The BD4NRG Reference Architecture for Big Data Driven Energy Applications

Katharina Aurelia Wehrmeister
RWTH Aachen University
Aachen, Germany
katharina.wehrmeister@eonerc.rwth-aachen.de

Efthimios Bothos
National Technical University of Athens
Athens, Greece
mpthim@mail.ntua.gr

Vangelis Marinakis
National Technical University of Athens
Athens, Greece
vmarinakis@epu.ntua.gr

Babis Magoutas
National Technical University of Athens
Athens, Greece
elbabmag@mail.ntua.gr

Alexander Pastor
RWTH Aachen University
Aachen, Germany
alexander.pastor@eonerc.rwth-aachen.de

Leonardo Carreras
RWTH Aachen University
Aachen, Germany
leonardo.carreras@eonerc.rwth-aachen.de

Antonello Monti
RWTH Aachen University
Aachen, Germany
amonti@eonerc.rwth-aachen.de

***Abstract*— The rising digitisation of the energy system and related services is unveiling an enormous opportunity for energy stakeholders to leverage on Big Data & AI technologies for improved decision making and coping with challenges emerging from an increasingly complex and interconnected energy system. Initiatives in the field of Big Data Reference Architectures, like IDSA, GAIA-X or FIWARE provide generic frameworks to share, manage and process Big Data. Through alignment among them and integration of missing aspects, an interoperable and secure framework for the energy comes into view. The Reference Architecture presented in this paper moves towards this goal and will be instantiated in a set of concrete use cases within the European Energy Sector. Structurally inspired by SGAM and the BRIDGE Reference Architecture, it puts concrete analytics processes and data source components into context, taking important issues of Data Governance, Security, and Value Creation into account.**

***Keywords*— Big Data for Energy, Data Exchange Platform, Data Value Creation, Reference Architecture.**

I. INTRODUCTION

The energy sector is transforming continuously with the steady growth of distributed renewable generation and energy storage systems [1]. With the additional expected rise of prosumers in demand response and electric mobility, new challenges arise for system operators. These include embedding new assets into the market and exploiting their flexibility while maintaining grid resilience, efficiency, and reliability. To tackle this, many data exchange platforms (DEPs), e.g., those examined by the BRIDGE initiative in Section II.A, have been developed in the context of their respective projects. These platforms are focused on data gathering, processing, and exchange to enable various services including market access. The BD4NRG project (https://www.bd4nrg.eu/) aims to develop a platform harmonising existing tools and systems from this landscape, to cater to a broad range of use cases covering topics across the energy sector; specifically, operation of the electricity network (BD-4-NET), management of distributed energy resources (BD-4-DER), and efficiency and comfort of buildings (BD-4-ENEF).

This paper presents the first iteration of the BD4NRG Reference Architecture (RA), detailing how existing approaches were brought together and extended to meet the needs of those use cases.

The remainder of the document is structured as follows. Section II presents approaches related to the BD4NRG Reference Architecture, providing an overview of the state of the art constituting the base of its development. Section III describes the set of use cases and resulting requirements for the RA. The Reference Architecture itself is presented in Section V, and Section V concludes with a summary of the contribution and a view of future work.

II. ANALYSIS OF EXISTING ARCHITECTURES

The following Sections introduce existing innovations in the field of Big Data Reference Architectures, describing their focus areas and features, and how they are relevant to BD4NRG RA development.

A. BRIDGE

BRIDGE is an initiative that started with the Horizon2020 program and now continues under Horizon Europe. Under its umbrella, various stakeholders in more than 90 projects cooperate to address cross-cutting issues in the areas of smart grid, energy storage, islands and digitalisation. These projects are concerned with technologies on different levels including consumer, grid, energy storage, and generation. BRIDGE is organised in four working groups: Data Management, Regulation, Consumer and Citizen Engagement, and Business Models [2] [3]. The Data Management Working Group in particular works on three main topics: communication infrastructure, cybersecurity & data privacy, and data handling. It has explored a dozen data exchange platforms to make recommendations towards building the business-agnostic BRIDGE Reference Architecture. The result is a multi-layered, cross-sectoral architecture model based on the Smart Grid Architecture Model SGAM. (cf. Figure 1) [4] [5].

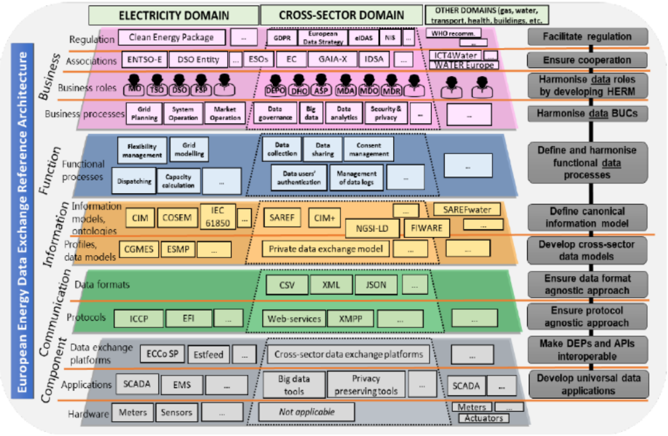

Figure 1 – BRIDGE Reference Architecture. Source: [2]

