FACULTY OF TECHNOLOGY

SYSTEMATIC SPECIFICATION OF REQUIREMENTS FOR ASSEMBLY PROCESS CONTROL SYSTEM IN PHARMACEUTICAL INDUSTRY

Teemu Vento

DEGREE PROGRAMME OF MECHANICAL ENGINEERING

Master's Thesis

June 2021

ABSTRACT

Systematic specification of requirements for assembly process control system in pharmaceutical industry

Teemu Vento

University of Oulu, Degree Programme of Mechanical Engineering

Master's thesis 2021, 62 pp.

Supervisor(s) at the university: D.Sc. (Tech.) Toni Liedes

Pharmaceutical manufacturing is one of the most strictly regulated fields in the world. Manufacturers of pharmaceutical products are juridically obliged to monitor the safety and quality of products. Any defects and manufacturing errors affecting the product are demanded to be traceable due to patient safety. Regulative bodies have set strict demands for data integrity in manufacturing records. The main objective of this thesis is to evaluate whether the proposed supervisory control and data acquisition software can adhere to current prevailing regulatory framework.

The evaluation of the proposed supervisory control and data acquisition software focuses on handling of electronic records and electronic signatures. Features like user management, alarm and event management, reporting, and locally set requirements in the target company are investigated and reflected to the prevailing regulations concerning data integrity.

The results showed that the proposed software is, when properly configured, compliant to prevailing regulations regarding electronic records and electronic signatures. In addition, the proposed software is capable of the requirements set by the target company.

Keywords: pharmaceutical industry, pharmaceutical regulation, industrial control system

TIIVISTELMÄ

Systemaattinen vaatimusmäärittely kokoonpanoprosessin ohjausjärjestelmälle lääketeollisuudessa

Teemu Vento

Oulun yliopisto, Konetekniikan tutkinto-ohjelma

Diplomityö 2021, 62 s

Työn ohjaaja(t) yliopistolla: Toni Liedes

Valmistava lääketeollisuus on yksi maailman eniten säädellyin teollisuuden ala. Lääkinnällisten tuotteiden valmistaja on lainmukaisesti vastuussa tuotteidensa laadusta ja valmistuksen valvomisesta. Tuotteiden laatu- ja valmistusvirheiden vaaditaan olevan jäljitettävissä potilasturvallisuuden vuoksi. Sääntelyviranomaiset ovat asettaneet tiukat vaatimukset tuotantokoneiden elektronisille tallenteille. Tämän diplomityön tavoitteena on arvioida noudattaako ehdotettu ohjausjärjestelmä nykyisiä säädöksiä.

Ohjausjärjestelmän arviointi keskittyy eletronisten tallenteiden ja elektronisten allekirjoitusten toteutukseen ohjelmassa. Arvioinnin perustana käytetään sääntelyviranomaisten viimeisimpiä säädöksiä. Arviointi kohdistuu ohjelmiston käyttähallintaan, hälytys- ja tapahtumahallintaan, raportointiin ja paikallisesti asetettuihin vaatimuksiin tiedon eheyden näkökulmasta.

Arviointi osoitti, että oikein konfiguroituna ehdotettu ohjausjärjestelmä noudattaa nykyisiä säännöksiä elektronisten tallenteiden ja elektronisten allekirjoitusten osalta. Ohjelmisto pystyy myös vastaamaan yrityksen paikallisesti asetettuihin vaatimuksiin. Ohjelmistoa voi kuitenkin käyttää vastoin nykyisiä sääntelyviranomaisten laatimia säädöksiä ilman riittävää asiantuntevuutta.

Keywords: lääketeollisuus, lääkevalvonta, ohjausjärjestelmä

PREFACE

This thesis was conducted during extra-ordinary times and while working full-time as an Industrial Control System IT Specialist in the pharmaceutical industry. The acquired knowledge has already been helpful in my current work tasks.

Big thanks to both my instructor and supervisor for prompt and periodic feedback. The periodic discussions were instrumental in finishing this thesis on schedule. I also want to thank my colleagues for giving me insight on industry regulations and their practical implementations.

Lastly, I want to thank my better half for emotional support and for putting up with me throughout the thesis process.

Turku, 29.06.2021

Teemu Vento
Teemu Vento

TABLE OF CONTENTS

ABSTRACT
TIIVISTELMÄ
PREFACE
TABLE OF CONTENTS
ABBREVIATIONS AND SYMBOLS
1 INTRODUCTION 7
2 PHARMACEUTICAL MANUFACTURING INDUSTRY 10
2.1 Regulatory agencies and requirements 11
2.2 Good manufacturing practice 12
2.3 Validation and qualification 14
3 INDUSTRIAL CONTROL SYSTEMS 16
3.1 Automation system pyramid 16
3.2 Supervisory control and data acquisition 18
3.2.1 Data acquisition 19
3.2.2 Communication of data 20
3.2.3 Data visualization 20
3.2.4 System control 21
3.2.5 Architecture and components 21
3.2.6 Current trend 24
3.3 Virtualization in industrial environment 25
4 DATA INTEGRITY 27
4.1 ALCOA+ standard 27
4.2 Legislation and guidance documents 30
4.3 Industrial control system data and records 32
4.4 Ensuring data integrity 34
4.4.1 Supplier assessment 34
4.4.2 Technical support features 35
4.4.3 Logical security 36
4.4.4 Physical security 37
5 ASSEMBLY PROCESS OF THE TARGET COMPANY 39
5.1 Role and responsibility of IT 39

5.2 Assembly process 40
6 CASE STUDY: EVALUATION OF PROPOSED CONTROL SYSTEM FOR TARGET COMPANY 43
6.1 Proposed software 43
6.2 User management 44
6.2.1 Authorization levels for segregation of duties 45
6.2.2 Connectivity to centralized user management software 45
6.2.3 Benefits of centralized user management 46
6.3 Audit trail and alarm management 46
6.3.1 Audit trail functionality 46
6.3.2 Time synchronization 47
6.3.3 Archiving of audit trail 47
6.3.4 Potential risks regarding audit trail archive 48
6.4 Reporting 48
6.5 Interfaces to other systems 49
6.5.1 Export module 49
6.5.2 Process gateway 50
6.5.3 Historian module 50
6.5.4 Open platform communication unified architecture 51
7 DISCUSSION 53
7.1 Result summary and benefits 53
7.2 Challenges 54
8 CONCLUSION 56
REFRENCES

ABBREVIATIONS AND SYMBOLS

ALCOA+	Attributable Legible Contemporaneous Original Accurate +
CPP	Critical Process Parameter
CSV	Comma-separated Values
DI	Data Integrity
EMEA	European Medicines Agency
ERP	Enterprise Resource Planning
EU	European Union
FDA	Food and Drug Administration
FIMEA	Finnish Medicines Agency
GAMP	Good Automated Manufacturing Practice
GMP	Good Manufacturing Practice
GUI	Graphical User Interface
GxP	Good ‘x’ Practice, general abbreviation for ‘Good Practice’
HMI	Human-Machine Interface
LDAP	Lightweight Directory Access Protocol
MES	Manufacturing Execution System
OPC UA	Open Platform Communication Unified Architecture
PLC	Programmable Logic Controller
QA	Quality Assurance
RTU	Remote Terminal Unit
SCADA	Supervisory Control and Data Acquisition
SQL	Structured Query Language
URS	User Requirement Specification
XML	Extended Markup Language

1 INTRODUCTION

Manufacturing of pharmaceuticals is one of the most heavily regulated industry due to products directly affecting individuals and communities. Major responsibility of ensuring patient safety lies firmly at the shop-floor-level of a manufacturing plant, where implementation of best practices and compliance is essential to maintaining efficiency and safety. Therefore, industrial control systems and the software within have strict requirements to comply with applicable local and international regulations. Compliance is achieved by demanding high standards throughout the manufacturing line, which includes many requirements for automated manufacturing from the company's personnel, processes, and control systems. The main goal of these regulations is to ensure patient safety (Mascia et al. 2013).

In modern world, manufacturing of drugs and medical devices is considered relatively safe. Some recalls of products happen from time to time, but it is most often related to wrong packaging or mislabels rather than the product being harmful. This is due to the safety work being done pre-emptively. Governing bodies also conduct inspections to manufacturing sites and they can be penalized if compliance is not met. Inability to prove sufficient compliance will lead to regulatory intervention in the form of a notification and in most severe cases even halt of production. International governing bodies can impose restrictions to exporting products in certain parts of the world. The responsibility of compliance always falls to the manufacturing process owner (FDA 2017).

Central part of compliance with regulations is how data regarding the manufactured product is handled. The regulations are somewhat ambiguous in nature and don't dictate exactly how they should be implemented. One can say that manufacturing process owner is required to keep records of most events taken place in the manufacturing process. This can be interpreted as an obvious and easy task, but the practical implementation of compliance to a sufficient degree that satisfies regulators can turn out to be time and resource consuming. The complexity of today's systems doesn't make managing the data compliantly any more convenient.

There are many established best practices in building a compliant system. Additionally, many out-of-the-box compliant solutions are in the market. Although this sounds convenient, such solutions can also be used non-compliantly. Therefore, extra attention must be targeted to implementing any solutions affecting the manufacturing process by the process owner.

This thesis evaluates COPA-DATA's Zenon, an industrial control system software, compliance to regulations and target company's local requirements. The main objective of this thesis is to evaluate how Zenon can be implemented into target company's automation system infrastructure in compliance with the regulatory demands of pharmaceutical industry. The implementation is scoped to target company's assembly process and its blueprint. Although a supervisory and data acquisition (SCADA) system generally consists of hardware and software (Macaulay, T. & Singer, B. 2016), only the software is in scope of this thesis. The thesis evaluates how the software can be used to meet demands of data integrity, availability, and accuracy. Key objective is to identify what kind of regulatory properties are built into the software and how easily it can be implemented to existing automation system infrastructure. Features like alarm management, user management, audit trail, and their execution are investigated and reflected to prevailing regulatory requirements. In addition, certain locally fixed requirements for SCADA are also investigated. The need for this thesis initiated from a machine investment project for a pharmaceutical manufacturing company where a new SCADA software was proposed.

Even though this thesis focuses merely on software interface of an industrial automation system, it is still seen to qualify as a thesis for the degree programme of mechanical engineering. The SCADA software is a vital part of the manufacturing system and it has direct impact on the processes affecting the product. The SCADA software is operated and interacted through a user interface displayed on a separate screen next to manufacturing line. The operation of sensors and actuators can be accessed through SCADA software.

A comprehensive academic literature review is conducted regarding the key subjects and applicable regulatory requirements are reviewed to evaluate the SCADA software. The

research conducted in this thesis is qualitive and a virtual test environment is established in the practical part of this thesis to investigate the SCADA software and its functionalities in depth. Afterwards, the functionalities are compared and reflected to regulatory and local requirements.

Chapters 2 presents a foundation to special characteristics of pharmaceutical manufacturing industry and its regulatory framework. Chapter 3 discusses industrial control systems in general and specifically SCADA including its subsystems, and virtualization of hardware and networks. These different technical oriented subjects listed before may seem as detached entities, but they are all interconnected in the context and scope of this thesis. Next chapter represent data integrity and the legislation within.

After a thorough academic literature review chapter 5 describes the commissioner of this thesis and the process in which the SCADA technology implementation is conducted.

Chapter 6 is the practical section of this thesis, where a SCADA software's test environment is described, and its functionalities are tested. Such features like alarm management, audit trail, user management and interfaces to other systems are investigated. The overall compliance of the software in its intended use and environment is researched and evaluation is performed based on compliance and combability.

Lastly, the results are summarized and discussed before the thesis is concluded.

