DATA COMPLIANCE IN PHARMACEUTICAL INDUSTRY
Interoperability to Align Business and Information Systems

Néjib Moalla[1,2], Abdelaziz Bouras[1], Gilles Neubert[1], Yacine Ouzrout[1]
[1]*CERRAL/PRISMa, IUT Lumière Lyon 2, 160 Boulevard de l'Université, 69500, Bron, France*

Nicolas Tricca[2]
[2]*Sanofi Pasteur / Campus Mérieux. 1541, Avenue Marcel Merieux 69280, Marcy l'Etoile, France*

Keywords: Pharmaceutical Sector, Information Systems, ERP, Marketing Authorizations, Compliance, Interoperability.

Abstract: The ultimate goal in the pharmaceutical sector is product quality. However this quality can be altered by the use of a number of heterogeneous information systems with different business structures and concepts along the lifecycle of the product. Interoperability is then needed to guarantee a certain correspondence and compliance between different product data. In this paper we focus on a particular compliance problem, between production technical data, represented in an ERP, and the corresponding regulatory directives and specifications, represented by the Marketing Authorizations (MA). The MA details the process for manufacturing the medicine according to the requirements imposed by health organisations such as Food and Drug Administration (FDA) and Committee for Medicinal Products for Human use (CHMP).
The proposed approach uses an interoperability framework which is based on a multi-layer separation between the organisational aspects, business trades, and information technologies for each involved entity into the communication between the used systems.

1 INTRODUCTION

The pharmaceutical industry is distinguished among process industries by the need to comply with regulatory constraints imposed by organizations like Food and Drug Administration (FDA) (FDA, 2004), Committee for Medicinal Products for Human uses (CHMP), the guidelines of International the Conference of Harmonisation (ICH) (ICH6, 2003). Further constraints are imposed by the conventions signed with national and international authorities, called Marketing Authorisation (MA) – Authorization to Make to Market (AMM) in Europe – for the manufacture of drugs.

In this operating context, the issue of product quality is one of high priority for a company in order to maintain its credibility compared to its customers.

One of the key factors of quality is the good management of product data. Product data comes in several types and formats specific to various business trades and are supported by several heterogeneous information systems. The challenge is to enable communication among these systems and the process of guaranteeing the validity and the conformity of exchanged information. This challenge is seldom addressed systematically. Indeed, considering the complexity of information systems architectures for the production, there is a general tendency to check conformance only between the MA files and the Standard Working Instructions (SWI).

Our Scope in this paper covers the problem of communicating product data between information systems supporting the MA and the ERP for structuring production data. Delivering a product according to its description in the MA requires the right information in the ERP. Otherwise, we risk manufacturing a non compliant product, to not deliver our product in time to respect customer commitments, and in final destroy these products and lose money.

The pivotal problem of medical data is the absence of machine readable structures (Schweigera, 2005). Most healthcare data is narrative text and

Moalla N., Bouras A., Neubert G., Ouzrout Y. and Tricca N. (2006).
DATA COMPLIANCE IN PHARMACEUTICAL INDUSTRY - Interoperability to Align Business and Information Systems.
In *Proceedings of the Eighth International Conference on Enterprise Information Systems - DISI*, pages 79-86
DOI: 10.5220/0002460300790086

often not accessible. Generally, related works (Schweigera, 2005) have a certain tendency to treat this problem in structuring drug and other information using XML standards. This is generally made using topic Maps (Schweigera, 2003), but presenting a product XML data models and connecting them is not sufficient (EBXML, 2001). Same Standard for the Exchange of Product Model data (like STEP-ISO 10303) addresses this through formats and programming interfaces derived directly from domain-related information models written in the EXPRESS information modelling language. However, these formats and programming interfaces are predetermined (Sang Bong, 2002), and not always well suited to current information processing technologies. We can find also Product Data Markup Language (PDML) (William, 2001) as an Extensible Markup Language (XML) vocabulary designed to support the interchange of product information among commercial systems (such as PDM systems) or government systems (such as JEDMICS), where the vocabularies are related via mapping specifications.

