

# *Technologie du "Knowledge Management"*

ou

## *Proposition d'approche pour la mise en place d'une infrastructure informatique pour le partage des expériences dans l'entreprise*

*Christophe DUPRIEZ, R & D Manager, Destin*

La création d'une nouvelle discipline telle que le "Knowledge Management" (ci après "KM"; je vous propose la traduction "Gestion du partage d'expérience") a pu sembler un phénomène de mode. Je pense plutôt qu'elle est naturelle tant il est nécessaire de créer des bases conceptuelles claires pour une activité essentielle de l'entreprise (et de toute autre organisation): conserver et développer son potentiel de connaissance, de décision et de création, traiter les informations qui sous-tendent ce potentiel, automatiser au maximum ces traitements pour les rentabiliser.

L'équilibre entre la valeur des informations et le coût de leur traitement est un dilemme vieux comme l'Histoire. C'est l'avènement d'outils informatiques, de plus en plus économiques et connus de tous, qui entraîne une remise en question des habitudes passées: nous sommes obligés de réinventer l'"économie" de l'information dans l'entreprise (le "business model" cher aux Américains) et donc de redessiner (re-engineering) les méthodes et les procédés.

En tant qu'ingénieur des systèmes d'information, ce sujet me passionne et je vais essayer de d'ouvrir des pistes de réflexion utiles pour celui qui se trouve devant un cas concret. Je voudrais toutefois éviter de promouvoir des méthodes précises: je pense que la révolution induite par les technologies de l'information est loin d'être terminée et que l'heure est encore à la réflexion des ingénieurs et pas encore à l'application directement utilisable par des "mécaniciens". Vu les efforts actuels, ceci viendra sans doute plus vite qu'on l'imagine.

Qui dit "ré-ingénierie", dit retour aux principes de base:

- la valeur d'une information est directement liée au coût d'une mauvaise décision due à son absence (à fortiori, au bénéfice de la bonne décision due à sa présence)
- dans l'utilisation d'une ressource informationnelle, l'utilisateur recherche deux choses:
  - \* l'exhaustivité (recall): la certitude d'avoir une information complète au vu de son problème
  - \* la précision: recevoir une quantité d'information telle que l'effort de lecture (ou le risque de se distraire!) ne vienne pas annuler la valeur de l'information requise
- les investissements dans l'organisation de l'information à priori doivent être établis en tenant compte de la valeur potentielle de l'information mais aussi des coûts d'interprétation / filtrage de l'information à posteriori

La définition de circuits de collecte, d'organisation et de redistribution de l'information, qui répondent à ces équilibres dans un contexte d'évolution technologique, est tout un défi!

Cette définition se fait en analysant:

- 1) les données internes déjà existantes
- 2) la collecte des expériences
- 3) les informations réellement utiles
- 4) les données et les traitements nécessaires pour créer ces informations

### 1) Les données internes déjà existantes

L'exploitation de données qui circulent déjà dans l'entreprise n'est pas seulement utile parce qu'elle réutilise des informations existantes déjà "payées". Elle est aussi essentielle car rien n'est plus néfaste à l'image-clientèle et à la "culture d'entreprise" que l'impression de devoir tout le temps répéter les mêmes choses (*quelle impression vous donne un fournisseur qui vous redemande votre numéro de téléphone? le bureau-conseil qui remplace un intervenant par un autre qui vous redemande les informations déjà fournies au précédent? l'employeur qui vous redemande votre numéro de compte financier? votre date d'embauche?*).

Le faible coût de la réutilisation des données existantes (et la nécessité de devoir être capable de le faire) permet de ne pas se demander à priori à quoi ces données peuvent servir (si elles existent déjà, c'est qu'elles sont déjà utiles quelque part, non?) mais d'abord de se demander comment les rendre réutilisables.

Ici plusieurs remarques s'imposent:

- le KM n'est pas que l'apanage des grandes sociétés "conseil" qui doivent rendre leurs experts "remplaçables": nous tentons ici d'élargir la problématique à tout type d'organisation.
- l'entreprise doit avoir un catalogue de ses "gisements" d'information: ce catalogue doit indiquer la nature et la localisation des informations, leur validité et leur étendue.

