Article

A Systematic Framework for Data Management and Integration in a Continuous Pharmaceutical Manufacturing Processing Line

Huiyi Cao [1], Srinivas Mushnoori [2], Barry Higgins [3], Chandrasekhar Kollipara [4], Adam Fermier [4], Douglas Hausner [1], Shantenu Jha [2], Ravendra Singh [1], Marianthi Ierapetritou [1] and Rohit Ramachandran [1,*]

[1] Department of Chemical and Biochemical Engineering, Engineering Research Center for Structured Organic Particulate Systems (C-SOPS), Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA; hc623@scarletmail.rutgers.edu (H.C.); douglas.hausner@rutgers.edu (D.H.); ravendra.singh@rutgers.edu (R.S.); marianth@soe.rutgers.edu (M.I.)
[2] Department of Electrical & Computer Engineering, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA; scm177@scarletmail.rutgers.edu (S.M.); shantenu.jha@rutgers.edu (S.J.)
[3] Johnson & Johnson, Co., Cork, Ireland; bhiggin2@its.jnj.com
[4] Janssen Research & Development, Spring House, PA 19002, USA; ckollip8@its.jnj.com (C.K.); afermier@its.jnj.com (A.F.)
[*] Correspondence: rohit.r@rutgers.edu; Tel.: +1-848-445-6278

check for
updates

Abstract: 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, and reliable manner is a big challenge for the pharmaceutical industry. In this work, an ontological information infrastructure is developed to integrate data within manufacturing plants and analytical laboratories. The ANSI/ISA-88.01 batch control standard 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. All the detailed information of the lab-based experiment and process manufacturing, including equipment, samples and parameters, are documented in the recipe. This recipe model is implemented into a process control system (PCS), data historian, as well as Electronic Laboratory Notebook (ELN) system. Data existing in the recipe can be eventually exported from this system to cloud storage, which could provide a reliable and consistent data source for data visualization, data analysis, or process modeling.

Keywords: data management; continuous pharmaceutical manufacturing; ISA-88; recipe; OSI Process Information (PI)

1. Introduction

For decades, the pharmaceutical industry has been dominated by a batch-based manufacturing process. This traditional method 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 economical production. The US Food and Drug Association (FDA) has recognized the advancement of this manufacturing mode and has been encouraging its development as part of the FDA's QbD paradigm [1]. The application of process analytical technology (PAT) and control systems is a very

useful effort to gain improved science-based process understanding [2]. One of the advantages of continuous pharmaceutical manufacturing process is that it provides the ability to monitor and rectify data/product in real time. Therefore, it has been considered a data rich manufacturing process. However, in the face of the enormous amount of data generated from a continuous process, a sophisticated data management system is required for the integration of analytical tools to the control systems, as well as the off-line measurement systems.

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 the pharmaceutical industry [3]. The ISA-88 Batch Control Standard [4–6] is an international standard addressing batch process control, which has already been implemented in other industries for years. Therefore, adapting this industrial standard into pharmaceutical manufacturing could provide a design philosophy for describing equipment, material, personnel, as well as reference models [7,8]. This recipe-based execution could work as a hierarchical data structure for the assembly of data from the control system, process analytical technology PAT tools, and off-line measurement devices. The combination of the ISA-88 recipe model and the data warehouse informatics strategy [9] leads to the "recipe data warehouse" strategy [10]. This strategy could provide the possibility of the data management across multiple execution systems, as well as the ability for 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 experimentation and a 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 interconnected unit operations with multiple execution layers. Therefore, it is 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 the 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 product [11]. These properties of relevance to continuous manufacturing are measured via many different analytical methods, including FT4 powder characterization [12], particle size analysis [13], Washburn technique [14], 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 [15]. Because of its many advantages including flexibility, security, and efficiency, cloud computing is suitable for data management in the pharmaceutical industry [16].

Objectives

In this work, 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 a 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.

2. Materials and Methods

2.1. 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.

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 [19]. XML is a metalanguage that defines a set of rules for encoding documents in a format which is both human readable and machine readable. 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].
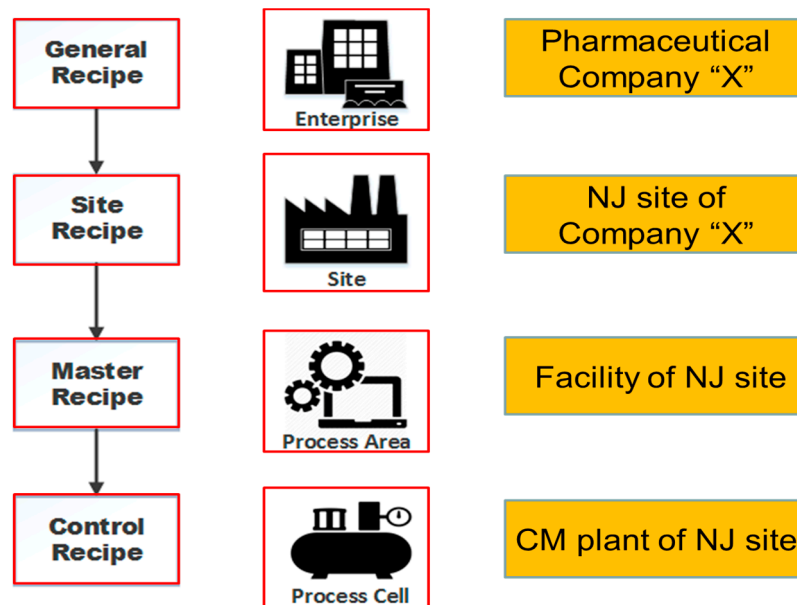
2.2. Recipe Model

In ISA-88, a recipe is defined as the minimum set of information that uniquely identifies the production requirements for a particular product [4]. 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.

