

# Using Classification in Common Information Spaces

Position paper for workshop on Common Information Spaces  
Copenhagen, Denmark, August 23-25, 2000

Kjeld Schmidt  
Center for Tele-Information  
Technical University of Denmark  
DK-2800 Lyngby, Denmark  
Email: schmidt@intermedia.dtu.dk

Ulrik Christensen  
Center for Tele-Information  
Technical University of Denmark  
DK-2800 Lyngby, Denmark  
Email: ulrik@kampsax.dtu.dk  
(ulrik@ics.uci.edu from September 11)

## 1 Introduction

1. There are, in the world, a large number of cooperative work settings where multiple actors interact through a collection of artifacts of some kind. An actor may for example place a document in a location where others can see it and read it. The actor may place it there for his or her own later use, but at the same time others may access it too. Or the actor may place it there deliberately for the benefit of others. Or the audience is not known yet. Anyway, whatever the reason for putting it there, since others have access to the document, other actors may retrieve and use it at some point in time, perhaps for purposes or in contexts that are quite different from those of the originator. It may also be the case that another actor may change the content of the document, change its location, rename it, delete it, etc. Such settings are ubiquitous in cooperative work.

2. The location of the collection may be a particular shelf, rack, room, or building, or the collection may be distributed over a number of shelves, racks, room, and buildings. The collection may comprise digital artifacts as well as paper and other physical artifacts. The items of the collection may be documents such as books, reprints of scientific papers, technical reports, invoices, design specifications, software code, blueprints, etc. Or the collection may contain parts and sub-assemblies, zoological or geological samples, forensic evidence, etc. Or the collection may be a heterogeneous assembly of such artifacts. However distributed and heterogeneous the collection may be, in their cooperative activities actors will (to some degree) interact and coordinate their activities through this collection, by changing its state.

3. Such collections of artifacts have been termed 'common information spaces' with a rather problematic choice of words. The term 'common' may give the impression that the artifacts in the collection mean the same to all actors and are used in the same manner. The problem arises from using the word 'common' as an adjective to 'information space'. What is meant is that the items in the 'information space' are used 'in common' (that is, in the same sense as 'the commons' in English villages before the Industrial Revolution). The term 'space' may similarly give the impression that the collection is in a room or some other kind of unitary location but using the term to denote a distributed assembly of items makes some sense. The term 'information', however, not only ambiguous and overworked; it is outright misleading, as it gives the impression that a 'common information space' denotes something like a 'shared understanding' or a 'shared culture'. The problem here is that the constitutive role of the artifacts is lost when the term 'information' is used as the key term.

(4) Semantics notwithstanding, the problem is real and important: How do cooperating actors manage to interact through a collection of artifacts in an orderly fashion without (a) access to the effortless awareness and continual recovery from misunderstanding that characterizes face-to-face work settings, and (b) without the agency of a specialized and putatively omniscient and omnipotent central agency such as librarians or a overseeing committee (as in the case of bibliographical databases).

Several issues are involved here:

(4a) The situations of producing the artifact, archiving it, retrieving it, using it, and perhaps modifying it, are (typically) temporally and possibly also spatially separated.

i. Nomenclature: When accessing an artifact produced by somebody else, the actor retrieving it will have to make sense of it. The item thus has to be named so that a user knows what it is. This may involve a convention for naming items (a nomenclature), such as ‘group technology’ in manufacturing.

ii. Standardized form: If the artifact is a written document, the document must be formatted according to a certain standard (e.g., the different standards for formatting scientific papers in the different disciplines, or the conventions for formatting agendas or minutes of meetings). Such standards will, *inter alia*, prescribe the inclusion of various contextual cues such as the identity and background of the author (e.g., affiliation, acknowledgements), the situation in which and the purposes for which the document has been produced, pointers to other documents (e.g., citations and references), etc.

In manufacturing, parts and sub-assemblies, are produced to meet strict specifications as defined by ‘templates’ such as drawings. In geological or zoological collections, samples are treated according to standard procedures of conservation, etc.

iii. When accessing an artifact produced by somebody else, the actor retrieving it will somehow have to assess its relevance, validity, veracity, etc. Actors will here rely on their knowledge of the procedures of origination, that is, the procedures according to which items are originally submitted and selected items (e.g., the review procedures of scientific journals, the acquisition procedures of museums, etc.)

Documents do not necessarily make sense to readers beyond the intended audience. In fact, to make sense of a document that originates in another ‘life world’ is a nontrivial task that involves sophisticated procedures and massive background knowledge (cf. historiography). The ‘organizational memory’ or ‘design rationale’ literature ignores this issue completely.

We do not really know or understand the practices involved in developing, applying, and following conventions of naming, formatting, and assessing artifacts in these collections in and are thus hardly in a position to design support systems for such practices.

(4b) For an actor to be able to contribute an item to the collection in such a way that others — in more or less foreseeable relevant situations — may be able to find it, requires that the collection is organized according to some relatively stable scheme. In simple cases such as the organization of utensils in a family’s kitchen, implicit conventions somehow emerge, supported by the physical constraints of the setting and the utensils, and transgressions can be dealt with rather fluently (and in some situations noisily). In typical cooperative work settings, far more elaborate schemes are required that are inscribed upon some artifact. These schemes offer a classification of the different categories of artifacts in the collection and thus provide competent members

(originators as well as users of the collection) with pointers to the correct physical or virtual location of such artifacts.

The problem of the distributedness is at play here as well:

i. The situations of indexation and retrieval are again temporally and spatially separated. To be able to apply the classification scheme, actors will have to make sense of it, and for actors to perform in an orderly fashion they will have to use the classification scheme relatively consistently. This in turn requires procedures of supervision and maintenance.

ii. The very processes of developing the classification scheme are temporally and spatially separated as well. In simple cases, a specialist (e.g., a librarian, custodian, etc.) may be able to impose a classification scheme. In most cooperative work settings, however, this is not possible. New classes are created or discovered; existing classes are discontinued, or merged, or split into two or more classes. The domain may change too rapidly for an external specialist to maintain the classification scheme, or the domain may be too complex for any external or too heterogeneous for any central agency to be able to cope with it. In such cases, the classification scheme is developed and maintained cooperatively, in a distributed manner.