Performing data mapping between regulatory and industrial product definition present a hard task that requires regrouping efforts from different sectors like regulatory affairs, industrial operations, information systems, etc.

Some pharmaceutical industries are specialized in biologic development of medicines. The implication of a deviation in manufacturing or subcontracting can run the gamut from very minor to catastrophic. Our challenge consists in delivering the right product data value through manufacturing states in the production information systems.

During manufacturing process, the product passes from one state to another. Each state may concern one or several components and we have to validate their corresponding specifications based on data coming from MA information system. The following Figure (Figure 1) presents a hierarchical structure for a product in the ERP (Enterprise Resource Planning) system of the company.

When we have to ensure compliance for one data from MA to ERP, it is necessary to find and validate product data value for each component through different product states.

Our main contribution in this paper is to use a modelling approach to handle the communication between information systems within a pharmaceutical context. We also propose a methodology for structuring and exchanging product data while ensuring their conformance. In the following section we present some modelling approaches and adapt them to our problem. In section 3, we propose a data exchange structure that ensures compliance between the information systems. Finally, by using our approach, we present a case study at Sanofi-Pasteur, a developer and producer of vaccines for human use.

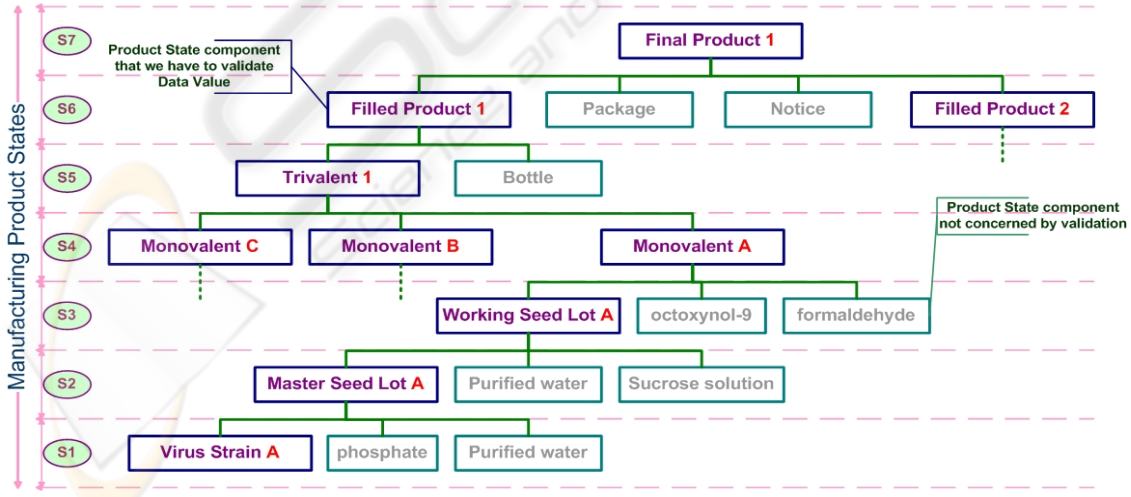

Figure 1: Manufacturing product states and state components.

2 INTEROPERABILITY IN PHARMACEUTICAL INDUSTRY

2.1 General Requirements for Interoperability

The IEEE standard computer dictionary defines interoperability as "the ability of two or more systems or components to exchange information and to use information that has been exchanged". The EU Software Copyright Directive (ATHENA, 2005) gives a similar definition and considers the interoperability between computing components as "the ability to exchange information and mutually to use the information which has been exchanged". This does not mean that each component must perform in the same way, or contain the same functionality as the other components– interoperability is not a synonym for cloning. Rather, interoperability means that components with different functionalities can share information and use it according to their needs.

