

IBM SolutionsConnect 2013

L'IBM TechSoftware nouvelle génération

28, 29 et 30 août
IBM Client Center Paris



#solconnect13

Transformez vos opportunités en succès



IBM SolutionsConnect 2013

L'IBM TechSoftware nouvelle génération

ISO03 - Découvrir les connaissances dans les sources de données textuelles

IBM Content Analytics with Enterprise Search

Jean-Marc Langé

Agenda

- **L'information dans l'entreprise**
- **L'analyse de contenu**
- **IBM Content Analytics**
- **Exemples d'utilisation**
- **L'interface utilisateur**
- **Architecture**
- **Administration**
- **Pour aller plus loin**

Le chaos de l'information

Augmentation du fardeau pour les utilisateurs qui se désengagent

Détruit trop tôt ...
gardé indéfiniment

Stratégies de rétention et de mise à disposition non définies ou non respectées

Silos multiples

Pas d'index

Incomplet, douteux, inutilisable

Pas localisable

Pas géré désorganisé

Perte d'information

Perte du contexte

Croissance des risques de litiges et des coûts

Explosion des volumes et organisation en silos

Emails, documents, fichiers, ... papiers

Volume - Formats - Emplacements

Pas réutilisable

Difficulté croissante à disposer d'une information de confiance

Pas localisable

Copies multiples

Pas traçable

Pas digne de confiance

Pas de contrôle. Risques opérationnels importants. Grands risques de mauvaise information. Aucune visibilité.

L'explosion de l'information

Régulation

70% des responsables affirment que la régulation justifie la croissance des capacités de stockage. Ils prennent conscience qu'ils **doivent** classifier, gérer, supprimer les données et non ajouter continuellement des ressources de stockage.

Volume

Chaque jour génération de **15 petabytes** d'information. 988 exabytes d'information digitale en 2010.

Diversité

80% de l'information générée est non structurée. Importance croissante des emails et des flux de réseaux sociaux

Efficacité

Une société de 1,000 employés dépense **5 millions €** par an pour trouver une information stockée sur ses serveurs. **42%** des managers affirment utiliser une mauvaise information au moins une fois par semaine.

Agenda

- L'information dans l'entreprise
- L'analyse de contenu
- IBM Content Analytics with Enterprise search
- Exemples d'utilisation
- L'interface utilisateur
- Architecture
- Administration
- Pour aller plus loin

Dévoiler un aperçu utile du contenu

Que faire avec Content Analytics ?

Comprendre la
demande clients
en anticipant.



Détecter les
déclarations
frauduleuses avant
réglement.



Redéployer
dynamiquement
les ressources aux
secteurs les plus
tendus.



Sauver des vies en
identifiant
rapidement les
risques.



Libérer la valeur de l'information non structurée !

Est-ce votre réalité quotidienne ?

Difficulté de retrouver et exploiter l'information ...

- Ne pas trouver l'information quand nécessaire ; des décisions basées sur une mauvaise information ... le modèle "rechercher et espérer".
- Garder tout pour toujours; l'incapacité humaine de traiter les volumes ... explosion des coûts et des risques.

Impossibilité de baser les décisions sur l'ensemble de l'information ...

- L'informatique décisionnelle est limitée aux données structurées ; pas d'exploitation des 80 % d'information non structurée.
- Solutions actuelles d'analyse de texte trop complexes, mise en oeuvre longue; manque de souplesse pour l'évolution.

Les approches traditionnelles convergent

La recherche par mots-clés n'est pas suffisante

“Pouvoir rechercher dans les informations non-structurées est une nécessité pour toutes les applications, de l'intranet aux référentiels de documents.”

– Whit Andrews, Rita Knox Gartner

Analyser le contenu non-structuré est indispensable

“L'accès aux données structurées, même valorisées par de l'Informatique Décisionnelle ou de l'Analyse Prédictive, manque d'indices contextuels pour la bonne prise de décision. Services clients, finance, ... impliquent une grande diversité de canaux”

– Craig Le Clair Forrester

Impact croissant dans l'activité

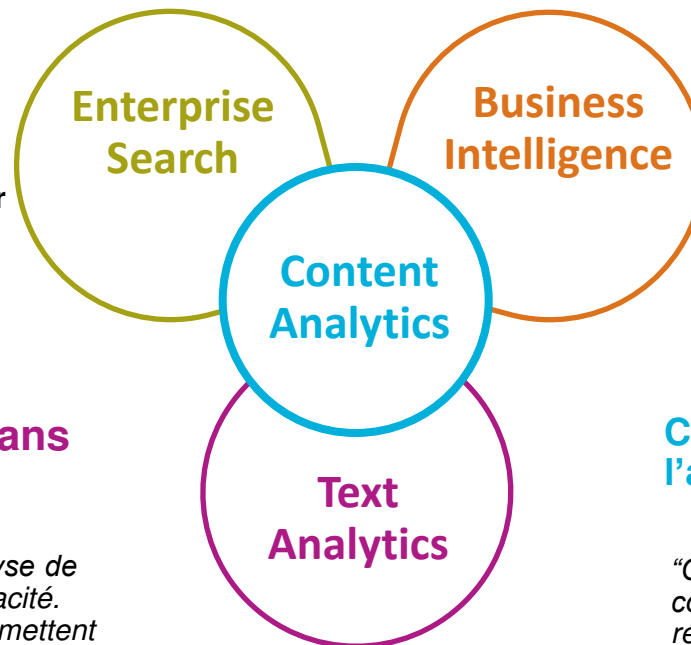
“Les pionniers utilisant l'analyse de texte ont déjà gagné en efficacité. Les autres organisations se mettent en retrait.”

– Sue Feldman IDC

Convergence vers l'analyse de contenu

“Chaque organisation doit comprendre comment l'analyse de contenu peut répondre à ses préoccupations cruciales; dès maintenant pour bénéficier de la disponibilité de ces nouveaux outils.”

– Rita Knox Gartner



De l'information brute à une appréciation efficace

Découvrir la valeur du contenu par une approche visuelle unique

Agréger et extraire de sources différentes

... pour générer des collections depuis des sources multiples, internes et externes, mais aussi de types différents : référentiels de documents, données structurées, réseaux sociaux, internet et autres.

Organiser, analyser et visualiser

... documents et données de l'organisation en identifiant des tendances, modèles, corrélations, anomalies, contextes.

Chercher et explorer pour puiser

... pour trouver confirmation ou découvrir de nouveaux éléments sans avoir à développer des modèles et déployer des systèmes complexes.

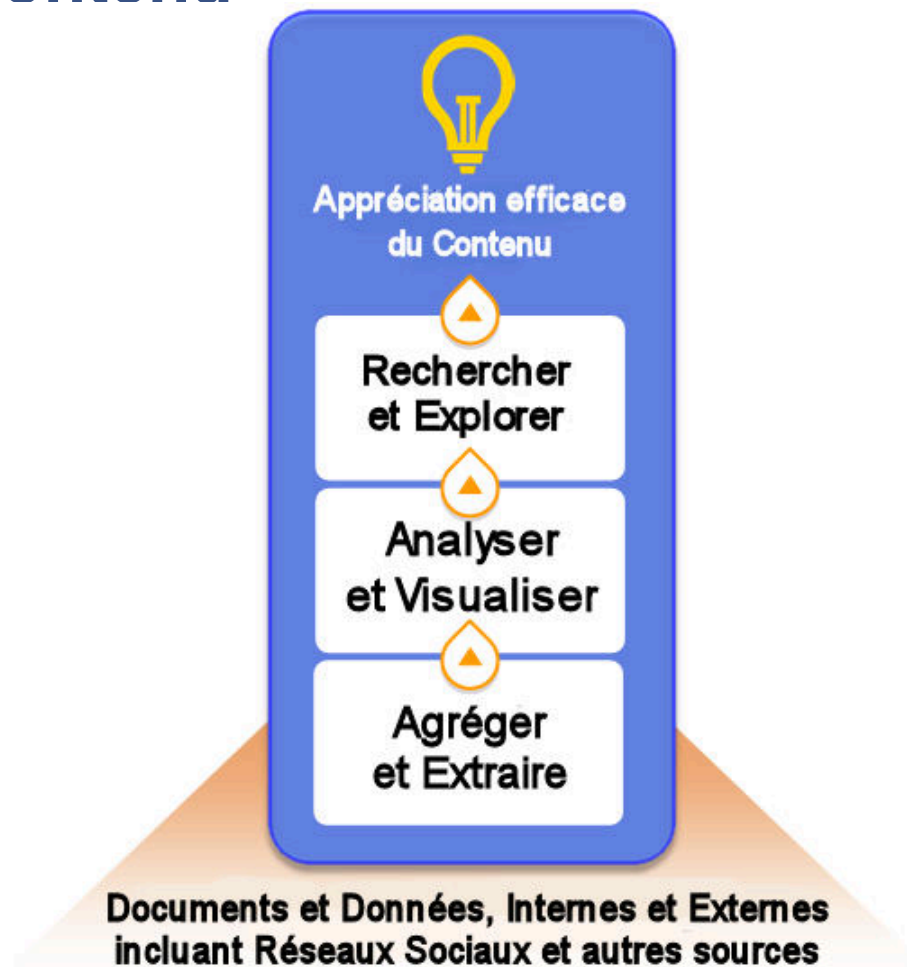


Agenda

- L'information dans l'entreprise
- L'analyse de contenu
- **IBM Content Analytics with Enterprise Search**
- Exemples d'utilisation
- L'interface utilisateur
- Architecture
- Administration
- Pour aller plus loin

IBM Content Analytics une solution pour une appréciation efficace du contenu

- Transformer une information brute en **contenu exploitable** sans avoir à développer des modèles et déployer des systèmes complexes.
- Dédire en quelques **heures** ou **jours** ... au lieu de semaines ou mois. Outil de **recherche** et **exploration** du contenu **d'utilisation aisée** par tout professionnel.
- **Flexible** et **extensible** pour toute évolution.



