Real World Data Identification and Classification for Support of Virtual Confidence

Iman Morshedzadeh[1], Jan oscarsson[1], Amos H.C. Ng[1], Manfred A. Jeusfeld[2] Anders Jenefeldt[3]
[1]University of Skövde, School of Engineering Science, Skövde, Sweden
[2] University of Skövde, School of Informatics, Skövde, Sweden
[3] Volvo Cars Corporation, Skövde, Sweden

Sweden

Iman.morshedzadeh@his.se

Abstract

Shortening of the product development process time is one of the main approaches for all enterprises to offer their products to the market. Virtual manufacturing tools can help companies to reduce their time to market, by reduction of the engineering lead time. Extensive use of virtual engineering models results in a need for verification of the model's accuracy. This virtual engineering usability and assessment have been named virtual confidence. The two main factors of the achievement of this confidence are the accuracy of the virtual models and the virtual engineering results.

For controlling of both above factors, a complete virtual model and related virtual model knowledge are needed. These knowledges can be tacit or explicit. For exploring explicit knowledge, a data and information collection from different disciplines in the organization is needed.

In this paper, a data map with focus on the manufacturing engineering scope will be presented. This data map is generated from different data sources at a manufacturing plant, and gives an overview of different data that exist at different data sources, in the area of manufacturing. Combining real world data from different sources with virtual engineering model data supports, amongst others, establishment of virtual confidence.

Keywords: Virtual confidence, manufacturing knowledge, data map, Product lifecycle management (PLM), explicit knowledge.

1. Introduction

Manufacturing enterprises use different virtual tools to create virtual models. A model is a representation of an object or occurrence. A virtual model can reflect items such as a product design, manufacturing resources and systems. These virtual models can be used to predict the results of a plan in advance, prior to the physical implementation. Virtual models can be used to identify failures or mistakes made in the early phases of the product lifecycle. With the virtual model usage, the need for testing in the real world has been decreased and which in turn reduces development lead-time and cost. Even if the virtual modelling is very beneficial for industry, it is not an easy job and it is time consuming. Creating a model, analysing it and generate results from a model consumes engineering working time, depending on the model type and many other criteria. Because of that, re-use of virtual models becomes an important factor in order to save time and money. Re-use of a model requires that the user is certain about the models reliability. Oscarsson et al. defined a concept for this model's reliability, named "Virtual confidence" [1].

They explained that each virtual model has a level of virtual confidence which defines to what extent results from modelling and simulation can be used for making business decisions. The level of confidence depends on the model's provenance data, and the correctness and accuracy of its results and predictions. This correctness and accuracy can be evaluated by comparing of results obtained from reality, after the physical implementation and those model's results and predictions. At their method for identifying the level of a virtual model's confidence, two types of knowledge should be clarified about that model. Knowledge about the creation of that virtual model and knowledge about how well the model worked in the real world (after implementation) are two types of knowledge.

For creating a virtual model, both tacit and explicit knowledge should be used. Tacit knowledge is a type of knowledge which is not codified, such as the knowledge which comes from human experience [2]. Explicit knowledge is knowledge which has been codified, and can be transferred to another person [3],[4]. Explicit knowledge has its origin in different sources, such as different documents or databases.

Lots of documents and data are involved during the creation of a virtual model. Whenever designers want to design models, in addition to their experiences, they are using different stated data and information. This information can explain limitations and other criterion's which affects the model design. If a model shall be reused, it is important to understand under which circumstances the previous model had been designed. Actually, related data and information can be considered as sources for designing a model, and they also support definition of the level of confidence for that model.

This highlights a necessity to identify and map different sources of knowledge, information and data, when building virtual models. The need for this is a motivation for this research. The paper focus on manufacturing data and virtual models related to manufacturing system design and development, and presents the result from a survey made on documents and information used during manufacturing system design. This survey does not cover all data and documents during the product life cycle, and it is limited to a part of the product development process. The result of the survey is a data map which specifies the relation of the data and documents as a support to find explicit knowledge which is one of the sources of knowledge for virtual models.

A similar research has been done in the manufacturing area, by Mott and Jack about classifying of manufacturing knowledge[5]. They define a manufacturing knowledge model and divided the knowledge area to four pillars. They explained some knowledge foundations for these four areas which consist of different types, including documents. These knowledge model had been not represented according to the PLC phases[5].

