# Semantic Interoperability in Power Companies

## An ontology-based asset dossier solution

Ismael Ávila
Diretoria de Soluções Tecnológicas e Consultoria
CPQD Research Center
Campinas, Brazil
avila_an@cpqd.com.br

Carlos Alberto de Sousa
Diretoria Adjunta de Estratégia, Meio Ambiente e Inovação
Companhia Energética de Minas Gerais - Cemig
Belo Horizonte, Brazil
casal@cemig.com.br

***Abstract*— One of the key competences in effective corporate management is the company's maturity level in asset management, which, largely, depends on the asset information management. In power companies, however, the diversity of asset types and the variety of the information required or produced in every phase of those assets life cycles has proven to be a big challenge for asset information management as this problem intrinsically relates to the information silos that characterize such large corporations. In this paper, we discuss the conception of an asset dossier solution aimed at semantic and chronological integration of asset information, throughout an asset's life cycle, by means of an ontology-driven semantic bus. We present here the proposed semantic asset dossier architecture and discuss a participatory evaluation process to assess its validity from the perspective of some key stakeholders.**

***Keywords*— *semantic interoperability; knowledge acquisition; knowledge representation; ontology; asset management***

## I. INTRODUCTION

Power companies operate in a highly regulated sector, in which a significant part of their revenues comes from the recognition, by the regulatory body, of the investments they make in their asset base. Therefore, an effective management of assets and their information is a regulatory requirement and the National Electric Energy Agency (Aneel) periodically audits the power companies to ensure that it is met.

The asset base (AB) of a power company comprises a wide range of equipment (transformers, switches, panels, regulators, etc.), structures (poles and towers), in addition to buildings and immaterial assets (patents, brands, etc.). The asset's life cycle undergoes phases such as design, acquisition, construction or installation, testing, commissioning, unitization, inspections, maintenances, de-commissioning, replacement and disposal.

Throughout these phases, information about the assets are created and gathered by different corporate areas, in often non-interoperable computer systems. This process consumes and generates different types of documents, which can make use of terminologies specific to each area of the company, in what one commonly defines as information silos.

Thus, in what refers to the document collection of an asset, this implies incompatibilities, redundancies and situations of inconsistency. In addition to hindering good asset management practices, a scattered and incompatible document base can also cause problems when, for example, the company is unable to retrieve all the information requested by auditors during a regulatory audit or the auditors find inconsistencies within the presented documentation.

In recent years, errors in the information provided to the regulatory body resulted in fines surpassing tens of millions of Brazilian Reals. In many cases, the companies were unable to locate the information requested by the regulatory body during the periodic audits. This stresses the importance of eliminating information silos, creating standardized and consistent access to information throughout the asset's lifecycle. This is also a requisite for a mature and effective asset management (AM).

In face of this problem, we here describe the conception and evaluation of an asset dossier solution that bridges some of the semantic gaps between informational silos in large power companies in order to streamline the asset management practices as well as eliminate or minimize regulatory sanctions due to information inconsistencies. A Knowledge Base (KB) of asset types and corresponding business processes combines expert knowledge to other sources of information in order to build an interoperable layer, a semantic bus among and above the relational databases currently in use in the company.

## II. MAIN CONCEPTS AND CONTEXT OF THE PROBLEM

### *A. Asset management versus information silos*

Asset management is an essential part of the organizational management and is endorsed by International Organization for Standardization (ISO) in a family of standards that "reflect the emerging international consensus on what is necessary to ensure the management of competent, integrated and sustainable asset life cycle" [1].

A company's AM maturity level is a standardized metric by the Institute of Asset Management (IAM), one of the leading authorities on the subject. The IAM recognizes six maturity levels that range from innocence to excellence, passing through the awareness, development, competence and optimization phases [2]. Within the 39 themes of the "Asset Management Horizon", the IAM included four directly related to asset information, in addition to 11 themes related to asset life cycle.

One AM challenge in power companies is to transform dispersed data, sometimes in incompatible formats, into useful

knowledge for the asset micromanagement and, above all, for the strategic management of the investments in new assets.

Thus, the study sought to conceive and prototype an Asset Dossier (AD) solution to organize both chronologically and semantically the documents and information for every asset type, in order to reduce information silos and discrepancies that could cause regulatory non-conformities as well as increasing the usability by centralizing access to and searches for the information. The improved usability comes from organizing the information according to each asset's life cycle phase and the corresponding processes. Finally, the AD should also consolidate the retrieved information, generating knowledge.