Faire du KM est impossible sans une parfaite organisation:

- de son **informatique de gestion**, ce qui comprend:

- \* la gestion de la production
- \* la gestion des ventes (et des clients)
- \* la gestion des achats (et des fournisseurs)
- \* la gestion financière

- de sa **bureautique**, ce qui comprend:

- \* le traitement de texte
- \* la production de documentation
- \* les analyses de données, tableaux et graphiques
- \* les bases de données personnelles
- \* le courrier électronique

- de sa **gestion documentaire**, ce qui comprend:
  - \* les manuels de procédure
  - \* le répertoire de la terminologie de la société (thésaurus)
  - \* l'indexation des ressources documentaires
  - \* la gestion des traductions
  
- du **suivi de ses archives**, ce qui comprend:
  - \* la définition des types de document
  - \* l'application des calendriers de conservation
  - \* le suivi des contrats et des titres de propriété
  - \* la conservation des documents
  - \* la conservation des logiciels pour y accéder
  
- de la **gestion des ressources humaines**, ce qui comprend
  - \* l'évolution de l'organigramme de la société
  - \* la gestion des CVs et des profils d'aptitudes
  - \* le suivi du temps passé dans différents projets
  - \* les indices de performance, les évaluations
  - \* la gestion de la paie et des avantages sociaux

(Cette énumération n'est certainement pas exhaustive)

Le KM ne pallie pas aux faiblesses de ces systèmes et de ces disciplines. Au contraire, il s'appuie sur leur parfait fonctionnement pour aller plus loin en partant d'une démarche d'intégration. Toute démarche d'implantation du KM passe nécessairement par une évaluation de la qualité et de l'adéquation des systèmes existants.

- o l'information existante se divise en information structurée et non structurée:
- \* l'information structurée doit être documentée non seulement dans sa localisation, sa validité et son étendue mais aussi dans sa structure (dictionnaires de données)

Malheureusement, les systèmes de gestion ont souvent une "histoire" dont il faut tenir compte. La signification d'une zone peut changer d'un système à un autre. Par exemple "date d'embauche" peut être la date de 1ère embauche dans un système et la date de la dernière ré-embauche dans un autre.

*Un rôle à la frontière entre le KM et l'informatique est celui qui consiste à établir les "contrats" entre un système informatique, les autres systèmes informatiques qui interagissent avec lui et les différents services de l'entreprise qui utilisent ce système. Ces "contrats" identifient les données à fournir, les traitements effectués et les données disponibles pour réutilisation.*

L'évolution récente de l'informatique tend à modéliser les systèmes comme des composants (objets) ayant des attendus et des responsabilités vis à vis des autres systèmes et des utilisateurs. Des normes comme CORBA (Common Object Requests Broker Application) commencent seulement à s'imposer comme dénominateur commun multi-plateforme et multi-application. Si l'implantation d'une telle norme se généralise, le Gestionnaire KM aura une garantie d'accès pérenne non seulement aux données brutes mais aussi, dynamiquement, aux données produites par tous les systèmes informatiques de l'entreprise et de ses partenaires. *Réveillons nous, ceci n'est peut être qu'un rêve!*

- \* **l'information dite non structurée** est peut-être plus structurée qu'on le croit ou pourrait l'être: l'utilisation des zones d'indexation des documents (auteurs, titre, mots-clés, date, etc.) est essentielle à toute bonne gestion documentaire et à l'archivage (voir remarque ci-dessus sur les bases essentielles avant de songer au KM). L'utilisation de feuilles de style standard permet parfois le repérage automatique d'éléments d'information à l'intérieur de documents (frontières et titres de chapitre, récupération d'informations structurées dans des tableaux, extraction de renvois à d'autres documents, etc.). L'extraction automatique d'éléments structurels de documents est évidemment facilitée par l'application de règles de présentation générales à l'entreprise. On dit que cette information représente plus de 80% de l'information produite par une entreprise.