This data collection is the first step toward reuse of virtual models. The main need for having this data map is to identify data and documents which are affecting a virtual model and also identify data and documents which are influenced by that virtual model.

Before exploring and collecting data, data sources were identified, and a list of features and attributes were defined. Whenever data was added to the data map, the feature list of that data also was filled in, to complete the provenance information about that data.

Since the data should be managed by a product lifecycle management (PLM) system, the data map was structured according to product lifecycle (PLC) phases, based on a manufacturing preparation perspective.

In the following, the product lifecycle phases will be explained briefly. In this research product lifecycle phases have been divided in to sub-phases for accurate location of the data in the PLC. Thereafter a data collection methodology will be explained and data and information sources will be classified. At the end of the article, the resulting data map will be presented. The survey was carried out in a factory in the automotive industry.

1.1. Product lifecycle management

There are lots of definitions for product lifecycle management. Saaksvuori and Immonen define PLM as "a systematic, controlled concept for managing and developing products and product related information"[6].

Grieves defines PLM as an information-driven approach to all phases of product life cycle. He mentions how PLM drive the generation of lean thinking by exchanging product information for wasted time, energy, and material across the entire organization[7].

In this paper PLM is defined as a business approach for management and use of corporate intellectual capital, such as product, process and production system related data and information including people's experiences and knowledge over the extended enterprise.

Table 1. PLC phases and sub-phases

	Product Life Cycle					
Phases	**Business idea**	**Requirement management**	**Development**	**Production**	**Operation and maintenance**	**Disposal**
Sub Phases	Perception	Requirement specifications	Conceptual development	Tooling	Operation	Delivery
	Assessment of market and technology	Requirement Analysis	System-level design	Machining	Maintenance	Disassembly
	Key requirement		Detailed design	Assembly		Recovery of recyclable materials
	Feasibility Study		Detailed requirement	Robotics		Disposal of no recyclable material
			Testing and refinement	Plant		
			Simulation	Quality control		
			Process planning	Production		
			Production ramp up	Supply		
			Quality			

PLM systems are intended to manage data over a product lifecycle. Crnkovic divides the product life cycle to six different phases [8]. He divides the PLC to; business idea, requirement management, development, production, operation, maintenance, and disposal phases. Grieves defines five phases for the PLC; plan, design, build, support and dispose. [6]. Stark divides the product lifecycle from the manufacturer of the product and from the user of product point of view. From the manufacturer perspective the PLC has been divided to; imagine, define realise, support (service) and retire phases. But the product lifecycle was divided to; imagine, define, realise, use (operate) and dispose (recycle), from a user or costumer point of view [9].

In this research, the PLC has been divided into six main phases, and each phase has been divided in sub-phases. This definition was made based on the type of activities which takes place in each phase, from a manufacturing point of view (*Table 1*).

In the first phase, the product will appear as a concept. The idea of the product will be developed and the market of the product will be evaluated. The key requirements for manufacture will be specified. In the requirement management phase, detailed requirements for manufacture of the product will be clarified. The third phase is the development phase. Most of data and information are generated in this phase. Product designs, process designs, simulations, and many other activities are undertaken in this phase.

The product life is entered into the production phase, where the product is produced at the production line. The tooling, machining, assembly, and many other activities are undertaken in this phase. The next phase is operation and maintenance, and in this phase, the product are used by costumers. Re-cycling or disposal is the last phase, and it is the last part of a product life cycle.

2. Data collection methodology

This research was made by collaboration of the university and the product life cycle management team from a manufacturing company, together with a consultant for PLM implementation in that company (*Fig. 1*). In this methodology, for data collection, some meeting and semi-structured interviews with each discipline in the factory were held. In those interviews, data, information and documents that are generated or used in that discipline was identified. Features of data were also clarified. Every week, a workshop with attending of different stakeholders was held, for discussing about collected data, and classifying them in a right position in the data map.

After several meetings with different disciplines and many workshops, the data map of the factory was prepared and presented.

