INTERNATIONAL DESIGN CONFERENCE - DESIGN 2008
Dubrovnik - Croatia, May 19 - 22, 2008.

DESIGN 2008

INITIAL STEPS OF DEVELOPING A PROCESS INTEGRATIVE STRUCTURING OF DOCUMENTS CONTAINING PRODUCT INFORMATION

W. M. Lauer and U. Lindemann

Keywords: structuring, process integration, information retrieval, document description, definition of attributes

1. Introduction

Today's companies need innovative products with more functionalities and decreasing time-to-market to survive in times of dynamic globalised markets. Thus, the products themselves and the required product development processes are very complex and difficult to manage. Therein, a major problem is to provide engineers with the right information at the right time [Weber 2002] which is aggrevated by the increasing amount of information. To fulfil the engineer's task, they need to search for existing information about a product (input) such as interface relations, weight or functionality of a component and have to evaluate its relevance for their request. Since product information is mainly stored in a variety of different documents, reasonable structuring of them is essential to control the development process [Stuffer 1993]. Researchers criticise the lack of common structure and a lack of precise semantics of documents [Del-Rey-Chamarro 2004]. Examples for structuring the documents are by function, form (geometrical) or manufacturing [Conrad 2007], but do not address the aspect of the document's integration into product development processes. In the framework of workflow planning, this research aims at the support of this process integration by defining a structure of documents which is geared to the different development process steps an engineer is working on. Companies are expected to reduce their effort on process workshops where documents are integrated into the process manually. The process integrative structuring approach of this research has the goal to support the integration of documents into different process steps and activities which leads to a better retrieval of product information contained in the linked documents. At the current stage of this research, the considered process has not been defined yet. However, by following the approved standard product development process (VDI 2221) the development of an exemplary process integrative structure is possible. The approach supports the allocation of relevant documents considering the engineer's process relevant demand and reduces time-consuming searching for relevant information.
For the development of this structuring method, an unstructured list of documents is gathered from companies and is analysed afterwards. The analysis is done by pairwise comparison of the documents. Therein, similarities and differences of product information is identified and converted into process oriented attributes later on. The attributes enable a clustering of the documents and their contained product information. To allow the support of the development process, especially process oriented attributes are to be selected which helps to increase the precision of relevance of the provided documents. The conclusion and outlook reflect first results of the structuring approach. Furthermore, an outlook is given on future work about the value scale of attributes, the mapping of these attributes to the product development process and the expansion of the data base the attributes will result from.

2. Motivation

As complexity of products and development processes increases, more and more documents are generated to make information available in the ongoing development process. In the context of this research, product information consist of helpful information about a product and its properties as a basis for further design steps of an engineer. They are stored in many different documents such as simulation reports, component descriptions or conceptual sketches. So, the engineers need to browse these documents to find the requested product infomation. Though common search engines support the retrieval of information, simple search mechanisms return a lot of unprecise results [Blocks 2004]. Experienced users improve the precision of their search results by modifying their search terms using synonyms. However, the search causes an overload of information which engineers cannot evaluate in the whole. In consequence, engineers select results which are the fastest and easiest to access (top of the list) [Fidel 2004]. The main reason for this circumstance is the lack of context and user specific search engines which could increase the relevance and precision of search results. Since the requested product information is closely related to the task and the process step the engineer deals with, a process integrative structure is expected to allow the retrieval of relevant information in advance which supports the fast access to needed information.

Especially in the framework of workflow planning, the evaluation of relevance for an engineer plays a major role. Usually, the relevance is defined in advance by analysts and results in a direct connection of documents and product information to the process steps. The required effort uses high expenses for doing process workshops. In contrary to the presented approach, these documents are linked rigidly whereas the new approach offers dynamic linkage by using attributes with varying values. The reason for the current rigid linkage are the used structures in product data management systems. These systems are used to provide documents to the process, but are mainly based on hierarchical product structures (see figure 1). This just addresses the demand of engineers to a small extent [Weber 2007], since engineers intuitively search by associations. The objective of this research is a process orientied integrative structuring which is believed to support the concept of workflow planning as well as the quality of results emerging from queries for product information. Therein, the intended approach supports the retrieval of documents and their product information in a process oriented manner.

