

S. Bødker, M. Kyng, and K. Schmidt (eds.), *ECSCW'99: Proceedings of the Sixth European conference on Computer Supported Cooperative Work* (Copenhagen, Denmark, September 12-16, 1999), pp. 41-60. Kluwer, Dordrecht.  
Preprint version

# Six Roles of Documents in Professionals' Work

Morten Hertzum

Centre for Human-Machine Interaction, Risø National Laboratory, Denmark

Documents are used extensively by professionals in their execution of their own work and to share information with others. Professionals use and manage their documents in ways that are woven into their work activities and leave most of the context unsaid because the documents are understood as belonging to a certain ongoing activity. Contrary to this, organisations have a strong interest in storing information in less person-dependent ways than simply relying on their employees' memory and personal files. To support document management effectively we need to balance the individual professionals' focus on their current activities against the long-term interests of the organisation, and we need a fuller understanding of the affordances and constraints of documents. This study identifies six roles documents play in professionals' work, namely that documents serve: (1) as personal work files, (2) as reminders of things to do, (3) to share information with some yet withhold it from others, (4) to convey meaning, (5) to generate new meaning, and (6) to mediate contacts among people. Painstakingly standardised and very time-consuming methods are required for documents to convey meaning but such efforts are rarely considered worthwhile compared to relying on other document roles or rework.

## 1. Introduction

Documents are used extensively by professionals in their execution of their own work and as a means of sharing information with others. The ways professionals use and manage documents for their own purposes are woven into their work activities and have been studied by researchers interested in how people organise

their individual information spaces (e.g., Malone, 1983). However, most work is co-operative and furthermore organisations have a large interest in storing information in less person-dependent ways than simply relying on the memory and personal files of their employees. The use of documents for information sharing has for example been studied by Bannon & Bødker (1997), Harper & Sellen (1995), and Star & Griesemer (1989), but at least with respect to paper documents little research has looked systematically at their role in organisations (Sellen & Harper, 1997). During the past 30 years computers have been assigned a key role in various efforts to support document management, but these efforts have time and time again failed to produce the expected results. It seems as if the role of documents in professionals' work has yet to be properly understood.

This study identifies six roles documents play in professionals' work. Professionals are subject specialists characterised by putting to work their intellectual skill learned in systematic education and through experience, and they are to a large extent paid to organise their own work, make sense of things, and pass judgements. The starting point in identifying the six document roles has been that documents are part of the context in which they are produced. This introduces a crucial distinction between subsequent use of the documents in this context and (re)use of them in other contexts (see Figure 1). The former use of documents include such situations as documents written for personal use or use within a project group, whereas the latter is exceedingly common in settings where the involved professionals are geographically dispersed or involved in projects that outlast their own involvement in them.

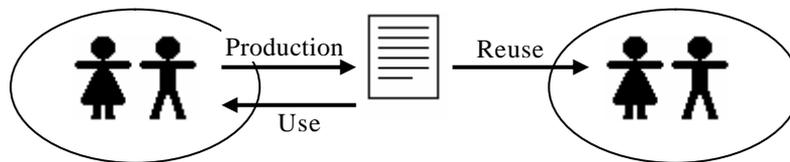


Figure 1. The general setting for production, use, and reuse of documents.

Document archives, whether personal or corporate, may serve many purposes, including accountability, operational continuity, planning, legal evidence, disaster recovery, research, and corporate history. Two types of document value can be distinguished in this connection. The evidential value of a document is its value in providing evidence of an organisation's structure, procedures, transactions and the like, whereas the informational value of a document is the value of its contents for reference, contemplation, and research (Bikson & Frinking, 1993). This study is concerned with the informational value of documents. Their evidential value is however important too, as illustrated by the central role of documents in proving ownership of ideas in patent applications.

It has been estimated that professionals spend 25% of their time distributing,

filing, and retrieving documents (Gordon, 1997). This study is intended to expand our understanding of *document management* and inform the design of systems that support this pervasive activity. The study also aims at contributing to the elaboration of the concept of *common information spaces* (CISs) as developed by Bannon & Bødker (1997), Bannon & Schmidt (1989), Schmidt & Bannon (1992). The CIS concept has mostly been discussed in connection with co-located, co-present persons such as air traffic controllers, but it is also intended to inform discussions of situations where people are distributed in time and space. This study investigates the work involved in creating and maintaining CISs and people's inclination to do it.

The next section outlines how documents enter into professionals' work at three levels, which differ markedly with respect to intensity, principles for organising the documents, and intentions of information sharing. This section also gives an introduction to the concept of common information spaces. Section 3 is about how the context, frozenness, and permanence of documents affect their ability to communicate meaning. Section 4 discusses how the document roles directed toward the professionals' individual information spaces and those directed toward information sharing can be brought together in the design of document management systems. Section 5 concludes the paper by summarising the six roles of documents in professionals' work. Please note that I make no claims as to the exhaustiveness of the six roles.

## 2. Individual and common information spaces

Several studies have found that professionals interact with three levels of information in their files (e.g., Cole, 1982; Hertzum, 1993): *Action information* which includes documents readily at hand and often piled up on the desk, *personal work files* which are within reach but usually put on shelves or held in other conventional filing devices, and *archive storage* which comprises information stored away from the office. Action information is to a large extent arranged on the basis of spatial clues, which require frequent interaction to stay functional. Personal work files are loosely systematised, and though some people maintain personal indexes to aid retrieval many people rely on their memory of the documents' location. Personal work files exist primarily to provide the individual professional with convenient access to the material, not to make it available to others. With respect to archive storage, Cole (1982) notes that the information items are rarely dealt with and when they are it is almost exclusively through extensive category structures. The change from a spatial, loosely systematised, memory-based organisation to category structures as the documents move from 'action' to 'archive' fundamentally changes the nature of the activities carried out to manage the documents. This change is also one from an individual to a common information space.

