Presented at the
ISPE – ISA Automation Forum - 2010

Global Serialization Markup Language

A Standardized Non-Proprietary Data Format and Architecture

Author Name	Charles Jabara
Title	Director
Company	Acadeus
Address	10975 Benson Dr
City/State/Country/Postal Code	Overland Park, KS 66213
Telephone	(913) 220-2882
Fax Number	(913) 663-2729
E-mail	charles.jabara@acadeus.com

KEY WORDS

Serialization Drug Global Markup Language Data Format

ABSTRACT

Abstract: Based on the FDA Standards for Securing the Drug Supply Chain, States such as California are pursuing efforts to develop additional technology standards for securing the drug supply chain against counterfeit, diverted, sub-potent, substandard, adulterated, misbranded, or expired drugs.

Development of these technology standards is based on the quality or stage in the supply chain process throughout the supply chain process. This is also described as the drug product pedigree. The pedigree of the drug spans many different parts in the supply chain from raw materials production and procurement, through drug product manufacturing, packaging, distribution, pharmacies, and finally the end user. In this light, the drug pedigree would be an electronic record containing critical quality data, including information pertaining to each transaction, change of ownership from producer to distributor, and finally to the pharmacy or person(s) administering or dispensing the drug.

The information that would need to be generated and retained with the drug supply chain process would come from many sources. For instance, the raw materials manufacturer provides quality and Lot information. The drug manufacturer incorporates the raw materials

manufacturer's data and add to it production quality data, as well as custody transfer data to the packager. The packager incorporates the tracking and Lot data while maintaining the integrity of the supplier quality and drug confirmation data. The packager also places the drug product in a transient state from shipper to distribution, then from distribution to pharmacies or authorized agencies to dispense or administer the drug to the end user. In all of these steps there is one common theme: data is collected in an electronic format. With all of this electronic data, there are multiple language protocols based on data information collection system.

If this data is to be collected and retained with the drug throughout the supply chain process, a common electronic information language protocol needs to be adapted. The data needs to be shared between all disparate systems owned by suppliers, manufacturers, distributors, and drug administrating agencies. The development of a Global Drug Serialization Markup Language (GDSML) schema would need to be developed and adopted by all groups participating in the drug supply chain. The language could be XML based. XML is a text base language designed for structuring data. XML was developed for information exchange. XML resembles HTML but XML features influence the development of other languages and lead to better standardization of formats. XML is a group of technologies with a long history and is a license free platform that is well supported.

GDSML will provide a bi-directional connector interface for systems such as ERP, packaging, labeling, and vision systems. Similar to the early paradigm of EDI (Electronic Data Interchange), GDSML will provide a more robust data schema. These language transfer protocols would need to be written though the supply chain process for transportation to raw materials quality data to administering of the drug product to the end user.

GDSML would provide the data collection and transport throughout the supply chain process. The very nature of XML is based on the openness of the protocol to work with other data collection protocols. The amount of data that XML can support, along with data encryption, only supplies the appropriate serialization data to the next source or confirmation point in the process.

INTRODUCTION

The United States is the largest producer and consumer of pharmaceuticals in the world. Concerns of increased pharmaceutical counterfeiting activities are resulting in the Food and Drug Administration (FDA) stepping up enforcement of supply chain control and issuing the Standards for Securing the Drug Supply Chain. Drugs that are counterfeited, diverted, sub-potent, substandard, adulterated, misbranded or expired pose a serious health risk to the public. In addition to the FDA, states such as California are pursuing efforts to require drugs to be serialized as well as providing the pedigree or genealogy information for that drug.

A brief explanation of the concepts of pedigree and serialization is appropriate. A pedigree means an electronic record containing information regarding each transaction resulting in a change of ownership of a given prescription drug, from sale by a manufacturer, through acquisition and sale by one or more wholesalers, manufacturers, or pharmacies, until final sale to a pharmacy or other person furnishing, administering, or dispensing the prescription drug.

Serialization is the process done at the point where the drug is produced and packaged to provide a unique number or code by which the "unit" can be tracked in the pedigree process. Serialization typically is done to the unit level which may be a bottle or even, in some cases, a single dose.

For those who were hoping that e-pedigree and serialization were concepts which would have to be done after the current generation retires, sorry, it is coming now.

The proposed solutions to implement drug serialization will require suppliers to incorporate better use of available technology. Utilization of identification tools such as data matrix labels and Radio Frequency IDentificaion (RFID) provide superior tracking of products. The use of these technologies will generate a data source, which will incorporate various formats or language types. The purpose of this paper is to suggest and outline a conversion methodology and commonality of data language to be adopted for that critical serialization and genealogy data. Global Drug Serialization Markup Language (GDSML) is one such possible solution. Implementation of a consistent data standard is a key component to the success of global tracking of drug products.