Following Figure 1, the layers of the BRIDGE RA are analogous to the SGAM interoperability layers: Business Layer, the Function Layer, the Information Layer, the Communication, and the Component Layer. The Business Layer encompasses the regulatory legal framework, business associations, roles, and processes. The Function Layer includes decision-making processes and services based on the data provided by the Information Layer. The Information Layer in turn is responsible for the structuring of the data in information and data models such that they can be exchanged through the Communication Layer in an interoperable fashion. The Communication Layer deals with standardising protocols and formats to facilitate data exchange. Lastly, the Component Layer includes all kinds of physical and virtual devices or entities. [5]

This Reference Architecture's basic structure serves as a starting point to the RA presented in Section IV. The components of each use case described in Section III.A were mapped to the BRIDGE RA's layers, resulting in an overview of the extent of its alignment with these real-world applications. Within this analysis, concerns of data governance as well as privacy and security emerged, covering the entire structure rather than just single layers, as well as the need for more extensive and specific Component- and Function layers.

B. *IDS-RAM*

The International Data Spaces Association (IDSA) aims to facilitate secure and standardised data exchange and data linkage through existing standards and technologies [5]. Their IDS Reference Architecture Model (IDS-RAM) focuses on data sovereignty and includes requirements for secure and trusted data exchange in business ecosystems. Clearly defined user roles and the so-called IDS connector enable a peer-to-peer network approach [6].

Core Participants of IDS are Data Owners, Data Providers, Data Consumers and Data Users. Further user roles include Intermediaries such as a Service Provider, which act as trusted entities between Provider and Consumer [5]and the Vocabulary Provider, which handles the IDS information model along with ontologies and reference data models [6]. Figure 2 depicts the relations between the available Roles.

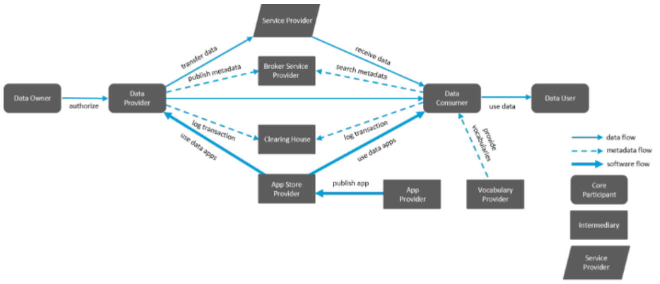

Figure 2 – IDS-RAM Participants. Source: [6]

The central technological building block of IDS is the IDS Connector. It acts as a gateway and allows the trusted execution of apps on the cloud [7]. These apps can cover functionalities ranging from secure bidirectional communication to system monitoring and logging, and may be extended through "Data Apps", custom software which would e.g., enable data processing or visualisation. [6]

The layers of the IDS-RAM are depicted in Figure 3. They consist of the business layer defining user roles and their interactions, the functional layer responsible for trust and sovereignty, the process layer facilitating data exchange, information layer with vocabularies, and finally the system layer defining the Connector's architecture along with Data Apps and the Broker. [6]

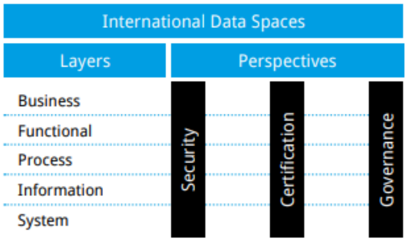

Figure 3 – IDS-RAM Layers and Perspectives. Source: [6]

The IDS-RAM and its information model are sector-agnostic, meaning they are not specialised to e.g. the energy sector. [8] However, the Vocabulary Provider may be extended by sector-specific vocabularies.

All in all, the IDS-RAM provides a clear framework for sovereign data exchange. Due to its goal of being technology-agnostic [8], it doesn't e.g. define specific standard APIs or data models. Similarly, being sector-agnostic means that the model stays at a rather high abstraction level [6] and doesn't include distinctions between different layers of analytics, or relations between different types of e.g. data users.

This generality is useful for the BD4NRG Reference Architecture: it provides a way to uniformly handle data governance issues across the broad range of given use cases. Specifically, the "Data Sovereignty and Trust" aspect of the RA is inspired by the IDSA structure.

C. GAIA-X and Integration with IDS

The GAIA-X project aims to tackle the current challenges to the European data environment, such as decentralisation, transparency and sector-specific data spaces and ontologies [5]. The GAIA-X Reference Architecture consists of two ecosystems: A Data Ecosystem for advanced smart services and data spaces, and an Infrastructure Ecosystem responsible for portability, interoperability, and interconnectivity as well as compliance. They are connected through GAIA-X federation services and define the three user roles Provider, Consumer and Federator [9].

Services within the GAIA-X system can be deployed on generic Nodes, which have a known certification level and are made available to the ecosystem via a Provider [10]. A service may be deployed on one or on multiple Nodes [9].

