DOCUMENT MANAGEMENT SYSTEMS AS AN ESSENTIAL STEP TOWARDS CIC

Assoc. Prof. Ziga Turk
University of Ljubljana, Department of Civil Engineering,
Jamova 2, Ljubljana, Slovenia.

Prof. Bo-Christer Björk,
Curt Johansson,
Kjell Svensson,
Royal Institute of Technology, Department of Building Economics and Organisation,
S-100 44, Stockholm, Sweden.

Abstract

Within the context of the construction industry Electronic Document Management (EDM) systems offer the means for rapidly achieving a shallow level of integration by providing process integration and information management on a coarse, document sized level of detail. A question which so far has received little attention is how construction document management (CDM) can provide a smooth transition toward computer-integrated construction, based on full-grown building product data models. An important research question for researchers developing a theoretical basis for CIC is to define hybrid conceptual models which synthesise these two levels of data management and describe accurately in an application independent way the CIC processes where a part of the information is managed by CDM systems and a part using product and project models.
As a pre-stage to the definition of formal models of the information manipulated by CDM systems we elaborate the functional requirements that such systems should fulfil. Next we describe a four dimensional CDM modelling space - product dimension (relation to product model and product decomposition), time dimension (document life cycle), organisation dimension and finally presentation dimension. In the end a document classification table which includes generic document properties sorted according to the four dimensions described above is presented and directions for further research are indicated.

1. INTRODUCTION

Electronic document management (EDM) is a new and rapidly growing branch of information technology (IT). Electronic document management systems focus on facilitating the management of documents pertinent to particular enterprises, projects and work groups in computer networks. In addition to the basic file management capabilities found in operating systems EDM systems contain enhanced features related to the life-cycle and versioning of particular classes of documents. Most systems are capable of finding and moving around documents regardless of the originating software. Some systems also help in managing the

scanning and later accessing of paper-based documentation. The increasing importance of such systems for the construction industry is indicated by the fact that the leading trade software show for this industry, the A/E Systems show organised in Washington D.C. in June this year, had a large section dedicated to EDM.

Within the context of the construction industry EDM systems offer the means for rapidly achieving a shallow level of integration by providing process integration and information management on a coarse, document sized level of detail. Integrated construction document management could thus be described as a first stage on the road towards full computer-integrated construction (CIC), which in later stages would consist of information and application integration on a deeper level. The most promising methods for achieving this include object-oriented product and project models. This does not imply that document management will, at some point in time, become obsolete. Documents have been around since the invention of the writing and the document centred communication paradigm is likely to dominate the ways of human doing in the future as well.

2. RESEARCH ISSUES

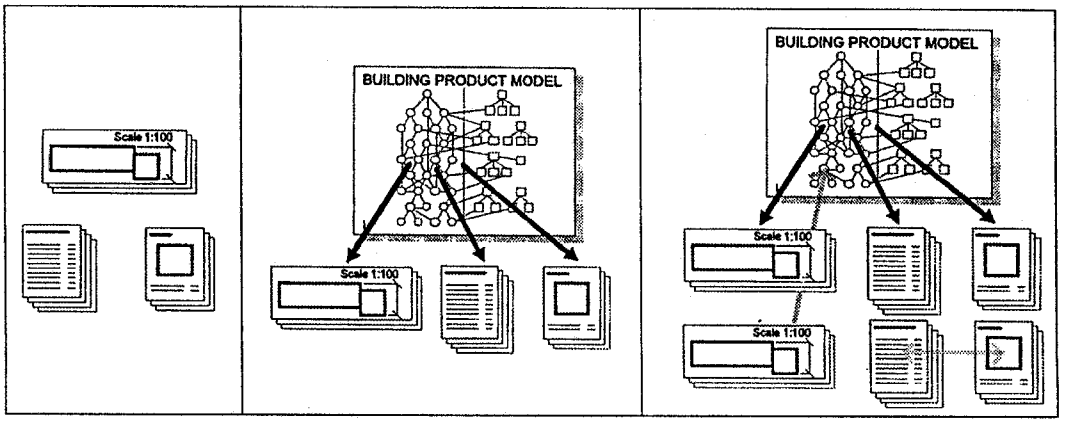


Figure 1: Three types of environments for storing building product information. Left: product information in a conventional environment is contained in documents; Centre: in a product model oriented environment documents are views into the product model which contains product information; Right: in a hybrid environment, some documents are views of the product model and some are original containers of information which may (or may not) point to a product model.

