



Mémoire technique

Datas et GED



Sommaire

1	Kentika, base de données documentaire : présentation générale	5
1.1	Les piliers de la solution Kentika	5
1.1.1	Méta-données, GED et gestion de contenu	5
1.1.2	Des profils d'utilisateurs précis.....	6
1.1.3	Le portail « Kentika-Atomic ».....	6
1.1.4	Emailing et portail : une combinaison gagnante	6
1.2	Logiciel ouvert et intégré à 100%.....	6
2	Les données, les fichiers, les contenus	8
2.1	Types d'informations	8
2.2	Dictionnaire de données	8
2.2.1	Natures d'informations.....	8
2.2.2	Attributs	9
2.2.3	Les points forts du dictionnaire de données de Kentika.....	9
2.3	Fichiers [GED]	10
2.3.1	Tout type de fichiers.....	10
2.3.2	Organiser les fichiers	11
2.3.3	Gestion d'alias.....	11
2.3.4	Zone de transit	11
2.3.5	Créer un nouveau document à partir d'un modèle	12
2.3.6	Gestion des versions.....	12
2.3.7	Points forts de la GED de Kentika.....	12
2.3.8	Un éditeur de texte html.....	13
2.3.9	Workflow de publication.....	14
2.3.10	Les points forts de la gestion de contenu(s)	14
3	Alimenter la base de données.....	15
3.1	Saisie manuelle via formulaire.....	15
3.1.1	Rubrique de type « Date »	16
3.1.2	Rubriques de type lien ou thesaurus.....	16
3.1.3	Rubrique de type « code » avec liste de valeurs.....	17
3.1.4	Rubrique de type « URL ».....	18

3.1.5	Contrôle des doublons	19
3.1.6	Reconnaissance des caractères et indexation	19
3.1.7	Dupliquer un enregistrement.....	19
3.2	Centrale d'importation.....	20
3.2.1	Email	20
3.2.2	Dossier.....	20
3.2.3	Fichier.....	21
3.2.4	URL.....	21
3.2.5	Bases externes	21
3.3	Fonctions avancées	21
3.3.1	Identifier les doublons / Fusionner	22
3.3.2	Piloter son scanner et composer une page au format pdf.....	22
3.3.3	Profil de conservation	22
3.3.4	Modification en lots.....	23
3.3.5	Importer des données.....	23
3.3.6	Recherche documentaire	23
3.3.7	Exporter des données	23
3.3.8	Impression et diffusion	24
4	Gestion des droits d'accès	25
4.1	Gestion des groupes d'utilisateurs	25
4.2	Définition de corpus	25
4.3	Rattachement des personnes aux groupes	25
4.4	Accès aux fichiers (GED)	25
4.5	Accès aux rubriques	26
4.6	Le portail s'autoconfigure.....	26
5	Thesaurus & plan de classement.....	27
5.1	Schéma entités-relations.....	27
5.2	Le thesaurus de Kentika : points clés & points forts.....	27
5.2.1	Spécialisation des branches.....	28
5.2.2	Importer – exporter.....	28
5.2.3	Thesaurus de référence dans la profession documentaire.....	28

5.2.4	Tag et mots clés libres	28
5.2.5	Thesaurus et indexation automatique	29
5.2.6	Plan de classement & recherche plein texte	29
5.3	Les outils de gestion du thesaurus.....	29
5.3.1	Mode hiérarchique.....	29
5.3.2	Mode alphabétique	29
5.3.3	Saisie d'un descripteur.....	30
6	Communiquer : email	31
6.1	Le générateur de emails : contextes d'utilisation.....	31
6.1.1	Messages à la demande.....	31
6.1.2	Messages sur événement	31
6.1.3	Messages programmés	32

1 Kentika, base de données documentaire : présentation générale

Kentika est une plate-forme de gestion de métadonnées, de contenu et de documents.

Kentika permet de créer des fiches descriptives en s'appuyant sur un dictionnaire de données personnalisé, ouvert et évolutif. Le thésaurus assure une indexation fine et précise.

La gestion de contenu permet de saisir des textes et des images avec tous les attributs du web d'aujourd'hui.

Les fonctions de gestion électronique de documents et d'indexation full text (plein texte) permettent d'accueillir et de traiter tout type de fichier et d'information, faisant de Kentika une solution à vocation universelle au sein de l'entreprise ou institution.

Pour chaque application développée avec le logiciel, des fonctions avancées permettent un haut niveau de productivité et de pertinence.

1.1 Les piliers de la solution Kentika

1.1.1 Méta-données, GED et gestion de contenu

La gestion des informations et des documents s'articule autour de trois pôles : les méta-données (une fiche descriptive par information identifiée), la gestion électronique de documents et la gestion de contenu.

La forte cohésion entre ces trois dimensions fait de Kentika un outil particulièrement efficace dans la gestion des informations, qu'elles soient structurées, structurables ou non.

Quelques exemples d'applications

Revue de presse : une fiche par article plus un fichier pdf ou un lien vers un site.

Catalogue d'ouvrages : une notice descriptive accompagnée d'une couverture et/ou un sommaire.

Photothèque : une fiche descriptive par photo ou par reportage.

Brevets : une fiche descriptive, le texte intégral joint en pdf.

Base de données juridique : les textes applicables avec un chaînage de type abroge / abrogé par.

Gestion de litiges : une fiche par pièce, informations permettant de construire l'argumentaire.

Site internet : une fiche par page du site, les textes saisis directement en gestion de contenu.

Blog – wiki : une fiche par billet, texte en gestion de contenu.

Pluridisciplinarité

Ces exemples illustrent la grande diversité d'application(s) que Kentika permet de mettre en œuvre, de plus : toutes ses applications peuvent cohabiter au sein d'une même application globale. Les avantages sont alors multiples :

- économie d'échelle : une seule application à mettre en place, à maintenir, à apprendre ;
- recherches transversales : toutes les ressources accessibles sur la même plateforme ;
- création de paniers hétérogènes ;
- système d'alerte unifié : permet de lutter contre la surabondance d'information et de disposer d'une charte graphique unique pour communiquer avec ses utilisateurs.

1.1.2 Des profils d'utilisateurs précis

Chaque acteur ou lecteur peut se voir affecter des corpus, des droits d'actions et s'abonner à des lettres d'information ou profils de veille. Kentika dispose d'un riche modèle de définition des droits permettant de faire face à toutes les situations.

1.1.3 Le portail « Kentika-Atomic »

Souple, ouvert et évolutif le portail Atomic permet de créer des applications présentant des corpus avec des fonctions et une charte graphique adaptés à chaque objectif.

1.1.4 Emailing et portail : une combinaison gagnante

Une diffusion efficace consiste à distiller la bonne dose d'informations à chacun en ciblant l'essentiel. Kentika dispose d'un outil de génération de emails à la demande ou en mode automatique permettant de réaliser tout type d'alertes.