In general, GAIA-X focuses more on cloud services and cloud infrastructure than IDSA, which puts emphasis on data and data sovereignty. Both initiatives aim to use existing technologies and tools [11].

This integration is possible by locating the Core Participants of IDS in the two different ecosystems of GAIA-X (Figure 4).

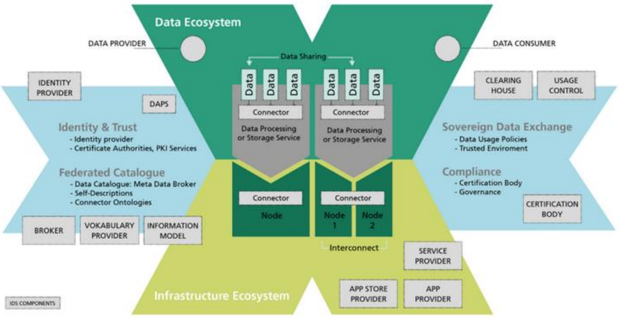

Figure 4 – GAIA-X Ecosystem. Source: [12]

In 2021, GAIA-X and IDSA came together and published a position paper on the integration of both approaches [11]. IDS Connectors as secure gateways can be integrated in GAIA-X nodes to cover both ecosystems. This integration forms an architecture combining the data approach of IDS with the cloud perspective of GAIA-X.

Regarding APIs and data models, an "Architecture of Standards", collecting the most relevant existing standards, is in progress [11] [13].

Leveraging the existing possible links with IDS, this combination includes concepts important to data value creation relevant within the BD4NRG Reference Architecture - such as marketplace functionalities, usage control, etc.

D. FIWARE Technological Ecosystem

The FIWARE Foundation is an EU-sponsored non-profit organisation that promotes the FIWARE technological ecosystem. This ecosystem aims at providing a modular, open, public, royalty-free software platform to enable a plethora of smart applications: smart energy, smart manufacturing, smart cities, smart agriculture etcetera to mention just a few. The initiative is backed by a great number of diverse partners – among others, universities, industry, and local communities [14]. From the technical perspective, three major factors contribute to broad interoperability and applicability: NGSI-LD, the so-called "Generic Enablers" and the smart data models.

NGSI-LD is an information model based on property graphs and an API that allows JSON-LD querying [15]. Its layered approach allows stacking custom sector-specific data models on top of the RDF grounding, NGSI-LD meta-model, and NGSI-LD Cross-Domain Ontology layers (cf. Figure 5).

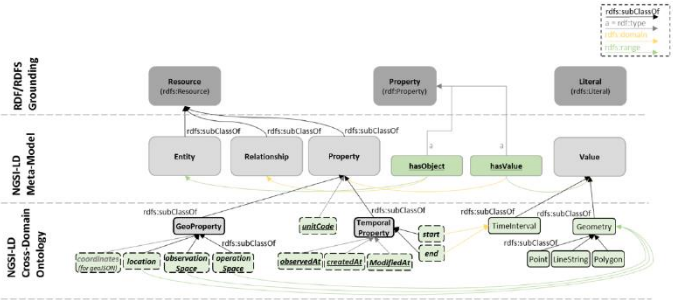

Figure 5 – NGSI-LD Information Model. Source: [15]

Generic Enablers are standalone modules that can be considered the building blocks of a data platform each fulfilling a well-defined purpose. Some building blocks are responsible for establishing the required trust between the data provider and data consumer, while ensuring the sovereignty of the provider using FIWARE's open-source implementation of the IDS connector. Another collection of blocks forms the so-called Business Application Ecosystem, generating value from data assets that can be published, discovered, traded, and consumed on the marketplace.. These modules are interconnected through a central NGSI context broker, which itself is considered a Generic Enabler. A platform can be considered "powered by FIWARE" simply by using a FIWARE context broker [5], [16].

The smart data models are grouped into subjects, e.g., street lighting or weather, each available as a GitHub repository. These subjects are each in turn referenced by one or more domains. For instance, the smart cities domain refers to both street lighting and weather [17]. To ensure quality, there is a multi-stage review process, i.e. models can be in the incubation or harmonisation stage before they are finally accepted or merged into an existing model, respectively. [18]

The FIWARE ecosystem provides a sound approach to interoperability, which is an important concern for the BD4NRG system, but does not directly offer concrete analytics or marketplace-related functionalities. As a result, while it cannot cover the entire span of the BD4NRG system, FIWARE compliance informed the development of the Reference Architecture's Interoperability Layer.

III. BD4NRG USE CASES AND REQUIREMENTS

A. Use Cases

The BD4NRG Reference Architecture has been developed to support a set of data analytics use cases covering different aspects of the energy sector. The specific use cases are grouped in 3 main clusters BD-4-NET, BD-4-DER, and BD-4-ENEF and provide requirements on the different layers of the BD4NRG RA as described in the following. Note that the use cases will be piloted in operational settings within the BD4NRG project.