The European Interoperability Framework definition identifies three separate aspects:

- Organisational – is concerned with defining business goals, modelling business processes and bringing about the collaboration of administrations that wish to exchange information, but that may have a different internal organisation and structure for their operations.
- Semantic – is concerned with ensuring that the precise meaning of exchanged information is understandable by any other application not initially developed for this purpose. Semantic interoperability enables systems to combine received information with other information resources and to process it in a meaningful manner.
- Technical – covers the technical issues of linking up computer systems and services. This includes key aspects such as open interfaces, interconnection services, data integration and middleware, data presentation and exchange, accessibility and security services

Identification and structuring of these interoperability types help to perform a better exchange between systems Therefore, it is necessary to identify the area of our investigation and its specifications: structures, business constraints, etc.

To achieve interoperability among divisions systems in collaborative enterprise, we consider three challenges (ATHENA, 2005):

- *Heterogeneity*, incoherent information, different systems and software infrastructures, different working practices, etc.
- *Flexibility*, information reuse, following of variations in documents versions, etc.
- *Complexity*, definition granularities, dependency between different components, etc.

Heterogeneity, flexibility and complexity must be managed at different levels:

- *Knowledge*, approaches, methods and skills needed for innovation, shared languages.
- *Process*, planning's, coordination and management of cooperative and interdependent activities.
- *Infrastructure*, information formats, software tools, interoperability technologies.

In an industrial framework, structuring business knowledge in an information processing system does not imply facilitation of communication with another business system. Data interpretation changes according to the business and the challenge is in the ability to preserve information semantics when communicating.

Building interoperability architecture for communication can align Business, Knowledge and ICT through semantic framework to ensure compliance when exchanging data. In the following section, we will explain a deployment of the interoperability framework to present a communication architecture adapted to our context.

2.2 Characteristics of Pharmaceutical Industry

Product data is compiled from various functional divisions which interact between each others for the creation and manufacture of the product. Each of the following divisions contributes by introducing different types of data and information:

- *Research division*: looks for new drugs or substances that can contribute to the creation of new drugs. At this stage conducted studies are reported and indexed in the form of technical reports.
- *Research & Development division*: conducts specific research, and is interested in the development of mixture processes of excipients, tests and stability conditions of the final solution that can be defined as a drug. The information system is used to structure data about clinical trials and tests for validity. At this stage starts the definition of an explicit product structure.
- *Industrialization division*: defines the industrial infrastructure which will support the

production of a defined product quantity on the basis of a definition of product solution. At this stage, we define technical data describing the product manufacturing operations and the used process and tools.

- *Production division*: deals with planning, scheduling and follow-up of production based on the data describing industrial infrastructure and product composition. At this stage, we identify static data compared to external dynamic data like work orders or those generated by the ERP such as buying orders of raw material.
- *Distribution division*: defines the conditions for handling the product for customer delivery in accordance with the description of the conditions of manufacture, which is given by R&D division. At this stage, product handling information is documented.

From one stage to another, product data are recorded using a specific structure and format. Each division information system is defined in accordance with the needs which are relevant to the business trades.

The definition of a product for pharmaceutical industry is not tied to physical shape except in the packaging stage.

The company submits to the Health Authorities entire product specifications along with documented information. These deposed documents constitute the request of Marketing Authorization. When health authorities approve this request, they give the Marketing Authorization. In the delivered documents to authorities, it is necessary to present all the information which justifies the product creation process, including pre-clinical tests, clinical trials, tests of validity and the appendices such as bibliography. Only after reaching the industrialisation stage that MA documents can get defined.

Once approved in one country, this MA is used as a reference document to manufacture the product. It is considered as a contract between the authority of a given country and the company, implying the respect of the regulatory constraints. For the American market for example, the Food and Drug Administration (FDA) is responsible for the checking of the adequacy of the delivered product and manufacturing processes to the acquired authorization.

The major quest for each pharmaceutical company is product quality. This objective is achieved only by ensuring a better degree of compliance between existing information in these MA documents and those used for the production.

We propose hereunder the means to use the MA data, which can be read only by pharmacists, to adapt them to logisticians needs. The used approach makes it possible to ensure interoperability between the supporting information systems, while satisfying some business constraints.

