

eDOC : A COLLABORATION INFRASTRUCTURE TO MANAGE KNOWLEDGE AND INFORMATION ON NUCLEAR PROJECTS AND RESEARCH ACTIVITIES

JM. Van Craeynest, F. Jacquemet, D. Chermette, S. Bonneau
Commissariat à l'Énergie Atomique, France

Email address of main author: jean-marie.van-craeynest@cea.fr

Abstract. After a brief recall of Knowledge Management issues and of the MKSM™ knowledge modelling method developed and used by the CEA, this paper focuses on the eDOC web collaboration platform designed to support a large range of cross-collaboration needs and a large spectrum of community types (from small units to large European Networks of Excellence). Online community members have different needs : knowing other members and their roles, accessing to reference information and documents, reviewing, annotating and publishing documents or sending information to all or part of members, being informed of events, scheduling shared jobs and manage tasks, discuss some questions using forums, etc. They also want to develop a common identity and portal look and structure customizations are very useful for that purpose. The Information System Department of the CEA has launched the eDOC project to supply a rapid demand growth for those tools. In the CEA, this demand is due to the growing importance and number of internal and external cross-collaborations and alliances with industrial and research partners. After a systematic evaluation campaign of both proprietary and open-source solutions, and after having defined the most pragmatic deployment strategy, we have chosen the Collaborative Portal Server (CPS) edited by Nuxeo. CPS is based on the Zope open-source object-oriented application server. Then we invested to improve functionalities and performances of CPS but also to design and implement a security policy adapted to different types of security requirements and information privacy levels. The eDOC web collaboration infrastructure is now used by 50 projects and this number keeps increasing. The main uses are internal documents repositories (to simplify classical “Intranet” building and maintaining process) and information repositories (including documents but also events, news, calendar, conferences, etc.) for external collaborations. The 6th european research framework program spread and stimulated the use of eDOC in european projects. At the end of this paper, we give some experience feedbacks and advises based on the different cases we had to treat. Main advise seems evident : keep it simple !

INTRODUCTION

Knowledge management is a management hot topic for more than ten years now. Information Technologies have had an important impact on all KM related areas. Expert systems and rule-based or case-based systems have had an important impact on both knowledge capitalization methods, decision helping systems and computer-aided learning. Document management and workflows systems have a direct impact on management and organisations. Search engines and databases as well as tag-based metalanguages (SGML, HTML, XML) have a decisive impact on document management, authoring and publication and on the concept of document itself introducing hypertext and semi-structured navigable content. At the same time, the use of web publishing on the Internet has changed the habits for information access and the massive use of electronic messaging has drastically changed the perception of distance between persons and the way of working together. Now, the maturity of document management technologies combined with the need to improve cross-collaboration methods and efficiency have lead to a recent issue : supporting “communities of practice”.

This paper gives an insight into eDOC, a web collaboration infrastructure deployed to help the CEA and its partners to work together. We will place this practical case study in the context of knowledge management and in the context of alliances and external collaborations that the CEA must set up to reach ambitious research and industrial objectives turned toward nuclear energy, defense and nuclear deterrence, information and healthcare technologies.

The European Union has shown the way at the Lisbon 2000 European Summit : to become the most competitive knowledge economy by 2010 !

KNOWLEDGE MANAGEMENT ISSUES

Knowledge Management encompasses various domains of business practices, related to human resources management, information, information technologies, strategy, and accounting. Facing such complex issues, especially in R&D organizations producing knowledge and innovations, knowledge management cannot only stand on a few organizational or technical solutions. All functions must be involved to achieve these strategic objectives: management must find realistic incentives and inscribe Knowledge Management as a core management objective (just as Quality Insurance has been). Human Resources departments and education institutes can benefit from new technologies to improve training methods. Research units have to launch knowledge capitalization projects to retrieve, save and transfer critical knowledge, technical skills and know-how. Scientific committees have to improve publication as well as evaluation processes.