1.2 Logiciel ouvert et intégré à 100%

Kentika : une grande simplicité d'installation et de maintenance. Toutes les fonctions de la base de données sont intégrées. Le serveur http, les clients ftp, smtp ou encore les webservices ne nécessitent pas d'installation complémentaire.

Kentika est fourni avec le SGBD 4D.

Kentika communique avec l'architecture informatique en place : l'interopérabilité est large et de nombreuses passerelles sont opérationnelles (LDAP, RES, Sharepoint...).

Client web et client riche

Pour les opérations au quotidien le portail fournit une grande qualité de présentation, privilégie l'ergonomie d'accès aux données et aux fonctions, bénéficie d'un haut niveau d'automatisme dans la génération des outils proposés à chacun.

Pour quelques fonctions avancées inaccessibles à un navigateur (exemple : piloter un scanner de manière efficace, surveiller le contenu d'un répertoire ou encore ocriser un document), le client riche donne accès à des outils sophistiqués.

2 Les données, les fichiers, les contenus

2.1 Types d'informations

Chaque « objet », au sens conceptuel du terme, peut être décrit au moyen d'une fiche dont les rubriques sont définies au sein d'un dictionnaire de données.

Une rubrique peut être partagée entre différents types d'objets (exemple : le champ « Titre » est commun aux « monographies », aux « films », aux « brevets » ...).

Exemples de types d'informations pouvant être décrits :

- monographie ;
- photographie ;
- texte de loi ;
- société ;
- auteur ;
- éditeur ;
- courrier ;
- article de presse ;
- ...

2.2 Dictionnaire de données

Le dictionnaire de données est ouvert et évolutif. L'administrateur peut, à tout moment et en toute autonomie intervenir sur le paramétrage (créer de nouveaux types de documents, déclarer de nouvelles rubriques). Les masques de saisie et de recherche étant complètement dynamiques, l'ajout d'une nouvelle rubrique ne nécessite pas d'intervenir sur le paramétrage de son portail.

2.2.1 Natures d'informations

- **Code** : les valeurs que peut recevoir la rubrique sont vérifiées par rapport à une liste de valeurs possibles.
- **Valeur thesaurus** : seuls les descripteurs appartenant au thesaurus peuvent être utilisés. Une rubrique peut être limitée à un thesaurus précis.
- **Valeur format libre** : aucun contrôle n'est effectué sur les valeurs saisies.
- **Date** : peut être précise (exemple : 18 mai 1959) ou bien floue (exemple : 1958-1960), Kentika détermine alors les « bornes ».
- **Nombre** : ne peut recevoir que des valeurs numériques.
- **Adresse email** : permet d'envoyer un email en utilisant l'adresse indiquée.
- **URL** : permet de recevoir une URL, possibilité de contrôle de sa validité et d'archivage de la page correspondante.

- **Lien** : permet d'établir un lien avec la même table ou une autre table.
- **Adresse** : adresse postale en vue d'envoi de courriers.

Libellé Nature du document

Commentaire Complète en le précisant le type de document

Contenu Code (valeurs ci-dessous)

Attributs

Valeur multiple Champs de recherche

Obligatoire Intégrer aux alertes si modification

Clé de dédoublement Inscrire les modifications au log

Saisie au clavier Bouton radio Trier Liste fermée

Code	valeur
03	Assurance qualité
04	Modèle
05	Document juridique

2.2.2 Attributs

- **Obligatoire** : lors de la saisie d'une nouvelle fiche, la validation est refusée tant que la rubrique n'est pas renseignée.
- **Clé de dé doublement** : lorsqu'un doublon potentiel est détecté, en bas de l'écran apparaît une alerte.
- **Valeur par défaut** : initialisation lors de la saisie d'une nouvelle fiche.
- **Valeur multiple** : permet de saisir plusieurs valeurs pour la même rubrique.

2.2.3 Les points forts du dictionnaire de données de Kentika

Kentika offre des fonctions avancées autorisant une indexation avec le juste niveau de précision recherché.

Rubriques groupées et multivaluées

Kentika dispose d'un mécanisme complètement intégré au moteur et ne nécessitant pas, pour le gestionnaire, de programmation pour en bénéficier : les rubriques groupées et multivaluées. Grâce à ce mécanisme, Kentika sait gérer des structures complexes comme :

- un ouvrage ayant plusieurs auteurs et chacun d'eux ayant un niveau d'intervention spécifique ;
- une société ayant des compétences avec, pour chacune, le niveau d'expertise ;

- une personne ayant occupé plusieurs postes avec pour chacun le nom de la société, le poste occupé, la date de début, la date de fin et un commentaire (exemple : base de cv).

Liens récursifs

Une rubrique permet de faire pointer un enregistrement vers un enregistrement de même nature, la fonction réciproque est alors automatiquement gérée. Ainsi, on peut préciser que le texte de loi du 12 mars 2012 abroge celui du 8 janvier 2007 ; que la société XYZ SAS est filiale du groupe WorldCompany SA.

Prise en compte de la confidentialité

Des restrictions peuvent être définies pour une rubrique afin que certains groupes d'utilisateurs ne puissent pas saisir, ou pas modifier ou pas consulter une information.

La valeur qu'un utilisateur peut saisir doit faire partie des valeurs qu'il peut consulter.

Traduction intégrée

Chaque libellé de rubrique peut être traduit, de même que chaque valeur des listes ouvertes ou fermées : une information sera alors présentée à chaque utilisateur dans la langue qu'il aura sélectionnée.

Exemple :

[en français] Descripteur : Etats-Unis

[en anglais] Keyword : United States

Evolutif, à tout moment et en autonome

Kentika rend ses utilisateurs autonomes : ceci est vrai aussi pour le dictionnaire de données.

Une rubrique à été ajoutée au dictionnaire ? Affectée à un type d'objet ? Les formulaires de saisie s'autoconfigurent automatiquement sans qu'aucune intervention ne soit nécessaire.

2.3 Fichiers [GED]

Une notice peut être porteuse d'une information dans sa globalité, elle peut aussi être un moyen d'accéder à une ou plusieurs ressources électroniques.

2.3.1 Tout type de fichiers

Kentika accepte tout type de fichier. En fonction du type d'information contenu, des traitements différenciés pourront être appliqués.

Document bureautique : analyse du contenu, extraction de concepts figurant dans le thesaurus (que ce soit sous forme de descripteur, de synonyme ou encore de traduction). Cette fonction est disponible avec le pack Search de Kentika.

Document issu de scan : reconnaissance de caractères en tâche de fond, ré-archivage automatique dès que le traitement est achevé.

Cette fonction est disponible avec un logiciel fonctionnant au niveau du serveur (exemple : ABBYY Recognition Server).

Document image : extraction d'une imagerie et d'une image basse résolution pour une diffusion optimisée de banque d'images.

Pour information : Kentika intègre les technologies de GED depuis plus de vingt cinq ans et s'adapte en permanence aux évolutions des technologies.

2.3.2 Organiser les fichiers

Un notice peut correspondre à un fichier (exemple : un texte de loi en pdf) ou à un ensemble de fichiers (exemple : la réponse à un appel d'offres).

