DATA MANAGEMENT AND INTEGRATION FOR CONTINUOUS PHARMACEUTICAL MANUFACTURING PROCESSES

By

HUIYI CAO

A thesis submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Master of Science

Graduate Program in Chemical and Biochemical Engineering

Written under the direction of

Dr. Rohit Ramachandran

And approved by

New Brunswick, New Jersey

October 2016

ABSTRACT OF THE THESIS

Data Management and Integration for Continuous Pharmaceutical Manufacturing Processes

by HUIYI CAO

Dissertation Director:

Dr. Rohit Ramachandran

As the pharmaceutical industry seeks more efficient methods for the production of higher value therapeutics, the associated data analysis, data visualization, and predictive modeling require dependable data origination, management, transfer, and integration. As a result, the management and integration of data in a consistent, organized, reliable manner is a big challenge for the pharmaceutical industry.

The S88 recipe model, an international standard for describing standard batch processes, has been adapted in this study to deliver a well-defined data structure that will improve the data communication inside the system architecture for continuous processing. This work has been divided into two parts due to differing requirements between laboratory-based analytical measurements and the pilot-plant continuous pharmaceutical process. In the laboratory platform, recipes have been developed for a sub-set of material

property tests that for instance, could be performed on the analytical instrument (e.g. FT4 for flow). Drupal, an open source content management system, is implemented on an Amazon web service for data transfer between the analytical devices eventually a data management platform. A recipe module for Drupal is developed for recipe management, in which users could create, import, and modify recipes. Scientists can access recipes through Drupal's web page interface and perform experiments following standard recipe steps. Research data can be recorded by manual input or automatically parsing data files on the backend of the server. This system works like a recipe based electronic laboratory notebook.

In the continuous manufacturing pilot plant, process data is generated by unit operation equipment and integrated process analytical technology (PAT) instruments [1]. A process control system (e.g. DeltaV (Emerson)) collects the data from equipment and a PAT data management tool (e.g. synTQ (Optimal Industrial Automation)). The PAT data management collects data from an inline/online measurement system [2]. The recipe for the whole continuous process is implemented in DeltaV. Data in DeltaV is collected according to the recipe and is transferred to a data storage hub (PI system (OSI Soft)) in the same structure. The Event Frame feature from PI system allows the possibility to create an individual recipe based on continuous data feeding. From PI system, the data is sent to online data storage box and cloud system. From the box/cloud, the data can be access at different physical company sites, can be analyzed and applied for various applications. This study is the first attempt to apply ISA-88, a batch control standard, to continuous pharmaceutical manufacturing.

All the detailed information of the lab-based experiment and process manufacturing, including equipment, samples and parameters are documented in the recipe. Recipes containing data can be exported from this system to be shared and transferred. After detaching the data from recipes, a reliable and consistent data source is provided for data visualization and process modeling. Another feature is the two-dimensional barcode labels that are used in this strategy. Every ingredient and equipment of the analytical experiment or manufacturing process will have a unique barcode, which can be used to identify the item and trace all the information related. This enforces material traceability, which is an essential requirement in the overall Quality by Design (QbD) initiative.

Acknowledgement

I would like to express my sincere appreciation to my advisor Dr. Rohit Ramachandran for his guidance and encouragement during the course of my master's degree at Rutgers. I would like to thank Dr. Marianthi Ierapetritou and Dr. Shantenu Jha for their invaluable suggestions on this research topic. I would like to thank Dr. Ravenda Singh and Dr. Jun Zhang for their great help with my work. My sincere gratitude to Janssen Pharmaceuticals, Inc. of Johnson & Johnson for funding this project, including a co-op with them for 6 months. I would like to extend special gratitude to Dr. Adam Fermier for his guidance and expertise throughout my project.

I would like to thank my parents for their encouragement and support over the years. I am privileged to have them.

Contents

List of Figures

List of Tables

List of Abbreviations

AF	Asset Framework
API	Active pharmaceutical ingredient
AWS	Amazon Web Service
CMS	Content Management System
DOE	Design of Experiment
ELN	Electronic Laboratory Notebook
ERP	Enterprise Resource Planning
FDA	Food and Drug Association
LIMS	Laboratory Information Management System
MES	Manufacturing Execution System
OLE	Object Linking and Embedding
OPC	OLE for Process Control
PAT	Process analytical technology
PFC	Procedural Function Chart
PLC	Programmable Logic Controllers
PLN	Paper laboratory notebooks
QbD	Quality by Design
R&D	Research and development
RDS	Relational Database Service
SFC	Sequential Function Charts
URL	Uniform Resource Locator
XSD	XML Schema Definition

Chapter 1 Introduction

For decades, the pharmaceutical industry has been dominated by a batch-based manufacturing process, which can lead to increased inefficiency and delay in time-to-market of product, as well as the possibility of errors and defects. Continuous manufacturing in contrast, is a newer technology in pharmaceutical manufacturing that can enable faster, cleaner and cheaper production. The US Food and Drug Association (FDA) has recognized the advancement of this manufacturing mode and been encouraging its development as part of the FDA's QbD paradigm [3]. The application of process analytical technology (PAT)[4] and control systems is great effort to gain improved science-based process understanding. In the face of the enormous amount of data generated from a continuous process, a sophisticated data management system is required for integration of analytical tool to control system, as well as the off-line measurement system.

In order to represent, manage and analyze a large amount of complex information, an ontological informatics infrastructure will be necessary for the process and product development in pharmaceutical industry [5]. ISA-88 Batch Control Standard [6-8] is an international standard addressing batch process control, which has already been implemented in other industry for years. Adapting this industrial standard into pharmaceutical manufacturing could provide a design philosophy for describing equipment, material, personnel, as well as reference models [9,10]. This recipe-based execution could work as a hierarchical data structure for the assembly of data from the control system, PAT tools, and off-line measurement device. The combination of the ISA-88 recipe model and the data warehouse informatics strategy [11] leads to the "recipe data warehouse" strategy [12], which will allow the data management across multiple execution systems, as well as the data analysis and visualization.

Applying the "recipe data warehouse" strategy to continuous pharmaceutical manufacturing provides a possible approach to handle the data produced via analytical experiment and process recipe execution. Not only the data itself but the context of data can also be well captured and saved for documentation and reporting. However, unlike batch operations, continuous manufacturing is a complicated process containing series of inter-connected unit operations with multiple execution layers. It will be quite challenging to integrate data across the whole system while maintaining an accurate representation of the complex manufacturing processes.

In addition to the data collection and integration of the continuous manufacturing plant, the highly variable and unpredictable properties of raw materials are necessary to be captured and stored in a database because they could have an impact on the quality of the production [13]. These properties of relevance to continuous manufacturing will be measured via many different analytical methods, including FT4 powder characterization, particle size analysis, Washburn technique, etc. The establishment of a raw material property database could be achieved by the "recipe data warehouse" strategy. Nevertheless, compared to the computer aided manufacturing used in production process, the degree of automation would vary a lot in different analytical platforms. Therefore, an easily accessible recipe management system will be highly desirable in the characterization laboratories.

Moreover, a cloud computing technology for data management and storage is adapted to deal with massive amount of data generated from the continuous manufacturing process. While traditional computational infrastructures involve huge investments on dedicated equipment, cloud computing offers a virtual environment for users to store or share infinite packets of information by renting hardware and storage systems for a defined time period [14]. Because of its many advantages including flexibility, security, and efficiency, cloud computing is suitable for data management in pharmaceutical industry [15].

In this study, a big data strategy following ISA-88 batch control standard is applied for data management in continuous pharmaceutical manufacturing. The ontology of recipe modeling and the design philosophy of recipe is elaborated. A data management strategy is proposed for the data integration in both continuous manufacturing processes and the analytical platforms used for raw material characterization. This strategy has been applied to a pilot direct compaction continuous tablet manufacturing process to build up data flow from equipment to recipe database on the cloud. In the characterization of raw material and intermediate blends properties, experiment data is also captured and transferred via a web-based recipe management tool.

Chapter 2 Recipe Definition

In ISA-88, a recipe is defined as the minimum set of information that uniquely identifies the production requirements for a particular product [6]. However, there will still be a significant amount and different types of necessary information, which is required to describe products and to make products. Holding all the information in one recipe will be complicated and cumbersome for human beings. As a result, four types of the recipe are defined in ISA-88 to focus on different levels and accurate information.

2.1 Recipe Types

Table 1 shows the four recipe types defined in ISA-88 and their relationship. Each of the four recipe types is described in Table 1 according to the ISA-88.

Table 1 ISA-88 recipe types

Type	Definition
General recipe	A type of recipe that defines raw materials and site independent processing requirements.
Site recipe	A type of recipe that usually derived from the general recipe to meet specific conditions or constraints of the site manufacturing the product.
Master recipe	A type of recipe that contains the information of an individual product and depends on equipment situation.
Control recipe	A type of recipe that defines the manufacturing process a single batch of product.