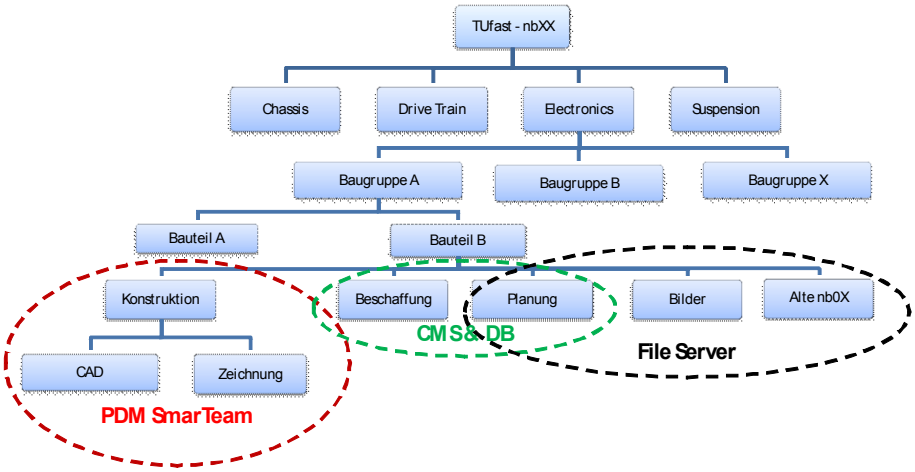

Figure 1. Exemplary hierarchical product structure of a vehicle

3. Objectives and Proceeding

As described in the previous sections, process integrative structuring is believed to improve the access to relevant information and to support the linkage of product information to the development process.

In this research, parameters have to be derived which are capable of defining a process oriented structure of documents containing these product information. This allows a dynamic linkage of documents to the development process steps which is needed for achieving more flexibility in the workflow planning and higher retrieval accuracy of product information. The parameters have to be understood easily by the using engineers, so they can modify the values of the parameters by themselves. The found parameters are developed further and defined as attributes. By using and evaluating these attributes it is possible to cluster documents and calculate their classification as well as their process affiliation. Furthermore, it allows the identification of product information which is closely related to other product information. Hence, the objective of this research is the identification of similarities and differences of documents and their content as well as the derivation of process oriented attributes which emerge from all examined documents. The result is a relation of product information and the development process by just dealing with documents.
In the initial step of the proceeeding, a list of documents is gathered (Step 1, see figure 2). The documents are analysed for similarities and differences (Step 2) by comparing them pairwise. This method is chosen, since it uses available information about the documents and does not require additional information which is not documented implicitly, but has to be added by further effort. The found similarities are transferred into respresentative attributes which are evaluated as being most adequate to describe a connection to the development process (Step 3). The exemplary application of the method to a single instance of a document demonstrates the intended evaluation process of the attributes (Step 4).

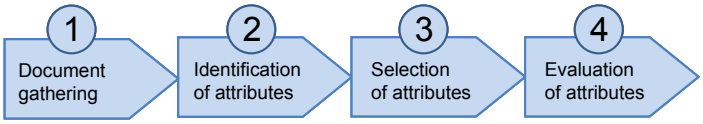

Figure 2. Proceeding of the document analysis

4. Document gathering (1)

Before gathering documents, the types of considered documents are defined below. Afterwards the level of integration is defined to allow gathering of suitable documents which can be integrated to the defined process level. Furthermore the sources of documents are explained and the resulting list of gathered documents is presented.

4.1 Types of considered documents and referred product developement process

In the context of this research, the considered documents include various product information such as form, function, maturity or quality status. This product information is stored in documents which are created in different stages of the development process. The consequence is that documents are expected to contain redundant product information. This is the main difference to partial product models defined by Anderl [Anderl 2000] which avoid these redundancies. In contrary, the process oriented structuring described in this paper aims at the support of already existing information sceneries which means it does not follow the same objective to introduce a new ideal information structuring like Anderl.
Since the number of all documents created during the development process of e.g. a surgical robot can easily exceed a few hundreds, the whole range cannot be analysed in this research alone. Furthermore there are different levels of the development process from abstract to concrete which influence the selection of proper documents. Hence, a level of the process has to be defined for the integration of the documents. For this process level decision, examples for different abstraction levels of the process are explained (see figure 3).
An existing workflow management system of the automotive industry serves as an example for the identification of possible process levels. The process modelled in this system divides into the process level and subprocess level (see figure 3). The process level includes the general approach of product development and consists of product planning, desing and construction, work preparation, production