1) BD-4-NET: Increasing the efficiency and reliability of the electricity network

The ability to analyse and manage vast volumes of data is increasingly important to improve operational efficiency and the reliability of the electricity network. The "BD-4-NET" use cases focus on predictive analytics that can forecast maintenance needs and increase the efficiency and reliability of the electricity network. They are applied in the following thematic areas:

a) Integrating off-grid Data with Condition based Monitoring for Enhanced Predictive Asset Maintenance: Asset management departments of system operators are traditionally data intensive and improved data analytics tools can be used to derive predictive models for assets maintenance. Predictive maintenance of Circuit Breakers and Overhead Lines depends on LIDAR data, power flows and incident reports that are processed to determine the probability of failure and generate a semi-automatic predictive maintenance plan.

b) Cross-functional Integration of Grid Operation with Predictive Asset Management: The operation of the power system with a large share of distributed variable Renewable Energy Sources (RES) requires new data driven services, as keeping the grid balanced becomes more complicated. The necessary scheduling of generation and consumption requires high availability of data about the status and load level of a large and diverse number of grid elements. This use case leverages SCADA real-time data, measurement instruments, periodic inspection reports and data from maintenance events for predictive asset management and outage planning.

c) Cross-Domain Cross-functional Integration of Predictive Asset Management and Grid Operation while coordinating with EVs smart charging: The increased share of intermittent power generation from renewable sources and new loads such as energy storage require optimised grid operation to rely on real-time data. Cross-functional predictive analytics provide the means for coordinating new loads from the e-mobility value chain, using data from phasor measurement units, EV chargers, smart meters, environmental and geographical conditions. The challenge is increased by the heterogeneous nature of the data and its sources, resulting in a need for harmonisation and standardisation of data- and information models as well as communication procedures.

d) Cross-functional Predictive Asset Management for MV Grid Planning: This use case focuses on predicting and analysing short term future maintenance and failure of network assets such as transformers, and hence failure along the grid. AI based predictive maintenance results allow operators to avoid breakdowns, connect new supplies at the optimal points of the network (for MV) and assess the effectiveness of investments.

2) BD-4-DER: Optimising the Management of Distributed Energy Resources (DER) Connected to the Grid

Over the last years, DER have continuously been introduced in energy grids. As a result, the electrical load has an entirely different character and in the emerging 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. The "BD-4-DER" use cases focus on optimising the management of assets connected to the grid in the following thematic areas:

a) Cross-stakeholder Transfer Learning for Flexibility Assets Forecasting: Within this use case, DSOs leverage residential data from devices like smart meters, EVs, and heat pumps as well as weather forecasts to perform data-driven and cross-context transfer learning. The result is a demand response approach supporting the grid by mitigating voltage and congestion issues.

b) Coordinating grid-owned and behind-the-meter assets management prediction for improved grid operation and near real time power market operation settlement: The escalating grid complexity necessitates data-driven transformation to instantiate situational awareness, allow the operation closer to the margin, and react on network events/disturbances. The goal of this use case is to enhance the flexibility and coordinated management of grid connected assets. The involved DSOs will leverage predictive analytics to enable proactive and corrective operation schemes.

c) Predictive Analytics Forecasting of storage optimised operation for improved market reserve participation: BESS (Battery Energy Storage Systems) plants integrated in the power system provide ancillary services. This use case aims to improve maintenance, observation, troubleshooting and monitoring operations on large-scale solar sites, by implementing predictive big data analytics. Using meter data as well as simulations, prediction of critical events and reduction of plant unavailability is achieved.

d) Predictive Demand and Generation Forecasting for Optimised Local Energy Community Management: This use case aims to provide improved analytics-based forecasting regarding local generation by RES and flexible/controllable loads, using weather service data, cross-domain data sharing, and non-grid data from controllable loads (e.g. EVs, water pumping systems). [19]. Analytics are used for optimal scheduling for peak shaving of deferrable loads, supportig real time control, fast response optimisation, improving energy availability security and preventing power interruptions or shutdowns for residential and commercial customers.

e) Collaborative aggregated energy generation prediction: This use case focuses on improving prediction of RES power production within energy communities, supporting predictive maintenance. Within a testbed including a large-scale solar plant and a self-consumption system with EV charging and storage components, the scheduling of maintenance operations (such as solar panel cleaning), predictive fault identification as well as energy production forecasts are implemented.

