

L'ID au secours de la GED

Jean-Claude Humbert
Docteur en Médecine
Laboratoire SPI-EAO
Faculté de Médecine de Nancy
B.P. 184
54505 Vandœuvre-lès-Nancy
humbert.j-claude@wanadoo.fr

Laurent Vigneron
Maître de Conférences en Informatique
LORIA – Nancy Université
Campus scientifique
B.P. 239
54506 Vandœuvre-lès-Nancy
laurent.vigneron@loria.fr

Résumé : L'accroissement du volume de l'information électronique à mémoriser concernant les systèmes d'information et de communication des entreprises et administrations est devenu un réel problème depuis quelques années. L'utilisation efficace de cette information est un casse-tête qui nécessite la mise en place de systèmes de GED (gestion électronique documentaire). Mais ces systèmes s'avèrent souvent insuffisants, les outils existants manquant de souplesse et d'adaptabilité aux modèles souhaités par les utilisateurs. Ainsi, tout comme l'intelligence économique est venue au secours de la gestion des systèmes d'information dans les entreprises, il est devenu indispensable de tenir compte d'une intelligence documentaire (ID) lors de la mise en place d'un système de GED.

1) La gestion électronique documentaire

Nous rappelons ci-dessous les caractéristiques et les enjeux de la gestion électronique documentaire, selon Jean-Yves Prax¹.

a) Les enjeux de la GED

À l'origine de la GED se trouve la volonté de créer une *connaissance collective* dans l'entreprise, véritable enjeu pour chacun. En effet, il est important pour un individu, comme pour un ensemble d'individus, de se préoccuper :

- des tendances de son domaine,
- des nouveaux produits,
- de la coopération ou de la concurrence,
- de l'émergence de nouvelles technologies,
- de la mémorisation du savoir et du savoir-faire.

L'acquisition de ces connaissances pouvant avoir un impact important sur la vie de l'entreprise, il est de première importance

- de mesurer ce que l'on sait et d'évaluer le risque encouru à cause de ce que l'on ne sait pas,
- de concevoir des organisations qui utilisent le patrimoine de connaissance et qui sont capables de le produire, de l'étendre, de le faire vivre en permanence,

¹ Jean-Yves Prax, *La gestion électronique documentaire, manager les flux d'information dans l'entreprise*, Interédition, 1999.

- de mettre en place des dispositifs sociaux ou techniques pour faciliter cette création, l'apprentissage et la propagation des savoirs.

Pourtant, un simple *constat* des pratiques montre le manque d'intérêt pour la gestion de cette connaissance collective. Nous produisons souvent un stock de normes, de valeurs, de procédures inutiles pour résoudre les problèmes nouveaux ou effectuer des choix. La fonction de mémorisation (documentation) devient une charge de travail supplémentaire, administrative et sans réalité économique. Cette fonction disparaît en premier lieu dès que la production s'accélère ou que les budgets diminuent. Elle est souvent reléguée au rang d'archives.

b) Organisation d'une GED

Lors de sa mise en place, la GED peut se concevoir comme la composante d'un système d'information et de communication. Le pilotage de l'organisation s'appuie alors sur deux facteurs structurants :

- le projet, moteur de l'action et de la convergence entre les acteurs ;
- les méthodologies, qui assurent la cohérence dans l'action.

L'organisation repose sur l'articulation entre la compétence et les savoir-faire en lieu et place des moyens. Souvent les besoins et les solutions ne peuvent être définis à l'avance. On ne peut pas exprimer des besoins par rapport à des usages que l'on ne connaît pas encore dans un contexte d'innovation, d'où la nécessité d'une certaine souplesse en même temps que la capacité d'adaptation et d'évolutivité.

La GED doit être capable de produire les connaissances et de mobiliser l'intelligence de manière permanente de façon à pouvoir formuler des réponses inédites à des problèmes nouveaux et imprévisibles.

Dans ce contexte la capitalisation des connaissances devient un produit de transformation des informations, du savoir-faire et des idées acquises dans l'expérience. Elle consiste à identifier, formaliser, et conserver la mémoire des activités tout en rendant cette mémoire accessible indépendamment des acteurs qui l'ont créée et de manière pertinente par rapport à un contexte d'intérêt donné.