2.1 A HYBRID MODEL

A question which so far has received little attention is how construction document management (for which we will use the abbreviation CDM in this paper) can provide a smooth transition toward computer integrated construction, based in part on full-grown building product data models and in part on information originating in documents. A likely scenario is that CDM, due to the availability of the basic technology and user's familiarity with the concepts, will develop quickly during the

next few years, whereas product data interchange will demand a longer saturation period from basic research to international standards and commercial use.

It is therefore an important research question for researchers developing a theoretical basis for CIC to define *hybrid conceptual models* which synthesise these two levels of data management. Such models should describe accurately in an application independent way the CIC processes where a part of the information is managed by CDM systems and a part using product and project models. An elaboration of such a hybrid model with a focus on documents is attempted in this paper.

2.2 INTEGRATED ACCESS TO DOCUMENTS RELETED TO A CONSTRUCTION PROJECT

Systems for electronic document management aid organisations and the individuals working within them in managing predominantly digital documents which are stored on hard disks and similar media in distributed networks. A typical feature of EDM systems is that they can manage documents of many different types (i.e. not only CAD-files). In addition to fundamental features increasingly found also in basic operating systems, EDM system can be tailored to the needs of particular enterprises or projects teams assembled from different enterprises. There are no strict boundaries between the technologies of EDM, group-ware, work-flow management, concurrent engineering etc.

There are two major contrasting ways of organising construction documents. One type of systems strives to provide integrated access to all the documents manipulated by a particular organisation (i.e. a design firm). Such systems should be able to deal both with project-specific documents from different projects and with company-specific documents presenting for instance standard technical solutions. General purpose EDM systems can easily be used for this purpose.

The other type of system provides integrated access to all the documents produced during a particular construction project, and is the focus of attention for this paper. When we talk of construction document management systems we imply this type of systems. The development of this second category of systems is much more difficult since it involves standardising the data structures used by several constantly changing organisations and the possible implementation of the data structures in several different EDM systems which should be capable of exchanging information. Thus the development of such systems offer more of a challenge for both research and standardisation.

3. RELATED WORK

Earlier theoretical work on this subject area is almost non-existent. Commercial development (Billing 1993) seems to precede the formulation of theory in this domain. Björk et al (1993) presents a first attempt at a conceptual model of meta-data about documents which would form the core of a CDM system. Otherwise most researchers developing formal data structures for construction computing seem to have concentrated more on product modelling. Recently there have also

been attempts at expanding the use of conceptual modelling to generic construction project models (Luiten et al 1993). Although such generic project information models recognise the fact that activities result in either physical building parts or information, they do not explicitly deal with separate categories for product model data and documents not to mention modelling the relationship between these categories.

On the level of the design and manufacturing industry as a whole the STEP standardisation (ISO 1993) effort includes generic models of product data and the presentation of this data in different types of views, which are the basis of documents. The problem is, however, that the notion of document management goes against the basic philosophy of STEP which strives for the exchange of product model data. Consequently there is little interest in STEP to move into the domain of explicitly modelling meta-data about documents. Implicitly many of the data structures contained in current STEP resource models could be of use in modelling document meta-data.

There is a growing concern for the integration of the different data structures used for EDI messages, product data models, general data bases for building material etc. (Neutebooms 1992), (Björk 1994). A recent workshop in Stockholm in April 1994 discussed the possibilities for such integration and there will be follow up activities involving the key standardisation organisations dealing with these issues. The development of formal framework models which integrate document management and product modelling, as outlined in this paper, should contribute to such efforts.

4. FUNCTIONAL REQUIREMENTS OF CDM SYSTEMS

A pre-stage to the definition of formal models of the information manipulated by CDM systems or the processes in which CDM is integrated is the elaboration of the functional requirements that such systems should fulfil.