Libérez l'information emprisonnée dans le contenu

IBM Content Analytics



Documents et Données, Internes et Externes
incluant Réseaux Sociaux et autres sources

- **Trouver rapidement l'information pertinente**
 - Enterprise Search
- **Evaluer le contenu et ne garder et gérer que l'information de valeur**
 - IBM Content Collector (optionnel)
Chargement dans Filenet P8 et CM8
- **Adapter efficacement aux besoins et spécificités**
 - IBM Content Analytics Studio
 - IBM Content Classification (optionnel)
- **Permettre les approfondissements par l'intégration avec d'autres applications**
 - Solutions IBM ECM et ACM
 - Systèmes IBM Cognos and SPSS
 - IBM InfoSphere and Netezza Data Warehouse Systems
 - Applicatifs métiers

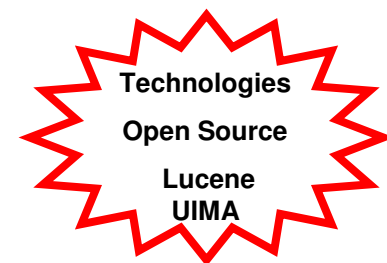
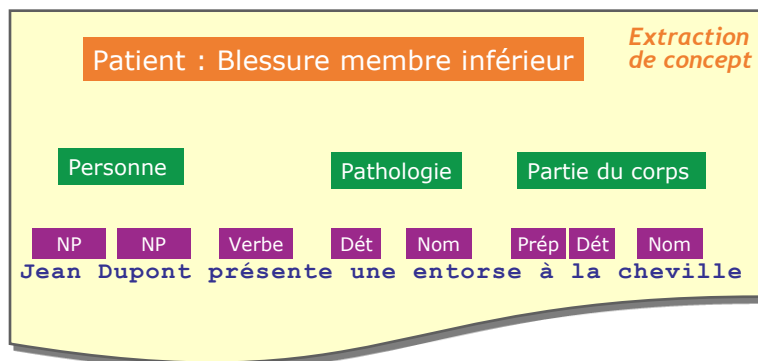
Content Analytics permet des analyses à grande échelle

Consolide les conclusions, étend la compréhension



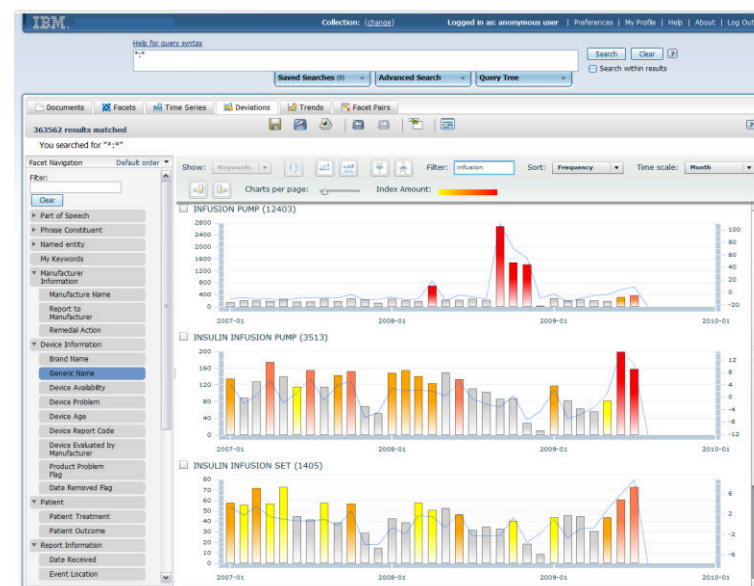
Les sources

(ECM, Fichiers, Web, BdD, ...)



Visualisation

Concepts et sources sont présentés dans l'interface utilisateur



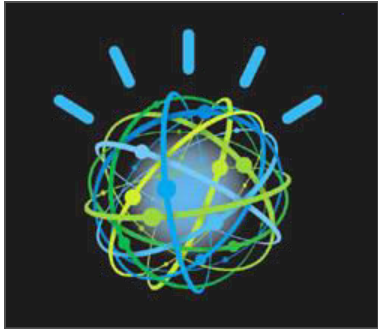
- **Analyse** document par document
- **Consolidation** des conclusions
- **Evaluation** sur de gros volumes d'informations
- **Inconcevable** par traitement manuel (hors coûts / délai)

Qu'est-ce que **NLP** et **UIMA**

- Natural Language Processing (NLP), **la pierre angulaire pour rendre possible l'interaction directe entre les ordinateurs et les langages humains**
 - Watson utilise **IBM Content Analytics** pour les fonctions NLP
- Unstructured Information Management Architecture (UIMA), **une architecture ouverte pour le traitement de texte et la définition de solutions d'analyse de texte**
 - Mise en oeuvre dans de nombreuses solutions IBM
 - Standard Open Source OASIS
 - Projet supporté par Apache Software Foundation



Redécouvrez votre contenu de façon plus intelligente



IBM Watson (Jeopardy)

Challenge Technique

Traitement de langage naturel (permettre à un ordinateur de comprendre une question exprimée dans une langue courante/usuelle et non pas spécifique de l'ordinateur).

"... un outil d'exploration de l'information capable de comprendre votre question et d'y répondre en langage naturel"

*Dr. David Ferrucci
Principal Investigator
Watson project*

"Smart"?

IBM Content Analytics (Studio) Cette analyse sémantique repose sur le composant IBM Analytics Studio, dispositif d'IBM Content Analytics spécialisé dans la création de règles / dictionnaires pour analyser du texte.

Solution

Elle s'appuie sur l'analyse sémantique – conversion d'un texte en un ensemble d'entités exploitable par l'ordinateur ; on effectue une analyse de la phrase : identification de la langue, des composants de la phrase – mots, verbes, compléments -, extraction d'entités (noms propres, lieux, ...), analyse de ces entités pour extraire des concepts ...).

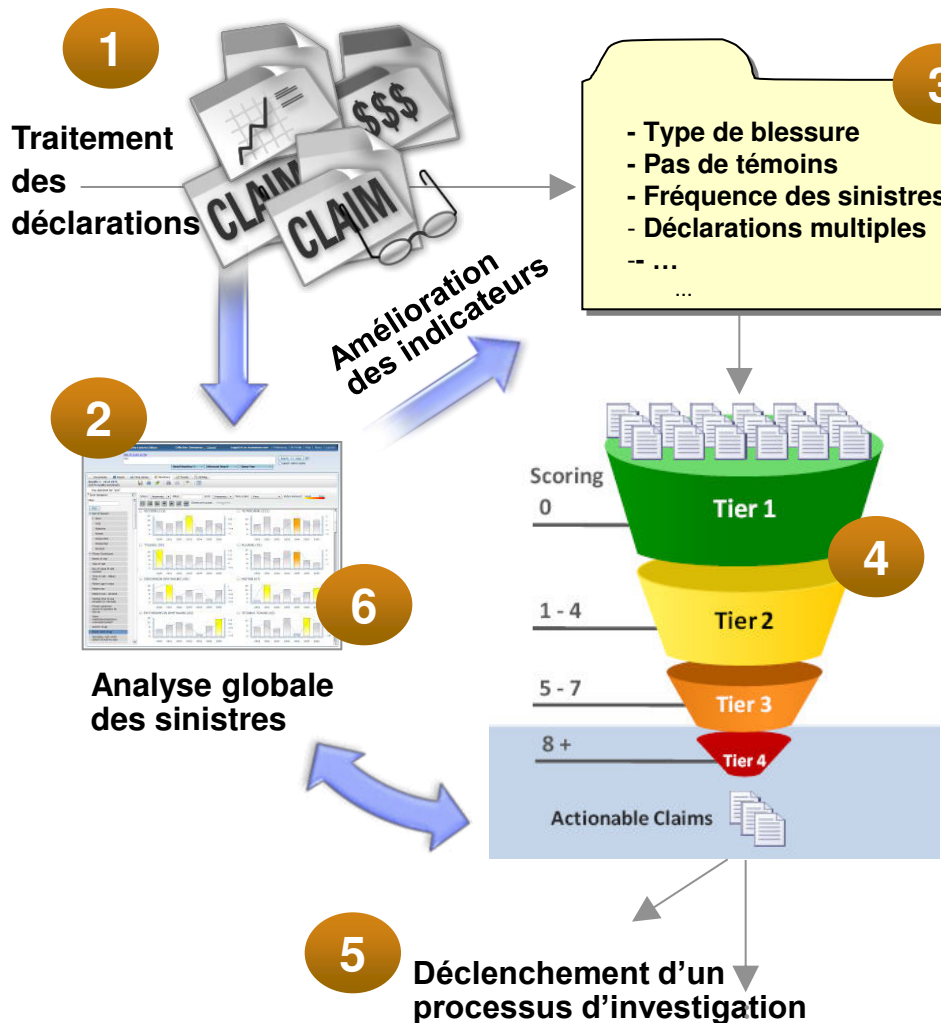
*Industrie : Toute industrie
Valeur: facilite la décision
Solution : content analytics*



Agenda

- L'information dans l'entreprise
- L'analyse de contenu
- IBM Content Analytics with Enterprise Search
- Exemples d'utilisation
- L'interface utilisateur
- Architecture
- Administration
- Pour aller plus loin

Dossiers d'assurance, détection de fraude et analyse des risques



1. Agrégation automatique de l'ensemble des informations, structurées et non structurées concernant le traitement des sinistres.
2. Détermination de corrélations, tendances, anomalies, comportements... visualisation et exploration pour de nouvelles perspectives du traitement des dossiers.
3. Indicateurs dans le traitement en ligne des dossiers. Identification de tentative de fraude et évaluation des risques.
4. Qualification des risques.
5. Traitement par exception des dossiers à risques.
6. Amélioration continue des indicateurs par réitération.

Les apports de Content Analytics

Banque de détail Satisfaction clients



- **Analyse** : Centre d'appels, commentaires en ligne
- **Pour** : Comportement clients
- **Bénéfices** : Satisfaction client, optimisation des offres, nouvelles opportunités

Marché de détail



- **Analyse** : Centre d'appels, commentaires en ligne
- **Pour** : Image de marque
- **Bénéfices** : Satisfaction clients, efficacité marketing

Délinquance



- **Analyse** : Centre d'appels, rapports de police ...
- **Pour** : Résolution des cas, tendances
- **Bénéfices** : Sécurité, déploiement des ressources

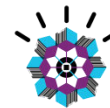


Santé



- **Analyse** : Rapports d'intervention
- **Pour** : Protocoles de traitement
- **Bénéfices** : Traitements, formulations, délais

Industrie automobile



- **Analyse** : Rapports interventions, centre d'appels, commentaires en ligne
- **Pour** : Image de marque
- **Bénéfices** : Coût de garantie, satisfaction clients,

Télécommunications



- **Analyse** : Centre d'appels, emails, commentaires en ligne
- **Pour** : Analyse attrition, FQA
- **Bénéfices** : Fidélisation, satisfaction, offres

Fraude à l'assurance



- **Analyse** : Déclarations de sinistre
- **Pour** : Détection des fraudes, détermination des profils
- **Bénéfices** : Coût, délai

...et plus !

