

Chapitre

Gestion du thesaurus



Durée 20 minutes

1. Définition :	2
1.1. Constitution d'un thesaurus	2
1.2. Les relations entre les termes	3
1.3. Exemple de thesaurus élémentaire	4
2. Utilisation du Thésaurus de Papyrus :	5
2.1. Création du thesaurus :	6
2.1.1. Création du micro thesaurus :	6
2.1.2. Ajout des mots clés :	6
2.1.3. Saisie du mot clé :	7

1. Définition :

Un **thésaurus** ou **thésaurus de descripteurs**, est un type de langage documentaire qui consiste en une liste **de termes sur un domaine de connaissances**, reliés entre eux par des relations synonymiques, hiérarchiques et associatives.

Le thésaurus constitue un vocabulaire normalisé. C'est une sorte de dictionnaire hiérarchisé ; cependant, un thésaurus ne fournit qu'accessoirement des définitions, les relations des termes et leur sélection l'emportant sur la description des significations.

Thésaurus est un emprunt au grec ancien θησαυρός « trésor » via le latin *thesaurus*.

Le thésaurus désignait, à côté du mot *trésor* lui-même, un dictionnaire de langue ancienne à but philologique ou archéologique, puis a désigné le thésaurus documentaire dans les années 1950.

Les deux orthographes *thesaurus* et *thésaurus* sont admises par les dictionnaires : la première reprend directement la forme latine, la seconde en est une francisation et semble la plus fréquente dans la littérature.

Un *thesaurus* est constitué d'un ensemble structuré de concepts représentés par des termes, pouvant être utilisés pour l'indexation de documents dans un système de Gestion Electronique de documents comme MicroGED .

Deux types de termes composent un thésaurus :

- les descripteurs utilisés pour indexer un document ;
- les non-descripteurs qui ne peuvent pas être employés pour indexer un document, et qui renvoient au descripteur à utiliser. Ils sont utilisés à la recherche.

Pour l'utilisateur d'un catalogue électronique ou d'une base de données bibliographique, le thésaurus peut constituer un instrument de recherche. L'utilisateur peut formuler des requêtes en ayant recours à des descripteurs ou un non descripteur (dans ce cas le système le renvoie au descripteur), ce qui peut, dans une certaine mesure, limiter le bruit et le silence documentaire.

1.1. Constitution d'un thésaurus

Un *thesaurus* s'élabore, soit manuellement par la voie d'une personne ou de plusieurs, grâce à une intelligence humaine, soit de manière automatique, par le biais de l'intelligence artificielle, grâce à des logiciels de construction automatique de thésaurus du type du SATO (Système d'analyse de textes par ordinateur), soit par un mélange de l'approche humaine et automatique.

Pour des raisons de simplicité le thésaurus de MicroGED sera rempli manuellement.

1.2. Les relations entre les termes

Les termes d'un thésaurus sont organisés hiérarchiquement (à l'intérieur de micro thésaurus souvent classés alphabétiquement). Cette hiérarchie permet de régler la précision de l'indexation ou de l'interrogation. L'indexation s'appuiera autant que possible sur l'identification des termes *spécifiques* (donc du niveau le plus bas possible), alors que la recherche selon les cas pourra faire appel aux termes *génériques* pour augmenter le nombre de réponses.

Les relations des termes sont de trois types :

- relation hiérarchique *stricto sensu* (entre descripteurs), base de la hiérarchie du thésaurus ;
- relation d'équivalence (entre descripteurs et non-descripteurs), base de l'univocité ;
- relation d'association (entre descripteurs), *enrichissement sémantique* ; *sujets connexes*.

Tout thésaurus comporte au moins trois catégories de termes : les termes génériques et les termes spécifiques qui doivent être utilisés comme descripteurs ; les termes équivalents qui sont considérés comme non-descripteurs selon les conventions du thésaurus.

- Les *termes génériques* sont repérés généralement par le sigle **TG** ; ils désignent les entités ou concepts principaux en référence aux autres termes et au domaine considéré ;
- Les *termes spécifiques* sont repérés généralement par le sigle **TS** ; ils précisent et identifient les entités ou concepts particuliers à l'intérieur du champ sémantique d'un terme générique donné ;
- Les *termes équivalents* sont repérés généralement par le sigle **EP** comme abréviation de *Employé Pour* ; ce sont des variantes des termes spécifiques ([synonymie](#) ou quasi-synonymie). Ils sont donc équivalents dans le langage courant, mais donnés pour subsidiaires dans l'emploi du thésaurus. Le terme à préférer au terme *employé pour* est indiqué par le symbole **EM** ou **EMP** comme abréviation de « Employer ».