### *B. Understanding the needs of the main stakeholders*

In order to understand the needs and points of view of the various stakeholders of asset management in the context of the company's information silos and given the current regulation, several research activities were undertaken, according to the following methodological chain:

- Interviews with internal stakeholders: we conducted 11 interview sessions, involving 15 employees, 7 of which holding a leadership role. The interviews covered the areas of compliance and risk, asset management and control, engineering, cost management, management and internal controls, IT, planning and operation management, maintenance planning and engineering, economic and technical/commercial regulations;
- Design sprint session with stakeholders to identify pains, unmet needs and to ideate features for the AD;
- Benchmarking in power companies with greater AM maturity: benchmark visits to three power companies taught us their best AM practices and strategies.

## III. Conception of the AD Architecture

The initial activities highlighted the requirements of the AD solution. The initial generic idea of an AD evolved to a more systemic and integrated view, which included:

1. A knowledge base (KB), such as a domain ontology, to promote the intended semantic interoperability between the various databases and information systems of the company;

2. A way to combine knowledge about asset types with information about individual assets (instances of those types);

3. An environment to be used by the company's employees to facilitate the collaborative acquisition of expert knowledge;

4. A mechanism for rewarding contributions (knowledge) made by the company's employees to the KB.

### *A. Development of the four main attributes of the AD:*

The concepts and ideas studied and discussed in the previous steps were implemented as software components and combined according to the schematic representation in Figure 1.

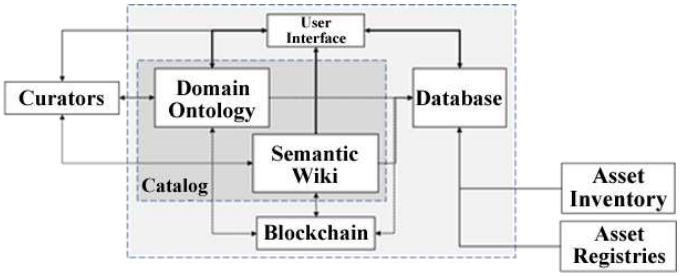


Fig. 1. Proposed architecture for the asset dossier solution

The AD therefore comprises:

• A domain ontology, which organizes the knowledge already accumulated and curated (about the assets) into a formal representation, readable by machine (and by people familiar with the ontological representation). It enables the automation of queries, logical reasoning and inferences and, to act as a semantic interface between different corporate systems, it was conceived and built following the guidelines of the main top ontologies BFO – (Basic Formal Ontology) [3] and IOF – (Industrial Ontologies Foundry). As a result, the AD can interoperate with other ontologies built according to the same guidelines;

• A semantic wiki [4], which gathers knowledge about assets (relationships, attributes and meanings), but in a format that is less formal than that of the ontology, in order to be readable by ordinary users (and machines). Hence, it is a gateway to new knowledge, brought by the user community. Knowledge is organized, matured and curated in the wiki before entering the ontology. Thus, the functions of the semantic wiki include:

- Organizing knowledge about assets in a friendly way, accessible to common users;
- Offering semantic query resources;
- Giving access to information about assets categories;
- Allowing users to input knowledge about assets, asset processes, etc.
- Defining the correct terminology to be used in corporate systems and documents;
- Containing glossaries/taxonomies of assets and systems;
- Describing business processes involving assets;
- Indicating and referencing procedures/manuals related to assets;
- Listing the best asset management practices;
- Allowing the users to share and collaboratively enrich the knowledge base;
- Consolidating and maturing every knowledge before incorporating it into the ontology;
- Encouraging the employees to participate in the use and evolution of the corporate knowledge.

• The blockchain application is responsible for registering events of KB usage, including contributions made to it, and running a code to assign values to them according to their relevance;

• The relational database (DB) is the KB's connecting point with other corporate systems that contain information on individual assets, such as inventory data and registry of events with each asset. With this integration, the received data become available for queries from the AD user interface;

• The user interface provides navigation and query options supported by KB and the DB so as to combine consolidated knowledge about general asset types (categories) with up-to-date information about specific individual assets.