2.2.1. 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. The different type of the recipe together with an example from pharmaceutical industry is shown in Figure 1.

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.

2.2.2. 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 has been displayed in Figure 2. As shown in the figure, the higher level consists an ordered set of the lower levels. The information of equipment, parameter, and material can be included in the attributes of the process actions.
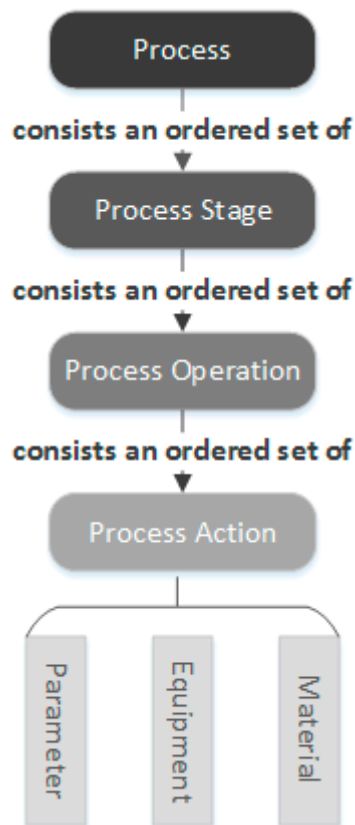
Figure 2. ISA-88 process model.

It is important to note 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 [21]. In this case, the 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.

2.2.3. 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 the recipe model is divided into several steps to perform the activities in a structured way.

- 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 offline characterization experiment of material to be applicable to the ISA-88 standard. In this situation, a technique-specific recipe containing detailed process steps could provide instructions for specific experiment type, as well as the designed structure for result documentation. As shown in Figure 3, after analytical tests are performed on samples according to the procedures in the recipes, experiment results will be recorded in the recipe as well.
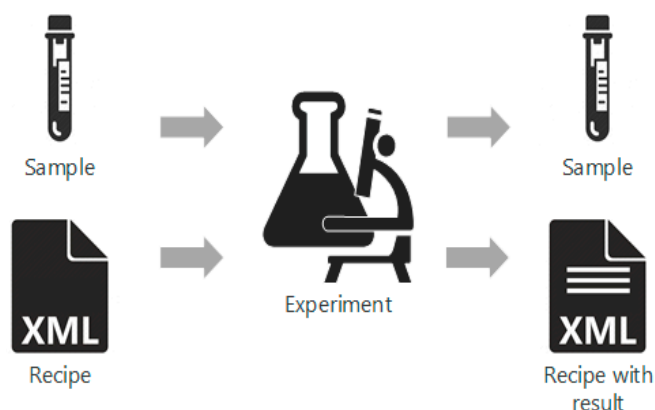
Figure 3. Material and data flow in an experiment process.

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.

- 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. This step needs to be completed 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.

- 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, depending on the equipment used. This does happen a lot in the product process development stage, especially in the early time period. Another barrier 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, and blending. Each unit operation could be performed by different equipment types that correspond to a different recipe module. In other words, each equipment type might take place in a continuous process and 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, Near-infrared 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.

- Step 4. Process mapping

XML-based master recipes will be developed in accordance with an XML Schema to represent the manufacturing process. Firstly, 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 of unit operations.

2.3. Systematic Framework of Integration

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 meaningful information [22]. For continuous pharmaceutical manufacturing, a data flow across the whole process is proposed in Figure 1 following the traditional design pattern, automation pyramid [23], to create an information and communication infrastructure.