and assembly. This process corresponds to the VDI-guidline 2221 which is approved as a standardized procedure. Though this does not reflect the individual process in the industry, it is sufficient to first trends. The actual integration will be done into a real process which still has to be defined in future work. The subprocess level details the mentioned process phases to subprocess modules which can be combined and varied to walk through the product development process. The subprocesses correspond to process steps which engineers have to carry out to fulfill their task. Examples are change processes, evaluation processes or approval processes for e.g. technical drawings.The documents required to fulfil these process steps are those considered in this research.Their integration into the process is illustrated as third level (Objects) in figure 3. For the gathering of documents the subprocess level is inspected whereby only official documents are selected. In this framework, official documents are defined as documentation of e.g. the development progress and reporting or milestone approvals. Inofficial documents like personal notes, unspecified drawings or auxiliary calculation tables are excluded from this analysis. The definition of a level of examination facilitates the intended analysis of later research, since it can be done with reasonable effort and still promises sufficient results to draw first significant conclusions.

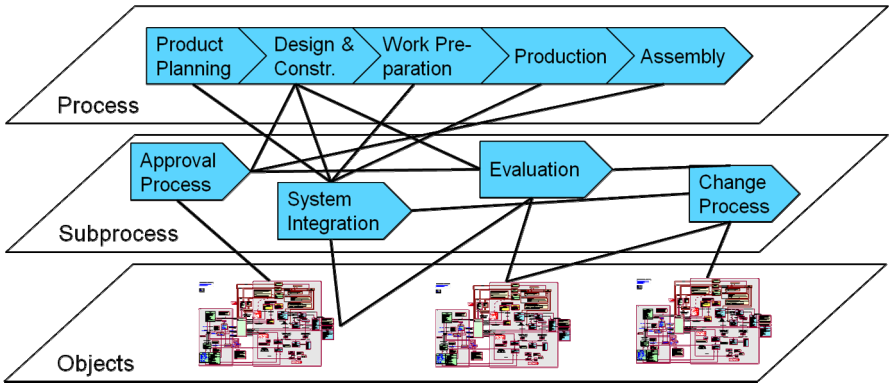

Figure 3. Level of integration and examination

4.2 Sources of documents

Concerning the level of examination, documents are gathered using different sources. Examples are literature like Anderl [AnTr00], a real product development process of student's team race car (Formula Student), and different original equipment manufacturers (OEMs) of the automotive industry. The usage and handling of documents are examined by analyzing existing information management systems and underlying structures. Especially a workflow management system of an OEM serves as a source for documents, since it represents an existing solution for the integration of documents into the development process. The student's racing team provides documents stored on a file server, in a product data management system as well as on a content management system.

4.3 Identified documents and their description

After searching for documents, they are collected in a list. Figure 4 indicates an extract of the list. The indicated table lists the documents in the left column and the corresponding decription of their content and properties in the right column. In total there are 72 selected documents. The description is given unstructured to achieve as much objectivity of the description as possible, but subjectivity cannot be avoided totally. However, the unavoidable partial subjective descriptions lead to unconscious categorization. Its effect is reduced by the unstructured way of gathering as many different criteria for the description of docuements as possible without being biased by a predefined structure. This helps to enlarge the pool of structuring criteria and enhances the probability of identifying most suitable process oriented attributes. The analysis for the identification of differences and similarities of the document's content and properties is done in section 5.

