IIoT TOOLS AND STRATEGIES FOR SYSTEMS INTEROPERABILITY AND FOR INCREASING THE INTELLIGENCE OF WATER INDUSTRY SOLUTIONS

PhD thesis – Summary
for obtaining the scientific title of PhD at
Polytechnic University of Timișoara
in the doctorate field of Systems Engineering
Author: inf. Andrei Mihai NICOLAE
Scientific leader: prof.univ.dr.ing. Ioan SILEA
TIMIȘOARA, September 2021

The thesis is structured in 9 chapters, the first of which having an introductory role, so that the second to describe the current state of research in the field of the thesis. The third chapter presents the research and development of an elementary software application, of historian type, destined for the water industry. Then, in chapter 4, a reference software architecture is exposed to elevate the software application to a higher, proactive level, which involves the autonomous, unassisted and non-invasive optimization of the monitored technical system. The next 4 chapters contain research, implementations and tests performed to obtain such a proactive tool, in accordance with the aforementioned architecture. The last chapter highlights the author's contributions and the future directions of research and development. The thesis graphics includes over 67 elements (figures, graphs, diagrams, tables) that illustrates the complexity of the topic, the test process, the solutions and the results.

1. Introduction

In a brief preface, the concept of Cyber Physical Systems (CPS) refers to a computerized system, within which a physical mechanism is controlled or supervised by software algorithms. The essential feature of this concept is defined by a symbiosis between the physical components and the cybernetic, digital elements. Against the background of a cumulation of recent developments, in directions such as communication technologies, but also of intelligent systems and sensors, this orientation is increasingly prevalent in the scientific world and in the applied research. In this regard, "doctrines" such as Industry 4.0 [1][2][3] or Industrial Internet of Things (IIoT) [4][5][6] only implement the ideas around CPS in the economic environment, of course, with specific customizations to different fields. Consequently, the research carried out as a foundation for the elaboration of the doctoral thesis was oriented in the proximity of the IIoT and Industry 4.0 paradigms, aiming, at the same time, at practically implemented components, with direct applicability in the water industry.

Regarding the water industry, one can observe a domination of the old, inherited technical systems. These were viable solutions at the time of their introduction, but the vast majority of those found today in the water industry are unoptimized and inefficient by modern standards, transforming this industry into the perfect environment to deliver the improvements promised by the Industry 4.0 and IIoT technologies. Moreover, the entire water sector faces extensive efficiency problems, one of the main causes being due to the fact that the current functional solutions, from the industrial “landscape”, consist of a wide variety of dispersed systems, both as a technical / technological / chronological level, as well as in terms of location. In addition, often the integrability and interoperability of local structures is obstructed by the basic interface with the physical equipment, which is achieved through proprietary protocols.

Thanks to recent advances in the area of connectivity, in the structure of IIoT, the emergence of the concept of data accumulation is signaled, interpreted in practice by historian type of applications, their objective crystallizing in the acquisition and storage of parameters and operating data from the supervised technical system. Although the data collected by the classical historian programs remain largely unused, they still open up new, unexplored opportunities for proactive historian software, where insights into stored data analysis algorithms and optimization strategies inaugurate perspectives on both distinguishing patterns as well as learning abilities or sound conclusions about real-world applicable efficiency improvements. *In reality, the latest trends are aimed at creating solutions that can use, in a skilled way, the already collected data, to solve certain specific problems*.

In the penciled context, the *research theme* is the development of both *tools* and *strategies* that can be successfully applied in industry, the *pursued goal* being to increase the interoperability of systems and the intelligence of solutions in the water industry. In terms of tools, the primary development of software applications that display the possibility of operating in a non-invasive manner, along with the solutions already implemented in the industry, was considered. Regarding strategies, the research theme aims at developing new algorithms and procedures to facilitate the creation of smarter software solutions, with potential to optimize various features of technical systems, in harmony with the ideology of IIoT and Industry 4.0.

Consequently, *the major objectives of the research theme* are: (i) investigate the field and the current stage of research in the proposed topic of the thesis; (ii) development of algorithms to optimize the performance of monitoring systems in the water industry; (iii) development and testing of an autonomous proactive type of historian application, capable of optimizing the supervised technical system from the water industry. However, the complexity of the third main objective requires the prior achievement of a series of subsequent objectives: (i) the preparation of an analysis of the requirements and performance of the systems from the water industry sector; (ii) identification and development of ways to store the values of the operating parameters of the systems related to the field; (iii) developing or improving methods for analyzing and identifying the dependencies between the parameters of the monitored system; (iv) developing or adapting algorithms for predicting the evolution of parameters, using the history of stored values and other information; (v) development of ways of recognizing recipes for optimization, applicable to the supervised system, in order to meet certain established optimization objectives; (vi) development of a reference software architecture and minimum hardware requirements for a proactive monitoring system, specific to the water industry; (vii) development and implementation of an experimental autonomous proactive historian.

Evaluating *the practical applicability targeted by the research results*, it is pronounced, the tools and strategies emerging under the umbrella of this thesis being validated by testing them on real systems. Specifically, the aim is to build an autonomous proactive historian software solution for the water industry, which has, in addition to the capabilities of storing the values of the operating parameters of the monitored technical systems, also capabilities to identify and understand the specific relationships between different parameters. These features facilitates for the application to implement the technical system optimization in a non-invasive manner. More precise, it is being developed an easy-to-use software tool, installed on an accessible hardware platform, located either physically close to the monitored system or remotely. In succession, after a short initial configuration period, the ensemble will be able to operate autonomously and automatically, providing, additionally to simple writing in a database of read values for various system parameters, also some optimizations, not requiring any human collaboration.

2. The current state of research in the field of thesis

The second chapter presents the current state of research in the field of interest, accounts for the main directions of development in Industry 4.0 and summarizes a review of the relevant literature in the context.

Of course, it is found that Industry 4.0 and IIoT are closely related, both being concerned with intelligent communication between different industrial entities, some researchers considering them as the same phenomenon, but manifested in slightly different domains of applicability.

Firstly, *the defining features of Industry 4.0* are: (i) the dissolution of the previous concept of the automation pyramid, replacing it with a structure within which entities can be directly connected to each other, thus allowing an entity to communicate with any other entity present; (ii) impacting multiple aspects of society (working conditions, new types of jobs, workforce qualification, education, quality of life), being already categorized as a true large-scale Industrial Revolution; (iii) more personalized products; (iv) automatic exchange of information; (iv) independent and autonomous production processes; (v) increased interoperability of production systems; (vi) predictive maintenance.

In other words, *the primary concepts of IIoT* are detailed, among which are: (i) the connection of physical objects through the Internet; (ii) high connectivity between different industrial entities; (iii) remote monitoring and control using intelligent software; (iv) autonomous system optimization; (v) improved communications.

Furthermore, 5 main lines of development of Industry 4.0 are identified, under each of them being highlighted the significant recent trends and contributions from the specialized literature: (i) networking and OPC UA; (ii) Plug & Produce, (iii) information and data models, (iv) Big Data and Cloud Computing, (v) standardization.

Thus, the networking area retains the greatest interest for the time being, probably enjoying, in the future, the fastest development of new technologies, where, through innovations such as the OPC UA protocol [7][8] or Time-Sensitive Networking (TSN) technologies [9] or Software-Defined Networking (SDN) [10], it is being tended to increase interoperability, reduce latency or improve interfacing and security.

In turn, the Plug & Produce trajectory [11] aims at: (i) achieving automatic integration of new devices into already existing production systems, without requiring manual intervention; (ii) more adaptable and flexible systems; (iii) self-discovery of devices and their capabilities through the network; (iv) self-configuration.

Studies on information and data models are directed at: (i) the variability of automated production systems [12]; (ii) the transformation of UML class diagrams into OPC UA information models [13]; (iii) generic model for configuring control devices; (iv) aggregation level of energy data [14]; (v) universal accessibility to device functionalities.