Enterprise resource planning (ERP) systems are placed at the top of structure shown in Figure 4. It provide the integrated view of the core business process, based on the data collected from various business activities [24]. Long-term resource planning—primarily human and material resources—will be its primary objective. Historian, on 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.
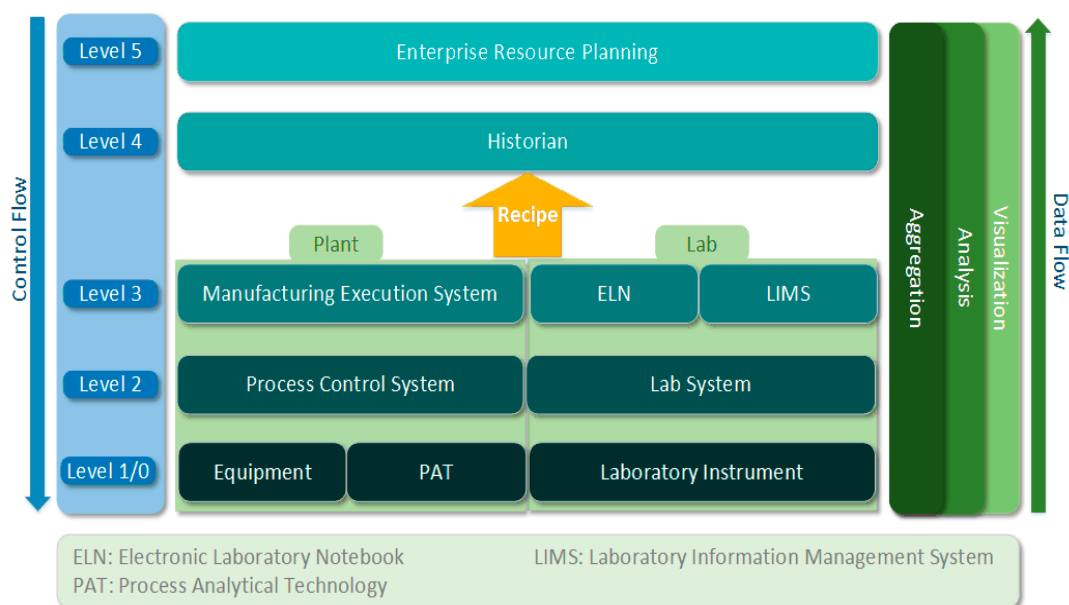

Figure 4. Architecture of data integration.

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.

2.3.1. Continuous Manufacturing

As shown in Figure 4, 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 manufacturing process, the 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 [25]. 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. The low-level network in this information architecture is mainly based on the communication over bus systems, such as field device to PLC. However, 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.

2.3.2. 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 is generated in recipe form according to the ISA-88 standard in Level 3.

The basic workflow of using ELN to support experiment documentation is illustrated in Figure 5. 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. After the execution of the control recipe, user will perform experiments according the steps in the recipe and recording test parameters at the same time. Eventually, all the data related to this single test—including operator, sample, instrument, and result—will be documented in the recipe and be transferred to cloud storage.
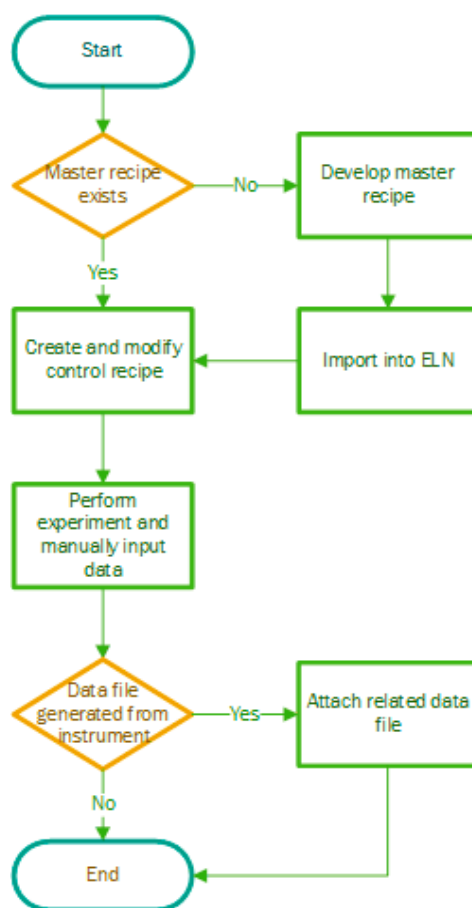
Figure 5. Workflow of using electronic lab notebook (ELN) to perform an experiment.

2.4. Transferring Data from Lab Computer to Cloud System

Transfer of data from a computing cloud to a storage cloud can be non-trivial, since cross platform support is not guaranteed to be available on the cloud service of choice. There is the option of using the application program interface (API) provided by the cloud service but the heterogeneity in the API's provided by different services makes this approach difficult. This issue is highlighted in our current case: the cloud platform of choice here is Box.com. The computing cloud, i.e., the Amazon Elastic Compute Cloud (EC2) runs a Linux distribution (Amazon Linux AMI) based on RHEL (Red Hat Enterprise Linux) and CentOS. Box.com, however, neither has support for, nor has any plans to include support for this platform. This puts the user in a situation where there is no direct method to easily and trivially push data to the cloud using a tool such as an app. Additionally, there is the difficulty in transitioning to a different platform should that need ever arise.

In order to address these issues, we have put in place a data transfer pipeline between the compute and storage clouds using davfs2 [26], a specific implementation of WebDAV (web-based Distributed Authoring and Versioning) [27], an extension of the hypertext transfer protocol (HTTP) that allows a user (client) to remotely create/edit web content. This method makes it possible to set up a directory on the Linux EC2 compute cloud that serves as a window into the cloud storage location. Now a rsync [28] may be set up to sync to this directory, and by extension, to the cloud.

3. Case Studies

Two different case studies have been performed to verify the proposed systematic framework in both pilot-plant and analytical laboratory.

