Jean-Pierre POITOU
poitou@aixup.univ-aix.fr

EDITING KNOWLEDGE WITH „ ATELIER FX “.

Introduction

The present paper has two main points : the first is to give an anthropological overview of organizational memory –or rather the *technological base* of the corporation, as well as of the corporate knowledge management ; the second is to draw the implications of this approach for the development of an assistance system to the collective management of knowledge in the firm. The two topics are articulated around the concept of *place.* This notion, introduced by Pêcheux (1969), is central to both the present author's sociological conception of knowledge distribution and management within the firm, on the one hand, and to the linguistic fundamentals of the ATELIER FX computer based tool-box on the other hand. According to Pêcheux, formulating as well as understanding a statement or a text, are contingent upon the individuals' *place* in the social division of labor. In line with Pêcheux's work on discourse analysis, it will be argued here, that the exploitation of *situated knowledge* should rest upon the actor himself. Accordingly, in order to satisfy this requirement, computer assistance to corporate knowledge management should consist neither in formalizing, nor in modeling, but rather in *editing* knowledge.

In the first part, a conception of knowledge and cognitive functioning will be introduced, germane to the "situativity theories" (Greeno & Moore, 1993), yet more deeply grounded in a sociological (and rather European) tradition. To that end, the notion of "intellectual object" suggested by Janet (1935) is proposed in order to include in an extended concept of documentation, not only verbal statements and graphic documents, but also tools and equipment as well. In the second part, organizational memory is considered in relation with the social division of labor. The concept of place is introduced and its implications exposed with regard to the construction of meaning on the one hand, and to the management of knowledge, individual as well as collective, on the other hand. This latter point is further elaborated in the third part. Two examples of a collective knowledge management system (CKMS) are briefly presented in the fourth part : DIADEME and SG2C. Finally, in the fifth part, ATELIER FX, the computer based tool-box used to implement SG2C is described.

Intellectual objects and objective knowledge

Since the 19th century, the clinical observations of speech troubles associated with brain injuries have supported the stereotype of localization of memories in the brain. A critical reconsideration of theses observations (Rosenfield, 1988) recently confirmed the idea that remembering should *not* be

regarded as the reading of a *mental* record, but as a cognitive reconstruction of a fact or event, through the examination of its context, as given by *external* records of various kinds (Halbwachs, 1925, Bartlett, 1932, Connerton, 1989, Middleton & Edwards, 1990). The mental activity of reconstructing facts or events out from the external context, we may call recollection (what Halbwachs calls *souvenir* or Bartlett *remembering*), while the term memory will be restricted to the very records or documents in the context - in accordance with the meaning of the term *memory* in computer science. Memory corresponds to a passive state of the information, while recollection refers to a cognitive activity, a processing of memories. In pretty much the same way, we will contrast *knowledge* with *expertise*. In fact, man's environment is artificial, and is used for recording knowledge. As the psychologist Janet (1935) suggested, man's environment is made of *intellectual objects.* Organizational memories can thus be considered as warehouses of intellectual objects, and knowledge management as the management of such storehouses.

Intellectual object

Tools are the clearest examples of intellectual objects. A tool is adapted not only to its use, but also to its user. It teaches the individual how to work efficiently. It does also teach him how to cooperate with his fellow workers. This is particularly the case with tools of artificial intelligence, but it holds true for any tool, however unsophisticated : every tool is an artificial thought device.

More precisely, every tool embodies what in ergonomics is known as a *model of the user*. Whether empirical or scientific, whether implicit or explicit, the user's model is the result of accumulated technological knowledge about the role and efficiency of both the tool and the worker in the production process. In order to be efficient, the worker has to conform to the user's model, inasmuch as he perceives and understands it (of course, the *actual* user's model does not necessarily coincide with the model which the designer of the tool had in mind.). Once having resolutely admitted the objective character of knowledge, the term document can be given in an extended meaning in order for it to include tools, as well as the organization of the work place, of the shop, of the plant. More generally the organized space (Bolt, 1984) can be considered as the objective form of knowledge.

The capacity to incorporate and transmit knowledge is specially developed in a special class of tools, mostly verbal (and/or graphic) in kind : instructions for use, recipes, norms, plans, which constitutes the technical literature. This class of tools constitutes an elicitation, only partial, of the knowledge embodied in the objective forms. Many situations, routines, and behaviors are set up for the elicitation of situated knowledge. For instance, training situations of various kinds, where some expert teaches an apprentice his trade, or instructs a newcomer to a new job, are meant to help the unskilled operator to "read" situated, objective knowledge. Acquiring skills means

mastering the methods for transforming objective knowledge into efficient productive action. That is what individual knowledge management is about : turning knowledge into expertise.

Knowledge management : a three-stage process

Knowledge, as the outcome of the prior development of expertise and as the condition for the developments currently in progress, is embodied in artifacts. It can be located, described, transmitted, and acquired. *Expertise* is the performance of knowledge. The output of this performance consists of intellectual objects, documents for more knowledge, and possibly new knowledge, which will be put into use in another performance. Expertise cannot be “encapsulated” or “boxed”. Only its output can. Knowledge can be viewed as raw material for expertise, or dead expertise. Reciprocally expertise is living knowledge.

To distinguish between these two notions whose terms are relatively arbitrary, is helpful, because it draws attention to a third stage, intermediate between the passive state of knowledge, and the active state of expertise. This transition stage of knowledge management consists of skills, a set of mental procedures, enabling the agent to exploit the intellectual objects displayed in the environment, in order to have passive knowledge realized into efficient productive action. Skills, whether social or technical, collective or individual are the basic components of the knowledge management process (Poitou, 1990). They correspond to the agents' ability to actualize, in practice, the knowledge stored in objective form, by adjusting ever more precisely to the actual user's model of the tools and other intellectual objects they are in charge of.