Il peut alors être très utile, voire indispensable, de pouvoir structurer les fichiers au sein d'un ensemble conceptuellement homogène.

Pour les niveaux de classement pré définis : des règles spécifiques peuvent être affectées (exemple : autorisation nécessaire pour voir le contenu).

Pour les niveaux de classement libre : création d'autant de niveaux d'arborescence que nécessaire, réorganisation possible à tout moment.

2.3.3 Gestion d'alias

Un fichier peut être pertinent pour différentes fiches. Ainsi, un appel d'offre peut utiliser un fichier déjà utilisé par ailleurs. La gestion d'alias permet de créer un pointeur vers le fichier à partager et d'éviter de dupliquer un fichier.

2.3.4 Zone de transit

Le versement d'un fichier peut se faire ponctuellement ou en lot. Lorsque cette deuxième méthode est utilisée, le serveur Kentika réceptionne un ensemble de données, les compare avec ce qui est déjà disponible dans la base, identifie les doublons ou quasi-doublons, propose d'enregistrer une nouvelle version d'un document existant dans ce dernier cas.

2.3.5 Créer un nouveau document à partir d'un modèle

Afin d'homogénéiser la présentation, le contenu, le formalisme et d'obtenir des gains de productivité, un ensemble de modèles peut être mis à disposition afin de créer de nouveaux fichiers répondant à des critères bien précis. De plus, on s'assure ainsi que chacun utilisera toujours la dernière version d'un document.

2.3.6 Gestion des versions

Pour chaque fichier, Kentika permet de conserver les différentes étapes en historiques et, le cas échéant, de revenir à une version plus ancienne (ie : annuler d'éventuelles modifications récentes non approuvées).

2.3.7 Points forts de la GED de Kentika

Adapter la granularité de l'information

Une application de GED doit avant tout apporter un gain de productivité et améliorer le confort de travail. Aussi, il est important de pouvoir soigneusement équilibrer l'indexation.

Documents sensibles protégés

Avant de fabriquer un lien vers un fichier, Kentika vérifie que l'utilisateur a bien le droit d'y accéder. Avant de fournir le fichier correspondant au lien, Kentika vérifie à nouveau que l'utilisateur a bien le droit d'y accéder.

De plus, ces droits peuvent se définir et s'appliquer pour chaque document, et non pour globalement les documents d'une collection.

En réseau local ou en SAAS : des documents physiquement localisés

Kentika sur votre réseau local : les documents sont regroupés au sein de répertoires, leur nom est banalisé, seul l'administrateur y a accès.

Traçabilité

Le journal d'événements de Kentika permet de savoir qui a saisi, modifié, téléchargé un document.

GED et Workflow

Transmettre un document pour lecture, pour avis, pour contribution, voire pour validation, permet de prendre en charge l'élaboration collaborative de son contenu et de s'assurer que chaque intervenant aura effectué la tâche qui lui aura été confiée.

Gestion de contenu

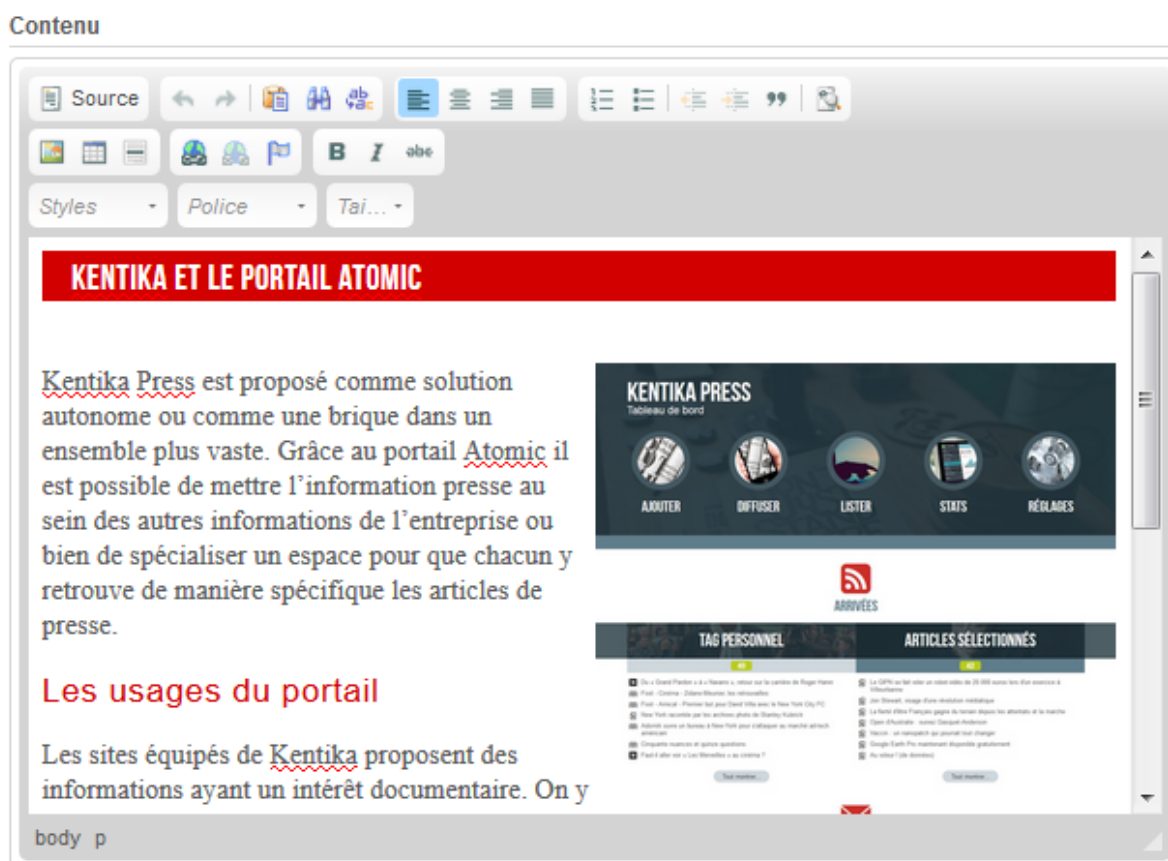
Chaque fiche peut recevoir un volet « gestion de contenu ». Ceci permet de gérer intégralement un site internet : c'est le cas du site de Kentika. De plus, la présentation s'adapte automatiquement au contexte dans lequel l'information est présentée.

Exemple d'une même fiche dans deux environnements (ou espaces) différents :
<http://www.kentika.com/KENTIKA-19108683124919268659-1-Kentika-et-le-portail-atomic.htm> / <http://press.kentika.com/KENTIKA-19108683124919268659-1-Kentika-et-le-portail-atomic.htm>

2.3.8 Un éditeur de texte html

La solution intégrée retenue par Kentika est une référence du monde du web : CKEditor. Elle donne accès à une large palette de possibilités.

