Common Informations Spaces: A few lessons learned about real-life archives and the structuring of information.

Peter Carstensen
The IT University of Copenhagen
carstensen@it-c.dk

1 Introduction

Many areas in industry are confronted by increasing demands for improved quality products, shorter lead-times, higher flexibility, etc. To cope with the demands engineering design is often undertaken by large groups of engineers organized as concurrent engineering settings. This implies that more actors become involved and an abundance of decisions have to be made by mutually interdependent actors. When the number of people involved in an engineering design project exceeds the limit of a few, they need to examine the state of affairs in the field of work, both with respect to the here-and-now ongoing activities and the state of affairs over a longer time horizon. Actors involved in complex cooperative work will apply different artifact based mechanisms supporting them in keeping track of state of affairs. Thus, common information spaces (e.g., files, archives, etc. and concomitant conventions and procedures for its use) are invented and used for structuring and filing information, and for retrieving relevant information when required. Common information spaces provide support for sharing information, keep track the process and the progress, support the coordination of the distributed activities, etc.

A common information space is here regarded as a structure containing information, objects, etc., used by collaborating actors, i.e., both the structure and the content are established, maintained and interpreted by the involved engineers, designers, etc. A common information space is — as discussed at the Copenhagen CIS'2000 workshop[1] — a discrete collection of items used in common by a number of collaborating actors. Information spaces are essential for performing cooperative work. As stated by Bannon & Bødker (1997): “To the extent that multiple actors can construct and maintain a common information space, they are able to articulate their work.”

In order to conceptualize and establish the required building blocks for designing IT-based common information spaces a deeper understanding of how common information spaces are established, maintained and used is required.

This paper reports an analysis of a case study addressing how the (paper- and folder-based) common archives (hereafter called the PVE II-file) of a large engineering design project were established, structured, maintained and accessed by a team of engineers with different background and expertise. The aim was to provide a basis for designing computer-based common information spaces through investigating what kind information the design engineers filed, how they structured the archives, and how they requested information from the archives.

---

1 A previous version of this paper was presented at 'CHP CIS'2000', a workshop on 'Cooperative Organization of Common Information Spaces', 23-25 August 2000 held at Center for Tele-Information, Technical University of Denmark.

Our major findings were that (1) the structures in a common information space evolves over time, (2) the structuring principles applied were very simple (at least in paper-based files), and (3) the content of a project group common information space is accessed by many different actors having different background, expertise and needs. This supports the dialectic approach to common information spaces suggested by Bannon & Bødker (1997) demanding that common information spaces on the one hand must be open and malleable within a local community of practice, and on the other must be boundary objects (Star, 1989) and immutables allowing sharing across contexts and different communities of practice.

Another interesting observation was that the most common type of information request concerns 'who will know something about...', i.e., the actors request for an 'intelligent' advice, rather than factual information. Furthermore, our studies illustrated that the actors did not aim at filing explicit descriptions of their design rationales. Instead, whenever needed the rationale behind certain decisions were interpreted from specifications of the decided solutions and from information on what had been considered, usually in terms of minutes of meetings. This might question the need for explicit capturing the design rationale, as suggested by different authors.

2 The work setting

One large engineering design project was studied. The aim of the project was to design and plan the production of a new and improved PVE. A PVE is a small component to be designed into fork-lifts, trucks and similar machinery to control the hydraulically based technology. It consists of a very complex electronic structure including a number of chips and flex-prints encapsulated by a light-weight plastic house.

The PVE II project lasted almost 4 years and there were naturally some changes in the manning of the project. Most of the time about 10 actors were involved in the project. Although they all had a background as engineers they came from different departments and had very different expertise and thus different perspectives on the design to be conducted. The typical manning included the project manager and three engineers from electronic engineering. Apart from these one actor from each of the areas of mechanical design, production planning, quality assurance (QA), marketing and purchase were more or less full time allocated. Furthermore, an 'environmental protection guardian' was involved. The project manager, the electronic engineers and the mechanical engineer were physically placed in the same room (called the PVE project office), and the person from the production planning had his office close to. The others had offices in buildings several hundred meters away. The group had a lot of interaction with other specialists, groups, departments, etc. both internally at Danfoss and with external partners including three major external sub-suppliers. Three major external sub-suppliers were involved in the project. These were responsible for respectively: 1) Designing a complex and essential integrated circuit; 2) Compiling the electronic components and encapsulating them into one component; and 3) Development of the flex-print technology to be used.