All crucial events in a life time of a building product are reflected in documents - the client's brief, tenders, bills of materials, design specification, blueprints, minutes from design team meetings, invoices, various legal documents, standards, permits ... We estimate that the total number of documents which relate to a single building structure is in the order of magnitude 10^4. This includes documents from all stages of a building product life cycle including financial documents. Because of the volume current practice partitions the manipulation of these documents according to the organisational structure of the companies involved in the construction project. Most of the documents are stored on paper. Although an increasing number is produced using computers, organised digital archiving using electronic archiving systems is rare.

There are several categories of requirements to a computer aided CDM system which, reflecting its position in a overall CIC framework, includes pure document management, support for building process management, support for building product management and finally a close integration with other CAx technologies.

The majority of **document management** functions are expected to be satisfied by the generic EDM systems becoming available by all major office automation software vendors. The requirements also important for the CDM systems include:

- Electronic archiving of documents or of electronic descriptors of documents not stored electronically.
- Creating, modifying, viewing and printing of documents.
- Getting or referencing external documents.
- Ensuring access security and possible confidentiality of documents.
- Auditing, supervising, approving, scheduling and discussing of documents.
- Management of the relations among documents - versioning, typing, dependencies ...
- Extracting collections of documents or data from documents for i.e. building inspection, facility management ...

The evolution of the document(s) follows the building process. Therefore, to support the **management of the building process** based on the information managed by a CDM system, the following capabilities are important:

- Relating a document to a construction activity;
- Relating a document to construction resource;
- Grouping of documents according to the break-up of the construction activities;
- The above relations are employed for planning, scheduling and supervision of the construction process; quality assurance of the information management of the construction process.

Management of the building product. Documents are the only written description of the building product and of the activities that lead to its construction. Therefore they are an excellent source of information for the usage, maintenance, adaptation and demolition stages in the product's life cycle. There is a significant discrepancy between a typical building and software product life cycle. The first is estimated at about 50 years but in practice there seems to be no fixed limit. Software is updated on a yearly bases. This poses some specific requirements:

- Legal, technical and common-sense rules on how long a particular document needs to be stored should be taken into account. Generally, documents describing technical details of a building should "last" as long as the building itself. Legal or financial documents are only of interest for the time defined by the law.
- Possibility to store information for decades therefore on a media with extremely long life.
- Possibility to discard some of the information over time.
- Documents should be stored in formats that are expected to survive the rapid development of IT. ASCII text and HPGL drawings are examples of such formats; Word for Windows text files or CAD Studio models are not.
- Documents should be stored in such a format that they could be used for adaptation, redesign etc. This may be contrary to the previous requirement or calls for multiple representation of a document.

Aid in automation of the construction process; CDM systems should become part-of, if not the core, of a construction project shell. In addition to the above functions interfaces and integration of engineering and office automation software should be possible. It should provide construction work-group environment for the activities of those stages in the life cycle of a building product before it is constructed and in which document manipulation is intense.

5. CDM MODELLING DIMENSIONS

The CIC universe of discourse may be functionally decomposed into a number of distinct modelling domains. These include models of the building as a product, models of project documentation and models of related and more general building information such as regulations. These models are all related to a model of the building process. An analysis of the dependencies between the models reveals that there exist client/server relationships between some of them. This enables autonomous modelling of the servers such as the product model. When modelling the clients, however, the properties of their servers must be considered.

In general, documentation is one of the views of the product model and thus a client of the product model. In theory, proper CDM systems could not be modelled and even less so implemented, without well-defined product models. In reality, however, it is possible to perceive the documents as stand alone objects. These are now created and used even though such product models are not used in practise. In the future some of those will be a result of some post-processing of the product model. CDM systems should be constructed in such a way that in the future this relationship could be established.

To encapsulate the future evolution of a CDM system model, its universe of discourse was organised into a four dimensional modelling space (Figure 2). The partitioning was also influenced by the belief that a document is a special kind of a product. And just like a building product is a result of a building process a document "product" is a result of a document creation process which is part of a building process. The whole product dimension relates a document to a product model.

The dimensions are:

- **Product dimension**. We understand the building product as the central concept onto which all other types of information like process data and documents are related. Therefore the product dimension defines an open ended interface to the product model and to the prospective process information. Design documents describe the product or its parts; financial documents relate to the prices of the product or the involved labour etc.
- **Life cycle dimension**. Just like a building product goes through its life cycles the same is true for a document but the life stages are different and are not, in general, related to the building life cycles.
- **Organisational dimension** reflects the impact that the organisation of the building industry has on the manipulation of the documents. The axis is partitioned into three sections - project, organisation (company) and general.