In combination, the main components of the AD provide some functionalities and enable typical use cases. Therefore, in order to enable a rich proof of concept (PoC), we developed a quite complete version of the AD. So, for example:

• The KB, whose two main components are the domain ontology and the semantic wiki, covered the 20 most relevant types of asset, therefore representing a large proportion of the investments that the company makes in its grid;

• While developing the ontology, those 20 prioritized assets were broken down into more than two hundred terms, which then had to be anchored and correlated with one another;

• The semantic wiki imported those basic terms from the ontology as an initial load (then originating 227 root pages);

• The wiki pages were then populated with content from official regulatory documentation that describe and categorize the assets from the regulator's perspective;

• In combination, the ontology and the semantic wiki form a sort of catalog of the company's asset types, defining what characterizes the assets in terms of hierarchical relations, similarities, functionalities, etc.;

• The user interface were organized to reflect the prioritized phases of the asset's life cycle and to provide fields for searching the KB and the DB;

•, Two external DBs were populated with data about individual assets and their corresponding documents in order to simulate data from an asset inventory, as exemplified in Table 1;

• In order to exercise the role of the AD as a semantic bus, the data populated in the two external DBs were incompatible with one another (had different formats) as discussed below;

• Finally, the blockchain application for registering and evaluating the AD usage was developed and integrated into the semantic wiki, focusing the contributions made by users to two specific fields.

TABLE I. DOCUMENTS POPULATED IN THE EXTERNAL DATABASES

| *Asset life cycle phases* | *Documents* |
|---|---|
| Project and design | Configuration, equipment and costs |
| Acquisition | Public call and technical specs |
| Construction | As built document |
| Unitization | Original Book Value |
| Operation/Maintenance | Register of maintenances, refurbishments, faults or retrofits |

### *B. AD integration and Proof of Concept*

Once the development of each component of the AD was completed, we proceeded with their integration for the PoC. The domain ontology was connected (via internal DB) to the external relational DBs, created to simulate typical information silos of a power company: a number of non-interoperable computational systems. In this sense, the used DBs had distinct data structures and information terminologically reciprocally inconsistent. Despite such intrinsic incongruences, a query mediated by the ontology should retrieved and combine the contents semantically in a complementary and consistent way.

The integration process began with the installation of the "Ontop" plugin in the Protégé ontology editor. This made it possible to map the entities in the ontology to the information in the DBs. This was done using SPARQL (Protocol and RDF Query Language), the standard language for querying RDF (Resource Description Framework) graphs.

Thus, once a SPARQL query is created and executed, it would return results in the form of a combination of the three sources: ontology, DB1 and DB2. However, to expand the integration beyond the scope of the Protégé tool, making the queries to the KB accessible to external applications, an "endpoint" had to be created. With this aim, we used a plugin of Protégé, named Ontop, which has a Java endpoint that allows exporting the results of the integrative query. In this way, any external application would be able to make a query via the endpoint and obtain, as a result, the aforementioned combination of knowledge mapped in the ontology with data from the two relational DBs. This allowed the connection of the user interface (a web application) to the KB and the successful execution of several query types.

The PoC involved five stakeholders, with different profiles and experiences, as well as representing different areas of the company, as shown in TABLE II.

TABLE II. PARTICIPANTS OF THE POC

| *#* | *Area* | *Gender* | *Company time* | *Manager* |
|---|---|---|---|---|
| 1 | Engineering | M | 13 years | No |
| 2 | Accounting | F | 21 years | Yes |
| 3 | Management control | F | 14 years | Yes |
| 4 | IT | M | 33 years | Yes |
| 5 | IT | M | 14 years | No |

In order to bring the AD (and its potential use) closer to the daily reality of the PoC participants, the session began with a description of a scenario that, although hypothetical, was based on reports of real situations and was considered plausible:

> “*The regulating agency requests, with great urgency, information about some of the company’s assets. A working group is created to answer the request and you are assigned as the representative of your area. The WG is coordinated by a director. S/he requests a preliminary inventory of assets on his/her desk within 30 minutes. Unfortunately, the information requested is dispersed in several databases, in different formats and terminologies not always consistent with one another.*”

The objective of this initial contextualization was to define a background against which the merits and limitations of the AD could be evaluated and highlighted by the participants.

Then, the use of the AD, its KB and some query cases were presented in individual sessions to each of the five participants.

We then presented the semantic wiki to each participant, emphasizing its resources and contents, such as the asset pages, the semantic searches and semantic representation of an asset management process in Business Process Model and Notation.