3 INTEROPERABILITY AND COMPLIANCE

In our context, the objective behind the establishment of the communication between the information systems is to ensure the conformity of the product data in one system in relation to each other. Based on the description of information in an Marketing Authorization, it is necessary to return the product data values, useful for the production, to the ERP.

3.1 From MA to ERP

As we mentioned before, the following systems are involved in our context:

- Marketing Authorization (MA) information system: generally managed by the regulatory affairs division of the company and constitutes a collection of different information. A MA is composed of electronic documents coming from several sources and contains, for example, scanned documents, reports and attached papers. The semantic structuring of these authorisations documents provides a format and content which are harmonized according to a pharmaceutical vision. It specifies the Common Technical Document (CTD), defined by the International Conference of Harmonisation (CTD, 2005) (ICH, 2000). In the CTD it is not always easy to find all the information needed for production, and some pharmaceutical background is necessary to find the needed information from regulatory data. Even with a very large number of MA documents – that's run into thousands of pages – it is very difficult to find all information needed for production. MA presents regulatory aspect of product data.
- ERP (Enterprise Resource Planning) system: related to different divisions of company and regroups complex functionalities of "provisioning and scheduling" and generates new dynamic data, such as working orders, based on the product definition. When the ERP data are non-conform to the right product data definition, it necessarily produces a non conforming product.

Each division presents a specific vision of the product with local knowledge tied to its business needs. To ensure the conformity at the product data definition level during its translation (from the regulatory systems to the ERP), it is necessary to define a communication platform to include the different viewpoints: organisational, business, informational, and technical (Gao, 2003).

3.2 Type of Data to be Translated

The product structure is defined in both MA and ERP systems as a specific series of "product states". The pharmaceutical description of the product and its various states related to the manufacturing phase are presented in the CTD "product quality" documents of the MA. These states are not necessarily coherent with the actual production states. To guaranty the product data coherence, it is suited to organize these data according to the product states. However, the problem still concerns the conformance of data values for each product states during the translation process. We should take care about the definition of these states and data semantic in each one. For example the shelf life of an intermediate product substance (state) is 3 years, at a storage temperature of -70°C if it was preserved with no alteration (as is) and 1 year if it was stored at 5 °C.

In the manufacturing phase, we assume that the product has a fixed number of states (reflected into the information system). It is necessary to identify from the ERP and the regulatory information system the entire specification of each state. This is achieved by what we call "product states reference frame". The reference frame represents the structuring of one product datum that assigns for each product state, the data value, rules applied to extract data from the information system, and business constraints helping to understand the choice of data value. For each product state, we need to define also some components of the bill of materials of this state. For example, when our final product is presented (at its final state) in the form of two substances (i.e. powder and liquid), we need to specify shelf life for these two substances.

The application of this reference frame to product data consists in seeking data values of all states in accordance to rules and business constraints already identified. Figure 2 illustrates this structuring.

This reference frame represents the data profile in both information systems. It must be updated during a potential modification of the structure of the product. It can also be published in the organization to ensure better comprehension and exploitation of the product data.

Each line of this reference frame contains the product state components and for each one of them, the value to be validated, the rules which allow to extract and transform data and business constraints. The interoperability process is supported by the link between these values, rules and constraints.

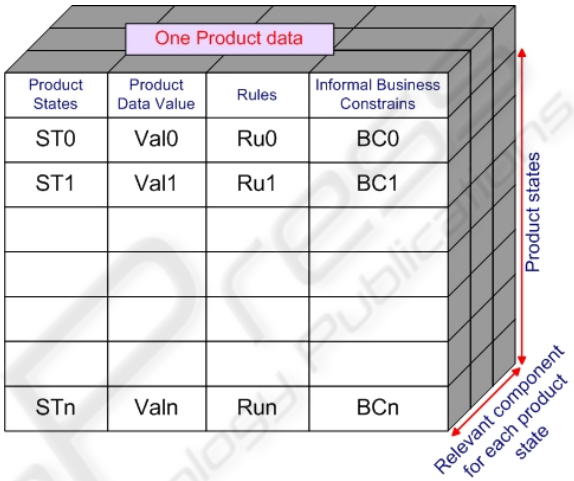

Figure 2: product states reference frame.