Converting knowledge into efficient behavior through mental operators

Converting situated knowledge into efficient behavior is geared through mental operators. They are the constituents of professional qualification. A skilled worker organizes mental operators in well-ordered sets in order to set up the most convenient storage of knowledge at his work place, and exploits the knowledge stored in the workplace according to the most efficient rules of stock management, that is : to alleviate the mental load by indexing knowledge in external storage under mental operators as easy to discriminate as possible; to order mental operators in accurate series, so as to retrieve adequate knowledge “just in time”; to re-externalize knowledge, when no more needed, under well-ordered mental operators in order to carry on mental activity with “zero-stock” and "minimal outstanding”. Training and professional achievement consist in improving the set of mental operators, by increasing their accuracy, and the adequacy of their ordering.

In order to improve the management of knowledge, Engelbart suggested to consider not the individual agent, but a larger system he called “H-LAM/T:

Human using Language, Artifacts and Methodology in which he is Trained" (1963, p.40). By proper assistance to the objective part of this system, intelligence could be amplified: *"Intelligence amplification* seems applicable to our goal [*of augmenting the human intellect*] in that the entity to be produced will exhibit more of what can be called intelligence than an unaided human could demonstrate. That which possesses the amplified intelligence is the resulting H—LAM/T system, in which the LAM/T augmentation means represent the amplifier of the individual's intelligence." (Engelbart, 1963, p. 45). While situated knowledge consists in artifacts, the language and methodology parts of the H-LAM/T system are the constituents of skills. "Language-Artifact-Methodology" is the part of the system where the best can be made of computer assistance. Basically, it means designing *methodologies* and tools for giving fast, exhaustive and reliable access to accurate and pertinent knowledge, in both its objective (*artifacts)* and its verbal (*language)* forms. Or, put otherwise, it means improving the management of documents (in the extended meaning of the term). However, before any attempt at improving the functioning of the H-LAM/T system, the division of labor has to be given consideration.

Organizational memory and division of labor

Technical cognitive processes occur within a technical and a social division of labor. The technical division of labor - both vertical (the segmentation of the operating cycle) and horizontal (multiple work stations for the same operation in the cycle) entails a distribution of knowledge. The situation created by the social division of labor (Poitou, 1997) is illustrated by this statement from Bonnet, Haton, and Truong-Ngoc (1986), "The expert withholds some capital. The knowledge engineer is paid to take it away from him." Such is the context in which cognitive processes in the firm are to be analyzed.

The management of knowledge and the division of labor

The technical and the social divisions of labor are conditioned by the changes in the economic circumstances. Accordingly, the distribution of knowledge in the firm is itself ever-changing, because the structures over which it is spread change continuously under the effects of financial, administrative, and technical reorganization in the firm and its branches.

The development of knowledge and diffusion of expertise are partitioned and segmented in nature. They are segmented in time, according to the ups and downs of competition between firms, between branches within firms, between departments within the branches, and as a function of individual social conflicts and rivalries resulting from competition between members of the work force. They are segmented in space, by the division into departments, each subjected to different, sometimes contradictory industrial policies. And finally, they are hierarchically segmented.

These divisions are a source of conflict due to power relationships between the economic agents, buyers, and sellers of labor. In such a context, pieces of knowledge are elements of professional qualification, and consequently, become bitterly fought-over goods. The outcomes of these conflicts determine the distribution of knowledge in the firms, and, as we shall see later, the absolute non-equivalence of places within the firm.

The current economic conjuncture causes further instability and discontinuity in the generation and use of knowledge in industry. Under these circumstances, the evolution of knowledge and skills is neither linear nor consistent, but dictated by the particular, rarely synchronized patterns of economic expansion, technical change, and product innovation.

The technological base

The notion of technological base (Poitou,1987) simultaneously refers to the lore of technical knowledge from which agents draw in order to carry out their production practices, and to the heterogeneous collection of pieces of equipment, in which it is embodied: shops, offices, work places and stations, tools and equipment, organizations as well as flow charts, documents and technical instructions.

The ongoing effort to organize production involves ordering and arranging the agent's environment so that it will be as controlled as possible while still being loaded with knowledge. This environment is deliberately pigeonholed with pointers, reminders, landmarks, and signs, which can later be used at low cost. These multiple and redundant cues mark the path for navigation in a controlled and meaningful environment, which might be called a knowledgeable workplace, insofar as it enables mental functioning without an excessive load of representations, for knowledge is already written in the environment. However, while the whole work environment can be considered as a collection of documents, navigation in this sort of hypertext is not equally open to every member of the firm. Neither the understanding of the knowledge embodied in work artifacts, nor the mutual understanding among fellow-workers are independent from the *places* of the agents in the network of social relationships woven through the division of labor. Actually, what sense is made out of knowledge situated at a given location is dependent upon who is trying to access

it, or rather upon his *place* in the social division of labor. Before considering the collective management at the level of the technological base, we shall thus first examine the concept of *place*, and its implications for the meaning of "documents".

The concept of place

The term "place" is used in its topographical sense, of course, as the physical work place, and in its cognitive sense, as a situation for the storage of knowledge. "A place" should also be understood in a social sense, as a position in the social relationships of production, indicated by wages, job assignment, qualification and skill requirements. The fundamental relationships between the instrument and the performer (Poitou, 1978) and between the instrument and the knowledge (Pêcheux & Fuchs,1975) are nested in the network of place relationships. Thus, the relationships between places determine the possibility for the performer of a task to capture the knowledge embodied in the environment, in order to actualize it as expertise, i.e. to perform the task.

Place and discourse : the construction of meaning

In developing his automatic discourse analysis program (A.A.D.), Pêcheux (1969) argued that every text or discourse could be understood as a selection of elements among an higher-level *master discourse*, which in turn could be viewed as a virtual hypertext, where parts are linked together and serve each other as references or rather as paraphrases. Understanding means thus navigating through this hypertext and using the associated discourses, or co-text, in order to enlighten the text actually read or heard. Enunciating a discourse also requires a navigation through the master discourse, in order to organize a selection of subparts.

Pêcheux further stated that the selection of the constituents of the master discourse, the linkage of its subparts, and the availability of the paths are ruled by the social conditions of discourse production, i.e. the socio-economical conjuncture and the social relations between speakers or writers, and recipients. Thus navigation through the master discourse is determined by:

(1) Morpho-syntactic markers : the linguistic cues serve as shifters from node to node. *What* is said in a document is understood through the *way* it is told or written. Not only the vocabulary, but the argumentative style, the grammar, the enunciation marks also serve to build meaning, by indicating how to navigate within the text, and between the text and its co-texts. Linguistic characterization of the text, using lexical and morpho-syntactic markers can be used to trace and assess relationships between different parts of the same text, and among texts between documents (Pêcheux, 1969, Stiegler, 1993).