2 PHARMACEUTICAL MANUFACTURING INDUSTRY

The medical industry is a sector where businesses manufacture medical devices or drugs, provide medical insurance or services to patients, and satisfies the healthcare needs of a community or individuals. The pharmaceutical industry is a component of health care system that comprises of multiple public and private organizations whose primary goal is to develop, discover, manufacture and market medicines for both human and veterinary health. The pharmaceutical industry bases its function on scientific research of pharmaceutical products that treat or prevent diseases or disorders under the supervision of jurisdictions and medicine agencies depending on geographical location. All medicines must be authorized in prior to entering the market. The objective of medicine agencies is to foster scientific excellence in evaluation and supervision of medicines to ensure the safety of patients by creating different regulations. In pharmaceutical manufacturing industry all regulations regarding the production are acknowledged in every process step or operation that influences product quality. The target company is a pharmaceutical manufacturing site thus regulatory requirements must be taken into consideration in the latter part of this thesis. Compliance of these regulations is investigated during regulatory inspections or audits. The audits are carried out internally or by an external operator (EudraLex 1999).

Due to the prevalent risks surrounding the developing and manufacturing of medical devices or drugs, regulations have evolved as a reaction to severe health defects to patients. Most famous of these tedious occurrences is the tragedy of thalidomide ($C_{13}H_{10}N_2O_4$) which took place in the 1960s. Thalidomide-based drugs were prescribed to pregnant women as an antiemetic to combat morning sickness and as sleeping aid. Unfortunately, it came apparent that thalidomide-based drugs caused severe birth defects and malformations such as phocomelia in approximately 10 000 newborns and lead to an iatroepidemic. Ban on sales was imposed and the product was withdrawn from the markets at that time. Causation of inappropriate research of the drug's health effects and nonexistent regulations was evident. Not so surprisingly, this thalidomide incident marks as a turning point in today's stricter regulation and governing of developing and manufacturing of drugs (Kim, J. H., & Scialli, A. R. 2011).

Furthermore, there are also many other possible reasons for health risks to arise besides understudied drugs. The health risks may be caused by user level misuse, mix-up of raw materials in manufacturing, contamination of the agent or using contaminated injectables. As many active substances have a rather narrow therapeutic window of use and are very potent, health risks may arise if a wrong amount of the active substance gets to the product. In addition, there is a potential risk of cross-contamination in manufacturing facilities where different active substances are handled (FDA 2021).

To prevent these kinds of tragic occurrences, there is a wide array of different regulations and different regulatory bodies formed in today's world. Generally, the regulations start all the way from the early steps of drug development and end to post-market actions. Most interesting and relevant regarding the scope of this thesis is the regulation focusing to the manufacturing phase of the pharmaceutical products particularly in the European Union (EU) and the United States of America.

2.1 Regulatory agencies and requirements

The target site is located in Finland thus the operations are monitored by Finnish Medicines Agency (Fimea). Fimea operates under the Ministry of Social Affairs and Health and acts as the national competent authority for regulating and licensing the pharmaceutical sector. Approximately 28,000 marketing authorization cases for medicines are processed at Fimea and about 170 inspections for pharmaceutical plants, laboratories and are made annually by 250 employees. Fimea's main objective is to promote rational use of pharmaceuticals to improve health of communities and people. Fimea is also a part of the European regulatory network (Fimea 2018). EMA is the European Medicines Agency and it acts as the regulatory authority in Europe. The national regulatory authorities in Europe are in close collaboration with EMA working together towards mutual goals. (EMA 2021).

The target site supplies products also to the United States which has its own pharmaceutical regulation authority. The U.S. Food and Drug Administration (FDA) is a federal agency founded in 1906 and it acts under the Department of Health and Human Services. Headquartered in Maryland, FDA oversees the pharmaceutical regulation,

guidance, and product approval nationwide. FDA also conducts inspections and audits to pharmaceutical manufacturing plants that supply products to the US therefore the target company is subject to these audits and compliance. In addition to regulating pharmaceuticals, FDA also fosters food supply, medical devices, tobacco products and naturally emerging public health threats (FDA 2017).

In addition, there are hundreds of smaller governing bodies or agencies around the world that regulate the pharmaceutical sector. The regulation set in major above-mentioned organizations descend to the smaller regulators who make minor adjustments according to local needs and policies. Practices regarding pharmaceutical manufacturing are somewhat objective and are applicable as such in major part of the world.

2.2 Good manufacturing practice

GxP is a collection of quality best practices and international guidelines initially established by U.S. Food and Drug Administration to ensure the safety of pharmaceutical operations such as manufacturing, control, storage, clinical trials and distribution. GxP stands for Good 'x' Practice where letter 'x' is a variant and refers to varying field; for example, GMP stands for Good Manufacturing Practice, and GLP stands for good laboratory practice. These regulations have the force of law, requiring the manufacturer their products safely, purely and effectively. To achieve compliance, manufacturers need a quality approach to manufacturing minimizing chance of contamination, mix-ups, or other errors. Failure to do so can result in very serious consequences including batch recalls, health defects to patients, fines and ultimately jail time (EMA 2021b) (EudraLex 1999) (EudraLex 2010).

GMP is often referred to as cGMP where "c" stands for "current" reminding manufactures to use up-to-date technologies and systems to keep up with the regulation. Production equipment that was commissioned 15 years ago and was state-of-the-art then may not be compliant with today's regulatory demands. To further emphasize this issue in the digitalized world, the GMP has evolved to cGMP (Gouveia, B. G et al. 2015).

cGMP guidelines focus on traceability, accountability, and data integrity. It is of great importance to be able to track down every step taken in the process and trace the action to a single product or batch and the people who operated these activities over the product's life cycle with end-to-end transparency. This helps to identify potential errors occurred in different phases of the process in case of recalls. The origin of an error is fundamentally important to investigate so proper adjustments can be done in advance. Accountability answers to the question who affected to the manufacturing process, how and when. Accountability complements traceability in many ways. Data integrity is a fundamental part of manufacturing in pharmaceutical industry and it will be presented more deeply in the next chapter. As stated in the introduction, many regulations are somewhat vague and open-ended leaving it for the manufacturer to decide individually how to best implement necessary controls and actions. Although this provides flexibility, it also requires the manufacturer to interpret the requirements in a business-case manner. As stated before, the responsibility always falls to the manufacturer. Ultimately, the main objective of cGMP guidelines is to ensure the safety of patients (FDA 2018b) (EudraLex 2010) (WHO 2016).

Different organizations have their own set of GMP guidelines, but they all share the same basic principles with minor differences. FDA GMP regulation covers the U.S. and EMA GMP covers the European Union area. In addition, World Health Organization (WHO) GMP is applicable mostly in developing countries. (WHO 2016) A big part of pharmaceutical manufacturers exports their products abroad and they need to comply with most of the jurisdictions around the world. If there are signs of uncompliant action, export bans can be imposed by the governing agency. Inspectors from various governing agencies perform audits to pharmaceutical companies to assure compliance to legislations, standards, and guidance. Consequently, incompliant activity is reported by the inspectors and warning letters are issued to the companies.

In consequence of regulations being somewhat vague, several established document collections representing best practices to clarify legislative requirements have been published by International Society for Pharmaceutical Engineering (ISPE). ISPE has published a guidance series called Good Automated Manufacturing Practices (GAMP) which explicitly concentrates on automated manufacturing for pharmaceutical industry

and includes many guidance documents how the regulatory demands can be met in practice. GAMP is neither a law nor regulation but a voluntary set of guidance documents for the industry based on best practices and standards in the industry. The set of guidance documents has been compiled by experts in the pharmaceutical manufacturing industry. In addition, the officials working for the regulative bodies in EU and the US have also contributed to the content of the documents. GAMP documentation was originally founded in Great Britain to further tackle and understand the demands set by the FDA. ISPE acquired GAMP documentation in 2000 and since then it has risen globally to a standard best practice documentation in many pharmaceutical companies in such an impactful way that regulative officials often reference to GAMP documents in their instructions (ISPE 2021). In scope of this thesis the most important guidelines and best practices focus on electronic records and electronic signatures. Many of ISPE's guidance documents are used as reference in this thesis.

2.3 Validation and qualification

Quality assurance (QA) consists of validation and qualification. Generally, qualification means acceptance or capability of approval and validation is described as quality assurance of repeatable action. The term qualification is used for systems or equipment opposed to validation for process. Both validation and qualification are an important part of any company's quality management. In pharmaceutical manufacturing qualification aims to create documented evidence that premises, systems, or equipment can achieve the predetermined specifications properly installed and work as intended and lead to expected results. The system or equipment is subject to design qualification (DQ), installation qualification (IQ), and operational qualification (OQ). The performance of this system or equipment is then determined by performance qualification (PQ). Process validation is the next step after qualification of equipment. Process validation is the act of documenting evidence that the process, procedure, or method actually and consistently leads to expected results. These different phases of qualification and validation build the foundation for achieving compliant manufacturing of pharmaceutical medical devices. As a result, a validation report is created summarizing the activities performed and it is often the first document of a system to be examined during regulatory inspections. When a system or procedure has achieved its validated state, any changes to the system are

managed by a process called change management. When any changes to any process or procedure is made it is reported by change control procedure and it is approved by the company authority. In addition, change control is precisely documented in case of failures through the made changes (ISPE 2008).

Every automated manufacturing system should have a validation master plan covering the areas of patient safety, data integrity and effect on product quality. The validation master plan contains high-level planning for the validation of the system defining the required activities, responsibilities, and requirements for acceptance. Engineering of this document should be started roughly at the same time as the development of the user requirements specification (URS) (ISPE 2008).

Since SCADA software is running in production environment and its functionality can affect the product quality, it is subject to validation and qualification. The target company has agreed locally that an individual SCADA software can be validated as a part of the manufacturing system qualification rather than as a standalone GxP computerized system, which has its own process for validation. The above-mentioned procedure for computer systems regarded as a component of automated manufacturing equipment is a best practice in the industry. However, this solution requires in-depth understanding of the process and the product and that the process parameters can be accurately and reliably predicted and controlled. In addition, documentation regarding this issue must be adequately demonstrated (ISPE 2008).

3 INDUSTRIAL CONTROL SYSTEMS

Industrial control system (ICS) is a higher-level term for describing controls and associated instrumentation such as networks, devices and systems used to monitor and control automated industrial processes. Due to strong linkage with automation and ICSs, The International Society of Automation (ISA) call modern ICSs industrial automation control systems (IACS). These control systems have been developed to make labor more safe, efficient and optimized. There are multiple types of ICSs but most common of them are Supervisory Control and Data Acquisition (SCADA) systems, Distributed Control Systems (DCS) and Process Control Systems (PCS). The term also overlaps more business-focused systems, such as Enterprise Resource Planning (ERP) and Manufacturing Execution System (MES). In the context of this thesis we are mostly interested in SCADA software and its interfaces to other systems (B. Galloway and G. P. Hancke. 2013).

3.1 Automation system pyramid

Industrial control systems are divided into five separate layers according to ISA-95 standard (ISA 2019) and multiple another source material shown in figure 1. Gathered information is communicated upwards from lower levels while planning and control instructions are sent downwards. The lower a level is, the more real-time computational capability is needed.

Starting from the bottom there is field level where devices such as robots and sensors interact directly with the actual process. Data is gathered and sent to upper level. The next level is control level where field devices receive signals from and send signals via I/O logic to control the process in real-time. This level consists typically from a programmable logic controller (PLC) or any other kind of a controller and is usually linked to one physical location in order to achieve real-time control and reliable data transfer with an emphasis on reliability. Surpassing the control level is supervisory level, where SCADA system gathers data over the lower levels and analyzes it. Next level is called the planning level, where MES manages and monitors the process from the raw

materials to the finished product. This allows management to get a clear picture what is happening with the manufacturing process and allows them to adjust shipment plans or make decisions based on information collected in prior levels. The top of the pyramid model is called the management level. ERP is used to oversee the company's different business-related processes such as inventory, finance, marketing, and customer relationship. Integration of ERP promotes efficiency and provides transparency in the company by keeping everyone in the same page.