These tactical objectives cannot be achieved without an efficient, usable and robust corporate information system. An a posteriori knowledge “fireman-type” saving action can be set up in the case of major events (for example, decision made to stop nuclear weapons experiments), but we must promote an ongoing capitalization effort as well as embed KM into projects and activities management methods. This effort, during and after the project, is implemented through a perennial information system. It should provide a wide range of services for scientific publications and patents management, corporate or local knowledge bases and document repositories, project management and collaboration tools, rich media authoring, etc.

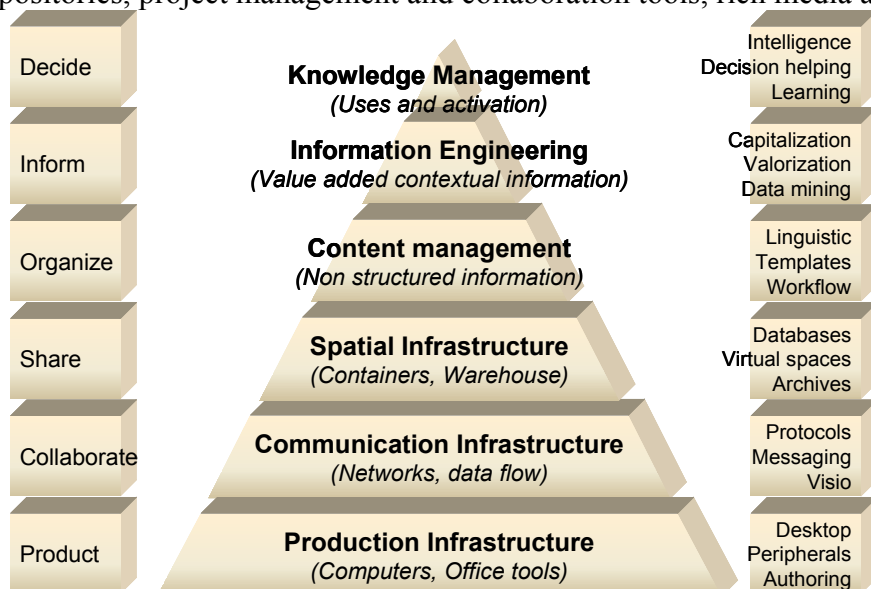


FIG. 1. Information Needs Pyramid.

To help understand that knowledge management initiatives depend on the accumulation of Information System layers, we re-used the famous Abraham Maslow’s “hierarchy of needs” model. The main idea is that certain lower level of needs have to be satisfied before higher needs can be satisfied (it was not useful to build networks before having distant CPUs to communicate with). The second main idea is that upper levels generate improved specifications over the lower ones. For example, Internet Protocol is taking new multicast protocol into account to allow multimedia real-time collaboration and video broadcast. In this

model, each level increases the information value (defined as the usability for an operational process, to decide and act in a particular way, depending on contextual objectives and resources).

Our experience has pointed out that it is very difficult to work at the upper levels when lower levels are not efficient enough. For example, a very important effort can be spent to build huge knowledge books. This effort can be justified more easily if collected knowledge can be disseminated. Using new semantic multimedia authoring tools could help for that. Such tools are related to the first level : production infrastructure. Knowledge books also get benefits from bibliography and contextual information links; such links are facilitated if document databases are already accessible and maintained. Document databases can conversely benefit from higher meta-information management and tools like taxonomies or linguistic resources, to facilitate searching and crawling among wide amounts of data.

SOME INITIATIVES FROM THE RECENT PAST AT CEA

A knowledge management strategy cannot only be founded on human management, information management or information systems improvement. All corporate functions should be involved using both traditional and new information technologies possibilities. Management staff should inscribe Knowledge Management objectives as a core-management objective (as Quality Insurance has been). Human Resources Management can improve knowledge dissemination introducing e-Learning methods for pedagogical and skill management purposes. Scientific Committees can use information technology to improve publication and reviewing process (what has been done for years). Business units and projects should launch knowledge capitalization projects to preserve critical expertises, skills or know-hows. They should also provide collaboration tools to facilitate synchronisation, information sharing and on-going corporate memory constitution. Of course, all these objectives cannot be met without an efficient, usable and robust information system. This is what we are involved in!