3.1. Direct Compaction Continuous Tablet Manufacturing Process

A direct compaction continuous pharmaceutical tablet manufacturing process has been developed at the Engineering Research Center for Structured Organic Particulate Systems (C-SOPS), Rutgers University [29]. This schematic of the process is illustrated in Figure 6. Three gravimetric feeders are used to provide raw materials, including active pharmaceutical ingredient (API), lubricant and excipient. The flow rate of powder is controlled via manipulating the rotational screw speed of 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. After the continuous blender, the homogeneous powder mixture is sent to tablet compaction. Some of the tablets produced will be sent for dissolution testing. In order to use the gravity, the pilot plant is built in three levels high for better material flow. The feeders (K-Tron) are located on the top level. The second level is used for delumping and blending and the tablet compactor (FETTE) is located on the bottom floor.
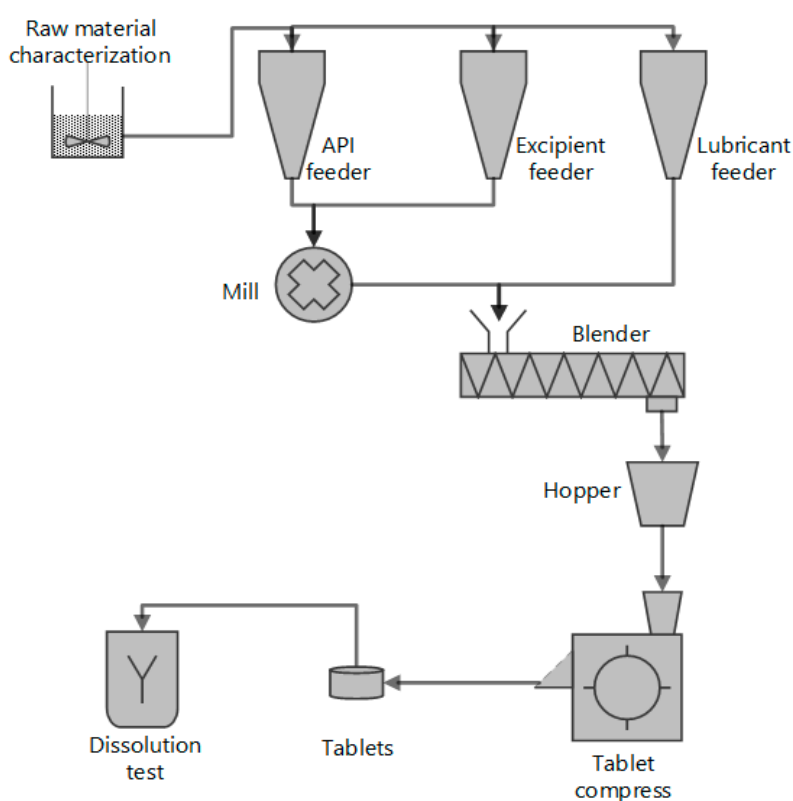
Figure 6. Continuous tablet compression process.

3.2. FT4 Powder Characterization Test

The input materials of continuous direct compaction tablet manufacturing process are powders whose physical and mechanical properties vary. Since this variation may have an impact on the performance of the final dosage, it is important to develop effective measurement methods on the critical properties. The FT4 Powder Rheometer of Freeman Technology developed over the last two decades is an ideal solution.

The FT4 Powder Rheometer (Figure 7) is designed to characterize powders under different conditions in ways that resemble large-scale production environments. It provides a comprehensive series of methods that allow powder behavior to be characterized across a whole range of process conditions. The methods include rheological, torsional shear, compressibility and permeability tests which can be performed using small bulk samples, such as 1, 10 or 25 mL.
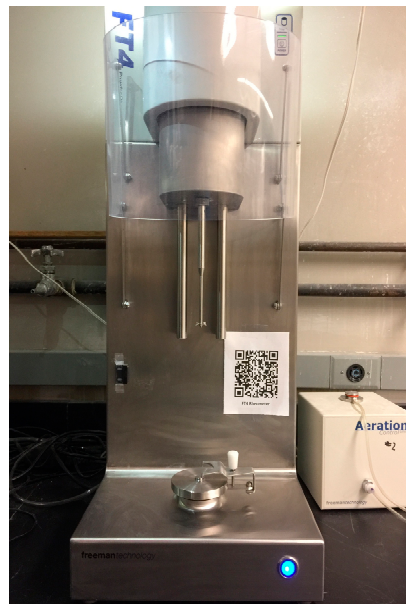
Figure 7. FT4 powder rheometer.