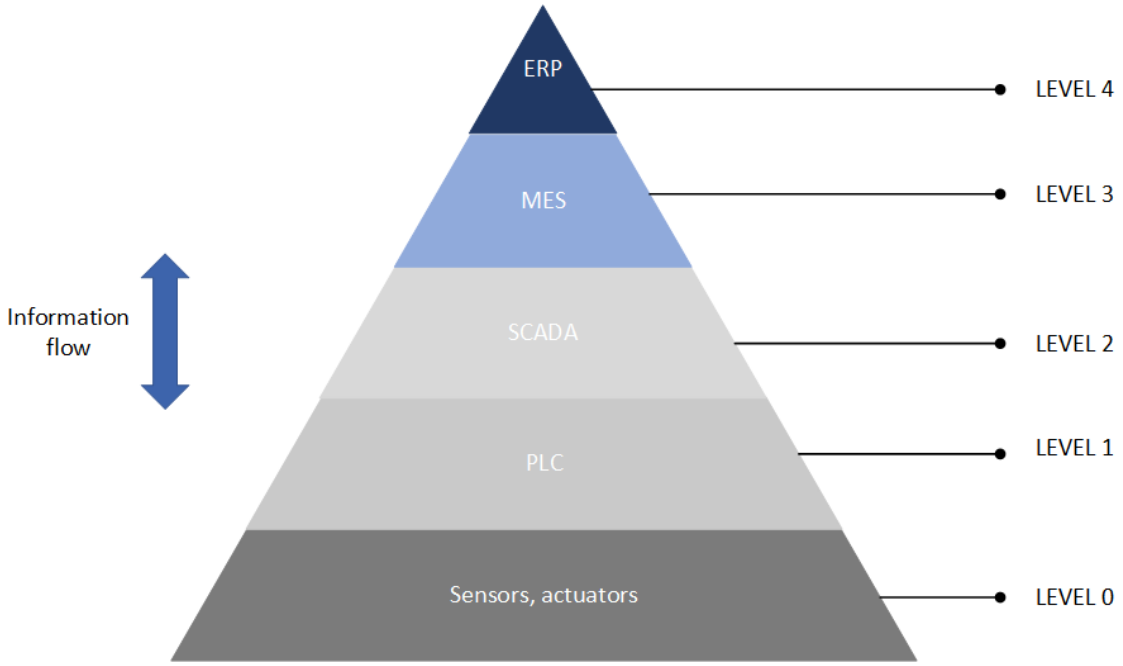

Figure 1. Industrial control systems represented in a multi-level functional hierarchical model (retell from Engell S. and Harjunkoski I., 2012)

With the introduction of Industry 4.0 and even Industry 5.0 (Industry 5.0. 2020), the hierarchical pyramid model of ICSs as such are seen to be disintegrated due to interconnectivity and the increased capability of lower level components. However, this distinct level-like thinking of ICSs is not vanishing in the near future, but the communication between these levels is becoming more and more flexible as time goes by (D. Schulz 2015).

ICSs have been developed to control performance-critical physical systems thus they have different performance and security requirements than purely software-based systems. A failure in an ICS can lead to devastating damage to personnel, environment, and the system. On the contrary, a failure in a software system may lead to inconvenience or a delay at its worst. ICSs are designed to run extensive periods of times and any failure or system stop lead to financial losses. In consequence, ICSs have far more strict requirements when it comes to real-time performance and reliability (Macaulay, T. & Singer, B. 2016).

Industrial automation systems also have long life cycles. This is due to high-class design quality and more strict requirements implemented successfully. The systems can run for several decades but they also must be maintained. This can prove troublesome due to rapidly developing technologies. Support for decade old hardware can become unavailable and the same goes for software. New software and hardware can be implemented to older systems, but they need additional integration work in order to operate reliably (Macaulay, T. & Singer, B. 2016).

3.2 Supervisory control and data acquisition

SCADA is the most commonly used subgroup of ICSs. As the name suggests, it is more data-driven system architecture of process data gathering from the field level devices and supervising rather than actively controlling the actuators and sensors in the field. Modern SCADA systems have been described as open-loop systems overseeing the process and providing the operator a complete overview of the process. However, there can also be some closed-loop control elements present. SCADA systems have been traditionally used in systems that are geographically wide, such as power station control, gas refining and transportation, telecommunications and water and waste control. In today's data driven world, SCADA systems are also used in local industrial automation systems in manufacturing sites due to their sophisticated data acquisition, data trending, advanced alarming, and connectivity capabilities. Looking back at the first SCADA system conceptualized in the early 1960s, many of these functions were not possible. For a long time, SCADA systems were lacking connections to other systems in the facility but last decades' rapid technology development and attitude change has made them more and

more interconnected. SCADA is popular due to its compatibility and reliability. Additionally, SCADA applications can range from supervising a temperature of one boiler tank in process control system to overseeing a dozen of geographically apart nuclear power plants (R. Radvanovsky & J. Brodsky 2013) (Mehta, B. R. et al. 2015).

In general terms, SCADA system architecture has four main functionalities which will be represented next. Additionally, the general architecture and components of SCADA system is depicted. Lastly, a brief introduction to current trend in SCADA technology is represented.

3.2.1 Data acquisition

SCADA systems can include hundreds or even thousands of field instrumentation devices such as actuators, valves, sensors, and drivers. All these measure inputs or outputs of the system. Inputs and outputs can be divided into analog and digital categories. In a case of digital input, a sensor can detect events by a straightforward on-off-switch. Generally, digital inputs and outputs are used to measure simple states. Digital inputs can also measure more complex solutions if needed and they are generally used in industrial control applications. On the contrary, analog devices detect continuous changes in voltage or current input and can also be used to measure more complex situations. For example, an analog sensor can detect the fluid level and temperature in a boiler. In analog applications, there is usually a normal range defined by bottom and top level. These field instrumentation devices produce great amount of data which can be presented on the SCADA system (Macaulay, T. & Singer, B. 2016).

In modern world, SCADA systems require real-time data visualization which means that the data must be collected instantly by a computer. In the field level, data acquisition involves an input scanner or analog-to-digital converter so that the data signals are transmitted from the devices. Data acquisition systems can be also used as a feedback control loop in process control systems to provide direct digital control of the system. Data acquisition is mostly used in SCADA systems to inform the operator about events and alarms, show critical process parameters on the display, and data logging for data analytics purposes (Yadav, G., & Paul, K. 2021) (Macaulay, T. & Singer, B. 2016).

3.2.2 Communication of data

To have the generated data from field devices communicated to the SCADA system, a sufficient communications network within the system is necessary. Early SCADA applications used radio, modem or dedicated serial lines as a method of communication, but nowadays modern SCADA systems communicate through Ethernet or even over the Internet. There are many industry standard protocols or manufacturer proprietary protocols. The controller in SCADA system operates autonomously on real-time basis using the latest command from the supervisory system. Consequently, the communication infrastructure is fundamentally important. Although, a disconnection with SCADA and the controller does not necessarily stop the plant process controls and on resumption of communications, the operator can continue with monitoring and controlling of the system. Many critical systems have built-in redundancy to overcome sudden communication breaks (Yadav, G., & Paul, K. 2021) (Macaulay, T. & Singer, B. 2016).

Older SCADA systems used closed proprietary protocols, but current SCADA systems facilitate more open and standard protocols. The controller in the SCADA architecture must have an interface between the field instrumentation and SCADA network. The controller encodes the field inputs into protocol format and forwards it to the SCADA control server. The SCADA systems usually have interfaces to other industrial control systems such as MES (Zhang, P. 2010) (B. Galloway & G. P. Hancke. 2013).

3.2.3 Data visualization

The most important data that is visualized in a SCADA system is alarms. Alarm is generated through a course of actions where a predefined setpoint is outside of its normal operating range thus creating an alarm. This alerts the operator to make necessary changes in the system via a siren or pop-up box on a screen. SCADA master station or human-machine-interface (HMI) also provides the operator a comprehensive view of the entire managed system. HMI also gives the operator access to change the parameters in the process. Additionally, the HMI can be linked to a database for providing trend data of the process, diagnostic data, or a view of maintenance schedule on the system, and detailed schematics of the process (Zhang, P. 2010) (Macaulay, T. & Singer, B. 2016).

3.2.4 System control

SCADA systems can control many kinds of industrial automation processes automatically. For example, in process automation a SCADA system can automatically open a valve in a boiler if a pressure sensor alerts of too high pressure. The operator can also adjust the valve manually through the operator screen. All in all, a SCADA system can make supervisory decisions based on plethora of inputs. The SCADA system usually provides control of the process through different means. Operators can change predefined setpoints for process parameters from the user interface. Depending on the application, operators can have wide system control possibilities, but usually they are restricted to enabling or disabling actuators and controlling the setpoints (Zhang, P. 2010).

3.2.5 Architecture and components

A general SCADA architecture begins with some kind of a controller that communicates with a wide array of objects, such as sensors and actuators sometimes referred to as intelligent electronic devices. The controller then routes this information to a computer running SCADA software by different means of communication depending on the application. This software processes, distributes, and displays the data to help the operators and other employees can analyze the data. Data such as alarms, measurement results and events can be easily presented in a graphical user interface (GUI) in SCADA software. A topology diagram of a case-specific SCADA system is depicted in chapter 5 figure 2.

SCADA system usually consists of the following main elements:

i) **Supervisory computer** is the core of the SCADA system. Supervisory computer refers to the computer and software that gathers the data from the controller on the field and sends control commands back. Supervisory computer can also be referred to as master terminal unit (MTU). In many modern applications, supervisory computer is composed of a single computer simultaneously acting as Human-machine-interface (HMI) for the SCADA software. In larger and more complex applications, the supervisory computer provides several HMIs hosted on client computers. In a modern application,

the supervisory computer can be a virtual machine with the SCADA software installed and the HMI is deployed as a runtime client to the computer located in an operator room.

ii) **HMI** is an important part of a SCADA system. HMIs are used by operators to monitor the process and adjust e.g. process parameters, control setpoints and it is usually located next to the actual physical machine. HMI allows for the control engineer to configure control algorithms and parameters in the controller. Live data of the controlled process is shown in many different graphical shapes. The visualization methods are highly customizable and sophisticated. For example, in a water supply system a valve can be shown as open or closed depending on its current digital state received from the field. The valve can also be configured to flash as a red icon if the system detects errors in the valve. Any abnormal conditions detected by the field instrumentation are registered at the central host as alarms. HMI can also represent all the latest analog values on the screen as some physical representation, such as a fluid level of a boiler monitored by a sensor. Additionally, the alarms are also presented in a dedicated alarm list notifying the operator. Alarms are usually depicted by pop-up boxes or flashy colors, sometimes even a siren. Depending on the criticality of the alarm, it usually must be acknowledged before continuing the process. If there are variables changing over time, the HMI can provide a trending graph to follow their behavior and allowing a human friendly view of the changes happening in the process.

iii) **Programmable Logic Controller (PLC) or Remote Terminal Unit (RTU)** is used as a controller in various types of industrial processes. They are connected to a SCADA system using feedback loops with sensors and actuators through fieldbus protocols to provide local management of processes and meet configured setpoints. PLC uses digital and analog inputs and outputs to control the processes. PLCs have an internal memory for storing instructions for the purpose of implementing specific functions such as logic, timing, counting, proportional-integral-derivate control, communication and data and file processing. Due to emphasis on real-time controlling, PLC has high-speed connection to SCADA. Also, PLC can perform individually in

less demanding and simpler control setups, but its data acquisition and alarm management properties are rather limited. PLC can be described as a specialized computer targeted to industrial automation with high focus on reliability and it can be accessed via a programming interface in an engineering workstation. RTUs in the other hand are mostly used in geographically widespread SCADA applications to collect real-time data from sensors and actuators and forward them to the supervisory computer. Due to absence of local infrastructure networks in remote locations, RTUs often use small solar panels or batteries as a powering system and they communicate via radio, GSM or satellite. RTUs are also ruggedized to withstand extreme conditions. In context of SCADA, the line between modern RTUs and PLCs has blurred, and the terminology is virtually interchangeable.

iv) **Communication infrastructure** connects the PLC or RTU to the master SCADA station. Control actions that are performed at the master station are treated as data sent to the RTU or PLC thus any control action initiates a communication link PLC or RTU allowing the command to be sent to field instrumentation. Modern SCADA solutions employ several layers of checking mechanisms to ensure that the transmitted command is received. The communication network can utilize wireless or wired technology. In geographically circulated areas, wireless transmission of data is used to communicate effortlessly. Traditionally, legacy automation communication protocols such as Modbus and Profinet have been used. In local applications, local area network technologies are used because of reliability and high-speed compared to long distance communication systems.

v) **SCADA software** is an important part of every SCADA systems. Depending on the application and the vendor, the software can be a significant cost item. A well designed and engineered SCADA software is the foundation of successful SCADA system. There is a wide range of commercial SCADA software available and many of them are highly configurable to suit different applications. Especially modern SCADA software developers aim for their product to be flexible and compatible with different types of hardware and software allowing the whole SCADA system to be non-vendor specific.