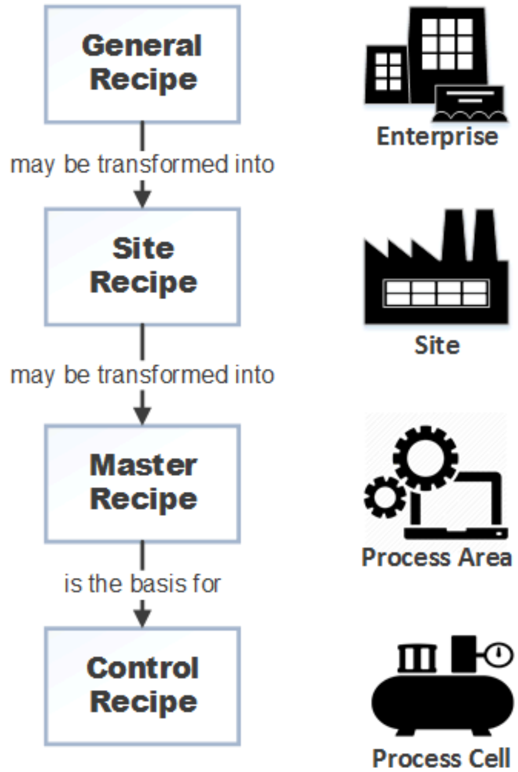

Figure 1 ISA-88 Recipe Model

2.2 Process Model

The ISA-88 standard defines a process model that has four levels, including process, process stages, process operations, and process actions. The structure of this process model can be displayed in Figure 2.

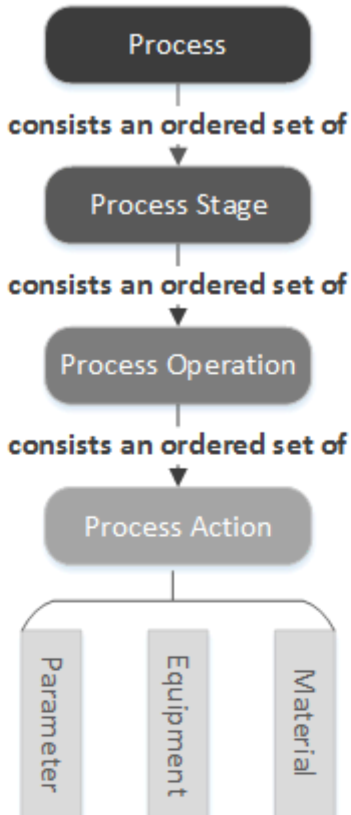

Figure 2 ISA-88 process model

It is important to be noticed that the reference model and guideline recommended in the ISA-88 standard are not to be strictly normative. This feature provides the sufficient level of flexibility to represent the current manufacturing process according to the ISA-88 standard, as well as the opportunity to expand the reference model to suit unusual manufacturing plant [16]. In this case, ISA-88 recipe model is extended to continuous manufacturing, instead of the batch manufacturing process, to provide a hierarchical structure for data management and integration.

Chapter 3 Recipe Modeling

3.1 Ontology

An ontology is an explicit specification of a conceptualization where conceptualization refers to an abstract, simplified view of the world that we wish to represent for some purpose [17]. A more detailed definition of an ontology has been made where the ontology is described as a hierarchically structured set of terms for describing a domain that can be used as a skeletal foundation for a knowledge base [18].

Ontologies are created to support the sharing and reuse of formally represented knowledge among different computing systems, as well as human beings. They provide the shared and common domain structure for semantic integration of information sources. In this work, a conceptualization through the ISA-88 shows the advantage of building up a general conceptualization in pharmaceutical manufacturing and development domain.

3.2 Ontology language

Information technology has already developed the capabilities to support the implementation of such proposed informatics infrastructure. The language used in an ontology should be expressive, portable and semantically defined, which is important for the future implementation and sharing of the ontology. The world Wide Web Consortium (W3C) has proposed several markup languages intended for web environment usage, and Extensive Markup Language (XML) is one of them.

XML is a meta-language that defines a set of rules for encoding documents in a format which is both human readable and machine readable [19]. The design of XML focuses on what data is and how to describe data. XML data is known as self-defining and self-describing, which means that the structure of the data is embedded with the data. There is no need to build the structure before storing the data when it arrives. The most basic building block of an XML document is an element, which has a beginning and ending tag. Since nested elements are supported in XML, it has the capability to embed hierarchical structures. Element names describe the content of the element while the structure describes the relationship between the elements.

As an XML Schema defines the structure of XML documents, an XML document can be validated according to the corresponding XML Schema. XML Schema language is also referred as XML Schema Definition (XSD) which defines the constraints on the content and structure of documents of that type [20].

3.3 Recipe model implementation

The purpose of this part is to illustrate the proposed methodological method to assess them implementation of ISA-88 recipe model into continuous pharmaceutical manufacturing. Proven evidence has shown that developing a new product manufacturing process might be a difficult task. Although ISA-88 standard provides a practical reference model for sure, adapting a batch model to the continuous process will still be challenging. The implementation of recipe model is divided into several steps to perform the activities in a structured way.

3.3.1 Step 1. ISA-88 applicable area identification

The initial scope and project boundaries will be defined in this step in order to discuss the main objectives. As mentioned above, the ISA-88 standard is not strictly defined and can be expanded to continuous manufacturing in contrast. Moreover, since analytical experiments are performed by each sample, one sample can be treated as a "batch" in the analytical process. Therefore, it is also reasonable to consider the off-line characterization experiment of material to be applicable to the ISA-88 standard. In this situation, a technique-specific recipe, which contains detailed process steps, could provide instructions for specific experiment type, as well as the designed structure for result documentation (Figure 3).

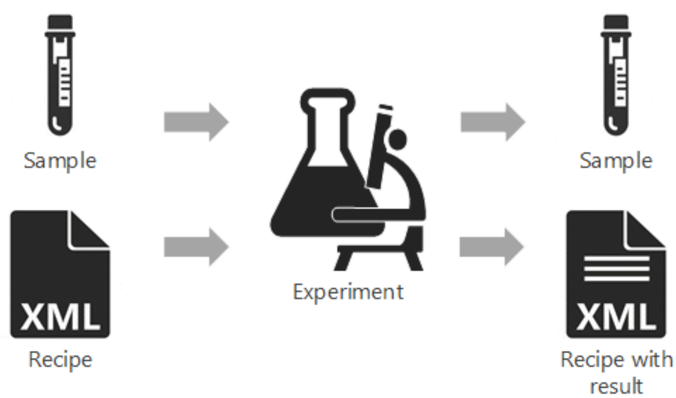

Figure 3 Material and data flow in an experiment process

Preliminary data gathering is performed to provide a clear understanding of the stream process and the requirement of each particular case. The component may be modified in ISA-88 recipe model could be recipe type, recipe content structure, recipe representation, etc. The in-depth knowledge of process architecture, manufacturing execution systems, product and material information, and analytical platform will be necessary to the accomplishment of this step.

3.3.2 Step 2. Recipe structure definition

An XML Schema is created for the implementation of an ISA-88 standard to provide the recipe model a structured, system-independent XML markup. The step needs to be done based on the data collected from the previous step. An XML Schema document works as a master recipe standard. The goals of this document include setting restrictions to the structure and content of the recipe, maximizing the recipe model's expressive capability and realizing machine's ability of recipe validation.

3.3.3 Step 3. Process analysis

Due to the master recipe's equipment dependent property, it will be complicated and inefficient to develop a single master recipe to describe the whole continuous pharmaceutical process. The process itself consists an ordered set of multiple unit operations, which may change from time to time, depend on the equipment used. This does happen a lot in the product process development stage, especially the early time period. Another barrier stands here is the transformation from research and development (R&D) to production. Usually, R&D performs the product development process on small-scale pilot plants, which could be quite different from the production plant in technology aspect. Moreover, the differences between technologies used in every plant are crucial to product manufacturing after the transformation. The entire problem mentioned above is critical to the implementation of ISA-88 to continuous pharmaceutical manufacturing.

A suitable approach to implementation of ISA-88 into continuous manufacturing might be developing a robust modular structure that could support a variety of equipment types and classes. In this case, a continuous process is divided into several unit operations, such as mixing, blending, and etc. Each unit operation could be performed by different equipment types, which correspond to different recipe module. In other words, every equipment type might take place in a continuous process will be mapped to a recipe module individually. The ordered combination of these recipe modules will form the master recipe of continuous manufacturing.