Documents	Description
Change Document	The Change Document is a standardized document for the propagation of planned or already applied changes of a product as well as for changes of a contract. Change Documents always consist of single isolated changes. They can include links to other Change Documents and are edited by standardized procedures.
CA-Document	A Computer-Aided-Document represents a geometrical model of designed construction. It is generated during the design phase, underlies various changes (iterative development process) and is used for later assemblies and product analysis. A CA-Document has the status "In Progress", "Denied" or "Approved". It refers to the CA-document, which organises all CA-Documents of the complete model of the product.
Weight	This sums up all data on weight, which is available at position variants or a toolbox position If there are several actual weights, they are prioritised as work weight. This automatic prioritisation can also be edited manually.
Material	Description of the used material with identification number and characterising properties
Part	A constructive part or workshop facility
Instance in assembly	Position-dependent place of assembly of a single part or subassembly
…	…

Figure 4. Extract of the gathered documents

5. Identification of attributes (2)

After the description of the selected documents, criteria for the classification and structuring of the documents have to be identified. Therefore, the documents are analysed by pairwise comparison of similiarities.

Each description is tailored into smaller, distinguishable description packages. These packages are equal to product information (e.g. weight of a component) and properties of the document (e.g. if it is standardized or not). The analysis is done by searching for similar entries about product information and document properties. The approach results in text fragments of each document description. The text fragments make the contained properties of a document and product information easy to analyse. Each text fragment is highlighted (see figure 5, bold italic) and is defined as an attribute if reasonable.

Documents	Description
Change Document	The Change Document is a ***standardized document*** for the ***propagation*** of ***planned*** or ***already applied changes of a product*** as well as for changes of a contract. Change Documents always consist of ***single isolated changes***. They can include ***links*** to other Change Documents and are edited by ***standardized procedures***.
CA-Document	A Computer-Aided-Document represents a ***geometrical model*** of ***designed construction***. It is generated during the ***design phase***, underlies various ***changes*** (iterative development process) and is used for later ***assemblies*** and product ***analysis***. A CA-Document has the ***status*** "In Progress", "Denied" or "Approved". It ***refers*** to the CA-document, which organises all CA-Documents of the complete model of the product.

Figure 5. Identifying similarities and differences of two exemplary documents

In an exemplary analysis for the identification of possible similarities and differences, two selected documents are compared (see figure 5, Change Document and Computer Aided Document). The synthesis of all description packages over all documents gives a list of attributes which is used to classify the documents e.g. a document for a geometrical model or a simulation result. At this point, it is considered non reasonable to reduce the number of identified attributes, since a higher number of attributes allows a higher degree of diversification and facilitates the definition of a larger number of different structuring possibilities. As already mentioned, so far the approach is only applied exemplary to two documents.

As indicated in figure 5, the Change Document (CD) is a "standardized document", "edited by standardized procedures" and includes a "single isolated change". In comparison to the Computer-Aided-Document (CA), the Change Document is passive, since it is only created if changes e.g. in the CA-Document are made (active). This means the CA-Document is the driving force for the Change Document and is the initiator for it. To describe this difference the attribute is defined as "influence".

The text fragments "planned or already applied changes" (CD) and "designed construction" (CA) represent the subjects the documents give information about. Since this gives a hint at the content of a document, this information can be used to propose a type of document which can be generated for the documentation of a similar content. Hence, the attribute "content" is introduced.

The terms “propagation” (CD) and “geometrical model” as well as “assemblies” and “analysis” (CA) are grouped by the attribute “purpose” of the document. Propagation stands for the support of the information flow, since the purpose of a Change Document is the propagation of changes to avoid inconsistencies. The CA-Document follows the purpose of visualising, virtualising (geometrical model), synthesising (assembly) and easing analysis of properties of physically not existing products.
Another difference is identified with the time of creation of the documents. While the CA-Document is related closely to the “design phase” of the development process, a Change Document is generated whenever a change is made during the whole process. So the time of creation indicates the probability of the “process relation” of a document.
A further difference is the degree of cross linking to other documents. The terms “links” (CD) and “refers” (CA) prove the existence of such a connection either to other Change Documents and those affected by the change. Furthermore the “dynamics” of a Change Document and a CA-Document differ much. Whereas a Change Document saves the information about a change and is fix from then, a CA-Document underlies many changes since it represents the ongoing development process. This shall not be mixed with versioning, because changes of a CA-Document which are made before an official approval are not considered here.
Altogether, the pairwise comparison of the documents results in the six explained attributes influence, content, purpose, process relation, cross linking and dynamics.