The general architecture of SCADA can vary greatly among applications based on the sophistication and setup of the individual system, but the main principle stays the same. Due to the trend of openness in SCADA systems, Ethernet and TCP/IP-based protocols are replacing the older proprietary protocols. There have been few characteristics of frame-based network communication such as determinism, synchronization, protocol selection and environment suitability that have restricted the adoption of Ethernet in few specific applications. However, it is now accepted in majority of different applications (Zhang, P. 2010) (Macaulay, T. & Singer, B. 2016).

3.2.6 Current trend

Modern SCADA systems have evolved from closed proprietary system into more sophisticated and complex open systems during the last decade. With the introduction of digitalization in manufacturing systems mainly by the philosophy of Industry 4.0, the current trend is pushing SCADA systems to integrate solutions like Internet of Things (IoT), cloud computing and big data analytics. Even before the introduction of IoT, SCADA systems and IoT already shared few characteristics e.g. access of data and visualization. However, where IoT differs is interoperability, scalability, and capability of big data analytics. IoT-based SCADA systems rely on effective information collection, analyzing and displaying data from heterogeneous devices using standardized protocols and wired or wireless network. The collected data is then stored and analyzed in cloud-based system. Moreover, open communications standards are used to collect and control data. The more connected nature of SCADA systems today has also exposed the systems to vulnerabilities (Yadav, G., & Paul, K. (2021).

With the implementation of IoT technologies in SCADA systems, a series of standards have been introduced to provide a homogenous system development utilizing open communication protocols. In the context of this thesis, key concepts of Open Platform Communication Unified Architecture (OPC UA) are briefly introduced in latter part of this thesis.

3.3 Virtualization in industrial environment

Virtualization is a term that encompasses various methods of creating virtual versions of something. In the context of this thesis, mainly complete virtual machines (VM) and network virtualization are seen relevant to be briefly introduced. Virtualization has been used in traditional ICT for a long time, but it has been increasing its use in industrial environments.

Hardware virtualization is made possible by software called the hypervisor which creates a virtualization layer on top of the physical hardware. Hypervisor communicates directly with the physical server's disc space and central processing unit (CPU) to manage the VM. The hardware is fully virtualized thus the virtualization of complete computers running operating systems is possible. VM consists of virtual hardware that the hypervisor has allocated and uses its dedicated share of physical resources, such as network devices, CPU, random access memory (RAM) and hard drive space. VM's can interact seamlessly with their physical counterparts. Multiple VMs can also be deployed on a single computer (Stouffer, Keith & Falco, Joseph & Kent, Karen. 2007).

Network virtualization is the act of decoupling a network from its physical equipment. Virtual networks can be created by altering network topology without changing the layout of physical network components. Virtual network can also be created entirely virtually of software components and VMs inside a server. The main driver for using virtual networks is to isolate existing traffic in different zones as a security precaution and higher utilization of the network. Additionally, management and design of networks is more flexible and controllable when using software to change the network's logical structure (Stouffer, Keith & Falco, Joseph & Kent, Karen. 2007).

The main advantages of virtualization in context of ICSs is security measures and adding redundancy. ICSs have very strict security requirements in regards of access. Virtualization presents great benefits to achieving isolation of systems and ease of maintenance. Additionally, virtualization can reduce points of failure and sectioning networks virtually in manufacturing environment is an efficient security hardening measure. The security of networks where ICSs reside can be easily hardened with creation

of logical subnets to the network infrastructure. The target company enables virtualization in their automation system blueprint (Macaulay, T. & Singer, B. 2016).

4 DATA INTEGRITY

Due to the growing use of information technology and computerized systems in the today's world, the digitalization has also impacted the pharmaceutical manufacturing industry (Markarian J. 2018). The integrity of critical records, data, decisions, and aspects concerned with physical attributes of the product ultimately affect patient safety. The term Data Integrity (DI) refers to the accuracy and validity of data over its lifecycle and it is fundamental in the pharmaceutical industry. DI has been the core focus in pharmaceutical industry in recent years and ensuring compliance has become more challenging with the implementation of more complex computerized systems. Many legislation and guidance documents regarding DI have been published recently by the FDA, UK Medicines & Healthcare products Regulatory Agency (MHRA) and WHO to clarify the regulatory framework and to help companies better understand them (Chan Wai Lah 2020).

However, the number of DI violation letters issued by FDA has quintupled from 2014 to 2017. In addition, large pharmaceutical companies have been cited for falsifying manufacturing related records, such as quality control results thus the effect of recent legislation and guidance documents regarding DI remains yet to be seen. The reason for such increase in DI violations may be due to pharmaceutical inspectors proactively searching for them, inspectors are better trained in DI issues, companies are taking risks in violating DI, or ignorance and carelessness of the operating staff. There has also been a rising prevalence of outsourcing manufacturing processes to third world countries to maximize productivity and business proficiency. In addition, the protocols used by parent companies in maintaining DI compliance have not been adopted by their subsidiary companies. Overall, there are plethora of reasons for the exponential increase in DI violations. The root cause is not easily analyzed because it appears to be such recent development (Chan Wai Lah 2020).

4.1 ALCOA+ standard

In order to regulate DI and security to achieve better processes and higher quality products, a standard called ALCOA was introduced by the FDA in 1990, although the

requirements for governing data have been in regulations for a much longer period of time. In pursue of keeping data consistent also a matter of data security come into play and even though data integrity and data security are somewhat overlapping they shouldn't be mistaken for one another. Other governing agencies have also slowly adapted this principle (FDA 2017).

According to the ALCOA principle data should pertain five qualities to maintain data integrity: Attributable, Legible, Contemporaneous, Original and Accurate. However, the original ALCOA has been updated to ALCOA+ adding four additions to the original framework of principles: Complete, Consistent, Enduring and Available. Any data that no longer fulfils these criteria is considered falsified, even if it's human error or deliberate. ALCOA+ is targeted towards GMP records and is applicable with electronic data but also with paper-based and hybrid applications. This thesis focuses on electronic records although ALOOA+ standard may differ depending on the type of data. The main principles of ALCOA+ are presented below, and a descriptive question of the said principle is conducted to better understand the meaning (ISPE 2019).

Attributable – When creating data, the input should be attributed to the person who generated it. This should include details who was the person who performed the action and an unambiguous timestep of the activity. This activity can be done physically signing and inserting initials and dating a paper document but also electronically through a digital system. Good Documenting Practice (GDP) guidelines having a signature or alias log in order to easily determine who changed or recorded new data. Answers to question "who performed the action or acquired the data and when?".

Legible – Data should be in clearly readable form. It must be possible for the data to be read and understood years and after its recording. This is relevant both to digitally recorded data as manually recorded data in notebooks. The best approach to fulfill this need is to use consistent and straightforward language throughout the organization regardless of locality. Additionally, materials used in recording and collecting data should be durable. Answers to the question "can the data be easily read and is it indelible?".

Contemporaneous – Data should be contemporary in nature. When recording a result, measurement, or any data, it is essential that individuals or systems make record of the activity at the exact time it takes place. Dealing with electronic data this is normal practice. Date and timestamps should flow chronologically in order of execution for the data to be credible. Answers to the question "is the data documented at the time of the activity?".

Original – Data should be original or a certified copy. Original data can be also referred as source data or primary data. Original records should be preserved, meaning the materials or standards used should be durable. In case of having duplicates e.g. a true copy, the creator of the original records should confirm the authenticity of the copy. Answers to the question "is the data recorded from an original observation or a certified, true copy?".

Accurate – For data to be accurate, it should be error free, complete, truthful, and reflective of its context. If editing data, documenting, and annotating the amendment is a must. Answers to the question "is the information complete, consistent and correct?".

+

Complete – Information that is critical to recreating and understanding an event. This would include any repeat or reanalysis performed on a laboratory test sample. All data recorded should be complete auditing all changes with respect to the source of change as well as time. Answers to the question "is all data, including modifications of data, included, e.g. testing, re-analysis, processing, reprocessing?".

Consistent – Data is presented, recorded, dated or timestamped in the expected and defined sequence so that it should be possible to create a chronology or sequence of events based solely on the data. The captured sequence must match the expected sequence. Answers to the question "is there consistent generation of records and application of timestamps?".

Enduring – Data must be maintained intact and accessible throughout defined retention period. This signifies the ability to store data in reliable form and place for a long duration. Answers to the question "is data recorded in such a manner which will enable them to last for the intended duration?".

Available – Data must be made accessible when needed. This emphasizes the ability to retrieve data at any point of time and not only storing it in a proper manner. Answers to the question "is data available for review and audit during their entire lifecycle?

In the context of SCADA software and its utilization, the above-mentioned principle must be taken into consideration when dealing with alarm management, events, audit trail, user management, process data, archiving of records, and interfaces to other systems. These functionalities of the said software will be reflected to ALCOA+ principle in the practical part of this thesis (FDA 2003) (ISPE 2008) (ISPE 2019).

4.2 Legislation and guidance documents

In this thesis, regulation from the European Union (EU) EudraLex in EudraBook and the FDA are seen most relevant thus they are discussed further. The legislation regarding DI in pharmaceutical manufacturing is concluded into the GMP regulatory framework. FDA's interpretation of GMP comprises of 21 Code of Federal Regulations (CFR) 210-212, 600, and 820 while EU's comprise of Commission Directive 2003/94/EC stated in EudraLex Volume 4 (EudraLex 2003). 21 CFR 210 provides very generic legislative requirements for pharmaceutical products, mostly on their quality and purity. EU Commission Directive 2003/93/EC states a generalized overview of GMP in pharmaceutical manufacturing companies. 21 CFR 211 and EudraLex Volume 4 are almost identical promoting and regulating the required documentation for equipment protocols, labelling and distribution processes, personnel qualifications, and training through standard operating procedures (SOP), protocols for recalls and corrective and preventive actions (CAPA) (FDA 2018a). More deeply, EudraLex Volume 4 chapter 7 is dedicated to contract requirements regarding outsourced manufacturing facilities and activities, whereas 21 CFR 211 does not explicitly state such. In addition, 21 CFR 212 and 600 regulate specifically radiological and biological pharmaceutical products (FDA

2018b). They generally state that radiological and biological pharmaceutical products require more accurate and attributable information to be retained for a longer time in order to retrace products in case of recalls because of errors in manufacturing. Lastly, 21 CFR 820 enforces to ensure the quality of the product is maintained throughout the manufacturing process by specifying the documentations required to validate such processes. A clear connection with GMP framework and DI regarding the importance of documentation can be seen without further investigation (FDA 2018a).

As stated previously, the legislation tends to be very generic and open to interpretation thus collections of guidance documents have been published to clarify legislative requirements and to introduce best practices for implementation. Generally, guidance documents promote voluntary compliance and can be integrated to any pharmaceutical company's culture and manufacturing processes.

There are multiple guidance documents promoting general GMP and specific portions of it, such as FDA Data Integrity and Compliance with CGMP Guidance for Industry, ISPE GAMP® Guide: Records and Data Integrity, ISPE GAMP® RDI Good Practice Guide: Data Integrity by Design, ISPE GAMP® 5: A Risk-Based Approach to Compliant GxP Computerized Systems, MHRA GxP Data Integrity Guidance and Definitions, the Pharmaceutical Inspection Convention and Pharmaceutical Inspection Co-operation Scheme (PIC/S) Guide to Good Manufacturing Practice for Medicinal Products. All these documents provide in-depth guidance to various aspects of GMP with consideration due for the respective country's legislation. The guidance focusing on DI issues have been just recently published and updated due to increasing number of DI violations reported by the FDA.