3) BD-4-ENEF: De-Risking Investments in Energy Efficiency and Increasing the Efficiency and Comfort of Buildings

Reducing energy consumption is of growing importance to the EU. In the field of energy efficiency measures it is absolutely critical to measure and integrate diverse heterogeneous sources of data [20]. Moreover, those sources must have a high degree of confiability with a focus on providing evidence to stakeholders, enabling both the analysis of actual consumption and energy efficiency and the prediction of the best pathways for investments with the maximum returns in terms of reduced energy costs and sustainability [21].

a) Predictive and Prescriptive Analytics for improved Energy Performance Certificates Reliability: EPC has been in place since the implementation of the Energy Performance Directive 2010 (2010/31/EU). Nevertheless, the lack on interoperability among data sources and the complexity of its management still make the implementation of the EPCs unreliable and, therefore, practically unfeasible. The aim of this use case is to enable scalable integration and management of vast heterogeneous cross-domain and cross-stakeholders amounts of data in a harmonised way by means of data-driven mechanisms. Automatic anomaly detection ensures cleanliness and reliability of the data and derived results.

b) Weather-enhanced building thermal comfort prediction to enable energy efficiency reliable forecasting: The thermal comfort of the occupants of buildings is very important for the general acceptance of measures towards higher energy efficiency.. This use case shows how cross-domain integration of building/district energy consumption data with off-grid weather data enables Big Data analytics applications. These predict the index of average comfort at building level, and related trade-offs with energy efficiency. Near real time comfort-enhanced decision-making by building or district operators is necessary for timely thermal comfort prediction. ML services will bridge the gap to optimise the tradeoff between the possible measures adopted and the perception of comfort.

c) Predictive Analytics for Energy Efficiency Investments de-risking: Energy efficiency projects are often fragmented, with high transaction costs, which makes it difficult to accurately asses financial risks of investments. Integrating cross-domain financial and energy consumption data is necessary to build the necessary market confidence in energy efficiency projects to make them an attractive investment asset class. Services like assessment of efficiency impacts of investments in energy savings, based on energy monitoring and financial data, will evaluate the quality and cost-effectiveness of projects [22]. This requires taking in account cross-sector information like legal and environmental regulations and requirements.

B. Requirements

The use cases presented in Section III.A provide requirements for a Reference Architecture that harmonises approaches like those presented in Section II and enables value creation and exchange of data across the different stakeholders in the energy sector.

In contrast to agnostic ideas like GAIA-X, the BD4NRG Reference Architecture needs to be specific to the energy sector, offering space for sector-specific business roles, analytics functionalities, and data sources. In the business layer, this includes acknowledging real-world stakeholders directly involved in technical procedures, such as Distribution- and Transmission System Operators, as well as regulations by e.g. the European Union, as use cases like the one described in Section III.A.3)c) have to interface with them.

Regarding analytics, the presented use cases have some specific needs, including complex tasks such as predictive maintenance and optimal scheduling for peak shaving. These procedures require layered functionalities that rely on each other. As a result, the RA should depict not only services performed on data directly, but enable layers of energy functionalities in turn using those services.

In addition to catering specifically to the energy sector, there is also a need to be less high-level and technology-agnostic than existing works. A high level of heterogeneity not just between different applications, but also within single use cases (see e.g. Section III.A.1)c)) makes it necessary to take clear steps toward interoperability by defining a layer for interfacing with data models, formats, and communication protocols.

Furthermore, compliance with a broad range of data sources is required. Implementations of this RA shouldn’t be restricted to data directly produced by hardware or gathered through control applications such as SCADA systems, but also be capable of acquiring e.g. weather data via public APIs or external databases.

Finally, large-scale data sharing always comes with risks. Use cases such as the one described in Section III.A.2)a) leverage residential data to improve analytics and control, and need to make sure this sensitive information is not misused or shared in ways disallowed by data owners. This responsibility needs to be reflected by ensuring the necessary functionalities for data space governance, security and trust throughout the whole architecture.

IV. THE BD4NRG REFERENCE ARCHITECTURE

The BD4NRG Reference Architecture aims to specify and structure stakeholders and components that need to be considered by smart and decentralised data-driven energy ecosystems to create value out of Big Data. Through a concise and detailed RA, BD4NRG foresees to overcome barriers hampering the exploitation potential of energy Big Data, offering a standardised way of managing, sharing and using data and related analytics services.

The BRIDGE RA described in Section II.A forms the base of the BD4NRG Reference Architecture. To asses its suitability for real-life applications, the BRIDGE RA was analysed with respect to a wide range of use cases discussed in Section III.A, which revealed additional requirements, as discussed in Section III.B. These requirements informed the development of the Reference Architecture, which addresses them by drawing on aspects of IDS, GAIA-X and FIWARE among others. The resulting RA is described in the following Sections, and graphically represented in Figure 6.

The architecture is structured in different building blocks: four layers and one pillar. The four layers represent those of the data value chain, while the vertical pillar accommodates data space enablers which are relevant across all the layers, allowing data to be shared in a trusted and sovereignty-preserving manner. In the following we provide more details for each of the building blocks of the BD4NRG RA.