Contenu



KENTIKA ET LE PORTAIL ATOMIC

Kentika Press est proposé comme solution autonome ou comme une brique dans un ensemble plus vaste. Grâce au portail Atomic il est possible de mettre l'information presse au sein des autres informations de l'entreprise ou bien de spécialiser un espace pour que chacun y retrouve de manière spécifique les articles de presse.

Les usages du portail

Les sites équipés de Kentika proposent des informations ayant un intérêt documentaire. On y

body p

Pour découvrir les nombreuses possibilités de cette solution : <https://ckeditor.com/>

2.3.9 Workflow de publication

Les modifications apportées à une zone de texte enrichie peuvent être mises en ligne directement ou bien être soumises à validation. De plus, les versions intermédiaires sont conservées.

2.3.10 Les points forts de la gestion de contenu(s)

Chercher dans les contenus

La zone de contenu est intégralement gérée dans l'application Kentika, elle est indexée au même titre que n'importe quel autre champ d'une notice. Elle peut être proposée dans les masques de recherche : seule ou en combinaison avec d'autres rubriques, voire avec les fichiers de la GED.

Serveur d'images

Kentika propose un serveur d'images intégré permettant d'intégrer des illustrations dans les contenus.

Envoi par email

La zone de contenu peut être jointe au corps d'emails. Il est également possible d'envoyer une liste de contenus : ils sont alors insérés les uns à la suite des autres dans un email unique et global.

Multilinguisme

Les textes de la base de données peuvent être traduits, les zones de contenus également.

Lorsqu'un utilisateur déclare une langue de navigation, toutes les traductions disponibles sont utilisées à la place de la version de la version d'origine, sans qu'il ne soit nécessaire de dupliquer les informations ou le site lui-même.

3 Alimenter la base de données

L'alimentation de l'ensemble des données est un facteur clés de succès d'une application. En effet, différents aspects doivent pris en compte.

- **Productivité** : la saisie doit requérir le moins de manipulations possibles.
- **Ergonomie** : la saisie doit être suffisamment intuitive pour que chacun puisse la réaliser sans formation excessive.
- **Contrôle de cohérence** : les valeurs saisies doivent, pour un certain nombre d'entre elles, répondre à des règles précises.
- **Respect des autorisations** : certaines rubriques doivent être réservées à certains utilisateurs.
- **Workflow de publication** : la mise en ligne d'un document peut nécessiter un niveau d'autorisation.
- **Interopérabilité** : à l'import ou à l'export ; en mode connecté ou déconnecté, les échanges avec d'autres applications doivent s'adapter à différents protocoles.

3.1 Saisie manuelle via formulaire

A partir du paramétrage en place et des droits de l'utilisateur, Kentika créé dynamiquement et à la volée des formulaires de saisie avec l'intégralité des contrôles en temps réel et de l'auto-complétion.

The screenshot shows a web form titled "Saisie de la notice" with a pencil icon. The form has several fields and a dropdown menu. The "Type" field is set to "Dépôt". The "Titre" field contains "Au musée del Duomo - Milano". The "Date de parution" field contains "08/07/2015". The "Commentaire" field is empty. The "Equipe" field is set to "Marketing" and has a dropdown menu open showing options: "Equipe", "Marketing", "Situé bâtiment A", "Afficher", and "Fermer". The "Propriétaire" field contains "Gestionnaire Kentika (Monsieur)". The "Statut" field has radio buttons for "Initialisé", "Brouillon", "Validation attendue", "Privé", "Visible des", and "Publié".

Une notice est composée de champs qui peuvent être mono-valués, multi-valués ou encore groupés. Les champs peuvent être attachés à une ou plusieurs bases. Les règles de contrôles définies au niveau des champs sont exploitées lors de la saisie, assurant ainsi une intégrité des données. Les masques de saisie sont générés dynamiquement avec toute aide et assistance nécessaire.

En envoyant un fichier (ou une URL) à Kentika, une analyse de ce dernier est effectuée et une fiche pré-remplie est proposée à la modification / validation. Les méta-données sont extraites de la source, le contenu est analysé par rapport au thesaurus.

Un bookmarklet permet, lors de l'identification d'une page web intéressante, d'en envoyer les méta-données et d'obtenir une fiche pré-remplie qu'il ne reste plus qu'à compléter et valider.

3.1.1 Rubrique de type « Date »

Les dates floues sont acceptées : Kentika détermine, grâce à un algorithme d'analyse syntaxique, les bornes de la période saisie.



3.1.2 Rubriques de type lien ou thesaurus

Les valeurs vérifiées par rapport à une table (exemple : un descripteur qui doit appartenir au thesaurus) peuvent être créées si elles n'existent pas encore.

The image shows a metadata entry form with the following fields:

- Référence: [Empty text box]
- Date de parution: Printemps-été 2012
- Nbre/N° de page: [Empty text box]
- Mots clés: Equipe | (An orange arrow points from the dropdown menu to this field)
- Cote: [Empty text box]
- Commentaire: [Empty text box]
- Niveau d'autorisation: Public (with a dropdown arrow)
- Auteurs: [Empty text box]
- Intervention: [Empty text box]
- ISBN/ISSN: [Empty text box]

The 'Mots clés' dropdown menu is open, showing a list of terms:

- FERROVIAIRE
- EQUIPEMENT INDUSTRIEL** (highlighted)
- THESAURUS MATIERE > GESTION DE 16-03 / bien d'équipement / bien de
- EQUIPEMENT URBAIN
- MATERIEL AGRICOLE
- MATERIEL DE BUREAU
- MOBILIER DE BUREAU
- SPECTACLE
- Créer "Equipe"

At the bottom of the dropdown, there is a 'Fermer' button. An orange arrow also points from the 'EN' language indicator to the 'Créer "Equipe"' option.

A noter que la liste des descripteurs proposés (auto-complétion) tient compte du descripteur, de ses synonymes (ou employé pour) ou encore de la traduction. De plus, l'utilisateur peut se déplacer au sein de la liste hiérarchique.

3.1.3 Rubrique de type « code » avec liste de valeurs

Les listes fermées peuvent être présentées sous forme de boîtes à cocher (valeur multiples possibles) ou sous forme de bouton radio.

Langue Allemand Anglais Espagnol
 Français Italien Russe

Cas d'une rubrique avec un nombre pré-défini de valeurs : possibilité (si la liste est semi-ouverte) d'ajouter une valeur en cours de saisie.

Auteurs	COELHO, Paulo
Intervention	Principal (Principal)
Auteurs	LACKBERG, Camilla
Intervention	
Auteurs	
Intervention	
ISBN/ISSN	
Prix	
Lieu d'édition	

Intervention

- Photographe (Photograph)
- Préfacier (Préfacier)**
- Principal (Principal)
- Producteur (Producteur)
- Réalisateur (Réalisateur)
- Réalisateur - Interprète (Real)
- Scénariste (Scénariste)
- Secondaire (Secondaire)
- Textes (Textes)
- Traducteur (Traducteur)

Nouvelle valeur

Fermer

A noter : dans la saisie ci-dessus, plusieurs auteurs ayant chacun un niveau d'intervention qui peut être différent (rubriques groupées). Ces groupes sont en nombre illimités (possibilité de saisir autant d'auteurs que nécessaire).