- **Presentation dimension** is concerned with form in which the documents are presented to the (human) users, to the software tools and to the IT environment.

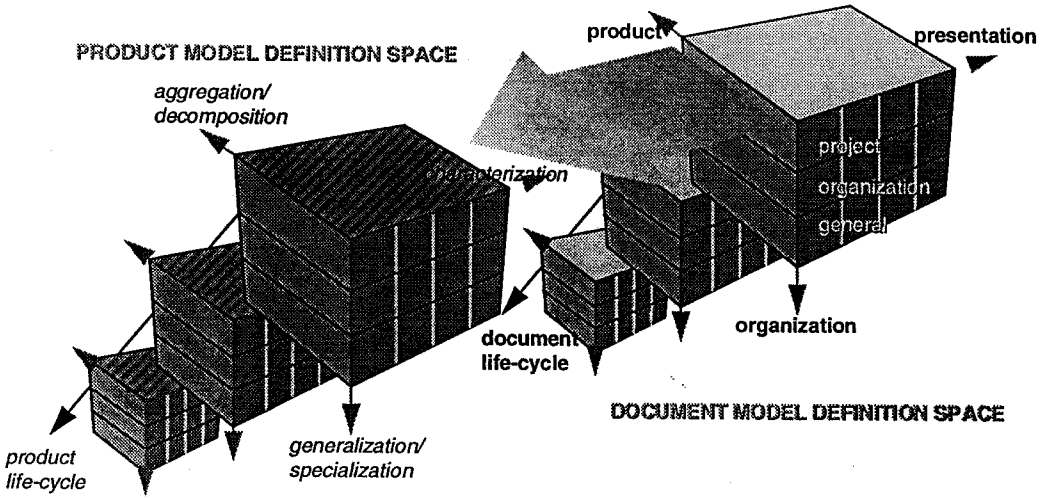


Figure 2: Product model (left) (Gielingh 1988) and document model (right) dimensions.

To ease the development of a CDM system conceptual model, a document classification table (see Appendix A) was defined. The table lists the criteria according to which the documents can be classified. At the same time, most of the criteria are attributes that need to be attached to the document so that it could be managed by the document management system. This is called document's meta-data (Bjork et al 1993). Meta data is created and changed by the CDM system. The documents themselves are manipulated by means outside the CDM system - if computerised by software, if not by hand.

The classifiers in the classification table are roughly sorted according to the dimensions of the modelling space. They include external references to documents like TH17, BSAB or KBS (Svensson 1991) which provide documented and quite standardised set of possible values for some of the criteria.

7. SYNTHESIS MODEL

The semantic network in Appendix B shows a CDM system conceptual model. A more generalised view is in Fig. 3 Underlying shading shows the areas of the network related to the model space dimensions. Synthesis model is presented which relates CDM objects to process and project models. Plain English description circling the network around the central concept clockwise starting at lower left section (Appendix B) is given below.

Documents differ according to their representation and content. Some are digital, some are not. The digital ones are manipulated with (software) tools by actors. Similarly to the solution in (Turk 94) documents are viewed as objects and tools as