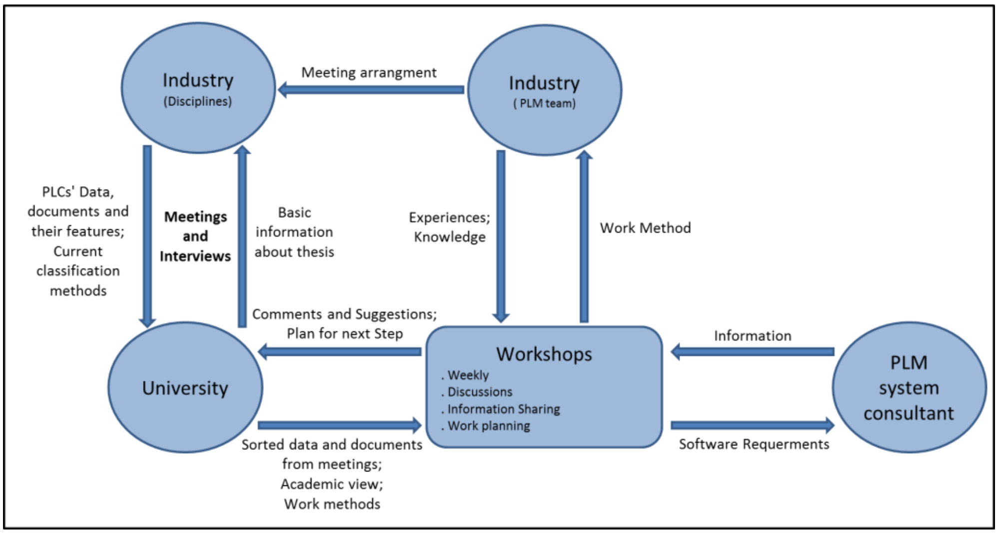

Fig. 1: *Data collection methodology*

3. Data and information sources

Lot of data is generated each day in a company, and added to the enterprise knowledge. This data is stored in different sway using different media, such as papers, computer files and pictures.

In this research, data and information sources were divided into three main groups (*Fig. 2*). This classification was made according to the theme and the storage type of these data and information.

The first group is documents, and most of data and information are placed in this group. For example, control plans, standards, drawings and instructions belong to this group.

The second group is data and information in engineering and specialized software programs. Nowadays, different software programs are used, and they generate many types of data and which are saved in different data storages. Different virtual models are one type of this data and information.

There are also different computer files which are informal, i.e. not part of any formal data structures. Lots of world files, excel files and photos belong to this group.

When the data sources had been defined, and before starting data collection, the data features needed to be specified.

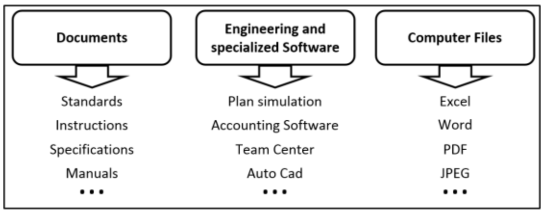

Fig. 2: *Data and information sources*

4. Data and information features

For creating a new model, different types of document and data have been used as background data. These data and documents also have been generated according to other data and documents. This relationship can be considered as the information flow, and different documents and data in this flow affect each other. For reuse of a virtual model, the information flow which the model have been generated according to need to be recorded. For both the virtual model and the related documents or data, some features should be clarified. These features can specify the provenance information of that data or model.

Ram and Liu describe a model for provenance data [10]. The model had been named 7W because it is based on seven questions about the data, which are; "What", "When", "Where", "How", "Who", "Which" and "Why". Recording the answers to these questions preserves the provenance of a particular document or piece of data. If an engineer have access to this information about a model and its related data and documents, then the conditions and basis for the model will be clear for him/her. Hence is it possible to decide if the model can be reused, or not in its current design, alternatively if it needs some modification before reuse

Because of this, in this project have data features been defined in order to cover these seven questions. When one type of data was identified, at the same time the features of that data were specified.

During this project were fifteen types of features defined (*Fig. 3*). The first feature is "Title", and it is the same as the name of the data. At the "Description" the matter of the data or document has been described. The "Type" feature is clarifying the type of a data such as: document, form, computer file, etc. For the computer files data, the "format" feature is applicable, and it is identifying the format of the computer file, like Microsoft Excel or Autodesk drawing.

