PART II

9

Digital Models for Industrial Automation Platforms

Nikos Kefalakis, Aikaterini Roukounaki and John Soldatos

Kifisias 44 Ave., Marousi, GR15125, Greece
Email: nkef@ait.gr; arou@ait.gr; jsol@ait.gr

This chapter presents the role and uses of digital models in industrial automation applications of the Industry 4.0 era. Accordingly, it reviews a range of standard-based models for digital automation and their suitability for the tasks of plant modelling and configuration. Finally, the chapter introduces the digital models specified and used in the scope of the FAR-EDGE automation platform, towards supporting digital twins and system configuration use cases.

9.1 Introduction

The digital modelling of the physical world is one of the core concepts of the digitization of industry and the fourth industrial revolution (Industry 4.0). It foresees the development of digital representations of physical world objects and processes as a means of executing automation and control operations, based on digital operations functionalities (i.e. at the cyber rather than at the physical world) [1]. The motivation for this stems from the fact that digital world operations can be flexibly altered or even undone at a low cost, while this is impossible in the physical world. Hence, plant operators can experiment with operations over digital models, run what-if scenarios and ultimately derive optimal deployment configurations for automation operations, while also deploying them on the field based on IT applications and tools, such as Industrial Internet of Things (IIoT) tools.

The concept of simulating and experimenting with automation operations in the realm of digital models of the plant is conveniently called "digital twin" and is a key enabler of the digitization of industrial processes. One of the automation platforms developed by the co-authors of this book, namely the FAR-EDGE edge computing platform, takes advantage of this concept based on the integration of digital models that represent the manufacturing shopfloor, as well as other physical and logical components of the automation platform such as edge gateways. In particular, the FAR-EDGE reference architecture and platform design specify a set of digital models as an integral element of the FAR-EDGE automation platform. In line with the Industry 4.0 "digital twin" concept, these digital models serve several complementary and important objectives:

- **Digital Simulation:** FAR-EDGE implements digital twins in order to support digital simulations of the plant, including what-if scenarios. The latter can be evaluated and used to decide optimal configurations of automation elements.
- **Semantic Interoperability:** The FAR-EDGE digital models provide a uniform representation of the concepts and entities that comprise a FAR-EDGE deployment, which boosts semantic interoperability across diverse digital systems and physical devices. The use of common data model provides a uniform vocabulary for describing various entities (e.g. sensors, CPS devices, SCADA Supervisory Control and Data Acquisition systems, production systems) across different applications in the automation, analytics and simulation domains of the platform.
- **Information Exchange:** The digital models in FAR-EDGE provide the means for exchanging information across different FAR-EDGE deployments. This is closely related to the above-listed semantic interoperability objective: By exchanging information in a common agreed format, two or more different FAR-EDGE deployments can become interoperable despite differences in their internal implementation details.
- **System Configuration:** The design and deployment of digital models is a key prerequisite for performing automation and control operations at IT (Information Technology) timescales. As part of the digitization of industrial processes, automation systems (i.e. Operational Technology (OT)) can be configured through IT systems and tools. The latter configure and update digital models, which reflect the status of the physical world. In this way, automation and configuration operations are performed at the level of IT rather than at the level of OT (Operational

Technology). This requires a synchronization between digital models and the status of the physical world, which is challenging to implement.

This chapter provides insights into the digital models that are used to support information exchange, digital simulations, semantic interoperability and digital operations as part of the FAR-EDGE platform. It first analyses the rationale behind the specification and implementation of digital models in the FAR-EDGE platform, along with some of the main requirements that drive the specification of the models. These requirements include standards compliance, extensibility, high performance, as well as support of FAR-EDGE functionalities in the platform's simulation domain. Following the review of these requirements for the FAR-EDGE digital models, we present a number of standards-based models (i.e. digital models and schemas specified as part of Industry 4.0 standards) against their suitability in supporting these requirements. As part of this chapter, we highlight the suitability of AutomationML and the standards-based schemas that it comprises (e.g. CAEX) for the simulation functionalities of the FAR-EDGE platform. Accordingly, we introduce a range of new proprietary models that can represent FAR-EDGE deployment configurations, based on concepts that cover the platform's edge computing model to automation and distributed data analytics. Specifically, we introduce new digital models that reflect concepts specified and used as part of the FAR-EDGE RA and the edge computing infrastructure of the project, such as edge gateways, data channels, measurement devices, as well as live data streams. These concepts can be blended with AutomationML and CAEX concepts as a means of putting plant models (e.g. CAEX instances) in the context of FAR-EDGE edge computing deployments.