Le concept de pertinence introduit la notion de temps. Une rupture temporelle constitue souvent un facteur de perte de connaissance. La capitalisation des connaissances a pour objectif le maintien du potentiel d'utilisation d'une connaissance et de sa pertinence dans le contexte futur.

Mais *quelles connaissances capitaliser ?*

- Comment s'assurer que les critères d'aujourd'hui (filtres, choix des documents référentiels, choix des descripteurs) correspondent aux préoccupations de demain ?
- Quelle est la demi-vie d'une connaissance ? Une capitalisation dynamique fait appel à la notion de flux et de processus. La création de savoir et de savoir-faire s'effectue de manière collaborative. Il faut distinguer :
 - o les flux entrants (informations acquises de l'extérieur, courrier, presse) ;
 - o les flux sortants (informations produites et diffusées pour faire connaître les savoir-faire, produits, méthodes ; les flux internes).

Il est nécessaire de mettre en place une intelligence documentaire (ID), qui devra non seulement stocker ces connaissances, mais également définir des dispositifs permettant

d'échanger et de partager les connaissances en permanence. Le rôle de la GED dans un processus d'ID est de piloter les flux entre producteurs et consommateurs, en s'appuyant sur les nouvelles technologies de l'information et de la communication (NTIC) qui ont connu un essor impressionnant ces dernières années :

- le document électronique est devenu multimédia, intégrant textes, graphiques, images, sons, vidéos ;
- les mémoires optiques et numériques améliorent sécurité et capacité ;
- des référentiels se sont imposés autorisant les échanges et assurant la pérennité des développements ;
- les interfaces hommes-machines se sont simplifiées ;
- les technologies informatiques de télécommunication et audiovisuelles se sont fusionnées ;
- la généralisation des réseaux fait émerger des concepts de groupes de travail virtuels.

c) Mise en place d'une GED

L'ID est devenue une composante des systèmes d'information dans lesquels elle introduit des fonctions de gestion ou de traitement des documents « vivants ». L'utilisateur est à la fois producteur et consommateur d'informations. Les exemples sont nombreux : blogs, wikipedia, HAL,...

Mais les *difficultés* sont nombreuses :

Toute organisation est naturellement hostile au changement et développe des routines défensives.

La méconnaissance de l'offre et de l'étendue de ses possibilités pour l'utilisateur rend nécessaire un accompagnement et une *expertise*.

Le mode de pilotage doit définir la mise en place de l'outil et de son appropriation, du projet (processus d'action qui génère les usages) et des valeurs directrices sous-jacentes au service d'une stratégie du changement, la prise en compte de l'évolutivité de l'environnement, du contexte et des besoins.

Grâce aux nouveaux outils de développement rapides (RAD) et à la disponibilité sur le marché de composants de base intégrables sous forme d'API (OCR, pilote de scanner, indexation), l'ID va se concevoir comme un ensemble de *maquettes jetables* et de processus itératifs incrémentaux.

Chaque étape d'expression des besoins est validée par une maquette

- qui structure la réflexion des acteurs et les aide à modifier leurs besoins au cours du temps en prenant en compte les champs de possibilité,
- qui sert de point d'accroche aux autres modules,
- qui constitue une réalisation concrète pour convaincre les hésitations.

Les *objectifs* sont variés, mais complémentaires :

- objectifs relatifs au stockage : réduction du volume de transactions et de l'encombrement, accroissement de charges, diminution des risques (vols, incendies, pertes), diminution des duplications ;
- objectifs liés à la gestion : diminution des délais de recherche, accès à l'information par critères multiples, amélioration de la confidentialité, réduction des coûts ;
- objectifs liés à la sécurité de la diffusion : diffusion interne ou externe, amélioration de l'accès à l'information et de sa pertinence, sécurité d'accès, accès partagé, accès distant, décentralisation du traitement ;

- objectifs de communication et de coopération : travail collaboratif, communication rapide et continue, structure des échanges, procédures d'orchestration des échanges ;
- autres objectifs : objectifs intermédiaires ou d'étapes, objectifs cachés sous-jacents, objectifs induits.