Sometimes data are in different languages, and it should be clarified in the "Language" feature. "Producer", "Sender", "Receiver" and "User" are features which specifies the roles or disciplines that are related to the data. If the data has some rights, confidentiality, or authorization to use or access, then it should be stated in the "Rights" feature. The "Size" feature gives an overview of the data size. If it is softcopy then it can be a size of the file, and if it is hardcopy then, it is the number of pages.

	Data Features	
Title	Producer	Size
Description	Sender	Version
Type	Receiver	Periodic/Some time
Format	User	Source
Language	Rights (access authorization)	Archive place

Fig. 3: *Data and document features*

The "Version" feature of a data can clarify the data's version, if it has some revisions. The "Periodic/Sometime" feature explains that data is generated one time or in some periods. If other data have been used for the producing this data, then it should be stated in the "Source" feature. The last feature is "Archive place" and it is specifying the final place, for archiving that data.

All of these features are not meaningful for all types of data, because of that, these feature's information should be specified as much as possible. For example, some kinds of data have been generated by an employee without any sources.

5. Data collection

As previously mentioned, the data collection was made discipline by discipline. At each meeting a semi-structured interview was held and the persons interviewed described their most used documents, data and information. Beside this data collection, the data features were specified according to the table. *Fig. 4* shows a sample of data features that have been collected for one type of data, in this case a production flow simulation using discrete event simulation technology.

Flow simulation	
Description Flow simulation	**Sender** productivity Eng
Type(Data/form/document/...) Data	**Receiver** Manufacturing EngSupliers
Periodic/Some time one time	**User** Manufacturing EngSupliers
Mean number per month	**Rights(access authorization)** no
Softcopy/Hardcopy softcopy	**Archive place** Prductivity Eng
Size -	**Source** Bill of process (Excel file), Technical Specification, prestudy report
Format Computer file	**Version** no
Producer productivity Eng	**Language** E

Fig. 4: *Data features for Flow simulation*

It should be highlighted that disciplines are adopted from different department in the organization, but they are not precisely the organizational structure.

5.1 Advanced engineering

This discipline produces information and data that can be used in the early phases of the PLC. Three main types of data are generated in this discipline. “Wanted industrial position”, “Technical strategic” and “Product content” are documents which are generated, and mostly used by manager and designers.

Since, these documents evaluate the manufacturing conditions and technical needs, they are placed in the “Business idea” phase of the PLC (*Fig. 5*).

Business idea
Assessment of market and technology
Wanted industrial position
Technical strategic
Product content file

Fig. 5: *advanced engineering data in business idea*

Since the focus of this data collection was on the manufacturing data, these documents are the only highlighted documents that belong to the business idea phase of the PLC.

5.2 Manufacturing engineering

Manufacturing engineering is a core engineering discipline in a manufacturing factory, according to the data usage and generation. Most of the other disciplines studied are somehow, related to this discipline, or can be considered as a part of this discipline. The data and information which are collected in this discipline are classified in two phases of the PLC. Most of them are placed in the “Development” sub-phases and some of them are located in the “Requirement” phase. The “Restriction model” and the “Process requirement description” are two documents that are used before developing a product, and states limitations and requirements for developing that product. Because of that, they are placed in the “Requirement specification” sub-phase.

Other types of manufacturing engineering discipline’s data are classified in different sub-phases of “Development” phase as shown in *Fig. 6*.

Requirements management	Development				
Requirement specifications	System-level design	Detailed requirement	Process planning	Production ramp up	Quality
Restriction Model	Flow chart	Fixturing	Process sheets	Activity instruction	Control instruction
Process requirement description		Standards		Capability Study	

Fig. 6: *Manufacturing engineering data*

At the “Flow chart” document, the information about production line and process orders are clarified. The “Fixturing” data describes how a product should be fixed, during the production. “Standards” are documents that explain different internal and external standards. “Process sheets” are drawings about the production process, and placed in the “Process planning” sub-phase. The “Activity instruction” is an instruction for building up a production line, and “Capability study” is a document that has been prepared for suppliers, to clarify the needs, from them.