Agenda

- L'information dans l'entreprise
- L'analyse de contenu
- IBM Content Analytics
- Exemples d'utilisation
- L'interface utilisateur
- Architecture
- Administration
- Pour aller plus loin

IBM Content Analytics

Un nouvel univers d'analyse

■ Exploitation de texte

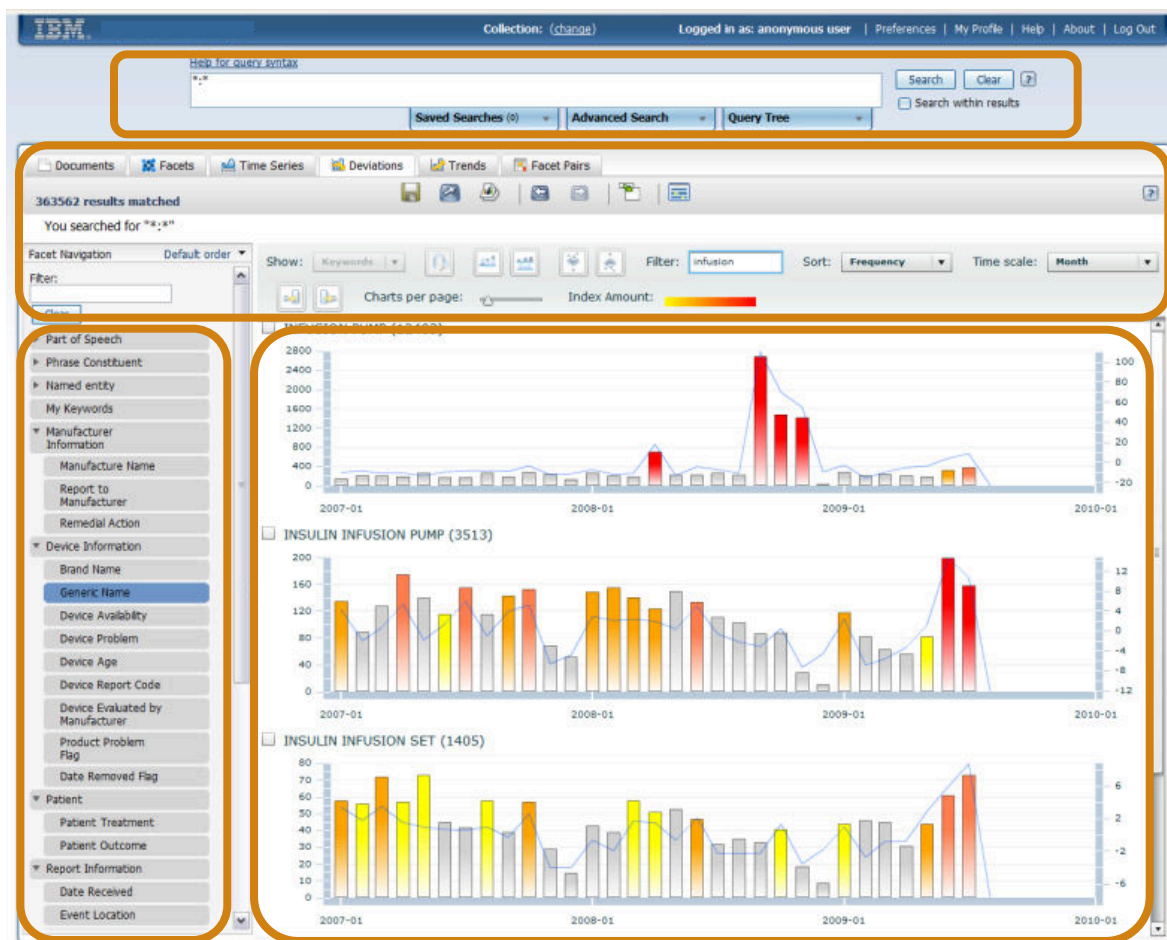
- Investigation par mots-clés
- Vues, filtre, exportation
- Extraction de concepts, relations, métadonnées, entités

■ Nouvelle compréhension du contenu des informations non structurées

- Mise en évidence de tendances, déviations, corrélations pour des recherches focalisées
- Fonctions d'extraction et de présentation standard et personnalisables
- Combinaison accès au contenu, entités, extraction de contexte, analyse, catégorisation avec fonctions d'exploration et de compte rendu

■ Intégration des informations structurées et non structurées

- Intégration et fonctions d'analyse pour Cognos 8 BI, InfoSphere Warehouse, IBM ECM, WebSphere Portal, et applicatifs métiers
- ETL disponible pour le non structuré
- Intégration du contenu avec applications, systèmes et processus



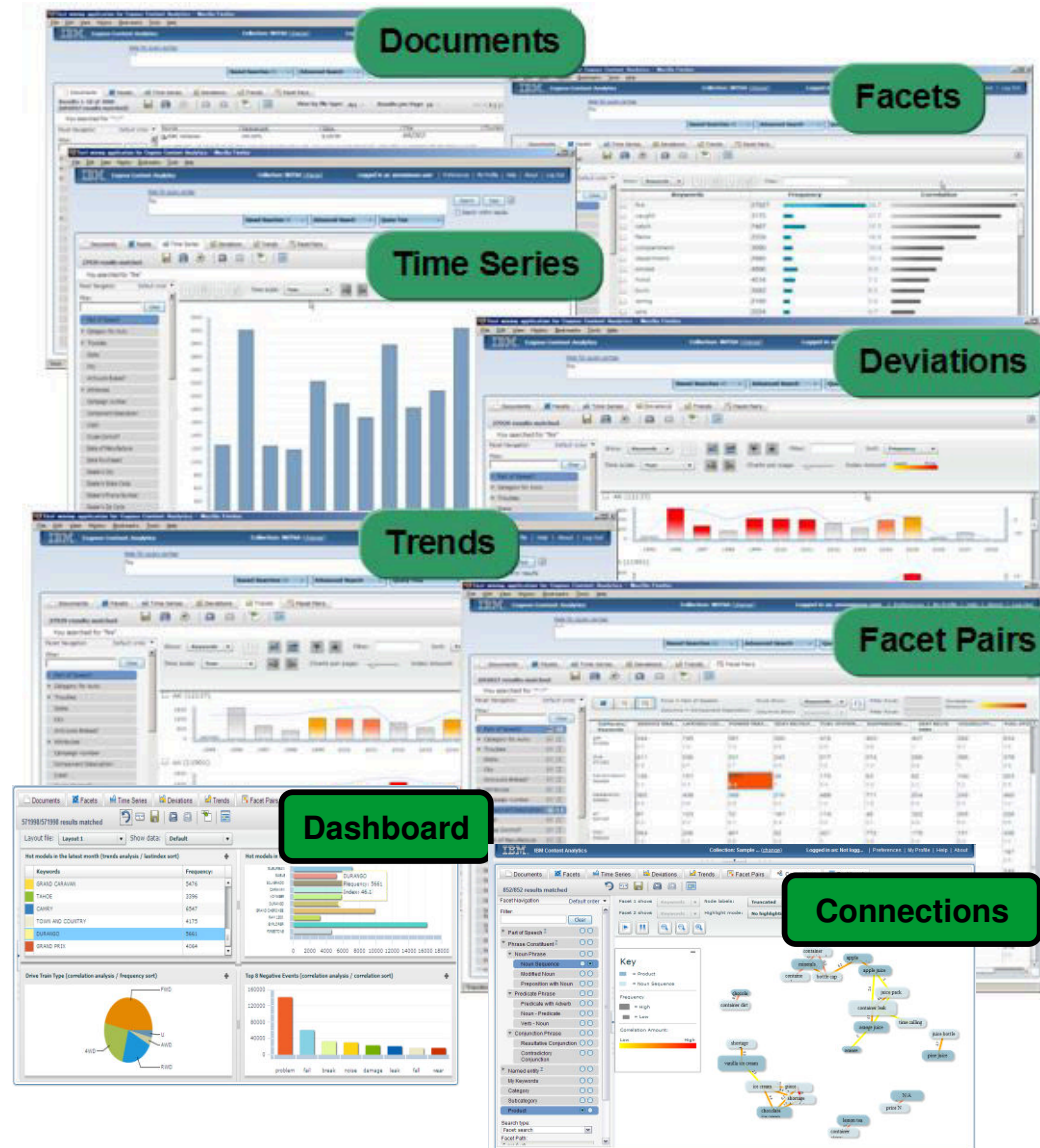
Explorer l'information

• Analyse interactive

- Faciliter la recherche de documents qui ont des valeurs d'index spécifiques par rapport à certaines facettes / dimensions
- Vues multiples en fonction de la nature des analyses
- Graphiques interactifs pour restituer l'information : zoom, tri, filtre

• 8 Vues

- **Documents** : liste les documents en fonction de la recherche
- **Facettes** : liste les mots-clés d'une facette
- **Série chronologique** : montre la fréquence dans le temps
- **Déviations** : montre les déviations entre les mots-clés
- **Tendances** : détecte les variations importantes dans le temps
- **Paires** : montrent la corrélation entre 2 dimensions / facettes
- **Connections** : montrent les connections entre les valeurs de facettes
- **Tableaux de bord** : vue de synthèse

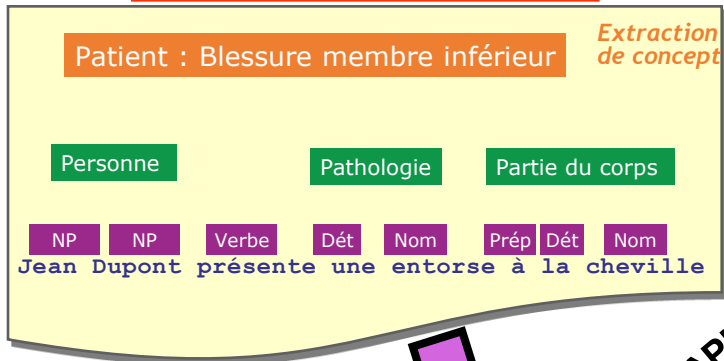


Agenda

- L'information dans l'entreprise
- L'analyse de contenu
- IBM Content Analytics with Enterprise Search
- Exemples d'utilisation
- L'interface utilisateur
- Architecture
- Administration
- Pour aller plus loin

Comment ça marche?

Content Analytics Studio



API REST Temps Réel

Annotateurs (UIMA)

Documents et données analysés
Avec identification de concepts

Crawlers

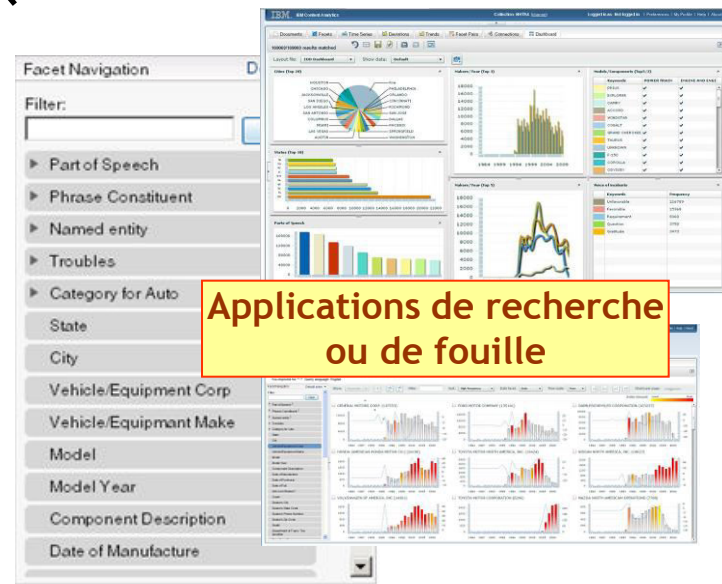
Crawlers



Données non structurées

Entreprise (Centre de Contact, logs de test, notes des concessionnaires / conseillers, messagerie, ECM, etc.) et externes (Web, forums utilisateurs, réseaux sociaux, etc.)

