"CAR WASH FOR DATA": BEST PRACTICES FOR INFORMATION AND CONFIGURATION MANAGEMENT APPLIED AT PALLAS NUCLEAR FACILITY

LEO VAN RUIJVEN
Principal Systems Engineer
CROONWOLTERENDROS B.V.

JIM NOVACK
President
DYNATEC, S.A.

With respect to a nuclear facility, information and configuration management is important throughout its life cycle, but in practice appears to have insufficient maturity. Many organizations within the industry in general and specifically within the nuclear industry struggle with implementing a sound, maintainable and integral Information and Configuration Management process. The design is fragmented and based on tools from different vendors and different versions, even by discipline, and the lack of adequate methods, technologies, and tooling wastes money, reduces plant efficiency, hindering reuse of design knowledge and can even cause safety issues. This paper presents a pragmatic and proven implementation of information and configuration management in the Pallas nuclear facility in the Netherlands that elegantly meets the needs of information and configuration management related to the equilibrium triangle specified by the International Atomic Energy Agency (IAEA). The result forms a solid starting point for developing a digital twin as well as unlocking and reusing design knowledge.

INTRODUCTION

The Pallas organisation is preparing the replacement of the ageing High Flux Reactor (HFR), as shown below, which produces medical isotopes in the Netherlands. Pallas aims to maintain and ensure the integrity and validity of design basis knowledge and information over time to support asset management including a safe and efficient operation. Pallas has opted for a linked data and graph database technology to realize a Common Data Environment (CDE) to classify, harmonize and integrate all relevant data from the fragmented tool set used over the life cycle, based upon ISO 15926-11 and a Reference Data Library (RDL). In this paper the solution will be presented which was developed to organize a set of related organizational structures (i.e., System-, Work- and Geographical Breakdown Structures) seamlessly merging the engineering environment supported by an Integrated Information Model (IIM) based on ISO 15926-11 and a RDL based on ISO 15926-4, the Capital Facilities Information Handover Specification (CFIHOS), Industrial Foundation Classes (IFC) and private classes in the context of the nuclear domain. A novel workflow was developed to clean and classify all relevant data derived from the engineering tools including 2D and 3D modelling software. This data-cleaning process, like a "***car wash for data***", is an essential part of the integral concept to arrive at a sound Common Data Environment (CDE) as defined in ISO 19650. The result is an effective hybrid solution with respect to traditional documents and data, with regular handovers of design and engineering results from the CDE to the Client, using neutral, structured data (complying with the IIM and RDL) and linked to additional content described in traditional documents. The result is the secure and free flow of reliable and updated information to all stakeholders of the facility.

THE CHALLENGE OF CONFIGURATION MANAGEMENT

From the perspective of an owner-operator of a nuclear installation, it is essential to always have reliable data, both current and historical, over the entire life cycle of the nuclear facility. To be more specific with respect to data in this context one can one can distinguish the next three categories of data that unambiguously represent:

- The approved '***Why***' information of the nuclear facility. Especially the capabilities 'as specified' of the nuclear facility should have a central emphasis. Capturing and assuring the design base of the facility (including stakeholder requirements, functions, system breakdown structure and mutual relationships) representing the 'Why'.

- The '***How***' information from the engineering phase, connecting traceably to the 'Why' information of the nuclear facility. Also, the internal interfaces between the systems that compose the facility need to be explicit and fully documented such as the FMECA and fault-tree analysis of all systems and their integrated analysis (including e.g., functions and related SSC's, failure modes and measures). The capabilities of the reactor should be unambiguously qualified and quantified both '*as designed*' and '*as commissioned*'.
- The '***What***' of the system, connected in a traceable manner to both the 'Why' and the 'How' resulting in integrated operator manuals and maintenance manuals, based on the PFD's, P&ID's and FMECA analysis concerning the 'How' of the Pallas Reactor to be able to monitor the performance of the Pallas Reactor and be able to monitor the capabilities 'as in use'. This should support decision making concerning maintenance and replacement programmes executing the maintenance program derived from and traceable to the FMECA and asset management plans. Furthermore, capturing all maintenance activities, including results and findings, representing the What, and enabling the analysis of measurements obtained from operations and maintenance to adjust and improve the configuration and safety of the nuclear facility.