Another important part of the chapter is the presentation of the linking between the above-listed models for edge computing configurations with the AutomationML-based models used for digital twins and digital simulations as part of the platform. The presented methodology is based on well-known concepts from the areas of data models linking and interoperability, including the concept of common repositories and registries for data models interoperability.

This chapter is structured as follows:

- Section 9.2 following the chapter's introduction presents the rationale behind the use of digital models in Industry 4.0 in general and in FAR-EDGE in particular;
- Section 9.3 reviews a set of standards-based digital models, which are commonly used for plant modelling and representation;

- Section 9.4 introduces the proprietary FAR-EDGE data models that are used for configuring the distributed data analytics functionalities of the platform;
- Section 9.5 presents a methodology for linking the FAR-EDGE proprietary data models with standards-based data models used for digital twins' representations in the platform's simulation domain.
- Section 9.6 is the final and concluding section of the chapter.

9.2 Scope and Use of Digital Models for Automation

9.2.1 Scope of Digital Models

Industry 4.0 applications are based on Cyber-Physical Systems (CPS). One of their main characteristics is that they implement automation functionalities at the cyber layer of production systems. In this context, they also take advantage of digital models as a pool of schemas and functions that are used for the digital representation of the factory, including the synchronization of their digital properties with the status of the real-world entities that they represent. At a finer level of detail, the functionalities of digital models support the operations and features that are described in the following paragraphs.

9.2.2 Factory and Plant Information Modelling

Primarily, digital models enable modelling of information at the factory and plant levels. In particular, the models provide a digital representation of the factory, which includes information about the elements (e.g. systems, devices and people) that comprise the plant. Automation and analytic applications can access the models in order to obtain information about the configuration of the plant, which they can use for implementing and validating automation processes. For example, in the FAR-EDGE project, digital models provide information about the hierarchical relationships between physical and logical entities in the scope of a FAR-EDGE deployment such as the sensors and devices that are associated with a given station.

In principle, a detailed and exhaustive description of the plant facilitates the implementation of many different processes and applications, including automation and analytics, as well as enterprise processes. However, developing and maintaining a detailed and exhaustive representation of the plant is very challenging. Therefore, FAR-EDGE and other digital automation platforms model only a subset of the plant, according to a "mini-world" that pertains to target automation and analytics use cases. Nevertheless, the digital modelling process can be open and extensible, in order to provide opportunities for supporting a broader set of functionalities and use cases, based on a fair additional effort.

9.2.3 Automation and Analytics Processes Modelling

Beyond a static representation of the structure of a factory and a plant, digital models should be able to represent the more dynamic automation and analytics processes, which form part of the plant's dynamic behaviour. Such processes should be represented based on the elements that are entailed in each processes, including their relationships and their evolution over time. Again, instead of an exhaustive modelling and representation of all possible workflows (e.g. through appropriate state machines), most automation platforms (including the FAR-EDGE automation platform) tend to focus on the processes that comprise a set of target use cases.

9.2.4 Automation and Analytics Platforms Configuration

The modelling of automation and analytics processing provides also a basis for their configuration and reconfiguration, as a means of changing the automation or the analytics logic based on IT functions. For example, using the digital model for an analytics process, it is possible to configure the devices and other data sources entailed in analytics processes, as well as the analytics (e.g. machine learning) algorithms applied on their data. This can provide increased flexibility in configuring and deploying different automation and analytics workflows in a factory. It can also support the implementation of the popular "plug and produce" concept [2].

9.2.5 Cyber and Physical Worlds Synchronization

As already outlined, the digital models can enable the configuration of automation functions and workflows at IT rather than OT times. In particular, automation operations can be configured at the IT layer of a digital automation platform, while being reflected in the physical world. The idea behind this configuration approach is that dealing with IT functions is much easier and more flexible than dealing with OT technology.

In order to provide this IT layer flexibility, there is a need to reflect changes in the IT layer to the OT layer (i.e. to the field) and vice versa. Hence, mechanisms for synchronizing the status of the physical world with its digital representation are needed based on digital models.