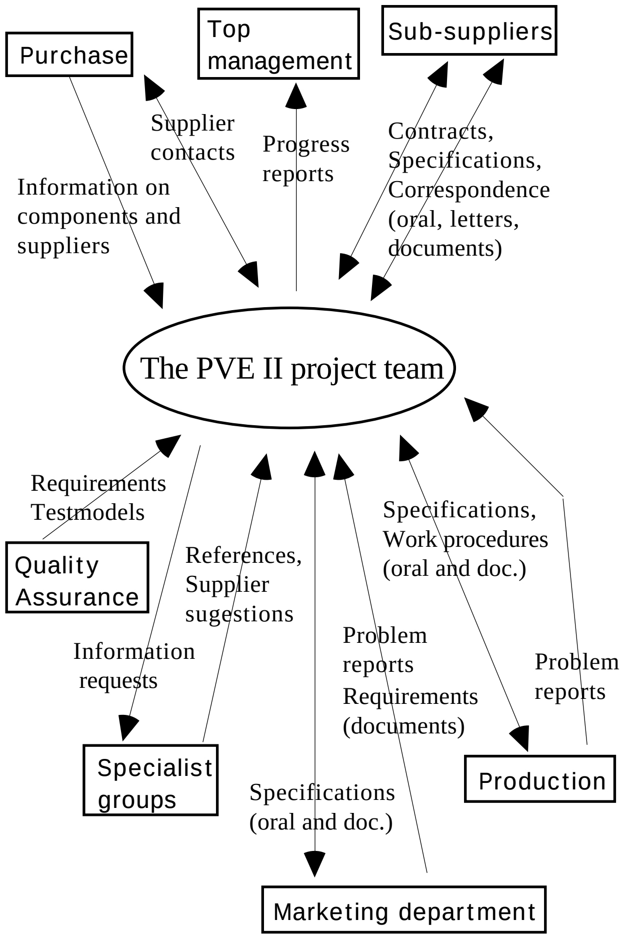

Figure 1: A "context diagram" illustrating the most important interaction between the project team and different stakeholders during the project. From (Carstensen, 1997).

A number of physical files were placed within the PVE II project office :

- The central PVE II engineering file containing binders including specifications, minutes of meetings, correspondence, test results, etc.
- The Drawings file including several shelves of binders containing drawings from many different products and components.
- Production specifications containing specifications and designs of the production lines, requirement specification for suppliers of standard components, operational procedures and manuals, etc.

All actors directly involved in the PVE II project or otherwise related to the project had, of course personal files that (usually) included documents related to the project. Furthermore, the administration department had copies of contracts etc.

The work setting and the overall aims of the project are described in some more detail in Carstensen (1997).

3 Observations on structuring, classification and retrieval.

According to the engineering designers we interviewed and observed, the PVE II archive has been very successful in the sense that it has actually, to a high degree, proved to be possible to structure, store, and re-find stored information. Whether this is due to a very systematic and structured effort from the project manager responsible for the maintenance of the archive during the project combined with a set of rigid procedures for the use of the file, or it is due to a well-chosen 'self explaining structure' of the file is not clear. The project manager put much focus and effort into documenting design ideas, problems, considerations, decisions, etc. during the project. This documentation together with all external and internal communication, prototype test results, etc. were filed in binders and organized in a structure which grew up ad hoc, as the project manager or others in the project thought there was a need. At the beginning of the project the archive structure contained only a few entries, mainly organized according to document types (specification, minutes of meetings, and in and out going mail). The refinements were initiated by the involved actors and implemented by the project manager. Normally changes to, and refinements of, the structure were debated on the weekly project meetings. Several minutes of meetings contains summaries of such debates. According to some of the actors one of the structure discussions was related to whether there was a need for a detailed structure or just a few rough categories.

Today one of the electronic engineers works with the maintenance of the PVE II design, i.e., when customers complain about a technical problem, he will investigate the reason and find a solution. Furthermore, he helps people retrieving the desired information from the archive. We call him the 'PVE file manager'.

As mentioned the archive was placed in the PVE project office. The most important file was the PVE II engineering file consisting of 29 binders arranged with 20 entries, i.e., the information is subdivided into 20 first level categories. Each entry was organized in from one to five volumes (binders). In the very first binder there was an overall table of content for all binders. In most binders there was a table of contents for that specific binder. Examples of first level categories were types of documents (e.g., specifications, photos, minutes of meetings), PVE II sub-systems (e.g., housing design, LVDT design), interaction (mails, faxes, etc.) with external partners, and company procedures and constraints (e.g., authorizations , norms and standard). Most of the entries were divided further and had a second level of categorization, e.g., specification document types, module #, document #, test #, half year time slots, and topics concerning economical status. Less than a third of all the binders had a third level or more of sub-categorization. All information at the final level of structuring were filed in chronological order if possible. In some of the binders, though, the final level order seemed to be arbitrary, e.g., binders with overheads and photos.