This view on data can be seen as both to ensure the possession of the license to operate and to be able to perform effective asset management. Sound asset management requires an integrated long-term view of several topics across the full plant lifecycle that have historically been largely managed separately such as (but not limited to):

- Stakeholder requirements
- Engineering data
- Maintenance and inspections, including surveillance and periodic testing
- Management of ageing and obsolescence
- Performance monitoring and feedback of operating experience
- Management of modifications

In this light it is essential to maintain and ensure the integrity and validity of design basis knowledge and information generated over time to support the safe and efficient operation of nuclear facilities, support effective decision-making, and mitigate the risk of knowledge and information loss. Decision making in this context concerns decisions at technical facility level as well as business processes. Therefore, the scope of life cycle information management covers not only design basis knowledge but also the engineering and physical construction of the various assets composing the facility, including their redesign and replacement. What has been described previously can be fully understood as configuration management of the nuclear facility, whereas information management can be seen on one hand as the process to retrieve information from the various sources and the delivery of required information to end-users and stakeholders on the other hand. IAEA expresses the essence of configuration management by means of the so-called ***equilibrium triangle*** as shown in Figure 1.

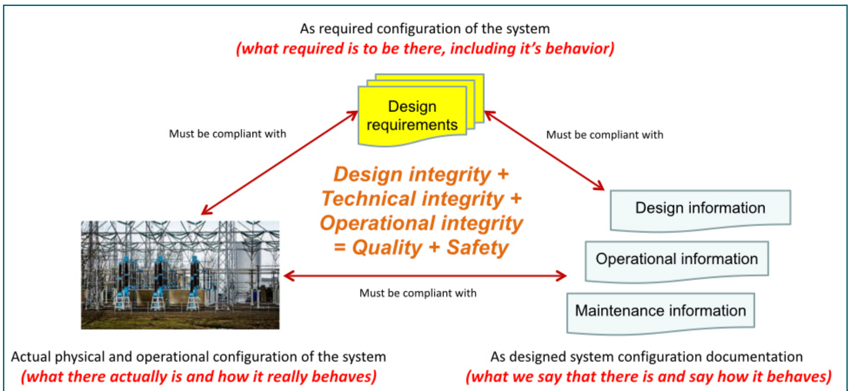

Figure 1. Equilibrium Triangle representing the challenge of configuration management.

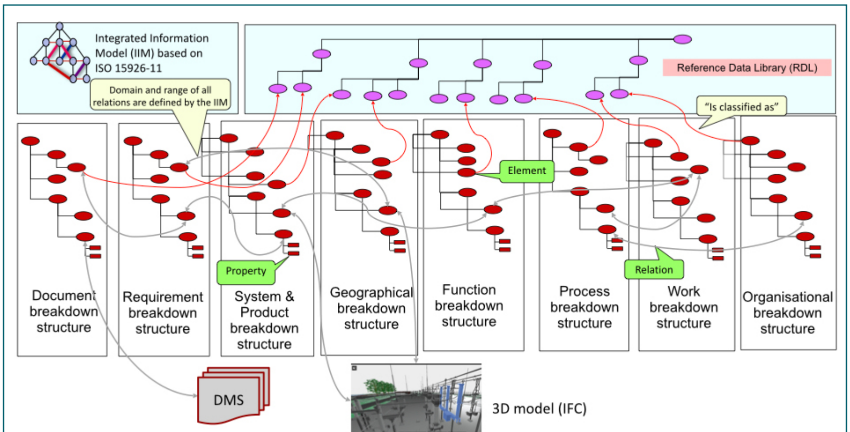

Figure 2. Breakdown structures, used to organize the data in the CDE and built with relationships defined in the IIM.

However, over the life cycle of a facility, in general practice, multiple information systems and databases from different vendors are used for different purposes. Most of these systems are not integrated with one another and cannot easily share plant data during different phases of the plant life cycle, such as design, operation, and decommissioning. This results in redundancies in capturing, handling, transferring, maintaining, and preserving facility data. This lack of interoperability stems from the fragmented nature of the construction and building industry, paper-based document control systems, a lack of standardization and inconsistent technology adoption among stakeholders. The end result is the inability to guarantee the integrity as presented in Figure 1 and wastes money, reduces overall plant efficiency significant and may even raise safety concerns.