DEVELOPMENT OF THE PROCESS

Development of these technology standards is based on tracking the quality data throughout the supply chain process. This is also described as the drug product pedigree. The pedigree of the drug spans many different parts in the supply chain from raw materials production and procurement, through drug product manufacturing, packaging, distribution, pharmacies, and finally the filling of a prescription for an individual patient. In this light, the drug pedigree would be an electronic record containing critical quality data, including information pertaining to each transaction, change of ownership from producer to distributor, and finally to the pharmacy or person(s) administering or dispensing the drug.

The information that needs to be generated and retained with the drug supply chain process must come from many sources. For instance, the raw materials manufacturer provides quality and lot information. The drug manufacturer incorporates the raw materials manufacturer's data, combines their production quality data, and transfers this data to the packager. The packager incorporates the tracking and lot data while maintaining the integrity of the supplier quality and drug confirmation data. The packager also places the drug product in a transient state from shipper to distribution, then from distribution to pharmacies or authorized agencies to dispense or administer the drug to the end user. In all of these steps there is one common theme: data is collected in an electronic format. With each of these entities potentially using a different method and different software for managing the data collection, integration of the information into a single traceable record becomes extremely complex and is unlikely to work.

THE USE OF XML

If this data is to be collected and retained with the drug throughout the supply chain process, a common electronic information language protocol needs to be adopted. The data needs to be shared among the disparate systems owned by suppliers, manufacturers, distributors, and drug administrating agencies. The development of a Global Drug Serialization Markup Language (GDSML) schema adopted by all groups participating in the drug supply chain is a

possible solution to this problem. A language based on Extensible Markup Language (XML) is recommended because of its versatility and the following advantages:

- Is a text based language designed for structuring data.
- Developed for information exchange
- Uses human readable language as opposed to computers which use binary and ASCII code.
- Specifically developed to enable information exchange among different systems.
- Resembles HTML but XML features influence the development of other languages and lead to better standardization of formats.
- Is a group of technologies with a long history and is a license free platform that is well supported.

There are many advantages to using XML for information exchange.

- Applicability to a wide range of software platforms (databases, e-commerce, Java, web development, searching).
- It is extendable, meaning you can create your own tags or use tags created by others that use the basic language operating on most domains with attributes that are logical.
- Can operate on any network and can be used as an instrument to share data and application models.

GDSML provides a bi-directional connector interface for systems such as Enterprise Resource Planning (ERP), packaging, labeling, and vision systems. Similar to the early paradigm of EDI (Electronic Data Interchange), GDSML provides a more robust data schema. These language transfer protocols need to be written to accommodate each data entry point in the supply chain process from raw materials quality data to the administration of the drug product to the patient.

GDSML will provide the data collection and transport throughout the supply chain process. The very nature of XML is based on the openness of the protocol to work with other data collection protocols. The amount of data that XML can support, along with data encryption, only supplies the appropriate serialization data to the next source or confirmation point in the process.

Adoption of GDSML industry wide will certainly have its challenges. Interoperability, a buzzword used since the early days of computer science, effectively means the ability for disparate systems to exchange and use data regardless of their differences -- a concept easier said than implemented. Technical Infrastructure, Security and Equipment are all challenges to a successful implementation of a common standard.

TECHNICAL INFRASTRUCTURE

The adoption of a global technology standard must first be well received from the major players involved in the supply chain business. Manufacturing, materials management, and logistics are often captured in the functionality of modern day ERP systems. While most ERP systems provide flexible connectors and components to interface with their systems, the challenges may lie in fundamental differences that may exist in the data schema. Customers who own these systems will most probably demand off the shelf functionality to avoid the

roadblocks most commonly encountered in custom software development; expertise, management oversight, time, and budget.

A major infrastructure consideration centers on data storage requirements. Although most will agree that storage today is inexpensive, cost is not the only issue. Record retention periods and data stored across multiple supply chain partners must also be considered. The question of where and how the data should be stored is critical.

Perhaps the solution to the storage dilemma is the migration to cloud computing. “It's become the phrase du jour," says Gartner senior analyst Ben Pring. But what does this mean for a standard such as GDSML? In broad terms cloud computing is an outsourced service, typically utility in nature, that is contained somewhere outside of your company firewall. Several roles could exist for a cloud computing provider. A cloud provider could host all of the data captured by GDSML documents and serve as official record holder, or they could simply perform data validation and migration services to existing systems. Either function eliminates or reduces new infrastructure requirements.

Cloud computing at the highest level moves the entire serialization process and application to an outsourced provider. Software as a Service (SaaS) providers aim to eliminate upfront software licenses, hardware investments, and simplify system rollouts. The majority of SaaS providers will even offer preconfigured solutions to meet the majority of your needs upon signup. SaaS based ERP from companies such as NetSuite indicates this model can be a viable option for GDSML as well.