PAT tools used in continuous pharmaceutical manufacturing, NIR, and Raman for example, are implemented in a continuous process to provide real-time monitoring of the material properties and are essential to manufacturing decision making. The PAT tools will also be treated as a unit operation that has a corresponding recipe module, in order to provide flexibility and reliability for the management of PAT.

3.3.4 Step 4. Process mapping

XML-based master recipes will be developed according to the XML Schema to present the manufacturing process. First of all, the equipment information, operation procedure, as well as process parameters of unit operations will be mapped carefully to the ISA-88 recipe model to generate respect recipe modules. These recipe modules could be transformed into the master recipe and used for the individual study and data capturing on unit operations.

Chapter 4 Big Data Recipe Architecture

The fast development of information and communication technology, such as big data and cloud computing, provides the capability to increase productivity, quality and flexibility across industries. The process industry is also seeking a possible approach to bringing together all data from different process levels and distributed manufacturing plants in a continuous and holistic way to generate meaning information [21]. For continuous pharmaceutical manufacturing, a data flow across the whole process is proposed in Figure 4 following the traditional design pattern, automation pyramid [22], to create an information and communication infrastructure.

On the top of this structure located the Enterprise Resource Planning (ERP) systems which provide the integrated view of core business process based on the data collected from various business activities [23] . Long term resource planning, primarily human and material resource will be its primary objective. Historian, on the Level 4, is intended for collecting and organizing data to provide an information infrastructure. Meaningful information in different representations, ISA-88 recipe document, could be generated from the organizational assets, as well.

In terms of the levels below, they are separated into two parts, continuous manufacturing process, and analytical laboratory platform. Because of the distinct data acquisition method used these two areas, they will be elaborated separately in the following section.

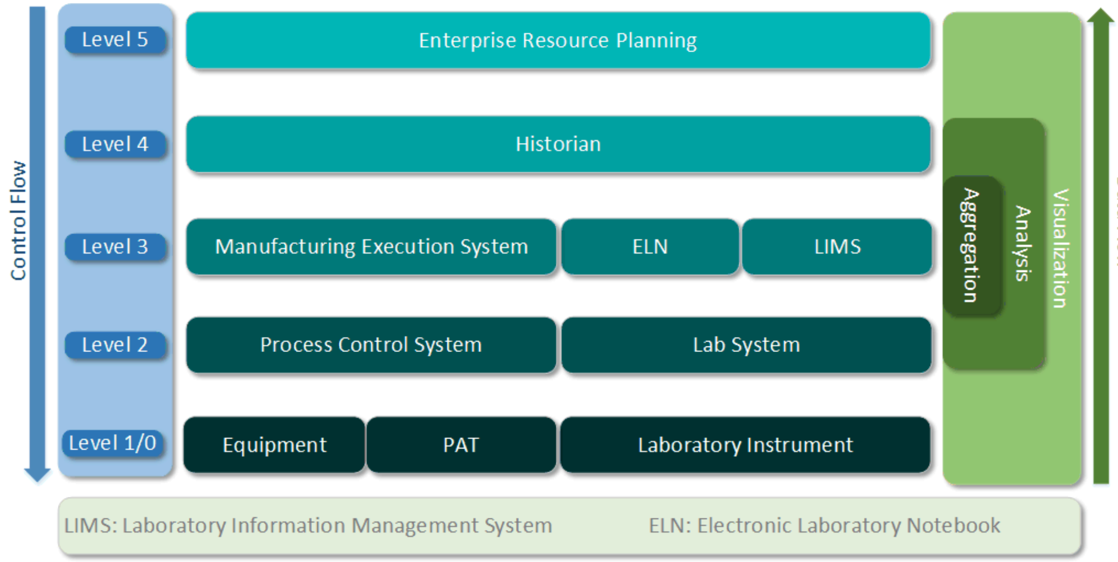

Figure 4 Architecture of data integration

4.1 Continuous manufacturing

According to the diagram above, there are three more levels in the continuous manufacturing process. Manufacturing Execution System (MES) performs the real-time managing and monitoring work-in-process on the plant floor. During the producing process, Process Control System plays a significant role in controlling the system states and conditions to prevent severe problem during operations. Process equipment located in Level 1/0 contains process equipment and PAT tools. Process equipment is always consisted of the field device and embedded Programmable Logic Controllers (PLC). PLC is an industrial computer control system which monitors input devices, such as machinery or sensors and makes the decision to control output devices according to its program [24]. The most powerful advantage of PLC is its capability to change the process or operation while collecting information. In terms of PAT tools, it has corresponding PAT data management tools that could communicate with the control platform. Above are the different levels of control systems across the manufacturing process.

Distinct from the control flows organized top-down, whereas the data capturing and information flow are bottom up. After generated from process equipment and PAT tools in the continuous process, various field data is fed into PCS to form a recipe structure which is accordance with ISA-88 recipe model. In other words, manufacturing data is collected, organized, interpreted and transformed into meaningful information starting from PCS level. This recipe structure will be mapped identically into historian for storage. While the low-level network in this information architecture is mainly based on the communication over bus systems, such as field device to PLC, most of high-level systems use connections based on Ethernet technology named Object Linking and Embedding (OLE) for Process Control, which is also known as OPC. OPC is a software interface standard that enables the communication between Windows programs and industrial hardware devices to provide reliable and performable data transformation.

4.2 Laboratory experiment

Unlike control systems used in continuous manufacturing process plant, the laboratory platform consists of light-weighted, but various types of instruments are generating analysis data in different formats. Therefore, on the third level located the Electronic Laboratory Notebook (ELN) system or Laboratory Information Management System (LIMS) that has the ability for support all kinds of analytical instruments. While material properties are measured by instruments and collected by laboratory systems, these data sets will be transformed into ELN or LIMS systems for further organization and interpretation. Meaningful information will be generated in recipe form according to the ISA-88 standard in Level 3.

Chapter 5 Results and Discussion

5.1 Data Integration in Continuous Tablet Manufacturing

5.1.1 Direct Compaction Continuous Tablet Manufacturing Process

A direct compaction continuous direct manufacturing pilot plant has been developed at the Engineering Research Center for Structured Organic Particulate Systems (C-SOPS), Rutgers University [25]. This schematic of the pilot plant, shown in Figure 5 is installed with several unit operations distributed on three levels from top to bottom. Located on the top floor are three feeders to provide raw materials, including active pharmaceutical ingredient (API), lubricant and excipient. The flow rate of powder could be controlled via the rotating screw inside the feeder. If necessary, a co-mill can be used to delump the API and excipient while lubricant will be directly added into the blender to avoid over lubrication. Both the mill and the blender are mounted on the second floor. After the continuous blender, the homogeneous powder mixture is sent to tablet compaction on the bottom floor.

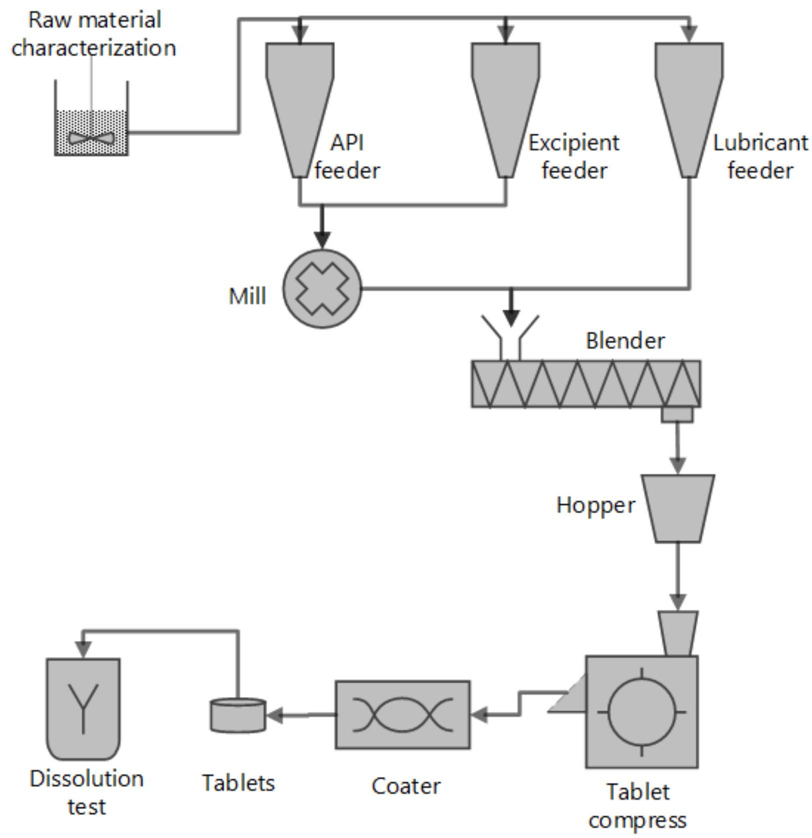

Figure 5 Continuous Tablet Compression Process

5.1.2 DeltaV recipe model

ISA-88 batch manufacturing standard is supported by DeltaV Distributed Control System and recipes can be created and maintained by DeltaV's Recipe Studio application. Recipe Studio supports two types of recipes: master recipes and control recipes. A master recipe Created and modified by process engineers, while control recipe can be Modified and downloaded by operators. All recipes have four main parts: a header, a procedure (sequence or actions), the parameters, and the equipment. They are further elaborated in Table

Table 2 Recipe parts

Name	Function
Header	Contains information about the exact version and the author of the recipe
Procedure	Contains the Procedural Function Chart (PFC) that defines the steps needed for the batch to run.
Parameters	Allow you to set values for different formulations of the product using the same recipe.
Equipment	Defined in DeltaV Explorer and associated with the recipe. Each operation in Recipe Studio is associated with a particular equipment unit.