3.1.4 Rubrique de type « URL »

Existante d'une URL : Kentika permet d'en vérifier immédiatement la validité. Un traitement peut être programmé pour vérifier régulièrement la validité des URL.

A Monaco, les yachts XXL font salon | Les échos 19 Sep 2015 Nouveau
 Article de presse | Panorama de presse (Economie) ★★★★★



Salle de cinéma, robinetteries en platine ou piscine avec vue sous-marine, les fabricants de yachts n'ont pas de limites pour séduire leurs clients. Le Monaco Yacht Show s'ouvre mercredi.

Salon

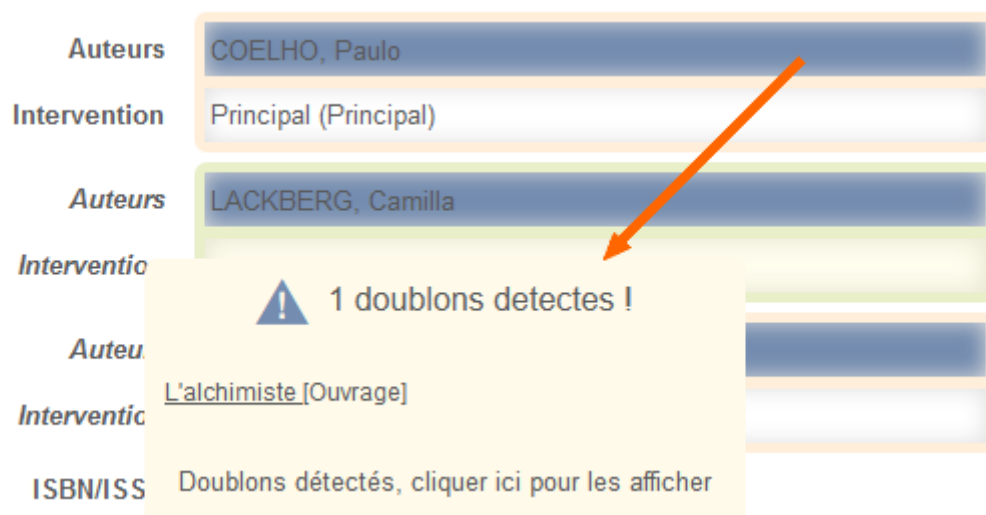
... Lire [+]

URL Image

Une URL peut être identifiée comme recevant l'URL d'une image : dans ce cas, c'est cette image qui peut être prise pour illustrer un enregistrement de la base de données. Ceci est particulièrement utile pour les articles issus des flux RSS ou encore des couvertures d'ouvrage.

3.1.5 Contrôle des doublons

Lors de la saisie, au fur et à mesure que les rubriques sont renseignées, Kentika effectue un contrôle et propose de consulter les fiches identifiées comme doublon potentiel (dans un nouvel onglet afin de ne pas perdre la saisie en cours). Les champs en commun avec des fiches existantes sont affichés avec un fond identifiable.



3.1.6 Reconnaissance des caractères et indexation

La reconnaissance optique des caractères peut être effectuée soit en amont de l'indexation (avant l'envoi du fichier au serveur), soit en aval (lors de la validation de la fiche).

Dans ce cas, Kentika archive le fichier non ocrisé et en effectue une copie en attente d'ocrisation. Lorsque le traitement est effectué, le fichier original est remplacé par celui ocrisé et peut être indexé en plein texte. L'avantage de ce process est que tous les fichiers sont ocrisés avec les mêmes réglages, aucune installation n'est nécessaire sur les postes clients et l'outil proposé dans la présente proposition (ABBYY RS4, Kentika est partenaire ABBYY) présente l'avantage d'optimiser les fichiers pdf.

Par ailleurs, si un document est déjà ocrisé : le copier / coller est tout à fait possible (une attention doit cependant être portée à cette opération qui peut coller des textes avec des césures non justifiées).

3.1.7 Dupliquer un enregistrement

Cette fonction est particulièrement utile lors de saisies répétitives.

3.2 Centrale d'importation

L'alimentation de la base de données peut se faire via un ensemble de connecteurs qui permettent d'automatiser les tâches d'alimentation.

Cinq modes de récupération de données à importer sont proposés. Pour chacun d'eux, quelques exemples vous sont proposés. Ils sont fournis à titre d'illustration et bien d'autres cas de figure peuvent être pris en charge.

3.2.1 Email

Vos collaborateurs adressent des documents à des clients par email. Le mail est adressé à son destinataire et est mis en copie à une adresse propre à votre application. Grâce à la centrale d'import, l'adresse recevant les copies est mise en relève automatique. Les emails reçus sont analysés, une fiche document est créée automatiquement et la pièce jointe (le compte-rendu de visite) est archivée. Le dossier du client est alors automatiquement alimenté.

Des abonnements à des bases de données font que des emails d'alerte vous sont adressés régulièrement à une adresse. Ces emails sont des listes de nouveautés. Lors de la relève de la boîte aux lettres, le contenu du message est découpé et autant de fiches documents sont créées à partir des informations trouvées dans le email.

Deux sites échangent en automatique des données via emails. Le premier site dépouille la presse quotidienne nationale et le deuxième, la presse quotidienne régionale. A midi, le dépouillement est terminé. Un programme batch (ie : sans intervention) identifie les nouveautés de la PQN sur le site 1, prépare un lot et l'envoie par email au site 2. Le site 2 surveille l'adresse email à laquelle le site 1 envoie ses nouvelles saisies, extrait les données et crée automatiquement les nouveaux articles... et réciproquement pour la PQR.

3.2.2 Dossier

Vos photos sont stockées dans des dossiers sur votre disque dur. Chaque dossier porte comme nom la date et à l'intérieur se trouvent des sous-dossiers avec les thèmes des photos.

La mise en surveillance de votre dossier de photos va permettre à l'application de détecter tout mouvement à l'intérieur de ce dossier, créer les fiches descriptives en extrayant automatiquement la date et le thème et procéder à l'archivage de vos photos qui seront alors immédiatement disponibles sur votre web (si vous disposez de cette fonction).

Des rapports sont produits et enregistrés sur un volume partagé dans des dossiers dont le nom correspond au nom du client. L'automate d'import surveillera les

nouveaux rapports et créera les fiches descriptives en rattachant le document au client concerné.

Un scanner est en libre service, chaque document est classé dans un sous-dossier. Les documents issus de scan sont archivés et traités en reconnaissance de caractères en tâche de fond. Le résultat est indexé en texte intégral et la fiche document placée dans un dossier pour complément et / ou diffusion.

3.2.3 Fichier

Une application de gestion du personnel permet de connaître tous les mouvements au niveau des salariés de l'entreprise. Tous les soirs, un export est programmé à partir de cette application avec les nouveaux entrants et les changements d'affectation. Un fichier est créé et placé à un endroit précis avec un nom précis. Un automate surveille si le fichier a été modifié et si c'est le cas, l'analyse et procède aux modifications des fiches personnes.