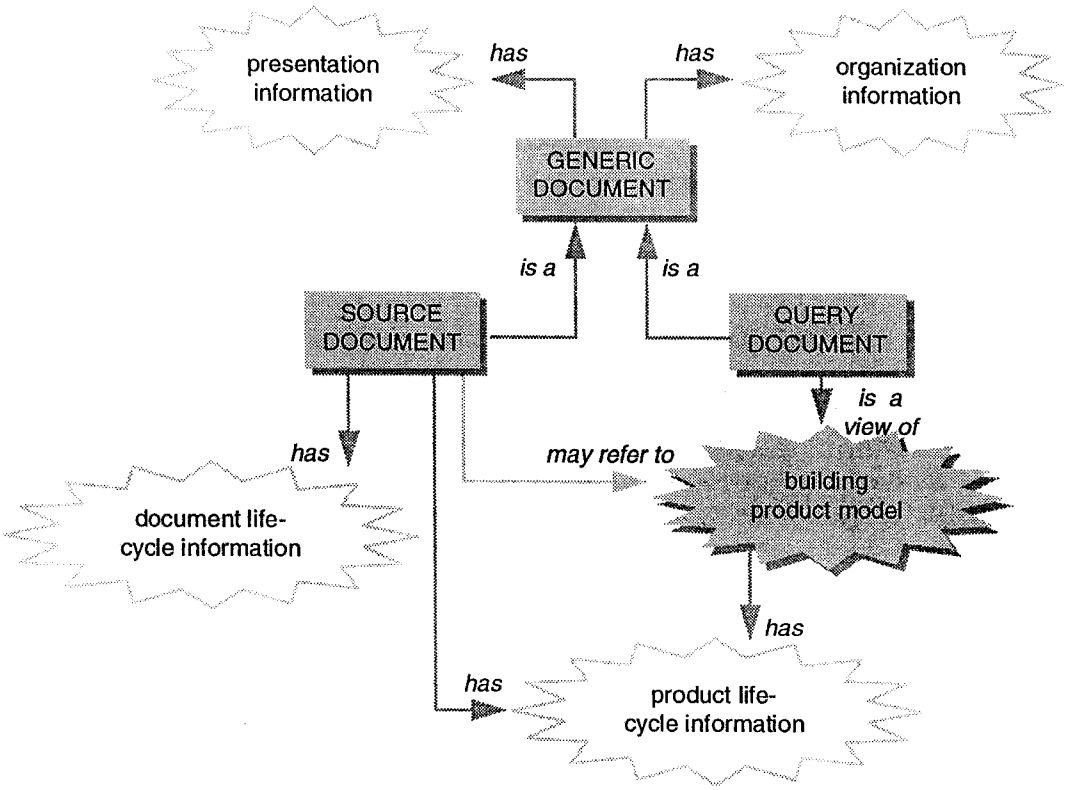


Fig. 3:The relations between the types of documents, the building product model and various document related information.

methods. Object oriented view on this part is further emphasised by a taxonomy of document types and inheritance of tools.

A document is either a simple document or a compound document - that is a collection of other (simple or compound) documents. Implementation of this feature will probably use some document centred user interfaces of modern operating systems like OLE 2.O or Open Doc. Documents are part of a document set. One set of documents may belong to a particular project, to an organisation (company or department within the company) or may be of a general nature (like standards and codes). It seems tempting to allow one document to belong to more than one project or both to a project and to a company's documentation. Literally this is not possible, but a meta data in one document set may reference a document in another document set.

Documents have current and past states. There are some generic kinds of states but some document types may have additional or more specialised states. For example an invoice may have a states "received" or "paid" or "disputed". These states are meaningless for a design document like a blueprint. Document states are changed by activities. Activity are planned with a schedule. A schedule is a special kind of "virtual" activity. Most attributes are common to both of them. An activity is part of a task. Each task belongs to one life cycle stage which is part of the building process in which building product is made or used

The most important subclass of construction documents describes a set of building product parts which are part of building product model. Alternatively, in the future

some of the documents (*the q-documents* in the diagram) will be results of queries into the building product model data - a special view of a set of building product parts. The grey boxes and lines in the diagram show the concepts that will evolve into use when the product model databases will be used. It is clear form the figure that there are not many new concepts and clearly not many new relations involved to make this step. The future *building product parts* entity has a lot in common with the present *document* entity - its grey part. The *q-document* entity has a lot in common with the *document* entity - the white part of it. Both inherit those features from the document entity.

That is why we consider the model as a suitable long lasting solution for CDM systems without product models and for hybrid systems where some documents will be views of the product model database and some creations of their own. A generalised taxonomy around the document entity is in Fig. 3.

8. CONCLUSIONS

Document management is one of the fastest growing sub domains of general computing. This is partly due to the fact that the basic hard- and software technology (computer networks, graphical user interfaces etc.) is available, partly due to end users realising the huge gap between the techniques used for generating single documents and for managing volumes of documents in organisations.