The “Control instruction” is the last document that is used in this discipline. It explains some quality control, at the last stages of a product development.

5.3 Equipment suppliers

In the most cases, equipment suppliers are not a discipline in the manufacturing enterprises, but they generate lots of data which are used during the PLC. Therefore they have been considered as a separate discipline in this investigation.

Equipment suppliers provides documents about the equipment’s that they supplied for a production line such as “Machine layout” and “Operation time”.

When a supplier customizes equipment for a specific process, then they use product drawings to generate their process information for using that equipment. As an example, the “Hole number specification” is a document that provides detailed information about setting equipment for drilling large number of holes, which should be done on that equipment.

These data and information are used and related to the "Process planning" sub-phase, from "Development" phase of PLC (*Fig. 7*).

Development
Process planning
Machine Layout
Operations time
Hole number specification

Fig. 7: *Equipment Suppliers data and information*

5.4 Equipment

When equipment and machinery have been installed in the factory, lots of data and information are generated for operation of the machinery and equipment's.

Fig. 8 shows different document and information about operating equipment and its functions. These documents are prepared through the close cooperation with equipment suppliers.

Development	
Detailed requirement	Production ramp up
Technical specification	Operation instruction
Functional description assembly equipment	FU operator instruction
Functional description machinery equipment	
Operational reliability electrical	
Operational reliability mechanical	
Functional test machining	
Functional test assembling	

Fig. 8: *Equipment's data and information*

All of these documents should be prepared before start of the production, hence they are placed in the development sub-phases.

5.5 Packing

For packing of the final product, two types of document have been used. The "Packing instruction" explaining how to pack a product and the "Packing drawing" is a 2D model of the packing process.

Since, these documents should be prepared and used, when the product is under development, they are classified in the development phase of product lifecycle. The *Fig. 9* shows the pacing data and information in the PLC.

Development	
Detailed requirement	Process planning
Packing drawing	Packing instruction

Fig. 9: *Pacing data and information*

5.6 Quality

In each sub-phase of the development process are quality control data and information generated. At the "Conceptual development" phase, failure mode and effect analysis documents for design and process are generated. These documents explain the failures that can occur in the product design and process design, and analysing the effects of these failures on the product. Since, the "Process failure mode and effects analysis" will be updated during the production phase, this document is placed in the development phase and production phase. *Fig. 10* shows other types of quality control data and information such as instructions, test procedures and control plans.

Development					Production
Conceptual development	Detailed design	Testing and refinement	Process planning	Quality	Quality
Design failure mode and effects analysis (DFMEA)	Quality control plan by product and process control	Measurement System Analysis	First piece control	PV Test	Quality control instruction
Process failure mode and effects analysis (PFMEA)					Process failure mode and effects analysis (PFMEA)
					KISA

Fig. 10: *Quality control's data and information*

5.7 Gauging and tooling

Gauging tools are instruments that are used for measuring tools or products. For gauging tools, two types of data are used. "Gauging drawings'' show the tool geometrical data, and "Gauging tool list" is a list of gauging tools with their specifications.

For cutting tools also, the "Cutting tool drawing" and the "Cutting tool list" are used. The "Wear parts" is a document that gives information about some specific parts in the operation. The wear parts have been consumed during the production such as some grippers.

These data and document should be prepared before production, hence are they placed in the development phase of the product lifecycle (*Fig. 11*).

Development	
Detailed requirement	Process planning
Cutting tool drawings	Cutting tool list
Gauging tool drawings	Gauging tool list
	Wear parts

Fig. 11: *Gauging and tooling data and information*

Tool's drawings should be generated at the "Detailed requirement" sub-phase and tool lists and wear parts, should be generated during the process planning.

5.8 Productivity engineering

At the "Productivity engineering" discipline, engineers try to manage and improve the production process. They work with two main groups of data and information. The first group is about simulation of production lines and the second group is data and information that are used in the process's operations.

Fig. 12 shows how these data should be classified in the PLC phases.

Development	
Simulation	Process planning
Flow simulation	Operator instruction sheet
	Work element sheet
	NC Program
	Measurement Program

Fig. 12: *Productivity engineering data and information*

The flow simulation data and operation related data should be prepared before production, and because of that, they are placed in the development phase.