En résumé, l'intelligence documentaire consiste à mettre en place une structure logique, d'utilisation simplifiée qui est également personnalisable, adaptative au besoin individuel, et évolutive.

2) L'Intelligence Documentaire (ID)

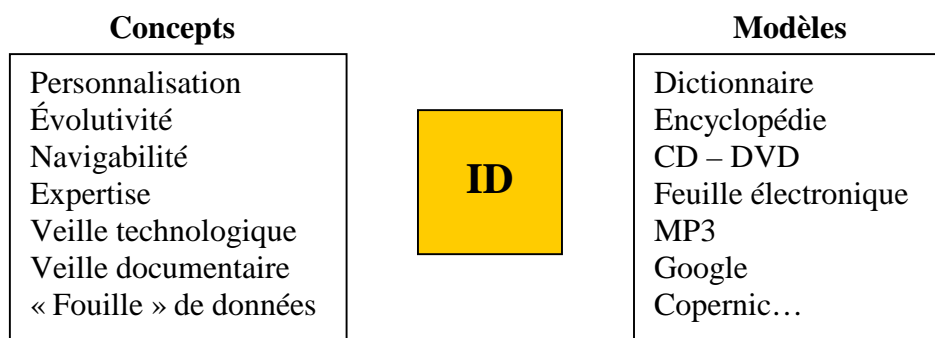
L'intelligence documentaire est constituée par l'ensemble des concepts, des outils, des méthodologies et des pratiques permettant de mettre en relation, de façon pertinente, différentes connaissances et informations dans la perspective de la maîtrise et du développement de la dynamique documentaire. Cette mise en relation implique en particulier :

- une mobilisation des hommes ;
- un traitement et une analyse de l'information et de la connaissance orientés vers une finalité opérationnelle ;
- une circulation efficace des informations et des connaissances au sein des organisations concernées.

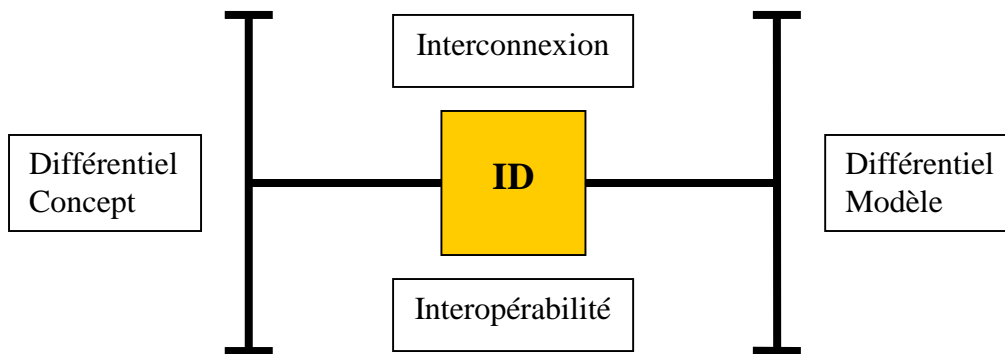
Ainsi, si l'ID peut être comparée à un véhicule tout terrain : le moteur représente les organisations socio-professionnelles, la roue avant-gauche la recherche fondamentale, la roue avant-droite la recherche appliquée et le transfert de technologie, la roue arrière-gauche la formation et l'évaluation, et la roue arrière-droite l'aide au diagnostic grâce à l'exploitation de données (par exemple dans le domaine médical, l'utilisation du DMP, dossier médical personnel) et l'utilisation d'algorithmes décisionnels. Enfin, la direction du véhicule peut se concevoir comme étant de nature politique, pour impulser une direction et une trajectoire.

Comme dans le cadre de l'intelligence économique, nous retrouvons

- d'une part les concepts : personnalisation, évolutivité, navigabilité, expertise, « fouille » des données, veille technologique et documentaire,
- d'autre part les modèles : dictionnaires, CD, DVD, MP3, feuille électronique, documentation web (Google, Copernic...).