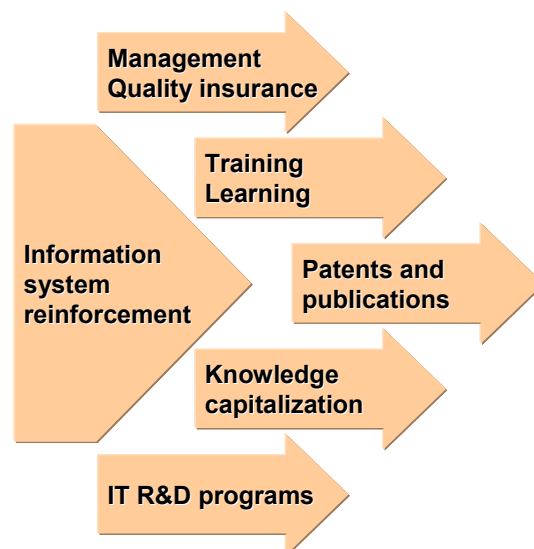


FIG. 2. Different axis of a Knowledge Management Strategy.

Among the various knowledge management initiatives and before focusing on the eDOC project, let's have a look at an upper level (in the KM pyramid) initiative. We want to talk about the electronic knowledge books authoring using a dedicated knowledge modelling method. This method, named MKSM™ (for "Method for Knowledge System Management") has been created ten years ago by Jean-Louis Ermine. It has been used and optimized within the CEA between 1996 and 1999. This methodology follows a systemic approach on a complex subject. It distinguishes six different viewpoints and provides six diagrammatic chart templates to help authors explicit knowledge and readers interpret information.

The *activity* model is used to describe the contextual and macroscopic processes in which the subject takes place. It presents the activities, the actors, the flows and the interactions between them

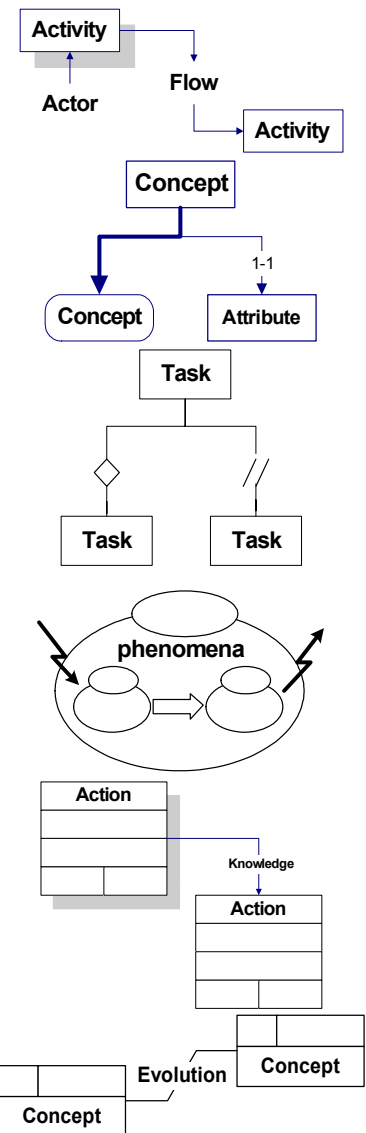
The *concept* model is used to describe the structure of the system in term of sub-systems and attributes

The *task* model is used to explicit checklists, procedures or problem solving methods

The *phenomena* model is used to explain interactions between phenomena in terms of causality and environmental dependencies

The *historical* model is used to tell the story of an invention, a specific version of a system. Focus is made on main phases of the project and return on experience

The *concept evolution* model is used to complete the history model to represent the different ways explored or the different species or versions realised



Many Knowledge Books have been written on various subjects (electron beam welding, radio-ecology, criticality risk management, software development, uranium metallurgy, weapon safety, etc.). Those books merged MKSM™ models with conventional presentations (texts, pictures, curves, equations, etc.).