Finally, we showed the blockchain reward mechanism for the contributions to KB according with their relevance.

After these presentations, we asked each of the stakeholders about the AD solution and its presented features and they all acknowledged the usefulness and the importance of such an AD and its KB, stressing the urging need for effective means to mitigate the information silos in the company. The participants also praised the use of a semantic wiki as a valuable channel for a collaborative development of the corporate KB.

The participants also approved the gamification aspect of the blockchain application for registering and evaluating the use of the AD and the inputs to the KB as something consistent with the objective of creating greater employee engagement. The transparency and auditability that the blockchain creates in this reward mechanism was also considered very suitable by all the participants.

Overall, the PoC results suggest that, with the proposed configuration and functionalities, the AD solution for semantic interoperability fulfills many of the current yearnings of the company's employees, moving towards a unifying enterprise solution. But possible improvements, extensions or evolutions of the AD were also identified in the PoC. They aim to cover a greater number of use cases, a greater variety of assets and other entities in addition to the assets themselves. Besides these justifiable improvements, it is important to highlight some others betterments that future developments of the study could address, namely:

• The extension of the ontology to incorporate a significantly larger set of terms, as a way to test the scalability of what is being validated on a medium scale in this study;

• Expansion of the PoC to cover other parts of the assets' life cycle in addition to the scope adopted in this study;

• The extension of the wiki's content in order to include the collaborative insertion of procedures and processes by the employees themselves, in order to contribute to this dimension of corporate knowledge;

• The extension of registration, valuation and recognition mechanisms for the contributions made to the KB, with the evaluation of new resources, such as bids, etc.;

• The evaluation of other semantic wiki solutions to see if any alternative would bring gains that justify its adoption;

• Evaluating other uses of blockchain, such as corporate digital currencies to reward contributions to the KB;

• Creation of an AD for the entire power distribution sector, comprising an ontology, a wiki and, mainly, a token that would motivate entities and individuals to contribute to the knowledge base of the Brazilian electricity sector;

• Integration with other existing ontologies, in order to expand the possibilities of interoperability between systems and machines.

## IV. Conclusions

The conception, development, semantic underpinning and integration of the AD achieved most of the objectives of the study, even though, as already mentioned, it was necessary to delimit a scope with the 20 most significant assets in terms of the company’s investments.

The participatory practices of identifying the real needs and expectations of the main stakeholders proved to be very important not only for defining the priority scope, but, above all, for the conception of an architecture with systemic value. The AD components also benefited from the contributions from the main stakeholders, either in the form of technical and procedural knowledge, either use cases or expectations.

With regard to the results of the PoC, it made it possible to assess what was considered most important: the perceived value of the AD. In this sense, there was unanimity among the participants in the understanding that the AD, in the way it was conceived and implemented, is justifiable if the background presented is considered, which, by the way, all the participants considered plausible and even frequent.

Participants also agreed on recognizing the value and relevance of all elements of the AD, although they were also able to see desirable advances in some of them, such as enriching the wiki with information from other sources.

Although these results are qualitative in nature, it should be noted that the five PoC participants have a total of 95 years of accumulated experience, and in different areas of the company. In this sense, they demonstrate very mature and refined views of the challenges and particularities of the company with regard to the management of its assets, in the different phases of its life cycle.

It is therefore possible to conclude that the study fulfilled its central objectives of designing an innovative AD proposal (with semantic integration) and implementing a preliminary version (prototype) of this concept in order to submit it to a

systemic and qualified validation from some of its key stakeholders.

It appears from the results that the AD demonstrates potential and merits that justify future developments in this study.

## ACKNOWLEDGMENT

This study was supported by the R&D Program of the National Electric Energy Agency (Aneel).

## REFERENCES

[1] ISO 55000:2014 Gestão de Ativos – Visão geral, princípios e terminologia.

[2] John Woodhouse Editor (Ed.). 2015. IAM – The Institute of Asset Management. Guia e Escala de Maturidade em Gestão de Ativos. Versão 1.0 junho de 2015. Disponível em: www.theIAM.org.

[3] ISO/IEC 21838-2: Information technology — Top-level ontologies (TLO) — Part 2: Basic Formal Ontology (BFO).

[4] Semantic MediaWiki: https://www.semantic-mediawiki.org/