Votre application est couplée avec une application dédiée à la surveillance de site. Lorsque cette dernière détecte des changements sur des sites, elle produit un fichier récapitulatif avec pour chaque site modifié, la date de modification et une indication de statut. L'automate analyse le fichier et change les informations sur les fiches correspondantes. Dans l'interface web, un dossier a été programmé afin de visualiser en un seul clic la liste des sites qui ont été modifiés.

3.2.4 URL

Un site "ami" publie régulièrement une liste de nouveautés sous la forme d'une page web. Pour chaque élément de la page, un lien est proposé sur le descriptif. La surveillance de l'URL permet de détecter les changements sur la liste, en découpe le contenu et créé autant de fiches dans la base que d'éléments dans la liste.

3.2.5 Bases externes

Un connecteur à un moteur de recherche permet d'effectuer des recherches. Des sujets particuliers vous intéressent et vous souhaitez que les nouveaux résultats soient automatiquement créés dans votre application. La requête est relancée régulièrement et les nouveaux résultats importés dans votre application.

3.3 Fonctions avancées

Travailler avec Kentika peut se faire via un navigateur (mode full web) et via un logiciel client dédié. Certaines fonctions très avancées qu'un navigateur ne peut

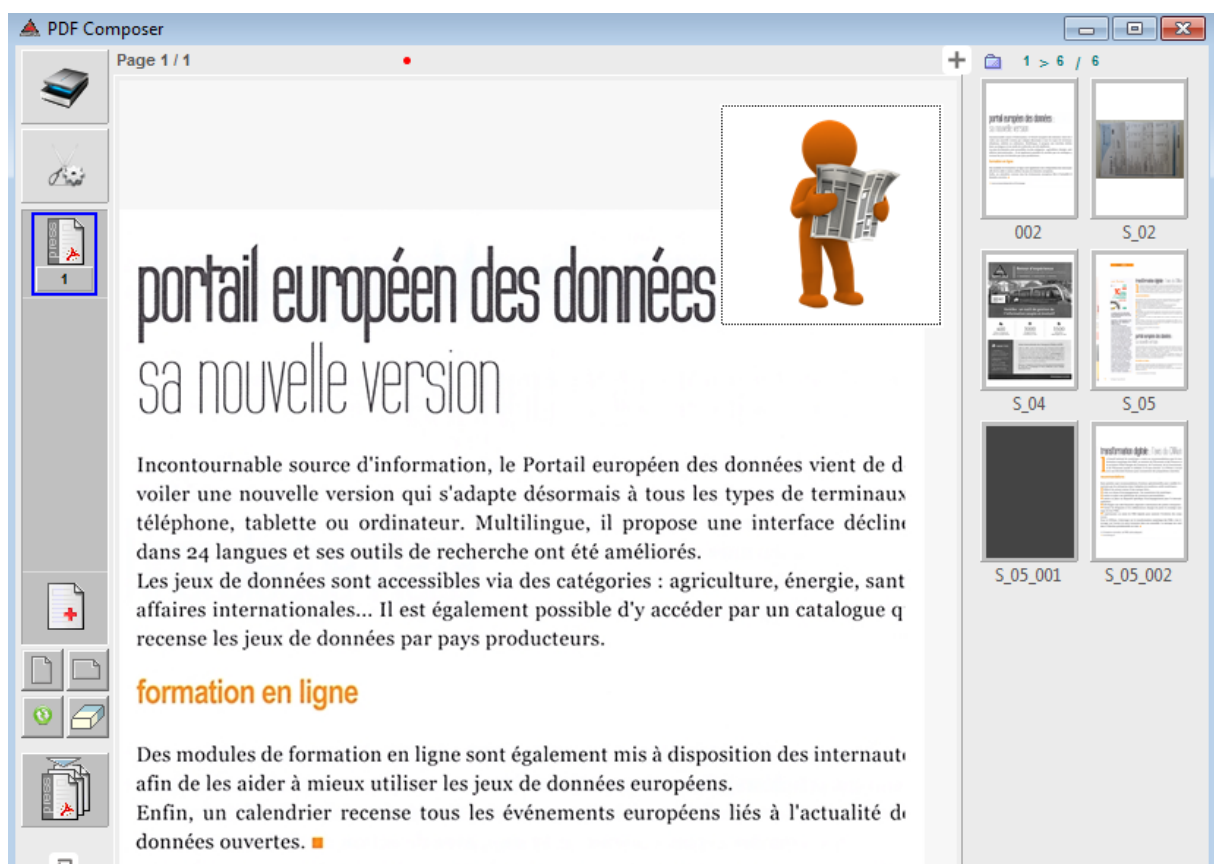
reproduire ou dont l'utilisation est occasionnelle ne sont disponibles que sur le poste client.

3.3.1 Identifier les doublons / Fusionner

Cette fonction a pour objectif de détecter dans la base de données les doublons et de proposer ensuite de ne conserver qu'une seule des deux valeurs en reportant les valeurs de l'un vers l'autre.

3.3.2 Piloter son scanner et composer une page au format pdf

Piloter son scanner, découper les parties de pages à conserver, recomposer, faire appel dans une maquette à des informations issues des notices, compléter avec des logos par thématique ou par support d'origine : cet ensemble de fonctions permet de créer des documents pdf bien présentés qui seront ensuite ocrisés et indexés en full text (plein texte) - suivant options de Kentika installées.



3.3.3 Profil de conservation

Certaines informations ont un intérêt pendant un temps très court (parfois quelques jours). Afin de les traiter au sein de produits communs (exemple : un panorama de

presse), il est nécessaire de les indexer mais de ne les conserver qu'un temps pré défini.

Un désarchivage est automatiquement effectué de ces données caduques et le résultat peut être stocké sur un support externe. Cette opération peut également être réalisée manuellement.

3.3.4 Modification en lots

Kentika propose un puissant outil permettant de réaliser des modifications sur des lots de notices. Ces dernières peuvent être simples (exemple : renseigner une date), conditionnelles (exemple : si une fiche contient un descripteur A, le remplacer par le descripteur B) voire complexe en exploitant le langage de script de Kentika (exemple : ajouter [OLD] en fin de cote si le document est de type article et antérieur à 2000).

La progression du traitement est visible et toutes les opérations sont immédiatement prises en compte dans la base de données.

3.3.5 Importer des données

L'outil d'import de données, disponible dans le logiciel et accessible aux gestionnaires permet d'intégrer tous les formats couramment rencontrés en documentation (texte tabulé, texte avec étiquettes, 2709, xml, csv, ris...). Depuis 32 ans, le logiciel Kentika a récupéré des bases en provenance de la quasi-totalité des acteurs du marché et intégré des données en provenance d'un grand nombre de fournisseurs.