The human cost (in terms of time needed and cognitive charge) of this knowledge capitalization approach implies that it must be limited to very critical knowledge preservation. To avoid or reduce needs of those kinds of expensive post-mortem actions, we must promote ongoing capitalization by all means : because they provide a common working space and a natural documentation and information repository, collaboration platforms actually contribute to this capitalization effort. That is the reason why the following parts of our paper describe collaborative virtual spaces implementation with eDOC.

ZOOM ON COLLABORATION ISSUES AND “REAL WORLD” CONSTRAINT

eDOC is designed to support “communities of practice” through document-oriented collaborative work.

By “community of practice”, we mean different kinds of group that can aggregate for a short or long period around common objectives. Such communities can be small or large, tightly or loosely connected, they can be organized hierarchically or as a network, but all share similar characteristics and communication needs.

Among these needs, we can point out the community identity assertion (often presented through the “home or public pages”), the asynchronous or synchronous discussion spaces, a membership directory with “skills or preferences”, methods to implement community structure (sub-groups, projects), one to one or broadcast information transmission features, document repository and knowledge base (including access-rights driven search engine of course), workflows and decision traceability tools for the collective memory, management tools to coordinate works and even control usage.

Serving such facilities for homogeneous and autonomous communities just seems to be difficult and the reality is much more complex: communities are often very heterogeneous because each member belongs to various other and more native communities. Differences in language, culture, local rules and tools, motivations, etc., are more often negative than positive factors for virtual workspaces success. We also have to deal with the staff and especially of top level management mostly reluctant to new information technologies which are seen as cost centers with uncertain return on investment.

We had to deal with security issues too. We had to take privacy protection, and of course, intellectual and industrial properties preservation into account. Working with industrial or research partners implies that the CEA is able to demonstrate an efficient security policy to protect everyone interests, with respect with agreements, rules and laws.

This is particularly difficult because there is a fundamental antagonism between sharing and protecting information. On one side, “Security by obscurity” is a very classic reflex of security officers but on the other side, knowledge managers and chief project officers want to enlighten and disseminate information.

We decided to offer four different hosting configurations according to four security levels to meet both security constraints and usability:

- “Internet” to host low security levels spaces
- “Extranet” to host inter-organization collaboration spaces with quite high security requirements, including data transmissions ciphering
- “Intranet” spaces for internal collaborations deployed on our intranet
- “Classified” spaces deployed on our classified private network.

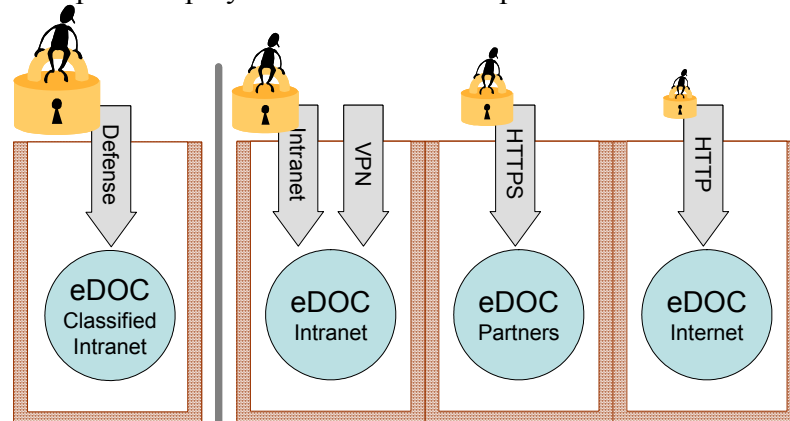


FIG. 3. Four different hosting configurations and four security levels

MARKET STATE OF ART

The eDOC project has begun with a quite important evaluation and benchmarking campaign. This campaign had two main objectives : 1 - detect the set of functionalities currently proposed in “collaboration” market solutions and 2 - find which of these solutions offers the best compromise between usability, cost, robustness and flexibility.