La généralisation des outils EXCEL, MS-WORD et Power-Point peut faire croire que la question de la représentation technique des documents est un sujet clos. Effectivement la compatibilité ascendante est assez bien assurée: *Microsoft assure assez bien la récupération des documents produits avec des versions antérieures pas trop anciennes (ou par des produits tiers) mais pas la compatibilité descendante (par exemple, sauver un document en format Word 6 avec Word 97 a causé bien des maux de tête à tout le monde)*. Un document sauvegardé il y a bon nombre d'années ne sera pas forcément récupérable (e.g. *on ne peut pas lire un fichier Wordstar avec Word 97*). Bref, la pérennité de l'information traitée par des produits externes (et donc en particulier les produits bureautiques) doit être gérée: *on ne doit pas changer de version d'un logiciel bureautique sans se demander si les documents en archive ne doivent pas être convertis et quelles sont les pertes de formatage entraînées par ces conversions*. On voit ici encore la nécessité d'archives bien gérées où:

- les règles de conservation des documents sont bien définies
  - le logiciel de traitement de texte associé au document est conservé opérationnel dans au moins un ordinateur.
- o Ceci est un peu une approche "Bottom-Up": *Que fait-on avec ce que l'on a déjà?*  
Corollaire: pour des informations jugées nécessaires (voir 2 et 3 ci-dessous), est il possible de trouver un biais pour que les utilisateurs la fournissent sans effort supplémentaire vis-à-vis de leurs opérations quotidiennes?

## 2) La collecte des "expériences"

Cette opération est très utile pour les firmes qui vendent le temps de leurs experts. Mais aussi:

- Il y a vingt ans déjà, l'industrie automobile mettait au point un programme permettant aux ouvriers de faire connaître leurs idées pour augmenter la productivité. Ce programme allait jusqu'à faire participer l'inventeur aux bénéfices de ses suggestions.
- Dans le domaine juridique, les bureaux d'avocats ont souvent des membres qui rédigent des articles de doctrine à usage interne ou publiés dans des revues spécialisées.
- On n'a pas besoin d'expliquer en détail le processus suivi par les bureaux de recherche et de développement, le corps médical ou les institutions scientifiques (articles ou communications diffusés internement ou publiés dans des symposiums ou des périodiques; prise de brevets).
- Plus récemment, les "Help Desk" se sont constitués des listes de questions ou de descriptions de problèmes avec en regard des textes expliquant les solutions à suivre.

D'une manière générale, l'expérience mise en oeuvre par tous les intervenants d'une entreprise, que ce soit pour atteindre les objectifs de gestion ou ceux de production, doit être "transférable" ou "remplaçable": le sort de l'entreprise ne doit pas pouvoir être pris en otage par un individu. Cette collecte doit être faite:

- \* par une identification systématique des domaines de connaissance essentiels au fonctionnement de l'entreprise
- \* par une analyse des moments où il est adéquat de demander au personnel les problèmes rencontrés et les solutions mises en oeuvre.

La saisie des "faits d'expérience" est donc intégrée au processus de production. Elle en augmente le coût et une évaluation serrée de ses retombées est nécessaire à sa justification. Toute une mécanique de gestion doit être mise-au-point pour savoir qui a contribué quoi et pour éventuellement apporter des primes à l'information réellement réutilisable / réutilisée.

Une base de données documentaires est un bon moyen de diffuser ces "articles" qui décrivent les "faits d'expérience". Un bon système de mots-clés (thésaurus) est essentiel pour l'indexation des articles. La liste des auteurs, le titre, le résumé sont des moyens classiques d'aider le lecteur à rapidement décider si l'article le concerne.

Il reste toutefois:

- que l'indexation de telles bases doit être facilitée au maximum (extraction automatique de la terminologie utilisée par exemple)
- que la rentabilisation de la constitution de telles bases d'expériences ou de "FAQ" (Frequently Asked Questions) passe par l'élargissement de leur utilisation à toutes les personnes concernées (quasiment tout le personnel de l'entreprise et bien souvent même les clients): chacun doit avoir le réflexe d'y faire appel avant d'entreprendre d'autres recherches (il est souvent nécessaire de mettre des contrôles ou des incitants à ce niveau).
- que les interfaces d'interrogation doivent donc être particulièrement soignées pour allier facilité d'apprentissage et efficacité d'interrogation. Les interfaces de type "Navigateur Internet" ont amené un grand progrès à la fois dans la facilité d'apprentissage et dans la "pré-connaissance" des utilisateurs. Il reste toutefois à en améliorer grandement le mode d'interaction tant au niveau de l'exhaustivité (élargissement des expressions recherchées) que de l'aide à la précision (spécification de critères additionnels).