Of course, on the Big Data and Cloud Computing axis, research initiatives were distinguished oriented to: (i) Fog Computing [15]; (ii) Edge Computing [16]; (iii) Cloud platforms; (iv) migrating existing systems to the Cloud [17]; (v) high-performance Big Data techniques.

In the fifth branch, namely standardization, the problems highlighted by the literature, whose solutions are the objective of the approaches in this area are: (i) the heterogeneity of the standardization field; (ii) gaps in standardization [18]; (iii) inability to keep up with the rapid progress of technologies; (iv) difficulty in finding the most appropriate standards; (v) the multitude of participating standardization organizations.

The last part of the second chapter of the thesis reviews the most important research studies carried out under the Industry 4.0 paradigm that focused on the water industry, evoking various research papers related to the treatment and distribution of drinking water [19], water sources management [20] and wastewater treatment [21]. Besides these, the concrete, practical situation of the historian applications in the water industry, the ways and locations of setup, the way of using these applications, as well as their typical problems were analyzed, all by reference to the water industry.

Although the water industry designates a wide and varied range of operations and technical systems whose central element is water, *only two types of such systems are of interest* for the present thesis, the further developments in this paper directly relating to drinking water treatment plants (DWTP) and wastewater treatment plants (WWTP). Therefore, a typical DWTP (Figure 1) has several sources of water, the water being filtered, a process that produces sludge and chlorine is injected at multiple points in the station, before the treated water is stored, so that, finally, it gets distributed to the final consumers. On the other hand, a WWTP (Figure 2) receives wastewater, which goes through a mechanical filtration, a sedimentation, a biological treatment and then a chemical treatment in order to be able to be released into the natural environment, all these stages producing residues, in various forms. The typical problems for a DWTP include a high consumption of energy and substances, changes in the quality of water sources over time, mechanical failures and maintenance, while for an WWTP there is a high consumption of energy and substances, equipment damage, overloading of some undersized WWTPs.

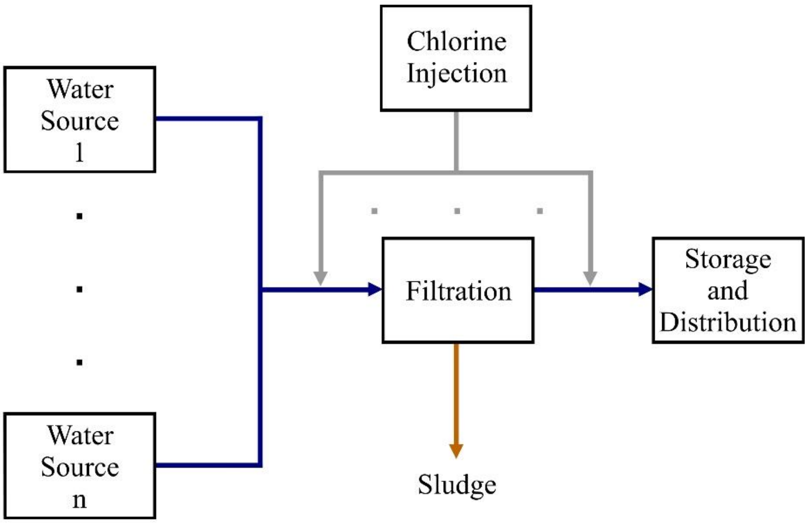


Figure 1. The typical processes from DWTP

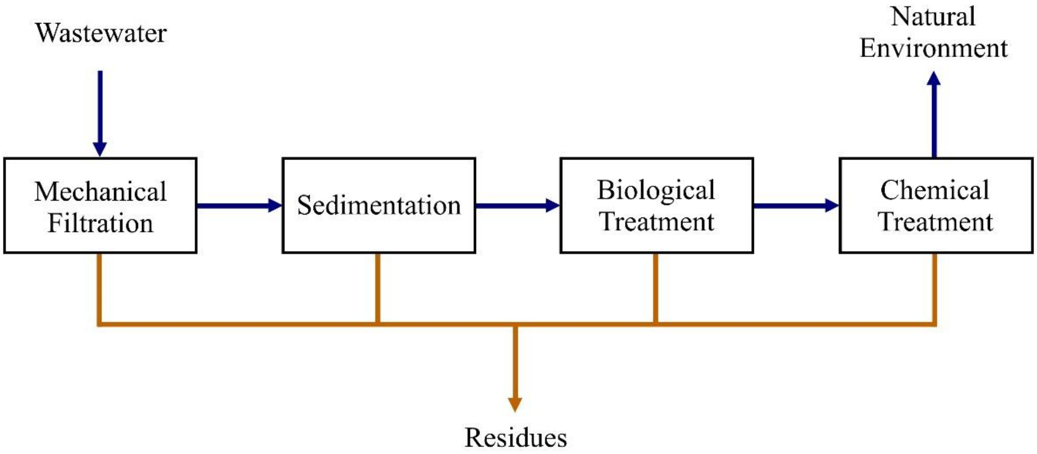


Figure 2. The typical processes from WWTP

3. Elementary historian application for the water industry

Since the third chapter of the thesis brings to the forefront the research, implementation, development and testing of a historian type of software application for the water industry, initially a practical perspective on these applications in the targeted industry (water industry) was detailed, being recorded a number of current issues, including: (i) the SCADA software comes from well-known manufacturing companies, which involves restricting access to the database itself, conditions and limitations on data handling, the requirement to purchase license extensions for the use of a historian, the marketing of historian software as a separate product at significant prices; (ii) in the vast majority of cases, the historian solution is placed only at a high level, at the top of the pyramid and not in the local SCADA control rooms, where the largest amount of data is available; (iii) the current historian solutions are platform dependent; (iv) the rare or no use of historian applications by local station operators due to the complexity of these software solutions and lack of knowledge; (v) lack of historian applications at the level of local HMI equipment. In these circumstances, it is discerned the need for a software component, with low costs and requirements, adapted for rapid integration at local level, with / through the automation panel, respectively easy to use by the local operators benefiting from low IT skills.

As a result, the general concept for an elementary historian application that meets the noted need is developed, listing the essential functionalities and features that such a solution must possess: (i) connecting to a local system with a role in controlling a process or even to devices in the industrial field that have the appropriate interface; (ii) low costs; (iii) high technological readiness level; (iv) low requirements; (v) platform independence; (vi) ease of use by operators; (vii) interfacing with local structures through OPC UA, with the possibility to extend the capabilities of the application to other protocols; (viii) automatic configuration and manipulation of a database; (ix) adaptability to changes in the set of monitored variables; (x) exporting data from the database in standard formats (Microsoft Excel, PDF and CSV); (xi) ensuring the continuous operation of the application in connection with the local system, with proper handling of all errors and exceptions; (xii) modularity of the application, in order to provide a framework, a platform for future developments.

Starting from these premises, a software solution was practically implemented that ticks all the requirements from the previous paragraph, *this application being further referred to as Historian* (Figure 3). For this, SQLite technology was used for the database, the application itself being implemented in Java, with the graphical part supported on Java Swing. The Node-Red technology was used to interface with the local system via OPC UA, the process in the operating system running this technology being controlled and monitored by the main Java application through interprocess communication techniques and capturing the output from the command line. Also, a finite state machine was developed, which the Historian application respects and increased attention was paid to the treatment of errors and exceptions, being implemented an automatic status monitoring system of all Historian components, a self-diagnosis system, a system of execution logs, for further analysis, as well as an automatic mechanism for re-establishing the connection with the local technical system, in case of interruptions. In this way, a software application has been obtained that meets all the above requirements and can connect to any OPC UA server, providing the user with the list of existing OPC UA tags on that server. The user can choose those tags whose values he wants to be stored by the Historian application, depending on the choice made, this automatically configuring the database. Additionally, the already stored data can be viewed inside the Historian, in graphical form of the evolution of values over time and can be exported outside it. For the hardware support of the solution, the Raspberry Pi 3 board was chosen.