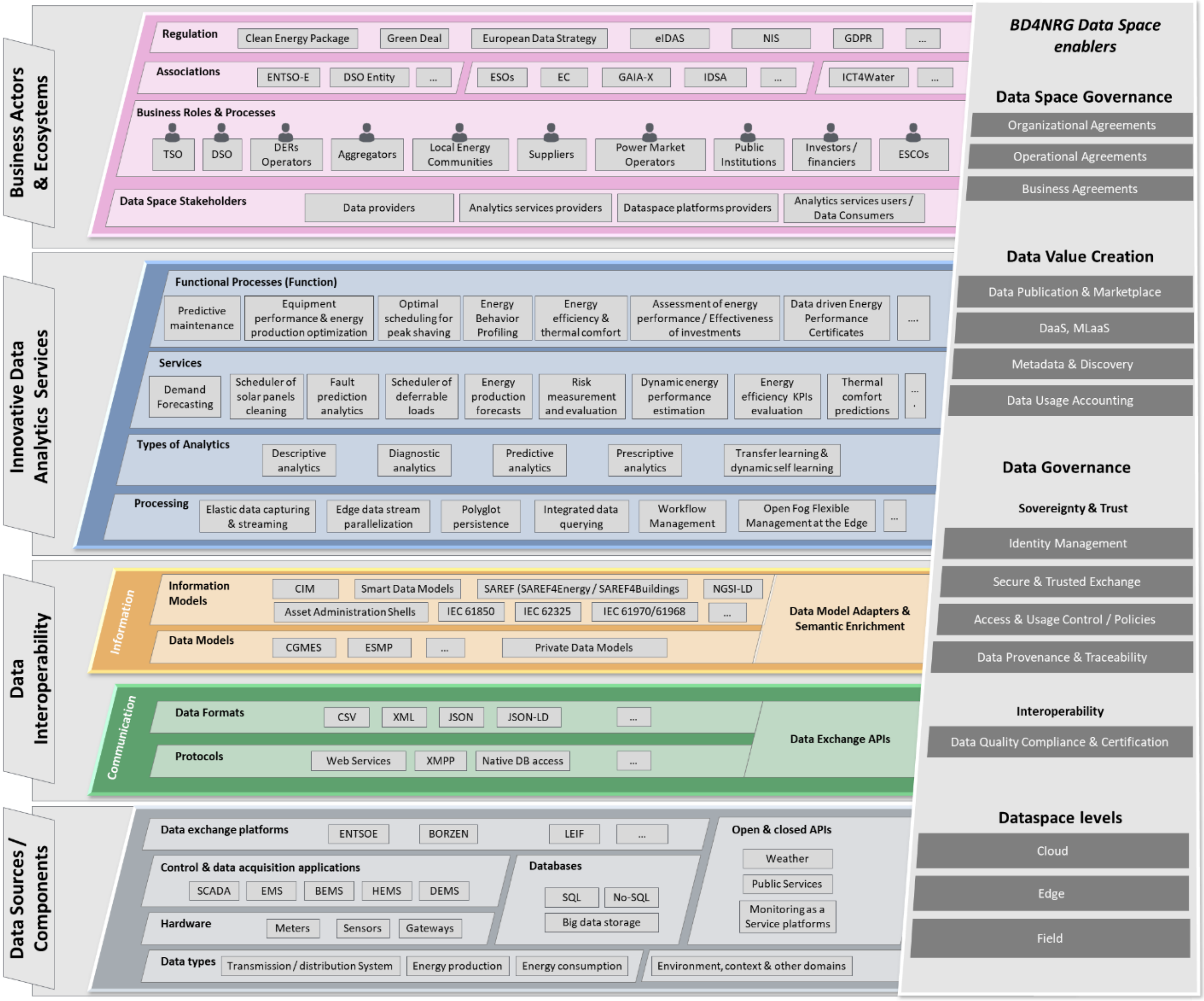

Figure 6 – BD4NRG Reference Architecture.

A. Layers

The **Data Sources Layer** represents the components for ingestion of data provided by numerous sources, including energy sensors and meters, data monitoring and acquisition platforms such as SCADA systems or Building Energy Management systems, databases with historical or real time data, smart grid data exchange platforms and cross-domain information such as environmental information and data coming from public administration services (e.g., related to energy performance contracts). Essentially, this layer includes all the data generating hardware, applications and platforms as described in the Component Layer of the BRIDGE architecture.

The **Data Interoperability Layer** identifies protocols and formats for communication between analytics services and data sources, as well as providing a set of data and information models to be used for data transformation. This addresses the diversity of data sources and structures as described in Section III, across different applications as well as within individual use cases, aiming to ensure interoperable data exchange and use. This layer integrates the “Communication” and “Information” layers of the BRIDGE architecture, increasing specificity by explicitly mentioning elements like NGSI-LD and Smart Data Models (see discussion of FIWARE in Section II.D).

The **Innovative Data Analytics Services Layer** includes and structures Big Data analytics functionalities, drawing on the individual analytics needs of the BD4NRG use cases. It is comprised of four sublayers, the lowest of which specifies Big Data processing infrastructure working directly on ingested data, followed by different "Types of Analytics" that are employed depending on user requirements and data analytics use cases. The "Services" sublayer represents the analytics services that utilise processed energy and cross-sector Big Data. Finally, the "Functional Processes" sublayer is comprised of the most complex applications, directly informed by use cases' goals and supported by the lower analytics services. This layer represents the "Function" layer of the BRIDGE architecture by consolidating aspects related to data analytics and correspondingly supported functions.