On trouve aussi très généralement les « termes associés » identifiés par TA (relation d'association : causalité, localisation, relations de nature temporelle, composition, etc). Étant eux-mêmes des descripteurs, ces termes connexes permettent au chercheur de modifier progressivement son interrogation ou de l'élargir sans faire appel aux termes génériques.

Divers types de relations et rubriques complémentaires peuvent être adjoints à cette structure de base pour enrichir le thésaurus ou améliorer son usage. On peut notamment prévoir des « équivalents linguistiques » pour des thésaurus multilingues ainsi que des *passerelles* avec d'autres thésaurus du même domaine ou de domaines différents.

1.3. Exemple de thésaurus élémentaire

Soient les rubriques principales d'un *micro-thesaurus* sur un système informatique collaboratif :

- Individus >
- Logiciel >
- Réseau >
- Ressources >

La rubrique *Individus* se composerait par exemple de :

- Lecteur (**TG**) ;
- Participant (**TG**) ; Auteur (EP) ; Contributeur (EP) ;
 - Éditeur non-inscrit (**TS**) ; Anonyme (EP) ; Adresse IP (forme métaphorique à éviter) ;
 - Éditeur inscrit (**TS**) ;
- Participant mandaté (**TG**) ;
 - Administrateur (**TS**) ; Sysop (terme usuel dans la communauté)
 - Gestionnaire (**TS**) ;
 - Représentant (TS) (chargé des relations extérieures) ;
- Utilisateur (terme imprécis : à proscrire) ; Internaute (imprécis : à proscrire).

Le responsable de toute contribution pourrait ainsi être spécifié par au moins un terme descriptif choisi parmi les cinq termes spécifiques (TS) ou parmi les trois termes génériques (TG), selon les besoins. Les termes (EP) seront par principe évités dans l'indexation, mais pourront être utilisés ultérieurement pour exploiter exclusivement tel ou tel type de contribution sans employer rigoureusement les termes propres de la description initiale.

2. Utilisation du Thésaurus de Papyrus :

Les mots clés peuvent être en saisie libre ou bien sélectionnés dans une liste de mots clés.

Le paramétrage de Papyrus (écran Infocompagnie) permet de forcer la saisie depuis le thésaurus :

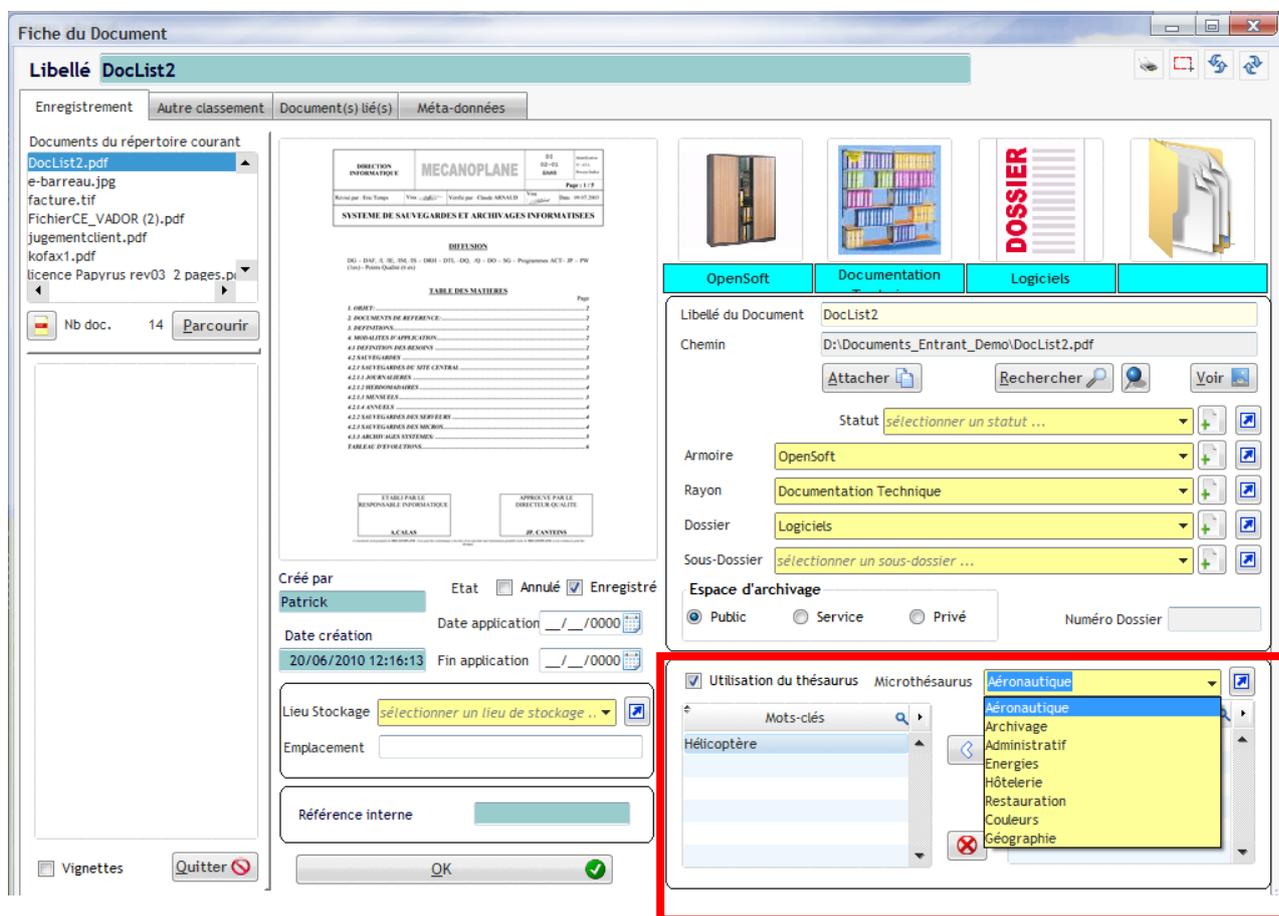
- Case à cocher Utilisation du thésaurus