5.9 Customers

After a product has been sold on the market, customers are using them and the manufacturer will receive some feedbacks from customers. In this paper, customers have been considered as a separate discipline, even though actually they are not a real discipline in the factory.

The “Customer remarks” are placed in the “Operation and maintenance” phase, but it can be used in the earlier phases of product lifecycle (*Fig. 13*).

Operation and maintenance
Operation
Customer remarks

Fig. 13: *Customer's data and information*

In this study have a limited portion of information and data used in a manufacturing company been surveyed, with focus on the manufacturing domain. In a company, a lot more data and information are generated. This wider scope is only partially covered in this report.

6. Conclusion

After the data collection, all data was placed together in the product lifecycle phases. *Fig. 14* shows the company’s manufacturing data and information. As previously mentioned, the main purpose of the data collection was to collect samples on manufacturing knowledge, for supporting the virtual confidence by preserving the data provenance.

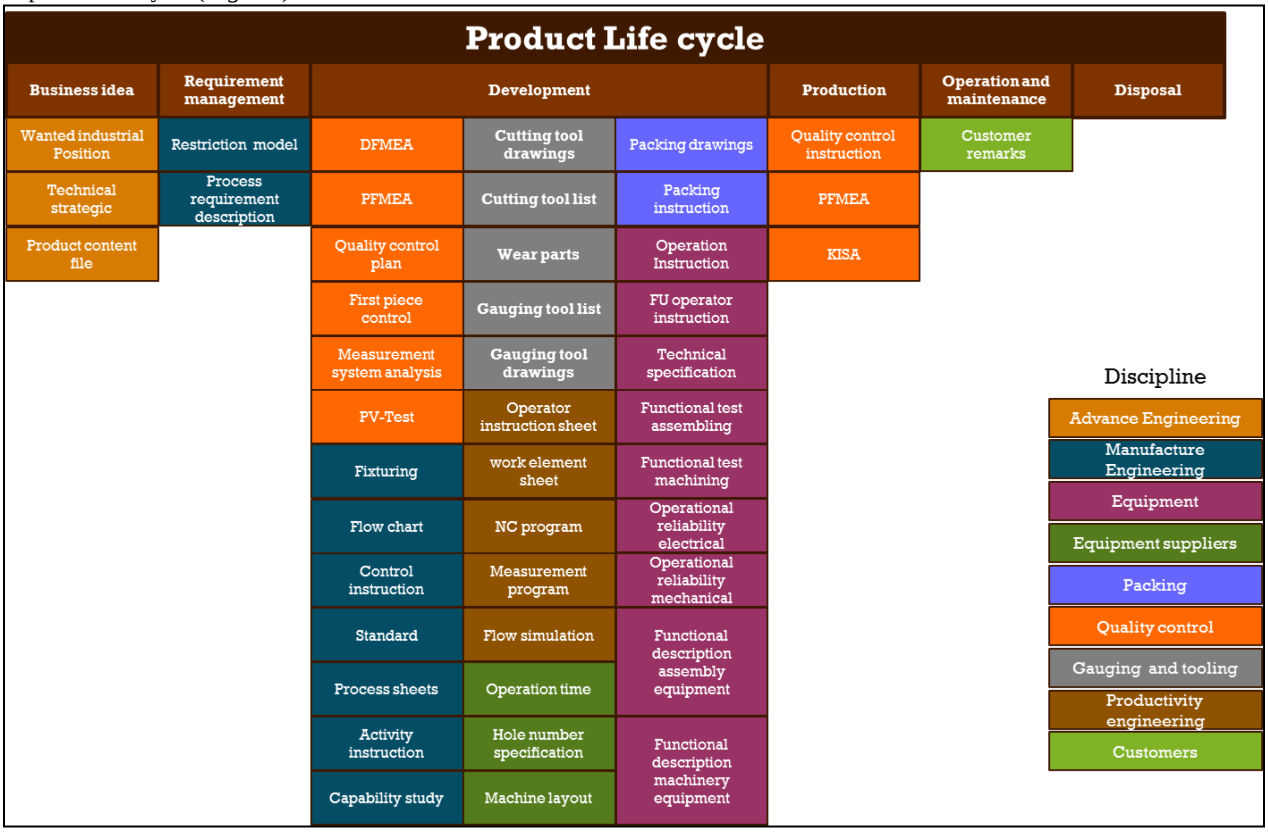

Fig. 14: *Company's Data Map with focusing on manufacturing engineering scope*

The data collection display different types of data sources in the manufacturing area. With data’s attributes specification, the provenance information for them can be explained, and this information can help to clarify different criteria which are involved in virtual modelling.