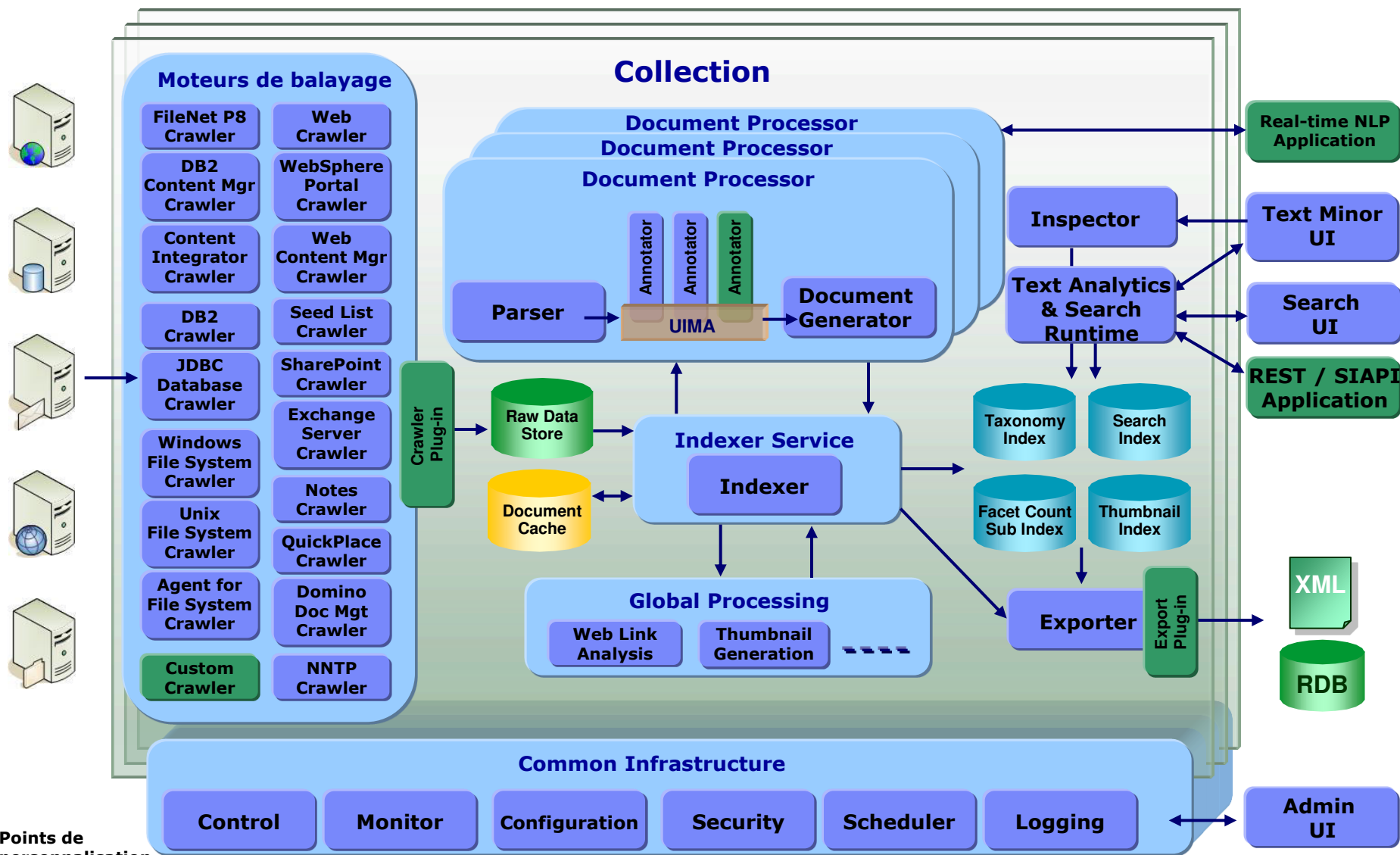
Applications de recherche ou de fouille



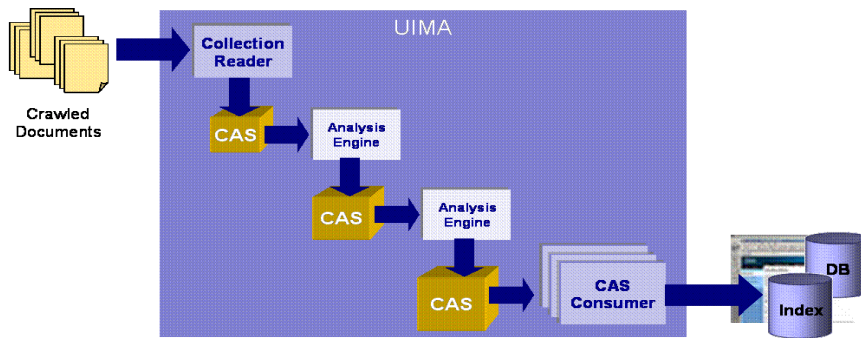
Applications tierces d'analyse / exploration



Architecture d'IBM Content Analytics

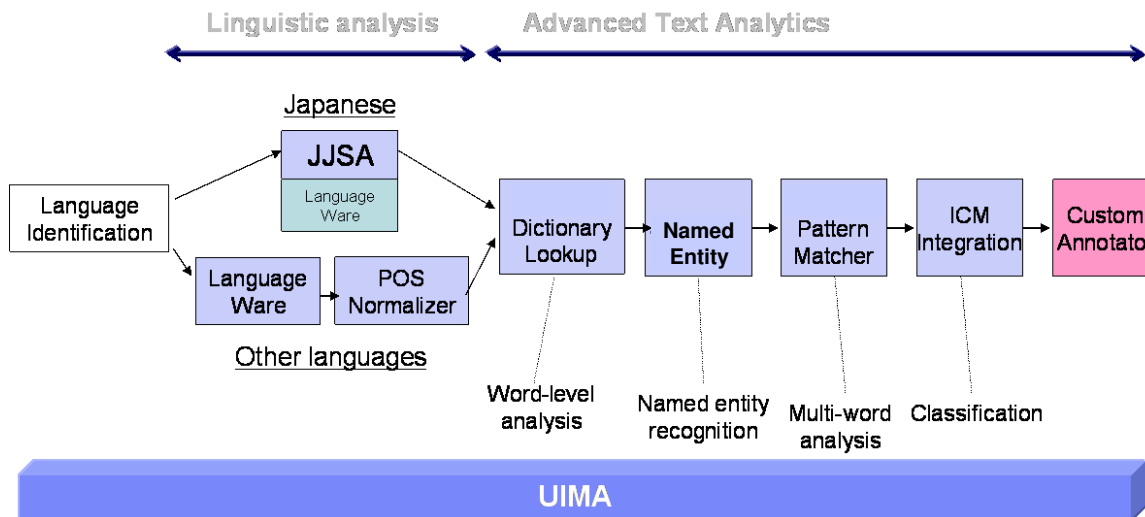


Unstructured Information Management Architecture



Contexte d'UIMA

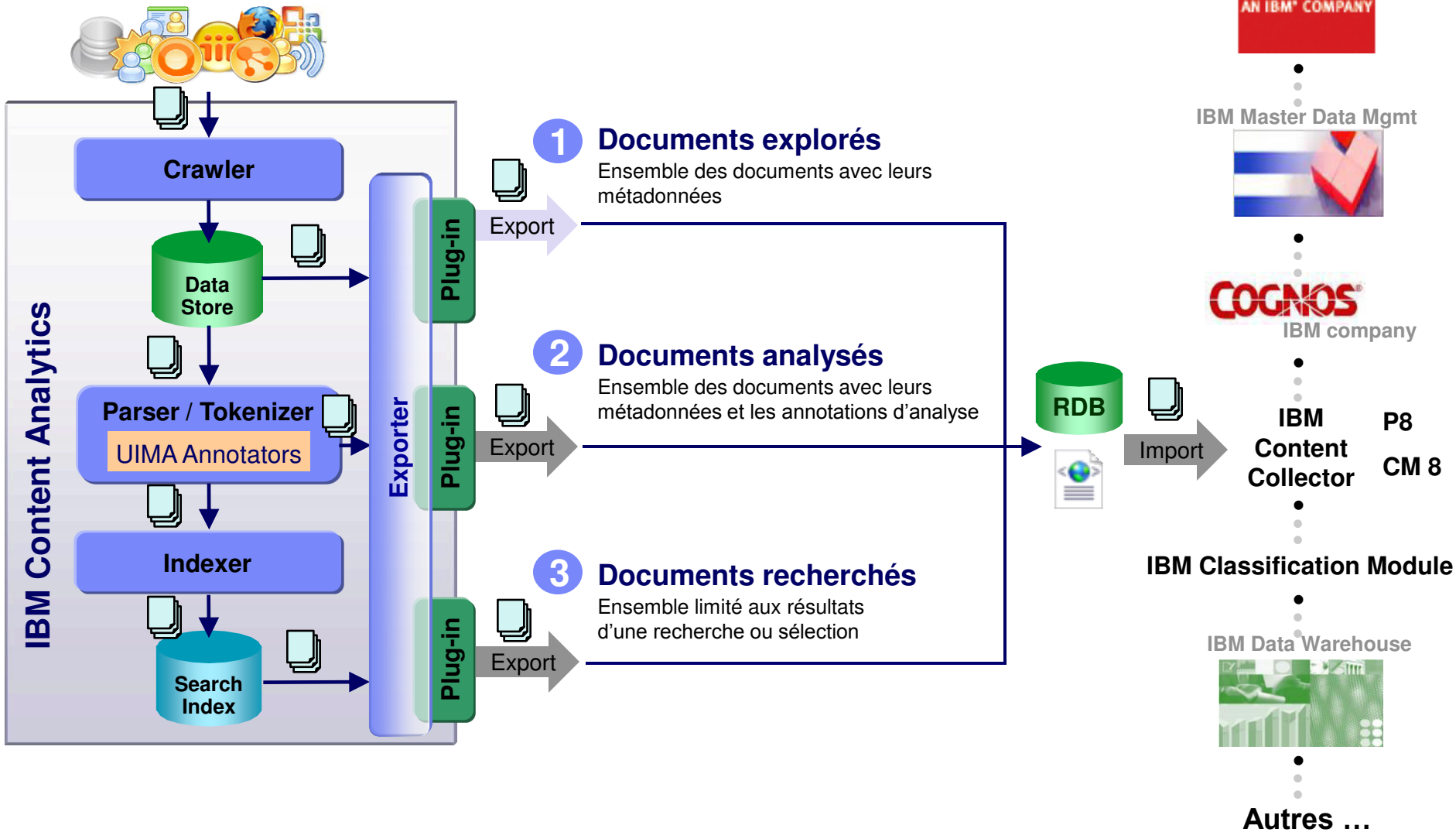
- Créé par IBM dans les années 1990
- Apache Incubator en 2006
- Standard OASIS en 2009