4. Results and Discussions

4.1. Data Integration in Continuous Tablet Manufacturing

Figure 8 presents the data flow within the continuous direct compaction tablet manufacturing 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.
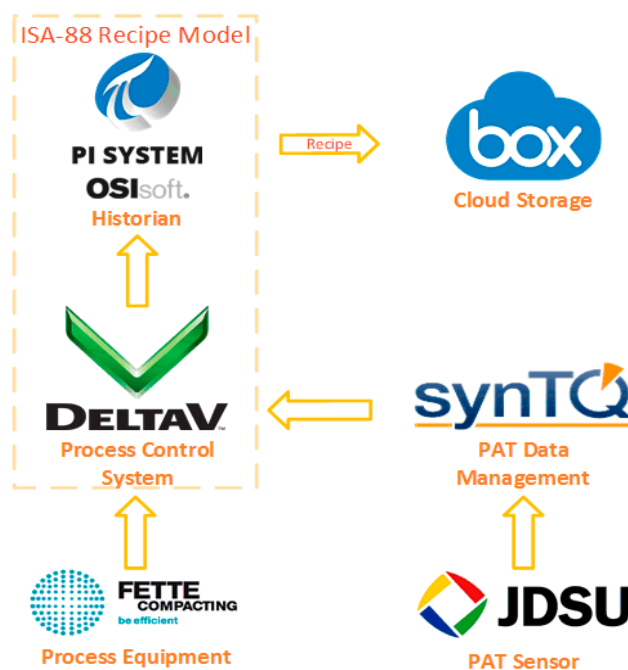
Figure 8. Data flow in continuous tablet manufacturing.

4.1.1. DeltaV Recipe Model

ISA-88 batch manufacturing standard is supported by DeltaV Distributed Control System. Recipes has been 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 2.

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 9, 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) [30].
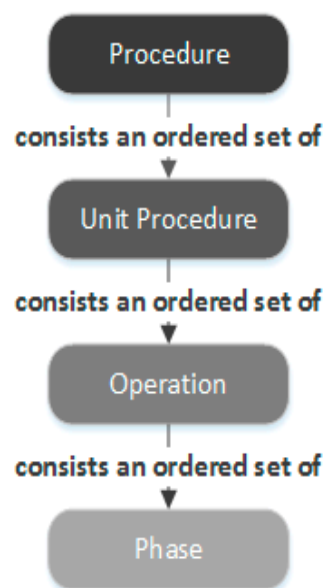
Figure 9. ISA-88 procedure model.

While the data generated from field devices is collected by DeltaV, PAT's spectrum data is captured and organized by the 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. This feature means data collected and managed in synTQ can be transferred into the DeltaV system for process control and data integration.

4.1.2. 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 10 illustrates the structure and data flow of the PI system, in which PI data archive and PI asset framework (AF) are the keys parts.
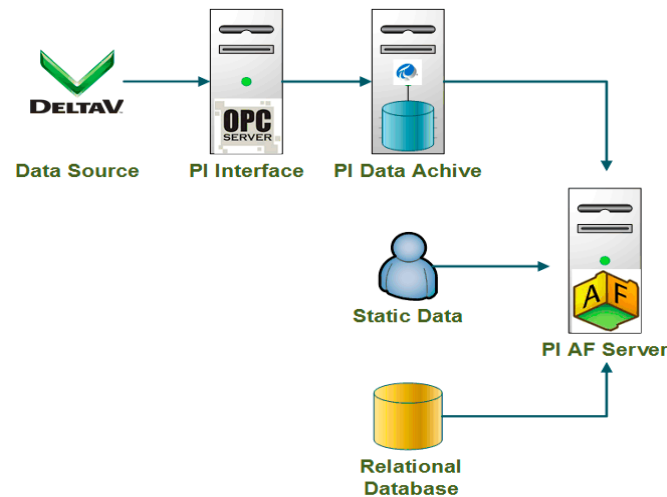

Figure 10. Structure and data flow of Process Information (PI) system.

After the PI interface receives data from a data source (DeltaV system in this case) via the OPC server, the 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 reference data from multiple sources including the 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. Figure 11 shows 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.
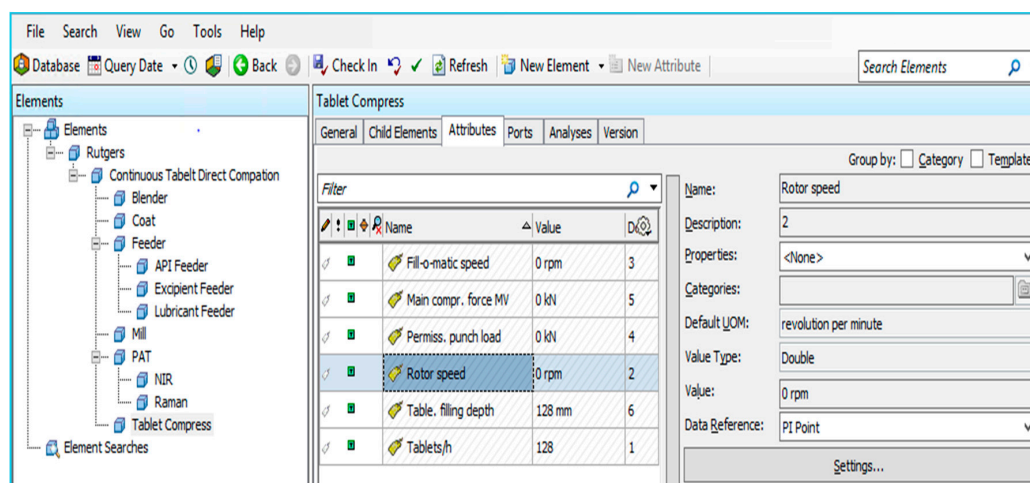
Figure 11. Include equipment into PI asset framework (AF) elements.