6. Selection of attributes (3)

Though a single pairwise comparison is not representative for attributes that will result from pairwise comparisons of all considered documents, the six identified attributes give a first hint at how to structure documents. Since it is the objective of this paper to identify attributes for the process oriented description of documents, the next step is the selection of suitable attributes. The process oriented structuring will be done by combining the attributes, since this allows the intended dynamic process relation of documents. Hence, attributes which seem to have no direct relation to the process on first sight are also examined for their eventual support of a process relation.

- **influence:** the attribute “influence” indicates the activeness or passiveness of a document. Though this information is valuable for matters of reasoning and logical relations, in the context of process orientation this attribute does not indicate when a document is relevant for a special process step.
- **content:** every process step uses documents as input and generates output. Hence, the process steps focus on different goals and each one uses a special kind of documents like e.g. CA-Documents in the process step of defining the geometry of a product. Though there are further documents relevant for the proceeding of steps, a main type of documents is likely to be used in special process steps. However, for their linkage to the process the content by itself is no adequate differentiation, since documents with different content can be important for both early or late process phases. A process oriented differentiation just based on the content of documents is evaluated as non-reasonable, but a combination with other attributes is supposed to achieve the process linkage.
- **purpose:** the explanation of the attribute “purpose” includes the application of a document in the sense of doing something like e.g. analysing or assembling. This close similarity to process steps which also have a certain purpose, is considered as a possibility for a direct linkage of documents and process steps sharing the same or at least similar purposes.
- **process relation:** this attribute reflects the probability of a special content to be important for a concerned process step (see section 5). The value of “process relation” is based on a previously defined output e.g. CA-Document of a construction step. The “process relation” itself is a rigid connection to a process step and cannot consider the design situation of an engineer. Here, this attribute is considered not to be useful, since the purpose of this research is to provide information dependent on the design situation. There, it is important to know which document should be generated next or to be reused in later/other process steps. Hence, “process relation” is evaluated as not suitable for the objectives of this paper.

- **cross linking:** the degree of "cross linking" plays a major role in the handling of the complex information flow between the documents and helps to arrange them to each other. The selection of this attribute allows to relate documents to each other and to link them to the process as a cluster of a high degree of cross linking. The higher the degree of cross linking, the more complex and concrete the product is. Hence, this delivers information about the progress of the process.
- **dynamics:** the frequency of changes indicates the degree of iterations and in consequence helps to connect it to process steps with similar frequency of appearence. However, the degree of iteration is hardly to be measured in development processes, because there are too many influences affecting the process. Hence, "dynamics" is not followed any further.

Previously, the attributes are discussed considering their qualification for a process oriented structuring. Since the attributes can not satisfy the goal of process orientation by themselves, a combination of attributes is proposed to help with this task. For the identification of most suitable combinations of attributes, as many combinations as possible have to be examined. The evaluation of isolated combinations will be done by applying the attributes to documents. Especially the attributes "content", "purpose" and "cross linking" are considered useful to describe documents in the context of process orientation. Their combination is applied exemplary in the following for a first rudimental evaluation. Further research will deal with the verification and identification of more attributes by examining further combinations. Furthermore, the scale of values has to be defined to describe the documents properly (see "Conclusion and Outlook").

7. Evaluation of selected attributes

Since the identified attributes are based on a single comparison and the value scale has not been defined yet, the evaluation can not be considered to be complete. However, for a first indication of the combination's effect on process integration, the descriptions of the compared documents presented in section 5 can be used to link them to the process exemplary. A special instance of a CA-Document serves as an example.

Geometrical model of a seat (CA-Document)

The content of this CA-Document mainly includes the assembly and integration of geometrical models of subcomponents. By defining the attribute "content" as assembly of geometrical models, they have to be linked to the model of the seat to provide the necessary parts. For an assembly, the geometrical models have to be at a certain development stage. This required development stage gives indirect information about the level of concretion of the models and therefore about the process relation. Hence, the attribute "content" facilitates the derivation of process relations. In addition to the content, the "purpose" of the geometrical model of a seat is the integration of parts and the analysis of e.g. kinematics. Process steps are potentially related to those which follow a similar purpose as the model of a seat e.g. system integration or property analysis. The degree of "cross linking" to other parts in an assembly is relatively high, since the involved parts are connected by many constraints whose number is an indicator for the value of the degree of "cross linking". Because documents of a similar degree of "cross linking" may be linked to the same process step, in this case it gives a hint on the level of complexity and indirectly on the level of concretion.