An important principle within configuration management is to structure and classify all data using various breakdown structures as described in e.g., ISO 18346 (Structuring principles and reference designations). This is shown in Figure 2 where the composing elements of the various breakdown struc-

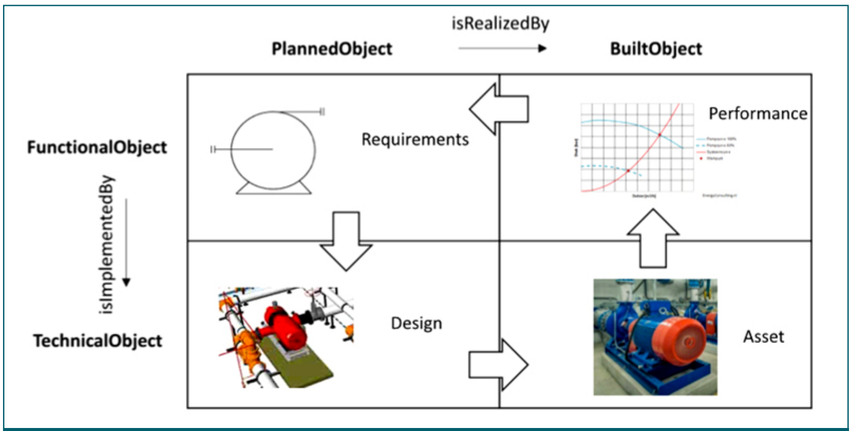

Figure 3. Life cycle model of facility elements, split into functional and technical and planned and as-built

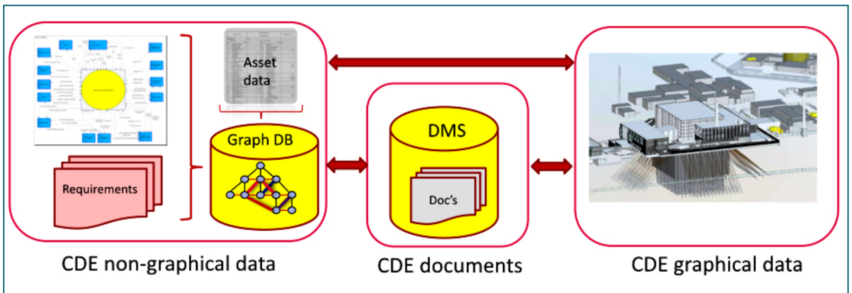

Figure 4. The three parts of the CDE, integrated by means of common identifiers.

tures are all in fact Configuration Items (CI) which are defined by their classifications, characteristics and relationships with other CI's. CI's can have relationships with one or more documents stored in a document management system (DMS) and one or more 3D objects in the integrated 3D model, representing the facility.

Specifying the composing elements of the System Breakdown Structure (including the seamless transition to and integration with the Product Breakdown Structure), complies with the life cycle model as shown by Figure 3. Respecting the life cycle model is extremely important for supporting asset management and enabling Digital Twin functionality by other applications that connect to the CDE. This life cycle model is fully supported by the relationships in the IIM and supports for example the traceability of which assets have been implemented from a planned technical object and how these different assets have functionally performed as a realization of the planned object. This is important for asset management and to be able to operate Digital Twins (for prediction of the effect of changes in configuration).

The life cycle shown in Figure 3 recognizes the planned functional object, its implementation by a design, the realization of the design by a built asset (mostly an instance of a manufacturer model) and finally the planned functional object as built. The latter represents the real performance of the functional object with respect how it was planned.

THE COMMON DATA ENVIRONMENT CONCEPT

Despite that the landscape of software tools during the project in practice will be fragmented and there will be a lack of interoperability between those tools, Pallas aims to maintain and ensure the integrity and validity of design basis knowledge and information over time to support asset management including a safe and efficient operation. Therefore, Pallas has opted for a linked data and graph database technology approach using the Resource Description Framework Schema (RDFS) from the World Wide Web Consortium (W3C) to realize a CDE, based on ISO 15926-11.