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 12, are designed in order to map the ISA-88 process model of continuous tablet press process.
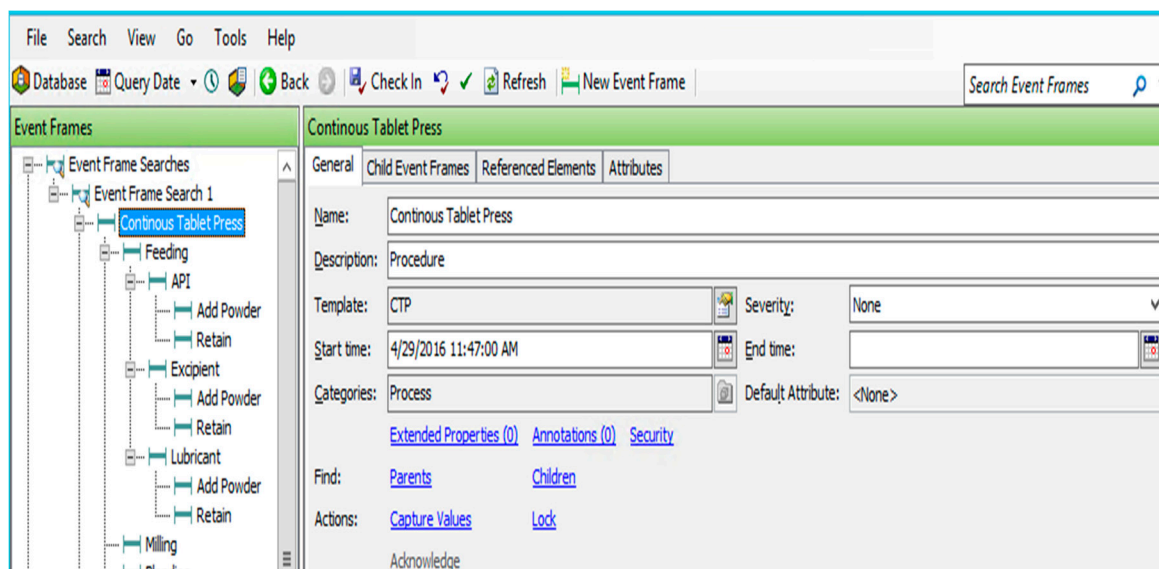
Figure 12. Using PI Event Frame to map ISA-88 recipe model.

4.2. 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. It could also provide material property information that complies with recipe model.

4.2.1. Information Management

This recipe-based ELN system is developed as a custom module for Drupal, an open-source content management system (CMS) written in Hypertext Preprocessor (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). AWS 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. Drupal runs in a Linux operating system on one of AWS' service, 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 13) 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. It will work as the experiment note to document all the experiment process and data. Figure 14 is an example of control recipe for FT4 compressibility test.
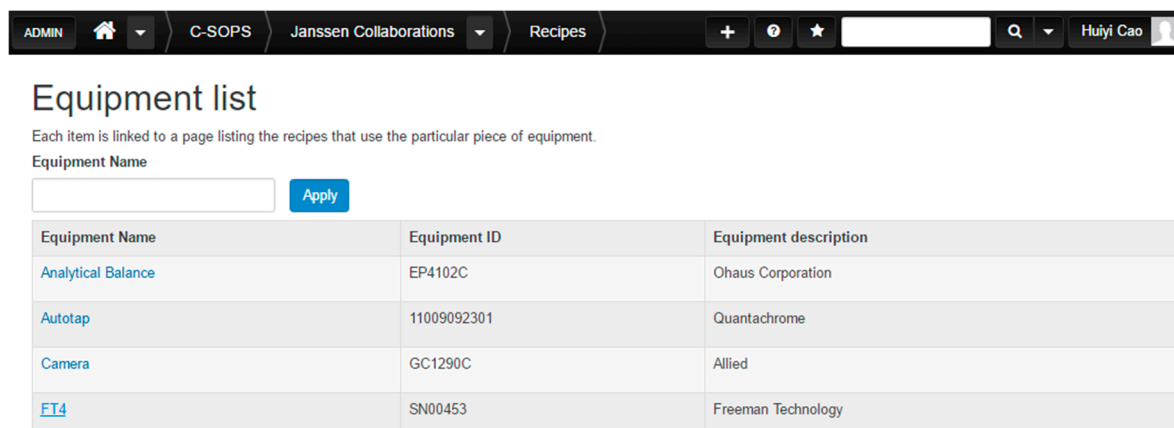

Figure 13. The web interface of ELN system.

Figure 14. Example of 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. Each item is linked to a page listing the recipes that use the particular piece of the component. Figures 15 and 16 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 15. Equipment list in ELN system.

Figure 16. Equipment page of FT4 power rheometer.

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. "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 17. Execution dialog of control recipe.

4.2.2. Other Features

- 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. Within this system, 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". It is designed to send the command of printing current web URL to the printer via the Google Cloud Print service.

- 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.