3.3 Rules Definition

The definition of the rules is a tedious phase and requires three levels:

3.3.1 Production Information Rules

These are rules specify when to extract or to insert data into the ERP. Some difficulties arise when attempting to insert data because ERPs are characterized by the re-use of product states information. Taking a close look into two drugs pharmaceutical solution, there is a great probability to find the same excipients. In this case, there are invariably one or more specific common production states with the same coding in the system.

In the ERP, and following a request for modification of a data value of a product state, it is necessary to check if the reference for this state is already used by another product. Considering the complexity of the ERP architecture and overlapping between the product states information, it is difficult to seek products by a simple indication of an "intermediate" state. For example, such identification can take up to two days to find all

concerned product states and theirs dependencies. If we schematize product states by a tree structure, the overlapping between branches can be possible everywhere except at the top level (tree leaves). Figure 3 shows an example of these overlapping. Each product has 6 states: S1 to S6.

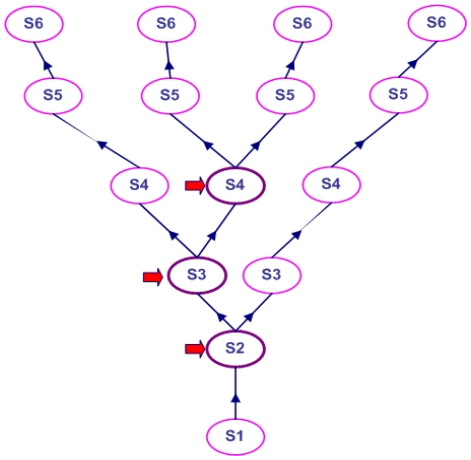

Figure 3: overlapping between product states for different products in the ERP system.

Integration rules are used to control the existence of any overlapping between the tree branches (resulting in common states) as well as the impact of any data modification on the product structure and its states. The impact of some modifications or transformations at the data level is sometimes governed by *informal business constraints*. For example, the manufacturing date of a product is notified as the starting date of the first valid test of stability. If we want to change the shelf life of a state, the expiry date must be revalidated. This aspect is important for data understanding. This is why we added informal business constraints to each product state. Moreover, these constraints help understanding the context in which the product states rules are used, and in mapping the ERP "product states reference frame" to its corresponding reference frame in the MA information system.

3.3.2 Mapping Rules

These are rules for mapping between "product states reference frames" by establishing links between "active product states". From all predefined product states in one reference frame, active product states present significant states with data value. Performing these links present a regulatory and pharmaceutical responsibility that is necessary to share with production, to ensure the coherence of rules. The product states are not the same across information systems and across reference frames. From one product to another, a state may or may not exist. We use different business knowledge as references to create these links of communication between active states. We notify these information on both MA and production reference frames (ERP).

The mapping rules allow the formalization of the fields of the data to be inter-connected (links n **..** n**).** Active states data values in regulatory reference frame generate corresponding values in the product states reference frame of the ERP. Figure 4 illustrates examples of connection modes. One state in each reference frame can correspond to one or more states in the second and vice versa. To generate mapping rules, we should analyze data and rules from the two reference frames. For example, mapping rules could be the adding of states data values, the calculation of their average or their minimum, etc.

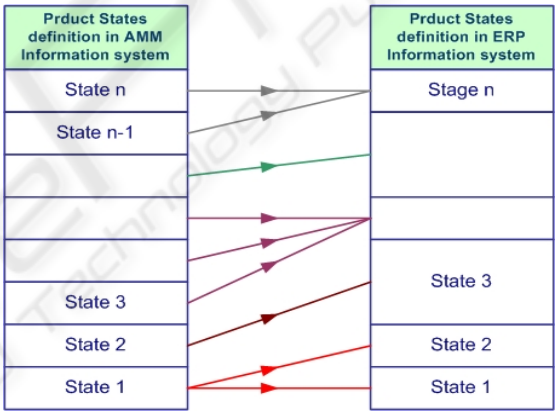

Figure 4: Mapping links.