Intentionally, only use of RDFS was chosen despite the available OWL extension on top of it. OWL appeared to complicate the whole matter significantly and does not outweigh the alleged benefits, and there is no off-the-shelf software available that supports this. The concept of the CDE itself was taken from ISO 19650 which is split for this project into three parts: a Document Management System (DMS), a geographical environment by means of a 3D model collecting platform (BIM 360) and the non-graphical data environment (Graph Database). Within the latter part of the CDE all relevant data from the fragmented software tool set used over the life cycle will be classified, harmonized, and integrated based on a common integrated information model (IIM) and a Reference Data Library (RDL). For creating the IIM, ISO 15926-11 is used (also based on RDFS) and the RDL is based upon ISO 15926-4, the Capital Facilities Information Handover Specification (CFIHOS) and Industrial Foundation Classes version 4 (IFC4). Recognising and respecting that each involved (sub)contractors has their own information management maturity and ideas about a roadmap to evolve from a document centric working company to a company that is capable to work in a data centric manner, the next principles and or rules have been used to validate the CDE concept:

- Make is as simple as possible but not simpler (Free to Einstein).
- Use international standards and dare to make simplifications to them.
- Offer a Single Source of Truth (SSOT) to ensure all parties involved base business decisions on the same, valid and consistent data.
- Distinguish and separate geographical related data, non-geographical data and documents.
- Don't force involved companies to change their work process in order to deliver their results.
- Accept the data created by the tools as is and focus on cleaning that data.

The idea behind this is that companies will gradually see the benefits of harmonized and integrated data of the whole facility, which during the project will probably lead to a higher maturity with regards to information management in general.

Breakdown Structures themselves as shown in Figure 2 are as much possible derived from the engineering tools (so defined by the engineers) but are made consistent and harmonised in the CDE and are supported by the IIM. Any breakdown structure can be applied in the CDE when useful, enabling specific aspect views (as long the aspects are modelled as such).

A novel workflow was developed to clean and classify all relevant data derived from the engineering tools including the output of 2D and 3D modelling software. This data-cleaning process is an essential part of the integral concept to arrive at a sound Common Data Environment (CDE) as defined in ISO 19650. The result is an effective hybrid solution (respecting data, documents, and 3D models), with regular handovers of design and engineering results from the CDE to the stakeholders including the Client, using neutral, structured data (complying with the IIM and RDL) and linked to additional content described in traditional documents and objects in the 3D models. The result is the secure and free flow of reliable and updated information to all stakeholders of the facility such as:

- all technical information that describes the requirements, the design and as-built of the facility,
- supporting Configuration Management by offering a Single Source of Truth (SSoT) to ensure that all parties involved base business decisions on the same and consistent data,
- technical information that enables the confirmation what is required to be there, that it conforms to what is stated in the documentation and conforms to what really is there.

The CDE including the DMS, 2D and 3D Models can be seen as a Plant Information Model which forms the basis for Asset Management Business Intelligence Solutions using the compendium of documents, data and 3D models and their mutual relationships (captured by the CDE). This should describe in an unambiguous way the requirements, design and construction of assets and must be keep updated and in sync during the operational lifetime and eventual decommissioning which is the main task of configuration management. This is especially important in the light that increasingly, the management of nuclear facilities has the following vision regarding asset management and digitalization, derived from their business model:

- management of facilities has a responsibility to its community to effectively and efficiently manage the safety and services provided by all its assets,
- asset management should support sustainable and flexible services delivery, community satisfaction, sustainable financial position and acceptable risk exposure.

THE ROLE OF INFORMATION MANAGEMENT

As stated before, information management is in the context of this paper seen on one hand as the process to retrieve information from the various sources and on the other hand the delivery of required information to end-users. With respect to the first part a method is developed to capture and integrate the output data of the various engineering tools in use. In general, engineering tools all have their own core model and data export functionality and use a tool-specific set of terms and definitions of these terms. This leads to multiple data streams each having their own structure, syntax and semantics to be interpreted by the CDE. Therefore, all data streams are treated by a cleaning process in which a seamless integration of all data streams is performed by applying a common language defined by the IIM and RDL. When processing these data streams to integrate the data in the CDE, data quality principles are applied as defined in ISO 8000 (Data Quality). The main data quality characteristics are defined in ISO 8000 as semantic quality, syntax quality and pragmatic quality. Figure 5 shows the principle of receiving baselined and signed datasets as they are received from the engineering environment passing through the cleaning process. In this cleaning process, syntax is checked and eventually corrected, the meaning of the data (semantics) is interpreted and agreed with the origin of the data and all data is classified according the RDL. All successfully imported data files are stored in the CDE on a file server as a back log. Each data element in the CDE has a signature (set of meta data) that refers to its original import file and containing knowledge such as