The synchronization between the physical and digital worlds can be also used to improve the results of digital simulations based on the so-called digital twins. In particular, it allows digital simulation applications to operate not only based on simulated data, but also with real data stemming from the synchronization of the physical and digital worlds. This can facilitate more accurate and realistic simulations, given that part of them can rely on real data that are seamlessly blended in the simulation application. The development of such realistic simulations is therefore based on dynamic access to plant information, which is illustrated in the following paragraph.

9.2.6 Dynamic Access to Plant Information

Digital models facilitate dynamic access to the status of the plant, at the cyber layer of the digital factory automation systems. Access to such information is needed in order to identify the configuration of automation processes, as well as the status of production processes and KPIs (Key Performance Indicators). Hence, digital models can serve as a vehicle for representing dynamic up-to-date information about the field, in addition to static (or semi-static) metadata of the shopfloor. One prominent use of the dynamic access to plant information involves the use of real-life data in order to boost the performance and accuracy of digital simulations, as outlined in the previous paragraph.

It should be underlined that the digital models specify the schema used to model the structure of the plant information. In the scope of the digital platform's operation, this schema is populated with instance data, which reflect the status of the plant at a given time instant. Hence, dynamic access to plant information is based on querying the instance of the plant database, which will follow the structure of the digital models. The concept of dynamic access to plant information and the importance of the synchronization between the digital models and the actual status of the plant is presented in Figure 9.1.

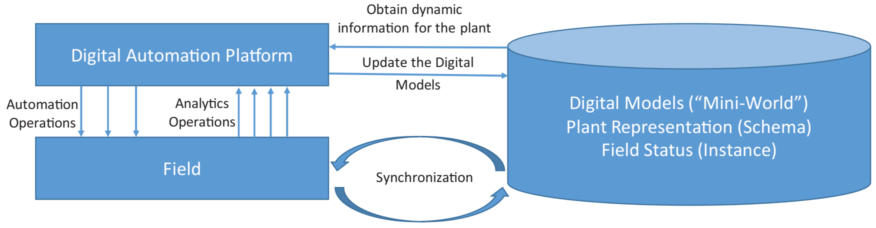

Figure 9.1 Digital Models and Dynamic Access to Plant Information.

9.3 Review of Standards Based Digital Models

In the following paragraphs, we provide a review of representative standards-based schemas that can be used for digital modelling in the scope of digital automation platforms. The review is by no means exhaustive, yet it covers some of the most popular models and schemas. Moreover, it provides some insights in terms of the ability of these standards to support the requirements and functionalities illustrated in the previous paragraph. Readers can also consult similar works on reviewing digital models (e.g. [3]), including works that have performed a comparative evaluation of alternative models [4].

9.3.1 Overview

For over a decade, various industrial standards have been developed, including information models that are used for information modelling in factory automation. Several standards come with a set of semantic definitions, which are typically used for modelling and exchanging data across systems and applications. These standards include, for example, the IEC 62264 standard that complies with the mainstream ISA-95 standard for factory automation. IEC 62264 boosts interoperability and integration across different/heterogeneous enterprise and control systems. Likewise, ISA-88 for batch processes comes with IEC 61512, and IEC 62424 supports exchange of data between process control and productions tools, while IEC 62714 covers engineering data of industrial automation systems [5]. Several of these standards are referenced and/or used by the RAMI 4.0 reference model [6], which is driving the design and development of several digital automation platforms. In the following paragraphs, we briefly describe some of these standards.

9.3.2 IEC 62264

IEC 62264 is a standard for enterprise-control system integration. It is based on the ANSI/ISA-95 hierarchy for automation systems. With reference to this hierarchy, the standard covers the domain of manufacturing operations management (i.e. Level 4) and the interface content and transactions within Level 3 and between Level 3 and Level 4. Hence, the standard is primarily focused on the integration between manufacturing operations and control, rather than on pure control (i.e. Levels 1, 2 and 3) operations only.

In practice, the standard defines activity models, function models and object models in the MOM (Manufacturing Operations Management) domain. The models are hierarchical and describe the MOM domain and its activities, the interface content and associated transactions within MoM level and between MoM and Enterprise level. Examples of entities that are modelled by the standard include materials, equipment, personnel, product definition, process segments, production schedules, product capabilities, production performance and more.