(2) Organizational cues, which help to identify places : knowing *where from,* that is *from which place* a speech or text comes, also matters in

understanding (Pêcheux & Fuchs,1975). Organizations are usually careful in making clear who addresses whom, and what the respective *positions* in the corporate structure are. Organizational indices concern the authorities (in terms of department, offices, positions) and agencies which issued the document, the people to whom it was directly and/or indirectly addressed, the specific functions the document fulfills in the organization (whether it is a memo, an agreement, a bill...). Industrial writings contain a great many markers of this sort, which are evidence of the fact that the construction of meaning by both the writer and the readers is heavily dependent on these organizational cues.

(3) Relationships between places. By selecting texts in the domains of the master discourse they have access to, speakers build a text relevant to the present and local conditions at their place. Readers take advantage of the various linguistic and organizational cues in order to identify the functional and *social* relationships between places, and to select accordingly in the domains of the master discourse they can access, the paraphrases relevant to the text. Comparing the text to its paraphrase allows to dispose of ambiguity.

Place and communication : relations between places

Information and knowledge in the firm are neither freely accessible nor freely activated by the agents, but are instead distributed according to the places agents hold in the work process and in the social relations of production. This is fairly obvious. Pêcheux's analysis rests on quite the same idea : the antagonistic nature of the social relationships of production result in an heterogeneous distribution of access to the paths and domains of the master discourse. Thus while different places determine different understandings of the same text, identical places under identical circumstances, result in similar understandings, through the selection of meanings which are "paraphrases" of each other (Pêcheux *et al*., 1979).

These very determinants regulate the forms of knowledge elicitation and circulation. As far as trade and job go, communication and understanding are totally possible between two agents whose places are identical. They are restricted to common sense, that is, for such matters, to almost nothing, when the places are completely disconnected. More generally, we may conclude that the way in which a given item (be it verbal or objective) is put to use will depend on the place of the performer who makes use of it. The reconstruction of expertise from available knowledge is circumstantial, and in most cases contentious. Attention is – or should be – usually paid to relationships of that sort in training situations : between the respective places of a trainer and a trainee, an optimal distance has to be found such that the former has something to teach, and a capability to make himself understood, and the latter has a desire to ask questions, and a capability to understand what he is taught. The Method

3A, to be presented later, has been designed for the elicitation and collection of expert knowledge on the basis of such relationship.

Place and representations

The circulation and processing of knowledge in a company is not a mere technical process but a social practice. Knowledge situated at a given work place is embodied in documents or artifacts, and elicited through speeches and texts which contradict each other because at that place they represent inconsistent, divergent, or antagonistic interests. Thus one cannot arbitrarily assign any inherent, specific and stable meaning to a document or artifact. This idea is clearly in opposition with current knowledge representation theories, which generally assume that knowledge can be represented and representations supplied irrespective of the place where knowledge is to be used. Situativity theories assume for their part that short-term transient representations of the situated knowledge are converted into expertise. Such representations are reconstructed on the spot, in a rather broad fashion, and continuously restructured in order to be adapted to the demands of the situation and of the work in progress. It has been here further suggested that situated knowledge available in documents of any form, is elicited in different ways according to the differences in places among the users. Accordingly, and in opposition with the representationalist view, any computer aid to knowledge management should *not* attempt to assign any meaning to a document nor to *re-present* (i.e. formalize or model) the meaning of a document. A more modest, but also more efficient and practicable goal is *helping* the user to reconstruct meaning *by and for himself*. What is actually needed in order to make the knowledge capitalized in the corporate technical heritage more readily available to members of the firm, is assistance in preparing, storing, retrieving and processing documents, i.e. *editing* knowledge, instead of *formalizing* knowledge. The processing and the integration of knowledge are and ought to remain the responsibility of the performers of the task.

Collective knowledge management

Because the technical division of labor causes heterogeneity in the distribution of knowledge in the firm, while heterogeneity of places in the social division of labor entails differential access to knowledge, it has always been necessary in cooperative production to develop a body of empirical practices and methods for "displaying" (Gibson,1979) elements of the technological base. This process of enumerating, locating, situating, diffusing, and routing knowledge, is endless because the firm is an ever evolving, antagonistic/cooperative body, where information does not flow through a network of open channels connecting equivalent positions, but between sets of places in the firm.

The aim of these practices is to establish, maintain, or restore some degree of mutual understanding between members of the firm, with respect to the technological base. They are the methods naturally developed, in order to capitalize knowledge and to circumvent (as much as possible) the obstacles to communication of knowledge, through the elaboration of accessible documents (either oral or written) This important activity of technological base management result in verbal elicitation (either oral or written) of objective knowledge. It transforms documents (in the extended meaning of the term) into documents in the ordinary restricted sense. Within the network of places where it is used, the knowledge embodied in artifacts and work places is documented, or more precisely *edited*. The conversion of objective knowledge into verbal knowledge can be observed in many situations such as training sessions. In-house training can suggest a methodology for the survey of the existing technical practices, their description, and their regular updating by the performers themselves.

In-house training, and more generally methods for the description of knowledge contribute to knowledge capitalization, and to the development of a shared corporate culture. In turn, sharing a common technical culture makes the description of activities and the elicitation of knowledge easier. The objective foundation of the corporate culture is the technological base, that is equipment as situated knowledge. As already stressed in the foregoing sections, access to the technological base is unequal. Some domains are more specific to certain shops and offices, while others are common to all the departments of the firm. However, as the actual instrument for cooperation, the technological base constitutes the deposit of a lore common to all members of the organization. Descriptions of knowledge are marked by idiolects, some specific to the whole organization, which they serve to differentiate from the outside, others only to some departments. In-house parlance is the vernacular which makes it possible to members of the organization to overcome misunderstanding. Idioms are markers useful for referring statements to their conditions of production. They point to the section of the organization where a statement comes from, and thus to the context from where to extract meaning.