Instead of the ISA-88 process model, The Procedural Control Model can be implemented in DeltaV Recipe Studio. This model, shown in Figure 6, focuses on describing the process as it relates to physical equipment. Procedures, unit procedures, and operations of the continuous direct compaction process are constructed graphically using IEC 61131-3 compliant Sequential Function Charts (SFCs) (shown in Figure) [26].

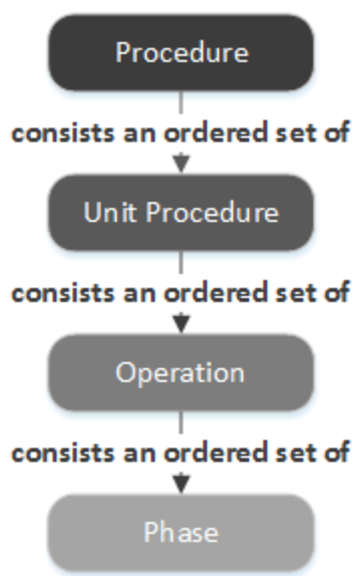

Figure 6 ISA-88 procedure model

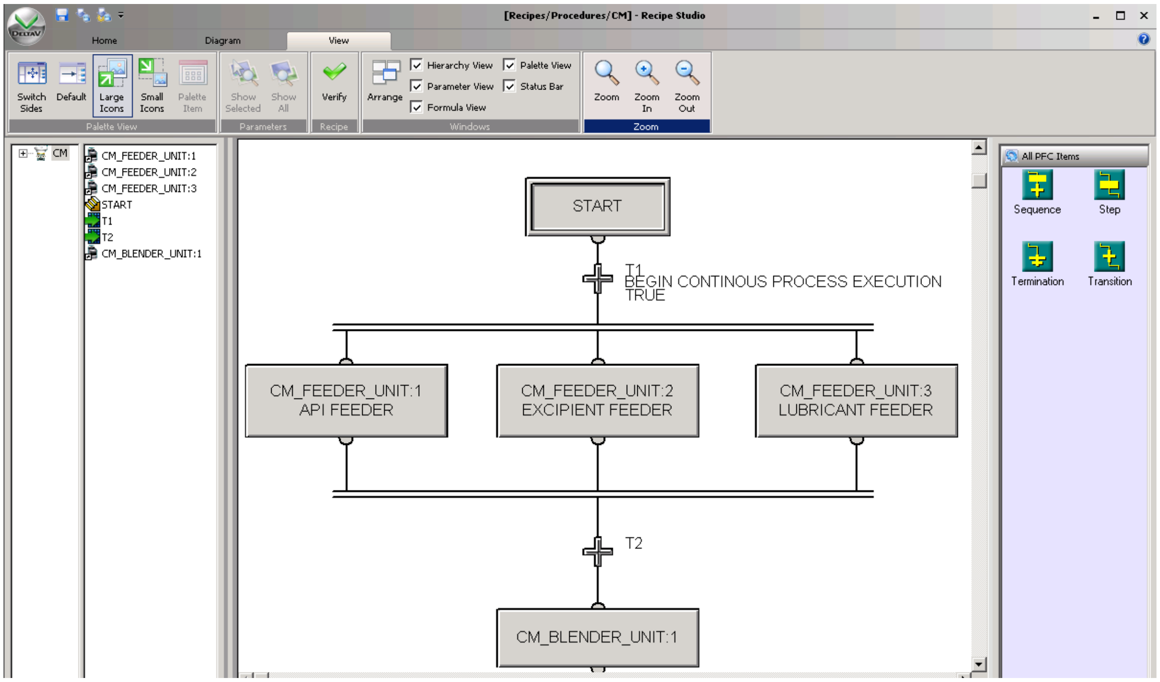

Figure 7 DeltaV recipe structure

While the data generated from field devices is collected by DeltaV, PAT's spectrum data is captured and organized by PAT data management platform, synTQ. synTQ has been designed to provide harmonization and integration of all plant-wide PAT data, including spectral data, configuration data, models, raw and metadata, and Orchestrations (PAT methods). It is worth noting that synTQ is fully compatible with the Emerson Process Management System, which means data collected and managed in synTQ can be transferred into the DeltaV system for process control and data integration.

5.1.3 OSI PI recipe structure

PI system is an enterprise infrastructure for management of real-time data and events provided by OSIsoft. It is a suite of software products that are used for data collection, historicizing, finding, analyzing, delivering, etc. Figure 8 illustrates the structure and data flow of the PI system, in which PI Data Archive and PI Asset Framework (AF) are the keys parts.

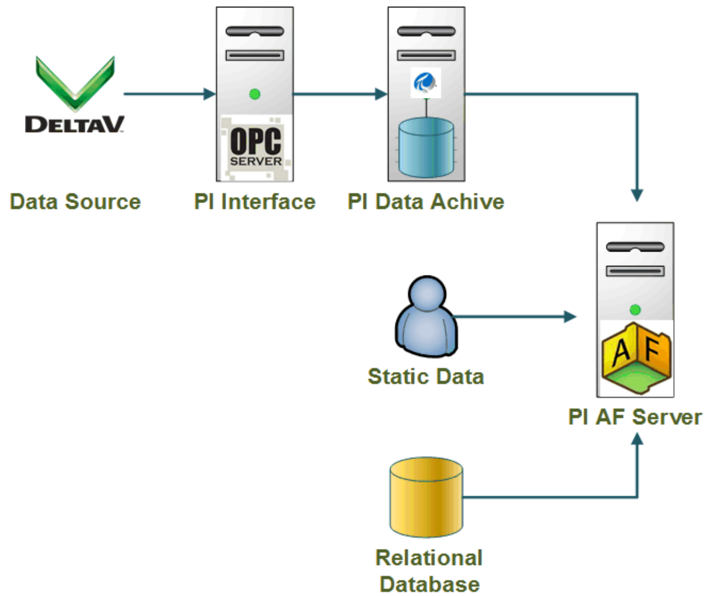

Figure 8 PI system structure

After PI Interface receives data from a data source (DeltaV system in this case) via OPC server, PI Data Archive gets the data and routes it throughout the PI System providing a common set of real-time data. PI AF is a single repository for objects, equipment, hierarchies, and models. It is designed to integrate, contextualize and references data from multiple sources including PI Data Archive and non-PI sources such as human input and external Relational Databases. Together, these metadata and time series data provide a detailed description of equipment or assets. What the hierarchical structure of elements showed in Figure 9 is how equipment from the continuous tablet direct compaction process is mapped in PI System. The attributes of element representing equipment performance parameters are configured to get value from the corresponding PI Data Archive Points.

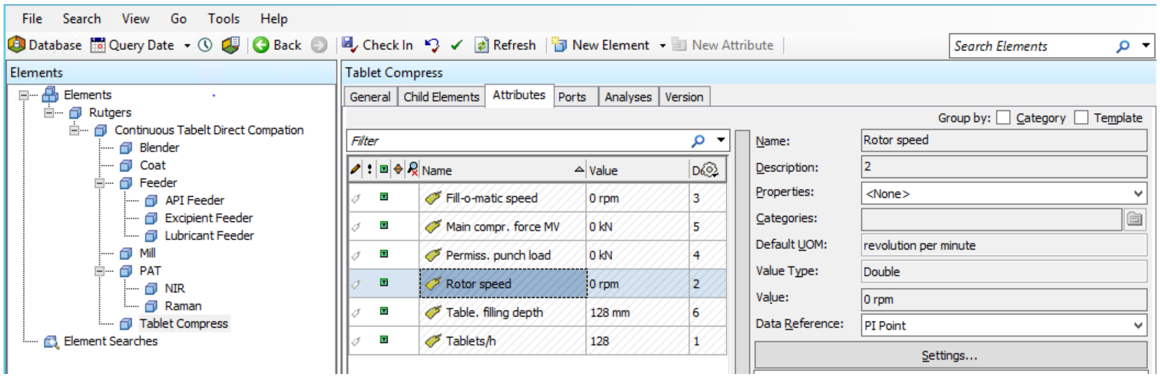

Figure 9 PI AF elements

Moreover, PI AF supports the most significant function of implementing ISA-88 Batch Control Standard, PI Event Frame. Events are critical process time periods that represents something happening and impacting the manufacturing process or operations. The PI Event Frame is intended to capture critical event contexts which could be names, start/end time, and related information that are useful for analysis. Complex hierarchical events, shown in Figure 10, are designed in order to map the ISA-88 process model of continuous tablet press process.