Note that IEC 62264 is among the standards referenced and used in RAMI 4.0. Due to its compliance with RAMI 4.0, IEC 62264 meets several of the requirements listed in the previous paragraph. However, it is focused on Level 3 and Level 4 entities of the ISA-95 standards and hence it is not very appropriate for use cases involving Levels 1, 2 and 3.

9.3.3 IEC 62769 (FDI)

The Information Model that is associated with the IEC 62769 (FDI) standard aims at reflecting the topology of the automation system. It therefore represents the devices of the automation system, as well as the communication networks that connect them. It includes attributes that are appropriate for modelling the main properties, relationships, operations of networks and field devices.

IEC 62769 is appropriate for modelling the field layer of the factory. This makes it appropriate for several of automation use cases, yet it does not provide the means for mapping and modelling some of the edge computing concepts of the FAR-EDGE automation platform (e.g. edge gateways and ledger services).

9.3.4 IEC 62453 (FDT)

IEC 62453 Field Device Tool (FDT) is an open standard for industrial automation integration of networks and devices. It provides standardized software to enable intelligent field devices that can be integrated seamlessly into automation applications, from the commissioning tool to the control system. FDT supports the coupling of software modules, which have been implemented as representatives for field devices and are therefore able to provide and/or exchange information. However, IEC 62453 is limited to the modelling of networks and devices and hence not suitable for plant-wide modelling.

9.3.5 IEC 61512 (Batch Control)

IEC 61512 – Batch control is also referenced by RAMI 4.0. It models batch production records, including information about production of batches or elements of batch production. IEC 61512 focuses on batch manufacturing and production processes.

9.3.6 IEC 61424 (CAEX)

IEC 61424 (CAEX) provides the means for modelling a plant in a hierarchical way and in an XML format (i.e. CAEX is provided as an XML Schema through an XML Schema language (XSD) file). CAEN abstracts a plant by considering it as a set of interconnected modules or components. CAEX models and stores such modules in an object-oriented way and based on object-oriented concepts such as classes, encapsulation, class libraries, instances, instance hierarchies, inheritance, relations, attributes and interfaces.

CAEX separates vendor-independent information (e.g. objects, attributes, interfaces, hierarchies, references, libraries, classes) and application-dependent information such as certain attribute names, specific classes or object catalogues. CAEX is appropriate for storing static metadata, but it is not designed to hold dynamic information. Nevertheless, it can be extended with special classes that could hold dynamic information and behaviour of the various modules.

IEC 61424 provides a sound basis for modelling the meta-data of a plant, which is one of the requirements for the digital models of an automation platform. However, there is also a need for supporting dynamic information as well, which asks for extensions to this model. CAEX is part of AutomationML compliant modelling, and as such, it is used in scope of FAR-EDGE in order to support the digital twins that are used from the simulation functionalities of the platform.

9.3.7 Business to Manufacturing Markup Language (B2MML)

B2MML is an XML implementation of the ANSI/ISA-95, Enterprise-Control System Integration, family of standards (ISA-95). As such, it is closely related to the above-listed IEC 62264 international standard, i.e. it provides a data representation that is fully compliant to the scope and semantics of IEC 62264. In practice, B2MML comprises a series of XML schemas, which are available as XML Schema language (XSD) files. Hence, B2MML supports the modelling of a large number of different entities, which represent MOM objects and transactions, as well as other interfaces between the enterprise and control layers.

B2MML is an excellent choice for supporting integration of business systems (such as Enterprise Resource Planning (ERP) and Supply Chain

Management (SCM) systems), with control systems (e.g. SCADA, DCS) and manufacturing execution systems (MES). This holds not only for B2MML compliant business systems (i.e. systems that support directly the interpretation of B2MML messages), but also for legacy ERP/SCM systems which can be made B2MML-compliant based on the implementation of relevant middleware adapters that transform B2MML to their own semantics and vice versa.

The language can be considered RAMI 4.0-compliant, given that RAMI 4.0 uses ISA-95 concepts and references of relevant standards (such as IEC 62264). It is also important that the B2MML schemas provide support for the entire ISA-95 standard, rather than a subset of it.