As we began the evaluation survey (in 2002), document and collaboration tools market was moving very quickly (it is not consolidated yet and it will not before years). First, we benchmarked market leaders’ solutions to have a good measure of cost/functionalities ratio.

Then, we looked at some new proprietary vendors surfing on the Internet e-business hype curve and then, to complete the round robin, we had a look at different open source web content management frameworks based on PHP, Java and Zope.

Market's top-notch commercial solutions design choices are more or less tied to one big application aiming at the whole organization. This leads to license pricing models that converge towards an economically acceptable per user price when buying a large number of tokens in one batch. Our context, in the case of a research and development organization, is quite different:

- many different cross-organization projects (for example, in the 6th EU's framework program, NoEs are involving ten or more partners with a large spectrum of financial, technical and organizational contexts). In the case of a commercial solution, we not only need to provide a technical solution but also to define a chargeback policy. It should be uncoupled from per user policy implemented by solution editors as some distant users are using the system on a daily basis while some others are using it from time to time. Defining significant and acceptable parameters driving this policy is very difficult.
- high reactivity level for new features implementation : complete commercial solution building project involves big amount of consultancy and therefore are intrinsically long projects, but needs are moving along...

Evaluation survey and deployment business model estimation conclusions were that even if some proprietary solutions were fonctionnally interesting, these were still too expensive for a large scale deployment in the context of research and development activities. Because of both the innovative aspects of these technologies and the complexity of the processes supported by such tools, its Return on Investment is impossible to evaluate and therefore, top management is reluctant to top-down sponsoring. At this time, it seemed obvious that we had to find another projet model to convert upfront capital investment into ongoing pragmatic and "on-demand" efforts. Open source free solutions appeared as a potential outsider solution.

As most open source solutions providers' business model is to mutualize technical layers and sell specific end-users solutions implemented on top of them, most of them were frameworks. Technical and functional modularity, code strength and configuration latitude were looking from naïve to very professional depending on project maintainers' long-term vision and open-source organisational know-how. Our open source functional and technological survey pointed out Zope-based CPS (Collaborative Portal Server) which includes:

- an end-user solution, implementing most of the desired functionalities (CPS)
- a full documentary framework (CMF), allowing high abstraction level for new functionalities implementation
- an application server including strong role-based security logic and object-oriented database persistence layer (Zope).

WHAT IS EDOC : MAIN FEATURES

The key concepts implemented in eDOC are documentary objects, workspaces and publication spaces.

A documentary object covers a functional spectrum from a file and its metadatas (Dublin-Core compatible) to complex forms and is often seen as 'content'. A documentary object has different viewpoints allowing to see it or edit it according to current user role. A workflow engine may drive user roles in its context, according to pre-defined workflow rules. Default basic workflows applied to documentary objects are versioning workflow and publication workflow. Documentary objects are subject to comments from users gluing a document-contextual forum feature.

A workspace is a role-driven container, containing documentary objects or other workspaces. Identified users may have three different local roles inside a workspace: 'visitor' may only view or comment contents while 'members' can view and modify contents and 'managers' may also give local roles to users. Mailing-list feature may listen to what happens inside a

workspace, sending emails to members when documentary objects are created or modified. Forums may be switched on *if* needed and *when* needed by workspace managers, allowing its use in a contextual way. While most collaborative platforms offer forums, they implement them as top-level feature somewhat disconnected from workspace context.

A publication space implements a workflow-driven publication process. Documentary objects are created inside a workspace, then submitted for review in one or more publication spaces. Users may have two different roles in that context : ‘Readers’ may submit and read contents while ‘Reviewers’ may read, submit and review contents. Mailing list may listen to what happens inside a publication space too, sending emails to reviewers when contents are submitted or emailing readers when contents are validated. Newsletter feature allows sending digests to subscribers periodically (daily, weekly, monthly), summarizing newly published contents.

Some other features are application-wide, including an access-rights driven search engine (what you see is what you are allowed to see) and a configurable directory which covers both applicative technical datas (login, password) and ‘community of practices’-centered data (laboratory website url, phone number, centers of interest, last publications, scientific background, etc).