Construction document management systems focusing on the integration of the documents manipulated by a particular enterprise will develop quickly in the next few years driven by commercial development. The need for explicit research work is not so strong for this type of applications. On the contrary the development of CDM systems for the integration within single projects across organisational boundaries is much more complicated. In this domain research can help practice by providing theoretical framework models and suggestions for conceptual models of document meta-information. Standardisation, both on the national and international level, could also have an important role to play.

In this paper, an attempt was made to define a schema framework for a hybrid - document and product model based - computer integrated environments and to semantically model the related information. It has been shown that by structuring the document modelling definition space into several dimensions the overlapping areas are virtually non existent. Detailing of the presented semantic network into a proper conceptual model in close correspondence with the existing process models would be a logical next step.

9. ACKNOWLEDGEMENT

The majority of work published in this paper has been done during a visit by the first author's to the Royal Institute of Technology in Stockholm. The visit was part of a pre-study of CDM systems carried out at KTH during 1993-94. This pre-study will be followed up by a more detailed research project involving other research institutions as well as commercial enterprises.

10. REFERENCES

Billing, K. (1993), *Engineering Document Management Systems*, Cadence, August 1993.

Bjork, B-C., Houvila, P., Hult, S. (1993), *Integrated Construction Project Document Management*, submitted paper, EuropIA'93 conference.

Björk, Bo-Christer (1994) *Conceptual Models of Product, Project and Document Data; Essential Ingredients of CIC*, Proceedings of the ASCE First Congress on Computing in Civil Engineering, Washington D. C., June 1994, 5 p.

Gielingh, W. (1988). *General AEC Reference model (GARM)*, in Christiansson, P., Karlsson, H., Conceptual Modelling of Buildings, CIB Proceedings, Publication 126, W74+W78 Seminar, Lund, October 1988.

ISO TC 184 SC 4 (1993) *The STEP standard.*

ISO TC 184. *Draft International standard DIS 10303.* Is continously since 1992 being published in several different parts.

Luiten, G., Froese, T., Björk, B-C., Cooper, G., Junge, R., Karstila, K. and Oxman, R. (1993) *An Information Reference Model for Architecture, Engineering and Construction* In Mathur, K., Betts, M. and Tham, K. (edts.) Management of Information Technology for Construction, World Scientific & Global Publication Services, Singapore 1993, pp. 391-406

Neutebooms, J. (1992) *Bridging the Gap Between EDI and PDI* Preproceedings of the seminar Technical Data Interchange in Construction, CSTB, Paris