The **Business Actors and Ecosystems Layer** identifies stakeholders who participate in the data analytics ecosystem and corresponding energy data spaces. They include data providers and analytics services users / data consumers, analytics applications providers and providers of data space enablers and related platforms. Note that an organisation or business role in the energy ecosystem can have one or more roles in data space, for example a TSO can be a data provider and a consumer of analytics services at the same time.

B. *Cross-Layer Elements*

The vertical pillar of BD4NRG **Data Space Enablers** identifies the components and functions required to realise the distributed BD4NRG data space and is aligned with the design principles specified by IDSA and GAIA-X.

A main element for smooth data space operation is the specification and provision of a set of rules or agreements which should be followed by all stakeholders. This set of **Data Space Governance** agreements define how stakeholders interact within it on organisational, operational and business levels. IDSA has already provided a rule book specifying blueprints of agreements which should be considered when implementing a data space [23].

The data space governance is complemented by a set of technical components for **Data Governance** that need to be in place in order to ensure data sovereignty and trust, track the origin (provenance) and usage of data, and enforce the interoperability specification by applying measures for data quality compliance and/or certification (note that in the BD4NRG RA, interoperability aspects related to communication interfaces and data/information models are part of the Data Interoperability Layer).

Data Value Creation within a data space is accomplished through Data Publication and Marketplace Services, which facilitate the exchange or transaction of data and the use of data analytics services. Marketplaces are spaces where data producers, analytics services providers and users of the services meet with the purpose of creating value from the available data. An efficient data marketplace is supported by services that allow data and metadata discovery as well as data usage accounting. Moreover, analytics services are provided in "as a service" manner (e.g., machine learning as a service). Essentially, energy stakeholders in need of analytics services buy such services instead of developing these inhouse, and benefit from reduced development and maintencance costs. Moreover they sell data to interested developers of services. Analytics services providers earn profit by offering innovative data analytics pipelines.

Finally, the BD4NRG RA integrates the various **Dataspace Levels** as specified by IDSA and GAIA-X. The first of which includes data storage, integration and analytics services operating on the *Cloud*. Scalable cloud resources provide the means for advanced Big Data driven services and the provision of DaaS/MLaaS. Mission critical processing is performed at the *Edge* or in the organisations' premises, whereas the cloud is used for long term Big Data storage and/or for offloading analytics tasks that can be executed on the cloud. Finally, other data analytics services are deployed in the *Field* (at the user premises) and are used in locally deployed infrastructure. Such an approach may be advantageous for organisations that are reluctant to make the transition to the cloud due to business, organisational or other restrictions.

V. CONCLUSIONS AND FUTURE WORK

The BD4NRG RA makes multiple contributions toward cross-domain Big Data Platforms in the energy sector.

Firstly, BD4NRG RA extends the BRIDGE RA in two ways: it adds sublayers, namely Services, Types of Analytics, and Processing to the Function Layer of BRIDGE comprising the Innovative Data Analytics Layer. It also adds concrete elements to these and other layers directly derived from the BD4NRG UCs that cover a broad range of topics from the electrical grid to distributed energy resources and energy efficiency in buildings. Furthermore, the BD4NRG Data Space Enabler pillar highlights the data space aspects previously not covered. These encompass data governance, data value creation, sovereignty and trust, interoperability, and dataspace levels. These aspects, which are to some extent covered in IDS-RAM, FIWARE, and GAIA-X, will be incorporated into the ongoing implementation of the BD4NRG system based on this RA. They are currently being developed in the form of harmonised and adapted implementations of components from, among others, these initiatives.

Future work on the RA will progress along with this development, updating as the continuation of use case trials reveals further requirements and challenges.

ACKNOWLEDGMENT

The work presented is based on research conducted within the framework of the H2020 European Commission project BD4NRG (Grant Agreement No. 872613). The content of the paper is the sole responsibility of its authors and does not necessary reflect the views of the EC.

REFERENCES

[1] V. Marinakis, H. Doukas, J. Tsapelas, S. Mouzakitis, Á. Sicilia, L. Madrazo und S. Sgouridis, From big data to smart energy services: An application for intelligent energy management, *Future Generation Computer Systems*, Bd. 110, pp. 572-586, 2020.

[2] K. Kukk, K. Kostalos, E. Lambert and O. Genest, "BRIDGE Reference Architecture Webinar," 07 2021. [Online]. Available: https://www.h2020-bridge.eu/register-for-the-reference-architecture-cross-border-and-cross-sector-energy-data-exchange-webinar/. [Accessed 10 12 2021].

[3] M. Jareño, “BRIDGE 2021 Brochure (Publications Office of the European Union),” 09 2021. [Online]. Available: https://op.europa.eu/en/publication-detail/-/publication/abf32809-143e-11ec-b4fe-01aa75ed71a1/language-en. [Accessed 10 12 2021].

[4] E. Lambert, G. Boultadakis, K. Kukk, K. Kotsalos and N. Bilidis, "BRIDGE - European energy data exchange reference architecture," BRIDGE Data Management Group, [Online]. Available: https://ec.europa.eu/energy/sites/default/files/documents/bridge_wg_data_management_eu_reference_architcture_report_2020-2021.pdf. [Accessed 10 12 2021].