Collective knowledge management systems

The foregoing considerations led to what might be termed an ecological approach to the management of organizational memory (Poitou, 1987, 1997). External intervention or assistance should not be permitted to alter the common lore grounded on the technological base. Increasing and complementing the descriptions of situated knowledge should be achieved according to in-house practice, with a concern for the conservation of local idiolects. Assistance to knowledge management should consist exclusively in enumerating, locating, describing and giving access to knowledge (Poitou, 1995). This approach has been validated through the implementation of the DIADEME system at Electricité de France (EDF), a French electrical power production and supply

company (Ballay & Poitou, 1996). It is presently further developed through SG2C.

DIADEME

DIADEME, is a computer assisted system for the capitalization and the management of research engineers' knowledge, designed, developed and implemented at EDF by Ballay. Considering that Poitou's conception of a system for the management of collective knowledge (CKMS) was quite germane to the rationale and the general outline of DIADEME, Ballay associated the present author to his project.

Ballay considered that knowledge-based systems are specialists' tool, hardly suited to capitalizing on knowledge collectively, in the business environment. The alternative, information system approach has much to commend it, because it apparently draws together the whole mass of information. There is a belief that such systems contain knowledge as such, and maybe even all transmissible knowledge. Unfortunately such knowledge is so all-encompassing and abstract that confusion arises, and useful know-how is hard to pinpoint, in a haystack of information. Therefore, the critical problem was for Ballay how to qualify, describe and locate the information, to retrieve it at the right time and interpret it in a practical context. He opted for a knowledge management system, a tool drawing from both knowledge-based and information systems. DIADEME comprises a compendium of knowledge, mostly contained in documents (the knowledge base). It also contains a set of tools through which to manage and bring such knowledge into play, by subjecting it to the appropriate processing (searching, browsing, viewing, diagnostic assistance). The knowledge base is a compendium of information, organized in a number of possible formats, and even on a number of media. The knowledge base may include decision-making tools, databases etc., if necessary. Even so, the document is the basis of the system, and most frequently used, as ideally suited to receiving information of various types, from text, images, tables and diagrams. The main management tools used by DIADEME are: tools for the production and querying of documents (WORD, ACROBAT, DYNATEXT); full text indexing and information tool (TOPIC) which enables natural language queries or requests to be generated ; essential communication tools for network operations (client/server architecture). Furthermore, DIADEME includes a specialist's tool for the assistance to diagnostic of overhead power lines (developed with SMALLTALK language and using fuzzy logic techniques to model qualitative judgments, based on "expert opinion").

During Spring 1995, the present author conducted at EDF a survey of the engineers' knowledge management behavior, of their attitudes toward a prospective CKMS, and the next year, a survey among the same engineers about their users' reactions to the implementation of DIADEME. Both studies, as well as reports from further users, have confirmed that DIADEME meets expectations. Actually, the system is being extended outside the department where it was originated. However, TOPIC, the full text indexing and information tool does not absolutely preclude external intervention in the

description of knowledge. Actually, at least in the so-called “concept” mode of request, it requires knowledge modeling. As argued in the foregoing sections, this might not be without disadvantages, which a tool like NOMINO, to be presented later, might hopefully help to circumvent.

SG2C

SG2C is an acronym of „ système de gestion collective des connaissances “, i.e. „ collective knowledge management system “ in French. The objectives of SG2C are not different from those of DIADEME. However, more stringent criteria are set with respect to the idiosyncrasies of the local corporate culture and of its sub-cultures. Thus, on the one hand, a method for the recording of expertise is proposed, which relies on in-house personnel and practices, and on the other hand, computer tools are used for full-text retrieving and indexing which avoid introducing surreptitious preconceptions in the knowledge base. To provide the easiest access to documents, together with fast, accurate, and exhaustive means for reworking the meaning of information when conditions and needs change, requires to explicitly delineate the heterogeneous conditions under which knowledge is used in the company, i.e. to locate knowledge in the technological base and to describe accurately how it is “started-up”; to thoroughly collect the knowledge in that base without making assumptions about its meaning, i.e. draw up a descriptive catalog (“*catalogue raisonné*”) of knowledge existing in the base, by formally describing the documents which record it; and to present the users of knowledge, who are the actual performers of expertise, with an instrument for retrieving knowledge in order to use it as *they* see fit.

Collecting and recording experience

In order to display a complete description of the technological base, the existing verbal documentation usually has to be supplemented by further descriptions of the objective knowledge and of its location in the organizational structure. This has to be done within the "place" relationships, by the performers themselves. As already noticed, members of the firm know how to explain a job, what has to be known in order to perform it, where lies the information, what are the decisive cues for appropriate action. Furthermore, they are familiar with the local idiolects, which provide the bests cues for locating situated knowledge. The methods thus used by the personnel for transmitting and eliciting knowledge should be observed and recorded, in order to document the technological base.

M3A, a method for the autonomous analysis of activities

To that end, we are currently developing a method for the “Autonomous Analysis of Activities (Method 3A, Poitou, 1991) which offers members of the personnel a procedure for describing their own activities, in line with the self-confrontation methods initiated by Von Cranach & Harré (1982). Our method is

based upon a trainer/trainee relation; it requires the participation of three persons: an expert, whose activity is to be described, an "observer-trainee", and a naive trainee.

In order to analyze the knowledge underlying or inherent to the activity of a given expert, at his work place, the first step is to select an “observer-trainee”. Performing a different task in the same office or shop would qualify a person for that role. It will permit enough familiarity with the job situation for basic understanding, while carrying enough uncertainty for accurate and relevant questioning about the particulars of the task to be described. The selected observer is then introduced to the observation and interviewing techniques of the Method 3A.