SECURITY

Why should one worry about security? Remember the controversy and interpretation discussions regarding open versus closed system in 21 CFR Part 11? That frames the exactly the concerns that have to be dealt with about security. No matter which implementation route is chosen, there will be multiple logistics partners involved with providing data into the system which means tight controls and encryption technologies must be considered.

Fortunately, this problem is not unique to the pharmaceutical industry. W3C.org published XML Encryption standards over two years ago that specify how an XML document and its children can apply an open encryption standard while retaining its XML data form. Although this standard already exists, there are still opportunities to mitigate broader XML security issues such as XML authentication, authorization and XML Signature.

EQUIPMENT

How will the data be collected that will become the drug pedigree? Automated data collection systems such as Machine Vision, Bar Code/Data Matrix, and RFID are all existing technologies that can be used for rapid and accurate data gathering.

RFID systems can collect large amounts of data for tracking and indentifying properties about an object. The data stream from multiple readers form complex data files. The processing of the complex data files, namely in the filtering of the data to collect the events that are pertinent

to the historical stream, can be problematic. One remedy for the sorting of the data is the use of a filtering database. Based on criteria for the pedigree the common or open XML protocol can locate data for the pedigree from the filtering database.

The data collection and management structure is not isolated to RFID. Machine Vision Systems incorporate the use of an internal relational database structure for the collection and management of the data. The data can be manipulated in the same manner as any relational database structure to create custom outputs.

The use of a secondary database for data collection and management makes GDSML a very viable option. GDSML as previously stated can operate on any network and can interface with many applications.

INDUSTRY TRAINING

Currently, the challenge to serialization, in general, stems from the fact that there are no clear cut governing guidelines. The FDA has compiled a nonbinding recommendation for securing the drug supply chain process. The guidance proposes standardized numerical identifiers (SNI) for prescription drug packages. California has approached this challenge by proposing a mandate for serialization of every drug product and requiring the provision of an electronic pedigree. These proposed approaches leave much to interpretation and offer very little guidance in the way of training.

What groups would need training and what group would provide that training? Who will train the trainer?

Different groups within an organization would be responsible for different aspects of training depending upon the subject matter expertise residing in that group. Each group will need to discover their own knowledge and learn to understand where the training opportunities are. Because of the uncertainty of what is to be done, initial training needs to play a lesser role to providing content, setting standards, and establishing consistent language across various organizational silos.

The use of GDSML could provide that common and consistent language. The GDSML approach of standardizing the data would be one step that can easily be communicated via vendors, manufactures, and regulators. The training then comes down to understanding the flow of data from each point in the production, distribution and end use. This is the process not the detail. GDSML takes into account the detail and provides a common and easily communicated language.

CONCLUSION

The challenges for the GDSML solution are varied.

- The technical infrastructure and where the data will be stored.
- The issue of security and integrity; how secure will the pedigree data and the encryption methods from collection to retention.

- The opportunity presented by the many types of equipment that will used to collect the data for the pedigree and how it will be collected and managed.
- Training will need to address the biggest challenge of what serialization is and how it will be implemented.

The main advantage of GDSML is that it was developed using XML. XML platform is:

- Applicable to a wide range of applications (databases, e-commerce, Java, web development, searching)
- Extendable meaning you can create your own tags or use tags created by others that use the basic language operating on most domains with attributes that are logical.
- Accessible on any network and can be used as an instrument to share data and application models.

These attributes incorporated into GDSML provide

- Bi-directional connector interface for systems such as ERP and MES Systems.
- A more robust data schema.
- Data collection and transport throughout the supply chain process.

Adoption of GDSML industry wide will certainly have its challenges. Interoperability, a buzzword used since the early days of computer science, effectively means the ability for disparate systems to exchange and use data regardless of their differences -- a concept easier said than implemented.

These challenges are more easily overcome when the various systems begin to communicate via a common language. With GDSML the data language will be easily interoperable due to the openness of XML. GDSML will be robust due the data structure and compatible with many network system applications making it a clear choice for standardizing the data format and architecture.

REFERENCES:

Standards for Securing the Drug Supply Chain - Standardized Numerical Identification for Prescription Drug Packages, March 2010

Background and Summary of the California ePedigree Law, California State Board of Pharmacy, March 2010

XML Encryption WG, http://www.w3.org/Encryption/2001/

XML Signature WG, http://www.w3.org/Signature/

What cloud computing really means, Eric Knorr, Galen Gruman, http://www.infoworld.com/d/cloud-computing/what-cloud-computing-really-means-031

Serialization and Pedigree Initiatives: Why Now?
http://www.scribd.com/doc/16930469/Pharmaceutical-Serialization-and-Pedigree-Planning0309

Estimated Rise In Serialized Drugs In The U.S. Supply Chain. August 3rd, 2010 | Author: Dirk Rodgers
http://www.rxtrace.com/2010/08/estimated-rise-in-serialized-drugs-in-the-u-s-supply-chain.html