All in all, the guidance documents and legislation seem to be quite comprehensive and up to date in terms of assuring and promoting DI. However, even the most encompassing guidance documentation is not enough to prevent DI violations alone. Major part of DI violations is discovered during audits or whistleblowing and by the time these violations surface, the non-compliant and potentially substandard pharmaceutical products have already been consumed by the patient putting their health at risk and even death. Therefore, other factors and actions must be obtained at higher levels to promote and

assure DI (Chan Wai Lah 2020). These factors are discussed generally and in context of industrial control system software in chapter 4.3.

4.3 Industrial control system data and records

Manufacturing systems have evolved tremendously in the last decade. Consequently, also the sheer number of data has increased exponentially. With today's advanced technology, it is possible to automatically track and record many kinds of data. Moreover, the companies are also pursuing a more data-oriented approach to manufacturing of pharmaceutical products which gives them great insight to quality assurance and process understanding. The focus in the regulative framework, most notably in FDA 21 CFR Part 11 shortly introduced before, is in electronic records and electronic signatures in pharmaceutical manufacturing industry. Next, some terminology surrounding data integrity in manufacturing records are explained.

Electronic records have a broad definition by the FDA, but it can be simplified as information in a digital state created, modified, archived, or used by a computerized system. In addition, electronic signature can be defined as a set of symbols that is as unique and legally binding as a handwritten signature, but that is used to sign records in a computerized system (FDA 2018a).

A copy of an electronic record, irrespective of the type of media used, can be used in GxP purposes as a "true copy" if the original record has been verified (e.g. by a dated signature) or that has been generated through a validated process to produce the copy having the exact same content and meaning of the original including the metadata. It should be possible to create a true copy of electronic data for the purpose of review, backup and archival (FDA 2018b).

The most essential electronic records in pharmaceutical automated manufacturing and the scope of this thesis are the following:

i) An **audit trail** provides documentary evidence of activities that have affected the content or meaning of a record at a specific time point. It is a secure

computer-generated, time-stamped electronic record that allows for reconstruction of the course of events relating to the creating, modification or deletion of GxP-critical records, such as process parameters, system logins, and batch start. Audit trail is a chronology of "who, what, when and why" of an electronic record. Audit trail needs to be available and convertible to human readable form. In a computerized system, audit trail should be enabled and appropriately configured to reflect the roles and responsibilities of personnel. Ability to modify audit trail settings should be restricted to authorized personnel only. Audit trails need to be available and convertible to a generally intelligible form and regularly reviewed.

ii) **Alarm and event management** is a system used to prioritize, group, and classify alerts and event notifications. It determines the function, need, priority and presentation of alarms and events to the operator. Alarms and events are used to notify the operator if setpoints are met or possible errors have arisen.

iii) **Critical Process Parameters (CPP)** are key variables affecting the product. CPP are attributes used to detect deviations in production process and product output quality. Tolerance of the acceptable limits should be set by the manufacturer according to product specifications. All data relating to CPP and their metadata should be recorded, stored, and analyzed.

iv) **Report generation**. Reports of manufactured batches or audit trails are needed for reviewing and inspection purposes. These are usually generated in text format into an external file.

v) **User data and management**. According to regulations and guidance, every individual should have personal login credentials to manufacturing systems. Also, the user role rights should be configured so that they reflect the responsibilities of the personnel. For example, a machine operator must not have administrator rights to a manufacturing system.

Other things to consider in terms of data in industrial control systems is their interfaces and communication protocols to other systems. Manufacturing systems may operate in total isolation, but majority of today's systems have interfaces where data is passed between systems e.g. SCADA to PLC and vice versa, SCADA to database, and SCADA to MES. It is therefore highly recommended to validate and qualify these kind of multi-interface systems in such a way, that data transfer between these interfaces can be stated secure and trustworthy and decisions based on them can be done in GMP purposes (ISPE 2019).

4.4 Ensuring data integrity

The reliability of data in manufacturing systems is dependent on the procedures, processes and controls used in ensuring data integrity. The integrity of production related data begins at the point of data creation and continues through its lifecycle including storage and retention to support the quality of products manufactured. Manufacturers must be able to ensure that relevant data is documented, it is attributable to the persons who captured it and that the data cannot be altered, omitted or deleted in any way to misrepresent what it actually is. DI is the cornerstone in maintaining and establishing product quality and patient safety. There are numerous ways to ensuring data integrity and the most relevant to the scope of this thesis are represented next (ISPE 2019).

4.4.1 Supplier assessment

In pharmaceutical manufacturing, it is important to make a careful supplier assessment when acquiring a manufacturing system running SCADA software. As part of the supplier evaluation process, it is important for regulated companies to assess whether the prospected supplier has taken data integrity issues and solutions into account when building the system. Generally, it is advised to use that kind of suppliers who have modeled data flows within their system and instruments. In certain cases, it may not be possible to fully map the data lifecycle or data flow to understand potential data integrity e.g. embedded control systems. In these instances, it is advisable to implement careful risk assessment and mitigation plan for the availing risks. Prevailing data integrity risks can often be solved by working with the supplier to fully understand and document the

data lifecycle and data flow map. Moreover, it is often found out that the supplier is unaware of potential data integrity problems unless this information is provided by the company (ISPE 2019).

4.4.2 Technical support features

There are several technical features of manufacturing systems and their SCADA technology that provide support for the integrity of data. In the table 1, some technical features are presented and reflected to the ALCOA+ principle discussed in chapter 4.1.

Table 1. Technical features of a manufacturing system to support ALCOA+ (ISPE 2019)

ALCOA+ Term	Technical feature
Attributable	• Restrictions of unauthorized access by security and access controls and clear segregation of duties between user right roles • Audit trail for GxP data • Ability to lock down validated configurations • Inactivity logout feature for systems giving access to GxP data or operations forcing the user to re-authenticate again
Legible	• Features to display the data in human readable form, such as sorting, trending, filtering and reporting by exception • Data audit trail presented in human readable form that allow previously used values to be seen
Contemporaneous	• Automatic time stamp of data upon capture • Time synchronization throughout the system

Original	• Automatic capture, process and archive of data • Automation of processes where a true copy is required (e.g. where data must be passed between systems or where data exists in redundant for di
Accurate	• Calibration routines for the process • format and range check for manual data entries • Support for semiautomated entries (e.g. bar core reader) in place of manual entry)
Complete	• Features to ensure that data does not become separated from metadata thus losing its context • Features that support understanding of displayed data, such as filter alarms and event history
Consistent	• Availability of standard interfaces • Time synchronization between systems
Enduring	• Archive of long-term data for the duration of the retention period
Available	• Support for backup and recovery of data • Support for automated retrieval of data • Ability to retrieve the data without affecting current data

4.4.3 Logical security

Information security controls are a key part in ensuring data integrity. Information security is well-known in traditional IT implementations, but it is extremely important in manufacturing systems. To uphold a manufacturing system's "validated" state, it is vital to control access to make changes.

Many modern manufacturing systems have user management of their own and they can be easily integrated with Lightweight Directory Access Protocol (LDAP) services to manage users, user rights and user groups through authentication against domain (ISPE 2019). Whenever possible, access permissions should be based on pre-defined roles and users should be allocated to appropriate roles. As an example, in the context of SCADA, an operator should only have access to those actions in the HMI he needs to operate the manufacturing system accordingly. Consequently, this allows actions to be restricted on a need-basis and allows actions that are tracked in audit trail to be attributable to the legally identified user. All needed user right groups and roles will be defined and described in user right specification documentation (ISPE 2017).

Despite pursuing uniquely identifiable users, some manufacturing systems have default "engineer" or "maintenance" modes and accounts in control system level, such as administrator, with default usernames and passwords. In systems that can provide equivalent access through uniquely identifiable user accounts, all the default accounts and modes should be disabled or removed. Operators and other employees using the system should be trained with proper data integrity culture subjects to fully understand the benefits of using their own credentials and not sharing accounts even if it would streamline running production. Manufacturing systems may also have active commissioning or test user accounts which are used to efficiently commission and qualify the system. This kind of accounts should be disabled prior to entering production run (ISPE 2017) (ISPE 2018).

4.4.4 Physical security

Assessing physical security begins with the access to the manufacturing site. Although, the people who have the strongest motive to manipulate data already have access to the site. In terms of industrial control systems, the amount of data stored locally in the shopfloor can be minimized. SCADA application can be executed in a way that only a runtime client is running next to the physical machinery and a proper network infrastructure enables real-time data saving to a virtual machine running the software in a server room. This way, many kinds of data integrity breaches can be evaded. With older and more simpler applications of SCADA some specially designed lockable cabinets can

be used to store control panels. However, these kind of physical access controls should be solved on the logical security level (ISPE 2017) (ISPE 2018).

5 ASSEMBLY PROCESS OF THE TARGET COMPANY

The commissioner of this thesis is one of the largest pharmaceuticals and life sciences companies in the world with focus in health care products, biotechnology, and crop sciences. In pursue of high digital maturity level throughout the production facility, the target company is enabling more automated and data-driven production. Additionally, the target site is one of five core sites in one of the leading pharmaceutical companies in the world.

The target site is one-of-a-kind production supply center and its Information Technology (IT) department acts as a local support for production, laboratories, and product development. The IT organization consists of different units such as industrial IT, business applications and IT infrastructure. Industrial IT is the only one in the scope of this thesis.

5.1 Role and responsibility of IT

The Digital Transformation and IT is responsible for the network connectivity, data gathering, and information security of the machines and devices used in production environment. The unit is actively developing processes used in production and laboratories thriving towards a digitally connected ecosystem with emphasis on IT and automation interfaces.

Team's main task is to maintain control systems of automated production machines and their interface layers between different automation levels. The unit also participates in diverse investment projects regarding industrial automation systems being responsible for their user requirement specifications covering data gathering, connectivity and cyber security.

5.2 Assembly process

The manufacturing of target company's product takes place in a clean room production environment. The purpose of clean room is to prevent contamination of the product. Clean room concept is accomplished by filtering the air entering the room and by controlling room's temperature, humidity, and pressure. The equipment used in clean rooms is designed so that surfaces attracting particles are avoided. Several tests are conducted periodically to measure particle counts and air flow. Personnel enter the clean rooms through an air lock, and they must always use high-grade protective clothing.

The assembly process is a central part in the target company's large manufacturing chain of the final product. In the assembly phase key components are connected to each other and the product is transferred onwards to the next process step. The machine operates every step of the process automatically apart from feeding material into different stations. Every component of the product has strict tolerances and demands precision mechanics while executing the process. The necessary steps taken in the automated assembly process are visualized in figure 2:

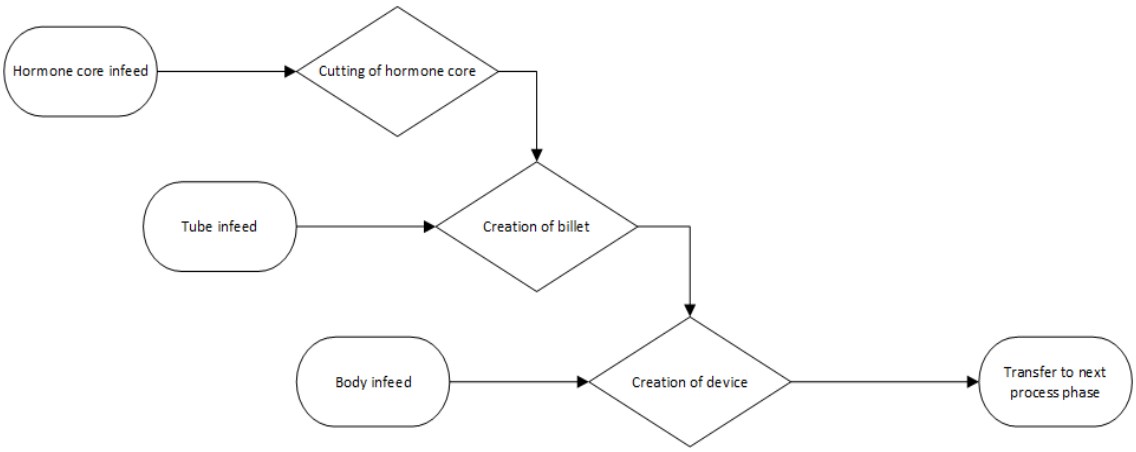

Figure 2. Process chart of the product assembly.