In order to be able to support the cooperative practices of managing ‘common information spaces’, we need to understand these very practices, that is, the social activities through which classification schemes are negotiated and made to work. We hardly do that as it is.

(4c) The problem of distributedness is recursive, in that it applies to any practical solution to deal with it. That does not make it impossible to deal with it, of course, but it means that each and every solution is local and temporary.

## **2 Empirical investigation**

In the rest of the paper the initial experiences from a field study of the use of shared information structures in a large pharmaceutical company are presented with focus on the role of classification. The investigation is aimed at getting a better understanding of the use of these structures and it thus a small step towards a more useful image of the issues discussed above.

First the research method is briefly commented. Then the experiences are discussed: first the overall organization setting and then six examples on areas of shared information, that each address several interesting aspects of classification use in relation to shared materials.

### **2.1 Investigation method**

As mentioned in the introduction it is believed that the present state of the area demands qualitative studies and that the primary focus should be on the basic mechanisms in use.

The study has so far been based on conversation like interviews with people representing various functions in DM and a small observation based study of the work in a laboratory. This can be seen as a pre-study that allows an organization wide view and allows investigation of the coupling of systems and representations across departments. The study has also served to locate candidates for the coming observation based field study that will get close to the practical everyday work in the organization for an extended period of time, allowing insight not possible with the method used in this part of the study.

The validity of the data has been rendered probable by comparisons with existing studies and with discussing written descriptions of interviews and observations with the sources.

## 2.2 *The organizational setting*

Danish Medical<sup>1</sup>, DM is a large medical company that develops and produces various types of medical products. The work is characterized by being accomplished in a huge organization with cooperation among many actors and groups. The process from idea to having a product on the market takes several years and among others involves such areas as development, clinical tests, production design, production, continuous testing, packing and distribution, government approval of the product, marketing, legal support, documentation and strategic management.

The cost intensive nature of the pharmaceutical industry brings with it a demand for large-scale production. The company thus markets the products globally.

Of central importance to the structure of the production system is the legal structures in the health area. Because of the extent (in number of affected persons and severity) of potential problems of a drug with unforeseen side effects, the area is highly regulated. Errors or inexpedient happenings can well lead to a company having to shut down. This leads to a structure where documentation of the processes and tests of the products are just as important as the product itself. On regular basis inspections of facilities and documentation are made by official institutions and representatives of collaborating partners.

Two systems meet the eye when approaching the information systems in DM. They are spread through most of the organization, are used often and in relation to areas that are essential to the organization, contain very large amounts of data and have separate support units. Furthermore they are introduced and sanctioned by the top-level management.

Using the two major systems and their relation to the other systems as the basis of the investigation is simply a question of choosing an approach that fits the objective of the investigation. It is however to narrow a view to consider only the central systems. In addition to the sharing made possible by these systems data are also shared by other means, some very similar to the central systems and some very different. Not all of these types of sharing are reasonably considered as using CISs, nevertheless it is necessary to include them in the considerations. For instance because work tasks will often involve systems of many types.

The investigation focussed on the use of two central storage facilities (material storage, MS and digital storage, DS), the units maintaining these (material storage unit, MSU, and digital storage unit, DSU), a laboratory unit, LabU, producing and using documents from the stores and a legal unit, LegU, using documents.

---

<sup>1</sup> All names are made anonymous. The used names signal the role that the bearers have in relation to the discussions in this paper.

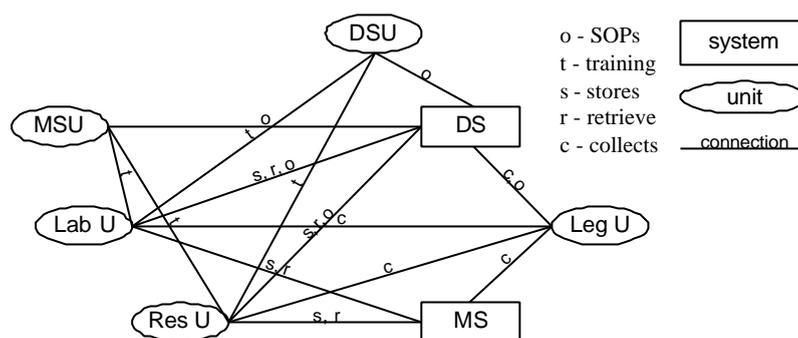


Figure 1: The investigation focused on the use of two central storage facilities, the units maintaining these, a laboratory unit producing and using documents from the stores and a legal unit using documents.

This is of course only a limited selection of the organization and will thus present only part of the needs, strengths and weaknesses of the information structure. It does however present a large part of the types of work done in the organization and represent what can be described as producers as well as users of the shared document bases as well as the units maintaining the supporting systems.

### 2.3 The roles of classification schemes

In this section a few examples of several persons using the same material are presented to point out some interesting aspects of classification in relation to CIS.

#### Study-related material in local archives

In LabU methods for measuring are developed and applied to samples from development and production. The work is done by trained laboratory assistants and chemists working in teams. The work is mainly done individually as part of a collaborative activity and for a laboratory assistant more than ninety percent of the time is used in the laboratory, and the rest of the time is then used to “paper-work”.

Producing documentation is inextricably bound up with producing test results and the work has to adhere strongly to procedures approved by quality departments.

Part of the work is closely related to study plans, which are documents that describe a study rather detailed and is approved by the quality department. In addition to providing ways to refer to work and outcomes and to associate procedures with specific tasks, they serve to divide up the individual work in somewhat self-contained parts and to allocate the work to be done across the department. The study plan does not describe the work to be done fully. Thus a document is made by the person doing the study that describes what is to be done in greater detail (e.g. which quantities to use to make certain solutions, how to set up the samples). This document also serves as a template in which to fill in results as they emerge during the study.

With each study a range of information types are associated.

1. Report, a description of the study results, made by a chemist after the study,
2. Study plan, the initial quality approved description of process and outcome,
3. Raw-data, data from the study, for instance the set up description, prints from laboratory equipment, used quantities,
4. Unforeseen events, any deviations from the prescription in the study plan,
5. Calibrators, documentation of production and approval of some substances,
6. Control, documentation of production and approval of some substances,
7. Correspondence. This classification is specified by quality department and is a result of external general standardization in the area.

