INTERNATIONAL DESIGN CONFERENCE - DESIGN 2002
Dubrovnik, May 14 - 17, 2002.

PRODUCT DEVELOPMENT PROCESS ONTOLOGY

Ade Mabogunje, Poul Kyvsgaard Hansen, Ozgur Eris and Larry Leifer

Keywords: Research Cumulation, Research Classification, Indexing Research

1. Introduction

In prior papers, we argued that the inherent complexity of product development research, and the variation in research methodologies and settings result in the lack of qualitative cumulation of research findings. We advocated the need of applying an ontological approach to overcome the problem [Eris et.al. 1999] and [Hansen et.al. 2001]. Specifically, we developed an initial ontological framework to facilitate the sharing and comparison of methods, variables, and results. Furthermore, we illustrated our approach by feeding the findings of published studies into the proposed framework as concrete examples. Our data set consisted of the findings from 19 papers published as the Delft Protocol Workshop [Cross et.al. 1996].

In this paper, we test and develop our initial ontology further by beginning to apply it to a much larger data set: all of the papers published in the proceedings of the International Conference on Engineering Design, 1999 and 2001 – 717 papers in total. Naturally, the broader and more comprehensive scope of the new data set challenges our assumptions and framework. In this study, our goal is to devise a repeatable validation method that improves the initial ontology. Our ultimate goal is to apply the resulting validation method repeatedly in order to make the ontology comprehensive and consistent.

2. The product development process ontology

Ontology can be defined as the specification of a conceptualization [Gruber 1993]. That is, an ontology is a description (like the formal specification of a program) of the objects, concepts, entities, and relationships that can exist in some area of interest. A conceptualization is an abstract, simplified representation of the area of interest. Its affordance is the consistent communication in a domain of discourse without necessarily operating on a globally shared theory.

The term “ontology” is mostly applied in the development of large, enterprise-wide information systems to support consistent product development across a geographically distributed organization. This is achieved through the use of a pre-defined ontology that enables enterprise-wide consistency and commonality through the interoperability of processes, systems, and databases [Uschhold et.al. 1998]. Other purposes are development of systems for supporting different disciplines of engineering design work by providing ontologies of for example functional concepts [Kitamura & Mizoguchi 1999].

Ontologies are often equated with taxonomic hierarchies of classes; however, ontologies need not to be limited to these forms. Within the field of product development, several attempts have been done to apply a taxonomic approach. Our initial attempts at composing a framework were built on a similar taxonomically understanding. However, after preliminary analysis of data, we decided that an ontological approach would give us more freedom in achieving our goal due to the looser and pragmatic structure it entails.

The conceptualization of the phenomenon "Product Development Process" has strong analogies to the problem of modeling processes in general. This is particularly the case when the aim is to simulate processes. In particular, we have got inspirations from the development of a computer-based project simulation tool, Virtual Design Team [Levitt et.al. 1994].
The conceptual model of a "Product Development Process" includes the five objects: Activity, Environment, Actor, Information, and Artifact. Although many researchers imply such a conceptual framework, the objects are rarely referred to explicitly. Rather than referring to an "Actor" most researchers refer to some of the "Attributes" of an actor, e.g., the skills, the roles, the knowledge, or more likely, to some "Sub Attribute" describing for example the specific skills of an actor. In essence, the lack of an explicit framework is what makes comparison between different research findings so difficult.
Our initial analysis of the Product Development Process topic consisted of a number of iterations where empirical sub attributes were identified and structured. The objects and the structuring of the main attributes were products of theoretical considerations; they were our own interpretations. The sub-attributes appeared explicitly in the published material and constituted the data. Although the structuring effort was rather labor-intensive, we believe that the richness of the manual procedure we employed proved to be more effective in making accurate distinctions than the employment of an automated approach.
However, the identification of sub attributes and the categorization of main attributes are not enough to construct an ontology; that would yield a taxonomy. An ontology differs from a taxonomy by highlighting the *non-linear relationships* and *dependencies* between the identified categories. Therefore, the structure presented on Table 1 does not quite constitute an ontology. Our conceptualization of the Product Development Process contains many undetermined links (Figure 1).
Therefore, the goals of this study are twofold: 1) Test the empirical validity for the sub attributes by applying the structure represented on Table 1 to a larger data set, and refines the sub attributes as necessary. 2) Identify relationships between objects.