B2MML is characterized by compatibility with enterprise systems (e.g. ERP and PLM systems), which makes it appropriate for supporting information modelling for use cases involving enterprise-level entities and concepts. Furthermore, B2MML can boost compatibility with a wide range of available ISA-95-compliant systems, while at the same time adhering to information models referenced in RAMI 4.0. Therefore, B2MML could be exploited in the scope of use cases involving enterprises systems and entities, as soon as it is used in conjunction with additional models supporting concepts and entities for the configuration of an automation platform (e.g. like edge node, edge gateways and edge processes in the scope of an edge computing platform like FAR-EDGE).

9.3.8 AutomationML

AutomationML is an XML-based open standard, which provides the means for describing the components of a complex production environment. It has a hierarchical structure and is commonly used to facilitate consistent exchange and editing of plant layout data across heterogeneous engineering tools. AutomationML takes advantage of existing standards such as PLCopen XML or COLLADA. It provides the means for modelling plant information and automation processes based on objects structured in a hierarchical fashion, including information about geometric, model logic, behaviour sequences and I/O connections. AutomationML comprises different standards that support modelling for various entities and concerns. In particular, it relies on the following standards:

- CAEX (IEC 62424), in order to model topological information.
- COLLADA (ISO/PAS 17506) of the Khronos Group in order to model and implement geometry concepts and 3D information as well as Kinematics (i.e. the geometry of motion). Support for Kinematics ensures

also the modelling of connections and dependencies among objects as part of motion planning.
- PLCopen XML (IEC 61131) in order to model sequences of actions, internal behaviour of objects and I/O connections.

Note that AutomationML and the three above-listed standards are also in the list of Industry 4.0 standards that are directly connected to RAMI 4.0 in order to boost semantic interoperability.

AutomationML satisfies several of the requirements of the digital modelling requirements in FAR-EDGE and is appropriate for supporting digital simulations based on the development of a "digital twin" of the plant. It is therefore the standards-based digital model that supports plant modelling at the FAR-EDGE simulation domain. Moreover, the proprietary digital models that are used in FAR-EDGE can be linked to instances of AutomationML/CAEX digital models, towards ensuring uniqueness of the referenced entities and bridging of the diverse concepts that are captured by the two models. This is further discussed in Section 9.5.

9.4 FAR-EDGE Digital Models Outline

9.4.1 Scope of Digital Modelling in FAR-EDGE

In line with the uses of digital models that are described in Section 9.2, the FAR-EDGE digital automation platform leverages digital modelling for a dual purpose:

- **Data persistence and plant modelling for digital simulation and digital twins**. This is the reason why FAR-EDGE uses digital models for its simulation functionalities. The respective digital models are based on AutomationML, which has been described in the previous section.
- **Configuration of the FAR-EDGE platform, including configuration of its automation and analytics functionalities**. In particular, the FAR-EDGE platform holds digital presentations of the logical and physical configurations of FAR-EDGE components such as data sources, devices and edge gateways. The FAR-EDGE platform makes use of these configurations in order to configure its analytics and automation functionalities, based on functionalities such as the definition and configuration of data sources, association of these data sources to gateways and more. Specifically, the platform offers APIs and tools for manipulating these data models towards configuring the platform. The respective data models are proprietary and complement the use

of AutomationML in the simulation domain. The following paragraphs present briefly the data modelling entities used for the configuring data analytics in FAR-EDGE. The latter models come with an open-source implementation of functionalities for their management and are part of the FAR-EDGE platform. They are outlined in the following paragraphs.

9.4.2 Main Entities of Digital Models for Data Analytics

The proprietary FAR-EDGE data models that are used for configuring distributed data analytics functionalities, model factory data and metadata, along with the analytics functions and workflows that process them.