These documents are kept in various places during the study and collected at the end of the study, where the report is written. For instance the calibrator documentation is used across studies and is renewed every half year. The documents play several roles in the workplace. For instance the study plan and the personal plan are used to: structure work, remind of where in the process you are, continuously check prescribed procedures and personal calculations and reflections on how to efficiently accomplish the work (quantities, equipment setup), enter results into, staple prints from laboratory equipment onto, check whether the results are correctly transferred to spread sheets. The last use is interesting in that the documents are here used by the laboratory assistant and a colleague who collectively point to and mark the document. The notes in the document are documentation that the data in digital form matches the data from the laboratory. The documents are collected in a study-binder ordered according to above structure as they emerge or become relevant.

The information containers are thus used individually for various purposes in the daily work, they are transferred as a collection between the laboratory assistant and the chemist who writes the report and are put together with the report as an entity that is stored to document the study.

The study plans are often made by the employees in the department themselves. Some degree of re-use is possible. Parts of study plans from similar previous studies can be copied when writing a new one. A local archive of all study plans written earlier is kept on the laboratory assistants PC. They are ordered according to first project number and then study number. Also parts of the personal plans can be re-used. The plans can be re-used by other persons in the department as well.

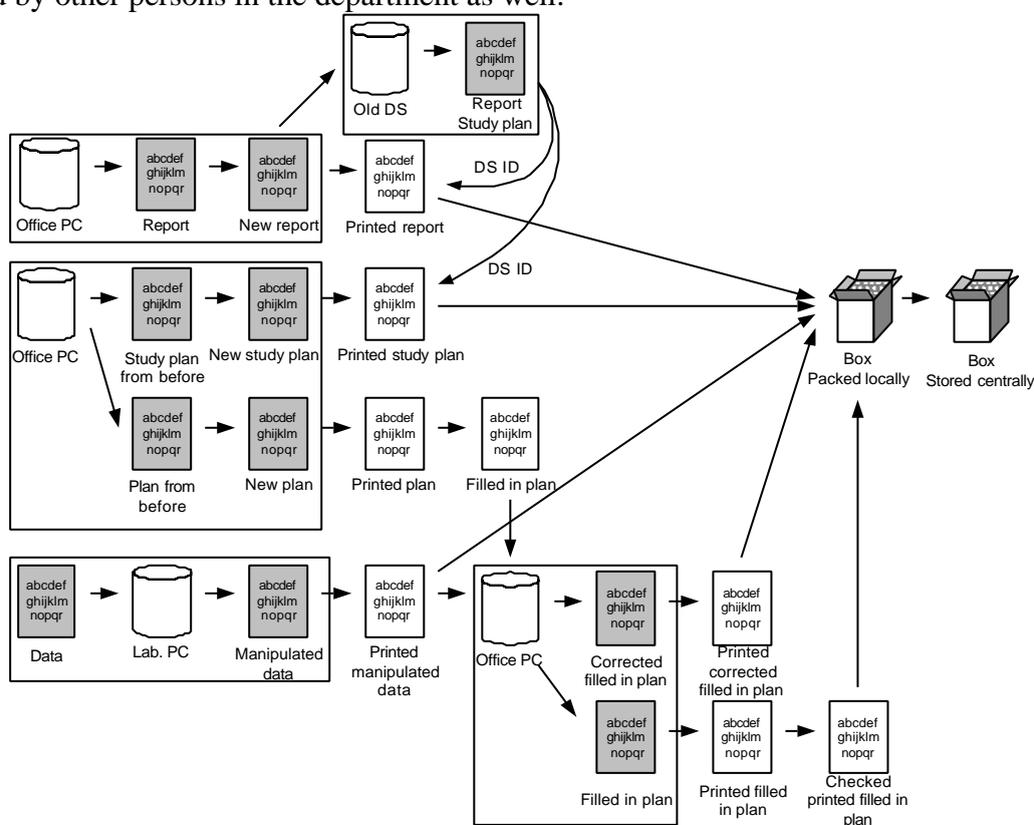


Figure 2: The figure provides an overview of the most important document types and their relations. Grey boxes are digital documents. White boxes are paper-based. Frames means that the documents are found in the system associated with the frame. Arrows indicate that the document itself or information or structure from the document is transferred into the next representation. Cylinders denote digital storage media, e.g. computers or databases.

Several classification schemes are in use to allow the multiple person access described above.

1. Project structure: Each study plans is associated with a project that runs across the organization from development to finished product. Thus a unique project number is associated with the study. A project typically has many study plans (fifty and five hundred are not unusual numbers).
2. Study structure: The study has a study number as well. This is however related to the quality department of the specific part of the organization. Different quality systems are used, for instance between clinical development and manufacturing. Thus the study number is related to the context of the part of the organization only.
3. Structure of material and binders: The types of material that are required as documentation and the way that the binders are ordered according to this allow people to find the relevant information.

Interesting issues:

- A. It is interesting to note how activities are supported by a range of classifications concerning substances, documents, equipment, projects, responsibilities etc.
- B. And that these schemes held by groups or individuals must meet the schemes of other individuals and groups. For instance in relation to the many coupled types of representation related to the production of test results.
- C. The way in which external standardization influences the binder structure is interesting as well.

The documentation is made available for people across the organization in central archives. The next two sections describe the use of first the digital and then the material storage.

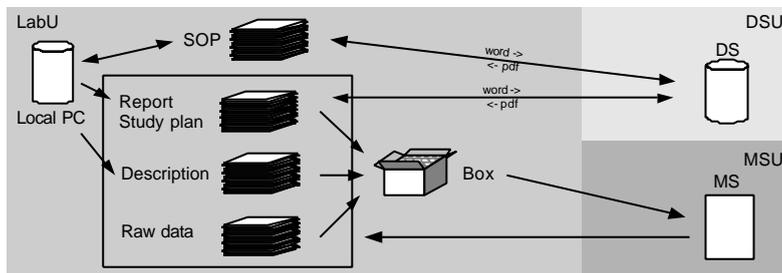


Figure 3: The study related material is kept in a central digital storage and in a central material storage.