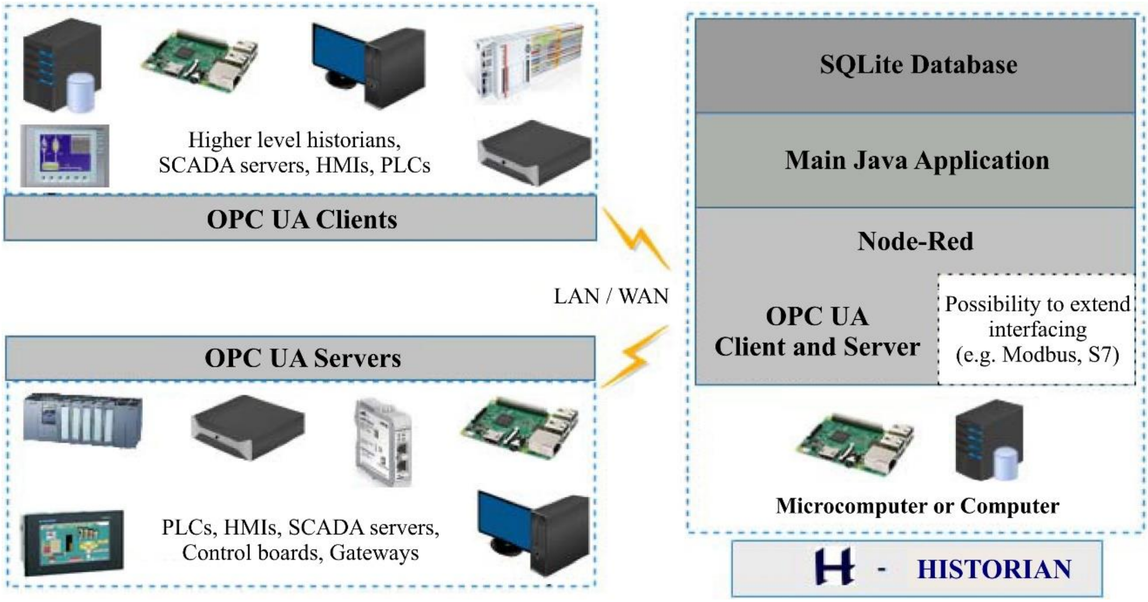


Figure 3. The general architecture of the Historian application and the relationship with the OPC UA structures

In order to test the resulting elementary Historian application, the proper operation was verified both on the Raspbian operating system (Linux), on the Raspberry Pi 3 board, and on Microsoft Windows, on a laptop with mid-level technical specifications. Subsequently, a pre-existing, solid partnership was used between the research team to which the thesis author is also affiliated, on the one hand, and the local water distribution company, on the other hand, which facilitated the rigorous testing of the Historian application, for longer periods of time, in connection with a real DWTP, which serves about 8000 inhabitants.

Although the testing of the application developed up to this point was done in the water industry, *it is stated that the implementation is a generic one, being able to fully exercise its functional characteristics for any type of OPC UA server, regardless of the technical process in which the monitored system is present*.

4. The reference software architecture for elevating the Historian application towards the proactive level

The fourth chapter marks the beginning of the elements of maximum interest, namely the elevation of the Historian application to a higher level, where it can contribute proactively, autonomously, automatically, non-invasively and unassisted to the optimization of the monitored technical system, intelligently leveraging the already recorded historical data from it, by methods in accordance with Industry 4.0 and IIoT.

For the materialization of the mentioned transformation of an elementary historian type application, which is endowed only for reading from a technical system and storing in a database some numerical values, category in which also falls the Historian application from the end of chapter 3, towards the proactive level, it is imperative to refer to a reference software architecture. As the search for solutions in the literature failed (in the attempt to identify an already existing model suitable for the particular situation), it was decided to develop such a reference architecture of its own, which can be applied to any elementary solution, not just to the Historian.

Therefore, a reference software architecture which is independent from the industry in which the historian application will operate has been proposed, the architecture being focused on the software elements that need to be added to an existing elementary historian application to convert it to a proactive one. Specifically, the architecture comprises a structure of software algorithms divided into 3 distinct levels (Figure 4).

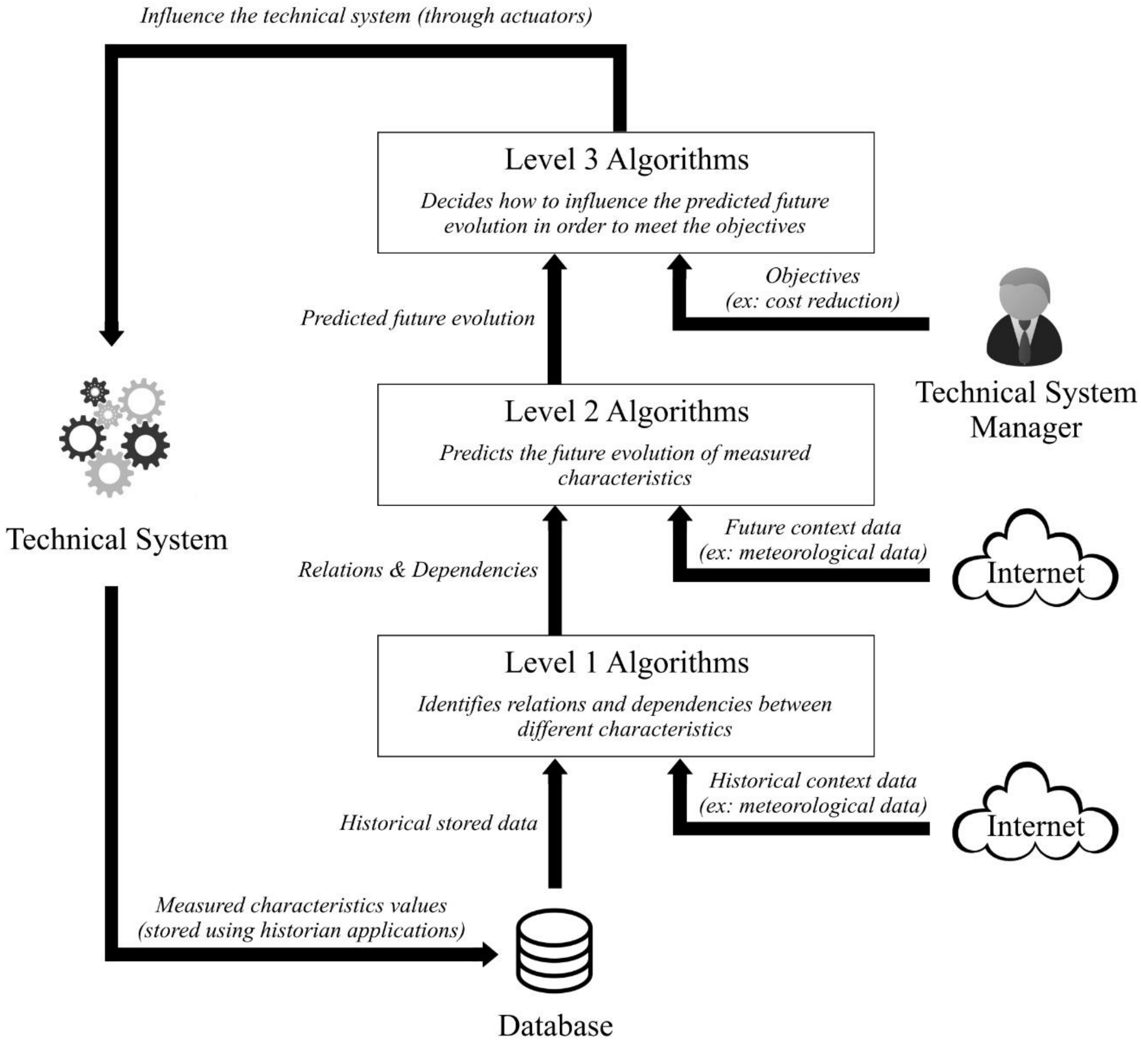


Figure 4. Reference architecture for a proactive historian application

Firstly, after the values of the parameters of the technical system are accumulated for a longer period of time, with the help of an elementary historian solution, in a database, the historical stored data in it will compose the entry for level 1 algorithms, along with this entry being optional to join, from the outside, some historical context data that could be relevant for certain particular situations (for example: historical meteorological data). In this way, the level 1 algorithms identify relationships and dependencies between different characteristics, tags, system parameters, the relationships and dependencies determined being the output of this first level.