Observers must first observe and record the behavior of their "instructor", i.e. their fellow worker in charge of the job to be analyzed. They will observe and question their trainer up to the point where they will be able to actually perform the task. The correct performance of the job by the trainees evidences that knowledge has been efficiently described in terms and statements clear enough for any member of the firm. Observer-trainees must further write down a notice of instruction about the job, technique or device they had to master. This verbal record of knowledge is then read by the second trainee, completely new to the job. When he can perform the job to the satisfaction of the expert, then the recording of knowledge is considered to be valid. Not only the device is correctly described, but the *modus operandi* is accurately recorded. Otherwise, the instruction notice is improved, and the validation procedure iterated until a correct performance is obtained from the naive trainee. On the basis of these purely empirical tests, it can be assumed that knowledge has been converted from its objective into its verbal form, accurately and completely. Such record increments the corpus of descriptions of the technological base. We can now turn our attention to the task of editing knowledge, whether available in documents, or through descriptions by the Method 3A.

Editing knowledge

Documents only become completely comprehensible if they are related to the work situations and operations to which they refer. The link has to be created between the organization and technical operation of the company on the one hand, and the descriptions of the in-house technical practices, on the other hand.

The descriptive catalog

A three-file descriptive catalog (*catalogue raisonné*) of the activities carried out in the firm, and of the knowledge associated with these activities, will be established and kept up to date. It will serve to locate and record knowledge in the technical base.

The first stage in editing knowledge from the technological base consists in creating linkages between documents or parts (excerpts) of documents. Links are established according exclusively to organizational and linguistic considerations. In a way, through this operation, the firm documentation is organized as the *master discourse* of and about the *technological base*. An indexing system based on organizational references and linguistic cues, will allow to extract a set of relevant quotations from all the texts dealing with an object of query. In other words, the user is not presented with rewritten knowledge, nor with a representation of knowledge, but with items already available in the technological base, edited according to the organizational and linguistic mores of the firm, in order to help *him* construct or improve *his* representation of the item of interest. Thus the descriptive catalogue rests basically upon three files: a document inventory, a data file, a notional file.

The document inventory

Existing documents are registered according to the usual rules of the documentation service, and indexed with respect to their organizational markers. The writing of new documents is computerized, and structured according to norms aimed at making their electronic edition and circulation as easy as possible. The record of each document in the inventory file includes, in addition to the usual archive indexes, complete references about : its origin, authors, addressees direct or lateral; its history (i.e. which documents have preceded, which it modifies); the data it refers to; the concepts it refers to. Documents are processed in order to select data and concepts they contain. It should be noted that the whole document is processed, including not only the text itself, but also the identification marks it bears (such as name of the authors and addressees, departments which released it, and so on). Excerpts are fed into the records of the relevant entries in the data file and in the notional file.

The data file

Items in the data file are the names of "individuals" (person, positions, part or subpart of the organization, object or collection of identical objects) which are recognized entities in the firm. Every data has its own record. An item record consists of the name of the data (for instance a registration number in a stock list); the category to which the data belongs, i.e. the relevant noun word from the lexical file; the excerpts where this very item is mentioned; the references to the documents where the excerpts come from, and key words, i.e. relevant concepts from the lexical file. Thus a data item is described in a file entry which presents the entity under consideration in precisely the terms used by the agents or in company documents, and only in those terms. The key words used to index the documents consist solely of the items in the notional file entries.

The notional file

The notional file is an in-house dictionary of corporate knowledge. Entries in this file are noun phrases parsed out from the firm documents. Under the heading of a noun phrase, there are : a definition made of the relevant excerpts, which allow for the construction of a concept of the heading; a list of related notions, i.e. other noun phrases co-occurring in some excerpt with the entry of the record; a list of “see also” or associated notions. Associated notions are the results of “navigation” in the file. A record is kept of the consultations made in the file. When, even through sheer “mental association”, several notional entries are “browsed through”, they are said to be associated. Of course, every excerpt quoted in a notional record is given with the references to the document it comes from. Updating the files is exclusively additive: deleting a previously recorded excerpt is strictly forbidden. The purpose of this rule is to make sure that every change in thinking, doctrine, or usage affecting in-house knowledge is recorded.

The descriptions that occur in the files are not the outcome of the file-entry writer's analysis or representation of what is written in the main documents. The descriptions are collections of excerpts taken from the main documents processed. This procedure makes it possible to avoid introducing into the texts any special vocabulary other than that found in the technological base under study (i.e. the lexicon of the parlance common to the members of the company); guarantees the local and historical validity of the definitions; preserves the significant idiosyncrasies of the company members' own syntax.

An item's description thus provides a means for linking together the various pieces of information about that item, while preserving the particularities or even contradictions among the agents who employ it and were consulted on the matter. This is possible only if the system provides an hypertext browsing system built automatically.

Implementing SG2C with ATELIER FX

From AAD to ATELIER FX

Earlier in this paper was presented the concept of place, which is central in our conception of a CKMS. The concept comes from the work of Pêcheux (1969) concerned with the early development of a program for discourse analysis. Pêcheux’s ideas have also been seminal in the research conducted at the Center for computer aided text analysis (Centre ATO) at the University of Quebec in Montreal, where NOMINO and ATELIER FX have been developed (Dumas, et.al. 1995,Plante, 1995, Plante & Dumas, 1995). It is thus perfectly understandable that both approaches meet in an effort to apprehend knowledge through the examination of its support and vehicle.

Atelier FX

Atelier FX is a tool-box for the building of knowledge based data bases. Knowledge is stored either as descriptions in natural language or as more formal categorizations. NOMINO, a linguistic analyzer, make it possible to automatically transform the former into the latter.

Automatic text analysis with NOMINO.

FX a programming language, developped in Allegro Common Lisp (version 3) environment, is specially well suited for the writing of parsers. NOMINO is a parser for the linguistic analysis of the French language.

The processing of a sentence with NOMINO results in the lists of all the lexical forms, lemmatized, together with the indication of their unequivocal syntactic category. For instance, for the following sentence : "Avec une carte à mémoire ou à puce, Jean transporte ses dossiers médicaux, scolaires ou autres, gagnant généralement de l'espace et du temps.", NOMINO builds a description with unequivocal syntactic categories as in Fig. 1. There, s-b stands for adverb, s-a for adjective, s-v for verb, s-n for noun, s-ucn for noun phrase, s-af for other forms. Lemmatized adjectives obtains as singular masculine, nouns and noun phrases as singular and verbs, as infinitive.