Considering the content of this CA-Document as geometrical model, its purpose is the integration of parts and the degree of cross linking. Therefore, this CA-Document is connected to the process step "System Integration". Though this result may not surprise, this example shows the possibility to describe documents by process integrative attributes and to use them for dynamic process integration. In consequence, this supports document retrieval dependent on the current process step a designer is executing.

8. Conclusion and outlook

The research described in this paper presents the proceeding of identifying and selecting attributes suitable for a process integrative structuring of documents which contain product information. This structuring is essential for purposeful integration of documents into the development process. It allows the dynamic allocation of relevant documents dependent on the current process step of the engineer.

After the gathering of documents and a brief description of the considered development process (section 4), two exemplary documents are compared to identify similarities and differences of their content and properties between them (section 5). This pairwise comparison delivers the six structuring attributes “influence”, “content”, “purpose”, “process relation”, “cross linking” and “dynamics”. Subsequently, these attributes are discussed to select the most appropriate attributes for process integrative structuring (section 6). The chosen attributes “content”, “purpose” and “cross linking” are applied rudimentary to an instance of a CA-Document (geometrical model of a seat). The result is the proposal to link this document to the process step “System Integration” (section 7).
Since a single pairwise comparison does not deliver representative features of similarity and differences for structuring documents, the examinations described in this paper are not sufficient for a finalisation of process integrative structuring. Therfore, a first set of 10 pairs of documents is scheduled to be compared in future research. However, it is essential for upcoming research to define an approach for the identification and selection of attributes which is presented by an exemplary parwise comparison. The results of this paper are the mentioned approach and a first set of attributes which results from the application of this approach. For the adequate evaluation of the attributes, the properties of the development process still have to be mapped to the attributes, so their value for process integration can be evaluated under given conditions. Futhermore, the value scale of the attributes has to be specified by extending the analysis of further existing documents.

Acknowledgements

The research described in this paper is publicly funded by the Bayrische Forschungsstiftung (BFS) in the framework of the joint project FORFLOW and supported by the Institute of Product Development of Prof. U. Lindemann (Technische Universität München).

References

Anderl, R., Trippner, D., “Standard for the Exchange of Product Model Data.”, Teubner, Stuttgart 2000.

Blocks, D., “A Qualitative Study of Thesaurus Integration for End-user Searching”, Dissertation, University of Glamorgan, UK, 2004.

Conrad, J., Deubel, T., Köhler, C., Wanke, S., Weber, C., "Comparison of Knowledge Representation in PDM and by Semantic Networks“, In Proceedings of the 16th International Conference on Engineering Design - ICED 2007, Ecole Central Paris, Paris, 2007.

Del-Rey-Chamarro, F., Wallace, K., “A Study of Information Retrieval in Aerospace Domain.”, In Proceedings of the 14th International Conference on Engineering Design - ICED 2003, Stockholm, 2003.

Fidel, R., Green, M., "The many faces of accessibility: engineer’s perception of information sources”, Journal of Information Processing and Management, Vol. 40, Issue No. 3, 2004, p. 563-581.

Lauer, W., Lindemann, U., "Purposeful Integration of Product Models into the Product Development Process“, In Proceedings of the 16th International Conference on Engineering Design - ICED 2007, Ecole Central Paris, Paris, 2007.

Stuffer, R., "Planung und Steuerung der integrierten Produktentwicklung”, Dissertation, Technische Universität München, Germany, Munich, 1993.

Weber, C., Werner, H., Deubel, T. "A Different View on PDM and its Future Potentials”, In Proceedings of International Design Conference - DESIGN 2002, Vol. 1, Dubrovnik, May 2002, p. 101-112.

Weber, H., Lenhart, M., Birkhofer, H., "Iterative Semi-Automatic Modularization of Documents and Topic Map Creation for Development Knowledge”, In Proceedings of the 16th International Conference on Engineering Design 2007, Ecole Central Paris, Paris, 2007.

Dipl.-Ing. Wolfgang Lauer
Technical University Munich, Institute of Product Development
Boltzmannstraße 15, 85748 Garching, Germany
Emial: wolfgang.lauer@pe.mw.tum.de