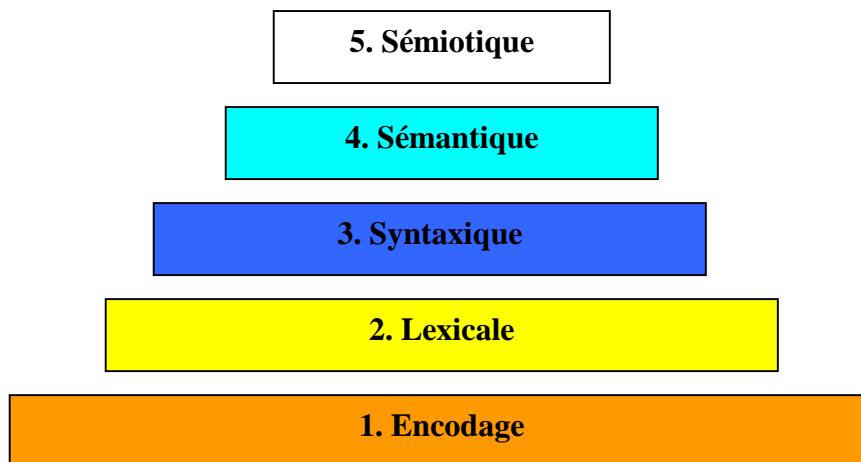


À ces notions de base vient s'ajouter un système de régulation permettant de maîtriser les stratégies de développement de l'ensemble documentaire. À l'image de certains procédés mécaniques, il est possible de concevoir un différentiel pour les modèles, un différentiel pour les concepts, et un différentiel intermédiaire faisant intervenir les notions d'interconnexion et d'interopérabilité.



L'interopérabilité, comme le souligne Salah Baïna², est nécessaire lorsque deux acteurs communiquent ou s'échangent des expressions, phrases, données : on peut alors identifier plusieurs types de communications, selon la compréhension mutuelle des éléments échangés. Jérôme Euzenat³ a ainsi décrit différents niveaux de communication :

- Niveau encodage : aptitude à identifier la représentation des caractères.
- Niveau lexical : capacité de reconnaître les mots ou les symboles.
- Niveau syntaxique : capacité de structurer les échanges en phrases, formules ou assertions.
- Niveau sémantique : capacité de reconstituer le sens des représentations échangées.
- Niveau sémiotique : faculté de restituer le sens pragmatique et le contexte des phrases échangées.



Cette représentation en couches exprime implicitement le fait que chacun de ces niveaux nécessite de remplir les conditions des niveaux inférieurs. Les trois premiers niveaux (encodage, lexical et syntaxique) sont généralement atteints par le biais de l'utilisation d'un langage commun entre les différents systèmes de communicants. La communication sémantique permet quant à elle, une compréhension et une interprétation unique des données, informations et services échangés entre systèmes communicants (émetteur et récepteur). Ainsi on appelle « interopérabilité sémantique » la faculté des systèmes d'interpréter les annotations échangées lors de la communication, comme par exemple assigner à chacun des éléments échangés le modèle correspondant ou une interprétation adéquate. L'interopérabilité

² Salah Baïna, *Interopérabilité dirigée par les modèles : Une approche orientée produit pour l'interopérabilité des systèmes d'entreprise*, Thèse de Doctorat de l'Université Henri Poincaré – Nancy I, 2006.

³ Jérôme Euzenat, *Towards a principled approach to semantic interoperability*, Workshop on Ontologies d Information Sharing, IJCAI, August 2001, Seattle, USA.

sémantique permet ainsi à des systèmes de combiner l'information reçue de la part d'autres sources d'information et de la traiter tout en conservant le sens.

En fait, il s'agit de gérer et de traiter dynamiquement des documents divers. L'ensemble de ces documents se classe en textes, images, tableaux, liens, etc.

La gestion documentaire connaît un développement important dans les secteurs industriels et commerciaux motivés par la contre-partie financière. Dans les domaines médicaux et scientifiques, l'aspect économique est moins évident à court terme, c'est la raison pour laquelle la mise en œuvre est plus laborieuse. Et en particulier en médecine, il faut faire face à des impératifs de sécurité et de confidentialité, et à des professionnels qui expriment souvent des réserves vis-à-vis des nouvelles technologies.

Certains secteurs économiques ont développé des logiciels de gestion documentaire adaptés à leurs besoins. C'est le cas pour les grands groupes de documentation :

- centres de documentations (INIST, NCBI, ONIM, HUGO),
- bibliothèques,
- médias (journaux, télévisions),
- éditeurs.