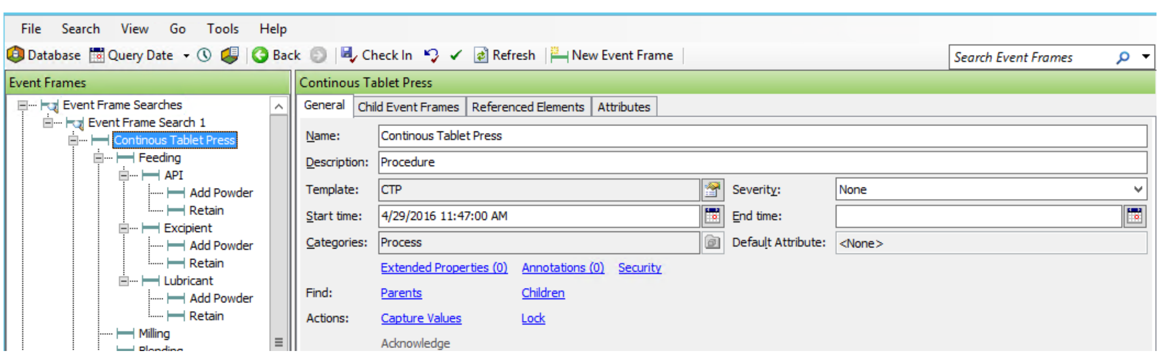

Figure 10 PI Event Frame

5.1.4 Dataflow

Figure 11 present the data flow within the continuous direct tablet compaction process. Data generated from both process equipment and PAT tools will be sent into DeltaV system and organized according to the ISA-88 recipe model. PI system, playing the role of historian, is able to receive data from PCS and build up recipe hierarchical structure using PI Event Frame. In the final step, cloud storage is chosen for permanent data storage and the portal of ERP.

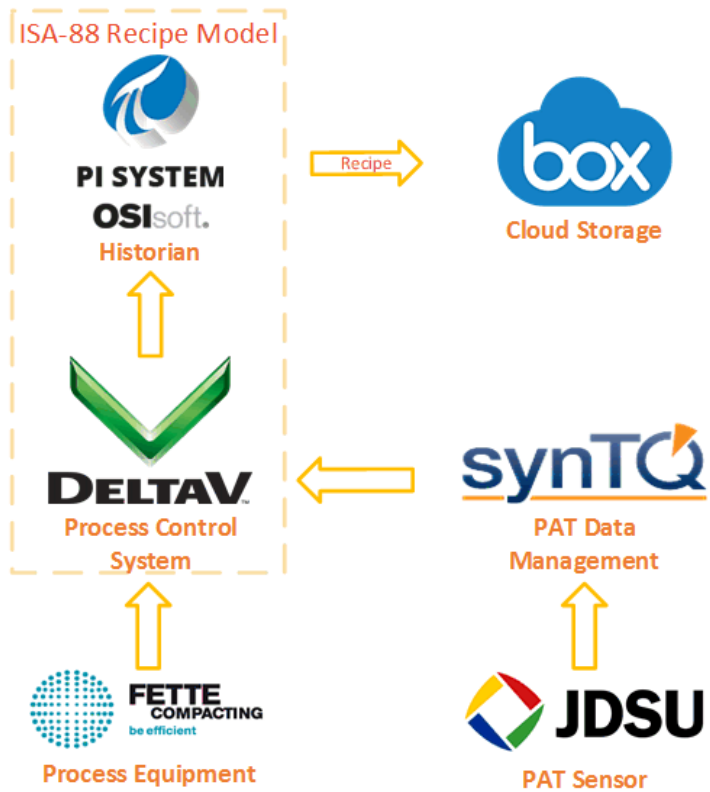

Figure 11 Plant dataflow

5.2 Recipe based ELN system

A novel ELN system is developed and implemented for data management in raw material characterization laboratories to replace paper laboratory notebooks (PLN). Scientists could create, import, modify recipes within ELN following the ISA-88 standards, as well as input data and upload related data files. This system has sufficient capability of documenting experiment process and gathering data from various analytical platforms in order to provide material property information that complies with recipe model.

5.2.1 Information management

This recipe based ELN system is developed as a custom module for Drupal, an open source Content Management System (CMS) written in PHP. The stand release of Drupal, also referred as Drupal core, contains the essential features of CMS, including taxonomy, user account registration, menu management, and system administration. Beside this, Drupal provides many other features, such as high scalability, flexible content architecture, and supporting mobile devices. One widespread distribution of Drupal, Open Atrium is selected as the platform for laboratory data handling because of its advanced knowledge management and security features.

In this case, Drupal is installed and maintained on a cloud computing service, Amazon Web Service (AWS) which is an on-demand computing platform, instead of setting up actual workstations in computer rooms. AWS provides large computing capacity cheaper and quicker than actual physical servers. While Drupal runs in a Linux operating system on Amazon Elastic Compute Cloud (EC2), the MySQL database which Drupal is connected to is supported by Amazon Relational Database Service (RDS). The adoption of cloud computing technology into laboratory data management benefits a lot, including economic computing resource expense, extensible infrastructure capacity, etc.

After installation and configuration, this Drupal system can be accessed in the form of the website via an Internet browser. Within Open Atrium, spaces can be created as the highest level of content structure. While the public space is open to all system users, the content of private space can only be accessed by website members that have been added to such space. After the ELN module is installed in Drupal, a notebook can be created as a section housed in every space. Via the section visibility widget, specific permissions for certain member group or people could be set for ELN section. Although the CMS is accessible across the world through Internet, the confidentiality of content in ELN is well protected.

The web interface of ELN system (Figure 12) contains two main parts: master recipe list and control recipe list. As mention in previous sections, master recipes are the templates for recipes used to perform experiments on individual samples, and they are analytical process dependent. Using ELN tool, master recipes can be created within the system or imported from external recipe document in XML format, as is shown in “Recipe Import” area. After clicking the green “Create Control” button, a control recipe will be generated as one copy of the master recipe with username and time in the name, which will work as the experiment note to document all the experiment process and data. Figure 13 is an example of control recipe for FT4 compressibility test.