Factory Data and Metadata

The representation of factory data and metadata is based on the following entities:

- **Data Source Definition (DSD):** This defines the properties of a data source in the shopfloor, such as a data stream from a sensor or an automation device.
- **Data Interface Specification (DI):** The DI is associated with a data source and provides the information need to connect to it and access its data, including details like network protocol, port, network address and more.
- **Data Kind (DK):** This specifies the semantics of the data of the data source, which provides flexibility in modelling different types of data. The DK is an XML specification and hence it can be used to define virtually any type of data in an open and extensible way.
- **Data Source Manifest (DSM):** A DSM specifies a specific instance of a data source in line with its DSD, DI and DK specifications. Multiple manifests (i.e. DSMs) are therefore used to represent the data sources that are available in the factory in the scope of the FAR-EDGE automation platform.
- **Data Consumer Manifest (DCM):** This models an instance of a data consumer, i.e. any application that accesses a data sources.
- **Data Channel Descriptor (DCD):** A DCD models the association between an instance of a consumer and an instance of a data source. It is useful to keep track of the established connections and associations between data sources and data consumers.
- **LiveDataSet:** This entity models and represents the actual dataset that stem from an instance of a data source that is represented through

a DSM. Hence, it references a DSM, which drives the specification of the types of the attributes of the LiveDataSet in line with the DK. A LiveDataSet is associated with a timestamp and keeps track of the location of the data source in case it is associated with a mobile (rather than a stationary) edge node. Hence, it has a location attribute as well. In principle, the data source comprises a set of name–value pairs, which adhere to different data types in line with the DK of the DSM.

- **Edge Gateway:** This entity models an edge gateway of a FAR-EDGE edge computing deployment. In the scope of a FAR-EDGE deployment, data sources are associated with an edge gateway. This usually implies not only a logical association, but also a physical association, i.e. an edge gateway is deployed at a station and manages data sources in close physical proximity to the station.

Based on the above entities, it is possible to represent the different data sources of a digital shopfloor in a modular, dynamic and extensible way. This is based on a repository (i.e. registry) of data sources and their manifests, which keeps track of the various data sources that register to it. The FAR-EDGE platform includes such a registry, which provides dynamicity in creating, registering and using data sources in the industrial plant.

Factory Data Analytics Metadata

In order to facilitate the management and configuration of analytics functions and workflows over the various data sources, the FAR-EDGE digital models specify a number of analytics-related entities. In particular:

- **Analytics Processor Definition (APD):** This specifies a processing function to be applied on one or more data sources. In the scope of FAR-EDGE, three processing functions are defined, including functions that pre-process that data of a data source (i.e. Pre-Processors), functions that store the outcomes of the processing (i.e. Store Processors) and functions that analyse the data from the data sources (i.e. Analytics Processors). These three types of processors can be combined in various configurations over the data sources in order to define different analytics workflows.
- **Analytics Processor Manifest (APM):** This represents an instance of a processor that is defined through the APD. The instance specifies the type of processors and its actual logic through linking to a programming function. In the case of FAR-EDGE, the latter is a class/programme implemented in the Java language.

- **Analytics orchestrator Manifest (AM):** An AM represents an entire analytics workflow. It defines a combination of analytics processor instances (i.e. of APMs) that implements a distributed data analytics task. The latter is likely to span multiple edge gateways and to operate over their data sources.

9.4.3 Hierarchical Structure

The FAR-EDGE Digital Models for distributed data analytics follow a hierarchical structure, which defines the different relationships between the various entities. For example, an edge gateway comprises multiple data source manifests. Each one of the latter is associated with a data source definition. Likewise, LiveDataSets are associated with instances of data sources, i.e. data sources manifests. As an example, Figure 9.2 illustrates a snapshot of the FAR-EDGE digital models structure, which shows the association of each edge gateway with data source manifests and data analytics manifests. A more detailed presentation of the hierarchical structure of our data models is beyond the scope of this chapter. Interested readers can consult directly our XML schemas, which are part of our open source implementation of the