Au niveau individuel, si l'on prend en compte l'abondance et l'importance de la littérature, la gestion des documents devient tout aussi nécessaire et utile pour :

- les enseignants et les étudiants,
- les médecins et les scientifiques.

Or, il n'est pas concevable de développer des logiciels adaptés à chaque individu, mais il est possible d'utiliser les logiciels du commerce pour mettre en place une structure documentaire simple, adaptée, évolutive et personnalisée tout en se conformant à la définition qui en est généralement donnée.

3) Base de connaissances médicales

Face à ces difficultés de mise en place de systèmes d'information et de communication, nous avons créé une base de connaissances dans le domaine médical, permettant de bien comprendre les différents concepts, les enjeux qui en découlent et de tester et/ou simuler le fonctionnement de l'ensemble en vérifiant que cette base de données permette de répondre aux diverses questions rencontrées en simulant les différentes possibilités en fonction des applications.

Cette « plateforme expérimentale »⁴ permet aussi de vérifier que toute évolution technologique permet de modifier les techniques utilisées, mais également la stratégie dans ses phases décisionnelles.

Les documents (textes, tableaux, photographies, films,...) peuvent être répertoriés à l'aide de fichiers créés par des logiciels standards tels que Word, Excel ou Access.

Un document regroupe des liens vers l'ensemble des documents disponibles sur une matière (et donc contenu dans le dossier « document... »), chaque lien étant associé à un ensemble de mots clés. Ce fichier permet de poser des filtres sur les mots clé et donc de regrouper l'ensemble des documents associés à certains mots clé choisis par l'utilisateur. Ceci répond à

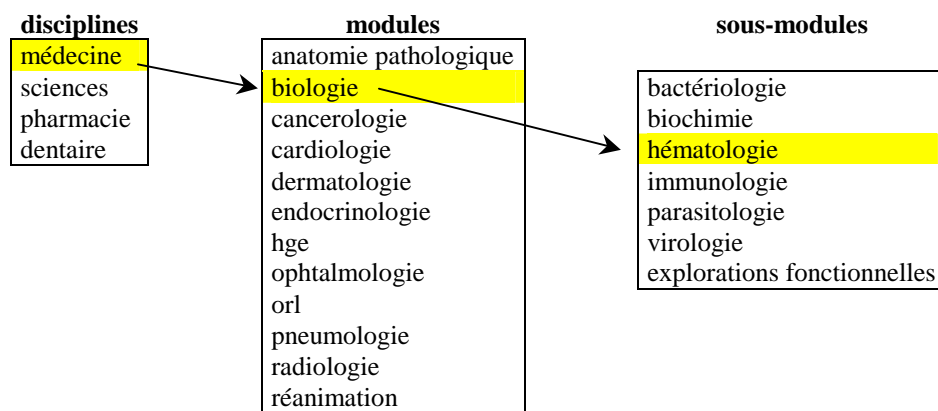
⁴ Cette plate-forme est consultable sur demande au premier auteur.

un vrai besoin de l'utilisateur de trier et de filtrer les documents lors d'une recherche, tout en restant efficace.

Un outil comme Access permet, par exemple, de lier les photos à un dossier photos par le biais de numéros d'ordre, de commenter ces photos, et donc de faire des tris et éventuellement de compléter ce commentaire par un lien avec un site internet. Un autre avantage d'Access est de pouvoir gérer des tableaux en provenance de différents fichiers. Enfin, les photographies et légendes peuvent être directement utilisées pour construire des diaporamas.

De la même façon, un traitement de texte ou un tableur tels que Word ou Excel permettent de gérer un ensemble de photos, de tableaux ou de textes. Les documents sont en outre hiérarchisés selon des niveaux des connaissances.

Exemple : hiérarchie permettant d'atteindre des cours d'hématologie.



Cette disposition hiérarchique facilite la mise en place d'interconnexions entre différentes disciplines et différents niveaux de connaissance.