Implementation ICA

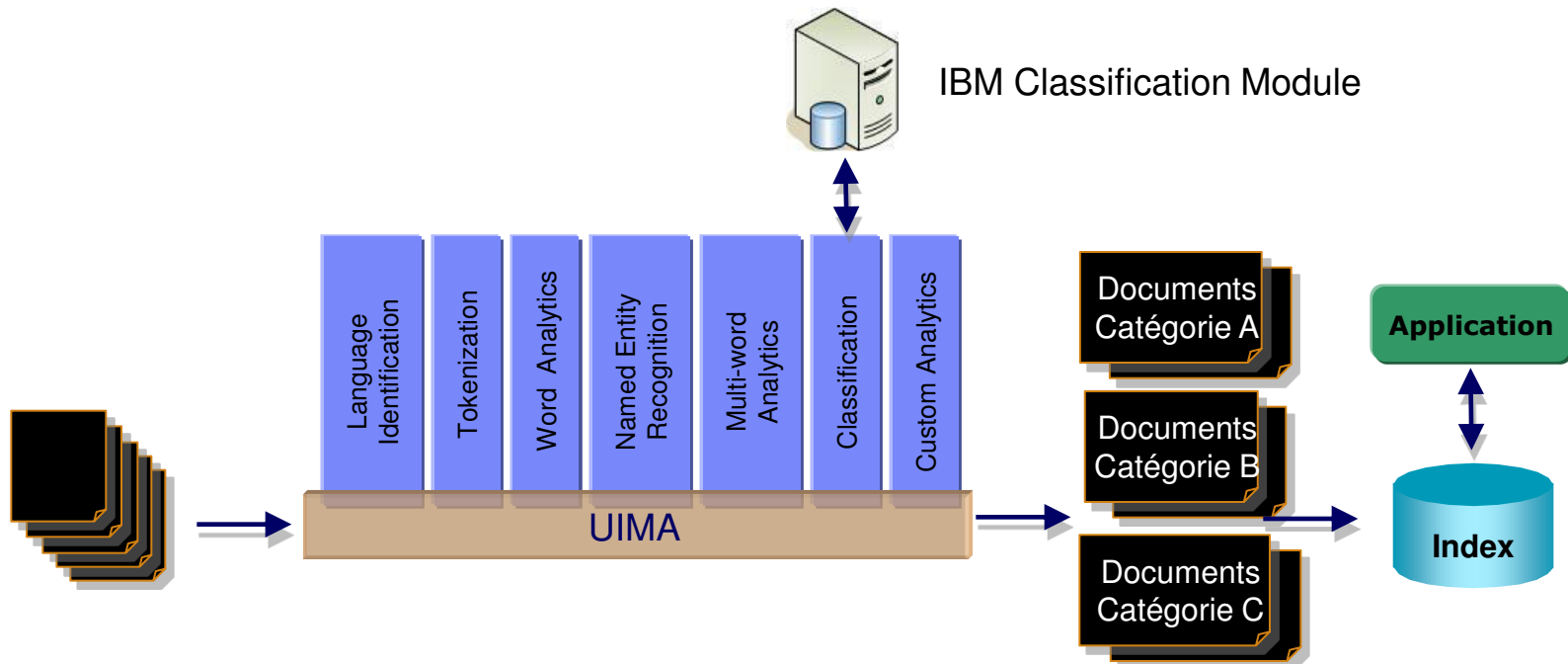
- Multi langues
- Combinaison des meilleurs technologies d'analyse
- Pipeline d'analyse de texte ouvert et personnalisable

Intégration en aval

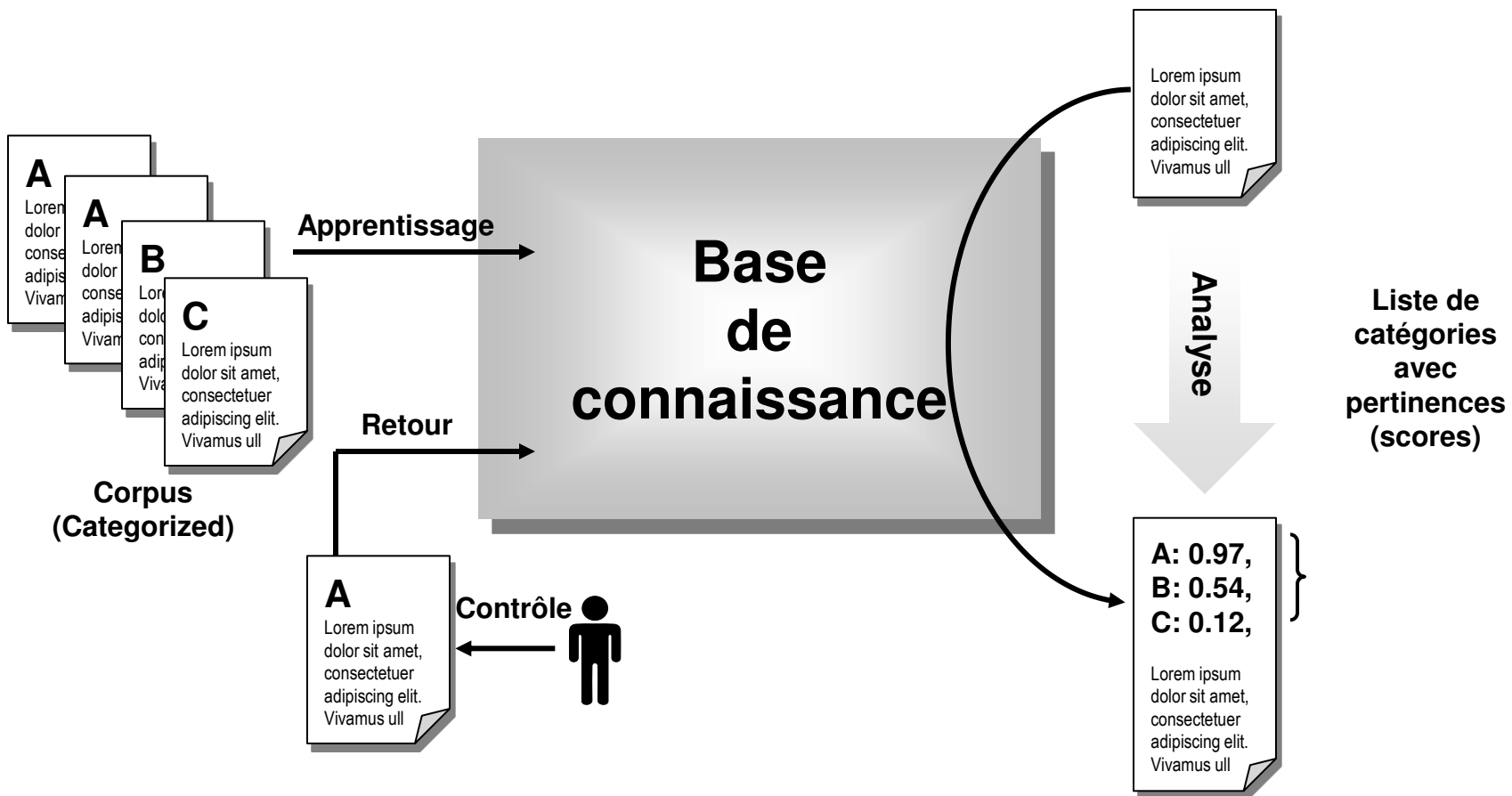


Intégration de IBM Content Classification

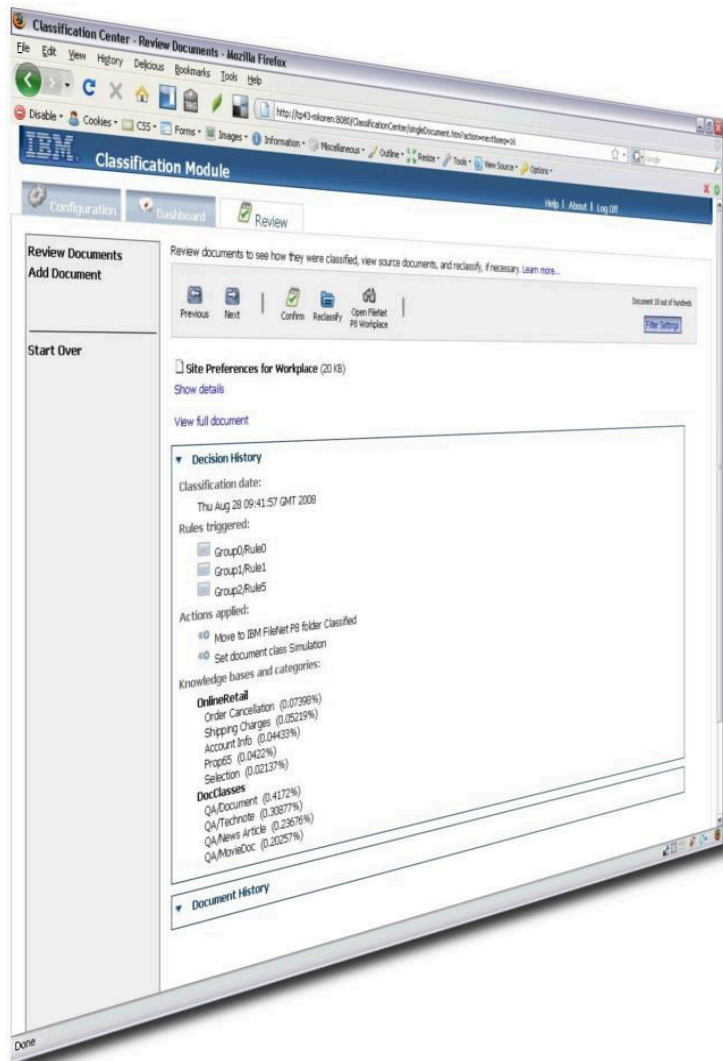
- **Classification automatique pour générer des métadonnées**
 - Classification basée sur des bases de connaissance et des règles
 - Exportation de documents de ICA pour alimenter une base de connaissance ICM
- **Catégories, coefficients de confiance et données extraites stockés dans des index**
- **Exemples d'utilisation**
 - Filtre de documents
 - Aide à l'analyse
 - Evaluation de pertinence et recherche conceptuelle basée sur les coefficients de confiance