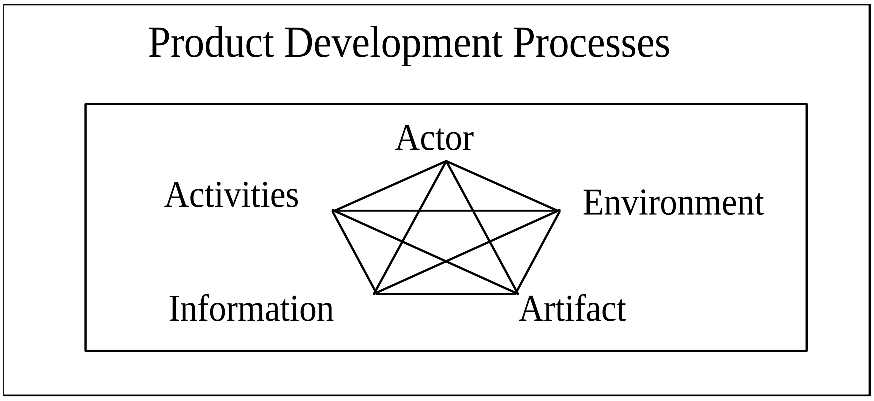

Figure 1. The undetermined links in the conceptualization of the Product Development Process

3. Utility of the ontology

The potential applications of the ontology are many, e.g.:

- Providing a tool for classifying papers and workshops for conferences within product development
- Providing a tool for adhoc search in recent publications within a particular field
- Providing a tool to assist new Ph.D. students in framing their focus for their study
- Providing a tool to support the formulation of research hypotheses
- Providing a tool to associate specific research methodologies with the particular focus of the research
- Providing online access for researchers to classify own papers and compare with matching papers within the particular field

- Support in classifying the different disciplines within the Product Development research area

However, we need to be modest at this stage of our work, and will not be addressing all of the potential applications. We have defined our contribution as a long-term research that has to mature slowly. During the development of the ontology we have interacted with, and intend to continue interacting with researchers both inside and outside of our own discipline. Even though our social network has grown considerably since we initiated this project, it has not yet reached a size we can consider critical mass.

In this paper we will focus on first application: utilization of the ontology to improve the classification of papers and workshops at conferences. Many researchers have experienced the feeling of being placed in a wrong session at a conference. This mainly results from the difficulties in classifying papers. The ICED (International Conference on Engineering Design) conferences have several hundreds of participants, and the chances of meeting the "right" people outside your traditional research community are limited during scheduled sessions if the classification of the papers fails to associate relevant topics. Within a contemporary research area such as Product Development, this is particularly true.

Ideally, we would like the ontology to support establishing a profile for each paper and to identify and associate papers with relevant topics. If that is done prior to a conference with all of the submitted papers, authors of papers sharing similar profiles can be notified in advance. Also, any conference participant can define a profile according to his/her interests.

As an example of this process, we will demonstrate the current functionality of the ontology by applying it to the past two ICED conferences in the final section of this paper.

4. The current status of the ontology

In this paper, we primarily focus on the "Activity" object. In a prior paper, we have focused on refining and verifying the "Actor" object according to the initial proposal [Hansen et.al. 2001].

The empirical basis for the analysis has been expanded to the Proceedings of the International Conference on Engineering Design 1999 and 2001. These included 717 papers covering the broad field of product development processes. Apart from providing an ideal updated data set for this study, the original categorization embodied in the proceedings confirmed the need we have been advocating for an ontology: Even though the 717 papers all have a title and three to four keywords, and are divided into predefined themes, it is difficult to identify papers with overlapping research topics.

We analyzed the papers by using a number of different text analysis tools. We found out that many of the available text analysis tools were too complicated and not transparent enough to provide the digest type of information we needed. The most useful tool was an indexing tool utilized by major Internet search engines.

In our initial proposal for handling the "Activity" object [Eris et.al. 1999], we categorized activities into three main attributes: cognitive, social, and physical. We then divided the three main attributes into eight paradigms. We identified the associated sub attributes by manually reviewing the Delft Protocol Studies [Cross et.al. 1996].

Table 1. Initial proposal for ontology related to the Activity object

Objects	Main Attributes		Sub Attributes/Thesauri
Activities	<u>Class</u>	<u>Paradigm</u>	<u>Action</u>
	Cognitive	Information processing	gathering, recognizing, accessing, visualizing, perceiving, reviewing, managing
		Problem-solving	adopting, referring, proposing, structuring, generating
		Decision-making	examining, analyzing, synthesizing, reasoning, inferring, deducing
		Function-evolution	decomposing, reinforcing
		Problem-solution interaction	evaluating, checking, testing, monitoring
	Social	Mediation	controlling, conflict handling, negotiating, persuading, arguing
		Information sharing	listening, speaking, documenting
	Physical	Object-engagement	gesturing, moving, touching, riding wearing