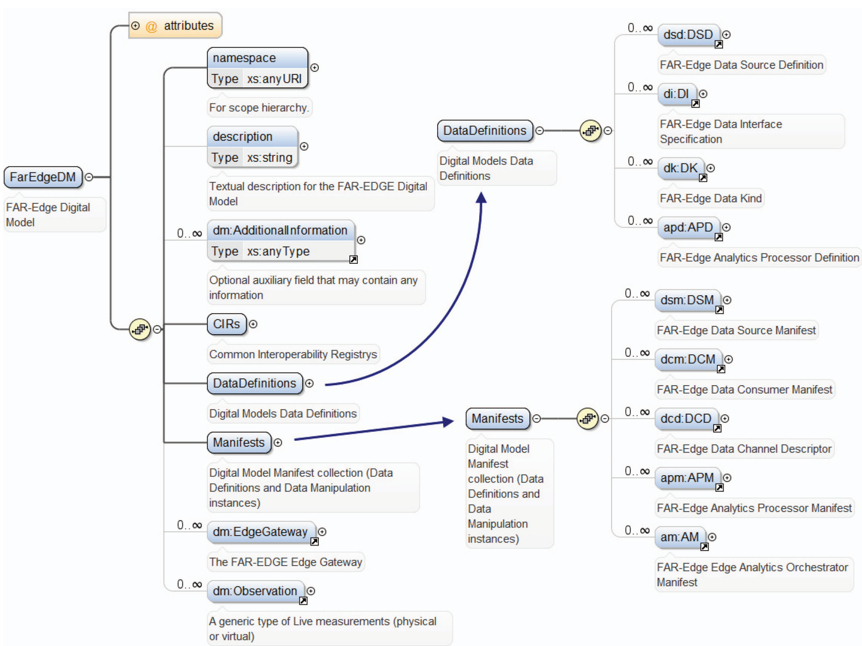

Figure 9.2 Snapshot of the FAR-EDGE Digital Models Structure.

FAR-EDGE digital models repository that is also an integral part of the FAR-EDGE platform.

9.4.4 Model Repository Open Source Implementation

As part of the open source implementation of the FAR-EDGE automation platform, we have implemented a data models repository, which provides support for the entities outlined in the previous paragraphs, including support for managing data kinds, data interfaces, data source definitions and analytics processor definitions. The implementation of the open source repository supports create, update, delete, get and discover functionalities, which are defined as follows:

- **Create:** This operation provides the means of creating an instance of the entity.
- **Update:** This allows updating an existing instance of the entity.
- **Delete:** This permits the deletion of an instance from the repository.
- **Get:** This fetches an instance of an entity based on its unique identifier.
- **Discover:** This helps model users to dynamically discover instances of one or more entities subject to given criteria.

The FAR-EDGE digital models repository implementation is available at the GitHub of the project at: https://github.com/far-edge/DigitalModels. The implementation comprises all schemata (i.e. see far-edge.dm.schemata) along with relevant ("generated") documentation in HTML (HyperText Markup Language) and PDF (Portable Data Format) formats. It also provides access to Java libraries, i.e. annotated libraries according to the JAXB (Java Architecture for XML Binding) framework in a proper Maven project (see far-edge.dm.commons). This open source implementation can provide a basis for researchers and engineers who might opt to implement their own digital models based on a similar approach. At the same time, they provide a means for implementing, using or even extending the FAR-EDGE analytics framework.

9.5 Simulation and Analytics Models Linking and Interoperability

The review of models in Section 9.3 justified the suitability of AutomationML for supporting the FAR-EDGE digital simulation functionalities. Furthermore, in Section 9.4, we introduced a digital model for representing and configuring the analytics functionalities of the FAR-EDGE platform. The use of a dedicated model for each of the two functional domains of the platform (i.e. analytics, simulation) provides flexibility to developers and deployers of analytics and simulation solutions, since they can use the model of their choice. Nevertheless, it could also create consistency and interoperability issues, especially in cases where functionalities and data from the two different domains need to be combined. To alleviate such problems, there is not only a need for linking entities in the two different models, so as to allow developers and deployers to access information for an entity in any of the two models, but also for combining information from the two models when needed. The merits of such linking become evident when considering the following examples:

- A digital simulation that needs to access information stemming from data analytics on real-life shopfloor data. For instance, a digital simulation may need to access maintenance-related parameters of a piece of equipment, following proper data analytics over sensor data (e.g. analytics vibration or ultrasound data for a machine). To this end, the machine representation in the simulation model (e.g. AutomationML) needs to be linked to the corresponding representation in the data model used for the distributed data analytics of the platform.
- Another digital simulation application that needs to analyse data sources using the distributed data analytics engine. In such a case, the simulation application needs to convey to the analytics engine the data sources to be used. To this end, there is a need for linking the representations of devices and data sources in the simulation domain, with the corresponding representations of the very same devices in the analytics domain.