Noun phrases (UCN, i.e. unité complexe nominale) are groups of words based upon a head word. For instance “base de données” (data base) has but one referent, and does not refer to any specific set of data, as it is the case with “sur la base de ces données” (on the basis of such data). The number of UCN is usually high, and so more so in technical documents. Accordingly, sorting out and assessing UCNs is indeed a drastic requirement.

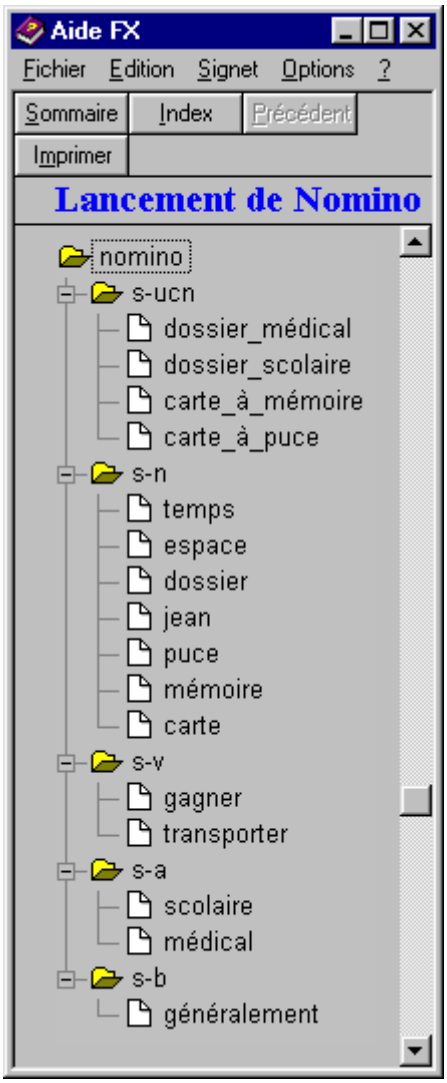

Fig. 1 : A sentence analyzed with NOMINO.

Descriptions in NOMINO make it possible to compare sentences in a text, through a matching technique, called “salience computation”, which assess the similarity between trees builds in FX.
This further allows for automatic hypertext navigation. A selected segment of the text can be compared with other parts of the same text. A selection of zones of the text is brought back, with indications (in terms of salience) about their similarity to the segment initially clicked upon.

The sentence " À la différence des cartes à piste, la carte à mémoire présente différents niveaux d'intelligence suivant l'utilisation visée." was selected. The corresponding NOMINO tree is addressed through FX as a request to the text. As a result a zone of sentences n° 79, 80, 81 in the text is obtained, together with an index of their salience (52.4300). Less salient and isolated sentences are also brought back.

52.4300 (79 80 81)
11.0700 83
10.2700 19
9.5700 85
8.3300 119
8.3300 37

79 Avantages et inconvénients
80 L'inconvénient majeur de la CAM tient à son coût .
81 La carte à piste ne coûte que 0,25 $, le coût de la CAM commence à près de 1 $.

Fig.2 : A request and the corresponding salient zone.

Of course, lexicons (together with item frequencies) are obtainable with NOMINO, and comparisons made through boolean operations (union, intersection, complement). Comparisons can be made between the lexicons of several texts, or between a text and a tree base, or a data base, build on an other FX program. The context of given word can be established through a local lexicon, giving the vocabulary surrounding that word within a given area of the text.

Text processing with NOMINO.

There are 4 modules in NOMINO.

- Edito processes the edition markers in the text to be analyzed
- LCMF lemmatizes and characterizes the lexemes context free
- DCS defines the syntactic categories
- DUCN assesses UCN (noun phrases).

That is, NOMINO successively performs an in depth morphological analysis of the lexems (LCMF), a local syntactic analysis (DCS) in order to obtain a unequivocal syntactic categorization, and identifies the relevant noun phrases

s(UCN). The file entry for such processing is a full text file. Indeed NOMINO does not require any text preparation, and accepts direct optic readings as input. Atelier FX, including NOMINO, is operated on PCs, with OS either Windows 95 or NT.

After proper identification of lexemes and phrases by Edito, LCMF (Lemmatisation et caractérisation morphologique du français), proceeds in order to give for every lexeme in the text, all the possible syntactic categories together with, for each category a possible lemma and its morpho-syntactic characteristic. For instance :

("réduit"
("nc" (cm "S" "M") "réduit") = noun : small room
("adj" (cm "S" "M") "réduit") = adjective : reduced
("vfle" (cm "S" "3") (ct "I" "PR") "réduire") = verb : reduces
("vppe" (cm "S" "M") "réduire")) = past participle : reduced

cm: morphological characterization
ct: tense characterization

Grammatical categories are attributed in LCMF, mostly with the help of a list of suffixes and inflections. For instance the inflection *ez* gives the grammatical category of about 35000 verbs. That rule has but some 10 exceptions, listed in a dictionary of 300 rules. Unlike other analyzers using only lexicons, LCMF has thus the capability of processing any neologism which abide by the inflectional and derivational rules of the French grammar. This is extremely helpful, considering the high number of neologisms, namely the so-called “franglais”, of American-English origin, in the French technical literature, as well as in dialogs among technicians.

The DCS module (Désambiguïsation des catégories syntaxiques) proceeds with a syntactic analysis of every lexeme, in its surroundings. For a given lexeme, DCS either declares unique category, or takes away one of the possible hypothetical categories. There are some 400 rules to that end in DCS. For instance : “ Elle est bien *mise*.”

If the lexeme is
noun,
and **adjective**
or **past participle verb**
and preceded (-1) by an **adverb**.
then
take the category **noun** away.

Several rules are used in order to proceed with coordination, a very important syntactic procedure, namely in technical literature.

The DUCN module set up the list of the noun phrases, providing the head word, together with its various extensions. For instance :
a) adjectival: "carte magnétique", "abonné téléphonique";
b) nominal: "carte EPROM", "carte multi-services";
c) nominal with preposition : "carte à mémoire", "bloc d'alimentation";
d) verbal infinitive with preposition "carte à jouer";
e) verbal with past participle: "carte embossée".