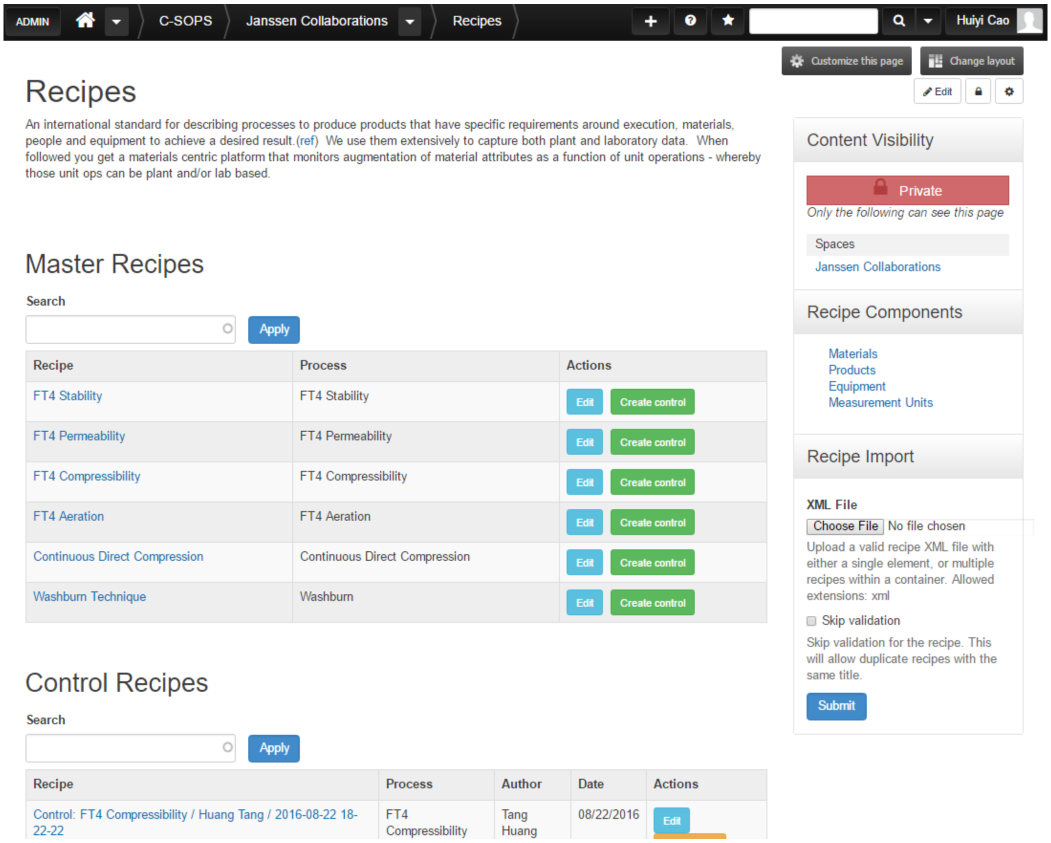

Figure 12 ELN web interface

Figure 13 Control recipe

The “Recipe Components” placed on the right sidebar of ELN front page contains four relevant information may be included in the recipe, and they are further introduced in Table 3. The list of each recipe component includes all the items, as well as their related information and attributes, in a table. Each item is linked to a page listing the recipes that use the particular piece of the component. Figure 14 and Figure 15 show the equipment list and the information page of FT4 Powder Rheometer, for example. The QR code displayed on the top right of the page is another feature of ELN. Each equipment or material existing in ELN system will have a unique QR code, which has the corresponding Uniform Resource Locator (URL) address, encoded. By scanning the QR code attached to the material or equipment, its information, and related recipes could be accessed on the website.

Table 3 Recipe components

Name	Definition	Attributes
Materials	Pharmaceutical ingredient used in manufacturing process	Material Name
		Role
		Aliases & Chemical Info
Products	Things produced from manufacturing process	Product Name
		Catalog
Equipment	The analytical instrument used in material characterization.	Equipment Name
		Equipment ID
		Equipment Description
Measurement Units	A definite magnitude of a quantity	Unit Name
		Unit Symbol

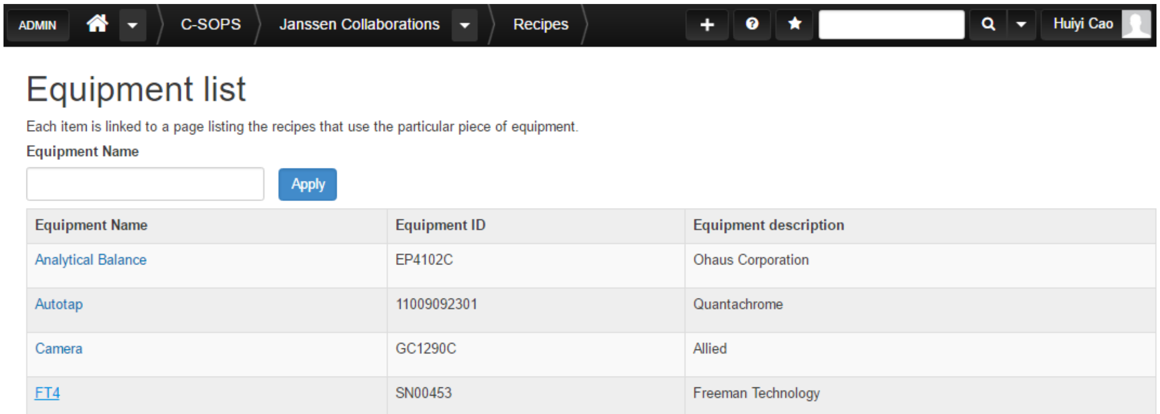

Figure 14 Equipment list

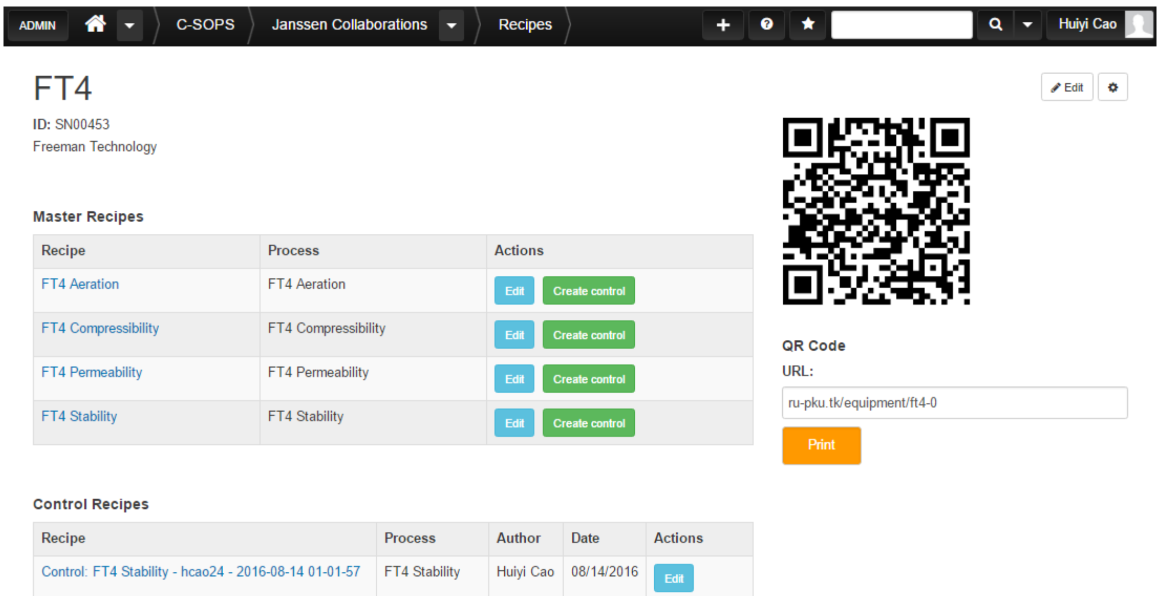

Figure 15 Equipment page of FT4

5.2.2 Workflow

The basic workflow of using ELN to support experiment documentation is illustrated in Figure 16. The first step is to select the master recipe developed for this specific test and create a control recipe. If such master recipe doesn't exist, this issue needs to be reported to laboratory manager because only lab administrators have permission to create and import recipes. However, users are able to change the control recipe, adding/deleting steps or parameters, according to their particular experiment plan. The execution of the recipe is designed to be conducted via a prompt window in the web browser when scientist begins analytical experiment process.

As shown in Figure 17, the left side of this execution dialog display the recipe steps from process to process action. Part of the recipe steps is marked with a green pencil icon which means there are step parameters, step ingredients, or step equipment linked to such steps. Related step information will be presented on the right side of the prompt window. In terms of the step parameters, there are two types of them regarding of the actions need to be taken, “input” and “output”. “Input” means this step parameter will be manually typed into ELN by a scientist while “output” indicates that such parameter is recorded in the output data file from instrument system. Once the step parameters are saved, the green pencil icon will turn into a check mark. The last action in the workflow is to upload the data file generated by analytical device system if there is any.