## 2.1 Individual information spaces

Professionals know the documents they keep in their offices, hence retrieval of these documents means re-locating them. The documents are typically organised according to their relation to the professionals' past and present work, so retrieval essentially means recalling this relation. Many professionals utilise this known-universe situation to emphasise retrieval over filing in that they minimise the up-front time spent on filing and base their ability to locate their documents on their memory of the relation between their documents and their work activities. These professionals are willing to spend more time on retrieval since retrieval is immediately valuable – it seems more worth the effort to spend five minutes retrieving a document than to spend five minutes filing it. Other professionals prefer to spend up-front time on filing to keep their items organised and reduce retrieval costs (Berlin et al., 1993). These professionals maintain orderly offices to make the known universe more efficient, and they may extend it with category structures that provide access to documents through for example authors, projects, or keywords.

Jahoda et al. (1966) found that 46 (61%) of their respondents maintained a personal index with category structures that provided access to their documents through one or more access points. The respondents reported a number of shortcomings of these personal indexes, the three major ones being: too time-consuming to prepare, inconsistencies in indexing, and not enough access points. These shortcomings reflect difficulties creating and maintaining the necessary category structures within the available time limits. Case (1991) found that 12 (60%) of his respondents maintained some kind of card file to index their documents, but very few of these indexes were like library catalogues in their purpose or exhaustiveness. The indexes included, for example, only the documents relevant to a project currently underway. Several of the respondents had tried to develop an index for their entire document collection but gradually abandoned the attempt. Similarly, Hertzum (1993) found that attempts to inventory everything or just papers from periodicals were not carried through or not considered at all. The respondents were somewhat frustrated about the ad hoc way many of their documents were organised but either overall structures were considered inappropriate or the overhead involved in creating and maintaining them was considered too big.

Rather than category structures consisting of document attributes such as author and subject, Kwasnik (1991) found that the factors people take into account in classifying their documents consist to a very important degree of situational factors such as the use to which the document is to be put. This reflects the short term, personal, and work-oriented motivation that underlies the organisation of professionals' individual information spaces. Probably, the most far-reaching of the situational factors is that much of the information on people's desks is there to remind them to do something, not just to be available when they

look for it. This way the documents on the desk are action information in a sense beyond being drawn upon as information sources in the course of the professional's activities: They play an active role in the management of the professional's work. The profound differences between category structures and the ways professionals organise their documents can be vividly illustrated by examples of organising principles used in individual information spaces:

- Keeping the material for this week's classes on the floor, where it would be impossible to overlook (Case, 1991).
- The top document in a pile, i.e. the immediately visible one, has a special status as pile representative (Hertzum, 1993).
- 'On the top shelf are books that are very seldom used' (Kwasnik, 1991).

Most documents are written on a computer but to a considerable extent computers are just used to produce documents, not to file them. The electronic version is stored for later elaboration, correction, and reuse but the authoritative copy is the paper copy on the shelves. One reason for this is that professionals often find it desirable to store their own documents together with documents available in paper copy only, for instance to collect correspondence pertaining to a project in one folder. Barreau (1995) and Barreau & Nardi (1995) find that people's behaviour in organising their electronic documents is consistent with the behaviour observed in organising physical offices. Thus, people prefer filing by location because it supports finding as well as serves a crucial reminding function, and they file documents according to the dictates and vagaries of their work because in the end carefully architected logical schemes do not yield enough value. It is however unknown whether the observed commonalities reflect genuine preferences of the studied people or stem from the desktop metaphor of the studied systems, a metaphor which relies heavily on a direct mapping to physical offices (Fertig et al., 1996).

## 2.2 Common information spaces

A key capability of documents is to facilitate the sharing of information among professionals who are not present at the same time or in the same place. Whereas oral communication is ephemeral and requires that the actors are co-present, the permanent nature of documents suggests that if professionals document their work in writing then organisations will retain the professionals' knowledge when the professionals retire, move to other jobs, or otherwise leave the organisation. However, empirical studies indicate that people tend to look for a person to ask rather than a document to read (Carstensen, 1997; Pinelli et al., 1993). It is often claimed, or tacitly assumed, that this state of affairs is the result of unsatisfactory documentation and that there is an urgent need for better documentation and, consequently, for improving the tools and practices used in documenting work (see, for example, Blair, 1996). In many cases these arguments convey the

impression that good documentation will make organisations far less dependent on the memory and continued presence of their employees. In this study we stress the context-dependent nature of written as well as oral communication, and how the frozenness of documents makes the context inherently underspecified.

Central to document sharing is the creation and maintenance of an archive or a database containing the shared information items. While archives and shared databases are effective at making documents accessible within a community or organisation it is crucial to note that shared access to documents does not imply that the meaning and implications of these documents are available in any complete or unambiguous way. Schmidt & Bannon (1992) emphasise that co-operative work ‘requires the active construction by the participants of a common information space where the meanings of the shared objects are debated and resolved, at least locally and temporarily.’ To share an information space involves that the local actors interpret the shared information items – make sense of them. The sum of the information items and the locally constructed, shared agreement about their meaning is termed a common information space or CIS (Bannon & Bødker, 1997; Bannon & Schmidt, 1989; Schmidt & Bannon, 1992). CISs come into existence only when a shared agreement can be reached, which is surely not always the case, and they cover just the points where the actors’ individual perspectives come together. In the most coherent exposition of the CIS concept to date, Bannon & Bødker (1997) stress that CISs are not confined to situations where people are co-located and co-present, it also includes situations where people are working separately and the main connection between them is a shared database:

Most discussions of shared spaces in CSCW have tended to confine themselves to situations in ‘real-time’, or near real-time. Our conceptualization of CISs however extends to situations where information is entered into a database at one point in time and subsequently accessed by others, perhaps months or even years later. In what sense can we characterize this situation as a CIS? In our view, the reason is because both the producer and the receiver consciously make an effort to understand each other’s context – of production and use, so that even though the efforts may be distributed over time and space, there is a form of communication, of “putting in common”, going on in such activity.