### Study-related material in central digital storage

When the study is through and the report is approved and signed the report and the study plan are added to the digital storage. This is done by a capture function in the interface program to DS. The word document is converted to pdf when added to the archives. This is by the employees considered a prevention of copying (copy-paste) documents. The document has a department number, a responsible person, the project number, the study number and a description attached to it and is now accessible from everywhere in DM. The system does provide various access restriction facilities.

This sharing allow people from other departments access to the results and methods of LabU. A set of classification methods is involved. For instance: 1. organizational structure, 2. project numbers, 3. study numbers, 4. descriptions (that are actually characterization and not classification) and 5. knowledge of persons that may be contacted and thus finding information without using classifications are indexation.

Interesting issues:

D. It is interesting to note the potential re-use in other contexts.

E. The way in which finding is made possible through stable organization wide classification.

F. And the fact that other methods are applied in addition to this (e.g. personal contact). Here the importance of the roles of the employees, their mutual relations and their knowledge of each other is central.

### Study-related material in central material storage

Five years ago it was chosen to keep the documentation in a central material storage, MS, to secure that documentation was not destroyed and that it could be accessed at all times.

1. Collecting documentation: The purpose with the MS is to collect all the relevant documentation in one place to allow easy access to needed documentation at all times and that it can be found without having to depend on persons or local archives.

2. Storing documentation in a secure way: The documentation is stored centrally in a secure facility protecting the documents from decay and physical damage as well as from access by unauthorized parties.

3. Providing access to documentation: The employees maintaining the MS mediate access to the documentation and keep a register with index information to be able to find documentation that meets certain description.

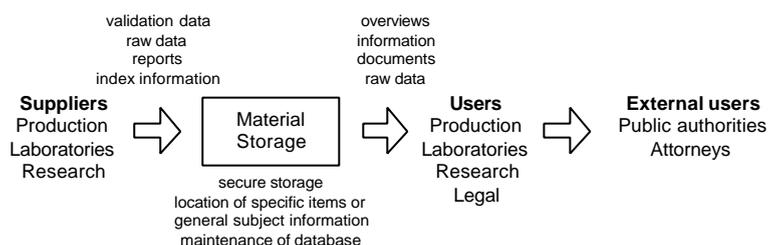


Figure 4: The figure describes part of the function of the material storage unit, inputs, outputs, and suppliers and users of information.

The material in the study-binder is sent to a central facility that protects and mediates access to the data. Especially in relation to the frequent inspections this is a useful structure. When asked for documents the departments order the documents from MS. MSU locates the data, makes copies and sends them to the department.

The documentation is also of use in relation to potential lawsuits where documents going back seventy years or more may be required. Having to keep documents for this long places demands on the storage facility. Humidity, acid free materials and data formats (MS for instance contains data on floppy discs and CDs) are important concerns.

The material is packed in boxes in the producing departments and a standardized form is filled in with information about the content of the box. It is interesting to note that a description of the content is given, but that this is not standardized in any way. The identification number from DS is added to the original by hand before it is sent to MS.

The material is accessed in other contexts than inspections. At times the departments need access to their own old raw data or logbooks (study related material is only a small

part of the documents kept in MS). An especially interesting use is what the MSU calls “knowledge bank”. Here MSU is approached with a request that is not directed towards documents known to exist (or studies known to exist) but is of more general nature. MSU here interprets and seeks information in the archive based on content as opposed to simply providing access as in the above cases. This last type of use is not very frequent but is becoming more widespread.

Again several classification structures are used: 1. organizational structure, 2. project structure, 3. study structure and 4. index information.

It is a deliberate choice not to implement a thesaurus. It is considered too much work and too little benefit (if any at all). In addition it is practically impossible to go back in time and mark the existing material because of the quantities of data.

It is interesting to note that MSU operates with several types of classification that are not completely similar to the classification schemes of the suppliers and other users.

The physically stored material is related to four overall categories. Each type of material is associated with a category of work. The material can be associated with pre-clinical work, clinical work, work related to the department or work related to a product (for instance packaging or labeling the product). Each of these types of material has attributes that allow classification according to various schemes. Some have study numbers, all have department numbers, some have a responsible person, some documents may have related biological/chemical samples that are stored separately and related digital documents that are stored in the shared information system. A distinction is also made between data that refer to specific individuals. According to Danish legislation this type of data can not be kept in a register of the type. Instead it is kept in a separate storage with special access control that meet the legislative demands.

<b>Storage</b>	<b>Work</b>	<b>Origin</b>	<b>Context</b>
Normal	Pre-clinical	Department	Project number
Wet	Clinical	Person	Study number
Digital	Department	Time	Description
Personal reference	Manufacturing		

Figure 5: The material in MS is classified according to several aspects.

The classification scheme used to add material to MS could be seen as an interface. It mediates between the classification used in the producing department and the classification used in MS. LabU produces pre-clinical and department data, but not clinical or manufacturing data and all the raw data is normal, that is not wet data and no personal references. Thus the LabU has one set of classifications describing their field of work and the way it matches with MS. This is only part of the classification used in MSU to order the material received from LabU, even at the same level.

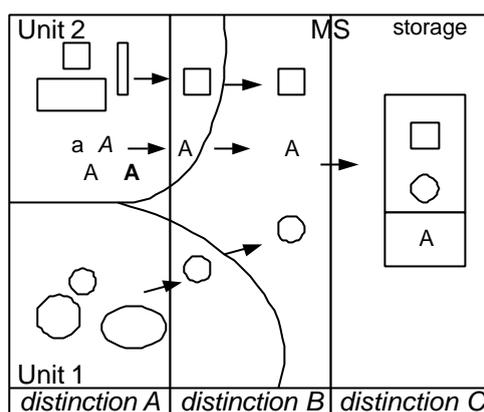


Figure 6: Classification schemes act as a medium between MSU and the documentation producing units.

Interesting issues:

G. It is interesting to note the frequent, unproblematic and very valuable use of central archives with old information.

H. It is also interesting to note that the information is not only of value to the producers but also in relation to similar functions.