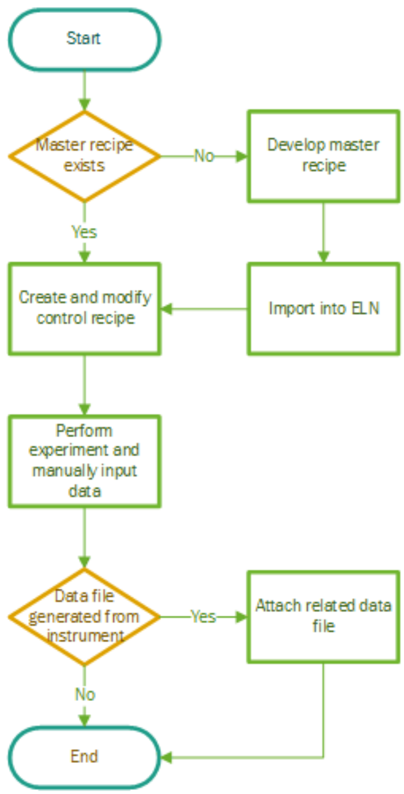

Figure 16 Workflow of using ELN

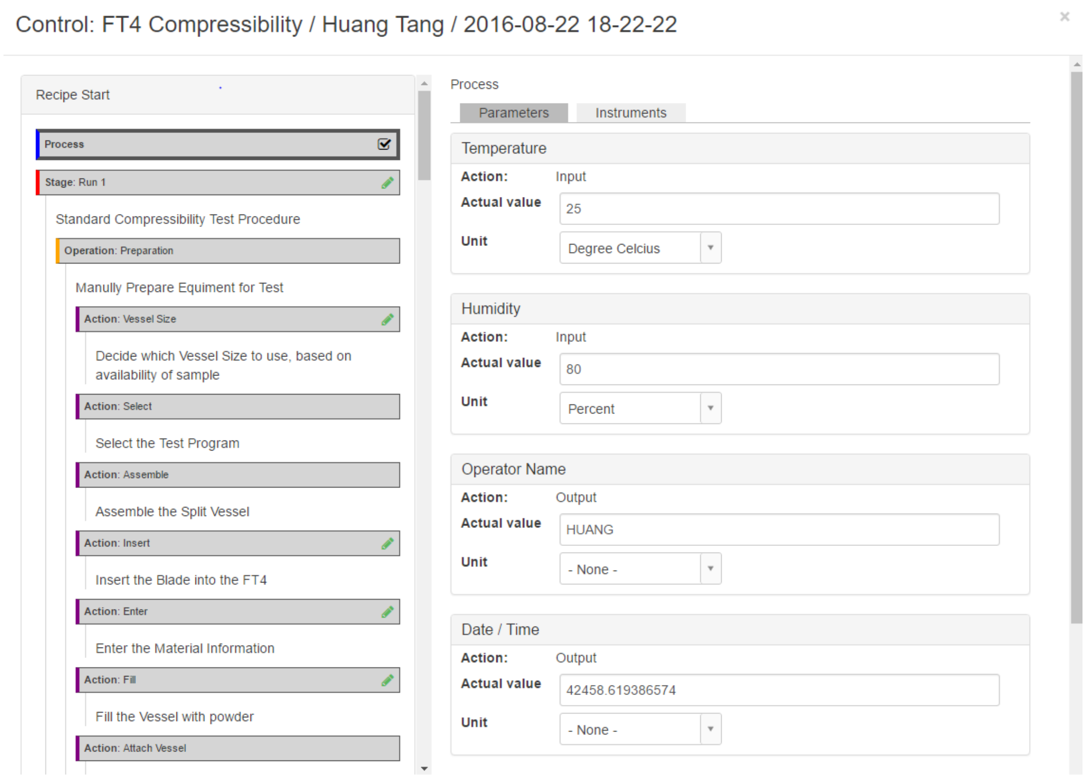

Figure 17 Execution dialog of control recipe

5.2.3 Data flow

There are two primary data sources for ELN system, human input and data file outputted by the instrument. Both of them will be loaded into control recipe of the experiment as the actual value of step parameters. As mentioned before, manually inputting data can be done via the execution dialog shown in Figure 17. However, the extraction of data from instrument-generated data file will be performed by ELN system's Excel Parsing function. Considering the fact that most of the analytical device systems have the ability to export data into Excel supported documents (xlsx and csv for example), ELN's Excel Parsing capability is developed based on the contributed Drupal module "PHPExcel". After the data file is attached to the recipe, PHPExcel module will convert the data into an array. Once there is a match within the step parameters, the value of an array item will be updated to that step parameter. Therefore, ELN supports various of analytical instruments, as well as manually data input.

Containing the information of the analytical process, test results and instrument information, each control recipe is an integral documentation of experiment. ELN system can export control recipe into XML document for archiving or data transformation. Drupal has the built-in functionality of dumping content into JavaScript Object Notation (JSON) format, which an alternative to XML for storing and exchanging data. Thus, the recipe is transformed into JSON document at first. To complete a fast and strict conversion between JOSN and XML format, a small application is developed in Haskell language. After these two steps, an XML file of control recipe is generated from ELN system, which is suitable for archiving and sharing information.

Figure 18 shows data flow within ELN system using FT4 Powder Rheometer Characterization for example. In this case, the final step is uploading recipe document into BOX database, which is intended for data warehousing. BOX is a cloud computing business that provides content management and file sharing service. Keeping data in cloud storage is a secure and economical method of data archiving. This is also the cornerstone of the data analysis and visualization that also happens on cloud computing.

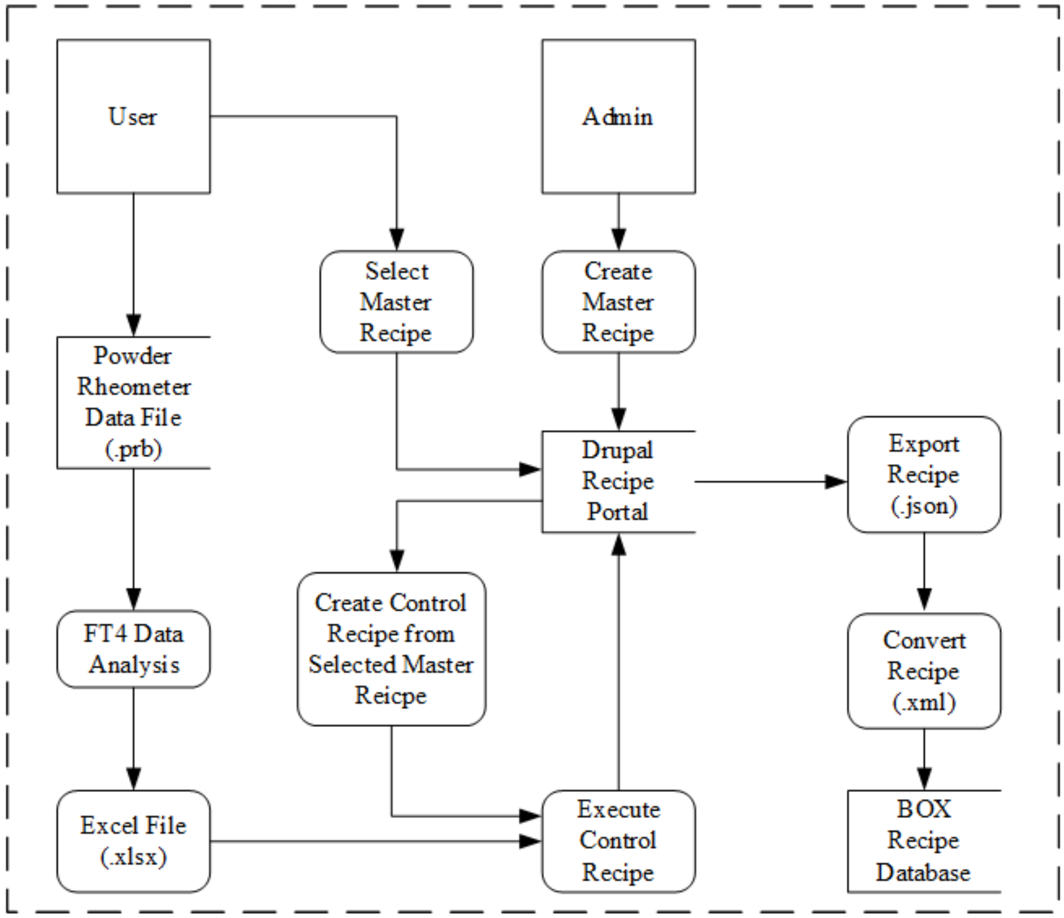

Figure 18 Data flow in FT4 experiment

5.2.4 Other features

5.2.4.1 QR code printing

It is worth mentioning that the ELN system has the capability of directly printing QR codes for recipe components via a barcode printer. While there always exists a unique QR code corresponded to the process material and equipment, a QR code sidebar can certainly be found on the material and equipment information web pages. After simply clicking the "Print" button, the QR code will be printed on a sticker from the barcode printer connected to the computer. Such convenient feature is enabled by the custom Drupal module “zprint”, which is designed to send the command of printing current web URL to the printer via the Google Cloud Print service.

5.2.4.2 Mobile device supporting

By reason of ELN system’s web interface, this recipe portal could be accessed via all kinds of mobile devices not only desktops or laptops. Scientists can open up recipes through the internet browsing apps on their smartphones or tablets, as well as executing recipe. Moreover, mobile devices are more convenient for scanning QR codes to acquire information of material, equipment, and samples.

Chapter 6 Conclusions and Future Work

An ontological informatics infrastructure for data management for continuous pharmaceutical manufacturing has been developed. ISA-88 Batch Control Standard is adapted to continuous manufacturing in order to provide a design philosophy, as well as reference models. The implementation of such recipe model into the continuous process is well summarized and assessed. The recipe data warehouse strategy is used for the purpose of data integration in both manufacturing plant and material characterization laboratory. While the proper communication among PLC, PAT, PCS, and historian enable the data collection and transformation across the continuous manufacturing plant, a recipe based ELN system is in charge of capturing data from various analytical platforms. Therefore, data from different process level and distributed location can be integrated and contextualized with meaningful information in order to support process control and decision making [27].

Considering the fact that pharmaceutical industry is contributing more efforts to this area, the number of technologies applied in continuous manufacturing will keep increasing for sure. Future work could focus on validating the possibility of using this data management strategy to support Design of Experiment (DOE). In terms of the laboratory platform, it is worth trying to include instrument configuration and calibration information into experiment recipes, which might be helpful to error analysis.