Une liste déroulante apparaît, proposant le choix d'un micro thésaurus.

Un micro thésaurus regroupant les termes spécifiques.

Le bouton de la flèche bleue permet la création du thésaurus.

Ce bouton est rendu disponible aux administrateurs et superviseurs.



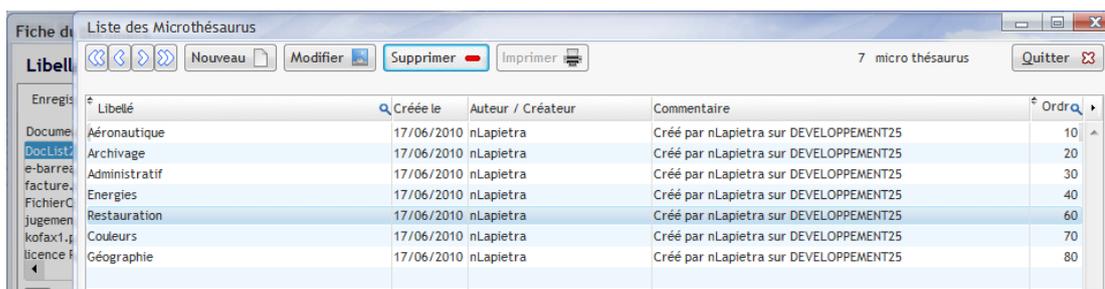
The screenshot shows the 'Fiche du Document' window for 'DocList2'. The interface includes a document preview, a table of contents, and a metadata form. The 'Utilisation du thésaurus' section is highlighted with a red box, showing a checked checkbox and a dropdown menu for 'Microthésaurus' with 'Aéronautique' selected. A search box for 'Mots-clés' contains 'Hélicoptère'. Other dropdowns for 'Armoire', 'Rayon', and 'Dossier' are also visible.

2.1. Création du thésaurus :

La création du thésaurus consiste en une somme de micro thésaurus.