Again, coordination is duly processed.
In the segment: "lecteur et encodeur de carte", DUCN will select:

"lecteur de carte"
"encodeur de carte"

Or, for "lecteur de carte et de puce":

"lecteur de carte"
"lecteur de puce “

The relations between UCN can be explored with the help of Atelier FX. For instance, in a document concerned with “carte à mémoire” (memory chip card), the word “carte” is included in several noun phrase as head word. We thus have a network of UCN, which is shown in Fig. 3, at two different levels.

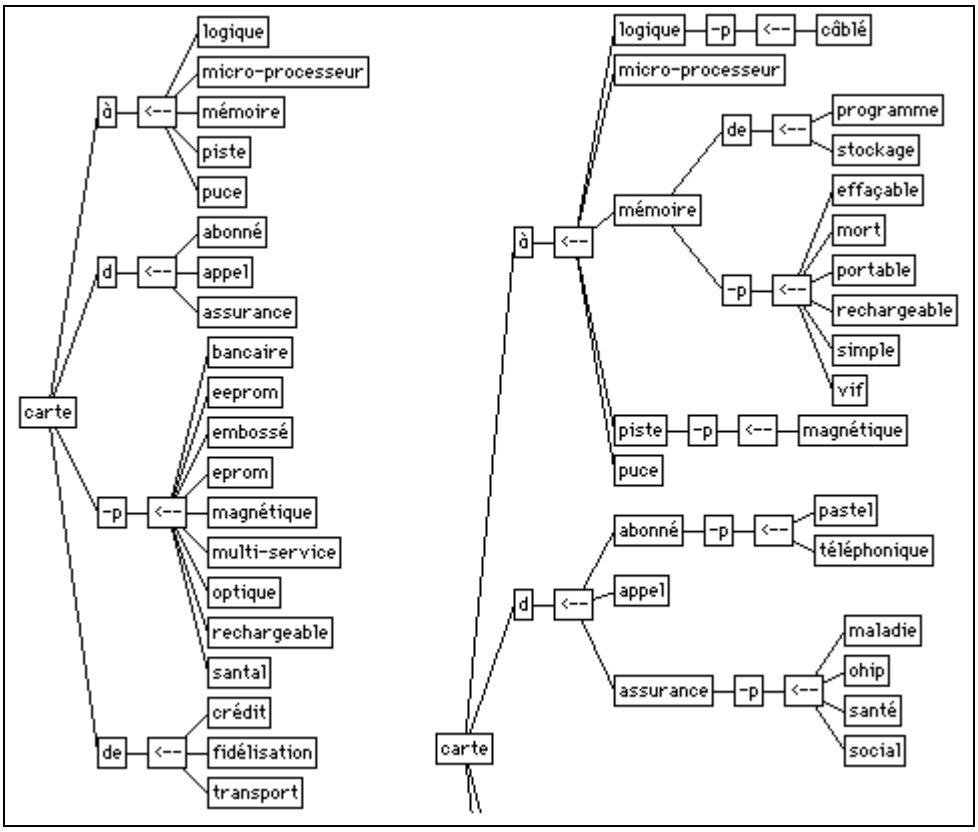

Fig. 3 : UCN network

Such graphs obtain automatically, and constitute actual semantic networks, which enable to zoom in and to navigate through the content of a documents. Several other tools make it possible to study the formal properties of the UCN networks, such as the lists of UCN ranked according to the number of extensions they either transmit or receive.

Thus the linguistic descriptions obtained with the help of NOMINO, together with the UCN networks based on such descriptions constitute a base for hypertext navigation within texts and/or documents.

The UCR (Unités complexes en réseaux, i.e. networks of UCN) bases give further possibilities to analyze texts and to built knowledge bases out of a set of documents. They are made of a special class of FX trees, known as referential network base. Trees in such bases (see Fig. 4) work with transmitter and receiver relations (← and→). Leaves (tree ends) are names of other trees in the base. Accordingly, trees in such bases are mutual references.

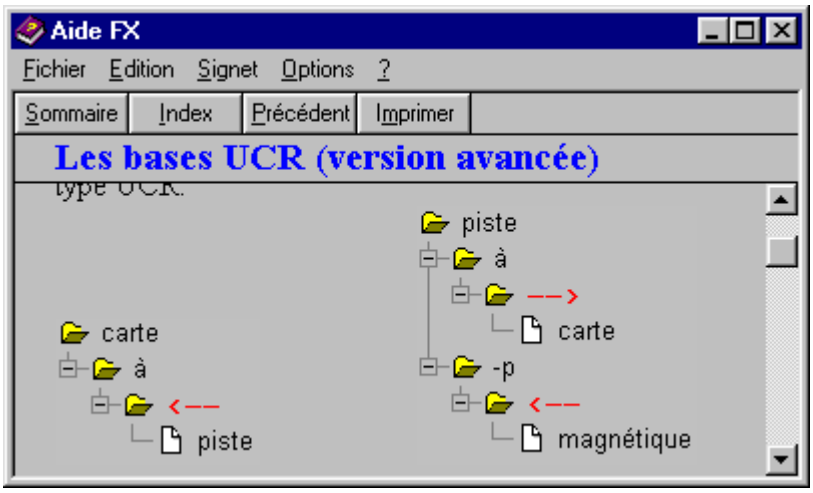

Fig.4 : UCR base

Knowledge bases

It is thus possible with Atelier FX, to develop a set of documents as an hyper-document, or as a knowledge base, where the structure of the texts is described, together with their relations, but not interpreted. Furthermore, the local idiolects are preserved, idioms are not translated, but used as specific entries within their local context. Furthermore the knowledge situated in the documents can be enlarged, through methods such as Method 3A, or commented by firm members, who can as well develop, on the basis of the present documentation, their own, more or less personal knowledge bases. In other words, Atelier FX gives the possibility of an actual collective knowledge management.

A reader may like to link his comment of a document to the text itself, in the form of a lecture note. He may use the “clip” procedure, for processing his observations or comment through NOMINO. The resulting linguistic description could be then matched with either other “clips” or documents, and documents selected on the basis of their similarity with the original comment. Files thus obtained can be organized into notional files according to their referential relationships (Fig.5).