Classification par contexte



IBM Content Classification

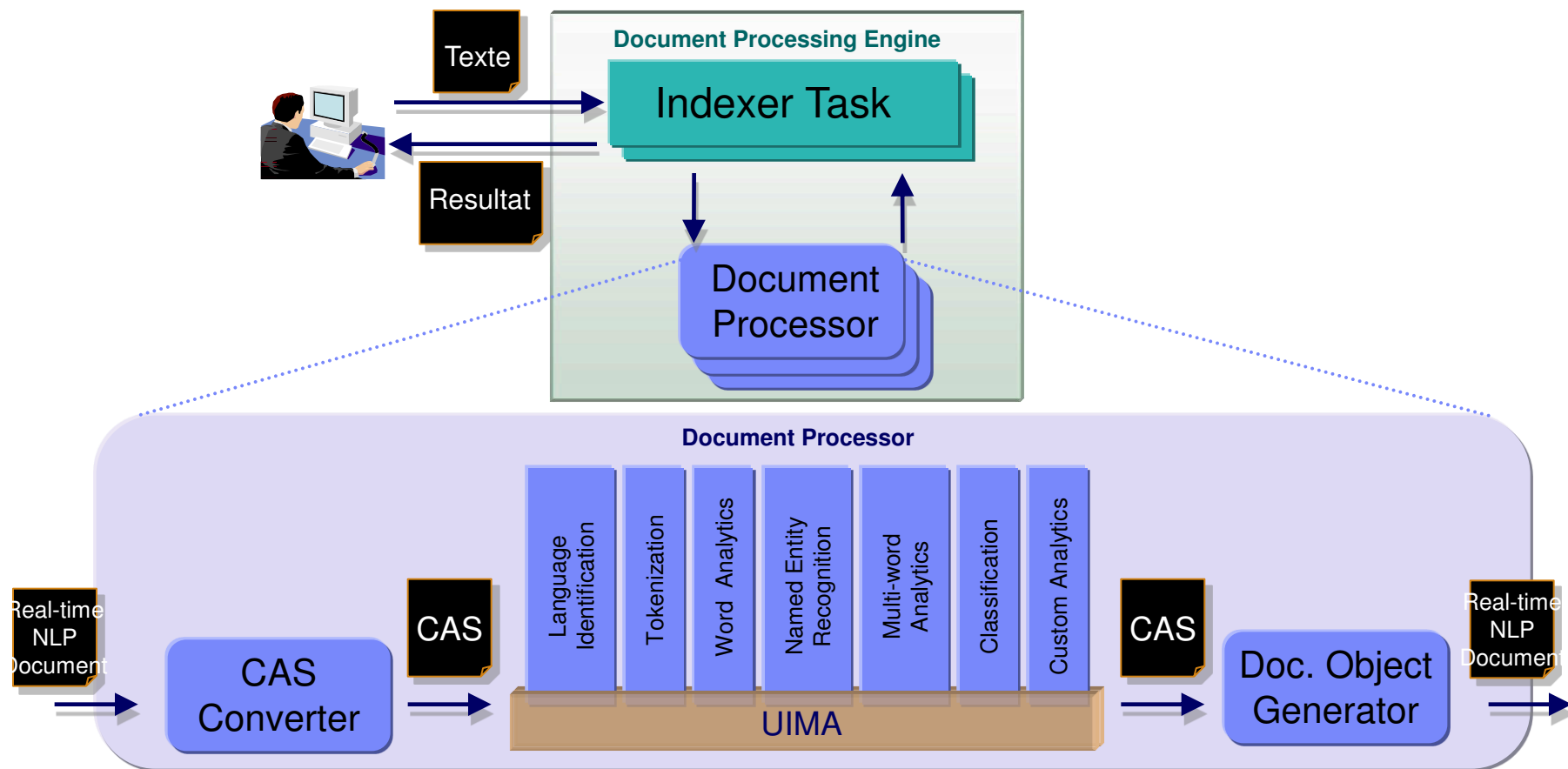


Diminution du délai de valorisation du contenu, réduction des coûts

- **Un serveur d'entreprise pour classifier textes, documents et emails**
- "Classifier" – Identifier le sujet ou la catégorie d'un document
- **Combinaison de règles et d'analyses statistiques avec bases de connaissance**
- Serveur d'entreprise : distribué, évolutif
- Intégré à l'architecture IBM ECM
 - **Content Collector,**
 - **Content Analytics,**
 - **P8 et CM 8**
 - **Datacap**
 - Ou utilisé seul, intégré applicatif métier
- **18 langues**

Analyse en temps réel

- Utilisation d'API REST pour analyse ponctuelle
 - Intégration applicatif métier



IBM Content Analytics Studio

- **Environnement de développement basé sur Eclipse**
 - Dictionnaires
 - Règles
 - Annotateurs
- Permet de se focaliser sur les concepts et les relations entre les points d'intérêt, de développer des analyseurs pour extraire ces éléments sans avoir à écrire du code.
- **Intégration avec IBM Content Analytics & Enterprise Search**
 - Exportation directe
 - Association de facettes

The screenshot shows the 'IBM Content Analytics Server connection' dialog box. It is divided into several sections:

- Server:** Contains fields for 'Host Name' (localhost), 'Admin Port' (8390), 'User Id' (esadmin), 'Port' (8394), 'Time Out' (100), and 'Password' (*****).
- Collections:** A list of available collections: 'IBM Products Content Assesment', 'IBM Products Search' (selected), 'Products Comments Analysis', and 'Sample Text Analytics Collection'. There are 'Update' and 'Remove' buttons next to the list.
- Collection:** A section for 'Collection details' with fields for 'Name' (IBM Products Search) and 'Id' (IBM_Products_Search).

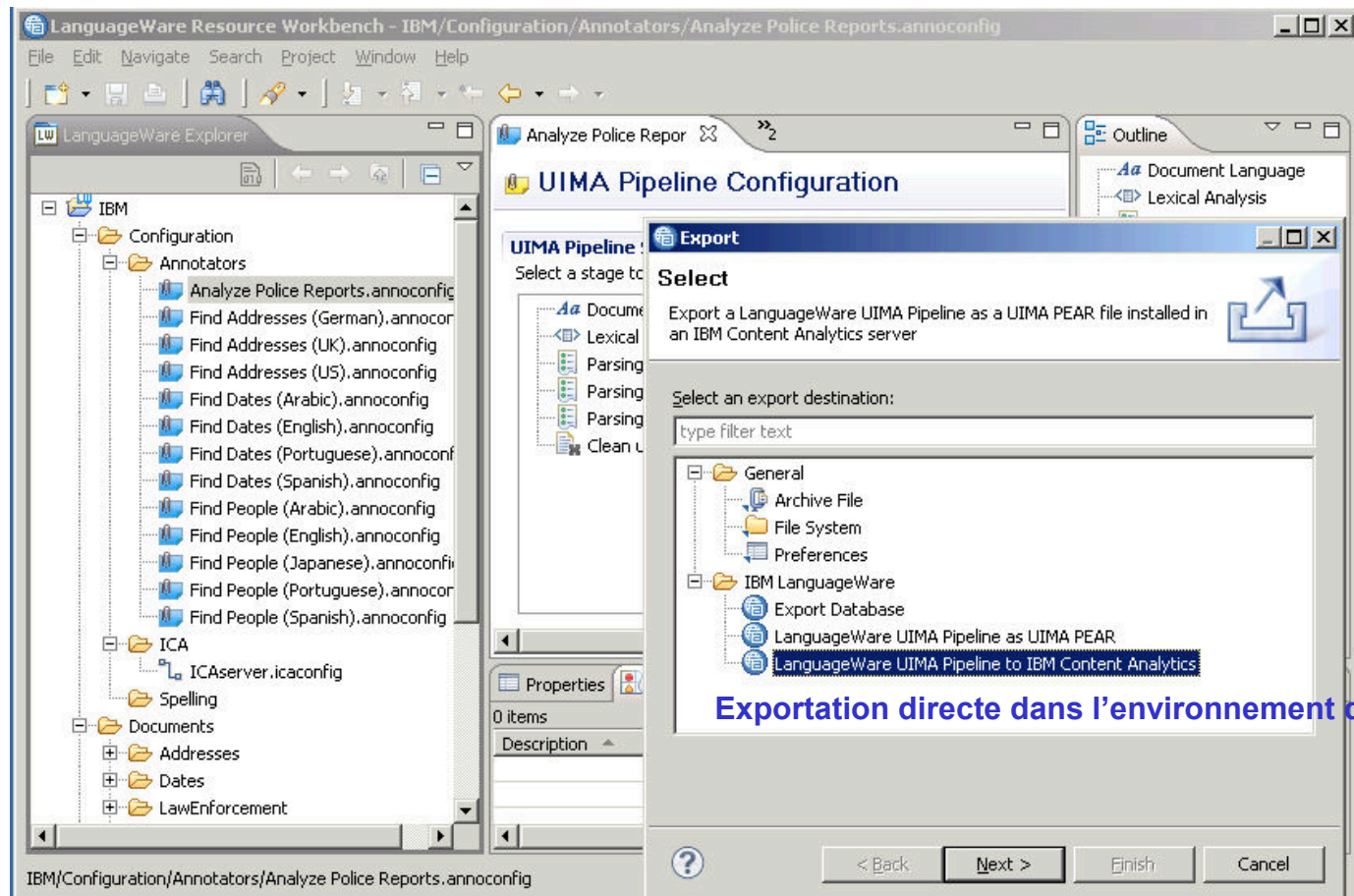
Les étapes de développement

1 Développement de l'analyse personnalisée avec IBM Content Analytics Studio

Dictionnaires
Règles
Annotateurs.

Développement et
gestion des
ressources

Test et validation



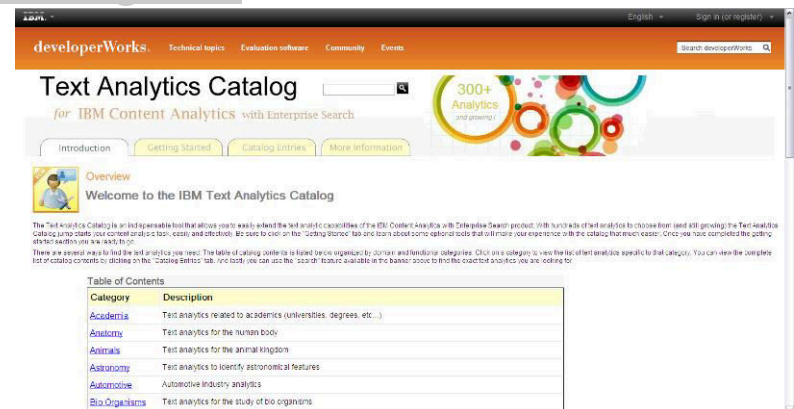
2 Exportation vers une Collection de ICA

En standard

- **Extraction d'entités**
 - E-mail
 - N° de téléphone
 - Code postal
 - URL
 - ...
- **Fonctions de standardisation**
 - Date
 - Montant
 - Distance
 - ...

Un catalogue de fonctions d'analyse

- Plus de 300 modules gratuits téléchargeables
 - Compatible UIMA
 - Un grand nombre de domaines
 - Accessible librement sur Internet
- <https://www.ibm.com/developerworks/mydeveloperworks/blogs/36db6433-2f12-4533-9c68-489067780bfd/entry/overview3?lang=en>