Apart from the applied visible structuring principles the PVE file manager obviously had a 'non-explicit hidden structure in his mind.' He knew the design history and the order in which activities took place. This was, for example, useful when entering the in and out-going communication, which was structured by chronology only.

The structuring principles were described (or at least listed) in the table of contents in the binders. Apart from this, the archive did not include any explicit support for how to use and maintain the structure. There were no further aids supporting information filing and information search.

First level categories:

- Type of documents
  Specifications, photos, overheads, minutes of meetings, reports and internal memos.
- Sub-systems in the PVE II product
  E.g., ASIC design, housing design, LVDT design, flex print design.
- Interaction with external partners
  In-going, out-going, offers from suppliers, etc.
- Company procedures and constraints
  Authorizations , managements statements, norms and standards, Q-releases and aspects concerning the production line specifications.
- Activities
  E.g., test of B-samples, housing integration.
- The design process
  Documents on project economy, introduction of the project, time schedules, work planning, etc.

Second level categories:

- None.
- Specification document types.
- Module #, document #, test #, or module #.
- State, i.e., “current, previous, miscellaneous”.
- Half year time slots.
- Non-orthogonal topics (e.g., photos, tests samples, calculations).
- Suppliers.
- Economical status.
- Reports “on topics” (e.g., status, economy, supplier, travel, tests)
- Reports “on company functions/department (e.g., engineering, quality, production).

Third level categories:

- None.
- Arbitrary.
- Report types.
- Chronology.

Figure 2: An overview of the types of categories applied for structuring the PVE project archive.

The files were mainly accessed by actors who participated in the project (engineering designers, marketing people, production planners, etc.). These people were, of course, characterized by having some knowledge about the structure and content of the file. Thus, most requests concerned issues which the requester knew could be found in the file. To phrase it differently: The files were usually used by actors searching for factual information concerning the specifications of the PVE II or about the project. However, in the more general analysis of the designers’ information needs it became clear that several types of information needs were not covered, for example information on others (colleagues, suppliers, etc.) with expertise and experience on specific topics. This type of information was actually considered the most important type of information. Several requests came from people outside the group. They usually accessed the archive through the PVE file manager. Most of the external requests came from external departments such as Purchase, Marketing, and Quality.

Prototypical requests to the files concerned aspects like:

- What is PVE II's specified resistance to water?
- What are the specified requirements for the PVE II — and the actual performances.
- Who will know why we chosed Beta as sub-supplier for the housing technology and why we specified the vibration resistance requirements as we did?
- Why did we chose flex print technology? And
- How was the pricing policy for a variant of the old PVE that was dropped a year ago?

It required a fairly detailed knowledge about the project — and understanding of the archive structure — to be able to formulate search criteria for requests that would result in useful information (apart from sequential search of the binders).

The two first request examples implicitly illustrate the problem of conflicting needs regarding the level of detail in the categorizations. The fourth request on the rationale for a specific decision was dealt with by browsing through the minutes of meetings for the period at which the decision on using flex print technology was taken. The production planner who made this request had been involved in the decision, and he stated that:

> "I spend 20 minutes browsing through the minutes of meetings, then I clearly remembered the discussions and why we made the choices we did. The most difficult was to find the right period - I couldn't remember whether it was in the winter or the spring that we had these discussions."

The last request example illustrates a general problem (dilemma) for the actors. For all requests they had to consider: How many resources are required for finding the information (if possible at all) compared to what it will cost to pass on the request, for example to a supplier? As the PVE file manager formulated it:

> "The problem is that in many situations I do not — at any time in the search process — know whether I am close to finding the requested information."

4 Lessons learned so far

According to our findings, common information spaces applied and maintained by engineering designers should have a compound nature supporting somewhat opposite requirements: On the one hand the structures must be simple, stable and fairly rigid in order to support actors with different background and knowledge about the field of work in sharing (generating, maintaining, filing and retrieving) the information. On the other hand, the structures must be open and flexible in order to support the evolving nature of a common information space. This confirms the findings of Bannon & Bødker (1997) indicating that common information spaces must be open and malleable on the one hand, and immutables on the other.

Even though the project archive studied was centrally maintained, there seems to be a need for structures that can be interpreted, developed and maintained in a decentralized manner by the actors. Furthermore, the study indicated that actors accessing shared information are often interested in references to other actors having knowledge about a certain field than in factual information, both inside and outside of the project team.