When applying the initial structure to the larger data set, it became clear that the distinction between different paradigms were too fine grained. Therefore, we merged some of the paradigms. We reduced the original five paradigms within the cognitive class to three: analyzing (including the gathering and retrieving of information), synthesizing (including the understanding of information), and evaluating, [Anderson & Krathwohl 2001]. We reduced the two paradigms within the social class to one.
There was some overlap between the sub attributes associated with the "Physical" main attribute (cf. table 1) in solely technical papers. Upon going through more such iterations, we were left with only one sub attribute: gesturing. We decided that gesturing is mainly associated with social activities, and included it. Consequently, we decided to leave the "Physical" main attribute open.
The analysis of the "Cognitive" main attribute revealed a number of problems. Our searches showed that many of the originally identified sub attributes were ambiguous; most authors used the terms to express their *own* cognitive activities as researchers as opposed to the cognitive activities of the actors they were studying. Distinguishing between the two is challenging and requires additional syntactic analyses that our current search and indexing tools do not support. We have applied such methods in prior studies [Mabogunje 1997]. We learned that they are extremely time consuming, and decided to make the necessary distinctions at the next stage. At this stage, we limited our effort to conducting manual reviews to exclude the use of the cognitive terms as references to the authors' *own* cognitive activities. Clearly, this limitation biases our findings. Therefore, we will not draw any conclusions from them in this paper. However, we will use them to illustrate how we would go about drawing conclusions if the biases were removed.
Furthermore, we concluded that terms such as "Analyzing" "Managing", "Generating", "Structuring" and "Testing" were too ambiguous, and excluded them from the sub attributes.
Table 2 shows the current version of the ontology including both the "Actor" and the "Activity" object. The current refinement of the "Actor" object is described in [Hansen 2001].

Table 2. Refined version of the Ontology including the Actor and the Activity Objects. The asterisk signals that also derived words would be included

Objects	Main Attributes		Sub Attributes/Thesauri
Actors	Skills, Knowledge, Roles		talent*, abilit*, skill*, experti*, routine*, experience*, aware*, understand*, tacit*, unfamiliar*, familiar*, eligib*, responsibi*, role*, novice*
	Motivation, Values, Emotions		enthusiasm*, incentive*, commitment*, motivat*, belief*, emotion*, curious*, relax*, resistance*
Activities	<u>Class</u>	<u>Paradigm</u>	<u>Action</u>
	Cognitive	Analyzing	identifying, retrieving, gathering, distinguishing, selecting, intergrating, outlining, parsing, deconstructing, decomposing
		Synthesizing	interpreting, clarifying, paraphrasing, representing, translating, exemplifying, illustrating, instantiating, classifying, categorizing, subsuming, summarizing, abstracting, generali*, inferring, extrapolating, interpolating, predicting, contrasting, mapping, matching, comparing, explaining, visuali*, proposing, structuring, perceiving, recogni*
		Evaluating	monitoring, detecting, testing, assessing, judging, evaluating, checking, reviewing, auditing
	Social		negotiat* arguing argue mediat* coordinat* listening speaking, gesturing
	Physical		?

Following our gradual development principle we will proceed in small steps, and perform checks between them. Given the size limitation of this paper, we will do not discuss the refinement process of the ontology in detail. However, we will demonstrate and test the current state of the ontology by applying it to the large samples of papers from the ICED 2001 conference.

5. Applying the ontology on the ICED 2001 papers: An illustration

When refining the structure and the elements of the ontology, we reviewed papers from the 1999 conference as well as the 2001 conference. However, to keep this discussion brief, we will restrict our focus to the 2001 conference.

Our current search tool facilitates a text search entered as a query (a combination of words, phrases, or keywords associated with the document type). The output of the search is ranked according to these criteria:

- Whether the words or phrases are found in the first few lines of the document (for example, in the title of a web page).
- The frequency of occurrence of a query word or phrase. Rare words in a query are weighted more heavily than common words (rarity is determined by the number of occurrences of the word in the index).
- Whether all of the specified words or phrases appear in a document. A document containing all three words specified in a three-word query would rank higher than a document containing only two or one of the words.
- Whether multiple queries words or phrases are found close to each other in a document.

One of us had a separate paper on the ICED 2001 conference [Mabogunje et.al. 2001]. According to our search the paper is ranking within the 20 highest at the following classifications: Actor Object Motivation etc., Activity Object Synthesizing, Activity Object Social. The five papers ranking highest within each of these classifications are listed in table 3.

Table 3.