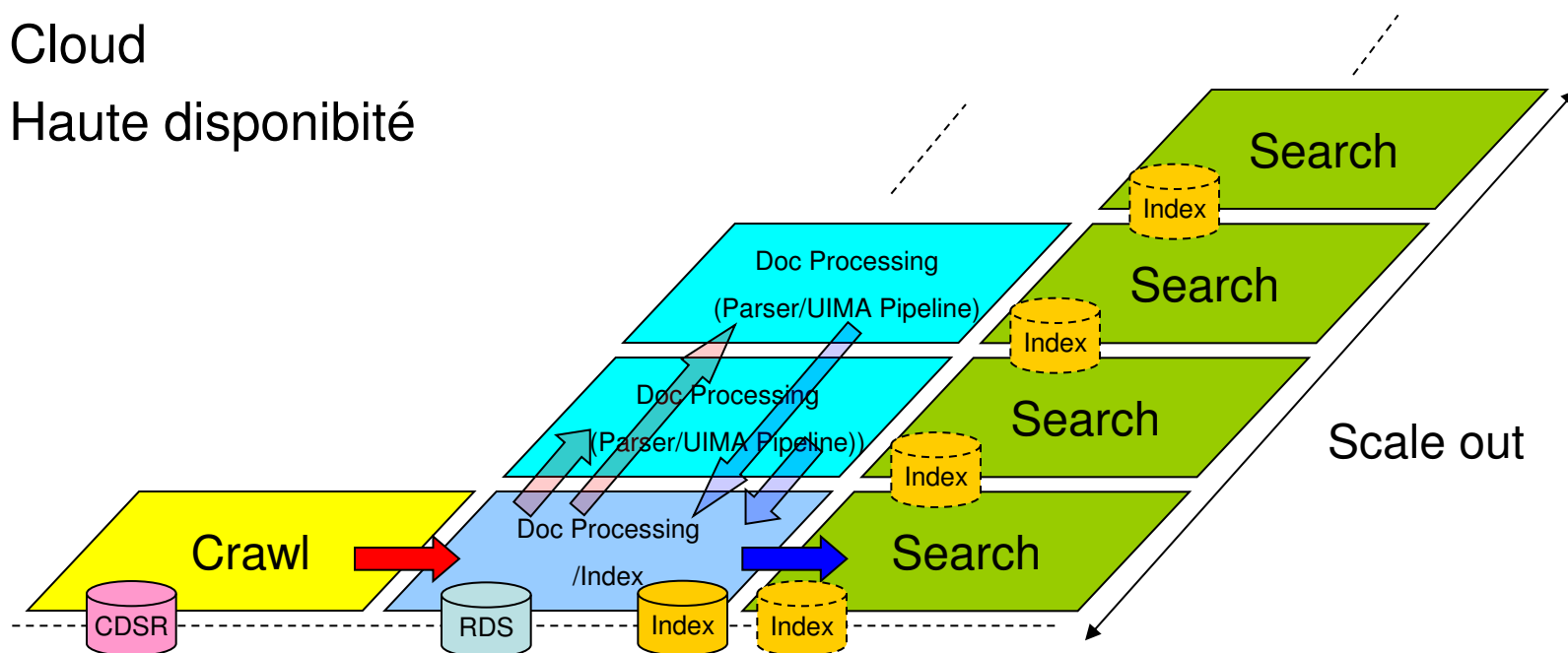


Category	Description
Academia	Text analytics related to academics (universities, degrees, etc.)
Anatomy	Text analytics for the human body
Animals	Text analytics for the animal kingdom
Astronomy	Text analytics to identify astronomical features
Automotive	Automotive industry analytics
Bio-Organisms	Text analytics for the study of bio organisms

Systeme évolutif

Environnement adaptable aux besoins (1 noeud à n noeuds)

- Traitement de documents n noeuds
- Recherche de documents n noeuds
- Cloud
- Haute disponibilité



Agenda

- L'information dans l'entreprise
- L'analyse de contenu
- IBM Content Analytics with Enterprise Search
- Exemples d'utilisation
- L'interface utilisateur
- Architecture
- Administration
- Pour aller plus loin

Administration du moteur de recherche

- Un tableau de bord de gestion
 - Collections
 - Système
 - Sécurité

IBM Content Analytics with Enterprise Search

Personnaliseur de recherche | Personnaliseur d'analyse | Déconnexion | Aide | A propos de

Collections | Système | Sécurité

* Créer une collection

Collections de recherche d'entreprise

	Explorer	Analyse syntaxique et index	Rechercher	
Douanes search	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Actions ▾
IBM Products Search	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Actions ▾
IBM Products Search with Classification	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Actions ▾
IBM Products search with security	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Actions ▾

Collections d'analyse de contenu

	Explorer	Analyse syntaxique et index	Analyse	
IBM Products Content Assesement	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Actions ▾
Products Comments Analysis	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Actions ▾
Sample Text Analytics Collection	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Actions ▾

La sécurité

• Une vue globale

Actions ▾

Sécurité de connexion à l'application

Paramètres de connexion à l'application

Activé

Préférences d'authentification
L'administrateur par défaut peut accéder à toutes les applications

Configuration du serveur LDAP

Nom d'hôte du serveur LDAP	
Numéro de port du serveur LDAP	389
Nom distinctif de base	
Utiliser les autorisations pour accéder au serveur LDAP	Non
Nom d'utilisateur LDAP	

Paramètres du jeton LTPA et du fichier de clés

Désactivé

Jeton LTPA pour la connexion unique

Valeur de délai d'attente du jeton

Nom de domaine du cookie

Mode interopérabilité LTPA Non

Chemin d'accès au fichier de clés LTPA

Domaine supplémentaire

Suffixe de nom d'utilisateur supplémentaire

Actions ▾

Sécurité au niveau des collections

Paramètre de gestion des identités

Activé

Intervalle d'actualisation: 3 jours

Tous les types de moteur de balayage qui prennent en charge la connexion unique

- Content Integrator
- Domino Document Manager
- Notes
- Quickr for Domino
- IBM Connections

Associations d'applications

All

Toutes les collections

- Analytics
- RestAPI
- Search
- Searchglobal

Paramètres de recherche sécurisée

IBM Products search with security

- Pré-filtrage activé
- Post-filtrage activé
- Modifier la sécurité au niveau du document

Actions ▾

Sécurité au niveau du système

Rôles d'administration

ID utilisateur	Rôle
esadmin	Administrateur par défaut

Droits utilisateur d'application

ID utilisateur ou groupe	Privilèges	ID applicati
utilisateur 3		Searchibn
utilisateur 1		Searchglc
user1	Sauvegarder les recherches, Gérer les indicateurs de documents	All
utilisateur 2		Searchsel

Données d'identification de connexion

Source de données	ID connexion
Système de fichiers Windows	C:\Program Files\IBM\ClassificationModule\Samples\SampleData\IBM Products
Système de fichiers Windows	C:\Program Files\IBM\es\samples\firststep\data\xml\xml-data
Système de fichiers Windows	c:\Autres_donnees\IBM Products
Système de fichiers Windows	c:\Autres_donnees\IBM Products 1

Répartition des rôles d'administration

- Selon la structure et l'organisation

The screenshot displays the IBM Content Analytics with Enterprise Search administration interface. The main window is titled "Ajouter un utilisateur" (Add user) and contains the following fields and options:

- ID utilisateur:** Administrateur
- Rôle:** Gestionnaire de collection
- Privilèges système:** La barre d'outils système est disponible
- Collections:** Toutes les collections, Collections spécifiques

Buttons for "OK" and "Annuler" are visible at the bottom of the form.

Overlaid on the right is a Mozilla Firefox browser window showing the "Centre de documentation d'IBM Content Analytics with Enterprise Search". The browser address bar shows the URL: `http://csld1cm:8889/help/index.jsp?topic=%2Fcom.ibm.discovery.es.nav.doc%2Fwelcome.htm`. The page content includes a "Sommaire" (Table of Contents) with the following items:

- Bienvenue
- Présentation du produit
- Installation
- Administration
- Sécurité
- Intégration
- Analyse de texte
- Maintenance
- Analyse du contenu
- Développement
- Exemples
- Référence
- Identification des incidents et support

The main content area of the documentation page features a "Recherchez des informations sur l'initiation, le dépannage, le téléchargement de correctifs ainsi que d'autres tâches" (Search for information on initiation, troubleshooting, downloading fixes, and other tasks) section. Below this, there is a sub-section for "Initiation" with links for "Notes sur l'édition" and "Configuration requise".

L'état du système

- Une vue globale

The screenshot displays the IBM Content Analytics with Enterprise Search management interface. The top navigation bar includes the IBM logo, the product name, and links for search and analysis customization, as well as help and about options. Below this, there are tabs for 'Collections', 'Système', and 'Sécurité'. A secondary navigation bar contains buttons for 'Configuration de serveur', 'Programme d'écoute de données', 'Analyse syntaxique', 'Rechercher', and 'Journal'. The main content area features a 'Ajouter un serveur' button and two status buttons: 'Démarrer les serveurs sélectionnés' and 'Arrêter les serveurs sélectionnés'. A table titled 'Informations sur le serveur' lists the server 'CSLD1CM' with columns for 'Nom d'hôte', 'Moteur de balayage', 'Contrôleur', 'Traitement de documents', 'Rechercher', 'De secours', 'Statut', and 'Actions'. All components for the listed server show a green checkmark, indicating they are operational. Below the table is a 'Vue de topologie' section, which currently shows a single server icon labeled 'CSLD1CM' with a status indicator.

Nom d'hôte	Moteur de balayage	Contrôleur	Traitement de documents	Rechercher	De secours	Statut	Actions
CSLD1CM	✓	✓	✓	✓		●	

Les applications de recherche et d'analyse

- Une vue de synthèse

IBM Content Analytics with Enterprise Search Personnaliseur de recherche | Personnaliseur d'analyse | Déconnexion | Aide | A propos de

[Collections](#) | [Système](#) | [Sécurité](#)

[Configuration de serveur](#) | [Programme d'écoute de données](#) | [Analyse syntaxique](#) | [Rechercher](#) | [Journal](#)

En savoir plus [»](#)

- Configurer des dictionnaires de synonymes
- Configurer les dictionnaires de mots vides
- Configurer les dictionnaires de mots avec degré de pondération
- Configurer la base de données des statistiques des requêtes

Statistiques de requêtes: [»](#)

Rechercher: **Nom du serveur:** CSLD1CM

Nom de collection	Débit de requête (5 minutes)	Réponses mises en mémoire cache	Résultats issus de la mémoire cache	Temps de réponse (millisecondes)	Statut
IBM Products Content Assesement	0	0	0%	0	 » ■
IBM Products Search			0%		 » ■
IBM Products Search with Classification	0	1	7%	0	 » ■
IBM Products search with security	0	0	0%	0	 » ■
Products Comments Analysis	0	0	0%	0	 » ■
Sample Text Analytics Collection			0%		 » ■
Douanes search	0	0	0%	0	 » ■

Les applications de recherche

- Un suivi détaillé et adapté

IBM Content Analytics with Enterprise Search Personnaliser

Collections Système Sécurité

Système : Rechercher > : Statistiques de requêtes

Statistiques de requêtes

[En savoir plus](#)

Spécifiez la période d'analyse et vos préférences d'affichage des statistiques.
Les informations sont actualisées automatiquement à l'intervalle que vous spécifiez.
Si vous sauvegardez l'historique des requêtes comme fichier CSV, vous pouvez importer les données dans d'autres applications pour une analyse supplémentaire.

Aujourd'hui | [Cette semaine](#) | [Ce mois](#) | [Période spécifique](#) | Options d'affichage | Affiner les résultats | Collection: Tous

Intervalle de rafraîchissement

Activer la régénération automatique

Intervalle [minutes] (1-60): 30

Graphique linéaire

Afficher la valeur: Nombre de requêtes

Echelle de temps: Heure

Serveur de recherche: CSLD1CM

Requêtes usuelles

Nombre de requêtes (10-50): 20

Historique des requêtes

Nombre de requêtes (10-50): 20

Améliorations

Affiner par ID utilisateur:

- Tous
- Douanes search
- IBM Products Content Assesment
- IBM Products Search
- IBM Products Search with Classification
- IBM Products search with security
- Products Comments Analysis
- Sample Text Analytics Collection
- Tous

L'historique de l'activité

- **Personnalisé selon les besoins**

IBM Content Analytics with Enterprise Search

Personnaliseur de recherche | Personnaliseur d'analyse | Déconnexion | Aide | A propos de

Collections | Système | Sécurité

Configuration de serveur | Programme d'écoute de données | Analyse syntaxique | Rechercher | Journal

[En savoir plus](#)

- Configurer des options de fichier journal
- Configurer des alertes
- Configurer les options d'adresse électronique pour les messages

Sélectionnez le fichier journal à afficher. Vous pouvez spécifier des filtres pour afficher des types spécifiques de messages ou des messages émis par des sessions spécifiques. Après avoir fait votre choix, cliquez sur **Afficher le journal**.