In a little further detail than these general conclusions, a number of lessons have been learned:

Lesson 1: Common information spaces are essential collaboration means for cooperating engineering designers. When work becomes complex and several engineers are involved the required horizontal information exchange and coordination activities increase and different types of common information spaces and artifact-based coordination mechanisms are introduced and applied. These spaces and mechanisms support keeping order of the processes and the intermediary products, and provides a basis for information exchange and negotiation.

When the work is conducted over a period of time, and a shared repository including specifications, case documents, descriptions of state of affairs etc. is considered vital for carrying out the work.

Lesson 2: The structures applied in the studied archive appeared surprisingly simple. Most entries in the files were organized in only two levels of categories, and the categories were a mixture of non-orthogonal dimensions (e.g., document types and sub-component names at the upper level). The search support was very simple, nothing but the categories at each level supported searching, and the files had very few cross-references to other entries or other files in the archive. These observations should not in themselves be used as arguments for designing computer-based common information spaces having similar characteristics. They are, however indications on the level of complexity acceptable for the engineering designers. Thus, a few guidelines should be obeyed:

- The classification structure must appear simple, and complex structures should be provided in ways so that it can still be inspected at a glance;
- In many situations rigid structures are easier to use;
- The common information space should support sequential (chronological) searching and browsing;
- And, since users are not used hyper-link structures the use of these need careful consideration. This is probably only a question of time until the users become familiar with www-searching, etc.

Furthermore, a common information space should also include classification features making it easy for workers to decide who has the information and knowledge needed for a particular work activity, and who would have interest in which piece of information or knowledge (a kind of applicability matrix or knowledge map). This implies support for managing and handling disparate know-how and heterogeneous viewpoints, and to make it accessible and suitable for all members of the organization. Also categories of different types of knowledge could be classified in a mixture of non-orthogonal dimensions.

Lesson 3: The structure of a common information spaces is not stable, it evolves. The pattern was that the initial structure was very simple with ambiguous categories that were frequently refined and changed. The refinement was usually triggered by a new and increased understanding (conceptualization) of the field of work, i.e., it was triggered by the involved actors. This calls for flexible structures that can be partially defined and then be regularly refined. Negotiation is a key issue here. Thus, support for refinement and evolution of the structure must also include negotiation support. This demand for flexibility is to some extent in opposition to the overall requirements described in the previous lesson.

Lesson 4: The PVE II project archive that we studied was officially maintained centrally, during the project by the project manager and later by the PVE file manager. Despite this, our study conformed that in order to be supportive a common information space must provide facilities for establishing structures that can be interpreted, developed and maintained in a decentralized manner by the collaborating actors, as also argued by Schmidt and Bannon (1992). Hence, support for (at least) decentralized maintenance of the structures in the common information space should be provided. Decentralized establishment should probably also be supported. Again this requirement will to some extent conflict with the requirements for simplicity and rigidity mentioned in lesson 2.

Lesson 5: The information filed in the PVE II archive could, at a very general level, be related to one of three categories: (1) Process-oriented information including both historical and future-related information on plans, activities, events, etc.; (2) Information concerning functionality and performance of the product (the design); and (3) Product-oriented information concerning and specifying the actual design. This structure might indicate that the aspects of accountability and reuse were considered more important than coordination support.

Lesson 6: When analyzing the types of information requests formulated by the collaborating actors, the most common type of information request concerned people. Actors search for another actor: ‘Who will know something about...’ Our interpretation of this is that most actors prefer an ‘intelligent’ conversation in which the requester gradually can refine his understanding of his problem and thereby refine his search request. A more thorough study of why, and to what extent, this is so is required. The file manager played an essential role when "outsiders" tried to access the archive. It became obvious from our findings that his ability so convert requests into "meaningful requests in relation to the structure" was essential to the success of the archive. This kind of support cannot be replaced by any existing kind of information technology

The need for access to search for other actors has at least two major implications for the design of common information spaces: (1) structures supporting establishing and retrieving information on ‘who knows something about’ in the common information space must be provided; and (2) step-vise refinement and partial definition of the problem that is addressed must by supported.

Lesson 7: Our analysis of the structure and content of the PVE II project archive have not provided concrete examples on filling (and retrieval) of explicit descriptions of the rationales behind specific design decisions. When the rationale behind certain decisions was needed, it was interpreted from descriptions of the design solutions and from information on what had been considered (e.g., from minutes of meetings). It is very difficult to come up with a clear-cut conclusion on this. It needs further consideration and dedicated studies.