I. And that the information is actually beneficially used in contexts that differ widely from the production context.

J. The role that MSU has as librarians is interesting. Activities such as marking information, maintaining indexes and classification filters, keeping overview and keeping order are here very valuable.

K. The role that classification schemes here play in the interface between the storage and the users is a fine example on classifications used as interpreters or mediators.

### Operational documents on department level

DS has two systems, which are actually two interfaces to the same system. The basic structure is used as described above. The newest structure handles operational documents (for instance Standard Operating Procedures, abbreviated SOPs, and Supplier Agreements). This part of the system allows the employees access to all relevant procedures in their valid form. The users thus can by logging on the system immediately check if any changes have been made to current procedures and if any changes are impending.

The operational documents are ordered according to level of applicability. Documents on departmental level are ordered by a numbering system locally controlled and specific for the department. This allows the departments to use already established numbering schemes in the central system. The more general documents are distributed through a mechanism described in the next section, it is not fine-meshed enough to be of use here. SOPs on department level are made in the LabU and approved by the quality department. All of the SOPs are revised at least every third year. Department specific templates are created to help create new SOPs.

The documents are added through an eight step process as described below.

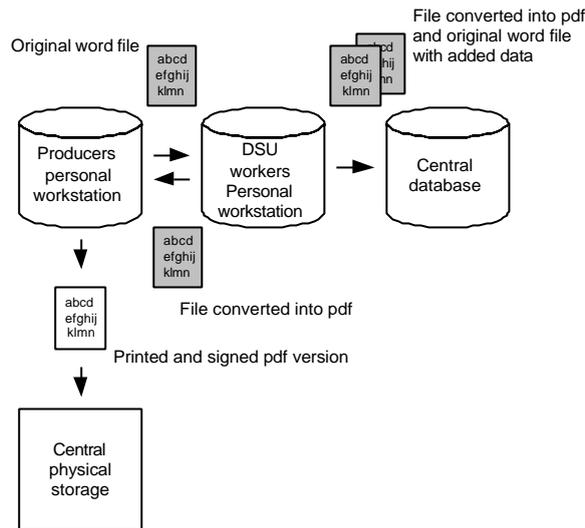


Figure 7: The process of adding SOPs to DS.

1. The document is created, usually as a word document and e-mailed to the DSU
2. It is converted to pdf in DSU and e-mailed back to the author
3. The author prints and signs the file
4. The area of application of the document is determined by the author. A form is filled in and sent with the paper version of the document
5. DSU adds signature date to the pdf version
6. DSU adds the electronic pdf and word files to the old system by “capturing” them and adding the necessary indexation
7. The printed version and the form are archived in DS
8. Correct distribution is verified by the DSU

The documents are kept in a local storage with a large shelf system with a file for each number. Only documents in the new system are kept like this, but both present and obsolete documents are kept there.

ISO has a set of standards that apply for DM. These standards were used to create a document with entries for the types of documentation needed. This was then meshed with demands from Europe, America and Japan to form a table with types of documents (according to subject) and associated “ISO-subject numbers”. In all about hundred and twenty numbers are found. This list serves two purposes. First it is a checklist that all the relevant documentation is produced and secondly it is used to identify documents.

The people making the documents find the numbers themselves. They can ask DSU for help. It is guessed that some of them have “homemade” solutions to the classification problem. For instance lists of often used ISO- subject numbers. A problem in relation to the ISO-subjects numbers is emerging. ISO is preparing a new set of standards and the question now is whether to use the new standard and thus change the system entirely, stick with the current one or make some compromise.

The main beneficiary or the electronic systems are the people responsible for quality assurance. Part of this function is spread throughout the organization and is typically part of the worktasks of an employee. With central, electronic systems it is a lot easier to introduce new procedure specification and to revoke old ones. In the old paper based days this could be very troublesome and involved many local binder archives. On the other hand the system is not necessarily very helpful in the normal work. To some

extend re-use could be done/is done anyway through shared local drives or through the individuals private copies. In the normal work the system can seem difficult to work with, inconvenient and unprofitable. It can be argued that the work of being ones own documentation center has been added.

Interesting issues:

- L. The two roles that documents play as internal work resources as well as documentation to outside parties are clearly illustrated here.
- M. Once again the use of classification in relation to finding documents is established.
- N. And again the importance of external standardization and the problems with being bound to a scheme not fully controlled are evident as well as the use of using such schemes.
- O. The old discussion about who benefits from using an arrangement, who has to do the work and thus the reward structure is once again shown to be of importance.
- P. The way in which the classification, in addition to allowing documents to be found by the right people, also provides a checklist that allows easy overview of the documents and thus of anything is missing, points to an important use of classification.

### Operational documents on cross-functional and top level

The cross-functional level and top level documents are mainly created in departments further up the management hierarchy. Proposals for documents can have an origin in the LabU. This type of documents usually go through a series of hearings on their way from draft to final version.

SOPs on cross-functional level or top level are distributed with the aid of a distribution profile associated with each document. The documents distribution profile is "filtered" through a set of department profiles so that the documents matching the department profile appears on the monitor when login onto the system with the name of the department as filter. The department filter is maintained by a dedicated administrator. This person is responsible of ensuring that the department profile is up to date. In addition to this the administrator can remove specific documents from the set that appear even though it meets the requirements. This activity must be cleared with the person responsible for quality in the department.

The distribution profile is a set of attributes that makes it possible to decide where it applies. These attributes are determined and entered into the system as the SOP is added. The general documents are distributed through a set of attributes provided by filling in a form with 104 check boxes divided in four categories as shown below.

Document no.: [ ]			
Made by:	Site: [ ] Department: [ ] Person: [ ]		
Prepared by:	Date: [ ] Initials: [ ] Signature: [ ]		
Entered by:	Date: [ ] Initials: [ ] Signature: [ ]		
Function	Product	Organization	Site/Unit
[ ] all	[ ] all	[ ] all	[ ] all
[ ] product dev.	[ ] raw materials	[ ] corp. staff	[ ] discovery
[ ] logistics	[ ] packaging	[ ] dev.	[ ] clinical dev.
[ ] GLP	[ ] product type I	[ ] discovery	[ ] foreign site I
[ ] marketing	[ ] product type II	[ ] quality	[ ] foreign site II
<i>in all 55</i>	<i>in all 13</i>	<i>in all 10</i>	<i>in all 26</i>

Figure 8: Operational documents are distributed by use of a "distribution profile" and corresponding department profiles. The square brackets show the fields to fill in.