Situations where people are co-present, co-located, and working together differ quite a lot from those where CISs are distributed in time and space. This span of the CIS concept is one of its attractions but also necessitates studies of CISs in a range of contexts to prevent the concept from becoming biased toward some situations and consequently describe others inaccurately. Tightly coupled co-operation such as air traffic control and loosely coupled co-operation such as communicating through documents in a shared database can be seen as extremes with respect to the nature of the CIS.

A major difference between tightly and loosely coupled co-operation is that in loosely coupled co-operation people must more explicitly attempt to include aspects of the context with the information items in an effort to ensure that in a

future use situation others will be able to deduce their intended meaning. Bannon & Bødker (1997) use the term *packaging* to denote this effort to put information items in common, i.e. to extend the stored items with some explanation of their context. Conversely, the reader's effort to recreate the context and get the intent of the message is denoted *unpacking*. This effort is as demanding and partial as the one that goes into packaging. Thus, packaging and unpacking are meant 'to draw attention to the myriad of ways in which people struggle to make sense of each other' (Bannon & Bødker, 1997) – they do not imply a smooth, perfect process.

### 2.3 Professionals' inclination to package and unpack

Several studies report considerable disuse or underutilisation of archived documents (e.g., Harper & Sellen, 1995; Kidd, 1994; Mintzberg, 1975). This may be due to practical circumstances such as lack of critical mass or slow retrieval facilities but other, more principal causes may also enter into it. Nardi & Barreau (1997) argue that old information is not, in general, useful information and ask what someone would do with all the old information even if they could find it quickly and easily. While they acknowledge that there are situations where old information is essential they argue that large-scale archiving is often promoted without a clear notion of what it should achieve. The lack of clarity regarding the purpose of such archives arises chiefly from vague relations to the primary work performed in the organisations. As long as the document archive is an appendix to the primary work, rather than a contributing part of it, many people will experience it as more or less pointless (Waters & Nagelhout, 1995).

Packaging also requires that the professionals suspend their normal way of looking at and working with their documents to take an outsider's look at them. This is, however, difficult because the individual professional has an inherently incomplete sense of whether his/her documents will eventually be of interest to someone else and, if so, to whom and in what context. Furthermore this outsider's look is to some extent even unpleasant for the professional. It becomes unpleasant because activities such as selecting keywords and setting the security level of a document form a detached view, which does not adequately reflect the professionals' own understanding of their documents or forces the professionals to state things, such as the importance of a document, on which they are not yet sure. Neither of these circumstances provide professionals with a strong inclination to carefully package their documents, rather the combination of these circumstances tends to create a situation where packaging is perceived as a tedious, low-priority task.

In summary, two roles of documents can be identified in the professionals' individual information spaces. Documents serve as *personal work files* to provide

the professionals with easy access to the documents they need in their current work. The essential aspect of this role of the professionals' documents is that the documents are collected by an individual professional and organised according to their relationship to this professional's primary work. Strict category structures are not required, and apparently seldom used, to organise documents in this role. As a consequence shared access to personal files must normally be facilitated by the professional collecting and maintaining the file because only this professional knows its organisation and contents. Moreover, a request for a document may involve querying not only the files but also their collector's expertise (Blomberg et al., 1996).

Documents also serve as *reminders of things to do*. Here the essential aspect is that documents play an active role in the professionals' management of their work tasks – the documents are not just passively available. Malone (1983) makes this point succinctly clear when he notes that 'a primary reason for placing tasks on the desktop in the first place is so that intentional search does not have to be relied upon.' Documents generally fulfil this role through their spatial location and visibility, rather than through their contents or an elaborate indexing scheme. This is evident in the organisation of individual professionals' offices and in the co-ordination of co-operative work, for instance when the documents pertaining to a task follow the person currently responsible for the task. In general, documents used in co-ordinating co-operative work serve to remind the involved actors of the co-ordination mechanisms that structure their work and the ways in which they are supposed to do or document their work (Schmidt & Simone, 1996).

In relation to professionals' use of document archives to share information across time and space boundaries the role of documents is less clear. On the one hand documents are an inadequate carrier of meaning in that packaging and unpacking are required but cannot make the full context available to the reader. On the other hand it is often claimed that documents are underutilised in that better documentation and more time spent reading documents should be a cost-effective strategy. For example, Repo (1987) reports that a group of energy researchers saved an average of just under \$1300 for every report they read. These savings are the estimated value of avoiding repeated investigations, but it is not clear whether the acquired information was extracted directly from the reports or, for example, provided by persons whom the reader contacted as a result of reading the reports. To better understand the roles that documents play in professionals' information sharing the next section introduces the notion that documents contain multiple 'voices'.

### 3. Text and context

Human communication is often conceptualised in terms of the transmission of information. This transmission model involves the encoding of an idea into a

signal by a sender, the transmission of this signal to a receiver, and the decoding of the signal into a message by the receiver (Wertsch, 1993). In their account of how documents can mediate common information spaces Bannon & Bødker (1997) suggest a very similar conceptualisation: The writer must package the information in documents which are then transferred to the reader who must unpack them to make sense of them. This view of communication is problematic because it tends to imply that documents have a single correct interpretation, which it is the reader's task to extract. While it is true that the wording of a document is frozen at the time the author completes it, the meaning of the document is not frozen. Documents are monologic conversations in the sense that the writer remains unaware of the concrete questions and intentions that cause subsequent readers to examine a document, but in terms of their meaning documents are dialogic.