3.3.3 Regulatory Information Systems Rules

According to pharmaceutical data structuring, the information system which manages the MA is not able to be directly interfaced to the regulatory product states reference frame. It is possible to have several MAs for only one product, and conversely, one MA for several products. These characteristics are relocated on product states, which increase the complexity of the information retrieval. It is very frequent to find for example two product authorizations with various destinations (country) or presentations (packaging containers) and having a common product state but with different data values. This difference is due to the history of the negotiations between the company and health authority about the MA content.

In the following, we will explain the need for defining different rules types and later (in a future work), we will present, through a multi level modelling approach, different kinds of rules we need to create.

4 CASE STUDY

This case study presents an illustration of an application developed within Sanofi Pasteur Company, a firm specialised on biologic development and the production of vaccines for human use. The purpose of this application is to ensure compliance, from the MA to the ERP, for three data: *Site of Manufacturing, Shelf Life*, and *Storage Condition.*

All MA data were structured in e-TRAC (Electronic Tracking of Registrations and Commitments) MA information systems. Access to these data is ensured through web interface allowing us to export the defined report from *RA-Cockpit* reporting module. As presented in figure 5, we can:

a) export data for one product line to create the report ,

b) distribute this report by product licence number as criteria to identify different product data,

c) for one product data, instantiate three reference frames for regulatory product states,

d) apply mapping rules to generate corresponding ERP (here SAP) product states reference frame,

e) use the same specific criteria for data structuring in SAP to validate data (comparing to those coming from SAP reference frame generated after mapping).

4.1 Validate Data in SAP

As mentioned before, there is a great probability to have the same product state in different product states decompositions. So, we can find the same value for the same product state in different SAP reference frames. In SAP system, we identify each entity, called item, by one code. That is why, in addition to the first SAP reference frame generated after mapping, we instantiated a second SAP reference frame with only SAP code and corresponding data value field. In this second reference frame it is necessary to find, from SAP, the code and value of each product state. Due to specific information structuring in SAP at Sanofi-Pasteur Company, we can find the item code for the last state (final product) and use "item code filiations" (Where-Used technology) to find the code for previous product states and their data values starting from the last.

Actually we have two SAP reference frames: one with data values generated after mapping from the regulatory reference frame, and the second with data values and item code coming from SAP. We define here some new rules of coherence:

R1: For the same product state, there is necessarily the same data value, otherwise notify a compliance exception,

R2: The same item codes in the second SAP reference frames (corresponding to different products) should have the same associated data values, otherwise notify a compliance exception,

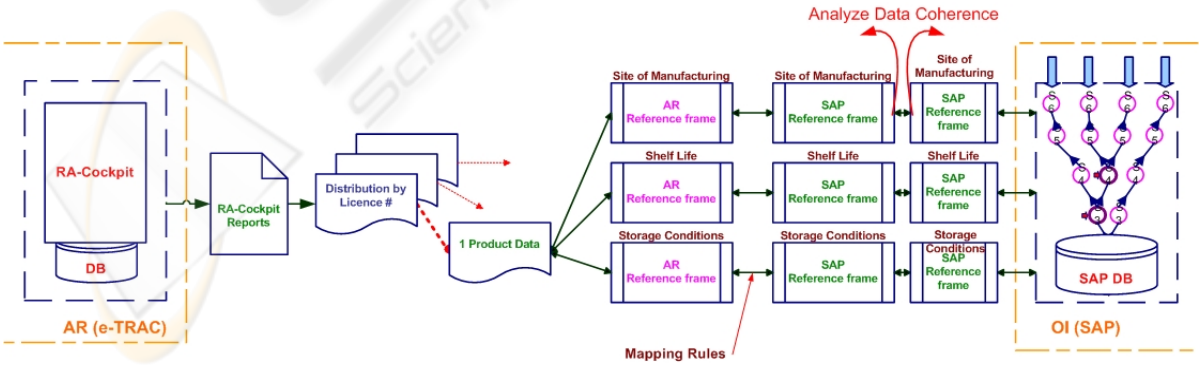

Figure 5: Communication scenario.