The relationships and dependencies from the first level are one of the inputs of level 2 algorithms, optional being another input, consisting of future context data (for example: weather forecast), from the outside. Undoubtedly, the purpose of the algorithms from level 2 is to predict the future evolution of the values of the monitored parameters of the system, this prediction being the output of the second level.

The input of level 3 takes over the predicted future evolution, provided by the lower level and, obligatorily, the other input is the optimization objective (for example: cost reduction), this being chosen by a manager of the technical system. Of course, this last level is responsible for deciding how to influence the system in order to achieve the objectives, the output being the very influence, the concrete change, which must be applied to the technical system.

In other words, it is noted that the proposed architecture consists of a repetitive loop through which the evolution of the technical system is recorded, analyzed, predicted and then altered from the prediction to achieve predefined objectives, after which the iteration is resumed. Certainly, the fixed objectives can change along the way, which supports the organization in the form of a continuous loop. From a divergent point of view, the algorithms that transpose a simple, elementary historian application into a proactive solution work in a *pipeline* architecture, where each level of algorithms uses the outputs of the lower-level algorithms as its own inputs.

The following 4 chapters advance, in turn, research and implementations of the components from the previously mentioned software architecture, which were the primary object of study during the doctoral period. Thus, together with this fragment, all 5 chapters are part of the “path” of elevating the Historian project from the elementary towards the proactive level.

5. Identification of dependencies between stored data

The fifth chapter deals with the first level of the reference architecture presented in the previous chapter, aiming at researching and developing an algorithm capable of providing the functionality claimed at this level of the architecture, so that, in the final part, the algorithm in question to be integrated in the Historian application and tested in the water industry, on a real DWTP.

In the development of the algorithm from this chapter, a generic approach was kept, the concept and implementation being detached from the particularities of a certain process or industry. Due to the test site (DWTP), but also the perseverance in the same environment in the next chapter, the typical processes that take place inside a DWTP were detailed, as well as the associated defining problems. Concisely, the water entering a DWTP comes from several water sources, for the beginning being possible, in some stations, to pass through an aeration basin (in which oxygen is introduced, with the help of a blower) to control of nitrite and nitrate levels. Beyond that, a pumping station sends the water through a sand filter and a charcoal filter, during which time the injection and measurement of chlorine also takes place at several points before the water is stored in tanks, reservoirs or water towers. Subsequently, another pumping station sends this water, through the drinking water network, to the final consumers, as needed. As a secondary process, the charcoal and sand filters periodically go through filter cleaning cycles, an operation that is performed either with water or air, using water pumps or air blowers, removing mud, which can be, in some cases, treated a priori disposal. Regarding the typical problems, the following are recalled: (i) high energy consumption due to blowers and pumps, high consumers, but which ensure the regulation of acidity, alkalinity and conductivity; (ii) high chlorine consumption; (iii) increased costs caused by equipment and filter maintenance; (iv) clogging caused by a high level of turbidity; (v) water losses by cleaning filters; (vi) premature wear due to frequent switching on and off of the station, especially during the night, one of the reasons being the water losses from the distribution network; (vii) disruption of the filtration and chlorination processes due to frequent start-up and shutdown of the station; (viii) the quality of water sources varies over time; (ix) although the water quality is not the same for all sources, it is not taken into account when choosing the mainly used sources.

Therefore, the developed stored data analysis algorithm identifies the dependencies between one OPC UA tag and another one, taken as a reference, establishes the measure, the degree of dependence and exposes functional patterns, being fundamental for obtaining proactivity. Specifically, the suggested algorithm uses a characteristic as a reference and, starting from the evolution of the measured values of the reference, determines whether the other characteristics are connected in any kind to the reference. In the event that the algorithm decides that the measured values of a characteristic are related, in terms of evolution over time, to the measured values of the reference, it also calculates the degree of impact on the dependency, the two characteristics can be very closely related or, on the contrary, could have a very small influence on each other. Examining the input of the algorithm, it receives a set of data consisting of measured numerical values, stored by the historian application, corresponding to various characteristics of the technical system (for example: water pressure, water flows, water levels in tanks, energies, etc.). In addition to this data set, it is also necessary to provide an indication of the characteristic, among those present in the input data, which will be used as a reference. On the other hand, the output of the algorithm returns two "pieces" of information for each characteristic analyzed in relation to the reference, hereinafter referred to as *Proportionality* and *Quantity*. The first of these shows whether the analyzed tag evolves in proportion to the reference, having as possible values: directly proportional, inversely proportional or not proportional. The second information, *Quantity*, is provided by the algorithm as a percentage, which indicates to what extent the evolution of the measured values of the characteristic is affected by the evolution of the reference values. Of course, this quantitative report is relevant only when the values of the first information, *Proportionality*, are either directly proportional or inversely proportional. In this respect, edifying examples are: (i) the *Quantity* information is 100% - it indicates a 1:1 ratio between the analyzed characteristic and the reference, so that if the reference value changes by 20%, then the value of the analyzed characteristic changes, in turn, also by 20%; (ii) the *Quantity* information is 50% / 150% - if the reference value changes by 20%, then the value of the analyzed characteristic changes, as appropriate, by 10% / 30%. Anyway*, it is observed that the algorithm in this chapter is oriented towards a numerical type of analysis process, independent of the significance of the values received at input, which gives it a general applicability, unlimited at the particularities of an industry or technical process*.

Predictably, the practical implementation of the conceptual development of the algorithm in question was integrated in the Historian application built in chapter 3, by attaching a software module for data analysis.

In order to test the results, the same partnership was used involving the local water company, which provided a real DWTP, from which different data sets were collected and analyzed through the Historian application. The results of testing the algorithm were promising, it successfully identified relationships and dependencies between stored data. Determining exactly the correctness and accuracy of the algorithm is a difficult task in the absence of benchmark data, a gap generated by the lack of an algorithm or software module with features and capabilities similar to the one presented, which could facilitate a direct comparison. However, the dependencies obtained are categorized as plausible, in view of the vast experience of water industry specialists and the research team.

In conclusion, the current chapter launched a variant of concretization of the first level of the reference architecture, describing a software solution endowed with data analysis skills, tested in an industrial environment, the current stage being the exponent of an essential step in the effort focused on acquiring a proactive historian application.

6. Reducing energy consumption in DWTP

Once the Historian application with data analysis capabilities was at hand, the opportunity to use it to procure results that would later allow research and refinement of an optimization strategy is perceived. Chapter 6 proposes a decision and control solution (FDC), placed in the Fog Computing area (Figure 5), which reduces energy consumption in a process of treatment and distribution of drinking water, using concepts based on IIoT, such as interoperability or non-invasive modification of local control systems, following the identification of recipes, after a long-term analysis of dependencies between data. In this sense, it is found that in DWTP the energy efficiency is strongly linked to the allocation and proper use of water sources. Of course, the interest of companies in such a reduction in costs is clear, being able to highlight benefits including in the ecological area, to reduce pollution.