The industrial automation system carries specific sensor inspections in different phases of the process and stores data in the database in synchronization with the manufacturing sequence. Products that don't meet product specification are rejected to separate rejection

stations. The machine is operated through an HMI by an operator. HMI is running a client runtime of a SCADA software and it allows parameter configurations and certain other control possibilities, such as setpoint. Separate screens are also provided for displaying critical process parameters and a statistical process control (SPC) trend view. The generalized automation system topology for assembly process is depicted in figure 3.

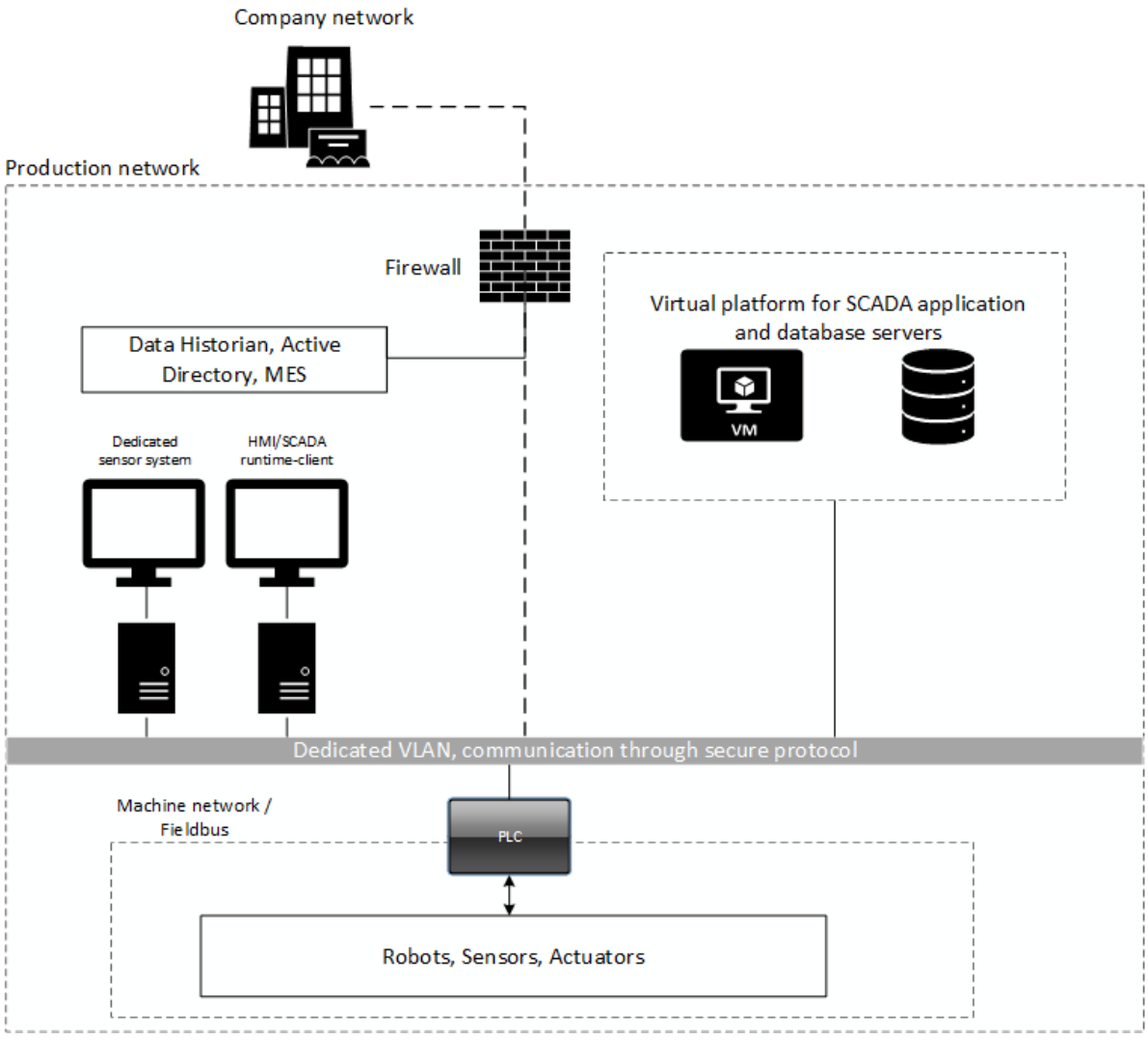

Figure 3. Control system topology in an automation system.

The whole automation system benefits from virtualization of hardware and network. The infrastructure is built in such a way that maintenance and managing of records is effortless. The data is kept off the shop floor by utilizing a server-client model for the

SCADA software for security and data integrity reasons. The servers are only accessible by authorized Industrial IT personnel.

6 CASE STUDY: EVALUATION OF PROPOSED CONTROL SYSTEM FOR TARGET COMPANY

This chapter begins the practical part of this thesis. The proposed SCADA software is installed on a dedicated VMware vSphere virtual platform simulating a real use case in the target company. The SCADA software's main features are evaluated and reflected to the regulatory framework of pharmaceutical manufacturing and data integrity presented in chapter 4, mostly adherence to FDA 21 CFR Part 11. In addition, the target company has its own locally set requirements and needs or wants for the SCADA software thus they are also discussed. A trial license for the software was kindly provided by the software company developing the SCADA software along with online training lessons and access to manuals. The practical implementation relies heavily on the material provided by the vendor.

6.1 Proposed software

The proposed software is called Zenon and it is a product of an Austrian technology company called COPA-DATA. Founded in 1987, their product is widely used in different applications of industrial automation, the most famous use case being the automotive industry. The Zenon software platform consists of different parts but only the Supervisor version 8.20 i.e. HMI/SCADA software is in the scope of this thesis. COPA-DATA has published different versions of Zenon tailored to specific industries, such as energy, automotive and pharma (Zenon 2021). However, the Supervisor version is used in this thesis to cover the main functionalities regarding regulations concerning pharmaceutical manufacturing which the software has to offer. The supervisor consists of two main components: Editor and Runtime. Editor is an administration module used to engineer individual projects, create user interfaces, and configure communication protocols among many other things. Afterwards, the runtime module is executed allowing for the previously engineered project to be used by the end users. Zenon software comes with a demo project demonstrating the key functionalities of the software. The scope of this thesis is not in programming an individual project thus the demo project is used as a solid basis for investigating the software's features.

The Zenon software claims to adhere to FDA 21 CFR Part 11 regulations regarding electronic records and electronic signatures through internal functionalities and that any project can be configured to be compliant at any stage of the project. To achieve compliance in an individual project, it is claimed to be as simple as selecting necessary options to activate user administration, audit trail, authorizations, and alarm management. As mentioned before, the target company also has locally set requirements for manufacturing systems with SCADA thus they are also discussed. Locally set requirements include using a Structured Query Language (SQL) relational database for archiving of GxP data, connection to Active Directory (AD) for centralized user management, and the use of Open Platform Communication Unified Access (OPC UA) as a communication protocol.

6.2 User management

Access to the operating runtime must be authorized by providing valid credentials. Zenon has a local user administration feature where users are established both in runtime and editor. The concept of Zenon user administration is based on different users having different operating rights in the system (i.e. authorization levels and function authorizations). Users and user groups, passwords and authorizations are defined in the editor. It is also possible for administrators to create users and issue rights in the runtime. Zenon has built-in procedures to enforce password aging and forcing system lockouts when unauthorized access is requested. Zenon also has the functionality to logout after a manually specified time to restrict unauthorized use of the machine. There are several settings under user administration that must be (de)activated in the project properties such as:

- Deletion of a user. This shall never be activated in GMP regulated environment because it creates a big contradiction with data integrity compliance. Deletion of a user will affect the traceability of data. If a user would be deleted, a modification of critical process parameters or login made by the user showing in audit trail would not be able to be traced back thus jeopardizing the integrity of the data.

- Maximum user identification errors. The system will lock out the user after a defined set of unsuccessful user identification attempts. The lockout time period can be defined. An administrator can unlock a locked user manually.
- Automatic logout of a user. If during a defined time period there is no operation on the machine interface, an automatic time triggered logout can be engineered to the project. Although, a method of merely locking the screen would be preferable instead of a logout which may interrupt on-going production run.

6.2.1 Authorization levels for segregation of duties

The user can be attributed authorization levels directly or the user can be attributed a user group, where authorization levels have been defined to groups organized by functionality. After attributing the authorization levels to a user or a group, only persons with the specific authorization level can execute the specific functionality. Different authorization levels attributed to user groups promote segregation of duties e.g. 'Operator' user group is attributed to authorization levels 1-20 and their corresponding functionalities, 'Engineers' have specific authorization levels of 70-80 and 'SuperAdmin' can have all the authorization levels attributed. Multiple groups can be selected therefore an 'Engineer' can obtain the functionalities of an 'Operator' and 'Engineer' attributed. There is a total of 128 (from 0 to 127) different authorization levels that can be configured. The functionalities within a project must be restricted to certain users to limit the range of activity and force certain workflow of operations. Authorization levels can be also assigned to dynamic elements to permit operations to be performed. The Zenon version 8.20 seems to have sufficient access restriction security capabilities to adhere to current prevailing regulations for restricting unauthorized access and use of the system. In addition, segregation of duties by attributing authorization levels is self-explanatory.

6.2.2 Connectivity to centralized user management software

Additionally, Zenon has a native support for Microsoft AD to manage user accounts centrally. The users in AD receive their corresponding user rights configured in Zenon. The communication between AD and Zenon is based on the user group name in AD and user group name in Zenon being the same thus attributing the user group authorization

levels. The settings regarding passwords e.g. period of validity, length, maximum of errors can be published to users via a group policy setting in AD.

6.2.3 Benefits of centralized user management

There are many benefits to managing user accounts centrally instead of local user management. In a modern factory, the industrial automation systems even within one department rarely are homogeneous but most of them can be connected to a service like AD. This removes much of the impractical burden for operators and other staff remembering different user credentials to different manufacturing systems. It is also much more convenient to have only the one or two credentials for logging into all systems the corporate has to offer. In terms of managing user accounts, the centralized model makes it far more agile. To make modifications to user accounts in a non-centralized model, it would be mandatory to make the changes through the local HMI thus interrupting the on-going production. In addition, if there are many systems alike, the procedure would have to be repeated accordingly. Also, any changes to user accounts i.e. adding a new operator or an engineer to the machine user interface, the same procedure mentioned above should be performed. This kind of behavior is very impractical and time and resource consuming in production environment where the on-going production is priority number one. With a centralized user management solution assigning modifying the user rights in the machine can be done with ease.

6.3 Audit trail and alarm management

All system events and user operations i.e. audit trail can be logged and displayed in runtime with a Chronological Event List (CEL) in Zenon. CEL records user related activity such as logins and logouts and start of the system. Additionally, critical process parameter changes can be configured to be recorded. The fields that are recorded in the Zenon audit trail can be modified in the project (i.e. in the Editor).

6.3.1 Audit trail functionality

By default, CEL records the date and time of the event, username, variable name, inserted and modified values. In case of modifying values, the audit trail should register the old

and the new value thus displaying them both. CEL and alarm management can be used independently but, in this case, they are combined in CEL. Additionally, the audit trail can be configured to display further information such as system name, project version/name, variable status, alarm area, alarm text, alarm class, variable information and user's full name. When a specific information is displayed on the CEL, the system can be configured to allow users add comments alongside the captured information. Zenon runtime CEL can also capture automated events and system events in the audit trail. The CEL can also filter the information in the runtime according to user's needs.

6.3.2 Time synchronization

Zenon uses built-in clock synchronization with the server operating system ensuring accurate recording of all date and time stamps in the audit trail. The time zone for the local time and time format used in runtime is taken over from the Windows regional settings. By default, the CEL does not show the time zone in the time stamp displayed in runtime. To prevent data integrity breaches concerning the audit trail, access should be restricted on the Windows user level. This is the responsibility of the customer using Zenon product to ensure any manipulation to the Windows level time and regional settings is restricted from unauthorized access. Additionally, time synchronization can also be executed with a centralized solution which syncs the time for all systems. This is very benefactory in a manufacturing site with more than one manufacturing system.