### 3.1 The Multivoicedness of documents

The dialogical nature of spoken as well as written utterances has been studied by Bakhtin (1981, 1986), a Soviet philosopher and semiotician. Dialogicality, the basic theoretical construct in Bakhtin's approach, concerns how one speaker's concrete utterances are a compound of her/his own voice and the voices of others. As Bakhtin (1986) says, 'the utterance is filled with dialogic overtones.' These overtones are carried by the speaker's utterance but they need in no way be related to the speaker's voice, i.e. to what the speaker is trying to communicate. For example, one professional's argumentation may gain additional credibility from being phrased in ways that carry well-esteemed scientific overtones, while another professional's argumentation may be blurred by his inadvertent use of a phrase that conveys the voice of a recent episode in a comedy series. Wertsch (1993) summarises Bakhtin's ideas about dialogicality:

A shorthand way of formulating Bakhtin's ideas about dialogicality for a sociocultural approach to mind is to pose a fundamental Bakhtinian question about forms of semiotic mediation: "Who is doing the talking?" From a Bakhtinian perspective, the answer will always be: "At least two voices."

The concept of dialogicality brings it to the fore that the context of an utterance is not a largely passive background against which the utterance is made. On the contrary, the context is actively present in the individual utterance as overtones of the speaker's voice.

In the case of clinic records, Garfinkel (1967) has studied how such records intertwine (1) a voice reporting who did what to the patients with (2) another voice whereby clinics demonstrate that they have honoured claims for adequate medical care. This involves a delicate balance between detail and intentional ambiguity. On the one hand, the records serve as an essential tool for the clinic personnel in their day-to-day work. On the other hand, for the records to make sense the reader must be able to correctly interpret the utterances and omissions

made with regard to the possibility that the records may have to portray the transactions with patients as having been in accordance with accepted medico-legal practice. That is, to make use of the records the reader must be able to tell who is speaking, and this requires detailed knowledge of the domain and of established local practices, i.e. of the context. By intertwining these two voices the clinic records allow for the appropriate use of informational materials by competent readers, and competent readers only, both in normal day-to-day practice and in unknown, future situations. Clement & Wagner (1995) give a similar account of how information is shared with some and at the same time deliberately withheld from others.

Wertsch (1993) distinguishes between the use of documents to convey meaning adequately and to generate new meaning. For a document to *convey meaning* the reader must be familiar with the context of the writer and thus able to distinguish the voice of the writer from the overtones induced by the context. Such agreement on the code used in preparing the document and that used in interpreting it presupposes some kind of standardisation of the employed language. This is clearly illustrated in Harper & Sellen's study of the International Monetary Fund. Harper & Sellen (1995) found that data in the statistical database were shared by a number of people because all data in this database were known to be derived from standard methods. As a consequence it sometimes took years for figures to be approved and added to the database, and there were numerous omissions in it since data were left out if they required any judgement to determine vagueness or inconsistency. Thus, data that did not adhere to the no-judgement rule governing the statistical database were not added to it, and the effort required to determine whether data items belonged in the database was by no means trivial – it could extend over several years. In contrast, collaboratively written reports were not shared, though facilities to do so were available. When economists were working on their own data, the data were unsuited for sharing and general use because they had not yet been through the social processes of validation and assessment. When the data had been through these processes, only the authors of the report were able to know when the data were usable, since only the authors could tell judgement from hard fact. Thus, the reports were of limited value to others because the data in them were either of unknown validity or inseparable from the judgements that interpreted them.

Harper & Sellen (1995) find that information items have to be packaged according to painstakingly standardised methods for people to consider them reusable, otherwise documents are considered useful to their authors only. The standardised methods are required to enable the words of a document to provide a passive link between the writer and the reader. Latour (1986) terms documents with this characteristic immutable mobiles to emphasise that they can be transported over long distances and convey unchanging information. In the absence of painstakingly standardised methods it becomes crucial to distinguish

between the wording of documents and their meaning, but this distinction is often blurred. For example, Levy (1994) argues that documents are both fixed and fluid but he restricts his discussion to the wording of the documents and seems to assume that the meaning remains fixed as long as the wording is not changed.

When documents serve to *generate new meaning* there is an inherent tension between the writer's world and the reader's world. This tension indicates that the document functions as a 'thinking device' rather than a passive link between the writer and the reader (Wertsch, 1993). In reading a document this way the reader brings a question or incoherence to the document and listens for a voice that will contribute to make coherent sense of his/her world. This involves that the reader enters into a dialogue with the text. On the one hand this dialogue can take many directions since documents are heterogeneous objects and readers have varying interests, on the other hand the wording of the document is frozen and severely bounds the possible dialogues. In this role documents are what Star & Griesemer (1989) term boundary objects, that is objects that are both adaptable to different viewpoints and robust enough to maintain identity across them.

Extensive use of condensed forms of communication, which leave most of the context unsaid because the document will be understood by the primary readers as belonging to a certain ongoing activity, preserves resources during document production but reduces the ability of documents to maintain identity across contexts. To make documents understandable to people who are not familiar with the context the condensed forms of communication must be elaborated, often to the exasperation of the primary readers who can see the elaboration as redundant (Brown & Duguid, 1996). When professionals are in a hurry or simply absorbed in their day-to-day work they are likely to document their work to support their own sense-making process; they are much less inclined to spend time expanding their writings into documents understandable to unknown future readers.