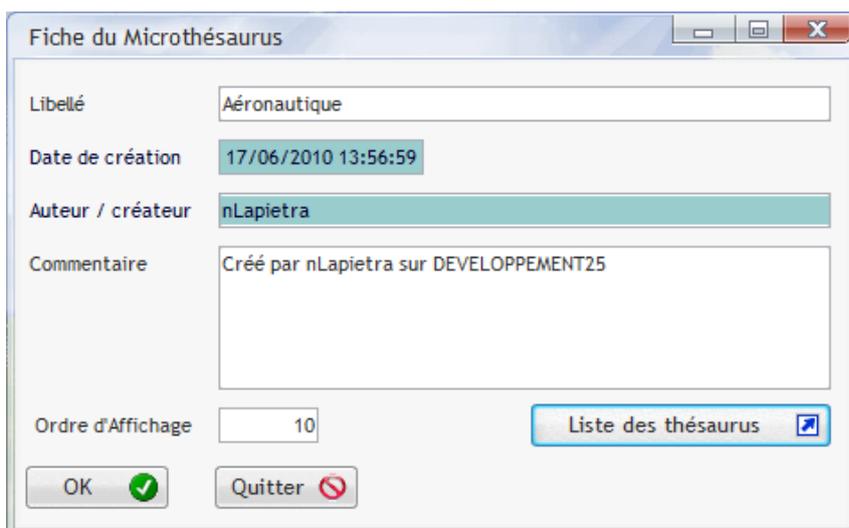
Liste des micro thésaurus

En cliquant sur la flèche bleue, l'administrateur ou le superviseur peut accéder à la création d'un micro thésaurus en cliquant sur le bouton nouveau



Libellé	Créée le	Auteur / Créateur	Commentaire	Ordre
Aéronautique	17/06/2010	nLapietra	Créé par nLapietra sur DEVELOPPEMENT25	10
Archivage	17/06/2010	nLapietra	Créé par nLapietra sur DEVELOPPEMENT25	20
Administratif	17/06/2010	nLapietra	Créé par nLapietra sur DEVELOPPEMENT25	30
Energies	17/06/2010	nLapietra	Créé par nLapietra sur DEVELOPPEMENT25	40
Restauration	17/06/2010	nLapietra	Créé par nLapietra sur DEVELOPPEMENT25	60
Couleurs	17/06/2010	nLapietra	Créé par nLapietra sur DEVELOPPEMENT25	70
Géographie	17/06/2010	nLapietra	Créé par nLapietra sur DEVELOPPEMENT25	80

2.1.1. Création du micro thésaurus :

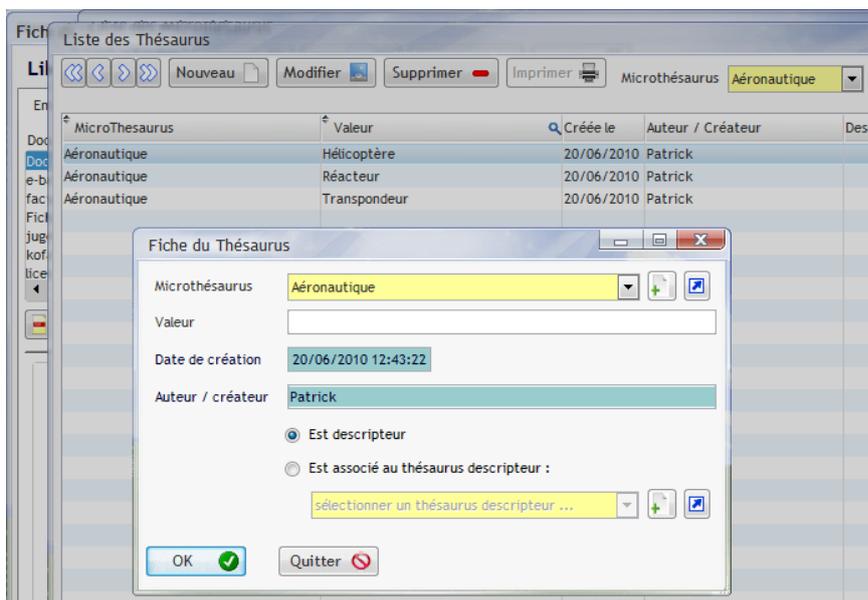


Fiche du Microthésaurus
 Libellé: Aéronautique
 Date de création: 17/06/2010 13:56:59
 Auteur / créateur: nLapietra
 Commentaire: Créé par nLapietra sur DEVELOPPEMENT25
 Ordre d'Affichage: 10
 Liste des thésaurus []
 OK [] Quitter []

2.1.2. Ajout des mots clés :

Il suffit de cliquer sur le bouton : liste des thésaurus

La liste des mots clés du micro thésaurus permet d'ajouter une entrée en cliquant sur le bouton nouveau.



2.1.3. Saisie du mot clé :

Le champ valeur contiendra ce mot clé, ce dernier pouvant être :

- ☒ descripteur ou
- ☒ associé à un autre thésaurus descripteur