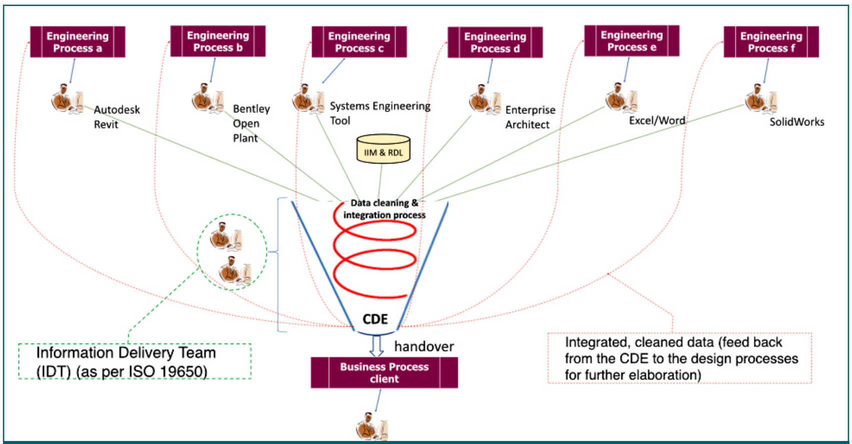

Figure 5. The principle of cleaning, harmonizing and integrating data streams based on a common language formed by the IIM and RDL.

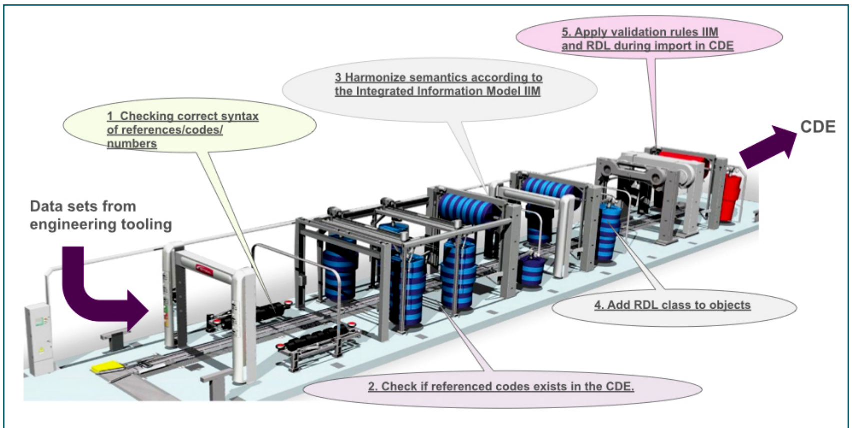

Figure 6. Data Car Wash which applies the data quality principles of ISO 8000.

when that data element was imported, which software tool was the origin and if it was a first creation, change or soft deletion. This all leads to a full “bookkeeping” functionality with respect to all imported data, necessary for supporting provenance, integrity, consistency checks and audit trails, similar to systems used in professional accounting software. In fact, it also supports basic functionality of block chain technology for greater security.

In this cleaning process, represented by Figure 6, syntax is checked and eventually corrected as far as possible, the meaning of the data (semantics) is interpreted and modelled according to the IIM and all data is classified according the RDL. If things cannot be done or completed due to ambiguousness, the data file is rejected and discussed with the originator. When this step has completed, the file will be imported in the CDE and references to documents and codifications like TAG code, location codes are checked whether they exist in the CDE when the data expects that they should be in the CDE, otherwise the import will be rejected and again checked with the originator of the data.

Also, the modelled data according to the IIM will be validated if the data modelling was done compliant to the domain and rang of the applicable relationships forming the IIM. This all is represented in Figure 6 by means of a “car wash for data”.

The CDE supports additions, changes (by means of explicit replace relations) and soft deletions of data to manage baselines and reporting of differences between any baseline and another baseline. This is made possible by adoption of the W3C resource description framework (RDF) standard that formalize “triples”. A triple is a linking structure that forms a directed, labelled graph, where the edges represent the named link between two resources represented by the graph nodes. Each triple is uniquely identified through a Unique Reference Identifier (URI) by means of a rdf:statement which enables to make statement about statements, necessary for configuration management. Each rdf:statement is also provided with a signature for traceability of its origin. In fact, the signature mechanism covers some block chain characteristics in a practical way. The combination of integration of all relevant facility life cycle data (multi-disciplinary, multi life cycle stages), strict use of a common language (IIM and RDL) and configuration management capabilities guarantees a neutral handover of any selection from the CDE content. The hand over architecture is shown in Figure 7.