### 3.2 Document reuse versus depth of understanding

A radical way to circumvent the limitations of documents with respect to conveying information is to abandon information sharing altogether and do the work anew. Reading often gets a frustrating experience, which does not seem worth the effort because the professional is left in doubt about some aspects of the document or discovers that it only partially provides the needed information. Doing the work anew may be more work but it is experienced as more satisfying because the professional is in control and gets the understanding of the topic that comes out of doing the work herself/himself.

In connection with systems development Naur (1985) explicitly argues that it should be seriously considered to build new versions of existing applications from scratch if the upgrade is to be made by persons who have not been deeply involved in the writing of the current version. 'Such a procedure is more likely to

produce a viable program than program revival, and at no higher, and possibly lower, cost' (Naur, 1985). Reuse is of questionable value in these cases because the program code and documentation cannot provide the necessary understanding of the system. Naur describes systems design as a process of theory building, a theory being an all-compassing, coherent understanding of the way a selected part of the real world is handled in a specific computer system. This theory is neither contained in the program code nor in the documentation but built gradually by the developer through his/her active involvement with the program code, the documentation, and their relation to selected aspects of the real world. A person having the theory is able to tell whether and how potential modifications of the system can be implemented as new elements fitting coherently into the structure and idea of the system, whereas a person who is not in possession of the theory is unable to preserve its coherence in face of modifications.

A very similar account of professionals' sense-making is given by Perby (1987) who studied meteorologists making local weather forecasts in an airport. The meteorologists maintained that taking over ready-made forecasts would degrade the quality of their briefing of the pilots since the process of making the forecasts was necessary in building the inner weather picture that provided the foundation of their work. What the meteorologists defended was an active assimilation of the continuous stream of information about various, evolving weather elements instead of passively receiving a lot of data. The crucial quality of this active assimilation was that it guaranteed a certain depth in the individual meteorologist's interpretation of the information. Again, doing the work anew has advantages over reusing documents produced by others – in a sense the choice between rework and document reuse resembles the choice between project-based and lecture-based teaching.

It seems that many efforts to emphasise documentation and document management have seriously overestimated the ability of documents to convey meaning or underestimated the amount of work required to package and unpack information contained in documents. The apparently ever-increasing storage capacity of computers has made it feasible to store literally every document, email, fax, letter, memo, minute, note, report etc. produced or received in an organisation. Such an archive would provide access to the text of the documents, but many archives have been influenced by a much more far-reaching idea of a vast information repository giving everyone within the organisation immediate access to the accumulated knowledge of past and present employees (see Ackerman, 1996). This amounts to considering filing documents an alternative to being informed by them (Kidd, 1994). The result is systems that are capable of storing masses of documents but largely fail to support professionals in reaching the understanding they need to perform their work. These systems focus on preserving a record of the professionals' past activities and thereby give priority to the evidential value of the documents as opposed to their informational value.

It seems warranted to suggest a stronger focus on the informational value, in terms of document management systems that aim more directly at the professionals' present activities.

In between reuse and rework professionals rely on oral communication with informed colleagues, though this involves both interrupting them in their work and succeeding in presenting the question in a way that triggers their attention and gets them constructively involved. While the wording of documents is frozen and thus does not adapt to the readers, oral communication affords mutuality and enables the actors to adjust to each other. Questions often lead to queries about the situation that gave rise to the question to provide the person being asked with a more solid basis for interpreting and answering the question. In the course of the conversation the person asking a question may encounter that it misses the real problem and rephrase it or ask additional questions. Also, the person asking the question will be interested in some background information about the experiences that gave rise to the answer in order to assess its reliability. Thus, neither the question nor the answer exists beforehand, both are products of the communication process. A number of studies have found that personal contacts are an essential source of information, which is often preferred to seeking written information (e.g., Pinelli et al., 1993).

While some of the persons professionals contact are part of their personal network documents play an important role in *mediating contacts among people* who do not know each other beforehand. This sixth role of documents emphasises that an important aspect of a document is to record that the author is a potential source of information about the document subject (Hertzum, 1993). Since documents often report from co-operative projects other people apart from the author(s) may also know something about the subject. Therefore information about where a document has been used is also of potential value – who participated in the project where the document was created, who were the document circulated to upon completion, etc. In paper archives some of this information is often stored haphazardly, for example information about who initially received a copy of a document may be handwritten on the cover page in the form of an instruction to the secretary who did the copying and mailing. In electronic archives such information must be explicitly captured to remain available. Another implication for document management systems is that they should provide users with an easy way to obtain current contact information for authors and other people mentioned in the documents, as opposed to contact information that was valid at the time the document was created.

In summary, the analysis has identified four roles documents play in professionals' information sharing. Documents serve to *share information with some yet withhold it from others*. These politics of sharing and withholding become necessary because the permanence of documents means that the writer

cannot know who will later obtain a document and what they will want to use it for. Thus, the very ability of documents to support information sharing makes it necessary for writers to take precautions against unintended sharing. This role of documents is not restricted to situations where documents may later be used as evidence about a course of events but has to be considered whenever ownership of information is important to a professional's position, power, or privileges.

Documents prepared according to painstakingly standardised methods can also serve to *convey meaning*. Standardised methods are required to reduce the multivoicedness of a document to a level sufficiently low to allow reading without continuously engaging in interpretation to establish the meaning of the document. While conveying meaning adequately is the ultimate type of information sharing the most important message regarding this document role is that documents rarely play it. Writers refrain from the required standardisation to avoid being overloaded, and readers may find that documents do not apply uninterpreted anyway because the world has changed since the document was written. In this sense documents change by virtue of staying the same while time passes and situations change.