These elements are rendered on screen as a configurable portal, allowing to show application navigation means and user relevant informations at first sight. Per-project portal customization leads to a better appropriation by project leaders and end-users.

Search

Shared

Latest

Next events

Static information contents

Navigation

FIG. 4. Example of eDOC application: main portal view of the “Nano2Life” European Project

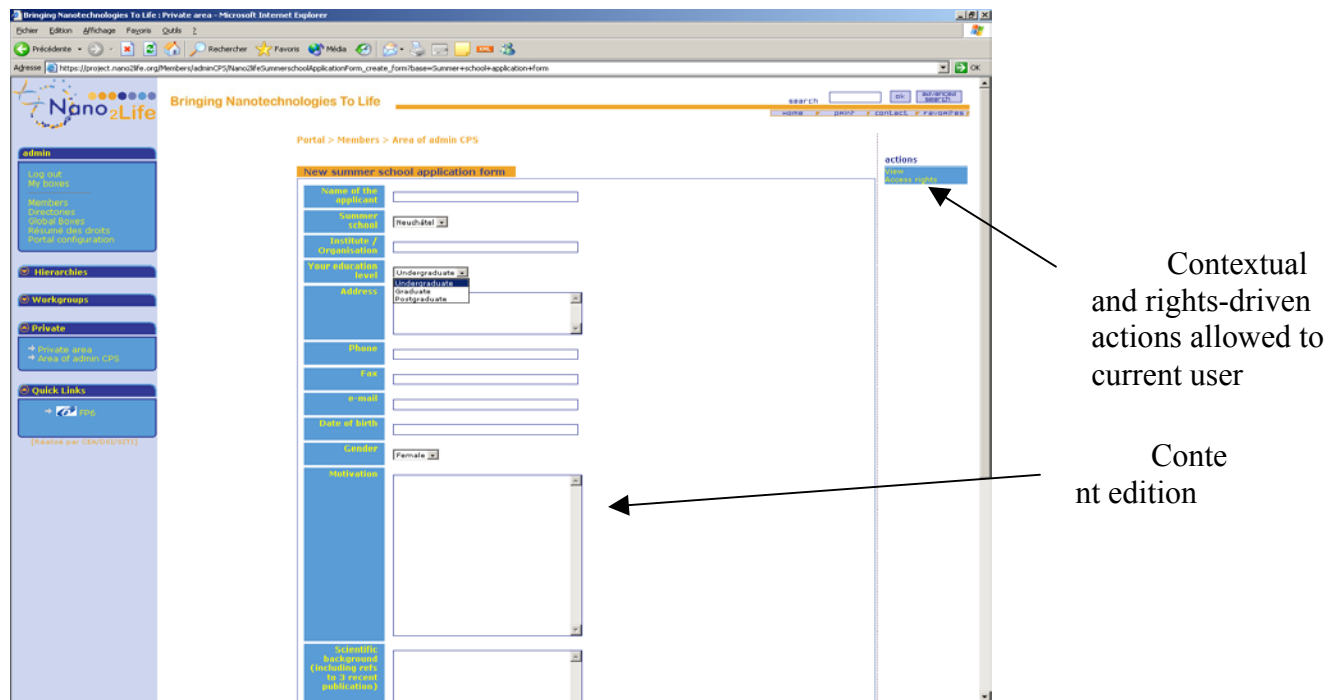


FIG. 5. Example of eDOC application: documentary object (summer school application form) edition view

CURRENT STATE OF USAGE

We host currently 50 eDOC applications and this number increases continuously. Those applications are mainly covering :

- Inter-organization projects : these applications are project tools aimed at helping research communities. They make full use of communication tools and documentary features. These projects are a strong source of new generic features. We included new features such as shared meeting managing application and generic tabular data managing tool (to provide specific finance and deliverables managing tables to project leader, in order to let them collect data from their partners).
- Document repositories: these are department-wide deployments. They use almost exclusively documentary functionalities and no communication tools. Version control features should be a key aspect but our return on experience shows us that it is not widely used. These applications are mostly seen as finalized document repositories more than document production assistants.