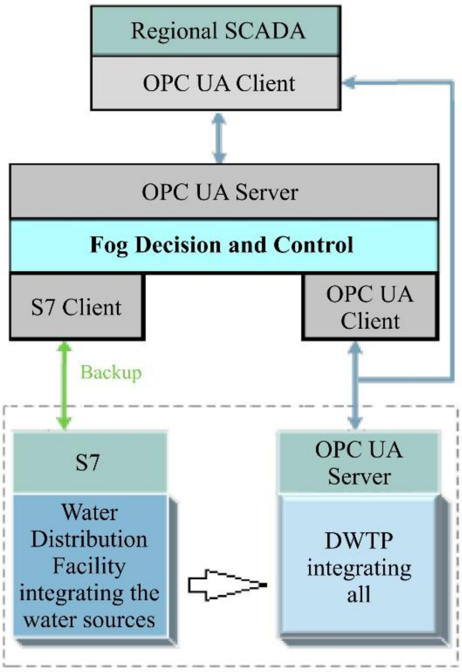


Figure 5. FDC placement relative to the targeted installation

In the first part of the chapter, the targeted water installation is detailed, which consists of both DWTP and a water distribution installation, having available 4 drillings (as water sources), each one with two automatic control loops, the primary one based on flow, and the secondary one based on level. In addition, the setpoints of the water sources have fixed values, set by the operators. Additional, the pumping station of the distribution facility has 3 tanks and 3 pumps, having implemented: (i) a pressure-based control loop for water distribution and pump rotation, depending on the functioning hours; (ii) a primary, level-based control loop that keeps the level in the tanks within the limits by requesting water from the sources; (iii) a secondary flow-based control loop for anticipating high water demands in the distribution network, at critical hours. The water sources are equipped / controlled with PLCs, while the DWTP is automated by two other redundant PLCs, with WinCC 7.2 type of SCADA software, with Connectivity Pack.

The FDC strategy presented in this chapter of the thesis has as a starting point the association of quality indicators for the water sources of the targeted installation, these indicators being determined by also using the Historian application. Also, a series of formulas are defined, with the help of which FDC establishes a general priority indicator for each source (this being calculated after determining a priority indicator based on the quality of the water from that source and a priority indicator based on the functioning hours of the source - degree of wear). Subsequently, the FDC calculates an optimal water flow for each source, depending on the priority indicator based on the quality of that source and the minimum and maximum flow rates of the source. Finally, an extensible flow distribution algorithm is developed that is adaptable to the possibility of adding new water sources. In fact, this algorithm decides how to divide the total flow requested by the existing automation of the station (the water flow at the entrance to DWTP) into specific flows for each source, an operation that takes into account the general priority of sources and the optimal flows calculated by FDC. In a simplified way, this algorithm chooses the minimum possible flow rate at the source with the lowest priority among the ones that needs to be started, the optimal flow calculated by FDC at the sources with the highest priority and the flow difference at the source with the penultimate priority among those started, in so as to cover the entire requested flow, the rest of the sources being stopped.

Intending to test the feasibility of FDC, *a model was developed, by the members of the research team, in Matlab-Simulink, of the targeted DWTP*, not presented in the thesis, and real data were used in it, obtained with the Historian tool from the real installation. Obviously, in the Matlab model a complete FDC was used, without constraints, being set as flow references (setpoint) for water sources the values calculated by FDC. By these means, a 9% improvement in energy consumption in DWTP was obtained when FDC was used, compared to operation without FDC. Taking advantage of this first proof of FDC efficiency, an authorization was allowed for some tests on the real DWTP by the owner water company, but with the imposition of constraints on FDC, not being accepted the change of setpoints of the flow of water sources from their individual fixed values or the activation of additional water sources, in addition to the 4 selected from the 6 available. However, the constrained FDC solution was applied to the real DWTP for a period of 2 weeks and a comparison of the total energy consumption with a 4-week posterior period in which no FDC was run was performed, *resulting in an impressive reduction in energy consumption, of 30%,* from a weekly average of 3.5 MWh without FDC to 2.7 MWh with FDC.

Although the effectiveness of the FDC has been proven, the members of the research team estimate that using an unrestricted FDC solution, at its full potential, on a real station, would bring even greater improvements, arguing that: (i) access has been limited to the actual drinking water installation, long-term testing, for several months, not being possible; (ii) during the tests there was a high demand for water from the network and a small number of drillings were used, conjuncture that diminished the degrees of freedom in operation; (iii) FDC was constrained, not reaching its full potential, the setpoints of the sources flows remaining fixed, at values decided arbitrarily by the operators, not by the FDC.

7. Predictive strategy based on meteorology in WWTP

Chapter 7 returns to the development of the Historian application, starting from its evolutionary stage reached at the end of chapter 5. The research cycle focuses on the second level of the reference software architecture discussed in chapter 4. Extensions, additions and improvements were also considered for the practical implementation in Historian of the first level of the reference architecture, important in this sense being the direction of integrating some contextual data in the implementation, in order to increase the performance of the solution. Due to the reporting to the water industry, the most appropriate such data for the domain in question are those related to weather conditions, and their greatest influence in the water industry is felt on wastewater treatment plants (WWTP).

Consequently, the studies in this phase were directed to a WWTP, a detailed characterization of the processes that take place in a typical WWTP being presented in the thesis. Briefly, the water entering the WWTP from the wastewater network begins the pre-treatment phase, in which the treatment of odors takes place, and then the water passes through a screening filter to remove large objects (tree branches, gravel, bottles, sanitary items, etc.), these objects reaching incinerators or landfills. Later, in the primary treatment phase, the water enters a sedimentation basin, where the sludge deposits at the bottom of the basin, while the fat and oil rise to the surface, those being extracted (the sludge for the sludge treatment process, while the grease and oil can be used in the manufacturing of soap). A priori to the secondary treatment, there is an optional bypass, which, when used, sends the treated water directly into the natural environment, without entering the next stages, thus protecting the station from hydraulic overloading during the receiving of very high water quantities (for example: heavy rain). The secondary treatment stage uses, in series, both a bioreactor basin, where oxygen and a biological mixture are introduced, and a clarification basin, the goal being the elimination of the remaining organic matter. Tertiary treatment neutralizes phosphorus through chemical additions. Downstream, the water passes through a sand filter and a charcoal filter, before reaching a disinfection tank, where another chemical mixture is added, at the end being another basin, for water dechlorination, in an attempt to avoid toxicity for the aquatic species. At the end, the treated water is released into the natural environment. At the same time, some stations may be equipped with a process of treating the sludge derived from water cleaning and filters washing.

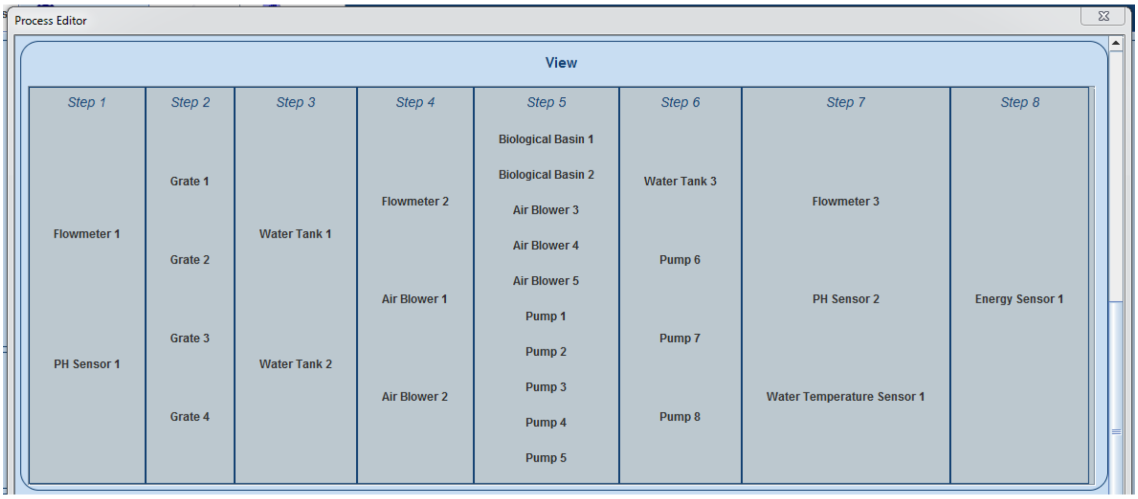
Regarding *the defining issues in the WWTP* that are of concern also in view of the theme of this research cycle, it is noted: (i) the overuse of the WWTP, which can reduce the efficiency of the secondary treatment or cause sludge leaks; (ii) high substance use; (iii) high energy costs; (iv) costly malfunctions of equipment or control algorithms; (v) undersized WWTPs due to lack of capacity increase over time.