It is frequent to find two or more MAs or registrations that differ just by product name from one country to another. For example we can define influenza (Flu) vaccines for entire Europe, but during the structuring of the product information in the e-TRAC, we should separate the products by country.

R3: Validating the three data (Site of Manufacturing, Shelf Life, *Storage Condition) for grippe in a particular region, requires the same data values in e-TRAC reference frame for all countries of this region, otherwise notify a compliance exception.*

Finally, within this Flu line product vaccines case study, the applied architecture and its rules provided an interesting solution by ensuring compliance of 94,6% of the final products for the used three data: *Site of Manufacturing, Shelf Life*, and *Storage Condition*. One of the reasons of non-total compliance is related to the existence of quality level information in the MA system that has no correspondence in the ERP system.

5 CONCLUSION

In this paper, we presented a methodology to communicate between information systems. We particularly focused on product structuring and explained dependencies between product data in the pharmaceutical field. Our main objective is to ensure data compliance between two information systems, one related to the Marketing Authorizations and the other related to production, through the establishment of communication architecture. We based our work on the mapping between product "states" information along the product manufacturing life cycle. In spite of differences in their business visions, both systems use the product manufacturing decomposition as guide-line for structuring the information.

Our methodology treats only the information coming from Marketing Authorizations systems to map and validate it in the ERP systems. However it does not treat product information that exists in the ERP systems and is not related to any MA system.

The next step of this work will focus on the generalization of the used rules and constraints, not only to extract or integrate data through reference frames, but also between product states in a same reference frame.

REFERENCES

ATHENA, 2005, "Second Version of State of the Art in Enterprise Modelling Techniques and Technologies to Support Enterprise Interoperability", *Version 1.3.1, 212 pages.*

ATHENA, 2005, "Interoperability Roadmap up date", *Version 1.0, 30 pages.*

CTD Format: http://www.aboutctd.com/resource.htm and http://www.ich.org/cache/html/1208-272-1.html, accessed on October 2005.

EBXML, 2001, "Technical Architecture Specification V1.0.4", *ebXML Technical Architecture Project Team*, 39 pages

Gao X., Hayder A., Maropoulos P. G. and Cheung W. M., 2003, "Application of product data management technologies for enterprise integration", *International Journal of Computer Integrated Manufacturing*, Taylor & Francis, Vol 16 N° 7-8, pages 491 – 500.

ICH6 - Sixth International Conference on Harmonisation, 2003, "New Horizons and Future Challenges", *Summary Report*, Osaka Japan, pages 13-15

International Conference on Harmonisation, Harmonised Tripartite Guideline, 2000 "Organisation Of The Common Technical Document For The Registration Of Pharmaceuticals For Human Use", *International Conference On Harmonisation Of Technical Requirements For Registration Of Pharmaceuticals For Human Use*, 14 pages.

Sang Bong Y. and Yeongho K., 2002, "Web-based knowledge management for sharing product data in virtual enterprises", *International Journal of Production Economics*, Volume 75, Issues 1-2, Pages 173-183

Schweigera R., Brumhardb M., Hoelzerc S., Dudecka J., 2003," Linking clinical data using XML topic maps", *Artificial Intelligence in Medicine*, Volume 28, Issue 1, Pages 105-115

Schweigera R., Brumhard M., Hoelzerc S., Dudecka J., 2004, "Implementing health care systems using XML Standards", *International Journal of Medical Informatics (2005)* 74, pages 267—277

U.S Food and Drug Administration, 2004, "Pharmaceutical CGMPS For The 21st Century — A Risk-Based Approach Final Report", *Department of Health and Human Services*, 32 pages

William C. Burkett, 2001, "Product data markup language: a new paradigm for product data exchange and integration", *Computer-Aided Design, Volume 33, Issue 7,* Pages 489-500