Fig.5 : Referential notion files.

Requests can be made either from a notion toward the documentation, in order to find the relevant documents, or from a segment of a text, in order to get back the organization of the relevant notions, and further on the address of the relevant documents. It is worth noting that such knowledge base does not constitute a representation of knowledge, but rather a map of the links actually woven within et between documents.
It could be further noted that Atelier Fx provides links with Microsoft Word and Internet Explorer according Dynamic Data Exchange in Wndows 95 and NT.

SG2C and ATELIER FX

It seems fairly obvious – at least in the eyes of the present author – that ATELIER FX can fulfill completely the requirement of SG2C. If SG2C constitutes in turn an appropriate response to the requirements of the collective management of the organizational memory of the firm, then the next obvious thing to do is to design, implement and assess an interface for ATELIER FX for SG2C. The further successful development – if any- of such a CKMS would lend some support to the anthropological considerations presented in the first part of this paper. This a rather lengthy way for testing social sciences theories, but would not it be right to have social thinking before the design of artifacts ?

References

Ballay, J.F., Poitou, J.P. (1996). DIADEME: a collective knowledge management system (CKMS). in J.F.Schreinemakers (Ed.) Knowledge Management. Organization, competence and methodology**.** Würzburg, Ergon Verlag, 1996, p. 265-285.

Bartlett, F.C. (1932) Remembering: a study in experimental and social psychology. Cambridge, Cambridge University Press, .

Bolt, R. (1984) The Human interface. Where people and computers meet. Belmont, Lifetime Learning Publications.

Bonnet, A., Haton, J.P., Truong-Ngoc, J.M. (1986). Systèmes experts, vers la maîtrise technique. Paris: Inter Editions. .

Connerton, P. (1989) How societies remember. Cambridge, Cambridge,University Press.

Dumas, L., Plante, A., Plante, P. (1995) Nomino ; Version 1.0b. (Tech. Rep.). Montreal: Centre d'Analyse de Texte par ordinateur ATO, University of Quebec at Montreal.

Engelbart, D. (1963) A conceptual framework for the augmentation of man's intellect. in P. Howerton, (ed.)Vistas in Information Handling Vol.1. Spartan books, Washington, DC, 1-29.reproduit in GREIF, I (Ed.) Computer-supported cooperative work : a book of readings**.** San Mateo, Californie, Morgan Kaufmann, 1988.

Gibson, J, J. (1979). The ecological approach to visual perception. Hillsdale, NJ: Lawrence Erlbaum Associates.

Greeno, J.G., Moore, J.L. (1993) Situativity and symbols: response to Vera and Simon. Cognitive Science. 17, 1, 49-59.

Halbwachs, M. (1925) Les cadres sociaux de la mémoire. Paris, Alcan.

Janet, P.,(1935) Les débuts de l'intelligence**.** Paris, Flammarion.

Middleton, D., Edwards, D. (Eds.) (1990) Collective remembering. London, Sage.

Pêcheux, M. (1969). Analyse automatique du discours. Paris, Dunod.

Pêcheux, M., & Fuchs, C. (1975). Mises au point et perspectives à propos de l'analyse automatique du discours in M.Pêcheux (Ed.) "analyse du discours. Langue et idéologies". Langages, 37, 7-80.

Pêcheux, M., Henry, P., Poitou, J.-P., Haroche, C. (1979) Un exemple d'ambiguité idéologique : le rapport Mansholt. Technologies, Idéologies, Pratiques**,1,** 3-83.

Plante, P. (1995) NOMINO [Computer program]. Montreal: Centre d'Analyse de Texte par ordinateur ATO, University of Quebec at Montreal.

Plante, P. Dumas, L. (1995) Nomino-signet - un assistant à la construction de bases de connaissances sur les textes. (Tech. Rep.). Montreal: Centre d'Analyse de Texte par ordinateur ATO, University of Quebec at Montreal.

Poitou, J.P. (1978). La dynamique des groupes, une idéologie au travail. Paris, CNRS.

Poitou, J.P. (1983). Savoir vif, savoir mort. Paper presented at the Symposium "Les cultures populaires" of the Société Française de Sociologie and the Société d'Ethnologie Française, Nantes.

Poitou, J.P. (1987). The expert and the system. ORIA 87. Artificial Intelligence and Sea(pp. E1-E12). IIRIAM, Marseille

Poitou, J.-P. (1990) Do new technologies upgrade knowledge and skills? 7th International Workshop of the International Study Group "New Technologies and Work - NeTWork" on the theme Managing New Technologies, Werner-Reimers Stiftung, 3-5 mai 1990, Bad Homburg.

Poitou, J.P. (1991). Définition d'une méthodologie de recueil et d'extraction des connaissances au service des systèmes experts en amont de la formalisation des connaissances. (Research rep. n° 89 D0041) Paris: Ministère de la Recherche et de la Technologie.

Poitou, J.P. (1995) Documentation *is* knowledge. An anthropological approach to corporate knowledge management. Proceedings of Ismick 95, International Symposium on the management of industrial and corporate knowledge. Compiègne. 23-24 octobre 1995. Compiègne, IIIA .

Poitou, J.P (1997).Building a collective knowledge management system: knowledge-editing versus knowledge-eliciting techniques.in G.C. Bowker, L. Gasser, S.L. Star, W. Turner (Eds.): Social science, technical systems and cooperative work. The great divide. Hillsdale, Lawrence Erlbaum, , p. 235-256.

Stiegler, B. (1993). Projet OPEN. In J.-P Barthès (Ed.). Proceedings of the Symposium COMETT 93 Gestion du savoir-faire et des connaissances de l'entreprise. Problématique, modélisation ; intérêt et limites de l'approche "objet". Compiègne, Institut International pour l'Intelligence Artificielle.

Rosenfield, I. (1988) The invention of memory, a new view of the brain. New York, Basic Books, Trad. française L'invention de la mémoire, Paris, Flammarion, 1994.

Von Cranach, M., Harré, R.(1982) The analysis of action**.** Cambridge, Cambridge University Press.