Among those use-cases, we find all the 6th european framework program projects coordinated by the CEA. Some externally coordinated projects use eDOC as a project Knowledge Management Platform, for the entire project or only for some workpackages.

THE EDOC PROJECTS (GOALS AND DEPLOYMENT STRATEGY)

eDOC project was designed as generic but simple solution. Being a generic solution, a large part of deployment work is a face to face work between us and project leaders. We map generic features on specific project needs as tightly as possible and inactivate useless features.

eDOC is deployed in full ASP mode including both technical and functional services (codebase maintenance, data backup, functional consulting,...). As a hosted service provider, we keep a unified codebase running multiple application. Therefore, we need to identify early what project can use eDOC and what projects should use a specific application because they have very specific needs.

Whereas many configuration tools are delegated to project leaders, this is where the application complexity mostly resides. We offer extensive documentation and training to

project leaders. A four hour long training session should give them full control on the application. Project leaders then act as functional administrators, giving fellowship training to project members. We also provide on-line video training sessions to project members, each video covers a generic use-case ('how to change my password', 'creating a document', 'reviewing a document', etc...).

eDOC is technically and functionally modular. This allows us to add new features as modules. We try hard to keep a global view on all applications deployed and detect what should become new generic features when we meet project leaders. Most of the time, there is a need in one project and its project leader agrees to help us in the specification process. Depending on the feature, we implement it "as is" or make it very configurable to meet a larger functional spectrum. Then this module is deployed on this specific application and made available to all other project thereafter.

RETURNS ON EXPERIENCE

The wide range of eDOC use-cases allow us to give some advises or present some facts to facilitate and promote collaboration web platform use.

First of all (and it looks obvious), the platform launch must be motivated by a real collaboration need. Even if it is very very "sexy", the platform never generates the need of use by itself !

Among the various information management tools, a rich members directory (with skills, identity pictures, roles in the project) is useful. It facilitates "knowing each other" when physical meetings are rare. However, physical presence can't be fully substituted by "virtual" communication tools.

Successful collaborative platform implementations need a strong project leader commitment organized with both IT and scientific team and KM team. The animation effort must not be underestimated. Self organisation is still a theoretical concept and in the real world, collaboration rules and guidelines have to be defined and applied. These guidelines also helps in mapping generic features to specific project processes.

Keep it simple! Project leaders often tend to define ambitious functional objectives, but this leads to rigid applications, difficult to use and maintain. Automating processes is tempting but while publication and reporting are widely used features, multi-stage validation and versioning are still rarely used. This is not a technical limitation, but a side effect of culturally heterogenous communities processes.

Worldwide public communication and internal community communication and collaboration are two different objectives and should lead to two different projects :

- a website as one of the communication tools to implement the communication strategy
- a collaboration platform as one of the management tools.

eDOC project shows that it is possible to build a generic low-cost service offer that covers a large part of specific projects needs. For more specific aspects, using a high-level object-oriented application server (Zope application server) allows a better reactivity to implement add-ons.

REFERENCES

- [1] CHAILLOT, M., ERMINE, J.C., Le livre de connaissance électronique, Document Numérique, vol.1, n°1, pp 75-98, Hermes (1997)
- [2] VAN CRAEYNEST. J.M., CHARLOT. J.M., MALAVIEILLE. D., BIGEON. P., CHAILLOT. M., Traçabilité des connaissances avec MKSM™, Journée Nîmes'TIC 2000, pp 12-22, Septembre (2000)
- [3] DRAKOS. N., Energy Agency Gains From Open-Source Collaboration Project, Case Studies, CS-21-1176, GARTNER RESEARCH, November 26 (2003)

- [4] VAN CRAEYNEST. J.M., Managing Knowledge at CEA : implementing documentary workspaces with eDOC, Topical Day on Knowledge Management in a Scientific Environment, pp 13-14, SCK.CEN, Mol, Belgium, March 25 (2004)