References

1. Singh R, Sahay A, Muzzio F, Ierapetritou M, Ramachandran R. A systematic framework for onsite design and implementation of a control system in a continuous tablet manufacturing process. *Computers & Chemical Engineering.* 2014;66:186-200.
2. Singh R, Roman-Ospino AD, Romanach RJ, Ierapetritou M, Ramachandran R. Real time monitoring of powder blend bulk density for coupled feed-forward/feed-back control of a continuous direct compaction tablet manufacturing process. *Int J Pharm.* Nov 10 2015;495(1):612-625.
3. Lee SL, O'Connor TF, Yang X, et al. Modernizing Pharmaceutical Manufacturing: from Batch to Continuous Production. *Journal of Pharmaceutical Innovation.* 2015;10(3):191-199.
4. FDA. Guidance for Industry PAT - A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance. 2004.
5. Venkatasubramanian V, Zhao C, Joglekar G, et al. Ontological informatics infrastructure for pharmaceutical product development and manufacturing. *Computers & chemical engineering.* 2006;30(10):1482-1496.
6. ISA. Batch Control - Part 1: Models and Terminology1995.
7. ISA. Batch Control - Part 2: Data Structures and Guidelines for Languages2001.
8. ISA. Batch Control - Part 3: General and site recipe models and representation2003.
9. Dorresteijn RC, Wieten G, Santen PTEv, et al. Current good manufacturing practice in plant automation of biological production processes. *Cytotechnology.* 1997;23:19-28.

10. Verwater-Lukszo Z. A practical approach to recipe improvement and optimization in the batch processing industry. *Computers in Industry.* 1998;36(3):279-300.
11. Kimball R, Ross M. *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*: Wiley Computer Publishing; 2013.
12. Fermier A, McKenzie P, Murphy T, Poulsen L, Schaefer G. Bringing New Products to Market Faster. *Pharmaceutical Engineering.* Vol 322012:1-8.
13. Ierapetritou M, Muzzio F, Reklaitis G. Perspectives on the continuous manufacturing of powder-based pharmaceutical processes. *AIChE Journal.* 2016;62(6):1846-1862.
14. Armbrust M, Stoica I, Zaharia M, et al. A view of cloud computing. *Communications of the ACM.* 2010;53(4):50.
15. Subramanian B. The disruptive influence of cloud computing and its implications for adoption in the pharmaceutical and life sciences industry. *Journal of Medical Marketing: Device, Diagnostic and Pharmaceutical Marketing.* 2012;12(3):192-203.
16. De Minicis M, Giordano F, Poli F, Schiraldi MM. Recipe Development Process Re-Design with ANSI/ISA-88 Batch Control Standard in the Pharmaceutical Industry. *International Journal of Engineering Business Management.* 2014:1.
17. Gruber TR. A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition.* 1993.
18. Swartout WR, Neches R, Patil R. Knowledge Sharing - Prospects and Challenges. *Proceeding fo the International Conference on Building and Sharing of Very Large-Scale Knowledge Bases '93*. Tokyo, Japan1993.
19. W3C. Extensible Markup Language (XML) 1.0 (Fifth Edition). 2008; https://www.w3.org/TR/2008/REC-xml-20081126/.

20. W3C. W3C XML Schema Definition Language (XSD) 1.1 Part 1: Structures. 2012; https://www.w3.org/TR/xmlschema11-1/.
21. Muñoz E, Capón-García E, Espuña A, Puigjaner L. Ontological framework for enterprise-wide integrated decision-making at operational level. *Computers & Chemical Engineering.* 2012;42:217-234.
22. IEC. IEC 62264-1:2013. *Enterprise-control system integration -- Part 1: Models and terminology*: IEC; 2013.
23. Hwang Y, Grant D. An empirical study of enterprise resource planning integration: global and local perspectives. *Information Development.* 2014;32(3):260-270.
24. Alphonsus ER, Abdullah MO. A review on the applications of programmable logic controllers (PLCs). *Renewable and Sustainable Energy Reviews.* 2016;60:1185-1205.
25. Singh R, Ierapetritou M, Ramachandran R. An engineering study on the enhanced control and operation of continuous manufacturing of pharmaceutical tablets via roller compaction. *Int J Pharm.* Nov 15 2012;438(1-2):307-326.
26. Godena G, Lukman T, Steiner I, Bergant F, Strmčnik S. A new object model of batch equipment and procedural control for better recipe reuse. *Computers in Industry.* 2015;70:46-55.
27. Meneghetti N, Facco P, Bezzo F, Himawan C, Zomer S, Barolo M. Knowledge management in secondary pharmaceutical manufacturing by mining of data historians-A proof-of-concept study. *Int J Pharm.* May 30 2016;505(1-2):394-408.

Appendix A. Xml Schema Document

```
<!-- An ISA88 "Recipe". -->
  <xs:complexType name="RecipeType">
    <xs:sequence>
      <xs:element name="Process" type="ProcessType" minOccurs="1" maxOccurs="1"/>
      <xs:element name="Product" type="ProductType" minOccurs="0" maxOccurs="1"/>
      <xs:element name="Material" type="MaterialType" minOccurs="0" maxOccurs="1"/>
      <xs:element name="Site" type="SiteType" minOccurs="0" maxOccurs="1"/>
      <xs:element name="Author" type="PersonType" minOccurs="0" maxOccurs="unbounded"/>
      <xs:element name="Approver" type="PersonType" minOccurs="0" maxOccurs="unbounded"/>
    <xs:attribute name="Name" type="RecipeNameType" use="required"/>
    <xs:attribute name="Description" type="xs:string" use="required"/>
    <xs:attribute name="Status" type="RecipeStatusType" use="required"/>
    <xs:attribute name="EffectiveDate" type="xs:date"/>
    <xs:attribute name="Version" type="xs:decimal" use="required"/>
    <xs:attribute name="Path" type="RecipePathType" use="required"/>
    <xs:attribute name="Type" type="RecipeTypeType" use="required"/>
    <!-- The "Type" field is necessary, as including the full recipe tree is
    optional. -->
  </xs:complexType>
  <!-- A "Process" is always the sole root of a "Recipe"'s step tree, and has only
   "Stage"'s for children. -->
  <xs:complexType name="ProcessType">
    <xs:complexContent>
      <xs:extension base="StepType">
        <xs:sequence>
          <xs:element name="Stage" type="StageType" minOccurs="0" maxOccurs="unbounded"/>
        </xs:sequence>
      </xs:extension>
    </xs:complexContent>
  </xs:complexType>
  <!-- A "Stage" has only "Operations"'s for children. -->
  <xs:complexType name="StageType">
    <xs:complexContent>
      <xs:extension base="StepType">
        <xs:sequence>
          <xs:element name="Operation" type="OperationType" minOccurs="0"
maxOccurs="unbounded"/>
        </xs:sequence>
      </xs:extension>
    </xs:complexContent>
  </xs:complexType>
  <!-- An "Operation" has only "Actions"'s for children. -->
```

```
<xs:complexType name="OperationType">
  <xs:complexContent>
    <xs:extension base="StepType">
      <xs:sequence>
        <xs:element name="Action" type="ActionType" minOccurs="0" maxOccurs="unbounded"/>
      </xs:sequence>
    </xs:extension>
  </xs:complexContent>
</xs:complexType>
<!-- An "Action" is always a leaf of the step tree. -->
<xs:complexType name="ActionType">
  <xs:complexContent>
    <xs:extension base="StepType"/>
  </xs:complexContent>
</xs:complexType>
```

Appendix B. Unit Operation Recipe Example


```
<?xml version="1.0" encoding="UTF-8"?>
<Recipe
  Description="Unit Operation Recipe"
  Name="Tablet Pressing"
  Status="Draft"
  Type="Library"
  Version="1.00">
    <Process
      Description="Tablet Compression Process"
      Name="Compression"
      <Stage
              Description="Tablet Compression Process"
              Name="Compression"
              <StepArgs>
              <StepParameter Action="Input" Name="Yield"/>
              <StepParameter Action="Input" Name="Lower Punch Drawing ID"/>
            <StepParameter Action="Input" Name="Upper Punch Drawing ID"/>
              <StepParameter Action="Input" Name="Die Drawing ID"/>
              <StepParameter Action="Input" Name="Feed Configuration"/>
        </StepArgs>
        <Operation
          Description=""
          Name="Store Product"
            <StepArgs>
              <StepParameter Action="Output" Name="Reconciliation"/>
              <StepParameter Action="Input" Name="Number of Tablets"/>
              <StepParameter Action="Output" Name="Average Tablet Weight"/>
            </StepArgs>
```