4.2.3. Data Flow

There are two primary data sources for the ELN system: human input and data file outputted by the instrument. Both will be loaded into the 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 18. 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.
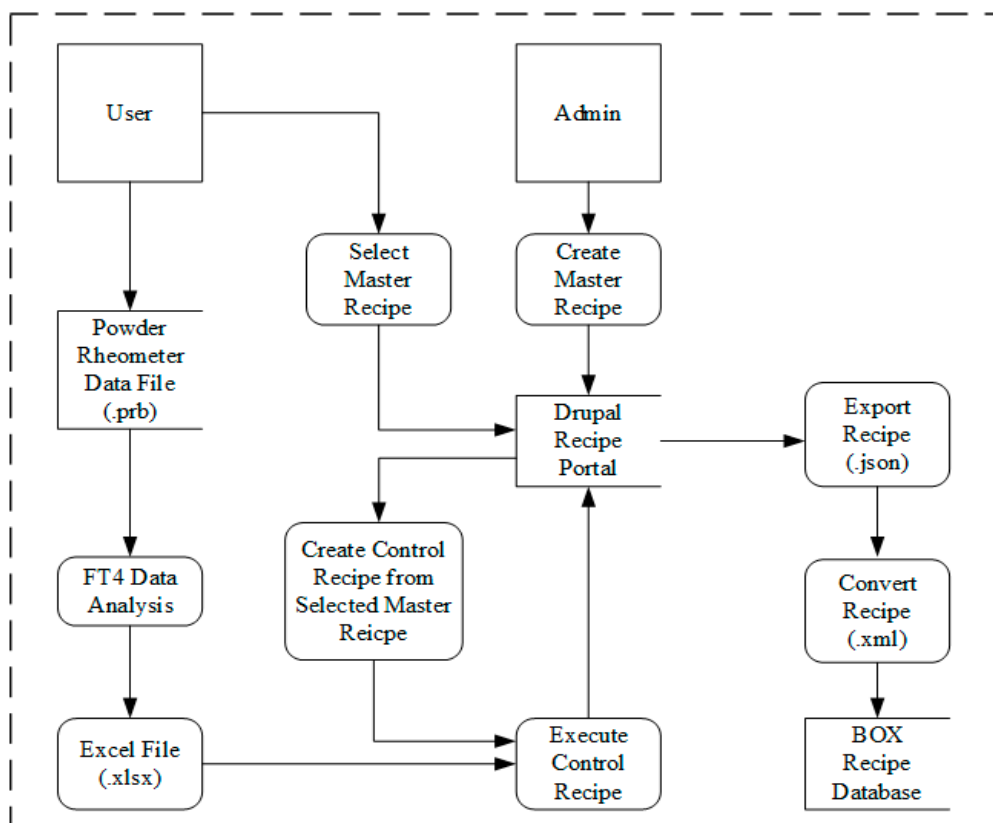

Figure 18. Data flow in performing FT4 experiment using the ELN system.

Containing the information of the analytical process, test results and instrument information, each control recipe is an integral documentation of experiment. The 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. JSON is 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 8 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.

5. Conclusions and Future Directions

A systematic framework for data management for continuous pharmaceutical manufacturing has been developed. The 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 a 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. 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 levels and distributed locations can be integrated and contextualized with meaningful information. Such knowledge plays an important role in supporting process control and decision making [31]. Future work includes validating the possibility of using this data management strategy to support the design of experiment (DOE). In terms of the laboratory platform, instrument configuration and calibration information will be included into experiment recipes, which are helpful for error analysis. The importance of this developed data management and integration is important since it will enable industrial practitioners to better monitor and control the process, identify risk, and mitigate process failures. This will enable the product to be produced with reduced time-to-market and increased quality, leading to potentially better and cheaper medicines.

Author Contributions: A.F., S.J., M.I., D.H and R.R. contributed theory and methods. B.H., C.K. and R.S. provided equipment and materials. H.C. and S.M. designed and performed the case studies. H.C., S.M. and R.S. wrote the paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Lee, S.L.; O'Connor, T.F.; Yang, X.; Cruz, C.N.; Chatterjee, S.; Madurawe, R.D.; Moore, C.M.V.; Yu, L.X.; Woodcock, J. Modernizing pharmaceutical manufacturing: From batch to continuous production. *J. Pharm. Innov.* **2015**, *10*, 191–199. [CrossRef]
2. Food and Drug Administration (FDA). *Guidance for Industry Pat—A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance*; Food and Drug Administration: Rockville, MD, USA, 2004.
3. Venkatasubramanian, V.; Zhao, C.; Joglekar, G.; Jain, A.; Hailemariam, L.; Suresh, P.; Akkisetty, P.; Morris, K.; Reklaitis, G.V. Ontological informatics infrastructure for pharmaceutical product development and manufacturing. *Comput. Chem. Eng.* **2006**, *30*, 1482–1496. [CrossRef]
4. Instrument Society of America (ISA). *Batch Control—Part 1: Models and Terminology*; ISA: Research Triangle Park, NC, USA, 1995.
5. Instrument Society of America (ISA). *Batch Control—Part 2: Data Structures and Guidelines for Languages*; SA: Research Triangle Park, NC, USA, 2001.
6. Instrument Society of America (ISA). *Batch Control—Part 3: General and Site Recipe Models and Representation*; SA: Research Triangle Park, NC, USA, 2003.
7. Dorresteijn, R.C.; Wieten, G.; Santen, P.T.E.V.; Philippi, M.C.; Gooijer, C.D.d.; Tramper, J.; Beuvery, E.C. Current good manufacturing practice in plant automation of biological production processes. *Cytotechnology* **1997**, *23*, 19–28. [CrossRef] [PubMed]
8. Verwater-Lukszo, Z. A practical approach to recipe improvement and optimization in the batch processing industry. *Comput. Ind.* **1998**, *36*, 279–300. [CrossRef]
9. Kimball, R.; Ross, M. *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*; Wiley Computer Publishing: Indianapolis, IN, USA, 2013.
10. Fermier, A.; McKenzie, P.; Murphy, T.; Poulsen, L.; Schaefer, G. Bringing new products to market faster. *Pharm. Eng.* **2012**, *46*, 1–8.
11. Ierapetritou, M.; Muzzio, F.; Reklaitis, G. Perspectives on the continuous manufacturing of powder-based pharmaceutical processes. *AIChE J.* **2016**, *62*, 1846–1862. [CrossRef]
12. Vasilenko, A.; Glasser, B.J.; Muzzio, F.J. Shear and flow behavior of pharmaceutical blends—Method comparison study. *Powder Technol.* **2011**, *208*, 628–636. [CrossRef]