The result of using a common language based on an integrated information model and RDL based on RDF triples is shown in Figure 8. In this figure, the central node is an instance of a document type, is the output of a work package, has as subject the Core structure, has two versions, is referenced in four other documents, has a unique identification, and has a signature. By deriving more and more explicit meta data and relations (according to the IIM) from traditional documents one can gradually move from a document-based way of working to a data-centric way of working.

EXTENDING THE CDE WITH DIGITAL TWIN CAPABILITIES

Since the CDE represents the complete configuration over the life cycle of the facility over time, it makes the CDE a perfect source for realizing a Digital Twin to support better operations and safety. In Figure 9 the Digital Twin concept realised by connecting a Computerized Maintenance Management System (CMMS), Common Data Environment (CDE) and Artificial Intelligence (AI), modelling and videogaming technology to really give substance to the principle of the equilibrium triangle as presented at the start of this paper. By collecting and linking dynamic data (based on IoT) to the breakdown structure elements in the CDE, a Digital Twin can be realised which is consistent with the configuration as defined in the CDE and reflecting the real configuration and performance of the facility in real time.

The potential of an approach as described in this paper does not only enable unambiguous communication in a new construction project but is extremely valuable when digitizing existing facilities to unlock the knowledge as contained in the design and to be able to realize more efficient asset management and operations. The CDE feeds the digital twin environment with a comprehensive configuration of the facility as well as harmonised calculation codes as a feed for simulators. These calculation code models can cover various aspects from neutronic

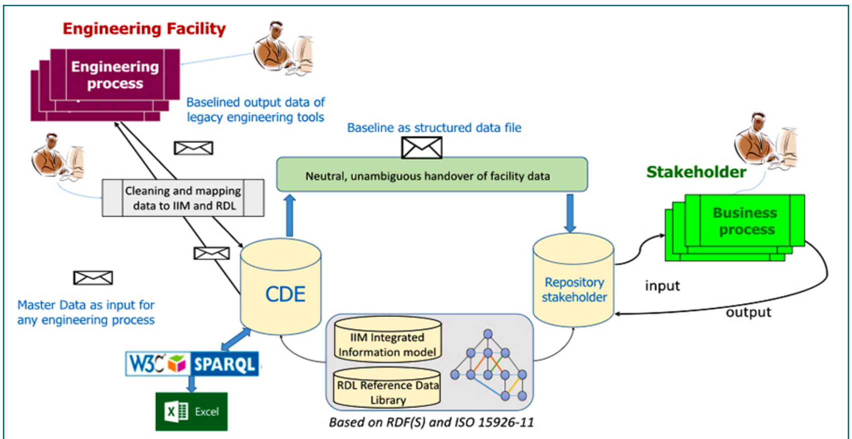

Figure 7. Data handover architecture.

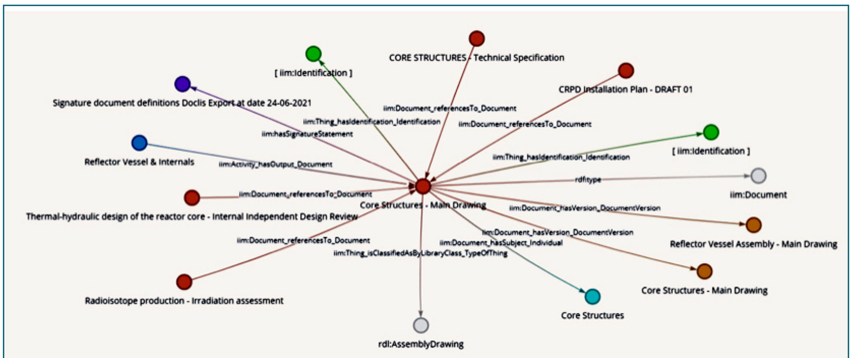

Figure 8. Example of a graph view on the content of the CDE.

aspects of the reactor core to thermal-hydraulics or chemistry. Some codes focus specifically on a particular component and its operation, while others deal with system-wide phenomena. The IIM and RDL both play a key role in ensuring the interoperability of all the various calculation codes. The application of open standards as applied in the CDE concept and enabling the integration of additional data sources and inviting third party extensions to the platform allows the creation of a fully explorable Digital Twin Simulator, even linked to Project Lifecycle Management (PLM) and other platforms. Companies in the nuclear industry in France like CEA's nuclear energy department, EDF and Framatome believe that digital twins will help in the process to continuous improvement of safety and quality of operations and will allow operators of nuclear facilities to better validate their action strategies in case of an event. The digital reactor is an important tool for the simulation to demonstrate the safety of equipment in a range of operating scenarios. With the collaboration of various originators of calculation codes, progress has been made in France realizing Digital Twins. As a concrete example, EDF has already trialled virtual reality devices to simulate opening or closing valves in the reactor building.