On the other hand, the operation of an WWTP is altered by the meteorological conditions in the area covered by the wastewater network that supplies water to the WWTP, the most prominent interference being caused by the amount of precipitation, which dictates the use of the bypass channel, but also the temperature, from the perspectives of odor treatment, biological treatment and dehydration of sludge, or storms during the fall, which can generate large amounts of leaves and branches, with potential to clog the filter for screening can be underlined.

In these circumstances, the implementation of the solution began with a series of improvements and changes to the Historian application already formulated in chapter 5. In addition to an increase in the accuracy of the first level algorithm, instead of choosing a reference tag, the algorithm was modified in such a way that this information is no longer requested and each of the analyzed characteristics plays, in turn, the role of the reference, in a succession of multiple executions of the algorithm. With this adjustment, the results of the algorithm execution, meaning the relationships and dependencies between the data, are now stored in an oriented and weighted graph, materializing a broader understanding of all relationships, connections and dependencies that exist within the monitored system. Moreover, the possibility of using historical meteorological data (7 different characteristics), procured from an online service, as input data for the level 1 algorithm was added, thus also calculating the dependencies of the technical system tags on these meteorological characteristics, the results being added to the same graph of dependencies. In order to acquire the corresponding meteorological data, the user is asked to include the address of the station, based on which, by calling an online service, the Historian solution determines the exact geographical coordinates for which the meteorological data is requested.

Along with the mentioned improvements, in continuation it was studied and implemented, by joining to the Historian application, an algorithm for predicting the values of the monitored tags of the WWTP based on a breadth-first search of the dependency graph and some specific numerical processing that start from the meteorological forecast, also obtained from the online service, and use the relationships identified at the first level, the execution of the prediction algorithm succeeding a successful execution of the algorithm at the lower level. In this way, at the output of the new algorithm, predicted numerical values are provided, for a period of 7 days from the prediction date, for each characteristic initially analyzed by the first level algorithm.

As the last part of the solution implemented in Historian in this chapter, an autonomous proactive software application must understand, from a certain stage of development ahead, the meanings behind the tags of a technical system, the concept of *process-aware historian application* being essential for correct predictions, recipes, relevant analyzes of dependencies between data, as well as for the interpretation of objective functions and constraints. On this side, for the practical implementation of this vanguardist hypothesis, *not yet found in literature or industry*, a new module was added to the Historian solution, which allows the user to define, in the graphical interface, a model of the process carried out in the monitored technical system, the model being made up of predefined objects (Figure 6), with predefined properties (for example: water pump with energy consumption, functioning time, status, etc.). For each property of each object, a tag monitored by the Historian can be associated, so that the software application will be able to understand the exact meaning of these tags, information particularly valuable for future steps in the direction of optimization. In addition, saving multiple models is allowed, facilitating an easy transition between monitoring of different systems by the Historian application, each model also allowing the attachment of specific constraints, joining a numeric value, a tag and a relationship (smaller, larger, equal).



Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
Flowmeter 1 PH Sensor 1	Grate 1 Grate 2 Grate 3 Grate 4	Water Tank 1 Water Tank 2	Flowmeter 2 Air Blower 1 Air Blower 2	Biological Basin 1 Biological Basin 2 Air Blower 3 Air Blower 4 Air Blower 5 Pump 1 Pump 2 Pump 3 Pump 4 Pump 5	Water Tank 3 Pump 6 Pump 7 Pump 8	Flowmeter 3 PH Sensor 2 Water Temperature Sensor 1	Energy Sensor 1

Figure 6. A process of an WWTP defined in Historian

To test the results, the Historian solution collected data from a real WWTP for a long time, after which the prediction algorithm was executed in 7 different test scenarios. Over the course of the predicted time period, the Historian application also stored the real data from the WWTP, so that, later, the prediction could be compared with the real data. However, it has been observed that the exact overall accuracy of the prediction algorithm is very difficult to assess because it is directly impacted by the accuracy of the weather forecast, but some promising results have been identified in the perspective of studying some future optimization strategies and techniques in WWTP. For example, the prediction algorithm predicted very precisely the usage of a pump used at the bypass channel of the WWTP, which has a massive footprint both on the overall electricity consumption of the WWTP and on the total costs with substances of the WWTP. Similarly, the turbidity of the water at the exit of the WWTP was predicted quite well, the anticipation of this landmark feature for the water quality favoring better estimates of future substances consumptions. Again, good results were also encountered in predicting the electricity consumption of the blower of the biological basin and the volume of water upon admission to the WWTP.

Looking from another angle, it has been managed to preserve a generic mentality in the augmentation of the Historian application, even at this level, so that it did not become limited to the water industry, being convenient a possible easy metamorphosis towards another industry, with all the features implemented up to this point.

The contributions presented in this research stage offer both improvements to the first level of the reference architecture, and an investment in the second layer of the same architecture, at the same time, also preparing the way for future innovative prospecting.

Therefore, being tested on a real WWTP, the transposition of the prediction algorithm in the Historian solution allows the glimpse of various possible optimization paths to be followed in the WWTP, in a future level 3 implementation of the reference architecture based on these predictions.

8. Complete automation of the strategy to reduce energy consumption in DWTP

The eighth chapter edges the last stage of research, aiming at a consonance with level 3 of the reference architecture, by integrating the optimization strategy from chapter 6 in the

Historian application, the purpose being to acquire the capability to influence the functioning of the water sources of a DWTP to reduce energy consumption, in a non-invasive, autonomous and unassisted by human operation. Specifically, the difficulty of identifying and automatically adapting the quality indicators of water sources was addressed, being designed a solution through continuous, automatic, long-term analysis within the Historian application. Of course, it is insisted that it was not studied a new automation of the central control loop of a DWTP, but a global automation in applying an optimization strategy on the already existing system, the solution implemented in this chapter being complementary to the main, centralized solution of automatic control of a DWTP, the latter having priority in operation. Therefore, the contributions of this chapter are guided towards filling the gaps that prevent the fully automatic, long-term operation of the strategy developed in chapter 6.

Next, some DWTP specific issues relevant in the context are brought to the attention. Consequently, *because invasive interventions at the PLC or SCADA level of a local system are avoided at all costs, for justified reasons, especially in critical infrastructure such as the water one, the only practically feasible solution to apply optimizations is in a non-invasive way, such as substituting references in a local control loop*. On the other hand, the total, mixed flow, directed at the DWTP inlet, in which all the running sources introduce water, is the one requested by the DWTP's automation, which propels towards the finding according to which both the way in which this total flow is divided between sources, as well as the choice of those sources that will be used, from the available whole, are decided, in practice, by taking into account only the functioning hours of the water sources. Besides this, the better the water quality at the entrance of the DWTP, the less treatment will be required, which means lower energy consumption. In terms of water quality, there are noticeable differences between water sources, at the inauguration of a new source a technical data sheet is drawn up, where several water parameters are passed, after sampling and performing laboratory analyzes. Even if, over time, the water quality of a source changes, the analyzes are not repeated for reasons of costs and difficulty of sampling. Anyway, the analysis returns values for a multitude of parameters, as there is no single water quality indicator parameter and no established method for determining it from other parameters. In these circumstances, the water quality is not taken into account in the decision to select water sources in the current operation, in none of the DWTPs, the reality being that all these operate over their optimal energy consumption.