In the table above some examples on the elements of each group are given. The “All” category are found in all of the groups. The form with the distribution profile also has fields for the document number in the system, the site and department where the document comes from as well as the name of the responsible person. It also has fields for the persons handling the SOP in DS where they can enter date, initials and sign for having handled the document.

The printed table of properties is sent in with the documents. Often as a suggestion for change, which must be approved by persons responsible for quality.

This arrangement makes it easy to add new departments. Once a profile is made for the department the right documents automatically appear. It is however difficult to add new classes and problems also occur when areas/departments change. This could make it necessary to reclassify the old documents.

The systems are between two and three years old and no major changes have been made till now. A function/process was added to the distribution profile. This area is harder to see through than for instance a site. It is more difficult to see where the documents “pop up”.

The classification scheme in the form is made by representatives from several departments. Various needs and perspectives led to a difficult decision process. For instance it was discussed how many categories should be used. Some found it important to limit the number of categories to ease the use of the system, especially in the introduction period and for the users who rarely adds documents, whereas other meant that very elaborate schemes would be valuable tools for their work. “Everyone think that what they do is special.” Sometimes these lists reflected the way that people were ordering their documents in the offices. Similar discussions were related to the allowed document types and their associated meta-data fields. A central issue here was that the larger the set of categories the larger the chance that people will not be able to categorize things consistently. The DSU worked with an estimate that more than between seven and ten options will make it hard for the users to choose consistently. It is acknowledged that categories overlap and that exact categories can be difficult to define. All of the categories are controlled by the DSU who are able to add and change the categories. New categories can be added by any person adding documents. It is however mainly done by DSU personnel.

The documents in the system with distribution profiles are among other things related to the sites or units they belong to. Some problems have arisen with a recent de-merger in DM. Some units are split up, some are joined and some units now refer to other units than before. Some units even use the same name as before even though the functionality has changed. This makes it impossible to see the line of command. This situation has led the DSU to consider the possibilities of categorizing the documents according to functions instead of to departments.

The impending changes are the first “real” update. The de-merger brings with it a need to adjust the profiles of existing documents and of the departments to meet the new category structure. The way in which the problems will be approached is not resolved as yet. DSU are working on ways to adapt to the new organizational structure. The administrators should be informed. The changes are proposed at hearings. It is then thought that the administrators should print a list of documents just before the change

and hold it up to the list after the change and in that way adjust the profiles so that only the entire set of relevant documents are distributed to the departments. It is still not clear how the potential reclassification of the documents should be done. It is at present not unrealistic to do it but it is a “tough job” and not made any easier by the necessary update of the paper version of the distribution profile lying with the document.

Interesting issues:

Q. Finding documents could be seen as bringing the right documents together with the right people. That is push as well as pull methods can be beneficially applied as we see with the way that SOPs are distributed. Configurable filtering mechanisms can be very useful tools in relation to shared information structures.

R. The coupling of the document classification to the organizational structure is central S. as are the problems that the classification meets when changes occur. Considerations of the use of coupling the scheme to functions instead of department codes or numbers may make the structures more stable, but this is far from a cure-all solution.

T. The origin and negotiations in relation to creation (and adaptation) of a classification scheme point to very important issues: various perspectives and uses, power struggles and negotiations, the coupling to the scheme to various organizational and other institutional factors, the difficulties of keeping schemes stable and the problems that arise if this succeeds.

### **Classification in the Legal Unit**

The legal unit in DM has several subdivisions. One of these, Legal unit, LegU, was studied. LegU is responsible for providing the attorneys with internal documents for use in American lawsuits. The function is a result of an increased presence on the American market, where the number of documents considered of importance in relation to a lawsuit is much larger than in Europe. This difference can be in the order of between four thousand and twenty thousand documents in a British case against between one hundred thousand and four hundred thousand documents in an American.

In LegU a distinction between documentation and documents is stressed. Documents are paper, electronic documents, video and the likes, that is “the *physical* the company has got”. Whereas documentation is related to the content of the documents or the information they contain, the things that LegU has to find are not interpreted. Another interesting aspect is found; an employee describes the documents of interest as follows: “We do not know what we are after. Someone has got to tell us what we are looking for.”

Central to the document procurement is a “document request”. This is a list of documents from the attorneys. In case of DM being sued this list is made by the suing party. The documents are described in various ways ranging from very specific descriptions to general characteristics (e.g. organizational charts, production methods, results from research and test, expected payoffs, anything related to knowledge of a specific subject).



procedures can be retrieved from DS. The systems are not seen as central sources of documents they are on the contrary almost invisible. It should be noted that this is considering documents only - not information. The documents found in the system may and may not be central to a case.

The procurement process may be made somewhat easier by having all documents in electronic form and this may be strengthened if they are kept centrally, but these two things in themselves will not lead to a change of the nature of the process. A range of reasons is found for the central archives not being sufficient to meet the demands of LegU.

1. Not knowing what documents are needed: The employees in LegU as a rule do not know what documents they are looking for. It is seen as an advantage that as many people as possible think the situation over. In that way it is more likely that all of the relevant documents are found. So LegU will inevitably visit people even though they can find the documents in a central storage. It is a lot easier to simply get the documents when they are there, than to get them from the central archives later. Here it should be noted that the present cases are patent cases that typically demand recent documents. It is expected that a product liability case will lead to stronger use of MS, because these cases go further back in time and thus may require documents that are not kept in the local archives but can only be found in MS.
2. Needing all versions of documents: All versions of a document must be obtained. If an employee has made a couple of marks or written something in a printed version this has to go into the documentation pool.
3. Some material is not kept in archives: The MS do not contain certain types of information (e.g. minutes, documents stored on local computers, floppy discs, e-mails, handwritten notes).
4. Getting an overview: For the work of the LegU it is central to be able to get an impression of the associated documents. This is possible in the local archives where the documents usually are kept with related documents in binder structures. To some degree this may be met by the way that related documents in different envelopes are sometimes placed in the same box in MS.
5. Documents can not be marked sufficiently: It is probably impossible to mark all data in relation to the potential later searches. For instance in early research it is not known which product the research will relate to. And the product may be the key with which the LegU has to identify relevant documents.