Pulling from a harmonized and verified data base stored in international standards, integrating IoT sensor data and merging this with the advanced visualization software of current videogames, it is very likely that future operators will manage complex industrial and nuclear facilities using immersive, real-time simulators, such as those being developed and implemented by Dynatec's Talent Swarm division as shown below in Figure 10.

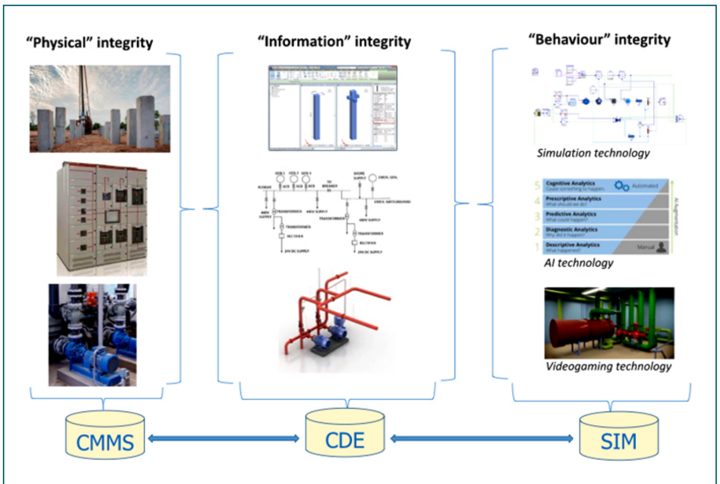

Figure 9. Digital Twin approach with a central CDE.

Figure 10. Rendering of Immersive Digital Twin Simulator (www.TalentSwarm.com).

CONCLUSIONS

Our parents always told us that key to an organized home was, "*a place for everything, and everything in its place.*" Similarly, the veritable Tower of Babel that is the collection of technical documentation that may accumulate during the decades of a facility´s operation screams for a well-thought system for arranging, validating, storing, sharing and displaying information.

Based on the standards as mentioned a knowledge-centric facility information model is developed and leveraged as a modern and efficient approach to better support, manage and enable seamless sharing, transfer, and use of sustainable design knowledge and information within and across all facility life cycle phases. The CDE is the core for creating an integrated foundation for making process-related, engineering and management decisions during plant operation and decommissioning. Data consolidation in a (virtual) single information storage increases efficiency and ensures transparency and safety of plant operation. The vision is that this only can be achieved by applying semantic modelling and linking technology and the use of data integration standards.

Essential in the CDE approach as described is the cleaning process of data retrieved from the varied landscape of software tools which require thorough "washing" of the data exported by the various software tools and then harmonizing and classifying it in accordance with the integrated information model and Reference Data Library. The "carwash for data" requires project roles of people with domain knowledge of engineering and nuclear facilities together with IT and semantic modelling skills. This best practice experience has shown that extending RDFS with OWL for example complicates the whole process significantly, such that at this moment in time this additional effort does not outweigh the alleged benefits. In fact, there is also no off-the-shelf software available that supports this. Also, sticking to RDFS rather than making extensive use of OWL makes the use of traditional IT knowledge feasible and increases the support of engineers which is essential to enlist and gain the support of the necessary domain knowledge. Finally, the good news appears to be that nuclear energy is now shedding its negative image due to the CO_2 challenge and the rising energy costs in Europe, so this may be a fantastic opportunity to digitize older facilities and attract and recruit new technology-savvy personnel to the nuclear industry.

REFERENCES

– Ruijven L. van; Ontology for Systems Engineering as a base for MBSE; 25th Annual INCOSE International Symposium Seattle USA 2015
– Ruijven L. van; Ontology based exchange mechanism for Systems Engineering; Fifth International Symposium on Life -Cycle Civil Engineering, IALCCE2016; Delft University 2016. ■