3.3.6 Recherche documentaire

Kentika propose trois types d'outils de recherche :

- **Recherche simple** : disponible sur toutes les pages, en haut à droite. Cette recherche peut être proposée avec des options permettant de restreindre ou d'étendre la portée de la recherche ;
- **Recherche avancée** : elle est paramétrée via un constructeur et proposée aux utilisateurs en fonction du contexte et des autorisations de chacun ;
- **Recherche experte** : permet de construire, par sélection de champs et d'opérateurs, des équations de recherche booléennes.

Les recherches sont détaillées dans le document « Kentika et le Portail Atomic ».

3.3.7 Exporter des données

Kentika répond à différentes situations :

- un utilisateur fait une recherche et souhaite obtenir simplement et rapidement le résultat de sa sélection sous forme d'un fichier sur son disque ;
- un gestionnaire effectue une extraction pour une exploitation dans un contexte autre que Kentika ;
- des exports réguliers de la base sont effectués pour alimenter un autre système.

Dans le premier cas, des formats prédéfinis peuvent être mis à disposition (exemple : texte tabulé pour une exploitation tableur ; format traitement de texte ou encore un mini-site composé d'un index et des documents GED associés).

Pour le gestionnaire : des constructeurs d'états sont mis à leur disposition. Les formats d'export proposés sont : txt, csv, xml, 2709, html, pdf, ris, doc,...

Récolement : il fait l'objet d'un traitement intégré au logiciel. Il rapproche l'inventaire déclaré (le catalogue), les exemplaires trouvés dans les rayons (par lecture de codes à barre) et les emprunts en cours.

Les **exports** réguliers peuvent être programmés sous forme de batch.

3.3.8 Impression et diffusion

Kentika dispose de différents outils pour réaliser des exports et impressions. Un générateur d'étiquettes ; un générateur de fichier pdf capable d'assembler un sommaire et les fichiers pdf.

Les possibilités offertes en export et impression peuvent être limitées à certains groupes d'utilisateurs. Les autorisations sont prises en compte dans les maquettes d'impression et d'export, que ce soit au niveau des corpus ou au niveau des rubriques.

4 Gestion des droits d'accès

La prise en compte de toute règle en termes d'accès aux données et aux fichiers, que ce soit en lecture et / ou en écriture, est un des points forts du logiciel.

Tous les cas de figure rencontrés, même complexes, d'organisation des autorisations ont pu être pris en compte dans Kentika !

4.1 Gestion des groupes d'utilisateurs

Kentika permet de créer autant de groupes que nécessaire.

Un groupe est caractérisé par un ensemble d'autorisations qui peuvent être globales ou limitées à un corpus.

Kentika propose un ensemble de 107 autorisations hiérarchisées avec un principe d'héritage. Chaque site peut compléter cette liste et se créer ses propres autorisations pour prendre en compte des situations particulières.

Exemples : un usager pourra voir les brevets lorsqu'ils sont au statut « publié » / un gestionnaire d'une équipe de travail peut modifier tous les documents de son équipe sauf s'ils sont au statut « archivé » / le thesaurus « géographie » est accessible à tous / le thesaurus « finance » a un accès restreint.

4.2 Définition de corpus

Des requêtes simples ou complexes permettent de définir des corpus de documents, une requête peut porter sur une information fixe (exemple : les documents de type « article » ou « ouvrage ») ou une information variant en fonction de l'utilisateur (exemple : les documents concernant une des zones géographiques auxquelles l'utilisateur connecté est affecté).

4.3 Rattachement des personnes aux groupes

Chacun peut appartenir simultanément à plusieurs groupes. Dans ce cas, les autorisations se cumulent. Si une personne n'appartient à aucun groupe, elle hérite des droits qui auraient été affectés aux groupes des invités.

4.4 Accès aux fichiers (GED)

En complément à l'ensemble des règles déjà actives, il est également possible d'affiner ces règles pour tenir compte de situations particulières.

Exemples : il est possible de permettre l'accès aux fichiers GED lorsque la base est accédée en intranet mais pas depuis l'extérieur ; les images basse résolution sont proposées aux visiteurs, les images haute résolution sont réservées au gestionnaire.

4.5 Accès aux rubriques

Chacun dispose d'un dictionnaire de données qui tient compte d'éventuelles restrictions qui auraient été mises en place. Si une rubrique lui est invisible, elle est masquée dans la restitution des informations à l'écran, dans les impressions, les exports ainsi que sur les masques de recherche.

4.6 Le portail s'autoconfigure

Les autorisations données à chacun ont un impact direct sur ce que le portail leur permet de consulter, saisir ou modifier. Ainsi, lorsque Kentika propose des options (exemple : un accès aux types de documents), il est vérifié préalablement que l'utilisateur a bien accès aux fonctions proposées (i.e. : ne sont proposés que les types de document auxquels l'utilisateur a accès).

5 Thesaurus & plan de classement

Kentika propose une structure arborescente de concepts permettant de disposer d'un plan de classement ou d'un thesaurus, particulièrement utiles pour structurer son indexation et offrir une grande souplesse et pertinence lors des recherches.

Kentika peut travailler avec plusieurs thesaurii simultanément.

5.1 Schéma entités-relations

Quelle que soit le formalisme que l'on adopte ou met en œuvre : le point commun est que l'on a toujours un schéma permettant de décrire des entités et d'établir des relations entre elles.

Les schémas couramment rencontrés en gestion de l'information répondent à des objectifs et formalismes variables mais qui, finalement, ont plus de points communs que de différences.

- **Thésaurus** : liste organisée de termes contrôlés et normalisés (descripteurs et non descripteurs) représentant les concepts d'un domaine de la connaissance.
- **Plan de classement** : système d'organisation de l'information consignée, par sujets, du général au particulier, reflétant les activités d'un organisme.
- **Ontologie** : ensemble structuré des termes et concepts représentant le sens d'un champ d'informations, que ce soit par les métadonnées d'un espace de noms, ou les éléments d'un domaine de connaissances.
- **Taxonomie** : organise l'information de façon hiérarchique, de la simple restitution de faits jusqu'à la manipulation complexe des concepts, qui est souvent mise en œuvre par les facultés cognitives dites supérieures.
- **FRBR** : modèle d'analyse d'une notice catalographique comme un ensemble d'informations portant sur quatre aspects distincts du document décrit : Item - Manifestation - Expression - Œuvre.

5.2 Le thesaurus de Kentika : points clés & points forts

Descripteur ou mot clé : si le langage courant retient souvent le terme de mot-clé, le formalisme utilisé dans le logiciel retient la dénomination descripteur.

Hiérarchie : la structuration arborescente est un atout important lors des recherches. Ainsi, une recherche sur un descripteur pourra être automatiquement étendue aux descripteurs qui s'y rapportent (termes spécifiques) ou qui sont en rapport (terme associés).

Synonymie (ou non-descripteur) : en enrichissant son thesaurus des formes non retenues d'une expression ou d'un concept, on donne une grande souplesse à l'utilisateur qui interroge la base de données.

Traduction : lors de la recherche sur une expression dans une langue étrangère, si cette expression correspond à la traduction d'un descripteur, elle renvoie automatiquement au descripteur. Lors de l'affichage d'un résultat, la traduction d'un descripteur dans la langue de l'utilisateur est automatiquement sélectionnée.