[5] K. Wehrmeister and et al, "BD4NRG Reference Architecture (First Version)," BD4NRG project. [Online]. [Accessed 10 12 2021].

[6] B. Otto, S. Steinbuß, A. Teuscher and S. Lohmann, “IDSA Reference Architecture Model Version 3.0,” International Data Spaces Association, [Online]. Available: https://internationaldataspaces.org/wp-content/uploads/IDS-RAM-3.0-2019.pdf. [Accessed 10 12 2021].

[7] International Data Spaces Association, "Enabling Data Economy Brochure," IDSA, [Online]. Available: https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/IDSA-brochure-International-Data-Spaces-Enabling-Data-Economy.pdf. [Accessed 10 12 2021].

[8] International Data Spaces Association, “Reference Architecture,” IDSA, [Online]. Available: https://internationaldataspaces.org/use/reference-architecture/. [Accessed 10 12 2021].

[9] Federal Ministry for Economic Affairs and Energy (BMWi), "GAIA-X: Technical Architecture," Project GAIA-X, Jun 2020. [Online]. Available: https://www.data-infrastructure.eu/GAIAX/Redaktion/EN/Publications/gaia-x-technical-architecture.pdf?__blob=publicationFile&v=5. [Accessed 10 Dec 2021].

[10] GAIA-X Project, “Glossary,” GAIA-X Website, [Online]. Available: https://www.gaia-x.eu/glossary. [Accessed 10 12 2021].

[11] B. Otto, “GAIA-X and IDS (1.0),” Zenodo, 01 01 2021. [Online]. Available: https://doi.org/10.5281/zenodo.5675897. [Accessed 10 12 2021].

[12] International Data Spaces, "Position Paper GAIA-X and IDS," IDSA, [Online]. Available: https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/IDSA-Position-Paper-GAIA-X-and-IDS.pdf. [Accessed 25 04 2022].

[13] Federal Ministry for Economic Affairs and Energy (BMWi), “GAIA-X: Policy Rules and Architecture of Standards,” GAIA-X Project, May 2020. [Online]. Available: https://www.data-infrastructure.eu/GAIAX/Redaktion/EN/Publications/gaia-x-policy-rules-and-architecture-of-standards.pdf?__blob=publicationFile&v=5. [Accessed 10 12 2021].

[14] FIWARE Foundation, e.V., "About us," FIWARE, [Online]. Available: https://www.fiware.org/about-us/. [Accessed 10 12 2021].

[15] ETSI Industry Specification Group, "ETSI GS CIM 006 V1.1.1," ETSI, [Online]. Available: https://www.etsi.org/deliver/etsi_gs/CIM/001_099/006/01.01.01_60/gs_CIM006v010101p.pdf. [Accessed 10 12 2021].

[16] FIWARE Foundation, e.V., “Developers Catalogue,” FIWARE, [Online]. Available: https://www.fiware.org/developers/catalogue/. [Accessed 10 12 2021].

[17] Smart Data Models Initiative, "Github Repository," Smart Data Models initiative, [Online]. Available: https://github.com/smart-data-models. [Accessed 10 12 2021].

[18] FIWARE Foundation, e.V., "Smart Data Models Initiative," FIWARE, [Online]. Available: https://www.fiware.org/developers/smart-data-models/. [Accessed 10 12 2021].

[19] E. Sarmas, E. Spiliotis, V. Marinakis, G. Tzanes, J. K. Kaldellis und H. Doukas, ML-based energy management of water pumping systems for the application of peak shaving in small-scale islands,“ *Sustainable Cities and Society,* Bd. 82, p. 103873, 2022.

[20] V. Marinakis, Big data for energy management and energy-efficient buildings,“ *Energies,* Bd. 13, Nr. 7, p. 1555, 2020.

[21] G. Hernández-Moral, S. Mulero-Palencia, V. I. Serna-González, C. Rodríguez-Alonso, R. Sanz-Jimeno, V. Marinakis, N. Dimitropoulos, Z. Mylona, D. Antonucci und H. Doukas, Big Data Value Chain: Multiple Perspectives for the Built Environment, *Energies,* Bd. 14, Nr. 15, p. 4624, 2021.

[22] E. Sarmas, E. Spiliotis, V. Marinakis, T. Koutselis und H. Doukas, A meta-learning classification model for supporting decisions on energy efficiency investments,“ *Energy and Buildings,* p. 111836, 2022.

[23] International Data Spaces Association, "IDSA Rule Book," IDSA, [Online]. Available: https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/IDSA-White-Paper-IDSA-Rule-Book.pdf. [Accessed 25 04 2022].

[24] V. Araujo, K. Mitra, S. Saguna and C. Åhlund, "Performance evaluation of FIWARE: A cloud-based IoT platform for smart cities," *Journal of Parallel and Distributed Computing,* pp. 250-261, 2019.