Documents that fail to convey meaning often serve to *generate new meaning*. However, such documents may mistakenly be believed to convey meaning because interpretation and personal judgement are so integral to professional work that they easily go unnoticed. When documents serve to generate new meaning the reader unpacks the document to make sense of it in her/his own context. The basic condition necessitating this unpacking is that the meaning of a document is 'hiding in the light' of its words. While the text of a document is tangible, sharable, and obviously visible its informational contents has to be brought out through the reader's active building of an understanding not conveyed in the text.

Writers have to significantly prune their dialogue with their readers to fit it into documents and this pruning leaves only weak traces of what the writers could have stated but did not. To circumvent this limitation documents frequently serve to *mediate contacts among people*. In this role a document is not read to provide the reader with the needed information directly but to determine whether the document can be used as a directory of people who may possess the needed information.

#### 4. A Janus-faced approach to document management

While several of the six document roles complement each other and can be supported side by side, there seems to be a conflict between individual and common information spaces. From the individual professionals' perspective the document management that goes on in their offices is part of their primary work and hardly experienced as a separate activity, whereas the document management

involved in putting documents in common is often experienced as a low-priority appendix to their work. Contrary to this, corporate efforts to improve document management practices have often included strong requests to abolish personal collections, which are seen as a threat to the authority and up-to-dateness of the shared archive. These two perspectives on document management are to a certain extent mutually exclusive in that adopting one perspective tends to make people blind toward the other.

Memex (Bush, 1945) and Dynabook (Kay & Goldberg, 1977) are seminal examples of visions about how computers could support individuals' document management. These systems present the collection as a browsable 'landscape' of inter-linked multimedia documents, but they also take the user to be an independent actor, i.e. not embedded in an organisational setting. Numerous commercial systems have attempted to canonise an organisational view on document management (Hertzum, 1995). These systems have introduced archiving requirements that are dependent upon the active co-operation of the professionals but have failed to provide any support for the professionals' personal document management.

It seems that most work in this area has focused on either the personal or the organisational part of document management. This study hypothesises that it should be a central consideration in the design of document management systems to strike the balance between supporting the individual professional in managing his/her personal documents and supporting information sharing among professionals. To achieve this it is suggested that systems be conceived of as two separate but interfaced parts: one directed toward the individual professional, the other toward the needs of the organisation. The interface between the two parts specifies what the individual professional must put in common and how this 'up-loading' is done as well as what the professional can 'down-load' from the corporate part of the system. Such a design brings personal and organisational document management together but maintains a boundary between them, a property with at least two advantages: (1) By accentuating that document management has two faces the boundary helps avoid designs that neglect the differences between them or are unduly biased toward one or the other. (2) Since the information that must be up-loaded is singled out the professionals will know when they use the system for their own benefit and when they fulfil their obligation toward the organisation.

One way to conceptualise an archive is as a number of intertwined, personal threads, each thread being the chronologically ordered sequence of documents written or received by an individual professional. This emphasises the personalised nature of the information and the possibilities for using documents to mediate contacts among professionals. The threads can be visualised as personal timelines where documents or clusters of documents constitute events in the person's corporate history (see Lansdale & Edmonds, 1992; Plaisant et al., 1996).

Such a visualisation retains some context in the sense that documents appear close to other roughly simultaneous events such as other documents or calendar notes. This enables document retrieval via searches for the context in which a document has been used. An additional feature of timelines is an element of automated document management in that old documents are gradually moved out of the user's immediate view as time passes (Fertig et al., 1996). Timelines belong in the personal part of a document management system because the only thing that adds useful structure to the chronological sequence of events is the searcher's memory of the temporal relationships between past events. Information sharing requires a different visualisation of the system, one that focuses on providing searchers with an impression of what they can expect to find.

Document management must strike a balance where it inflicts minimal inconvenience on the individual professional and yet ensures the quality of the shared archive. To keep packaging manageable and avoid information overload on the part of the receivers it is probably reasonable to package selected documents only. It is, in other words, counterproductive to aim at a level of documentation that makes everything that has been done once available for reuse. In connection with the production of documents – a pervasive activity in the environments of most professionals – it may even be beneficial to opt for reuse of fragments of document text rather than for preservation and reuse of experience, insights, and other kinds of information. Carefully made documents often contain well-structured explanations of complex concepts or meticulously thought-through formulations of critical issues such as contractual terms and legal obligations. Blomberg et al. (1996) describe how the personal files of an attorney in a large law firm were used in this way, i.e. as templates for new documents, by himself and his colleagues. Shared access to the documents was provided through personal contact to the attorney who also contributed comments on the specific qualities of the documents. Thus document fragments were frequently reused but their meaning was established through discussion among colleagues, not through strict, corporate-wide category structures. This may indicate that document management systems should supplement the facilities that take a whole document as the entity with facilities for handling document fragments (Levy, 1993). However, document fragments are exactly like whole documents in that they too have multiple voices unless made according to painstakingly standardised methods, and therefore reuse of document fragments is subject to essentially the same obstacles as reuse of whole documents.

## 5. Conclusion

Professionals have made extensive use of documents for centuries but the role of documents in professionals' work is still not well understood, as evidenced by the meagre results of the past 30 years of work on providing effective computer

support for document management. This study has investigated the ways professionals use and manage documents with special emphasis on the interactions between the use of documents in the individual professionals' execution of their current activities and the use of documents for information sharing among professionals who are separated in time or space. This investigation has identified six roles documents play in professionals' work:

- *As personal work files.* Documents serve in this role to be readily available to the individual professional and through her/him to colleagues within the organisation. Generally, such documents are organised according to their relation to the professional's work, rather than according to category structures.
- *As reminders of things to do.* Through their spatial organisation documents in this role serve a crucial function in the professional's management of his/her work activities. Also, such documents are kept visible to cue the professionals back into a line of thinking they have been engaged in but not yet finished.
- *To share information with some, yet withhold it from others.* Access to information may affect the distribution of power and privileges in an organisation. As a precaution against unintended sharing many documents are intentionally ambiguous and thereby understandable to competent readers only.
- *To convey meaning.* Painstakingly standardised methods are required to reduce the multivoicedness of a document sufficiently for it to convey meaning. In practice, this substantial effort is rarely considered worth it compared to doing the work anew or relying on people's ability to actively build a coherent interpretation of the document – i.e. generate new meaning.
- *To generate new meaning.* Rather than relying on the writer to produce a document that can be moved to other contexts and provide unchanged information, it is usually left to the reader to reinterpret the document in her/his own context. In this role, what the writer provides is a thinking device to be used by readers in their making sense of their world.
- *To mediate contacts among people.* Since documents are a restricted form of communication and reading is a quite slow process, readers frequently use interesting documents as pointers to people. This way documents serve to direct readers in their search for people with specific competencies.

Organisations have an obvious interest in storing information in less person-dependent ways than simply relying on the memory and personal files of their employees. However, using written documentation to store and share information involves two time-consuming activities, packaging and unpacking, and seems to result in a reduced depth of understanding compared to doing the work oneself. Furthermore, the sole purpose of packaging is to support sharing; for those involved first hand packaging is superfluous and in fact tends to interfere with efficient execution of their ongoing work. As a result packaging and unpacking

shall be applied selectively, that is in situations where actual time savings are achieved and these time savings outweigh the costs in terms of reduced depth of understanding.

This study hypothesises that document management systems should avoid canonising either the professionals' individual information spaces or the use of documents to support the creation and maintenance of common information spaces. Both perspectives seem necessary to understand the dynamics of document management and to strike a balance that avoids overloading the professionals in their current work and yet achieves effective information sharing. However, the relative importance of the six document roles depends on the concrete circumstances and must thus be established through analysis of the work domain. Science differs from engineering, safety-critical domains from transaction-processing ones, bureaucratic institutions from entrepreneurship etc., and this calls for future work into the concrete implications of the document roles in different domains of professional work.

## Acknowledgements

This work was supported by a grant from the Danish National Research Foundation. The major part of this study was done while I was at the University of Limerick in the context of the EU Training & Mobility of Researchers project COTCOS. My stay there was co-funded by a grant from Christian and Ottilia Brorson's trust for the endowment of young scientists. I wish to thank the people in the Interaction Design Centre, University of Limerick, for their support.

## References

- Ackerman, M. S. (1996): 'Definitional and contextual issues in organizational and group memories', *Information Technology & People*, 9, 1, pp. 10-24.
- Bakhtin, M. M. (1981): *The dialogic imagination: four essays by M. M. Bakhtin* (edited by M. Holquist), University of Texas Press, Austin.
- Bakhtin, M. M. (1986): *Speech genres and other late essays* (edited by C. Emerson and M. Holquist), University of Texas Press, Austin.
- Bannon, L. and Bødker, S. (1997): 'Constructing common information spaces', in J. Hughes, T. Rodden, W. Prinz and K. Schmidt (eds.): *ECSCW'97: Proceedings of the 5th European Conference on CSCW*, Kluwer, Dordrecht.
- Bannon, L. and Schmidt, K. (1989): 'CSCW: four characters in search of a context', in: *EC-CSCW'89: Proceedings of the First European Conference on CSCW*. Reprinted in J. M. Bowers and S. D. Benford (eds.): *Studies in Computer Supported Cooperative Work: Theory, Practice and Design*, North-Holland, Amsterdam, 1991, pp. 3-16.
- Barreau, D. K. (1995): 'Context as a factor in personal information management systems', *Journal of the American Society for Information Science*, 46, 5, pp. 327-339.
- Barreau, D. and Nardi, B. A. (1995): 'Finding and reminding', *ACM SIGCHI Bulletin*, 27, 3, pp. 39-43.