Note d'application : précise le contexte d'utilisation du descripteur. Peut également servir à la constitution d'une base de terminologie.

5.2.1 Spécialisation des branches

Kentika permet de créer plusieurs branches ou plusieurs thesaurii. Il permet également de créer plusieurs rubriques utilisant le thesaurus : chacune peut être alimentée soit par tout descripteur, soit par les descripteurs appartenant à une branche bien précise.

Ainsi, la rubrique « Zone géographique » pourrait n'être alimentée que par des descripteurs de la branche « Géographie ».

5.2.2 Importer - exporter

Kentika permet d'échanger tout ou partie du thesaurus. Différents organismes ont élaboré des thesaurii : nombre d'entre eux sont disponibles au format Kentika. C'est le cas du thesaurus d'Interdoc, du thesaurus du management de l'Aciege, de celui des Chambres de Commerce : Delphes ou encore de la santé : MeSH.

5.2.3 Thesaurus de référence dans la profession documentaire

Existant depuis 1987, le thesaurus de Kentika a été utilisé par un grand nombre d'institutions. Le respect du formalisme, la souplesse offerte, les contrôles d'intégrité font du gestionnaire de thesaurus de Kentika un outil précieux dans la constitution de thesaurii métier.

5.2.4 Tag et mots clés libres

Les indexations libres, proposées aujourd'hui et largement répandues, donnent de plus en plus la main aux utilisateurs qui peuvent alors participer en affectant des tags à des documents.

Gérer les tags via une branche spécialisée du thesaurus permet de disposer d'un vocabulaire contrôlé, de retenir des formes spécifiques (exemple : minuscule et singulier), de bénéficier de synonymes et de traduction.

5.2.5 Thesaurus et indexation automatique

Lors du versement d'un fichier à la base de données ou de l'indexation à partir d'une URL, Kentika analyse le contenu du texte proposé par rapport au thesaurus et en extrait les expressions rencontrées.

Dans une application de veille : il sera important de disposer d'un corpus de termes large et pertinent afin que l'extraction soit la plus précise possible.

Cette analyse est aussi couramment appelée « agents intelligents ».

5.2.6 Plan de classement & recherche plein texte

Lorsque la productivité est un objectif et que l'indexation est participative (i.e. : non systématiquement réalisée par des professionnels), l'indexation doit être plus légère et sera complétée lors des recherches par une recherche directement dans le contenu.

Dans ce cas, le thesaurus peut être vu comme plan de classement qui pourra être plus compact.

Si un thesaurus professionnel comporte une moyenne de 10 000 expressions, un plan de classement pourra être limité à quelques centaines.

5.3 Les outils de gestion du thesaurus

L'affichage du thesaurus bénéficie de diverses présentations, suivant l'usage que l'utilisateur veut en faire.

5.3.1 Mode hiérarchique

Particulièrement adapté à une utilisation comme un index permettant d'accéder au contenu du fonds documentaire. En option lors de l'affichage en mode hiérarchique : ne présenter que les branches pour lesquelles il existe des documents dans la base de données.

5.3.2 Mode alphabétique

Permet un accès aux notices des descripteurs : affiche l'ensemble des relations, modifie la fiche du descripteur (pour les gestionnaires autorisés à gérer la branche à laquelle le descripteur appartient).

5.3.3 Saisie d'un descripteur

Les relations gérées sont : termes génériques (poly-hiérarchie autorisée), termes spécifiques, employé pour (ou synonyme), terme associé et note d'application.

6 Communiquer : email

Kentika dispose d'un outil de génération de emails permettant d'envoyer des messages à la demande, sur événement ou à périodicité régulière.

La mise en page d'un email exploite toutes les possibilités offertes par le générateur de pages du portail : modèles ouverts et personnalisables, insertion de balises, génération de liens, prise en compte des autorisations du destinataire.

6.1 Le générateur de emails : contextes d'utilisation

6.1.1 Messages à la demande

En différentes circonstances, un envoi de email est proposé.

Quelques exemples :

- Un utilisateur sur le portail identifie des informations à signaler à un autre utilisateur ou à toute une liste de diffusion : une option lui permet de générer un email en indiquant l'objet et en sélectionnant une maquette de présentation.
- Un gestionnaire identifie un numéro de revue non reçu, il peut immédiatement le réclamer au diffuseur en lui adressant un email.

6.1.2 Messages sur événement

Le logiciel exploitant son propre générateur : des emails pré définis sont automatiquement émis sur certains événements.

Quelques exemples :

- Une potentielle anomalie est détectée par l'outil intégré d'auto-surveillance : un email est adressé à l'administrateur.
- Un ouvrage réservé vient d'être retourné par le précédent emprunteur : un message prévient de sa disponibilité.

Le scriptage de l'application permet de créer des emails suivant des événements particuliers.

Exemple : un utilisateur fait une recherche sur le portail et ne trouve aucun résultat : un email est alors généré pour que l'administrateur soit prévenu. Il peut alors lui porter assistance.

6.1.3 Messages programmés

Ils concernent principalement deux cas de figure : les alertes pour signaler des nouveautés et les messages de relance.

Kentika signale les nouveautés

Kentika propose deux processus pour signaler des nouveautés. L'une est très personnalisée : la diffusion sélective d'informations ; l'autre est plus générique : la veille diffusée via Kentika Press.

La DSI

Elle utilise des requêtes (une requête correspond à un axe de veille) qui sont soit préparées par les gestionnaires, soit réalisées par les utilisateurs eux-mêmes. Les informations envoyées à chaque destinataire tiennent compte des autorisations d'accès aux informations. Des outils statistiques comptabilisent le nombre de clics générés par chaque envoi de DSI.

NB : il est également possible de s'abonner au flux RSS correspondant à la requête de DSI.

Email de DSI ou de produit presse

L'exemple ci-dessus exploite une des maquettes proposées en standard.

En cliquant sur un article présent dans le message (DSI ou produit presse), l'utilisateur accède à l'article lui-même sur le portail. Si l'article est au format PDF, c'est le fichier pdf qui s'ouvre. S'il comporte un lien externe, c'est la cible de ce dernier qui est présentée.

Point important ici : l'utilisateur n'a pas besoin de revenir sur son email d'origine pour visualiser les autres articles : ils lui sont restitués dans le même ordre.

Les produits de veille ou de presse

Des documents (exemples : des articles de presse, des événements ou agenda, des brèves...) peuvent figurer sur différents produits de veille ou de suivi de presse (exemples : bulletins de veille ; revue des sommaires ; panorama de presse...).

Pour en savoir plus : voir le document « VeillePress ».

Emails de relance

Différentes situations peuvent engendrer des messages de relances.

Exemples :

- un emprunteur est en retard ;
- un numéro de revue n'est pas arrivé ;
- une commande auprès d'un fournisseur n'est toujours pas livrée.

Des emails de relance peuvent être générés automatiquement auprès des « retardataires ».