According to this data map, most of virtual models are placed in early phases of PLC, before production. Restriction model, flow simulation, and different drawings are some types of virtual models.

As an example, by specifying the “Flow simulation” model’s provenance data, it has become clear that the “technical specification” is used as one of data sources.

Therefore, technical specification is one of the items that affect the confidence level of the flow simulation model.

When a designer wants to reuse this model, according to its virtual confidence level he/she can determine how to reuse the model. These virtual model's provenance data can also specify places where the results of the virtual models have been used. For instance, the "Flow simulation model" will be used by equipment suppliers, at the "Machine layout" document.

This data map can also be used for other purposes. This data map shows that most of the manufacturing data and information are generated and used in the development phase. The map reveals the fact that few documents are generated in the early stages of the product lifecycle, which means that manufacturing engineering has an inconspicuous role at the "Business idea" and "Requirement management" phases of the PLC. This is a weakness for manufacturing companies, especially when they decide to develop a new product, and causes lots of problems when it comes to manufacturing.

This data map can help companies to obtain an overall picture, over their data and information. Manufacturing companies can use their data map to find the best way for managing their data, and if they are using a PLM system they can decide which of these data should be managed by the PLM system. It is also helpful for implementation of PLM systems in their enterprises.

As previously mentioned has, this data collection been done with a limited scope. For the future work, this data collection can be expanded to other areas in the company, such as product design or maintenance. Based on a map which presents sources of data and relations between different sets of data, defining flow of explicit knowledge within a company will be possible. This would be an important part of a knowledge network which is the basis for product and production system realisation.

References

[1] J. Oscarsson, M. A. Jeusfeld, and A. Jenefeldt, "Towards Virtual Confidence - Extended Product Lifecycle Management," in *Product Lifecycle Management in the Era of Internet of Things*, A. Bouras, B. Eynard, S. Foufou, and K.-D. Thoben, Eds. Springer International Publishing, 2015, pp. 708–717.

[2] M. Polanyi, *Personal Knowledge*. Routledge, 2012.

[3] I. Nonaka, "The Knowledge-Creating Company," *Harvard Business Review*, 01-Jul-2007. [Online]. Available: https://hbr.org/2007/07/the-knowledge-creating-company. [Accessed: 24-Mar-2016].

[4] D. A. Guerra-Zubiaga and R. I. M. Young, "A manufacturing model to enable knowledge maintenance in decision support systems," *J. Manuf. Syst.*, vol. 25, no. 2, pp. 122–136, 2006.

[5] R. Mott, H. Jack, V. Raju, and S. Mark, "The Four Pillars of Manufacturing Knowledge," *http://www.ncmeresource.org/*, 2015. [Online]. Available: http://www.ncmeresource.org/lcbp/pdf/webinar1/The%20Four%20Pillars%20of%20Manufacturing.pdf. [Accessed: 06-Sep-2016].

[6] A. Saaksvuori and A. Immonen, *Product Lifecycle Management*. Springer Science & Business Media, 2005.

[7] M. Grieves, *Product Lifecycle Management: Driving the Next Generation of Lean Thinking*, 1 edition. New York: McGraw-Hill Education, 2005.

[8] I. Crnkovic, U. Asklund, and A. P. Dahlqvist, *Implementing and Integrating Product Data Management and Software Configuration Management*. Boston: Artech Print on Demand, 2003.

[9] John Stark, *Product Lifecycle Management: 21st Century Paradigm for Product Realisation*. London: Springer-Verlag, 2005.

[10] S. Ram and J. Liu, "Understanding the Semantics of Data Provenance to Support Active Conceptual Modeling," in *Active Conceptual Modeling of Learning*, P. P. Chen and L. Y. Wong, Eds. Springer Berlin Heidelberg, 2006, pp. 17–29.