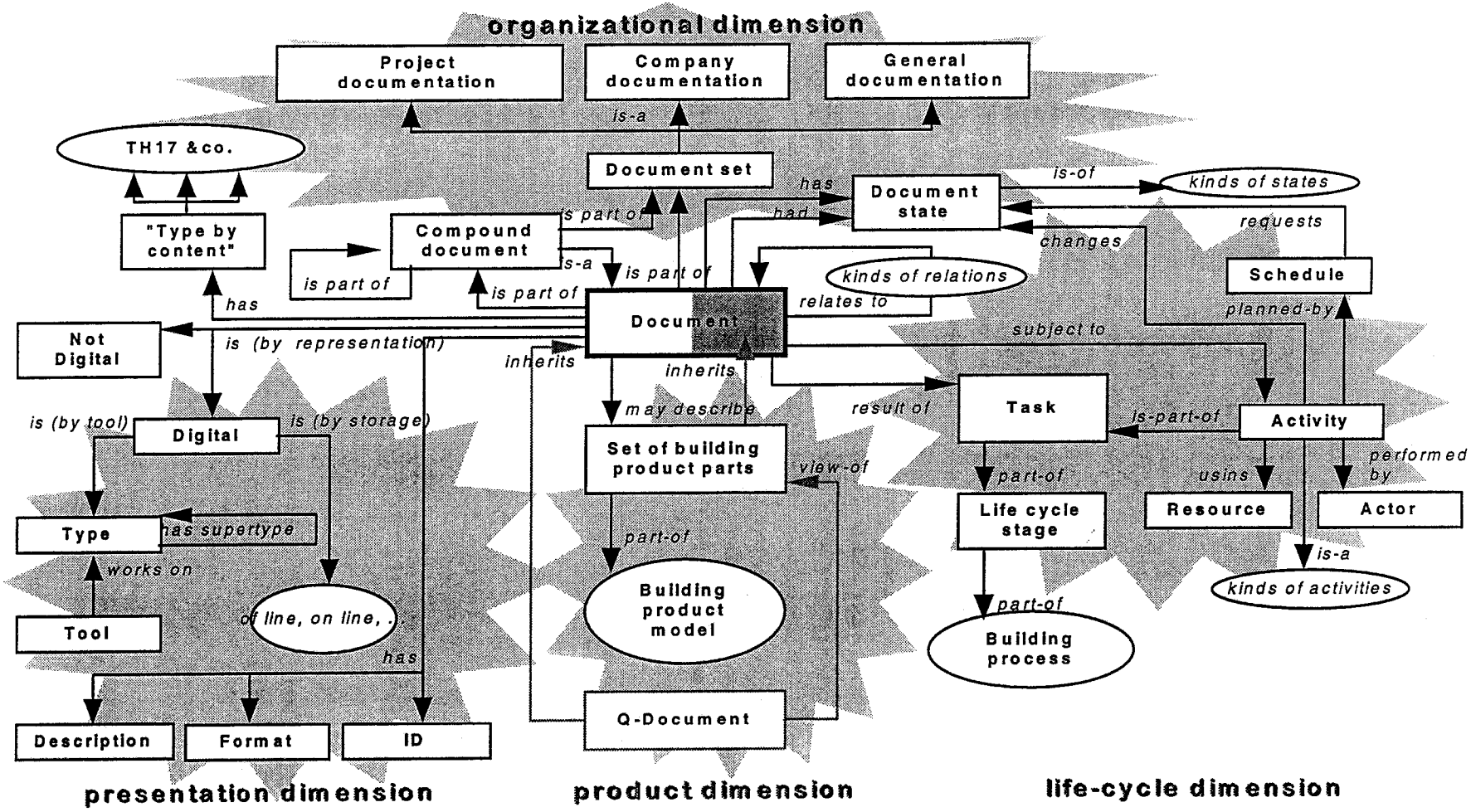
Svensson, K. (1991) *Neutral Building Product Model - "The KBS Model"*. Report T:123E, National Board of Public Building, Stockholm, Sweden.

Turk, Z. (1994) *Three prototype CDM systems*, accepted paper; CIB W78 Workshop on Computer Integrated Construction, VTT, Helsinki, Finland, August 22-24, 1994.

APPENDIX A

CRITERIA	CARDI-NALITY	RESTRICTIONS	SOURCE	SAMPLE CLASSES	COMMENT
content					
Type by content	1	none	TH17/AG1		
Product related	0-n	none	BSAB, KBS	wall, groudwork, door, opening ...	
time					
Building life cycle phases	0-1	none	(Gielingh 1988), (Svensson 1991).	conception/feasibility, design, tendering, construction, facility management, demolition/recycling	Break the phases into tasks ?? yes Subclass these phases?
Event management	1	Type by content ?		needed, not needed	May be very defined for some legal or accounting documents!
organization					
Scope	1	none		project company general	Depends on how you "group" documents, can 1 doc. be in more than one "scope" at the same time ?
information technology					
Representation	1	Storage "on line" implies digital		digital not-digital	
Format	1	some valuse at "type by content" imply certain formats		paper (fax, book, sheet, blueprint, roll, form), audio, video, physical object ...	
Type by tool	1	applies only to "digital" representation		ascii test file, WW document, DXF file, ...	creates inheritance tree
Storage	1	none		on-line (local or remote) off-line	

APPENDIX B


organizational dimension
Project documentation
Company documentation
General documentation
is-a
TH17 &co.
Document set
has
Document state
is-of
kinds of states
"Type by content"
Compound document
is part of
had
changes
requests
Schedule
is part of
is part of
is part of
kinds of relations
relates to
has
Document
Not Digital
is (by representation)
subject to
planned-by
inherits
inherits
is (by tool)
Digital
is (by storage)
may describe
result of
Task
is-part-of
Activity
Set of building product parts
view-of
performed by
part-of
usins
Type
has supertype
Life cycle stage
Resource
Actor
works on
part-of
is-a
Tool
of line, on line, .
Building product model
kinds of activities
part-of
has
Building process
Description
Format
ID
Q-Document
presentation dimension
product dimension
life-cycle dimension