Mais on est ici dans le domaine déjà existant de l'"ingénierie documentaire" où l'expert en KM trouvera l'aide dont il a besoin.

### 3) Les informations réellement utiles

Sans décisions bénéficiant de sa présence, l'information est sans valeur. Elle doit être livrée là où on en a besoin quand on en a besoin (désintermédiation maximum tant que cela n'affecte ni l'exhaustivité ni la précision; télécommunications; qualité de présentation; facilité d'utilisation). *Croire que les technologies "grand public" de l'Internet vont supprimer le besoin d'analyser:*

- *qui a besoin de ...*
- *quelle information ...*
- *sous quelle forme ...*
- *avec quelle précision / fraîcheur*

*est dangereux. Croire que l'analyse peut tout révéler à l'avance l'est tout autant.* Il faut donc s'assurer de répondre aux besoins à plusieurs niveaux:

- des outils pointus répondant aux besoins pré-identifiés;
- des outils "grands publics" offrant une solution universelle de dernier recours.

Il est important de se rappeler qu'en dehors de la réutilisation de l'information déjà produite par l'entreprise, il est nécessaire de calculer le coût/bénéfice de toute chaîne **collecte / traitement / diffusion** d'informations offertes aux utilisateurs. On doit nécessairement partir de leurs besoins, remonter aux sources possibles et identifier les traitements nécessaires pour que l'information soit présentée de manière utile.

Il faut noter que beaucoup d'espoirs ont été mis dans les EIS ("Executives Information Systems"). Certains espoirs ont été déçus, mais ils doivent continuer à évoluer pour répondre aux besoins cruciaux d'archivage, de consultation et d'analyse des données structurées de l'entreprise. L'adjonction des fonctions de "Data Mining" et surtout la possibilité pour l'utilisateur final d'obtenir une vue intégrée de toutes les données de l'entreprise manipulable à souhait, font des EIS de haut niveau des outils indispensables. Ici encore la norme CORBA est porteuse d'espoir mais la norme SQL apporte déjà l'essentiel: avez-vous étudié la possibilité d'accéder au moyen des EIS déjà implantés (implantables) aux bases de données SQL sous-jacentes à vos applications de gestion des archives, de gestion des ressources humaines, de gestion documentaire? Quelles informations pourraient être ajoutées facilement aux systèmes existants pour répondre aux vraies questions de la Direction?

On évoque de plus en plus souvent des outils de "Business Modeling" permettant d'explicitier la réalité dans laquelle évolue une entreprise et d'établir les informations utiles pour contrôler cette réalité ou s'y adapter. Cette démarche permet d'appuyer le processus de définition des informations utiles au fonctionnement de l'entreprise.

Ceci est une approche "Top-Down": *De quoi a-t-on besoin pour répondre aux contraintes de la réalité, aux besoins de la production et aux vœux de la direction?*

Reste à trouver "Comment fait-on pour répondre aux contraintes ou aux vœux imprévus ou imprévisibles?"

Ici les technologies "grand public" telles que les index généraux (moteurs de recherche Altavista par exemple) ont une grande utilité comme "parachute de secours".

#### 4) Les données et les traitements nécessaires pour créer les informations

Une fois les questions circonscrites, il devient plus facile de définir les éléments d'information qui permettent d'y répondre puis de déterminer des sources possibles pour ces éléments. Mentionnons:

- les banques de données (ou services d'information) externes:
  - il y a-t-il des questions répétitives ? peut-on automatiser la livraison des éléments requis et leur intégration dans les banques de données de l'entreprise (formulation des requêtes de recherche; récupération des résultats; identification des ajouts / modifications / suppression; reformatage/recodage) ?
  - pour les questions ponctuelles, il est intéressant de les conserver (questions et résultats) et de les indexer: les personnes d'une entreprise concernées par le même sujet peuvent se découvrir par les questions qu'elles posent à l'extérieur, la communication peut s'établir, les expériences peuvent se partager.