Interesting issues:

- U. An example on the use of documents across organizational structure is found here, that stress the potential use of documents in very different contexts.
- V. The above described reasons, for the information structures not being useful, contain material for lengthy discussions of the use of these structures, their limitations, their forces and the ways in which to improve the current available technologies.
- W. The importance of history - of documents, organizations, departments, persons related to documents - clearly appears from the situation the LegU faces.

### **3 Classification as a tool for working with CIS**

The empirical investigation offer material for a wealth of considerations on the nature of common or shared information and the use of classification in relation to this as well as in general. Currently only a few of these issues are investigated in any depth in the project. This chapter contains some of these general considerations of the use of classification as a tool for working with CIS.

It has been shown that classification schemes are part of several types of useful functionality in CISs (social-technical structures not limited to computer systems). They allow:

- distribution, push
- finding specific data, pull
- finding data related to a subject - overview
- statistic impressions of an area - overview
- certainty that things are as they should be - overview
- re-use - the structure of data collections and documents are known and it is possible to have a good idea of what the creator of the data meant, because of the place the data is put contains information of the data, support interpretation
- interaction, interface (to MS), output standardization that allow cooperation - packing is done in a way that makes storing and indexation easy/possible
- in addition sorting things (in accordance with a scheme) may be a valuable activity in itself (for instance in relation to storing in the classes normal, wet and referring to individuals)

In addition it was noted that:

- schemes are growing wild and that they are used layered and meshed in relation to the various areas of shared material. Figure 10 on the last page shows a list with some examples on shared resources and some of the schemes that are applied
- the shared information and the schemes are heterogeneous and involve several levels
- historical layers are of importance both in relation to evolution of schemes and in relation to use of information - as the studied systems are quite young not many stories exist of these problems but the current challenges related to the adaptation to the new organizational structure will be interesting to follow in relation to this
- the importance of the surrounding structures in shape of projects, departmental organization, production types, sites, ISO etc. is fundamental to understanding classification
- in some areas DSU and MSU has a role as mediators. They teach about using the system, the classification and MSU even helps people with finding non-specific information.

More generally classification is a valuable tool in various task types:

- It is a natural part of human cognition
- It allows overview (for instance used to ascertain that things are as they should be, sort of a checklist)
- It can be a valuable activity in itself
- It allows finding and re-finding (with push as well as pull methods, specific items as well as related to something)
- It allows delegation of responsibilities
- It allows communication and interaction, standardization
- It allows re-use (by creator as well as other persons)

#### **4 Conclusions**

A field study has provided some insight into the ways in which information structures are constructed, used and maintained and the role that classification in various shapes plays in relation to this. The heterogeneous nature of shared information and the no less heterogeneous nature of classification, may be what makes these structures the useful tools they are. They however also result in a vast and complex area of investigation and have thus provided us with an interesting and hopefully profitable challenge. A challenge that to some degree may be made more manageable by choosing different focus areas for several strands of research.

#### **Acknowledgements**

The work was part of the distributed multimedia project financed by Danish Center for Distributed Multimedia.