Classification	Highest ranking papers in classification
Actor – Motivation etc.	436-Ono; 412-Lloyd; 446-Muir; 233-Viktorsson; 411-Link
Activity – Synthesizing	221-Lim; 206-Eckert; 231-Liang; 417-Steel; 219-Samuel
Activity – Social	256-Stacey; 412-Lloyd; 261-Eckert; 186-Johansson; 124-Bender

To illustrate how we might check the validity of our approach in future, we read through all these papers with the aim of finding the areas they had in common with the above referenced paper. Our review demonstrated a similarity of the papers in the Activity-Synthesizing and Activity-Social classes and a difference in the Actor-Motivation class. This class was different because while the reference paper contained the sub-attributes in the motivation class, a qualifier to the effect that the author had not addressed these issues preceded the relevant sentence. This has two implications: First, it shows that one cannot infer an author's interest by looking at the sub-attributes alone because the context of the sub-attributes is equally important; Second, it demonstrates a way by which we can cumulate research by identifying related studies. For example, consider the case where author A looked at issue X but not Y, and author B looked at issue Y but not X. As a whole, the community could be said to have studied X and Y, but neither A nor B are necessarily aware of that. When we associate their work this way like a jigsaw, we can accelerate the resolution of the research puzzle.

6. Next step

We anticipate tackling two major issues in order to reach a position where we can conduct a conclusive verification of the ontology: expanding the sub attributes and identifying and/or developing software tools for text analysis.

Initial tests showed that the current number of sub attributes associated with each main attribute category does not provide the required degree of discrimination during article analysis. Searching articles for a limited set of sub attributes does not yield enough hits, neither does it allow us to identify a meaningful context for what is discussed. We propose to overcome this by continuing our current efforts to extracting more sub attributes from a richer sample of design research articles such as the ICED 2001 proceedings.

Initial tests also demonstrated that the software search tools we currently utilize are inadequate. In order to overcome the significant false attribution error discussed earlier, we need to be able to alter

the search heuristic of the tool. We also would like to have the search tool to have embedded statistical analysis capabilities in order to increase overall accuracy and decrease analysis time.
Regarding the representation of the results we would like to be able to generate an individual profile for each paper. This again calls for new tools.

7. Conclusion

This paper presents a step in a continuous effort to define an ontology for product development processes. The particular goal of this step has been to capture the "Activity" object of the ontology. On our first attempt, we analyzed and classified activities from the large sample of papers from the ICED1999 and ICED2000, and uncovered a major limitation of our approach: most authors used the cognitive activity terms to express their *own* cognitive activities as researchers as opposed to the cognitive activities of the actors they were studying. In order to overcome this limitation so that we can reliably classify and associate the papers, we need new and improved search, indexing, and filtering tools.
Regardless, we believe that the limited illustration we were able to construct demonstrates that the ontological approach has significant potential for classifying and associating research contributions. This belief is also supported by the results we obtained in the previous effort where we addressed the "Actor" object [Eris 1999].

References

Anderson, L.W & Krathwohl, D.R., "A Taxonomy for Learning, Teaching, and Assessing", Addison Wesley Longman, 2001

Eris, O., Hansen, P., Mabogunje, A. & Leifer, L., "Toward a Pragmatic Ontology for Product Development Projects in Small Teams", Proceedings of the 12th International Conference on Engineering Design, Lindemann et.al. (ed.), Technische Universität München, München, 1999, pp 1645-1650

Cross, N., H. Christiaans & K. Dorst, "Analyzing Design Activity", John Wiley & Sons, Chichester, 1996

Gruber, T.R., "A translation approach to portable ontologies", Knowledge Acquisition, 5(2), 1993, pp 199-220

Hansen, P., Mabogunje A., Eris, O., Leifer, L., "The Product Development Process Ontology: Creating a Learning Research Community", Proceedings of the 13th International Conference on Engineering Design, Culley et.al. (ed.), The Institute of Mechanical Engineers, London, 2001, pp 171-177

Kitamura, A.M. & Mizoguchi, R., „Meta-functions in Artifacts", Papers of 13th International Workshop on Qualitative Reasoning (QR-99), 1999, pp 136-145

Levitt, R.E., Cohen, P.G., Kunz, J.C., Nass, D., Christiansen, T., & Jin, Y., "The Virtual Design Team: Simulating How Organizational Structure and Communication Tools Affect Team Performance" in Carley & Prietula (Eds.), "Computational Organization Theory", Lawrence Erlbaum Associates, 1994

Mabogunje, A., Carrizosa, K., Sheppard, S.D., Leifer, L., "Towards a Science of Engineering Design Teams", Proceedings of the 13th International Conference on Engineering Design, Culley et.al. (ed.), The Institute of Mechanical Engineers, London, 2001, pp 171-177

Mabogunje, A., "Measuring Conceptual Design Process Performance in Mechanical Engineering: A Question based Approach", Dissertation, Stanford University, 1997

Uschhold, M et. al. "The Enterprise Ontology," The Knowledge Engineering Review, Vol. 13, Special Issue on Putting Ontologies to Use, 1998.

Ade Mabogunje
Stanford University, Center for Design Research
Building 560, 424 Panama Mall
Telephone: 650-725-5010
Telefax: 650-725-8475
E-mail: ade@cdr.stanford.edu