*Formel vs Informel, Structuré vs Non structuré, Répétitif vs Ponctuel, Top-Down vs Bottom-Up : des outils différents sont nécessaires pour chacune des approches, à chacun des niveaux.*

- la numérisation (OCR) de documents:

La productivité de cette technique, par rapport à la resaisie manuelle, n'est pas facile à obtenir: l'injection de "structure" dans le document numérisé est difficile à automatiser et le taux d'erreur pour des documents-source non normalisés est important. Une solution intermédiaire: présenter l'image du document mais chercher sur le texte reconnu. Il est de toutes manières beaucoup plus intéressant de s'assurer de la disponibilité des documents désirés sur support électronique.

- la saisie manuelle: cette opération, surtout si elle pré-suppose une enquête pour obtenir l'information, est lente et couteuse. Mais, c'est ce que nous faisons maintenant tous en tapant nos rapports, mémos, E-mails, etc. Y avons nous vraiment gagné par rapport au dictaphone et à la secrétaire/dactylo efficace? Quelle est la perte de productivité due à l'attention parfois exagérée aux détails de présentation?

L'intégration de la donnée obtenue manuellement, dans les banques de données de l'entreprise, est évidemment essentielle. Le fait-on suffisamment pour tout ce que nous tapons? Les politiques de catalogage (gestion documentaire) et d'archivage sont elles bien appliquées?

Une fois les données colligées, elles doivent être conservées. Ce sont soit des documents, soit des données intégrables dans des bases de données. Ici deux niveaux de systèmes doivent être prévus:

- le niveau "structuré" qui intègre les informations colligées de manière ad-hoc dans l'architecture des données de l'entreprise
- le niveau "non-structuré" qui conserve les informations sous la forme de documents indexés comme tout autre document concerné par la gestion documentaire.

## Conclusion quant aux technologies qui soutendent le KM:

Le KM s'appuie sur les systèmes déjà en place pour:

- l'informatique de gestion,
- la bureautique,
- la gestion documentaire,
- le suivi des archives,
- la gestion des ressources humaines.

Il est certainement utile de vérifier avec les concepteurs de ces systèmes ce qu'ils proposent comme évolution pour appuyer une démarche d'implantation du KM.

L'infrastructure matérielle et logicielle permettant les échanges d'informations, entre les utilisateurs et les systèmes et entre les systèmes eux-mêmes, est évidemment essentielle. On ne doit pas sous-estimer les bandes passantes nécessaires aux "Intranet" (un facteur 10 par rapport aux réseaux "données" et 2 par rapport aux réseaux "bureautiques" n'est certainement pas exagéré). Le passage à la norme TCP-IP peut aussi alourdir le fonctionnement des réseaux.

Les banques de données documentaires intégrant un thésaurus et le texte intégral de courts rapports, mémos ou communications scientifiques, sont un outil très correct pour la conservation et la diffusion des "faits d'expériences" recueillis auprès du personnel. Les navigateurs Internet ont rendu ces banques de données beaucoup plus conviviales à l'emploi. Il reste toutefois de grands progrès à faire au niveau de l'assistance à la recherche.

La norme CORBA est une norme fédératrice qui, si elle s'impose, permettra d'établir des contrats de livraison d'informations validées avec les différents systèmes déjà en place.

Les systèmes EIS en place dans l'entreprise peuvent déjà rendre de grands services en permettant l'analyse de données obtenues des différentes applications par des requêtes SQL. Les outils de "Data Mining" en sont le complément naturel.

Les logiciels de "Business Modeling" facilitent le travail d'identification des informations utiles aux décisions de l'entreprise.

Les technologies "grand public" telles que les index généraux (e.g. moteurs de recherche Altavista) ont une grande utilité pour répondre à des besoins imprévus.

Le KM, c'est finalement très concret:

- s'assurer de la cohérence des systèmes d'information dans l'entreprise et de la possibilité d'en réutiliser les données produites
- s'assurer de la normalisation de la terminologie, de sa traduction et de son utilisation pour indexer les informations non structurées
- encourager (au besoin financièrement) la rédaction d'articles sur des "faits d'expérience" qui sont partagés avec toutes les personnes concernées
- identifier les besoins en information et les données à obtenir
- s'assurer du stockage efficace des nouvelles données
- s'assurer de l'accessibilité des "gisements d'information" et de l'utilisation de bons outils pour les exploiter ou en présenter le contenu.
- encourager l'utilisation des nouvelles ressources en information