In order to realize this linking, the FAR-EDGE data models include placeholders for data linking entities, i.e. linking of two representations of the same object/entity in different domains. In particular, both DSMs and logical entities in the simulation domain are linked based on a Universally Unique IDentifier (UUID). DSMs are assigned a UUID in an analytics domain whenever they are created and introduced to the system. Likewise, simulation applications assign a UUID to the main entities entailed in the simulation. The linking and harmonization of these UUIDs provide the means for linking the entities of two models.

This linking concept resembles to the concept of a Common Interoperability Registry (CIR), which is very commonly used in O&M (Operations and Maintenance). This registry is destined to provide "Yellow-Pages" lookup for all systems. This facilitates location of an object in any of the systems where it is registered, as soon as it is referenced with its UUID. Hence, different systems and models that have different identifiers for the very same entity or objects are glued together and are able talk "on-line". The main vehicle for this gluing is the specification and use of globally unique identifiers, which are linked to "local" object identifiers, i.e. identifiers pertaining to each one of the models.

9.6 Conclusions

This chapter has analysed the rationale behind the specification and integration of digital models in emerging digital automation platforms, which included a discussion of the main requirements that drive any relevant digital modelling effort. Moreover, it has presented a range of standards-based digital models, notably models that are used for semantic interoperability and information exchange in Industry 4.0 systems and applications. Following this review, it has illustrated why AutomationML is suitable for supporting the digital simulation functionalities of the FAR-EDGE platform.

The chapter has also introduced a proprietary model for representing and configuring the analytics part of the platform. This model provides the means for modelling and representing data sources and analytics workflows based on appropriate manifests. The respective models are implemented and persisted in a models repository, which is provided as a set of schemas and open source libraries as part of the FAR-EDGE digital automation platform. Hence, they can serve as a basis for using the FAR-EDGE digital models in analytics scenarios, as well as for implementing similar digital modelling ideas.

As part of this chapter, we have also outlined how globally unique identifiers can be used to link different models that refer to same entity or object in the factory based on their own local identifiers. The use of such global identifiers permits the association of entities referenced and used in both the AutomationML models of FAR-EDGE simulation and the FAR-EDGE models of the analytics engine. As part of our implementation roadmap, we also plan to implement a Common Interoperability Registry (CIR) that will keep track of all global identifiers and their mapping to local identifiers used by the digital models of the simulation, analytics and automation domains. This will strengthen the generality and versatility of our approach to digital model interoperability.

Overall, this chapter can be a good start for researchers and engineers who wish to start working with digital modelling and digital twins in Industry 4.0, as it presents the different use cases of digital models, along with the specification and implementation of a digital model for distributed data analytics in industrial plants.

Acknowledgements

This work was carried out in the scope of the FAR-EDGE project (H2020-703094). The authors acknowledge help and contributions from all partners of the project.

References

[1] H. Lasi, P. Fettke, H.-G. Kemper, T. Feld, M. Hoffmann, 'Industry 4.0', Business & Information Systems Engineering, vol. 6, no. 4, pp. 239, 2014.

[2] G. Di Orio, A. Rocha, L. Ribeiro, J. Barata, 'The prime semantic language: Plug and produce in standard-based manufacturing production systems', The International Conference on Flexible Automation and Intelligent Manufacturing (FAIM 2015), 23–26 June 2015.

[3] W. Lepuschitz, A. Lobato-Jimenez, E. Axinia, M. Merdan, 'A survey on standards and ontologies for process automation', in Industrial Applications of Holonic and Multi-Agent Systems, Springer, pp. 22–32, 2015

[4] R. S. Peres, M. Parreira-Rocha, A. D. Rocha, J. Barbosa, P. Leitão and J. Barata, 'Selection of a data exchange format for industry 4.0 manufacturing systems,' IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society, Florence, pp. 5723–5728, doi: 10.1109/IECON.2016.7793750, 2016.

[5] 'IEC 62714 engineering data exchange format for use in industrial automation systems engineering - automation markup language - parts 1 and 2', in International Electrotechnical commission, pp. 2014–2015.

[6] K. Schweichhart, 'Reference Architectural Model Industrie 4.0 - An Introduction', Deutsche Telekom, April 2016 online resource: https://ec.europa.eu/futurium/en/system/files/ged/a2-schweichhart-reference_architectural_model_industrie_4.0_rami_4.0.pdf