La navigation est facilitée par une disposition de plans de travail sur des feuilles de tableur tel que Excel. En effet la mise en place d'un plan classique permet de « structurer » la connaissance, car ce plan correspond à une accumulation des connaissances et des expériences. Il devient alors très facile de lier les documents mis en formes précédemment, et de « naviguer » d'une discipline à l'autre. Un plan est constitué de documents divers provenant de dossiers personnels ou d'Internet.

Microsoft Excel - repertoire tm.xls

Fichier Edition Affichage Insertion Format Outils Données Fenêtre ?

B11 = tableaux virologie

	A	B	C	D
1	plan simplifié			
2	generalites	herpesvirus	virus des hepatites	retrovirus
3		herpesvirus simpley		vih
4		virus varicelle-zona		hbu
5		cytomegalovirus		
6		virus d'epstein barr		
7		6è herpesvirus humain		
8		cours virologie pasteur.htm		
9		Mononucléose infectieuse.doc		
10				
11		tableaux virologie		
12	plan detaille			
13	virologie generale	virus: structure et classification	multiplication virale	genetique et evolution des virus-facteur
14		structure	cycle de multiplication virale	methodes d'etude de la variabilite des virus
15		acide nucleique viral	cellule	souches virales et diées-souche de reference
16		capside	sensibilite	marqueurs phenotypiques
17		nucleocapside à symétrie helicoidale	permisivite	formation des plages de lyse
18		nucleocapside à symétrie icosaedrique ou cubique	cycle de multiplication virale	thermosensibilite du developpement
19		capside à symétrie complexe	phase d'eclipse	specetre d'hoté
20		enveloppe virale	phase de maturation	pouvoir pathogene experimental
21		classification et nomenclature des virus	particules virales neoformees	proprietes biologiques
22		virus à ADN	etapes initiales	resistance à l'inactivation par les agents physico-chimiques
23		virus à ARN	attachement	resistance aux antiviraux
24		exemples d'interaction virus-recepteur cellulaire	penetration	methodes d'etude de la variabilite antigenique
25		picomavirus	fusion des membranes virales et cellulaires	methodes d'etude de la variabilite genetique
26		virus de la grippe	endocytose apres liaison du virus aux recepteurs cellulaires	determinisme de la variabilite des virus
27		virus de l'immunodefience humaine	transfert du materiel viral à travers la membrane cellulaire	les mutations
28			decapsidation	les recombinaisons genetiques
29			expression et repliation du genome viral	les reassortissements genetiques
30			transcription	facteurs de virulence
31			proteines non structurales	generalites
32			proteines structurales	exemples
33			traduction	poliovirus
34			repliation	virus rabique
35			virus à ADN	cas de la grippe
36			virus à ADN double brin	consequences de l'evolution genetique chez les virus
37			transcription	apparition de souches nouvelles
38			phase precoce	facteurs d'origine humaine influençant la variabilite virale
39			phase tardive	immunodepression d'origine therapeutique ou infectieuse
40			repliation	utilisation de therapeutiques antivirales
41			initiation	certaines vaccinations
42			elongation	consequence de l'evolution sur le diagnostic
43				consequence de l'evolution sur la prophylaxie
44				consequences sur la therapeutique, la desinfe

Exemple de lien vers le document correspondant

Exemple d'une table des matières du sous-module : 'multiplication virale', du module : 'virologie'

Liste des différents modules de la base de connaissances

physiologie respiratoire / pneumologie / rhumatologie / urgences / virologie /

4) Conclusion

Toute entreprise ou administration doit maintenant gérer de nombreux documents électroniques, et cette gestion doit passer par différents niveaux :

- l'accumulation de documents, qui correspond à la phase proprement dite de *documentation*,
- la gestion de la documentation,
- et l'ID, qui correspond au système de communication.

Ainsi, l'intelligence documentaire devrait permettre de mettre en œuvre pour mieux les résoudre les questions posées actuellement pour la gestion des dossiers dans le domaine médical : sécurité, mobilité, efficacité, statistiques, évaluation, mémorisation, etc.

La plate-forme que nous avons mis en œuvre et qui est consultable sur demande devrait permettre de simplifier et de simuler l'approche du modèle évolutif décrit dans ce document.

Illustration de Philippe Geluck⁵ :

⁵ Album « Et vous, chat va ? », de Philippe Geluck, Casterman Edition.