6.3.3 Archiving of audit trail

Zenon automatically records and maintains the audit trail as a Zenon-specific proprietary binary file during runtime. This data is only accessible, filterable, and available to query through the user interface in the HMI. Zenon encourages for the customer to establish policies and procedures to prevent unauthorized access to these audit trail files through Windows file security system. In other words, the binary files can only be accessed via Zenon software. The path where these binary files exist on the target system is user definable thus these files can be stored in a secure location in a server rather than local control server on the field.

To maintain the audit trail records, Zenon will automatically create one audit trail file per day. The naming format for these files is the letter “C” followed by time format YYMMDD (e.g. C210530.cel). The size of these audit trail files depends directly on the number of occurred events within the configured project and the event frequency. It is up to the manufacturing process owner (i.e. customer) to handle the access and archive of these files for audit and reviewing processes. However, the audit trail can be exported in several file formats such as comma-separated values (CSV), extensible markup language (XML) (i.e. CSV, XML, SQL etc.) and printed in PDF as a report. This way, the audit trail is in clear human readable form if not accessed via the user interface. The export of audit trail data can be executed incrementally to a SQL relational database. Unfortunately, Zenon software is engineered in a way that the recording of audit trail to closed proprietary binary files cannot be bypassed in any way.

6.3.4 Potential risks regarding audit trail archive

Regardless of wide exporting possibilities for the audit trail, there are some challenges regarding data integrity and the criteria of original records or true certified copy regarding the audit trail. When dealing with the binary .cel audit trail file and the exported file, a decision must be made which is the primary source (i.e. source record) of audit trail and how the archival of the files should be performed for auditing purposes throughout the lifecycle of the manufacturing system. As the audit trail binary file is generated automatically on a daily basis, the number of dividual binary files in 15 years would be enormous. There is also a risk that a newer version of the software cannot open the old binary files. Thus, the software should be archived. As such the binary file is not in human readable form. Although, the binary file can be read by Zenon and displayed in human readable form. A preferred solution to archiving audit trail through its retention period for the target company is further discussed in chapter 7.

6.4 Reporting

Zenon offers clear printed copies of electronically stored data in form of reports for archiving and auditing purposes. Individual or combined documents for audit trail, alarm lists, historical and online values can be provided by default. Thus, explicit batch related

reports containing for example measurement data can be configured via Zenon. The reports are provided as a screen report, hardcopy or external PDF file format encouraging for the reading and archiving of data. The export file path can be configured in the project properties.

The report function can be operated during runtime. Depending on the configured report definitions, reports for alarms, CEL, archives and online value can be display. Through runtime, it is possible to print or export the report as a PDF file in the user-configured export folder.

In terms of data integrity, the reporting of audit trail for auditing functions is important. As stated before, the time stamp format corresponds to the format on Windows level. The audit trail time stamp in the report does not include the time zone as a default. In a global company that has manufacturing facilities in different countries and the data is interpreted by different people, this can be an issue. By not including the time zone in audit trail timestamp, there is a big risk for the integrity of the data, and it can be interpreted in a wrong way. However, the issue of not having time zone in the timestamp can be documented in the validation documents of the machine and further assessed in risk assessment.

6.5 Interfaces to other systems

This chapter briefly discusses the target company's criteria for interfaces to other systems. Only a connection to SQL database and the use of OPC UA communication protocol are in scope of this thesis. These technical criteria functions are not further explained in depth in this thesis.

6.5.1 Export module

Zenon offers several ways to interface with SQL servers. One of them is Zenon SQL Export module, which is an optional module of Zenon Editor and Runtime that allows archive data, event data and alarm data to be exported to an SQL database. This interface is only one directional export i.e. the exported data cannot be read back into the Zenon.

This module can be beneficial in a use case where it is only necessary to export Zenon historical data, alarm data or event data to an SQL database. The export module can be configured with pre-defined Zenon functions. The configuration includes defining the SQL server, database and tables to use for the SQL export. Zenon will automatically generate the tables and columns needed in the database. The event of exporting the data to SQL database can be event-triggered, time-triggered or on demand.

6.5.2 Process gateway

Another type of interface from Zenon to SQL is via Zenon supervisor process gateway, also known as SQL Online. The basic principle is to use a table in an SQL database to hold the current timestamps, value, and status of any defined Zenon variables in a predefined format. This is used to provide online values of tags derived from Zenon drivers to any external systems or applications in general format. Process Gateway makes it possible for the values of all current Zenon variables to be shared with applications capable of reading data from SQL database. The process gateway is configured with OLE-DB (Object Linking and Embedding Database). connection to the SQL server database. The user can configure which individual variables are exposed thus allowing only desired values to be exposed in the SQL table. The refresh rate of updating the values from Zenon to the database can be configured by the user. Zenon will automatically generate the needed tables and columns. The information consists of variable ID, variable value (strings), variable value (float), timestamp second, timestamp millisecond and status of the variable.

6.5.3 Historian module

An SQL server can also be used as a historian server for Zenon via Historian module. It allows Zenon to transfer archive data to a centralized SQL server to promote transparency between different departments and to store it in an open format where data can be queried with simple SQL statements. In addition, the data transferred to the SQL can be read back into Zenon for trending or reporting purposes. This module also allows to export alarms and events i.e. CEL to a centralized database. However, the alarm and event logs cannot be read back into Zenon from the SQL server after exporting. This module is used in companies who believe in open integration of process data to other organizations in the

corporate structure. When process data is stored to an SQL database, it can be easily queried, reported, and analyzed by many people and by third party software. The exportation of archived data in runtime to an SQL database is called evacuation in Zenon. When configured, this can be performed automatically. For example, the evacuation cycle can happen at the end of a batch or after every month. To evacuate the archival data to an SQL server, a time period has to be configured after which the older records will be automatically evacuated to the database. To configurate this behavior, database connection must be created, credentials need to be provided pointing to the SQL server and database which is used. Zenon will automatically create the necessary tables and formats. The historian interface also provides a buffer for the archived runtime data in case of connection issues to the SQL server. Zenon will archive the historical data locally on the runtime server and when the SQL server comes available, the data which should have been previously evacuated is evacuated automatically.

As the target company uses a custom-made database model for assembly process, Zenon must be used without the above-mentioned default tables. It can be used this way, but the included functions of data evacuation and export cannot be executed automatically. Therefore, the custom-made database model can be used with Zenon.

6.5.4 Open platform communication unified architecture

Open Platform Communication Unified Architecture (OPC UA) is a vendor independent communication protocol for industrial automation systems released in 2008. Developed by OPC Foundation, the current version of the specification is 1.04. The target company is pursuing the use of this communication protocol as a standard solution in manufacturing systems to homogenize communications and data structure in production departments. The use of OPC UA as a communication protocol offers benefits regarding data integrity and security in production environments. The most important benefits in scope of this thesis are discussed after evaluating Zenon's combability with the protocol.

OPC UA server for Zenon has been certified by the OPC Foundation starting from October 22, 2012. Zenon can communicate with OPC UA servers with OPC UA Client driver. The OPC UA server has been available for use in Zenon since version 5.50 (Zenon 2021).

The use of OPC UA as a communication protocol offers many benefits such as platform independency and scalability. OPC UA can be operated with basically any operating system from Linux distribution to Apple OSX. It can also be scaled from an embedded micro-controller to all the way to cloud-based infrastructure (OPC Foundation 2021).

Third major benefit is the security aspect of OPC UA protocol. OPC UA presents session encryption where messages are transmitted securely at plethora of encryption levels. The messages can also be signed by the sender to verify the origin and integrity of received messages. With this feature, it can be verified that the messages are not altered by anyone in the middle of the transaction. OPC UA also uses sequenced packets to prevent message replay attacks. Additionally, the protocol uses X590 certificates in identification of each UA client and server to protect which system or application can connect with each other. Moreover, the protocol represents user control, where any application can require user authentication to restrict access rights and address-space views. OPC UA also inhabits auditing possibility to activities by the system or user to be logged providing an audit trail (OPC Foundation 2021).

In addition, OPC UA provides one fundamental benefit to other protocols in the field which is the modeling of information and access. The framework of information modeling in OPC UA turns data into information with complete object-oriented capabilities. This behavior sets the rules and foundation block necessary to use any information model with OPC UA. Despite OPC UA already defines several core models for many industries, many organizations can build their own information models upon them. The information modeling makes it possible to standardize and homogenize the information models used in different manufacturing systems. With OPC UA the information can be already modeled in a pre-defined way even before it is transmitted onwards from the field instrumentation (OPC Foundation 2021).

7 DISCUSSION

This chapter summarizes the evaluation and discusses both benefits and challenges regarding the proposed SCADA software called Zenon Supervisor version 8.20.

7.1 Result summary and benefits

The first evaluation step of the proposed SCADA software, Zenon, was user management. The user management in Zenon enables proper restriction of unauthorized personnel and clear segregation of duties once properly configured. Zenon has a native connection to AD enabling the use of centralized user management as the target company wishes. By using the application level segregation and duties with a centralized user management system, the target company can restrict access to authorized personnel only and the management of user groups is easy. This combination of user management adheres to the current prevailing regulatory framework.

Second evaluation target was the inspection of audit trail. Zenon has the necessary features regarding audit trail to promote transparency and traceability of data once properly configured. In addition, the time synchronization was investigated. Zenon 8.20 adopts the time from the server where it is running. Thus, it is on the responsibility of the manufacturing system process owner to configure the server accordingly. With proper time synchronization capabilities, the audit trail feature can capture all the needed data in manufacturing system to adhere to current prevailing framework.

Third evaluation step was the inspection of reporting services integrated to the software. The reporting services of Zenon Supervisor 8.20 is very configurable and rather easy to use. To promote traceability and transparency in auditing of manufacturing records, report generation with Zenon is straightforward. Zenon offers easy tools to modify different report templates for different purposes. The reports can be exported from the Runtime to PDF and the file path can be user configured, thus keeping data away from the shopfloor.

Next evaluation point was the connection to SQL relational database. The connection to SQL can be established by three different interfaces. All of them provide easy access to SQL after proper configuration. Zenon promotes the connection to external SQL relational database thus the target company can make use of it in storing GxP data centrally in standardized form. The data can be easily queried from the database with third party applications across departments in the organizations adding transparency. Zenon can also utilize a custom-made database model.

Lastly, the possibility to use OPC UA as a communication protocol with Zenon was investigated. As it turns out, Zenon is officially certified by OPC Foundation to use OPC UA server. The use of OPC UA enables many benefits to the target company regarding data integrity concerns and the overall security architecture in automated manufacturing systems. For example, the use of certificates and encryption of transmission for security measures is highly enabled in the target company. Additionally, OPC UA presents information modeling and access to standardize the data in the shop-floor level even before it is transmitted to upper automation levels. Therefore, the target company can standardize their information models in different automated manufacturing systems even across departments.

7.2 Challenges

Although the Zenon Supervisor 8.20 can be declared adhering to current regulative framework regarding electronic records and electronic signatures, the software does not come without challenges.

The biggest challenge was found out regarding the archiving of audit trail. As previously introduced, the Zenon software archives audit trail data during runtime to Zenon-specific proprietary binary files. Every day an audit trail archive file will be generated to user configurable file path in Zenon project directory. The audit trail archive file can be only read by executing it through runtime and it is displayed on the HMI in human readable form. The challenge is to properly manage and archive these files. As stated previously, a common lifecycle for an industrial automation system is roughly 15 years. Therefore, in the time span of 15 years Zenon will generate approximately 5435 binary files during

this time. The management and archival of these files for auditing purposes seems very impractical. For a sophisticated pharmaceutical company promoting data integrity, it is not ideal. The audit trail can be exported to an SQL relational database, but the generation of these Zenon-specific binary files cannot be bypassed by any means. If the solution is using SQL relational database for archiving audit trail data, then the same data is in two different places. This rises the question of the originality of the audit trail data. In this case, it must be sufficiently documented which of the data sources is used as a primary source so it can be used for GxP purposes. If the data is copied from Zenon to SQL, the integrity of the data must be carefully validated through comparing that the data has been transferred unchanged. The target company uses storing of manufacturing records including audit trail data to SQL relational database as a best practice. The best and only approach in terms of data integrity would be to copy the audit trail data from Zenon to SQL and declaring the Zenon binary files as the primary source of audit trail data and SQL as a true copy. With careful validation of the data transmitting process and documentation the solution represented here should be satisfactory to the needs of the target company and compliant to prevailing regulatory framework. There is other SCADA software available where the audit trail data can be written directly to the SQL relational database without generating any binary files to project directory. The archiving and management of this centralized SQL database model is superior to the binary file archival.