The study presented here has, in the line with others indicated that the very basic nature of engineering design implies that the systems must be flexible from the point of view of the daily users. The actors must be supported in managing and maintaining the structure of the common information space in a distributed manner. IT-based common information spaces can therefore not be based upon one pre-defined given structure only. What is needed are malleable and controllable structures that can be applied by the actors themselves in order to establish a useful common information space.

The common information space should offer both access to very specific and concrete information and a more general overview of information on the same topic. Furthermore, similarity search was a very common type of search (Carstensen, 1997). This indicates that common information spaces should include facilities for requesting information like: "give me information on designs similar to this...." An engineering design common information space must, furthermore support a seamless shift and navigation between the different sources, i.e., problems with the different formats, different search machines, different database hosts, etc. should not be visible to the user.

These preliminary requirements have implications for the technological platform upon which the common information spaces are build, e.g., a high degree of flexibility in mixing access tools, access to and across different media, step wise refinement of access profiles, similarity search support, mechanisms for distributed maintanence of the structures and classification schemes in the information space, etc.

5 Open questions

The study presented raises more questions than it provides answers when it comes to considering design of common information spaces. Some of the core questions are:

- What does it mean that information is common?

  Although this study is based on the assumption that the information in the PVE files is common (shared in some sense), I don't know whether any of the actors actually had a common understanding of the information. Perhaps it is only the data that are common, and what are the implications of that?

- Have the classification structures and the information retrieval strategies impact on what can be considered the "organizational knowledge" handled by the work arrangement?

  Other studies have aimed at deeply understanding what knowledge management or organizational memory in an organization really are (e.g., Bannon and Kuutti, 1996; Ackerman and Halvorsen, 1998). It is, however not clear what the constructive implications of these studies are. This problem complex also includes aspects on how to ensure and provide context sensitivity facilities in our common information spaces.

- How do we — in a constructive manner — handle the problem of providing simple, stable and fairly rigid structures, and be open and flexible.
- How to cope with the demand for a distributed maintanance of the classification structures?
- Structures and mechanisms supporting the "information and knowledge creation process". In a similar study of how a group of quality assurance specialists gathered, structured, inspected and distributed information ("knowledge") (Carstensen and Snis, 1999) we found that a core problem in using the tools for structuring the information to distribute was that there were no means support the process of creating information.

- Can certain features of a common information space support better re-use of ideas, plans, decisions, etc.? How, and what are the implications for the classification structure? The study of the PVE II file indicated that much was done to support re-use, but it was not clear to what extent this was actually achieved.

etc. etc.

Acknowledgements

This work could not have been conducted without the invaluable help from numerous people at the Danfoss A/S. A special thanks to Jacob Buur and Flemming Thomsen for supporting the field study work. Part of the data collection and analysis work has been undertaken in collaboration with Annelise Mark Pejtersen and Thomas Miller.

Bibliography

Ackerman, Marc S., and Kristine Halvorsen: “Considering an Organization's Memory,” *CSCW'1998 - Proceedings of ACM Conference on Computer Supported Cooperative Work, Seattle, Washington, US*, ACM Press, 1998, pp. 39-48.

Bannon, Liam, and Susanne Bødker: “Constructing Common Information Spaces,” in J. A. Hughes at al. (eds.): *The fifth European Conference on Computer Supported Cooperative Work, Lancaster*, Kluwer Academic Publishers, 1997, pp. 81-96.

Bannon, Liam, and Kari Kuutti: “Shifting Perspectives on Organizational Memory: From Storage to Active Remembering,” *HICSS-29*, , 1996, pp. 156-167.

Carstensen, Peter H.: “Towards information exploration support for engineering designers,” in S. Ganesan (ed.): *Advances in Concurrent Engineering — CE97, August 20-22, Oakland, Michigan*, Technomic, 1997, pp. 26-33.

Carstensen, Peter H., and Ulrika Snis: “On Knowledge Management - A Field Study,” in D. G. Schwartz, M. Divitini, and T. Brasethvik (eds.): *Internet-Based Organizational Memory and Knowledge Management*, The Idea Group, 1999, pp. 170-198.

Schmidt, Kjeld, and Liam Bannon: “Taking CSCW Seriously: Supporting Articulation Work,” *CSCW*, vol. 1, no. 1-2, 1992, pp. 7-40.

Star, Susan Leigh: “The Structure of Ill-Structured Solutions: Boundary Objects and Heterogeneous Distributed Problem Solving,” in L. Gasser and M. Huhns (eds.): *Distributed Artificial Intelligence*, vol. 2, Pitman, London, 1989, pp. 37-54.