## References

- Ackerman, 1998; Ackerman, Mark S. and Christine Halverson; Considering an organization's memory; in Computer Supported Cooperative Work, Proceedings of the ACM 1998 conference on Computer Supported Cooperative Work, November 14 - 18, 1998, Seattle, WA USA, pp. 39-48
- Bannon, 1997; Bannon, Liam and Susanne Bødker; Constructing common information spaces; in W. Prinz, T. Rodden, J. Hughes, and K. Schmidt (eds.), Proceedings of ECSCW '97, Sept. 7-11, Lancaster, England, Kluwer Academic Publishers, Dordrecht, 1997, pp. 81-96
- Bowker, 1999; Bowker, Geoffrey C and Susan Leigh Star; Sorting things out: classification and its consequences; MIT press, 1999
- Carstensen, 1999; Carstensen, Peter H., Kjeld Schmidt and Uffe Kock Wiil; Supporting shop floor intelligence: A CSCW approach to production planning and control in flexible manufacturing; in Group'99: International Conference on Supporting Group Work, Phoenix, Arizona, 14-17 November 1999, ACM, New York 1999, pp. 111-120
- Celentano, 1991; Celentano, A. M. G. Fugini and S. Pozzi; Classification and Retrieval of Documents Using Office Organization Knowledge; in Proceedings of the Conference on Organizational Computing Systems, ACM, 1991, pp. 159-164
- Dourish, 1999b; Dourish, Paul, Richard Bentley, Rachel Jones and Allan MacLean; Getting some perspective: Using process descriptions to index document history; in Conference on Supporting Group Work, Proceedings of the international ACM SIGGROUP conference on Supporting group work, November 14 - 17, 1999, Phoenix, AZ USA, pp. 375-384
- Grinter, 1996; Grinter, Rebecca E.; Supporting Articulation Work Using Software Configuration Management Systems; in Computer Supported Cooperative Work: The Journal of Collaborative Computing 5: 447-465, 1996, Kluwer Academic Publishers, 1996, pp. 447-465
- Heath, 1996; Heath, Christian and Paul Luff; Documents and professional practice: "bad" organisational reasons for "good" clinical records; in Computer Supported Cooperative Work, Proceedings of the ACM 1996 conference on Computer Supported Cooperative Work, November 16 - 20, 1996, Boston, USA, pp. 354-363
- Hertzum, 1999; Hertzum, Morten; Six roles of documents in professionals' work; in ECSCW '99 - Proceedings of The Sixth European Conference on Computer Supported Cooperative Work, 12-16 September 1999, Copenhagen, Denmark, Kluwer Academic Publishers, 1999, pp. 41-60
- Lakoff, 1987; Lakoff, Georg; Women, fire and dangerous things; The University of Chicago Press, 1987
- Mark, 2000; Mark, Gloria; Conventions and Commitments in Distributed groups; to be published in Computer Supported Cooperative Work (CSCW) an International Journal
- Marshall, 1998; Catherine C. Marshall; Making Metadata: A study of metadata creation for a mixed physical-digital collection; in I. Witten, R Akscyn and Frank M. Shipman (eds.), Proceedings of digital Libraries '98. Pittsburgh, PA, June 23-26, ACM Press, 1998, pp. 162-171
- Orlikowski, 1992; Orlikowski, Wanda J.; Learning from NOTES: Organizational Issues in Groupware Implementation; in J. Turner and R. Kraut (eds.) CSCW '92. Proceedings of the Conference on Computer-Supported Cooperative Work, Portland, Oregon, September 26-28, 1988, ACM Press, New York, 1992, pp. 362-369
- Rouncefield, 1994; Mark Rouncefield, John A. Hughes, Tom Rodden and Stephen Viller; Working with "constant interruption": CSCW and the small office; in Computer Supported Cooperative Work, Proceedings of the ACM 1994 conference on Computer Supported Cooperative Work, October 22 - 26, 1994, Chapel Hill, USA, pp. 275-286
- Schmidt, 1992; Schmidt, Kjeld and Liam Bannon; Taking CSCW seriously: Supporting articulation work; in Computer Supported Cooperative Work (CSCW), An International Journal, vol. 1, no. 1-2, 1992, pp. 7-40
- Schmidt, 1997; Schmidt, Kjeld; Of maps and scripts: The status of formal constructs in cooperative work; in S. C. Hayne and W. Prinz (eds.): GROUP'97, Proceedings of the ACM SIGGROUP Conference on Supporting Group Work, Phoenix, Arizona, 16-19 November 1997, S. C. Hayne and W. Prinz (eds.), ACM Press, New York, N.Y., 1997, pp. 138-147
- Simone, 1999; Simone, Carla, Gloria Mark and Dario Giubbilei; Interoperability as a Means of Articulation Work; In D. Georgakopoulos, W. Prinz, and A. Wolf (eds.) Proceedings of ACM Conference on Work Activities Coordination and Collaboration (WACC'99), ACM Press, Feb. 22-25, 1999, San Francisco, pp. 39-48
- Star, 1994; Star, Susan Leigh and Karen Ruhleder; Steps Towards an Ecology of Infrastructure - Complex Problems in Design and Access for Large-Scale Collaborative systems; in CSCW '94 Proceedings of the Conference on Computer Supported Cooperative Work, Chapel Hill, North Carolina, Oct. New York: ACM press, 1994, pp. 253-264
- Suchman, 1987; Suchman, Lucy A.; Plans and Situated Actions: The Problem of Human-Machine Communication; Cambridge University Press, Cambridge, 1987
- Trigg, 1999; Trigg, Randall H., Jeanette Blomberg and Lucy Suchman; Moving document collections online: The evolution of a shared repository; in ECSCW '99 - Proceedings of The Sixth European Conference on Computer Supported Cooperative Work, 12-16 September 1999, Copenhagen, Denmark, Kluwer Academic Publishers, 1999, pp. 331-350
- Van House, 1998; Van House, Nancy A., Mark H. Butler and Lisa R. Schiff; Cooperative knowledge work and practices of trust: sharing environmental planning data sets; in Computer Supported Cooperative Work, Proceedings of the ACM 1998 conference on Computer Supported Cooperative Work, November 14 - 18, 1998, Seattle, WA USA, pp. 335-344
- Wittgenstein, 1998; Wittgenstein, Ludwig; Philosophical Investigations; Blackwell Publishers, 1998

Material	Within functional area				Across functional areas	
	Within department		Across departments		use	schemes *
	use	schemes	use	schemes		
<i>Kept in local stores</i>						
samples and substances	+	chemical knowledge, labeling	+	chemical knowledge	-	-
personal notes	+	various personal classification	-	-	l	org. and project structure
articles, books	+	various personal classification	~	?	-	-
minutes	+	various personal classification	~	?	l	org. and project structure
logbooks	+	binder structure	-	-	il	org. and project structure
manuals	+	binder structure ?	~	?	il	org. and project structure
raw data	+	data types, binders	-	-	il	org. and project structure
study binder	+	data types, binder structure	-	-	il	org. and project structure
study report PC	+	project structure, folders	-	-	il	org. and project structure
<i>Kept in central digital storage</i>						
manuals	+	org. and project structure	+	org. and project structure, index, types of documents	il£	org. and project structure
study report	~	org. and project structure	+	org. and project structure, index, types of documents	s£	types of material, types of use
item specific	+	product or process structure	+	product or process structure	+lis£	product or process structure
department level SOP	+	#	~	#	si£	types of material, types of use
cross-functional level SOP	+	#	+	#	+lis£	# + types of material, types of use
top level SOP	+	#	+	#	+lis£	# + types of material, types of use
<i>Kept in central material storage</i>						
logbooks	~	org. and project structure, index, types of documents	-	-	ils\$	org. and project structure, types of material, types of use
raw data	+	org. and project structure, index, types of documents	+	org. and project structure, index, types of documents	ils\$	org. and project structure, types of material, types of use
study binder	+	org. and project structure, index, types of documents	+	org. and project structure, index, types of documents	ils\$	org. and project structure, types of material, types of use

Signature: + : used, ~ : unlikely, ? : unknown, - : not used, s : statistics, i : inspection, l : legal.

The most often found use is marked with bold types.

\* : In case of inspection or legal document request the use also requires that the local users are able to locate the material and thus involves their classification tools as well.

# : Schemes: organization structure, ISO structure, read/ unread, in effect/ obsolete/ to be updated/ impending, types of work, roles and competencies

\$ : The MS access the material and use special classification.

Figure 10: The table shows a few of the shared information-containing artefacts and some of the classification schemes used in relation to them. The primary use of the table is an illustration of the large amount of objects and the way that the many, heterogeneous classification schemes overlap and find use in relation to several types of information artefacts.