- Berlin, L. M., Jeffries, R., O'Day, V. L., Paepcke, A. and Wharton, C. (1993): 'Where did you put it? Issues in the design and use of a group memory', in: *Proceedings of the ACM/IFIP INTERCHI'93 Conference*, ACM Press, New York, pp. 23-30.
- Bikson, T. K. and Frinking, E. J. (1993): *Preserving the present: Toward viable electronic records*, Sdu Publishers, The Hague.
- Blair, D. C. (1996): 'STAIRS redux: Thoughts on the STAIRS evaluation, ten years after', *Journal of the American Society for Information Science*, 47, 1, pp. 4-22.
- Blomberg, J., Suchman, L. and Trigg, R. H. (1996): 'Reflections on a work-oriented design project', *Human-Computer Interaction*, 11, pp. 237-265.
- Brown, J. S. and Duguid, P. (1996): 'The social life of documents', *First Monday*, 1, 1, <http://www.firstmonday.dk/issues/issue1/documents/index.html>.
- Bush, V. (1945): 'As we may think', *Atlantic Monthly*, 176, 1, pp. 101-108.
- Carstensen, P. H. (1997): 'Towards information exploration support for engineering designers', in S. Ganesan (ed.): *Advances in Concurrent Engineering - CE97*, Technomic, Lancaster, Pennsylvania, pp. 26-33.
- Case, D. O. (1991): 'Conceptual organization and retrieval of text by historians: The role of memory and metaphor', *Journal of the American Society for Information Science*, 42, 9, pp. 657-668.
- Clement, A. and Wagner, I. (1995): 'Fragmented exchange: Disarticulation and the need for regionalized communication spaces', in H. Marmolin, Y. Sundblad and K. Schmidt (eds.): *ECSCW'95: Proceedings of the 4th European Conference on CSCW*, Kluwer, Dordrecht, pp. 33-49.
- Cole, I. (1982): 'Human aspects of office filing: Implications for the electronic office', in: *Proceedings of the Human Factors Society 26th Annual Meeting*, Human Factors Society, Santa Monica, pp. 59-63.
- Fertig, S., Freeman, E. and Gelernter, D. (1996): "'Finding and reminding" reconsidered', *ACM SIGCHI Bulletin*, 28, 1, pp. 66-69.
- Garfinkel, H. (1967): "'Good" organizational reasons for "bad" clinic records', in H. Garfinkel: *Studies in Ethnomethodology*, Prentice Hall, Englewood Cliffs, NJ, pp. 186-207.
- Gordon, M. D. (1997): 'It's 10 a.m. Do you know where your documents are? The nature and scope of information retrieval problems in business', *Information Processing & Management*, 33, 1, pp. 107-121.
- Harper, R. and Sellen, A. (1995): 'Collaborative tools and the practicalities of professional work at the International Monetary Fund', in: *Proceedings of the ACM CHI'95 Conference*. ACM Press, New York, pp. 122-129.
- Hertzum, M. (1993): 'Information retrieval in a work setting: A case study of the documentation part of chemists' work', in J. P. Bansler, K. Bødker, F. Kensing, J. Nørbjerg and J. Pries-Heje (eds.): *Proceedings of the 16th IRIS. Information Systems Research Seminar in Scandinavia*, DIKU report 93/16, University of Copenhagen, Copenhagen, pp. 786-798.
- Hertzum, M. (1995): 'Computer support for document management in the Danish central government', *Information Infrastructure and Policy*, 4, 2, pp. 107-129.
- Jahoda, G., Hutchins, R. D. and Galford, R. R. (1966): 'Characteristics and use of personal indexes maintained by scientists and engineers in one university', *American Documentation*, 17, 2, pp. 71-75.
- Kay, A. and Goldberg, A. (1977): 'Personal dynamic media', *IEEE Computer*, 10, 3, pp. 31-41.
- Kidd, A. (1994): 'The marks are on the knowledge worker', in: *Proceedings of the ACM CHI'94 Conference*, ACM Press, New York, pp. 186-191.

- Kwasnik, B. H. (1991): 'The importance of factors that are not document attributes in the organisation of personal documents', *Journal of Documentation*, 47, 4, pp. 389-398.
- Lansdale, M. and Edmonds, E. (1992): 'Using memory for events in the design of personal filing systems', *International Journal of Man-Machine Studies*, 36, pp. 97-126.
- Latour, B. (1986): 'Visualization and cognition: Thinking with eyes and hands', *Knowledge and Society: Studies in the Sociology of Culture Past and Present*, 6, pp. 1-40.
- Levy, D. M. (1993): 'Document reuse and document systems', *Electronic Publishing*, 6, 4, pp. 339-348.
- Levy, D. M. (1994): 'Fixed or fluid? Document stability and new media', in: *ECHT'94: Proceedings of the European Conference on Hypertext*, ACM Press, New York, pp. 24-31.
- Malone, T. W. (1983): 'How do people organize their desks? Implications for the design of office information systems', *ACM Transactions on Office Information Systems*, 1, 1, pp. 99-112.
- Mintzberg, H. (1975): 'The manager's job: Folklore and fact', *Harvard Business Review*, 53, 4, pp. 49-61.
- Nardi, B. A. and Barreau, D. (1997): "'Finding and reminding" revisited: Appropriate metaphors for file organization at the desktop', *ACM SIGCHI Bulletin*, 29, 1, pp. 76-78.
- Naur, P. (1985): 'Programming as theory building', *Microprocessing & Microprogramming*, 15, 5, pp. 253-261.
- Perby, M.-L. (1987): 'Computerization and the skill in local weather forecasting', in G. Bjerknes, P. Ehn and M. Kyng (eds.): *Computers and Democracy*, Avebury, Aldershot, pp. 213-229.
- Pinelli, T. E., Bishop, A. P., Barclay, R. O. and Kennedy, J. M. (1993): 'The information-seeking behavior of engineers', in A. Kent and C. M. Hall (eds.): *Encyclopedia of Library and Information Science*, vol. 52, supplement 15, Marcel Dekker, New York, pp. 167-201.
- Plaisant, C., Milash, B., Rose, A., Widoff, S. and Shneiderman, B. (1996): 'Lifelines: Visualizing personal histories', in: *Proceedings of the ACM CHI'96 Conference*, ACM Press, New York, pp. 221-227.
- Repo, A. J. (1987): 'Economics of information', in M. E. Williams (ed.): *Annual Review of Information Science and Technology*, vol. 22, Elsevier Science Publishers, Amsterdam, pp. 3-36.
- Schmidt, K. and Bannon, L. (1992): 'Taking CSCW seriously - Supporting articulation work', *Computer Supported Cooperative Work (CSCW)*, 1, pp. 7-40.
- Schmidt, K. and Simone, C. (1996): 'Coordination mechanisms: Towards a conceptual foundation of CSCW system design', *Computer Supported Cooperative Work: The Journal of Collaborative Computing*, 5, pp. 155-200.
- Sellen, A. and Harper, R. (1997): 'Paper as an analytic resource for the design of new technologies', in: *Proceedings of the ACM CHI'97 Conference*, ACM Press, New York.
- Star, S. L. and Griesemer, J. R. (1989): 'Institutional ecology, 'translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39', *Social Studies of Science*, 19, pp. 387-420.
- Waters, P. M. H. and Nagelhout, H. (1995): 'Revolution in records: A strategy for information resources management and records management', *American Archivist*, 58, 1, pp. 74-83.
- Wertsch, J. V. (1993): *Voices of the mind*, Harvard University Press, Cambridge, MA.