation. *Autom. Constr.* **2014**, *41*, 96–105. [CrossRef]

86. Smart Appliances REFerence (SAREF) Ontology. Available online: https://sites.google.com/site/smartappliancesproject/ontologies/reference-ontology (accessed on 14 June 2021).

87. Brick. Available online: https://brickschema.org/ (accessed on 14 June 2021).

88. Project Haystack. Available online: https://project-haystack.org/ (accessed on 14 June 2021).

89. European Commission. Energy Use in Buildings. Available online: https://ec.europa.eu/energy/en/eu-buildings-factsheets-topics-tree/energy-use-buildings (accessed on 14 June 2021).