Fichier journal:

- system_20120529.log
- system_20120503.log
- system_20120502.log
- system_20120430.log
- system_20120406.log
- system_20120405.log
- system_20120322.log
- system_20120222.log
- system_20120203.log
- system_20111202.log

Nombre d'entrées de journal par page:

Filtres

Gravité:

- Tous les niveaux de gravité
- Erreur
- Avertissement
- Texte informatif

Session:

- Toutes les sessions
- Converter (node1)
- DSCConfigurator (node1)
- Customcommunication Service (node1)
- Search Manager (node1)
- Utilities Service (node1)
- ConfigManager Core Service (node1)
- Search Server (node1)
- ES Admin (node1)
- Resource Service (node1)
- Discovery (node1)
- Importer Service (node1)
- Search Application (node1)

Administration des collections

- Paramétrage
- Pilotage
- Suivi

The screenshot displays the 'Administration des collections' interface. At the top, the header includes the IBM logo, 'IBM Content Analytics with Enterprise Search', and navigation links: 'Personnaliseur de recherche | Personnaliseur d'analyse | Déconnexion | Aide | A propos de'. Below the header, a navigation bar contains 'Collections', 'Système', and 'Sécurité'. The main content area is titled 'Collections de recherche d'entreprise' and features a 'Créer une collection' button. A table lists search collections with columns for 'Explorer', 'Analyse syntaxique et index', and 'Rechercher'. The 'IBM Products Search with Classification' collection is highlighted in blue. Below the table, three numbered panels (1, 2, 3) show the configuration steps: 1. 'Explorer et importer' (stopped), 2. 'Analyse syntaxique et index' (stopped), and 3. 'Rechercher' (in progress). Panel 1 shows the 'Moteurs de balayage' section with 'IBM Products Crawler' (stopped) and an 'Importer' button. Panel 2 shows the 'Détails' section for 'Processus globaux' and 'Annotateurs'. Panel 3 shows the 'Détails' section for 'CSLD1CM' (in progress). The bottom of the interface shows a breadcrumb for 'IBM Products search with security'.

Suivi d'une collection

- Toutes les informations sur l'activité

IBM Content Analytics with Enterprise Search

Personnaliseur de recherche | Personnaliseur d'analyse | Déconnexion | Aide | A propos de

Collections | Système | Sécurité

* Créer une collection

Collections de recherche d'entreprise

	Explorer	Analyse syntaxique et index	Rechercher	
Douanes search	■	●	●	Actions ▼
IBM Products Search	■	■	■	Actions ▼
IBM Products Search with Classification	■	■	●	Actions ▼

IBM Products Search with Classification

Explorer et importer ► arrêté 1

Moteurs de balayage +

- IBM Products Crawler ► arrêté

Importer

Analyse syntaxique et index ► arrêté 2

Détails

Processus globaux

Annotateurs

CSLD1CM en cours d'exécution

- Afficher les fichiers journaux
- Configurer des options de fichier journal
- Configurer des alertes
- Configurer les options d'adresse électronique pour les messages
- Paramètres
- Consignation
- Cloner cette collection
- Supprimer cette collection

IBM Products search with security

Les fonctions d'analyse

- Zones
- Facettes
- Annotateurs

The screenshot displays the IBM Content Analytics with Enterprise Search interface. The top navigation bar includes the IBM logo, the product name, and links for 'Personnaliseur de recherche', 'Personnaliseur d'analyse', 'Déconnexion', 'Aide', and 'A propos de'. Below the navigation bar, there are tabs for 'Collections', 'Système', and 'Sécurité'. A button labeled 'Créer une collection' is visible.

The main content area is titled 'Collections de recherche d'entreprise'. It lists several search collections with their respective analysis and search status:

Collection	Explorer	Analyse syntaxique et index	Rechercher	Actions
Douanes search	■	●	●	Actions ▾
IBM Products Search	■	■	■	Actions ▾
IBM Products Search with Classification	■	■	●	Actions ▾

The 'IBM Products Search with Classification' collection is expanded, showing three main sections:

- 1 Explorer et importer**: Includes 'Moteurs de balayage' and 'Importer'. Status: 'arrêté'.
- 2 Analyse syntaxique et index**: Includes 'Détails', 'Processus globaux', and 'Annotateurs'. Status: 'arrêté'.
- 3 Rechercher**: Status: 'en cours d'exécution'.

A context menu is open over the 'Annotateurs' section, listing the following options:

- Configurer...
- Options d'analyse syntaxique
- Zones d'index
- Facettes
- Traitement global
- Annotateurs
- Plus...

Les fonctions de recherche

The screenshot displays the IBM Content Analytics with Enterprise Search interface. The main content area is titled "Collections de recherche d'entreprise" and lists several search collections: "Douanes search", "IBM Products Search", and "IBM Products Search with Classification". The "IBM Products Search with Classification" collection is expanded, showing a workflow with two steps: "Explorer et importer" (1) and "Analyse syntaxique et index" (2). The "Explorer et importer" step is currently "arrêté" (stopped) with 1 item. The "Analyse syntaxique et index" step is also "arrêté" (stopped). Below the workflow, there are sections for "Moteurs de balayage" (Search engines) and "Importer" (Import). A "Rechercher" (Search) button is visible. A "Configurer..." (Configure...) dropdown menu is open, listing various configuration options such as "Options du serveur de recherche", "Liens rapides", "Caractères génériques", "Portées", "Options de la saisie automatique", "Masques URI pour influencer les scores", "Résultats de la recherche réduits", "Importance du document", "Indicateurs de documents", "Règles d'optimisation des requêtes et des résultats", "Mémoire pour les serveurs de recherche", and "Exportation vers une collection d'analyse de contenu". The interface also includes a navigation bar with "Collections", "Système", and "Sécurité" tabs, and a footer with "IBM Products search with security" and "Actions" menu.

Des fonctions pointues d'optimisation

- Extensions de requêtes
- Filtres
- Agrégation

Règles d'optimisation des requêtes et des résultats

[En savoir plus](#)

Vous pouvez améliorer la précision des résultats en configurant des règles qui étendent automatiquement les requêtes lorsque des utilisateurs effectuent une recherche dans cette collection. Vous pouvez également configurer des filtres pour contrôler la manière dont les documents sont classés et des règles pour définir les priorités des groupes de documents dans les résultats.

Extension de requête basée sur des règles

Configuration des règles d'extension de requête

Activer l'extension de requête à base de règles

[Modifier les règles](#) [Importer des règles](#) [Supprimer les règles](#) [Exporter des règles](#)

Règles	
Nombre de règles	-
Date de la dernière modification	-

Configuration de dictionnaire

Sélectionnez les dictionnaires à utiliser pour l'extension de requête.

[Créer un dictionnaire](#) [Importer un dictionnaire](#)

Nom	Actions
-----	---------

Filtres de classement des documents

Configuration des filtres de classement des documents :

Activer les filtres de classement des documents

[Editer les filtres de classement des documents](#)

Groupe de filtres de classement des documents : **Premier**

Filtrer	Classement personnalisé
---------	-------------------------

Groupe de filtres de classement des documents : **Au début**

Filtrer	Classement personnalisé
title & Par défaut	

Groupe de filtres de classement des documents : **A la fin**

Filtrer	Classement personnalisé
---------	-------------------------

Groupe de filtres de classement des documents : **Par défaut**

Agrégation des résultats

Configuration de l'agrégation des résultats

Facile d'utilisation

Créer un moteur de balayage

[Aide concernant cette page](#)

Sélectionnez le type de données à ajouter à la collection.
Définissez ensuite les valeurs par défaut à utiliser.
Complétez les pages de l'assistant pour créer le moteur.

* Type de moteur de balayage:

[Sélectionner un élément]

- [Sélectionner un élément]
- Agent pour les systèmes de fichiers Windows
- Base de données JDBC
- Case Manager
- Content Integrator
- Content Manager
- DB2
- Domino Document Manager
- Exchange Server
- FileNet P8
- Gestion du contenu Web
- Liste de départ
- NNTP
- Notes
- QuickPlace
- SharePoint
- Système de fichiers Windows**
- Web
- WebSphere Portal

Précédent

Suivant

Terminer

Annuler

Sélectionner des sous-répertoires Windows à explorer

[Aide concernant cette page](#)

Indiquez si vous souhaitez explorer des sources locales ou éloignées. Pour les sources éloignées, spécifiez un nom de serveur et un chemin de données à explorer.
Utilisez un caractère générique (*) pour indiquer des noms de répertoire, puis cliquez sur **Rechercher des sous-répertoires disponibles**.

Sélectionnez des répertoires spécifiques et ajoutez-les à la liste **Sous-répertoires à explorer**. Répétez cette procédure jusqu'à ce que tous les sous-répertoires locaux et distants à explorer figurent dans la liste.

Explorer des sources locales

Nom de répertoire principal ou masque: (par exemple, c:\dir1\dir2*)

* [] Rechercher des sous-répertoires

Niveaux de sous-répertoires à explorer:

1

Explorer des sources distantes

Sous-répertoires disponibles:

c:\
c:\63e29af8875efa97b436b41a
c:\Autres_donnees
c:\Config.Msi
c:\Documents and Settings
c:\EXPORT
c:\HOME
c:\Program Files

* Sous-répertoires à explorer:

c:\Autres_donnees [Retirer]

Précédent

Suivant

Terminer

Annuler

Définition d'un moteur de balayage en quelques cliques

Agenda

- **L'information dans l'entreprise**
- **L'analyse de contenu**
- **IBM Content Analytics**
- **Exemples d'utilisation**
- **L'interface utilisateur**
- **Architecture**
- **Administration**
- **Pour aller plus loin**

Pour aller plus loin

- **IBM Content Analytics with Enterprise Search**

- Portail IBM Software

- <http://www.ibm.com/software/ecm/content-analytics/>

- IBM Redbooks

- <http://www.redbooks.ibm.com/redpieces/abstracts/sg247877.html?Open>

- **IBM Content Analytics Studio**

- Portail IBM Software

- <http://www.ibm.com/software/ebusiness/jstart/textanalytics/infocenter.html>

- IBM alphaWorks

- <http://www.alphaworks.ibm.com/tech/lrw>

- **IBM Content Classification**

- Portail IBM Software

- <http://www.ibm.com/software/ecm/content-classification/>

La plate-forme d'échange **Vivastream™**

- Développez votre réseau
- Découvrez les experts sur les sujets qui vous intéressent
- Echangez avec les speakers et les experts
- Regardez qui participe aux sessions auxquelles vous êtes inscrit
- Évaluez les sessions auxquelles vous avez participé



Inscrivez-vous sur le web ou avec votre smartphone

<http://www.vivastream.com/events/ibmimt-solutionsconnect-frfr>