Having available, from previous research undertaken in the doctoral framework, a historian type application for storing the operating parameters of a technical system (chapter 3), a module for identifying the dependencies between data (started in chapter 5 and improved in chapter 7), possibilities to know the significance of the monitored tags (chapter 7) and, most importantly, a strategy to reduce the energy consumption of a DWTP by ranking its water sources and setting the flow reference for each local control loop (chapter 6), the complete automation of the optimization strategy from chapter 6 is blocked by: (i) the absence of a fully automated method to determine the water quality indicator for each water source in DWTP; (ii) the lack of a robust software implementation of both the sequence of equations from FDC and the algorithm for allocating the total required water flow between sources. Therefore, the discontinuities evoked above substantiate the primary interest of this last phase of research, and the solutions to these problems monopolizes the main scientific and practical contributions of the stage in question.

In order to be able to use the automatic tool to reduce energy consumption, any DWTP must meet the following conditions: (i) have at least 2 water sources; (ii) certain data, such as flow rates or energy consumption, are provided for a period of time before the optimization is applied; (iii) to the previous data must be attached the particular meanings, through the corresponding module of the Historian; (iv) the automation already present in the DWTP is mandatory to have control loops implemented in such a way that they use, as a setpoint for the water flow that each source delivers, the numerical values assigned to an OPC UA tag.

In this sense, it was decided that the best approach for obtaining the desirable complete automation of the strategy to reduce energy consumption would be its integration in the Historian application, aiming at a concordance with level 3 of the reference software architecture.

Firstly, a method has been developed to calculate the water quality indicator for each source, without requiring any human assistance or complicated analysis and sampling, by studying the flows and energy consumptions recorded while the DWTP has operated in the past. Then, the calculations according to the FDC formulas were implemented in the Historian application, so that, finally, the logic for dividing the total requested flow into individual flows for water sources is implemented. Knowing which of the tags correspond to the flow references of the DWTP sources, the Historian application can write the optimal flow values, calculated by the energy reduction strategy, as values for the respective tags, thus optimizing, non-invasively, the monitored system.

In this way, *for the first time, the proactive Historian solution closes the loop* characterized in the reference architecture established in chapter 4, becoming able to monitor a technical system, analyze stored data and use the conclusions to plan a reduction of energy consumption in DWTP, the necessary operations being administered directly to the system, the operation of which is influenced by the proactive Historian application.

In regards to testing, data from a real DWTP were used and, through successive changes in the database of the Historian application of the total flow requested at the entrance to the station, the correctness of the calculations was verified, as well as the exact observance of the flow division algorithm proposed by FDC in this complete automation, assimilated in the Historian solution. In fact, the efficiency of FDC in reducing consumption has already been demonstrated in chapter 6, at the level of this chapter being tested only the compliance of the automation from inside Historian with FDC.

In conclusion, this research stage must be interpreted as the acquisition of a tool integrated inside a proactive historian type of software solution, which can apply, in a fully automated manner, without requiring human cooperation, a strategy to reduce energy consumption within a DWTP, in a non-invasive way in terms of the local automation of the station.

9. Conclusions of the thesis

Against the background of the emergence and continuous refinement of the IIoT and Industry 4.0 principles in recent years, there is an increase in the amount of data conveyed through the decor of industrial automations, which, together with the use of historian type of software applications, have generated the ocurrence of the data accumulation phenomenon. In this context, the approach of the thesis in discussion investigated *a number of methods and strategies through which the multitude of stored data can be utilised, in a useful way*, through wise and innovative means, with the purpose of optimizing various aspects of the supervised technical systems. In practical terms, *the development of a complex software application, of proactive historian type*, fully functional, tested and validated on real systems in the water industry, capable of optimizations in an autonomous, automatic and unassisted by man manner was detailed, *this application being without equivalent in the literature or industry at the time of implementation*. Also, the efforts summarized in the thesis put their shoulder to the renewal and advance in the context of this Industrial Revolution maintained by the notions of Industry 4.0 and IIoT, contributing *a small step forward in the area of increasing the intelligence of industrial software solutions*.

With regard to *personal contributions*, the author of the thesis reclaims, on his own behalf, the following: (i) the carrying out of a detailed study of the current state of research in the field of the thesis [22]; (ii) the distinguishing and cataloging of some main development directions in the branch of competence of the paper, together with conducting a review-type analysis of existing scientific papers [22]; (iii) the practical implementation and testing on real systems from the water industry of an elementary historian type of software application [23]; (iv) the conceptualizing and developing of a reference software architecture for elevating an elementary historian solution towards the proactive level [24]; (v) the conducting of a research to recognize the typical processes and problems found in a DWTP [24]; (vi) the conceptualizing, developing, implementing, integrating into the Historian and practical testing of an algorithm for identifying dependencies between stored data [24]; (vii) the installation on the hardware platform, the configuration and the launch in operation, in the industrial climate, of the Historian application, in the version available at the respective phase of its evolution, the results of which were subsequently fructified by the rest of the team in the optimization from chapter 6 [25]; (viii) the conducting of a study to discover and observe the particular processes, the defining problems and the usual meteorological influence in a typical WWTP [26]; (ix) the conceptualization, development, implementation, integration into Historian and practical testing of a prediction algorithm, based on the influence of meteorological characteristics on an industrial technical system [26]; (x) the development, implementation and integration in Historian of a software module for the insertion and association, in the Historian application, of a precise specific meaning to each tag of the monitored system [26]; (xi) the developing and integrating into the Historian of a complete automation of the process of applying the energy consumption optimization strategy in a DWTP [27]; (xii) the collection of a large volume of data, with the application in question, from real technical structures in the water industry, which have been used to test, at various times, the latest additions to the Historian project [23][24][25][26][27]; (xiii) the searching for and inspecting, on the one hand, of the software alternatives that include reuse potential, and, on the other hand, of the available tools and services that could have been used to fulfill the purposes of each research phase [23][24][25][26][27].

Regarding the dissemination of the information, during the current doctoral studies, 7 scientific articles of specialty were published, out of which 2 at conferences and 5 in journals, all being indexed in the Web of Science Core Collection databases.

The future directions of research and development are numerous, the most important of which being: (i) the evaluation of some approaches based on Machine Learning and artificial intelligence technologies at the first level of the reference architecture, for the identification of dependencies between data; (ii) the addition of other contextual data than meteorological data, both at the first level of the architecture and at the second level, with a view to improving the accuracy of the prediction algorithm; (iii) the studying and adding to the Historian application of new optimization objectives, in addition to reducing costs in DWTP by prioritizing water sources; (iv) the study of some tactics and methods to always keep the water from the exits of the stations in the vicinity of the legally accepted lower extremity of quality, but without exceeding this limit, with the purpose of optimizing the operational costs; (v) the investigating of means of self-triggering warning of qualified personnel in relation to overloads or overuses that will occur in the future at equipment, using the existing prediction of future evolution in the WWTP for the predictive protection of industrial equipment; (vi) the implementation of an OPC UA server inside the Historian solution to make the stored data available to other external agents, including computers, scripts, automations; (vii) the development of the Historian application so as to master a wide range of other protocols, in addition to the OPC UA, through which communication with the monitored technical system can take place; (viii) the determination of an automatic strategy for the periodic running of the necessary algorithms, being indispensable the establishment of the ideal time intervals between these runs and between the interventions, through new adjustments, on the supervised system; (ix) the evaluating of existing alternatives for replacing the usage of Node-Red with a composition that reads the values of the monitored tags directly in the main Java application, preferably with a single complex delivery, in block, of all the requested values, instead of the current sequential manner from Node-Red, thus obtaining an improvement of the performance of the Historian application by eliminating the management of another operating system process; (x) the implementing of the establishment of the predefined objects and their characteristics outside the Historian application, in an XML file, verified in the application by DTD technology, so as to facilitate the adaptation of the solution for other industries by removing, in this case, the imperativity of making changes in the source code; (xi) the updating and reimagining of the artistic, coloristic part of the general appearance of the graphical user interface of the Historian application, which leaves little to be desired, compared to today's standards, in order to be able to take advantage of an eventual potential of commercial harnessing of the application; (xii) the research into ways in which the functionalities of the Historian application could bring benefits in other industries, by deepening the particular processes in the respective industries and an eventual adaptation of the solution for the necessary distinctive notes, together with the pursuit of some own optimization objectives for that sector of activity.