There was also a challenge regarding the time stamp of audit trail. By default, the audit trail data does not include time zone. As stated before, this can be an issue in an international company where people residing in different countries are in different time zones and they analyze the same data differently. However, the time zone can be added to the audit trail report template by external means. In addition, the time zone which the system uses must be clearly stated in the documentation regarding the machine.

8 CONCLUSION

In this thesis, the compliance of a SCADA software was studied. To give the reader a clear understanding of the environment in which the target company operates, the special characteristics of pharmaceutical manufacturing were introduced. Next, industrial control systems were discussed in general terms after focusing on SCADA technology. After that, data integrity was introduced by discussing legislation and guidance documents and manufacturing records. In addition, general ways to ensuring data integrity were discussed.

After four chapters of academic literature review the target company and its relevant organization and assembly process was introduced. The commissioner of this thesis is one of the largest pharmaceutical companies in the world and the topic of this thesis initiated from a machine investment project. Next, the proposed software was represented and its functionalities regarding alarm and event management, reporting services, user management and interfaces to other systems were investigated and reflected to regulations. After discussing the functionalities, a statement was made if the SCADA software can be used compliantly in the target company specific implementation.

In today's digitalized world the use of data acquisition and data visualization software in manufacturing industry is growing. Moreover, the pharmaceutical industry has strict requirements on handling of manufacturing records. This thesis offers a practical guideline on which SCADA functionalities should be inspected closely when operating in a regulated environment. The practical part of this thesis outputs a consideration structure of regulation adherence for any SCADA software, not only Zenon Supervisor.

In conclusion, COPA-DATA Zenon Supervisor version 8.20 offers the necessary features to achieve compliance to current prevailing regulatory framework. However, the functionalities must be properly configured to achieve compliance. After all, the responsibility of compliance always falls to the manufacturing system owner. In addition, the compatibility of the software with best practices used in the target-company-specific assembly process was investigated and deemed satisfactory. Some challenges regarding the archival of audit trail data were found but a sufficient solution for the challenges was

proposed. Also, one challenge regarding time stamp not including time zone was found during inspection of audit trail data. This challenge can be solved by external means.

This thesis provided only a theoretical approach to the features in Zenon Supervisor 8.20 used to adhere to prevailing regulatory framework surrounding electronical records and electronic signatures. Other features of Zenon Supervisor 8.20 were left untouched.

All the things evaluated in this thesis would benefit from a practical test project implementation, where all the necessary features would be presented and configured as the target company plans to use them in virtual environment. A relevant test case environment would be possible using server and PLC virtualization and HMI programming to further demonstrate the features of the software in action. This thesis would also benefit from practical implementation and configuration of the SQL server, AD and OPC UA configuration. Especially the configuration of OPC UA communication protocol would require more in-depth research. However, this thesis does not take a stand how some features should be configured. This thesis merely points out that the proposed software can be used compliantly in the target company specific assembly process.

REFRENCES

Chan Lai Wah, Sia Chong Hock, Vernon Tay, Vimal Sachdeva. (2020) Pharmaceutical Data Integrity: issues, challenges and proposed solutions for manufacturers and inspectors. Generics and Biosimilars Initiative Journal (GaBI Journal). 2020;9(4):171-82. DOI: 10.5639/gabij.2020.0904.028

Schulz D. (2015). FDI and the Industrial Internet of Things IEEE 20th Conference on Emerging Technologies & Factory Automation (ETFA), pp. 1-8, doi: 10.1109/ETFA.2015.7301513.

EMA. (2021). European Medicines Agency – What We Do. [online] Available at: https://www.ema.europa.eu/en/about-us/what-we-do [Accessed 6 Jan. 2021]

EMA. (2021). European Medicines Agency – Good Manufacturing Practice [online] Available at: https://www.ema.europa.eu/en/human-regulatory/research-development/compliance/good-manufacturing-practice [Accessed 3 Jan. 2021]

Engell, S. and Harjunkoski I. (2012). Optimal Operation: Scheduling Advanced Control and their Integration, Computers and Chemical Engineering, 47, pp. 121-133. doi: https://doi-org.pc124152.oulu.fi:9443/10.1016/B978-0-444-63456-6.50072-7

EudraLex (1999) Pharmaceutical legislation. Medicinal Products for human and veterinary use: Good manufacturing practices. European Commission, vol 4, ISBN 92-828-2029-7.

Eudralex. (2003). 2003/63/EC. Official Journal of the European Union. [online] Available at: https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-1/dir_2003_63/dir_2003_63_en.pdf [Accessed in 25 Feb. 2021]

EudraLex (2010). The Rules Governing Medicinal Products in the European Union, Volume 4: Good Manufacturing Practice, Medicinal Products for Human and Veterinary Use. European Commission. [online] Available at

https://ec.europa.eu/health/sites/default/files/files/eudralex/vol-4/chapter4_01-2011_en.pdf [Accessed 10 Jan. 2021]

FDA (2003). Guidance for Industry Part 11, Electronic Records; Electronic Signatures - Scope and Application. [online] Available at: https://www.fda.gov/media/75414/download [Accessed 4 Feb. 2021]

FDA. (2017). What we do - FDA. [online] Available at https://www.fda.gov/about-fda/what-we-do [Accessed 1 Feb. 2021]

FDA. (2018). Code of Federal Regulations Title 21. Part 11 Electronic records; Electronic Signatures. [online] Available at: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=11 [Accessed 2 Feb. 2021]

FDA. (2018) Facts About the Current Good Manufacturing Practices (CGMPs) [online] Available at: https://www.fda.gov/drugs/pharmaceutical-quality-resources/facts-about-current-good-manufacturing-practices-cgmps [Accessed 3 Feb. 2021)

FDA. (2018) Code of Federal Regulations [online] Available at: https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/Databases/ucm135680.html [Accessed 2 Feb. 2021]

FDA. (2021) Recalls, Market Withdrawals, & Safety Alerts [online] Available at https://www.fda.gov/safety/recalls-market-withdrawals-safety-alerts [Accessed 6 Jan. 2021]

Fimea. (2018). About us – fimea englanti – Fimea, [online]. Available at: https://www.fimea.fi/documents/542809/841791/Facts+about+Fimea.pdf/c5645118-de5c-24ad-18e1-65188d77654b?t=1530873000789 [Accessed 2 Feb. 2021]

Galloway B. and Hancke G. P. (2013). Introduction to Industrial Control Networks IEEE Communications Surveys & Tutorials, vol. 15, no. 2, pp. 860-880, doi: 10.1109/SURV.2012.071812.00124

Gouveia B. G., Rijo, P., Gonçalo, T. S., & Reis, C. P. (2015). Good manufacturing practices for medicinal products for human use. Journal of pharmacy & bioallied sciences, 7(2), 87–96. https://doi.org/10.4103/0975-7406.154424

ISPE. (2021). About ISPE. [online] Available at: https://ispe.org/about [Accessed 12 Jan. 2021)

ISPE (2008, pp 196). GAMP 5: A Risk-based Approach to Compliant GxP Computerized Systems, 3rd ed. ISPE, ISBN: 1-931879-61-3.

ISPE (2018, pp 196). GAMP RDI Good Practice Guide: Data Integrity - Key Concepts. International Society for Pharmaceutical Engineering. ISBN 978-1-946964-11-3

ISPE (2017, 152 pp). GAMP Guide: Records and Data Integrity. International Society for Pharmaceutical Engineering. ISBN 978-1-936379-95-8

ISPE (2019, pp 156). GAMP RDI Good Practice Guide: Data Integrity - Manufacturing Records. International Society for Pharmaceutical Engineering. ISBN 978-1-946964-20-5

Industry 5.0. (2020). [online]. Available from: https://ec.europa.eu/info/research-and-innovation/research-area/industrial-research-and-innovation/industry-50_en [Accessed 5 Jan. 2021]

ISA. (2019) ISA-S95-1: Enterprise-control system integration part 1: Models and terminology [online] Available at https://isa-95.com/isa-95-enterprise-control-systems/ [Accessed 5 Jan. 2021]

Kim, J. H., & Scialli, A. R. (2011). Thalidomide: the tragedy of birth defects and the effective treatment of disease. Toxicological sciences: an official journal of the Society of Toxicology, 122(1), 1–6. doi: https://doi.org/10.1093/toxsci/kfr088

Macaulay, T. & Singer, B..(2016 pp 204). Cybersecurity for industrial control systems: SCADA, DCS, PLC, HMI, and SIS. ISBN 978-1-439801-96-3

Markarian J. (2018). Modernizing Pharma Manufacturing. Pharmaceutical Technology. 42 (4) Available at: https://cdn.sanity.io/files/0vv8moc6/pharmtech/2910f5c9beedad8d8c3787d3f849073c6579aa70.pdf

Mascia, S., Heider, P.L., Zhang, H., Lakerveld, R., Benyahia, B., Barton, P.I., Braatz, R.D., Cooney, C.L., Evans, J.M.B., Jamison, T.F., Jensen, K.F., Myerson, A.S., Trout, B.L. (2013). End-to-end continuous manufacturing of pharmaceuticals: Integrated synthesis, purification, and final dosage formation. Angewandte Chemie - International Edition, 52 (47), pp. 12359-12363. [Available at https://www.scopus.com/inward/record.uri?eid=2-s2.0-84887553366&doi=10.1002%2fanie.201305429&partnerID=40&md5=8b19d94e8a4c511b0ff0af05a502a9d7]

Mehta, B. R., & Reddy, Y. J. (2015). Chapter 7 - SCADA systems. In B. R. Mehta, & Y. J. Reddy (Eds.), Industrial process automation systems (pp. 237-300). Oxford: Butterworth-Heinemann. doi: https://doi-org.pc124152.oulu.fi:9443/10.1016/B978-0-12-800939-0.00007-3

Radvanovsky R. and Brodsky J. (2013) Handbook of SCADA/Control Systems Security. Auerbach Publications, NIST Special Publication 800-82, Revision 2. http://dx.doi.org/10.6028/NIST.SP.800-82r2

Rule D. & Dittner R. (2011). The Best Damn Server Virtualization Book Period: Including Vmware, Xen, and Microsoft Virtual Server

OPC Foundation (2021). Unified Architecture. [online] Available at: https://opcfoundation.org/about/opc-technologies/opc-ua/ [Accessed in 5 June 2021]

Stouffer, Keith & Falco, Joseph & Kent, Karen. (2007 pp. 960). Guide to Supervisory Control and Data Acquisition (SCADA) and Industrial Control Systems Security ISBN 9781597492171

WHO. (2015). Medicines Good manufacturing practices. [online] Available at: https://www.who.int/news-room/q-a-detail/medicines-good-manufacturing-processes [Accessed 11 Jan. 2021]

Yadav G. & Paul K. (2021). Architecture and security of SCADA systems: A review. International Journal of Critical Infrastructure Protection, 34, 100433. doi: https://doi-org.pc124152.oulu.fi:9443/10.1016/j.ijcip.2021.100433

Zenon (2021). Zenon software platform. [online] Available at: https://www.copadata.com/en/product/zenon-software-platform-for-industrial-automation-energy-automation/ [Accessed 25 Apr. 2021]

Zhang P. (2010). CHAPTER 10 - industrial control networks. In P. Zhang (Ed.), Advanced industrial control technology (pp. 363-427). Oxford: William Andrew Publishing. doi: https://doi-org.pc124152.oulu.fi:9443/10.1016/B978-1-4377-7807-6.10010-5