13. Ramachandran, R.; Ansari, M.A.; Chaudhury, A.; Kapadia, A.; Prakash, A.V.; Stepanek, F. A quantitative assessment of the influence of primary particle size polydispersity on granule inhomogeneity. *Chem. Eng. Sci.* **2012**, *71*, 104–110. [CrossRef]

14. Llusa, M.; Levin, M.; Snee, R.D.; Muzzio, F.J. Measuring the hydrophobicity of lubricated blends of pharmaceutical excipients. *Powder Technol.* **2010**, *198*, 101–107. [CrossRef]

15. Armbrust, M.; Stoica, I.; Zaharia, M.; Fox, A.; Griffith, R.; Joseph, A.D.; Katz, R.; Konwinski, A.; Lee, G.; Patterson, D.; et al. A view of cloud computing. *Commun. ACM* **2010**, *53*, 50–58. [CrossRef]

16. Subramanian, B. The disruptive influence of cloud computing and its implications for adoption in the pharmaceutical and life sciences industry. *J. Med. Mark. Device Diagn. Pharm. Mark.* **2012**, *12*, 192–203. [CrossRef]

17. Gruber, T.R. A translation approach to portable ontology specifications. *Knowl. Acquis.* **1993**, *5*, 199–220. [CrossRef]

18. Swartout, W.R.; Neches, R.; Patil, R. Knowledge sharing—Prospects and challenges. In Proceedings of the International Conference on Building and Sharing of Very Large-Scale Knowledge Bases 93, Tokyo, Japan, 1–4 December 1993.

19. W3C. Extensible Markup Language (XML) 1.0 (Fifth Edition). Available online: https://www.w3.org/TR/2008/REC-xml-20081126/ (accessed on 1 March 2018).

20. W3C. W3c XML Schema Definition Language (XSD) 1.1 Part 1: Structures. Available online: https://www.w3.org/TR/xmlschema11-1/ (accessed on 1 March 2018).

21. De Minicis, M.; Giordano, F.; Poli, F.; Schiraldi, M.M. Recipe development process re-design with ansi/isa-88 batch control standard in the pharmaceutical industry. *Int. J. Eng. Bus. Manag.* **2014**, *6*. [CrossRef]

22. Muñoz, E.; Capón-García, E.; Espuña, A.; Puigjaner, L. Ontological framework for enterprise-wide integrated decision-making at operational level. *Comput. Chem. Eng.* **2012**, *42*, 217–234. [CrossRef]

23. International Electrotechnical Commission (IEC). IEC 62264-1:2013. In *Enterprise-Control System Integration—Part 1: Models and Terminology*; IEC: Geneva, Switzerland, 2013.

24. Hwang, Y.; Grant, D. An empirical study of enterprise resource planning integration: Global and local perspectives. *Inf. Dev.* **2014**, *32*, 260–270. [CrossRef]

25. Alphonsus, E.R.; Abdullah, M.O. A review on the applications of programmable logic controllers (PLCS). *Renew. Sustain. Energy Rev.* **2016**, *60*, 1185–1205. [CrossRef]

26. Baumann, W. Davfs2. Available online: http://savannah.nongnu.org/projects/davfs2 (accessed on 1 March 2018).

27. Whitehead, J. Webdav: Versatile collaboration multiprotocol. *IEEE Internet Comput.* **2005**, *9*, 66–74. [CrossRef]

28. Tridgell, A. Efficient Algorithms for Sorting and Synchronization. Ph.D. Thesis, Australian National University, Canberra, Australia, 1999.

29. 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.* **2012**, *438*, 307–326. [CrossRef] [PubMed]

30. Godena, G.; Lukman, T.; Steiner, I.; Bergant, F.; Strmčnik, S. A new object model of batch equipment and procedural control for better recipe reuse. *Comput. Ind.* **2015**, *70*, 46–55. [CrossRef]

31. 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.* **2016**, *505*, 394–408. [CrossRef] [PubMed]