Selective bibliography

[1] M. Pinzone, P. Fantini, S. Perini, S. Garavaglia, M. Taisch, G. Miragliotta, “Jobs and Skills in Industry 4.0: An Exploratory Research”, IFIP Advances in Information and Communication Technology: Advances in Production Management Systems. The Path to Intelligent, Collaborative and Sustainable Manufacturing, vol. 513, pp. 282–288, 2017.

[2] A. Ustundag, E. Cevikcan, “Industry 4.0: Managing The Digital Transformation”, Springer International Publishing, Switzerland, 2018.

[3] Y. Liao, F. Deschamps, E. Loures, L. Ramos, “Past, present and future of Industry 4.0 - a systematic literature review and research agenda proposal”, International Journal of Production Research, pp. 1–21, 2017.

[4] S. Jeschke, C. Brecher, T. Meisen, D. Ozdemir, “Industrial Internet of Things and Cyber Manufacturing Systems”, Springer International Publishing, Switzerland, 2016.

[5] H. Xu, W. Yu, D. Griffith, N. Golmie, “A Survey on Industrial Internet of Things: A Cyber-Physical Systems Perspective”, IEEE Access, vol. 6, pp. 78238–78259, 2018.

[6] A. Gilchrist, “Industry 4.0: The Industrial Internet of Things”, Apress, New York, USA, 2016.

[7] I. Gonzalez, A. J. Calderon, J. Figueiredo, J. M. C. Sousa, “A Literature Survey on Open Platform Communications (OPC) Applied to Advanced Industrial Environments”, Electronics Journal, vol. 8 (issue 5), 2019.

[8] H. Haskamp, M. Meyer, R. Mollmann, F. Orth, A. Colombo, “Benchmarking of existing OPC UA implementations for Industrie 4.0-compliant digitalization solutions”, IEEE 15th International Conference on Industrial Informatics (INDIN), Emden, Germany, pp. 589–594, 2017.

[9] A. Gogolev, F. Mendoza, R. Braun, “TSN-Enabled OPC UA in Field Devices”, IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA), Turin, Italy, pp. 297–303, 2018.

[10] M. Ehrlich, D. Krummacker, C. Fischer, R. Guillaume, S. S. P. Olaya, A. Frimpongk, H. de Meer, M. Wollschlaeger, H. D. Schotten, J. Jasperneite, “Software-Defined Networking as an Enabler for Future Industrial Network Management”, IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA), Turin, Italy, pp. 1109–1112, 2018.

[11] T. J. Schoepf, “White paper: The road to plug-and-produce”, Lumberg Automation, Tech. Rep., 2016.

[12] B. Vogel-Heuser, J. Mund, M. Kowal, C. Legat, J. Folmer, S. Teufl, I. Schaefer, “Towards interdisciplinary variability modeling for automated production systems:

Opportunities and challenges when applying delta modeling: A case study", IEEE 13th International Conference on Industrial Informatics (INDIN), Cambridge, United Kingdom, pp. 322–328, 2015.

[13] F. Pauker, S. Wolny, S. M. Fallah, M. Wimmer, "UML2OPC-UA — transforming UML class diagrams to OPC UA information models", 11th Conference on Intelligent Computation in Manufacturing Engineering (ICME), Naples, Italy, 2017.

[14] A. Wurger, K. H. Niemann, A. Fay, "Concept for an Energy Data Aggregation Layer for Production Sites: A combination of AutomationML and OPC UA", IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA), Turin, Italy, pp.1051–1055, 2018.

[15] M. Aazam, S. Zeadally, K. A. Harras, "Deploying Fog Computing in Industrial Internet of Things and Industry 4.0", IEEE Transactions on Industrial Informatics, vol. 14 (issue 10), pp. 4674–4682, 2018.

[16] J. H. Huh, Y. S. Seo, "Understanding Edge Computing: Engineering Evolution With Artificial Intelligence", IEEE Access, vol. 7, pp. 164229-164245, 2019.

[17] R. Khan, K. McLaughlin, B. Kang, D. Laverty, S. Sezer, "A Secure Cloud Migration, Monitoring and Analytics Framework for Industrial Internet of Things", IEEE 6th World Forum on Internet of Things (WF-IoT), New Orleans, USA, 2020.

[18] O. Meyer, G. Rauhoeft, D. Schel, D. Stock, "Industrial Internet of Things: covering standardization gaps for the next generation of reconfigurable production systems", IEEE 16th International Conference on Industrial Informatics (INDIN), Porto, Portugal, pp. 1039–1044, 2018.

[19] H. Mala-Jetmarova, N. Sultanova, D. Savic, "Lost in Optimisation of Water Distribution Systems? A literature Review of System Design", Water Journal, vol. 10 (issue 3), 2018.

[20] S. Chowdhury, "Water quality degradation in the sources of drinking water: an assessment based on 18 years of data from 441 water supply systems", Environmental Monitoring and Assessment, vol. 190, 2018.

[21] A. Durán-Sánchez, J. Álvarez-García, E. González-Vázquez, M. C. Del Río-Rama, "Wastewater Management: Bibliometric Analysis of Scientific Literature", Water Journal, vol. 12 (issue 11), 2020.

[22] A. Nicolae, A. Korodi, I. Silea, "An Overview of Industry 4.0 Development Directions in the Industrial Internet of Things Context", Romanian Journal of Information Science and Technology (ROMJIST), vol. 22 (issue 3-4), pp. 183-201, 2019.

[23] A. Nicolae, A. Korodi, "Node-Red and OPC UA Based Lightweight and Low-Cost Historian with Application in the Water Industry", IEEE 16th International Conference on Industrial Informatics (INDIN), Porto, Portugal, pp. 1012–1017, 2018.

[24] A. Nicolae, A. Korodi, I. Silea, "Identifying Data Dependencies as First Step to Obtain a Proactive Historian: Test Scenario in the Water Industry 4.0", Water Journal, vol. 11 (issue 6), 2019.

[25] A. Korodi, R. Crişan, A. Nicolae, I. Silea, "Industrial Internet of Things and Fog Computing to Reduce Energy Consumption in Drinking Water Facilities", Processes Journal, vol. 8 (issue 3), 2020.

[26] A. Nicolae, A. Korodi, I. Silea, "Weather-Based Prediction Strategy inside the Proactive Historian with Application in Wastewater Treatment Plants", Applied Sciences Journal, vol. 10 (issue 9), 2020.

[27] A. Nicolae, A. Korodi, I. Silea, "Complete Automation of an Energy Consumption Reduction Strategy from a Water Treatment and Distribution Facility, Inside an Industrial Internet of Things-Compliant Proactive Historian Application", Sensors Journal, vol. 21 (issue 7), 2021.